Supplementary Materials for: From Prediction to Decision: The Decision Integration Deficit Index (DIDI) and Structural Imbalance in AI-Driven Digital Health Systems

Scope of Supplementary Materials. This document provides methodological details, corpus construction procedures, a domain taxonomy, and coding inputs that support the structural analysis presented in the main manuscript. The formal definition of the DIDI metric and its computation are reported in the Methods section of the main manuscript. Interpretative analysis and system-level implications are presented exclusively in the main text. All terminology follows the standardized conventions adopted in the main manuscript, including ordinal integration strengths (0–3) and directed source–target pairs.

Supplementary Methods S1. Search strategy and PRISMA 2020 study selection

Databases and limits. Searches were conducted in Web of Science Core Collection (SCI-EXPANDED, SSCI) and Scopus. The primary corpus covered publications indexed between 2012 and 2025. To capture the most recent evidence, an update search was performed from 2025-12-01 to 2026-01-10, with the final search executed on 2026-01-10. Only peer-reviewed publications in English reporting empirical or methodological contributions relevant to digital health systems, AI-based inference, self-monitoring, self-testing, or the decision layer (C) were considered eligible.

Search strings (exact).

The following search strategies were used:

Scopus (TITLE-ABS-KEY):

("self-monitor*" OR wearable* OR "remote monitoring" OR telemonitor* OR "personal informatics") AND (health OR clinical OR patient) AND (system OR framework OR platform)

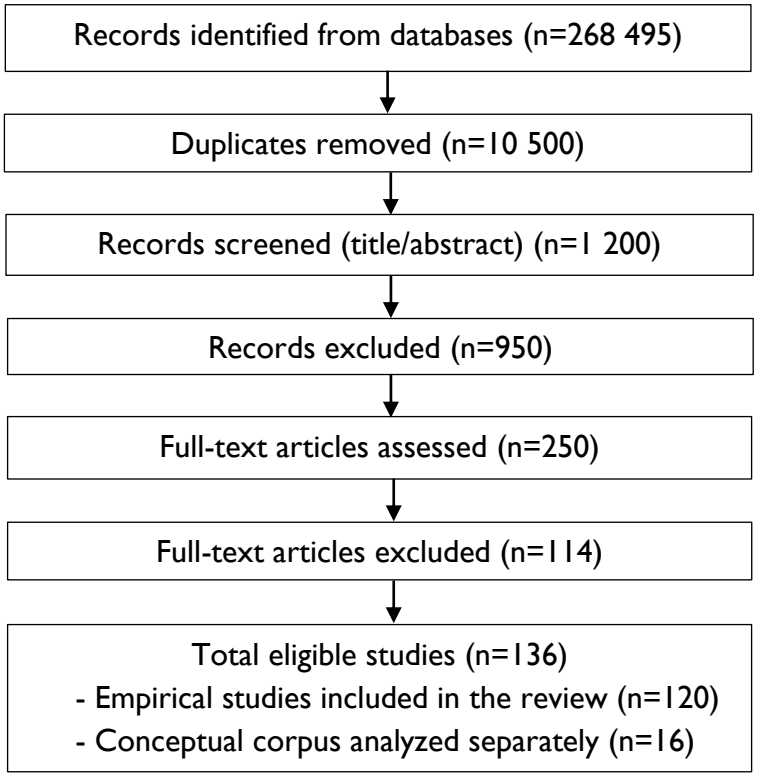
Web of Science (TS=):

("self-monitor*" OR wearable* OR "remote monitoring" OR telemonitor* OR "personal informatics") AND (health OR clinical OR patient) AND (system OR framework OR platform)

Screening and inclusion process. Records retrieved from the two databases were compiled into a unified dataset and de-duplicated using manual reference reconciliation in Microsoft Excel. Titles and abstracts were screened for relevance, and potentially eligible studies were assessed in full text according to the inclusion and exclusion criteria described in the main manuscript (Methods section). Disagreements between reviewers were resolved through discussion and consensus.

PRISMA 2020 reporting. Study selection followed PRISMA-based screening procedures. Due to the extensive initial retrieval, database-specific filters (including subject area restrictions,

document type, language, and relevance-based ranking) were applied to ensure feasibility while retaining representative coverage of relevant studies before title and abstract screening, reducing the dataset to approximately 1,200 unique records after deduplication. Approximately 10,500 duplicate records were removed, leaving around 1,200 records for title and abstract screening. Of these, approximately 950 records were excluded during screening. Full-text articles assessed for eligibility numbered approximately 250, of which around 114 were excluded with reasons, resulting in 136 eligible publications. Of these, 120 empirical studies were included in the structured review corpus, while 16 publications were designated as a separate conceptual corpus and analyzed independently. The number of included empirical studies (n = 120) reflects the predefined analytical design of the review, which required the construction of four balanced domain corpora of equal size (30 studies per domain: Self-Monitoring, Self-Testing, Artificial Intelligence, and Decision layer (C)) to ensure structural comparability across factors in the integration matrix. The detailed PRISMA flow diagram is provided as Supplementary Figure S1; no separate supplementary table was required for this stage, as all selection information is contained within the diagram.



Supplementary Figure S1. PRISMA flow diagram summarizing study identification, screening, eligibility assessment, and inclusion.

Supplementary Methods S2. Domain taxonomy and classification framework

1. Self-Monitoring Factor

1.1 Overview and Rationale of the Self-Monitoring Factor

The identified studies describe continuous behavioral and physiological data streams used as inputs within digital health systems, supporting trend analysis, early detection, personalized guidance, and adaptive interpretation within AI-enabled and Decision layer (C) – enabled decision frameworks [1, 2, 3, 4]. The framework positions self-monitoring as the primary source of real-world, user-generated inputs, which are subsequently transformed by AI models into structured features and interpreted through formal decision-analytic procedures. In this way, self-monitoring supports core functions such as behavior change, self-regulation, early risk identification, and informed daily decision-making.

1.2 Corpus Formation

To ensure methodological symmetry with the other three analytical factors, the Self-Monitoring corpus was constructed as a 30-study dataset encompassing foundational self-monitoring theory, wearable-based activity and sedentary behavior monitoring, and continuous physiological or stress monitoring. Publications were identified through structured searches in Scopus and Web of Science and were restricted to peer-reviewed articles published between 2013 and 2025. Eligible studies examined real-world or near-real-world monitoring practices in which individuals continuously or intermittently tracked behavioral, physiological, or contextual indicators using digital devices, wearables, smartphone sensors, or IoT-based systems.

The inclusion logic mirrored that used for the Self-Testing and AI factors. Only studies providing objective, sensor-derived, and behaviorally or physiologically meaningful monitoring signals, validated against reference standards or supported by methodological and behavioral analyses, were retained. Conceptual or purely technical sensor-engineering papers were excluded unless they directly informed the interpretation or validity of monitoring data. This approach ensured that the corpus represents the actual monitoring layer of contemporary digital health architectures rather than hypothetical or experimental sensing systems. The complete set of included studies and their domain classification are detailed in Supplementary Table S2.

Representative sources include Swan [5], Kersten van Dijk et al. [6], Strain et al. [4], Brickwood et al. [2], Duncan et al. [7], and Hovsepian et al. [8].

1.3 Inclusion Criteria

Consistent with the criteria applied to the AI and Self-Testing factors, studies were included if they:

1. Objectively measured behavioral, physiological, or contextual indicators using consumer-grade or research-grade wearables, smartphone sensors, or digital reporting applications.
2. Evaluated validity, reliability, feasibility, adherence, or behavioral impact under real-world conditions.
3. Generated interpretable indicators relevant to health behavior, physical activity, recovery, stress, or physiological state.
4. Were conceptually or empirically linked to health decision-making, early detection, or self-regulation processes.

5. Reported outcomes relevant to downstream AI or Decision layer (C) integration, such as continuous time-series metrics, behavior-change mechanisms, or methodological constraints.

Studies were excluded if they focused solely on technical sensor engineering, lacked a behavioral or health-related outcome, or did not evaluate the validity of monitoring approaches.

1.4 Conceptual Structure

Following the tripartite structure used for the Self-Testing and AI factors, the Self-Monitoring corpus is organized into three domains, each representing a distinct layer of person-generated data with implications for behavior change, early detection, and adaptive decision processes.

Domain 1 – Foundational and Conceptual Self-Monitoring

Foundational self-monitoring frameworks describe how individuals interact with personal informatics systems and how feedback, reflection, and awareness formation shape self-regulation processes. Core mechanisms include insight development, motivation, iterative adjustment, and reflective cycles driven by longitudinal data streams [5, 6]. This literature also highlights challenges such as fluctuating adherence and monitoring burden, where cognitive load, feedback fatigue, and perceived usefulness influence long-term sustainability. Collectively, this domain establishes the behavioral logic through which monitoring signals contribute to downstream interpretation in AI-enabled and decision-layer (C) – enabled health decision systems.

Domain 2 – Wearable-Based Activity and Sedentary Behavior Monitoring

This domain encompasses empirical evaluations of activity trackers, accelerometers, and hybrid mHealth systems used to monitor daily steps, moderate-to-vigorous physical activity (MVPA), sedentary behavior, and mobility in both clinical and general populations [2, 9, 10]. These studies provide evidence regarding validity, adherence, wear-time bias, real-world usability, and intervention effects. They demonstrate how wearable-captured movement data function as behavioral signatures linking moment-to-moment activity patterns with broader health trajectories (Supplementary Table S2).

Domain 3 – Continuous Physiological, Stress, and Clinical Monitoring

The third domain focuses on multimodal, high-frequency physiological monitoring using wearables, smartphones, and biodigital systems. Studies examine heart rate variability-based recovery tracking, stress detection models [8], physiological correlates of mental health [11], and behavior-change techniques embedded in commercial trackers [7, 12]. This domain forms the

closest conceptual bridge to AI, as the resulting data streams are granular, multivariate, and context-rich, providing a foundation for early detection and adaptive guidance.

These three domains are summarized in Supplementary Table S2. Together, the table links conceptual structure to concrete empirical evidence, maintaining methodological parallelism with the Self-Testing, AI, and Decision layer (C) factors.

1.5 Corpus Summary (30-Study Overview)

Across the corpus, self-monitoring research demonstrates three primary contributions. Foundational studies explain behavioral mechanisms underlying engagement with self-monitoring systems [5, 6]. Wearable-based investigations document the validity, usability, and behavioral impact of activity and sedentary monitoring in real-world contexts [2, 9, 10]. Research on continuous physiological and stress monitoring demonstrates the feasibility of high-resolution assessment of heart rate (HR), HRV, sleep, and stress using mobile and wearable platforms [7, 8, 12, 13]. The reviewed studies indicate that self-monitoring generates multidomain time-series data for early detection, trend analysis, and personalized decision-making, and that it serves as the foundational input layer for AI-based interpretation and Decision layer (C)–based prioritization.

2. Self-Testing Factor

2.1 Overview and Rationale of the Self-Testing Factor

The Self-Testing factor captures episodic, user-administered physiological, fitness, and functional assessments that provide structured calibration points within decentralized digital health systems. Unlike continuous monitoring, self-testing generates standardized measurements that can be interpreted against reference values, enabling physiological quantification, functional benchmarking, and validation of longer-term behavioral signals. Within the four-factor architecture, Self-Testing complements continuous Self-Monitoring by anchoring AI-based inference and supporting transparent prioritization within the Decision layer (C) – based decision frameworks.

2.2 Corpus Formation, Conceptual Structure, and Justification

To construct the Self-Testing corpus, 30 studies published between 2012 and 2025 were selected based on their evaluation of self-administered, home-based, or remote physiological and functional assessments. These studies demonstrate that individuals can independently generate quantifiable health indicators relevant to decentralized monitoring and personalized decision-making [14, 15, 16]. The corpus was designed to maintain methodological symmetry with the Self-Monitoring, AI,

and Decision (C) layer factors by focusing on validated, user-led measurement methods that yield interpretable data suitable for downstream computational modeling and structured decision-making.

2.3 Inclusion Criteria

Studies were included if they met all of the following criteria:

1. Self-administered measurement procedure. The primary measurement was performed by participants rather than clinicians (e.g., HbA1c testing, blood pressure measurement, 6-minute walk test, Timed Up and Go). Prior research indicates that many such procedures can be completed reliably by users when supported by validated tools and clear instructions [14].
2. Objective physiological or functional indicators. Only studies producing quantifiable metrics relevant to physiological status, aerobic performance, or functional mobility were retained. Infectious disease self-tests were excluded because they do not assess physiological functioning or physical capacity.
3. Demonstrated validity, reliability, or feasibility. Included studies reported evidence of measurement accuracy, reproducibility, or user acceptability under real-world or near-real-world conditions. Existing evidence shows that remote or self-administered tests can approximate laboratory or supervised measures with clinically acceptable precision [16, 17].
4. Relevance to AI-enabled monitoring. Most studies utilized smartphones, sensors, or connected devices that enable automated data capture and computational interpretation [18, 19]. In a small number of cases, supporting studies were included when they provided essential physiological or technological benchmarks for calibrating or contextualizing self-administered assessment protocols, even if they were not direct validation studies.

Studies were excluded if they did not involve user-led measurements, relied primarily on clinician-administered procedures, or focused exclusively on infectious disease diagnostics, as these do not align with the physiological or functional scope of the Self-Testing factor.

2.4 Conceptual Structure

The Self-Testing corpus is organized into three complementary domains:

1. Physiological self-testing: Examples include home-based assessments of HbA1c, glucose, INR, heart rate variability, and blood pressure.

2. Physical fitness self-testing: This domain includes submaximal aerobic tests, GPS-guided walking tests, and wearable-based fitness assessments.
3. Functional capacity self-assessment: Examples include self-administered 6-minute walk tests, Timed Up and Go protocols, and gait-related mobility metrics.

This tripartite structure aligns with contemporary digital health evaluation models that integrate internal physiological indicators, performance-based fitness markers, and real-world functional capacity measures to produce a multidimensional representation of individual health status [20, 21].

The three Self-Testing domains are summarized in Supplementary Table S3. Together, the table links the conceptual framework to concrete empirical evidence, mirroring the structured mapping applied to the Self-Monitoring, AI, and Decision layer (C) factors.

3. AI Factor

3.1 AI Factor: Corpus Formation, Conceptual Structure, and Justification

To construct a coherent AI factor aligned with contemporary digital health research, a corpus of 30 empirical and methodological studies published between 2013 and 2025 was assembled. These studies evaluate how AI models process multimodal clinical, physiological, wearable, and behavioral data to support prediction, diagnostics, and personalized health decision-making. The aim was to capture validated approaches that demonstrate how heterogeneous, user-linked health signals can be transformed into reliable computational indicators suitable for adaptive, individualized guidance systems.

The corpus was developed through structured searches in Scopus and Web of Science, restricted to peer-reviewed publications from 2013 to 2025. Eligible studies were required to include empirical evaluations of AI models applied to real clinical, sensor-based, behavioral, or physiological datasets. Conceptual or purely algorithmic studies without demonstrated health relevance were excluded. This corpus construction strategy mirrors the approach used for the Self-Monitoring, Self-Testing, and Decision (C) layer factors, ensuring methodological symmetry across domains and enabling consistent integration into the overarching AI–Decision (C) decision framework. The complete list of included studies is provided in Supplementary Table S4.

3.2 Conceptual Framework

The 30-study corpus is organized into three complementary domains representing the principal AI functions relevant to personalized health systems:

1. AI infrastructure, methods, and multimodal integration

   This domain encompasses large-scale clinical datasets, harmonized data environments, and scalable deep learning architectures for integrating multimodal health data, including electronic health records, imaging, biosignals, and wearable data [22, 23, 24]. These studies provide the methodological foundation for robust, generalizable modeling across diverse populations and clinical contexts.
2. AI for clinical and physiological diagnostics and risk prediction

   This domain includes image-based diagnostic systems, AI-enabled ECG interpretation, and models using physiological or wearable data to estimate cardiovascular, metabolic, or functional risk [25, 26, 27, 28]. These studies demonstrate how AI transforms raw physiological or clinical signals into actionable indicators of disease risk, early dysfunction, or physiological instability.
3. Digital phenotyping, behavior analytics, and adaptive interventions

   The third domain incorporates smartphone-based passive sensing, mobility traces, high-resolution wearable features, behavioral indicators, and context-aware analytics to infer momentary states such as stress, depressive symptoms, cardiometabolic risk, or functional decline [13, 29, 30]. These approaches capture real-world variability and contextualize physiological signals within daily routines, thereby supporting adaptive and personalized interventions.

Together, these domains span data infrastructure, diagnostic modeling, and behavioral-contextual analysis, reflecting the tripartite structural organization applied in the Self-Testing factor [22, 23, 25, 31]. Representative studies within each domain are listed in Supplementary Table S4. Collectively, these materials translate conceptual domains into a balanced 30-study corpus, ensuring methodological parallelism with the Self-Monitoring, Self-Testing, and Decision layer (C) factors. The study selection process, including database search, screening, and eligibility assessment, is documented in Supplementary Figure S1 following PRISMA 2020 guidelines.

3.3 Inclusion Criteria

Studies were included if they:

1. Applied machine-learning or deep-learning models to clinical, physiological, wearable, or behavioral data [23, 32, 33];
2. Reported empirical validation or predictive performance [25, 26, 27];

3. Generated outputs relevant for monitoring, early detection, or personalized decision processes [28, 30];
4. Demonstrated clear relevance to real-world physiological or behavioral interpretation.

Purely algorithmic studies without demonstrated health applications were excluded to maintain alignment with the Self-Monitoring and Self-Testing factors.

4. Decision layer (C) Factor

4.1 Corpus Formation, Conceptual Structure, and Justification

To ensure methodological symmetry with the AI and Self-Testing factors, the Decision layer (C) corpus was constructed as a 30-study dataset representing the development and application of structured decision-making approaches in modern health systems. The selection spans foundational methodological contributions, health-technology assessment (HTA) models, clinical decision applications, and emerging digital and AI-linked prioritization frameworks. This approach mirrors the logic applied in the other factors: diverse inputs, whether physiological, functional, behavioral, or AI-generated, require an interpretive layer capable of weighing, integrating, and translating evidence into actionable recommendations. The resulting corpus, therefore, reflects both classical value-based decision structures and contemporary extensions operating in digital, sensor-rich environments.

Studies were eligible for inclusion if they met the following criteria:

1. Implementation of a recognized Decision layer (C) method, such as AHP, TOPSIS, VIKOR, ELECTRE, DEX, fuzzy, or hybrid approaches;
2. direct relevance to real healthcare decisions, including HTA, treatment selection, medical device procurement, program evaluation, or patient prioritization;
3. transparent specification of criteria structures, weighting logic, and scoring procedures;
4. demonstration of decision-relevant outcomes.

These criteria parallel the inclusion logic used for the Self-Testing and AI factors, ensuring conceptual and methodological alignment across all four analytical components. The whole corpus and its domain allocation are reported in Supplementary Table S5.

Duplicates, purely theoretical decision-layer (C) papers, and non-health applications were excluded. The final corpus includes foundational methodological syntheses [34, 35], value-based frameworks and HTA methodological developments, systematic reviews summarizing the evolution of healthcare Decision layer (C) [36, 37], and selected contemporary applications in

clinical decision-making and digital health. These applications include, for example, device procurement [38], selection of chronic pain treatments [39], and hybrid AI-linked prioritization frameworks [40].

4.2 Conceptual Framework

Consistent with the tripartite structures used in the AI and Self-Testing factors, the Decision layer (C) corpus was organized into three complementary domains:

1. Decision layer (C) for Health Value, HTA, and Priority Setting

   This domain includes foundational methodological papers such as Thokala et al. [35] and Marsh et al. [37], which established ISPOR good-practice standards for Decision layer (C); the Advanced Value Framework [41]; methodological reflections on integrating Decision layer (C) into HTA; and systematic reviews summarizing the development of the field [36, 37]. These studies form the conceptual "infrastructure layer," analogous to the AI infrastructure domain.

2. Decision layer (C) for Clinical Technologies, Treatments, and System-Level Interventions

   Representative applications include Decision layer (C) tools for implantable device procurement [38], HTA prioritization processes, and treatment comparison using multi-criteria procedures [39]. This domain mirrors the transition from conceptual frameworks to operational decision pipelines, paralleling the mid-level application studies in the AI factor.

3. Decision layer (C) for Digital Health, Self-Care, Remote Monitoring, and AI-Linked Prioritization

   This domain comprises hybrid AI– Decision layer (C) approaches for multimorbidity prioritization [40] and home-based self-care decision models (Kapun et al., 2023) [42]. These studies operate most closely with the AI and Self-Testing ecosystems, emphasizing sensor-based data, dynamic updates, and algorithmically prioritized decisions.

The corresponding 30-study corpus is summarized in Supplementary Table S5. Together, these materials clarify how value frameworks, clinical decision models, and digital health applications converge into a unified interpretive decision layer within the four-factor architecture.

Supplementary References

1. Ajana, B. Personal Science and the Quantified Self Guru. In *Digital Wellness, Health and Fitness Influencers*; Routledge: London, UK, 2022; pp. 52–72. https://doi.org/10.4324/9781003256021
2. Brickwood, K.J.; Watson, G.; O'Brien, J.; Williams, A.D. Consumer-Based Activity Trackers Increase Physical Activity: Systematic Review and Meta-Analysis. *JMIR Mhealth Uhealth* 2019, *7*, e11819. https://doi.org/10.2196/11819
3. Sim, I. Mobile Devices and Health. *N. Engl. J. Med.* 2019, *381*, 956–968. https://doi.org/10.1056/NEJMra1806949
4. Strain, T.; Wijndaele, K.; Pearce, M.; Brage, S. Considerations for the Use of Consumer-Grade Wearables and Smartphones in Population Surveillance of Physical Activity. *J. Meas. Phys. Behav.* 2022, *5*, 8–14. https://doi.org/10.1123/jmpb.2021-0046
5. Swan, M. The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data* 2013, *1*, 85–99. https://doi.org/10.1089/big.2012.0002
6. Kersten van Dijk, E.T.; Westerink, J.H.D.M.; Beute, F.; IJsselsteijn, W.A. Personal Informatics, Self-Insight and Behavior Change: A Critical Review of Current Literature. *Hum. Comput. Interact.* 2017, *32*, 268–296. https://doi.org/10.1080/07370024.2016.1276456
7. Duncan, M.; Murawski, B.; Short, C.E.; Rebar, A.L.; Schoeppe, S.; Alley, S.; et al. Activity Trackers Implement Different Behaviour Change Techniques. *Interact. J. Med. Res.* 2017, *6*, e13. https://doi.org/10.2196/ijmr.6685
8. Hovsepian, K.; al'Absi, M.; Ertin, E.; Kamarck, T.; Nakajima, M.; Kumar, S. CStress: Towards a Gold Standard for Continuous Stress Assessment. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*; ACM: New York, NY, USA, 2015. https://doi.org/10.1145/2750858.2807526
9. Compernolle, S.; DeSmet, A.; Poppe, L.; Crombez, G.; De Bourdeaudhuij, I.; Cardon, G.; et al. Effectiveness of Interventions Using Self-Monitoring to Reduce Sedentary Behavior. *Int. J. Behav. Nutr. Phys. Act.* 2019, *16*, 63. https://doi.org/10.1186/s12966-019-0824-3
10. Ferguson, T.; Olds, T.; Curtis, R.; Blake, H.; Crozier, A.J.; Dankiw, K.; et al. Effectiveness of Wearable Activity Trackers: Systematic Review of Reviews. *Lancet Digit. Health* 2022, *4*, e615–e626. https://doi.org/10.1016/S2589-7500(22)00111-X
11. Sano, A.; Taylor, S.; McHill, A.W.; Phillips, A.J.K.; Barger, L.K.; Klerman, E.; et al. Identifying Objective Physiological Markers for Self-Reported Stress and Mental Health Using Wearable Sensors. *J. Med. Internet Res.* 2018, *20*, e210. https://doi.org/10.2196/jmir.9410
12. Mercer, K.; Li, M.; Giangregorio, L.; Burns, C.; Grindrod, K. Behaviour Change Techniques Present in Wearable Activity Trackers. *JMIR Mhealth Uhealth* 2016, *4*, e40. https://doi.org/10.2196/mhealth.4461
13. Sano, A.; Picard, R.W. Stress Recognition Using Wearable Sensors and Mobile Phones. In *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*; IEEE: Geneva, Switzerland, 2013. https://doi.org/10.1109/ACII.2013.117
14. Brooks, G.C.; Vittinghoff, E.; Iyer, S.; Tandon, D.; Kuhar, P.; Madsen, K.A.; et al. Accuracy and Usability of a Self-Administered 6-Minute Walk Test Smartphone Application. *Circ. Heart Fail.* 2015, *8*, 905–913. https://doi.org/10.1161/CIRCHEARTFAILURE.115.002062
15. Heneghan, C.; Ward, A.; Perera, R.; et al. Self-Monitoring of Oral Anticoagulation: Systematic Review and Meta-Analysis. *Lancet* 2012, *379*, 322–334. https://doi.org/10.1016/S0140-6736(11)61294-4
16. Mak, J.; Rens, N.; Savage, D.; Nielsen-Bowles, H.; Triggs, D.; Talgo, J.; et al. Reliability of a Smartphone-Based Six-Minute Walk Test. *Eur. Heart J. Digit. Health* 2021, *2*, 77–87. https://doi.org/10.1093/ehjdh/ztab018
17. Oronti, I.B.; Lopez-Perez, L.; Piaggio, D.; Fico, G.; Pecchia, L. Effects of Home Telemonitoring Using Non-Invasive versus Conventional Approaches on Patients Affected by Congestive Heart Failure: A Systematic Literature Review and Meta-Analysis. *Health Technol.* 2023, *13*, 535–555. https://doi.org/10.1007/s12553-023-00764-0
18. Maddison, R.; Nourse, R.; Daryabeygikhotbehsara, R.; et al. Digital Home-Based Self-Monitoring System for Heart Failure: Protocol. *JMIR Res. Protoc.* 2025, *14*, e62964. https://doi.org/10.2196/62964
19. Zaharieva, D.P.; Addala, A.; Prahalad, P.; Leverenz, B.; Arrizon-Ruiz, N.; Ding, V.Y.; et al. An Evaluation of Point-of-Care HbA1c, HbA1c Home Kits, and Glucose Management Indicator: Potential Solutions for Telehealth Glycemic Assessments. *Diabetology* 2022, *3*, 494–501. https://doi.org/10.3390/diabetology3030037
20. Pepera, G.; Karanasiou, E.; Blioumpa, C.; et al. Tele-Assessment of Functional Capacity Using the Six-Minute Walk Test. *Sensors* 2023, *23*, 1354. https://doi.org/10.3390/s23031354
21. Piatrikova, E.; Willsmer, N.J.; Altini, M.; Jovanović, M.; Mitchell, L.J.; Gonzalez, J.T.; et al. Monitoring the Heart Rate Variability Responses to Training Loads in Competitive Swimmers Using a Smartphone Application and

the Banister Impulse-Response Model. *Int. J. Sports Physiol. Perform.* 2021, *16*, 787–795. https://doi.org/10.1123/ijspp.2020-0201
22. Beam, A.L.; Kohane, I.S. Big Data and Machine Learning in Health Care. *JAMA* 2018, *319*, 1317–1318. https://doi.org/10.1001/jama.2017.18391
23. Johnson, A.E.W.; Pollard, T.J.; Shen, L.; et al. MIMIC-III, a Freely Accessible Critical Care Database. *Sci. Data* 2016, *3*, 160035. https://doi.org/10.1038/sdata.2016.35
24. Rajkomar, A.; Oren, E.; Chen, K.; et al. Scalable and Accurate Deep Learning with Electronic Health Records. *npj Digit. Med.* 2018, *1*, 18. https://doi.org/10.1038/s41746-018-0029-1
25. Attia, Z.I.; Kapa, S.; Lopez-Jimenez, F.; et al. Screening for Cardiac Contractile Dysfunction Using an Artificial Intelligence–Enabled Electrocardiogram. *Nat. Med.* 2019, *25*, 70–74. https://doi.org/10.1038/s41591-018-0240-2
26. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature* 2017, *542*, 115–118. https://doi.org/10.1038/nature21056
27. Gulshan, V.; Peng, L.; Coram, M.; et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy. *JAMA* 2016, *316*, 2402–2410. https://doi.org/10.1001/jama.2016.17216
28. Mishra, T.; Wang, M.; Metwally, A.A.; et al. Pre-Symptomatic Detection of COVID-19 from Smartwatch Data. *Nat. Biomed. Eng.* 2020, *4*, 1208–1220. https://doi.org/10.1038/s41551-020-00640-6
29. Onnela, J.-P.; Rauch, S.L. Harnessing Smartphone-Based Digital Phenotyping. *Neuropsychopharmacology* 2016, *41*, 1691–1696. https://doi.org/10.1038/npp.2016.7
30. Zhou, W.; Chan, Y.E.; Foo, C.S.; Zhang, J.; Teo, J.X.; Davila, S.; et al. High-Resolution Digital Phenotypes from Consumer Wearables and Their Applications in Machine Learning of Cardiometabolic Risk Markers: Cohort Study. *J. Med. Internet Res.* 2022, *24*, e34669. https://doi.org/10.2196/34669
31. Sutton, R.T.; et al. An Overview of Clinical Decision Support Systems: Benefits, Risks, and Strategies for Success. *npj Digit. Med.* 2020, *3*, 17. https://doi.org/10.1038/s41746-020-0221-y
32. Kline, A.; Wang, H.; Li, Y.; et al. Multimodal Machine Learning in Precision Health: A Scoping Review. *npj Digit. Med.* 2022, *5*, 171. https://doi.org/10.1038/s41746-022-00712-8
33. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep Learning for Healthcare. *Brief. Bioinform.* 2018, *19*, 1236–1246. https://doi.org/10.1093/bib/bbx044
34. Marsh, K.; Ijzerman, M.; Thokala, P.; et al. Multiple Criteria Decision Analysis for Health Care Decision Making: Emerging Good Practices. *Value Health* 2016, *19*, 125–137. https://doi.org/10.1016/j.jval.2015.12.016
35. Thokala, P.; Devlin, N.; Marsh, K.; et al. Multiple Criteria Decision Analysis for Health Care Decision Making: An Introduction. *Value Health* 2016, *19*, 1–13. https://doi.org/10.1016/j.jval.2015.12.003
36. Gongora-Salazar, P.; Rocks, S.; Fahr, P.; Rivero-Arias, O.; Tsiachristas, A. The Use of Multicriteria Decision Analysis to Support Decision Making in Healthcare: An Updated Systematic Literature Review. *Value Health* 2023, *26*, 780–790. https://doi.org/10.1016/j.jval.2022.11.007
37. Marsh, K.; Lanitis, T.; Neasham, D.; Orfanos, P.; Caro, J. Assessing the Value of Healthcare Interventions Using MCDA: A Review. *Pharmacoeconomics* 2014, *32*, 345–365. https://doi.org/10.1007/s40273-014-0135-0
38. Elezbawy, B.; Fasseeh, A.N.; Németh, B.; et al. Development of a Multiple Criteria Decision Analysis Tool for Implantable Devices in Egypt. *BMC Med. Inform. Decis. Mak.* 2022, *22*, 289. https://doi.org/10.1186/s12911-022-02025-y
39. Sir, E.; Batur Sir, G.D. Evaluating Treatment Modalities in Chronic Pain Treatment by the Multi-Criteria Decision-Making Procedure. *BMC Med. Inform. Decis. Mak.* 2019, *19*, 191. https://doi.org/10.1186/s12911-019-0925-6
40. Mohammed, K.I.; Jaafar, J.; Zaidan, A.A.; Albahri, O.S.; Zaidan, B.B.; Abdulkareem, K.H.; et al. A Uniform Intelligent Prioritisation for Solving Diverse and Big Data Generated from Multiple Chronic Diseases Patients Based on Hybrid Decision-Making and Voting Method. *IEEE Access* 2020, *8*, 91521–91530. https://doi.org/10.1109/ACCESS.2020.2994746
41. Angelis, A.; Kanavos, P. Multiple Criteria Decision Analysis for Evaluating New Medicines. *Soc. Sci. Med.* 2017, *188*, 137–156. https://doi.org/10.1016/j.socscimed.2017.06.024
42. Kapun, M.M.; Hlebec, V.; Zajc, M.; Kramberger, T. A Multi-Criteria Decision Model for Assessing the Health and Self-Care of Patients with Chronic Diseases in the Home Environment. *Cent. Eur. J. Oper. Res.* 2023, *31*, 1147–1176. https://doi.org/10.1007/s10100-022-00823-3

Supplementary Methods S3. Cross-factor and domain-level coding procedure

Each relationship was represented as a directed source–target pair reflecting the dominant orientation of information flow reported in the primary studies, such as data generation, model input, or formal decision processes (Supplementary Table S1). The direction of each linkage reflects the dominant functional role described in the source study, distinguishing upstream data generation and inference processes from downstream evaluation or prioritization mechanisms within the decision layer (C). These relationships were assigned ordinal integration strengths on a 4-level scale (0–3). Integration strengths were derived through independent dual-rater coding followed by structured consensus resolution, as detailed in the coding procedure described in this section, and are presented descriptively to reflect the frequency and explicitness with which each linkage was operationalized in the reviewed literature.

Inter-rater agreement was examined using percent agreement prior to consensus resolution, indicating high consistency across independently coded relationships before the consensus stage. All linkages were initially coded independently by two reviewers, with discrepancies resolved through structured consensus procedures. Individual pre-consensus coding records were not retained due to the consensus-based coding design; however, high pre-consensus agreement and fully specified coding rules ensure consistency and reproducibility of the final integration matrix. The finalized consensus matrix serves as the dataset for the structural analysis.

Methodological rigor was supported through predefined coding rules, independent dual coding, transparent consensus procedures, and citation-based weighting applied separately from the qualitative coding stage. Together, these procedures ensure that the resulting integration matrix reflects documented methodological relationships rather than subjective interpretation.

Supplementary Table S1. Complete domain-level integration matrix representing the whole structural space of directed relationships. This matrix constitutes the full analytical basis for all DIDI calculations reported in the main manuscript, including both observed and structurally possible domain-level relationships.

AI → Decision layer (C)

Source domain	Target domain (Decision layer (C))	Strength	Status
AI infrastructure and multimodal integration	Health Prioritization Systems	0	Missing
AI infrastructure and multimodal integration	Clinical Decision Models	0	Missing
AI diagnostics and risk prediction	Health Prioritization Systems	0	Missing
AI diagnostics and risk prediction	Clinical Decision Models	3	Observed
Digital phenotyping and behavioral analytics	Health Prioritization Systems	2	Observed

Digital phenotyping and behavioral analytics	Clinical Decision Models	0	Missing

Self-Monitoring → Decision layer (C)

Source domain	Target domain (Decision layer (C))	Strength	Status
Foundational and conceptual monitoring	Health Prioritization Systems	0	Missing
Foundational and conceptual monitoring	Clinical Decision Models	0	Missing
Behavioral tracking	Health Prioritization Systems	0	Missing
Behavioral tracking	Clinical Decision Models	0	Missing
Continuous physiological monitoring	Health Prioritization Systems	2	Observed
Continuous physiological monitoring	Clinical Decision Models	0	Missing

Self-Testing → Decision layer (C)

Source domain	Target domain (Decision layer (C))	Strength	Status
Physiological self-testing	Health Prioritization Systems	3	Observed
Physiological self-testing	Clinical Decision Models	0	Missing
Physical fitness self-testing	Health Prioritization Systems	0	Missing
Physical fitness self-testing	Clinical Decision Models	0	Missing
Functional capacity self-assessment	Health Prioritization Systems	0	Missing
Functional capacity self-assessment	Clinical Decision Models	0	Missing

Supplementary Table S2. Summary of the 30-Study Corpus for the Self-Monitoring Factor

No.	Reference	Description
DOMAIN 1. FOUNDATIONAL and CONCEPTUAL SELF-MONITORING		
1	Swan, M. The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data* 2013, *1*, 85–99. https://doi.org/10.1089/big.2012.0002	Establishes a conceptual foundation for self-monitoring by framing personal data streams as continuous, multimodal inputs that connect everyday behavior with biological discovery. Positions self-tracking as a methodological paradigm influencing individual decision-making and data-intensive health research.
2	Kersten van Dijk, E.T.; Westerink, J.H.D.M.; Beute, F.; IJsselsteijn, W.A. Personal Informatics, Self-Insight and Behavior Change: A Critical Review of Current Literature. *Hum. Comput. Interact.* 2017, *32*, 268–296. https://doi.org/10.1080/07370024.2016.1276456	Critically examines empirical evidence on personal informatics systems to assess whether self-tracking reliably produces self-insight and supports behavior change. Identifies methodological constraints and articulates key mechanisms through which monitoring data shape reflective processes.
3	Feng, S.; Mäntymäki, M.; Dhir, A.; Salmela, H. How Self-Tracking and the Quantified Self Promote Health and Well-Being: Systematic Review. *J. Med. Internet Res.* 2021, *23*, e25171. https://doi.org/10.2196/25171	Synthesizes findings on how self-tracking technologies influence health and well-being, highlighting mechanisms such as awareness formation, motivation, and behavioral insight. Shows that continuous monitoring contributes to early detection of health-relevant behavioral patterns.
4	Bartels, S.L.; van Knippenberg, R.J.M.; Dassen, F.C.M.; Asaba, E.; Patomella, A.-H.; Malinowsky, C.; et al. A Narrative Synthesis Systematic Review of Digital Self-Monitoring Interventions for Middle-Aged and Older Adults. *Internet Interv.* 2019, *18*, 100283. https://doi.org/10.1016/j.invent.2019.100283	Provides a structured synthesis of digital self-monitoring approaches among middle-aged and older adults, detailing monitored domains, technologies, feedback modalities, and adherence patterns. Demonstrates how self-monitoring supports self-management while noting gaps in sustained engagement and social health tracking.
5	Strain, T.; Wijndaele, K.; Pearce, M.; Brage, S. Considerations for the Use of Consumer-Grade Wearables and Smartphones in Population Surveillance of Physical Activity. *J. Meas. Phys. Behav.* 2022, *5*, 8–14. https://doi.org/10.1123/jmpb.2021-0046	Outlines methodological considerations for interpreting self-monitoring data from consumer wearables in population-level surveillance. Addresses sampling representativeness, wear-time bias, device validity, and inter-device comparability as factors shaping data quality.
6	Del-Valle-Soto, C.; López-Pimentel, J.C.; Vázquez-Castillo, J.; Nolazco-Flores, J.A.; Velázquez, R.; Varela-Aldás, J.; et al. A Comprehensive Review of Behavior Change Techniques in Wearables and IoT: Implications for Health and Well-Being. *Sensors* 2024, *24*, 2429. https://doi.org/10.3390/s24082429	Reviews behavior change techniques embedded in wearable and IoT-based monitoring systems, detailing how monitoring, feedback, and personalization mechanisms shape user engagement. Highlights the central role of continuous and intermittent monitoring in facilitating sustainable behavior modification.
7	Sanders, J.P.; Loveday, A.; Pearson, N.; Edwardson, C.; Yates, T.; Biddle, S.J.H.; et al. Devices for Self-Monitoring Sedentary Time or Physical Activity: A Scoping Review. *J. Med. Internet Res.* 2016, *18*, e90. https://doi.org/10.2196/jmir.5373	Maps the landscape of commercial devices enabling the self-monitoring of sedentary behavior and physical activity, categorizing sensor modalities and device functions. Identifies common technical limitations and functional gaps that shape real-world monitoring utility.
8	Strath, S.J.; Rowley, T.W. Wearables for Promoting Physical Activity. *Clin. Chem.* 2018, *64*, 53–63. https://doi.org/10.1373/clinchem.2017.272369	Reviews evidence on the accuracy and behavioral impact of wearable activity monitors, emphasizing self-monitoring as a core mechanism driving awareness-based physical activity change. Summarizes the technical strengths and limitations that shape their effectiveness.

9	Piwek, L.; Ellis, D.A.; Andrews, S.; Joinson, A. The Rise of Consumer Health Wearables: Promises and Barriers. *PLoS Med.* 2016, *13*, e1001953. https://doi.org/10.1371/journal.pmed.1001953	Examines consumer wearables through a monitoring-centered lens, outlining engagement barriers, adherence challenges, and data-interpretation issues. Argues that the benefits of monitoring depend on user motivation, the contextualization of feedback, and long-term device use.
10	Russell, E.; Kirk, A.; Dunlop, M.D.; Hodgson, W.; Patience, M.; Egan, K. Digital Physical Activity and Sedentary Behavior Interventions for Community-Living Adults: Umbrella Review. *J. Med. Internet Res.* 2025, *27*, e66294. https://doi.org/10.2196/66294	Synthesizes systematic reviews of digital PA/SB interventions, highlighting self-monitoring as a central behavior change component across modalities. Notes that while monitoring-driven interventions show promise, heterogeneity and reporting gaps limit definitive inferences.
DOMAIN 2. WEARABLE-BASED ACTIVITY and SEDENTARY-BEHAVIOUR MONITORING		
11	Brickwood, K.J.; Watson, G.; O'Brien, J.; Williams, A.D. Consumer-Based Activity Trackers Increase Physical Activity: Systematic Review and Meta-Analysis. *JMIR Mhealth Uhealth* 2019, *7*, e11819. https://doi.org/10.2196/11819	Synthesizes randomized and quasi-experimental studies evaluating commercial wearables for increasing physical activity. Demonstrates consistent short-term increases in steps and MVPA, reinforcing self-monitoring as a primary mechanism driving behavior change.
12	Ferguson, T.; Olds, T.; Curtis, R.; Blake, H.; Crozier, A.J.; Dankiw, K.; et al. Effectiveness of Wearable Activity Trackers to Increase Physical Activity and Improve Health: A Systematic Review of Systematic Reviews and Meta-Analyses. *Lancet Digit. Health* 2022, *4*, e615–e626. https://doi.org/10.1016/S2589-7500(22)00111-X	Aggregates evidence from 39 systematic reviews evaluating wearable-based physical activity interventions across clinical and general populations. Shows robust though modest improvements in steps, walking time, and select health outcomes, positioning wearables as scalable self-monitoring tools.
13	Larsen, R.T.; Christensen, J.; Juhl, C.B.; Andersen, H.B.; Langberg, H. Physical Activity Monitors to Enhance the Daily Amount of Physical Activity in Older Adults: Systematic Review and Meta-Analysis. *Eur. Rev. Aging Phys. Act.* 2019, *16*, 7. https://doi.org/10.1186/s11556-019-0213-6	Evaluates wearable-assisted interventions targeting older adults and quantifies their effects on steps per day and MVPA. Confirms that continuous self-monitoring combined with feedback and goal-setting yields measurable increases in activity.
14	Franssen, W.M.A.; Franssen, G.H.L.M.; Spaas, J.; Solmi, F.; Eijnde, B.O. Can Consumer Wearable Activity Tracker–Based Interventions Improve Physical Activity and Cardiometabolic Health in Patients with Chronic Diseases? *Int. J. Behav. Nutr. Phys. Act.* 2020, *17*, 57. https://doi.org/10.1186/s12966-020-00955-2	Meta-analysis of 35 RCTs showing that wearable-based interventions increase steps and yield modest improvements in cardiometabolic indicators. Highlights self-monitoring and real-time feedback as key components driving behavior and physiological change.
15	Elfouly, T.; Alouani, A. A Comprehensive Survey on Wearable Computing for Mental and Physical Health Monitoring. *Electronics* 2025, *14*, 3443. https://doi.org/10.3390/electronics14173443	Surveys recent advances in sensor modalities, computational architectures, and analytics for wearable-based monitoring of physical and mental health. Emphasizes the shift from passive data capture to personalized, context-aware self-monitoring frameworks.
16	Butte, N.F.; Ekelund, U.; Westerterp, K.R. Assessing Physical Activity Using Wearable Monitors: Measures of Physical Activity. *Med. Sci. Sports Exerc.* 2012, *44*, S5–S12. https://doi.org/10.1249/MSS.0b013e3182399c0e	Methodological overview of wearable-based physical activity assessment; details sensor modalities, output metrics and analytic considerations affecting validity and interpretation.
17	Sun, Y.; Chen, J.; Ji, M.; Li, X. Wearable Technologies for Health Promotion and Disease Prevention in Older Adults: Systematic Scoping Review and Evidence Map. *J. Med. Internet Res.* 2025, *27*, e69077. https://doi.org/10.2196/69077	Maps wearable use cases in older adults across mobility, mental health, arrhythmia detection, sleep, and preventive monitoring. Identifies opportunities for long-term preventive health support and

		challenges related to data quality, usability, and real-world adoption.
18	Compernolle, S.; DeSmet, A.; Poppe, L.; Crombez, G.; De Bourdeaudhuij, I.; Cardon, G.; et al. Effectiveness of Interventions Using Self-Monitoring to Reduce Sedentary Behavior in Adults: Systematic Review and Meta-Analysis. *Int. J. Behav. Nutr. Phys. Act.* 2019, *16*, 63. https://doi.org/10.1186/s12966-019-0824-3	Synthesizes intervention studies using self-monitoring to reduce sedentary time and shows significant reductions when objective monitoring tools are used. Demonstrates that focusing directly on SB rather than combined PA+SB targets yields more potent effects.
19	Compernolle, S.; Cardon, G.; van der Ploeg, H.P.; van Nassau, F.; De Bourdeaudhuij, I.; Jelsma, J.J.; et al. Engagement, Acceptability, Usability, and Preliminary Efficacy of a Self-Monitoring Mobile Health Intervention to Reduce Sedentary Behavior in Belgian Older Adults: Mixed Methods Study. *JMIR Mhealth Uhealth* 2020, *8*, e18653. https://doi.org/10.2196/18653	Assesses a self-monitoring mHealth intervention for reducing sedentary behavior in older adults, combining usability metrics with objective accelerometry outcomes. Shows high engagement and acceptability, though behavioral effects over a short follow-up were minimal.
20	Aguiar, M.; Trujillo, M.; Chaves, D.; Álvarez, R.; Epelde, G. mHealth Apps Using Behavior Change Techniques to Self-Report Data: Systematic Review. *JMIR Mhealth Uhealth* 2022, *10*, e33247. https://doi.org/10.2196/33247	Reviews mHealth applications incorporating behavior change techniques to improve adherence in self-reported monitoring data. Identifies standard BCT configurations and emphasizes the need for clearer design principles to sustain reliable longitudinal data entry.
DOMAIN 3. CONTINUOUS PHYSIOLOGICAL, STRESS and CLINICAL MONITORING		
21	Lyons, E.J.; Lewis, Z.H.; Mayrsohn, B.G.; Rowland, J.L. Behavior Change Techniques Implemented in Electronic Lifestyle Activity Monitors: A Systematic Content Analysis. *J. Med. Internet Res.* 2014, *16*, e192. https://doi.org/10.2196/jmir.3469	Content analysis of lifestyle monitors mapping behaviour change techniques reveals that self-monitoring, feedback, and goal-setting dominate, while planning/maintenance features vary.
22	Mercer, K.; Li, M.; Giangregorio, L.; Burns, C.; Grindrod, K. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. *JMIR Mhealth Uhealth* 2016, *4*, e40. https://doi.org/10.2196/mhealth.4461	Codes seven commercial trackers using the CALO-RE framework to quantify embedded self-monitoring and self-regulation mechanisms. Identifies a consistent emphasis on monitoring and feedback while noting limited integration of advanced planning-oriented BCTs that may constrain long-term behavior change.
23	Gjoreski, M.; Gjoreski, H.; Luštrek, M.; Gams, M. Continuous Stress Detection Using a Wrist Device: In Laboratory and Real-Life Settings. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Adjunct)*; ACM: New York, NY, USA, 2016; pp. 1185–1193. https://doi.org/10.1145/2968219.2968306	Evaluates continuous stress detection using wrist physiological sensors in both laboratory and real-world settings; multi-sensor fusion enables the discrimination of stress states outside controlled settings.
24	Sano, A.; Taylor, S.; McHill, A.W.; Phillips, A.J.K.; Barger, L.K.; Klerman, E.; et al. Identifying Objective Physiological Markers and Modifiable Behaviors for Self-Reported Stress and Mental Health Status Using Wearable Sensors and Mobile Phones. *J. Med. Internet Res.* 2018, *20*, e210. https://doi.org/10.2196/jmir.9410	Uses continuous multimodal wearable and smartphone data to identify physiological and behavioral correlates of self-reported stress and mental health. Machine-learning models achieve strong classification accuracy, demonstrating the feasibility of digital phenotyping for real-world psychological-state monitoring.
25	Gabbett, T.J. The Training–Injury Prevention Paradox: Should Athletes Be Training Smarter and Harder? *Br. J. Sports Med.* 2016, *50*, 273–280. https://doi.org/10.1136/bjsports-2015-095788	Articulates the training–injury paradox, emphasizing that systematic monitoring of acute and chronic load trajectories can identify elevated injury risk—positions continuous workload surveillance as a

		central mechanism in athlete resilience and injury prevention.
26	Drew, M.K.; Finch, C.F. The Relationship between Training Load and Injury, Illness and Soreness: A Systematic Review. *Sports Med.* 2016, *46*, 861–883. https://doi.org/10.1007/s40279-015-0459-8	Synthesizes evidence on the relationship between training-load metrics and injury and illness outcomes across athletic populations. Highlights consistent associations between rapid load increases and adverse events, underscoring the importance of continuous load monitoring for risk management.
27	Dunn, J.; Runge, R.; Snyder, M. Wearables and the Medical Revolution. *Pers. Med.* 2018, *15*, 429–448. https://doi.org/10.2217/pme-2018-0044	Reviews wearable sensor applications across metabolic, cardiovascular, neurological, sleep, and environmental health domains. Emphasizes how continuous physiological monitoring enables longitudinal, real-world insights that extend beyond traditional clinical assessment.
28	Quer, G.; Radin, J.M.; Gadaleta, M.; Baca-Motes, K.; Ariniello, L.; Ramos, E.; et al. Wearable Sensor Data and Self-Reported Symptoms for COVID-19 Detection. *Nat. Med.* 2021, *27*, 73–77. https://doi.org/10.1038/s41591-020-1123-x	Analyzes multimodal wearable and self-report data from over 30,000 participants to distinguish COVID-positive from negative cases. Demonstrates that combining passive physiological signals with symptom reports improves discriminative accuracy over symptom-based models alone.
29	Au, R.; Kolachalama, V.B.; Paschalidis, I.C. Redefining and Validating Digital Biomarkers as Fluid, Dynamic Measures of Health. *Front. Digit. Health* 2022, *4*, 751629. https://doi.org/10.3389/fdgth.2021.751629	Reframes digital biomarkers as temporally dynamic indicators requiring continuous, context-rich data for meaningful interpretation. Discusses methodological and regulatory challenges and underscores the importance of wearable-derived longitudinal signals for precision monitoring.
30	Babatunde, A.O.; Ogundijo, D.A.; Afolayan, A.O.; Awosiku, O.V.; Aderohunmu, Z.O.; Oguntade, M.S.; et al. Mobile Health Technologies in the Prevention and Management of Hypertension: A Scoping Review. *Digit. Health* 2024, *10*, 20552076241277172. https://doi.org/10.1177/20552076241277172	Synthesizes evidence on mobile health approaches for hypertension prevention and management, detailing how BP tracking, medication adherence monitoring, and lifestyle surveillance integrate into multimodal care. Highlights continuous and remote physiological monitoring as foundational to scalable hypertension control models.

Supplementary Table S3. Summary of the 30-Study Corpus for the Self-Testing Factor.

No	Reference	Description
CATEGORY 1: Physiological Self-Testing (Biomarkers, HRV, BP, Glucose, HbA1c, INR)		
1	Verougstraete, N.; Stove, V.; Stove, C.P. Remote HbA1c Testing via Microsampling: Fit for Purpose? *Clin. Chem. Lab. Med.* 2024, *62*, 3–17. https://doi.org/10.1515/cclm-2023-0228	Validates a microsampling-based home HbA1c testing approach and demonstrates that results meet clinically acceptable accuracy thresholds. Supports the feasibility of patient-led glycemic monitoring in decentralized care settings.
2	Van Beek, A.; Moeyaert, M.; Ragheb, B.; Price, E.; MacEwan, J.P.; Ahmed, N.; et al. Outcomes of Warfarin Home INR Monitoring versus Office-Based Monitoring: A Retrospective Claims-Based Analysis. *J. Gen. Intern. Med.* 2024, *39*, 1127–1134. https://doi.org/10.1007/s11606-023-08348-4	An extensive retrospective analysis comparing home INR monitoring with clinic-based care. Home testing was associated with improved therapeutic control and fewer thromboembolic events, supporting the real-world effectiveness of patient self-testing modalities.
3	Heneghan, C.; Ward, A.; Perera, R.; et al. Self-Monitoring of Oral Anticoagulation: Systematic Review. *Lancet* 2012, *379*, 322–	This meta-analysis shows that trained patients can safely and accurately perform self-testing for anticoagulation. Demonstrates clinically meaningful reductions in

	334. https://doi.org/10.1016/S0140-6736(11)61294-4	thromboembolic risk when patients engage in structured self-management.
4	Bereznicki, L.R.E.; Jackson, S.L.; Peterson, G.M. Supervised Patient Self-Testing of Warfarin Therapy Using an Online System. *J. Med. Internet Res.* 2013, *15*, e138. https://doi.org/10.2196/jmir.2255	Evaluates a digital system supporting at-home INR self-testing and remote clinician oversight. The platform enabled stable therapeutic levels and highlights how digital feedback loops enhance safe self-management.
5	Piatrikova, E.; Willsmer, N.J.; Altini, M.; Jovanović, M.; Mitchell, L.J.; Gonzalez, J.T.; et al. Monitoring Heart Rate Variability Responses to Training Loads in Competitive Swimmers Using a Smartphone Application. *Int. J. Sports Physiol. Perform.* 2021, *16*, 787–795. https://doi.org/10.1123/ijspp.2020-0201	Validates daily athlete-led HRV self-measurement as a sensitive marker of physiological load and recovery. Demonstrates that smartphone-based HRV testing can reliably guide individualized training adjustments.
6	Greenwood, D.A.; Blozis, S.A.; Young, H.M.; Nesbitt, T.S.; Quinn, C.C. Overcoming Clinical Inertia: A Randomized Clinical Trial of a Telehealth Remote Monitoring Intervention Using Paired Glucose Testing in Adults with Type 2 Diabetes. *J. Med. Internet Res.* 2015, *17*, e178. https://doi.org/10.2196/jmir.4112	RCT of telehealth intervention using structured paired glucose self-testing plus remote feedback; reports improved glycemic outcomes and reduced clinical inertia.
7	Beck, R.W.; Connor, C.G.; Mullen, D.M.; Wesley, D.M.; Bergenstal, R.M. The Fallacy of Average: How Using HbA1c Alone to Assess Glycemic Control Can Be Misleading. *Diabetes Care* 2017, *40*, 994–999. https://doi.org/10.2337/dc17-0636	Shows that relying solely on HbA1c can misrepresent actual glycemic patterns, highlighting the importance of patient-generated glucose profiles and home monitoring data. Supports integrating self-testing glucose information into more robust, individualized decision-support systems.
8	Zaharieva, D.P.; Addala, A.; Prahalad, P.; Leverenz, B.; Arrizon-Ruiz, N.; Ding, V.Y.; et al. An Evaluation of Point-of-Care HbA1c, HbA1c Home Kits, and Glucose Management Indicator. *Diabetology* 2022, *3*, 494–501. https://doi.org/10.3390/diabetology3030037	Evaluates performance of home-kit HbA1c testing and CGM-derived Glucose Management Indicator (GMI) compared with point-of-care HbA1c in youth with type 1 diabetes. Finds strong concordance between home-kit and POC HbA1c, supporting the feasibility of remote glycemic assessment in telehealth settings — acknowledges limitations of pediatric population and sample size.
9	Gal, R.L.; Cohen, N.J.; Kruger, D.; Beck, R.W.; Bergenstal, R.M.; Calhoun, P.; et al. Diabetes Telehealth Solutions: Improving Self-Management through Remote Initiation of Continuous Glucose Monitoring. *J. Endocr. Soc.* 2020, *4*, bvaa076. https://doi.org/10.1210/jendso/bvaa076	Demonstrates that remote initiation of CGM improves glycemic outcomes and patient engagement. Shows how telehealth-supported home monitoring strengthens diabetes self-management.
10	Lee, P.A.; Greenfield, G.; Pappas, Y. The Impact of Telehealth Remote Patient Monitoring on Glycemic Control in Type 2 Diabetes: A Systematic Review and Meta-Analysis of Systematic Reviews. *BMC Health Serv. Res.* 2018, *18*, 495. https://doi.org/10.1186/s12913-018-3274-8	Meta-review evaluating telehealth remote monitoring effects on glycemic control in type 2 diabetes; reports improvements in HbA1c with structured monitoring and feedback vs standard care.

CATEGORY 2: Physical Fitness Self-Testing (Cardiorespiratory fitness, performance tests, self-administered field tests)		
11	de Jesus, M.O.; Ostolin, T.L.V.D.P.; et al. Self-Administered Six-Minute Walk Test Using a Free Smartphone App in Asymptomatic Adults: Reliability and Reproducibility. *Int. J. Environ. Res. Public Health* 2022, *19*, 1118. https://doi.org/10.3390/ijerph19031118	Validates a smartphone-guided self-administered 6MWT, demonstrating strong agreement with clinician-supervised protocols. Confirms feasibility and reliability in free-living environments.
12	Mak, J.; Rens, N.; Savage, D.; Nielsen-Bowles, H.; Triggs, D.; Talgo, J.; et al. Reliability and Repeatability of a Smartphone-Based 6-Minute Walk Test. *Eur. Heart J. Digit. Health* 2021, *2*, 77–87. https://doi.org/10.1093/ehjdh/ztab018	Shows excellent repeatability of home-based smartphone 6MWT metrics in cardiovascular patients. Supports remote functional capacity monitoring using digital self-testing tools.
13	Ferguson, T.; Rowlands, A.V.; Olds, T.; Maher, C. Validity of Consumer-Level Activity Monitors in Healthy Adults Worn in Free-Living Conditions. *Int. J. Behav. Nutr. Phys. Act.* 2015, *12*, 42. https://doi.org/10.1186/s12966-015-0201-9	Assesses the validity of consumer wearable activity monitors in free-living conditions; shows several devices provide sufficiently accurate step/activity estimates for self-administered monitoring.
14	Lee, K.; Kim, D.; Shin, S.; Choi, H.; Jang, A.; Chung, J. Smartwatch-Derived VO_2max Prediction Model for Korean Adults. *Healthcare* 2025, *13*, 722. https://doi.org/10.3390/healthcare13070722	Develops and validates a smartwatch-based VO_2max prediction model using user-generated heart-rate and GPS data. Demonstrates strong agreement with laboratory VO_2 metrics, supporting remote aerobic fitness assessment.
15	Lin, I.I.; et al. Test–Retest Reliability of Home-Based Fitness Assessments Using the R Plus Health App in Healthy Adults. *JMIR Form. Res.* 2021, *5*, e28040. https://doi.org/10.2196/28040	Evaluates test–retest reliability of multiple home-based fitness assessments conducted via a smartphone application in healthy adults. Demonstrates good reliability and feasibility of unsupervised functional and mobility tests in non-clinical settings.
16	Van Hooren, B.; Fuller, J.T.; Buckley, J.D.; Miller, J.R.; Sewell, K.; Rao, G.; et al. Is Motorized Treadmill Running Biomechanically Comparable to Overground Running? A Systematic Review and Meta-Analysis. *Sports Med.* 2020, *50*, 785–813. https://doi.org/10.1007/s40279-019-01237-z	Systematic review/meta-analysis comparing treadmill vs overground running biomechanics/physiology; provides reference parameters for translating lab/treadmill intensity metrics to field/self-administered assessments.
17	Nordgren, B.; Fridén, C.; Jansson, E.; Österlund, T.; Grooten, W.J.A.; Opava, C.H.; et al. Criterion Validation of Two Submaximal Aerobic Fitness Tests, the Self-Monitoring Fox-Walk Test and the Åstrand Cycle Test, in People with Rheumatoid Arthritis. *BMC Musculoskelet. Disord.* 2014, *15*, 305. https://doi.org/10.1186/1471-2474-15-305	Validates the Fox-walk and Åstrand cycle self-administered fitness tests in clinical populations. Demonstrates that self-paced submaximal tests can accurately estimate aerobic capacity
18	Villar, R.; Beltrame, T.; Ferreira dos Santos, G.; Zago, A.S.; Bocalini, D.S.; Pontes Júnior, F.L. Test–Retest Reliability and Agreement of Remote Home-Based Functional Capacity	Shows excellent test–retest reliability of remote, self-administered functional capacity tests in socially isolated older adults. Supports digital monitoring of mobility and endurance.

	Self-Administered Assessments in Community-Dwelling, Socially Isolated Older Adults. *Digit. Health* 2024, *10*, 20552076241254904. https://doi.org/10.1177/20552076241254904	
19	Simmich, J.; Andrews, N.E.; Claus, A.; Murdoch, M.; Russell, T.G. Assessing a GPS-Based 6MWT for People with Persistent Pain. *JMIR Form. Res.* 2024, *8*, e46820. https://doi.org/10.2196/46820	Validates a GPS-enabled, self-administered 6MWT in individuals with persistent pain, demonstrating strong agreement with clinic-based measures. Confirms the feasibility and accuracy of unsupervised distance tracking in real-world environments.
20	Altini, M.; Casale, P.; Penders, J.; Amft, O. Cardiorespiratory Fitness Estimation in Free-Living Using Wearable Sensors. *Artif. Intell. Med.* 2016, *68*, 37–46. https://doi.org/10.1016/j.artmed.2016.02.002	Uses wearable sensor data collected independently by users in free-living contexts to estimate cardiorespiratory fitness. Demonstrates the feasibility of passive, user-generated fitness monitoring without laboratory testing.
CATEGORY 3: Functional Capacity Self-Assessment (6MWT, TUG, gait, mobility tests—home or self-administered)		
21	Brooks, G.C.; Vittinghoff, E.; Iyer, S.; et al. Accuracy and Usability of a Self-Administered 6-Minute Walk Test Smartphone Application. *Circ. Heart Fail.* 2015, *8*, 905–913. https://doi.org/10.1161/CIRCHEARTFAILURE.115.002062	Demonstrates high accuracy and usability of a self-administered 6MWT smartphone app. Shows strong agreement with clinician-administered protocols.
22	Storm, F.A.; Heller, B.W.; Mazzà, C. Step Detection Accuracy of Physical Activity Monitors. *PLoS ONE* 2015, *10*, e0118723. https://doi.org/10.1371/journal.pone.0118723	Evaluates seven wearable sensors for accuracy in detecting walking and running steps. Provides foundational evidence for the use of consumer wearables in home-based functional assessments.
23	Mellone, S.; Tacconi, C.; Chiari, L. Validity of a Smartphone-Based Instrumented Timed Up and Go. *Gait Posture* 2012, *36*, 163–165. https://doi.org/10.1016/j.gaitpost.2012.02.006	Validation study showing smartphone inertial sensors can instrument the Timed Up and Go test with agreement vs reference systems; supports minimally supervised mobility self-testing.
24	Scherrenberg, M.; Valkenborghs, K.; Vissers, D.; et al. A Mobile Application to Perform the Six-Minute Walk Test (6MWT) at Home: A Random Walk in the Park Is as Accurate as a Standardized 6MWT. *Sensors* 2022, *22*, 4277. https://doi.org/10.3390/s22114277	Shows that home-based, app-guided 6MWT performance is comparable to standardized corridor testing. Supports decentralized mobility monitoring.
25	Pepera, G.; Koufaki, P.; Batalik, L.; Vrana, A.; Svoboda, Z.; Littnerová, S. Tele-Assessment of Functional Capacity through the Six-Minute Walk Test in Patients with Diabetes Mellitus Type 2: Validity and Reliability of Repeated Measurements. *Sensors* 2023, *23*, 1354. https://doi.org/10.3390/s23031354	Demonstrates excellent reliability of remote 6MWT in individuals with type 2 diabetes. Highlights the feasibility of telemonitoring functional capacity
26	Salvi, D.; Olsson, C.; Molloy, J.; Orchard, E. Clinical Usefulness of a Smartphone-Based	This mixed-methods study examines the clinical usefulness of a smartphone-based self-administered 6MWT during pandemic-related service restrictions.

	Self-Administered Six-Minute Walk Test in a Hospital Outpatient Clinic within the Constraints of the COVID-19 Pandemic: Mixed Methods Study. *JMIR Form. Res.* 2025, e70495. https://doi.org/10.2196/70495	While not a pure validation study, it provides practical insights into integrating remote mobility assessments into outpatient workflows. Its inclusion reflects the corpus principle that empirical implementation evidence can be used when it directly informs the feasibility and design of self-administered functional testing.
27	Tsiakiri, A.; Plakias, S.; Giarmatzis, G.; Tsakni, G.; Christidi, F.; Papadopoulou, M.; et al. Gait Analysis in Multiple Sclerosis: A Scoping Review of Advanced Technologies for Adaptive Rehabilitation and Health Promotion. *Biomechanics* 2025, *5*, 65. https://doi.org/10.3390/biomechanics5030065	Synthesizes technological approaches for gait and mobility evaluation in MS, including wearable and smartphone tools. Identifies opportunities for self-administered functional capacity assessments.
28	Silva, B.M.; Rodrigues, J.J.P.C.; de la Torre Díez, I.; López-Coronado, M.; Saleem, K. Mobile Health: A Review of Current State in 2015. *J. Biomed. Inform.* 2015, *56*, 265–272. https://doi.org/10.1016/j.jbi.2015.06.003	Reviews mobile-health technologies supporting remote monitoring and functional assessment. Highlights sensor-based mobility evaluation relevant to decentralized 6MWT and TUG implementations.
29	Lachant, D.J.; Lachant, M.D.; Haight, D.; White, R.J. Cardiac Effort and 6-Min Walk Distance Correlate with Stroke Volume Measured by Cardiac Magnetic Resonance Imaging. *Pulm. Circ.* 2024, *14*, e12355. https://doi.org/10.1002/pul2.12355	Shows strong physiological coupling between cardiac effort and 6MWT distance. Supports the interpretation of remotely collected functional capacity metrics.
30	Torres-Castro, R.; Pascual, H.; Alonso, A.; Gimeno-Santos, E.; Palomo, M.; Barberà, J.A.; et al. Use of Automatic 6-Minute Walking Test Recording System in Patients with Chronic Respiratory Diseases. *Pulmonology* 2024, *30*, 390–392. https://doi.org/10.1016/j.pulmoe.2023.08.011	Evaluates a wearable-assisted automated 6MWT recording system. Confirms the accuracy of distance and oxygenation tracking in remote functional testing.

Supplementary Table S4. Summary of the 30-Study Corpus for the AI Factor

No	Reference	Description
1. AI infrastructure, methods and multimodal integration		
1	Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep Learning for Healthcare: Review, Opportunities and Challenges. *Brief. Bioinform.* 2018, *19*, 1236–1246. https://doi.org/10.1093/bib/bbx044	Broad review of deep learning architectures in healthcare using EHR, omics, and clinical data. Summarizes dominant architectures, data modalities, and practical barriers that now define the AI-in-health "infrastructure layer".
2	Johnson, A.E.W.; Pollard, T.J.; Shen, L.; Lehman, L.W.H.; Feng, M.; Ghassemi, M.; et al. MIMIC-III, a Freely Accessible Critical Care Database. *Sci. Data* 2016, *3*, 160035. https://doi.org/10.1038/sdata.2016.35	Describes the MIMIC-III ICU database, which became a core open data resource for hundreds of AI and predictive modelling studies in critical care. Establishes the data backbone for many later deep learning models.
3	Rajkomar, A.; Oren, E.; Chen, K.; et al. Scalable and Accurate Deep Learning with Electronic Health Records. *npj Digit. Med.* 2018, *1*, 18. https://doi.org/10.1038/s41746-018-0029-1	Demonstrates end-to-end deep learning on raw, FHIR-formatted EHR from multiple hospitals, predicting mortality, readmission, and other outcomes better than

		classical models, with explicit attention to scalability and deployment.
4	Rasmy, L.; Wu, Y.; Wang, N.; et al. A Study of Generalizability of Recurrent Neural Network–Based Predictive Models for Heart Failure Onset Risk Using a Large and Heterogeneous EHR Data Set. *J. Biomed. Inform.* 2018, *84*, 11–16. https://doi.org/10.1016/j.jbi.2018.06.011	Tests RNN-based risk models for heart failure across ~400 hospitals. Shows that deep models can outperform logistic regression but still face substantial generalizability and domain-shift issues across health systems. Highlights that this is one of the most extensive cross-system generalizability evaluations for clinical deep learning models, underscoring the challenge of deploying AI across heterogeneous healthcare networks.
5	Ravì, D.; Wong, C.; Deligianni, F.; Berthelot, M.; Andreu-Perez, J.; Lo, B.; et al. Deep Learning for Health Informatics. *IEEE J. Biomed. Health Inform.* 2017, *21*, 4–21. https://doi.org/10.1109/JBHI.2016.2636665	The article reviews how deep learning architectures are applied to multimodal health data, including imaging, physiological signals, and EHRs. It outlines core methodological advances that enable the modeling of complex clinical patterns. The paper also notes key implementation challenges that shape the development of AI infrastructures in healthcare.
6	Topol, E.J. High-Performance Medicine: The Convergence of Human and Artificial Intelligence. *Nat. Med.* 2019, *25*, 44–56. https://doi.org/10.1038/s41591-018-0300-7	A foundational perspective piece outlining how AI infrastructures, multimodal clinical data, and machine intelligence can augment diagnostic accuracy, workflow efficiency, and precision medicine. The paper is widely recognized as a translational blueprint that specifies policy, workforce, and system-level requirements, and highlights algorithmic reliability and practical challenges for safe AI adoption in healthcare.
7	Beam, A.L.; Kohane, I.S. Big Data and Machine Learning in Health Care. *JAMA* 2018, *319*, 1317–1318. https://doi.org/10.1001/jama.2017.18391	Perspective piece outlining what is required to translate big-data ML into routine care: robust evaluation, integration with clinical workflows, and attention to equity and regulation.
8	Rudin, C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nat. Mach. Intell.* 2019, *1*, 206–215. https://doi.org/10.1038/s42256-019-0048-x	Argues that in high-stakes domains such as healthcare, inherently interpretable models should be preferred over post-hoc explanations of black-box models, thereby reframing how AI should be designed for clinical decision-making.
9	Kline, A.; Wang, H.; Li, Y.; Dennis, S.; Hutch, M.; Xu, Z.; et al. Multimodal Machine Learning in Precision Health: A Scoping Review. *npj Digit. Med.* 2022, *5*, 171. https://doi.org/10.1038/s41746-022-00712-8	Reviews multimodal ML architectures integrating EHR, imaging, and wearable data for precision health. Maps out fusion strategies, evaluation gaps, and open challenges in handling heterogeneous longitudinal signals.
10	Krones, F.; Marikkar, U.; Parsons, G.; Szmul, A.; Mahdi, A. Multimodal Machine Learning Approaches in Healthcare: A Review. *Inf. Fusion* 2025, *115*, 102690. https://doi.org/10.1016/j.inffus.2024.102690	Provides a methodological framework for multimodal fusion (early, late, hybrid) across imaging, physiologic, and clinical data, with healthcare case studies on diagnostics and prognosis. Focuses on model design rather than specific diseases.
2. AI for clinical and physiological diagnostics and risk prediction		
11	Esteva, A.; Kuprel, B.; Novoa, R.A.; et al. Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks. *Nature* 2017, *542*, 115–118. https://doi.org/10.1038/nature21056	Trains a CNN on 129,450 clinical images of skin lesions and shows dermatologist-level performance for key binary tasks (melanoma vs nevi, carcinoma vs seborrheic keratosis), illustrating end-to-end image-based diagnosis.
12	Gulshan, V.; Peng, L.; Coram, M.; et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016, *316*, 2402–2410. https://doi.org/10.1001/jama.2016.17216	Uses thousands of fundus images to train and externally validate a DL system that reaches specialist-level sensitivity and specificity for referable diabetic retinopathy screening.

13	De Fauw, J.; Ledsam, J.R.; Romera-Paredes, B.; et al. Clinically Applicable Deep Learning for Diagnosis and Referral in Retinal Disease. *Nat. Med.* 2018, *24*, 1342–1350. https://doi.org/10.1038/s41591-018-0107-6	Applies a DL architecture to 3-D OCT scans to support real-world referral decisions in ophthalmology. Achieves expert-level referral performance and demonstrates device-independent representations.
14	Attia, Z.I.; Kapa, S.; Lopez-Jimenez, F.; et al. Screening for Cardiac Contractile Dysfunction Using an Artificial Intelligence–Enabled ECG. *Nat. Med.* 2019, *25*, 70–74. https://doi.org/10.1038/s41591-018-0240-2	Trains a CNN on ECG-echocardiogram pairs to detect reduced ejection fraction from standard 12-lead ECGs, enabling low-cost screening for asymptomatic ventricular dysfunction.
15	Hannun, A.Y.; Rajpurkar, P.; Haghpanahi, M.; et al. Cardiologist-Level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network. *Nat. Med.* 2019, *25*, 65–69. https://doi.org/10.1038/s41591-018-0268-3	Applies deep neural networks to long single-lead ambulatory ECGs, achieving cardiologist-level performance in detecting and classifying common arrhythmias, with implications for wearable rhythm monitoring. Notably, the model was trained on one of the most extensive annotated ECG datasets (>90,000 recordings), setting an early benchmark for wearable-compatible arrhythmia detection.
16	Frizzell, J.D.; Liang, L.; Schulte, P.J.; et al. Prediction of 30-Day All-Cause Readmissions in Patients Hospitalized for Heart Failure: Comparison of Machine Learning and Other Statistical Approaches. *JAMA Cardiol.* 2017, *2*, 204–209. https://doi.org/10.1001/jamacardio.2016.3956	Compares multiple machine learning algorithms with traditional statistical models for predicting 30-day readmissions in heart failure patients using extensive registry data. Shows that ML can modestly improve discrimination, but highlights calibration and implementation challenges.
17	Al'Aref, S.J.; Anchouche, K.; Singh, G.; Slomka, P.J.; Kolli, K.K.; Kumar, A.; et al. Clinical Applications of Machine Learning in Cardiovascular Disease and Its Relevance to Cardiac Imaging. *Eur. Heart J.* 2019, *40*, 1975–1986. https://doi.org/10.1093/eurheartj/ehy404	Synthesizes ML applications in cardiovascular disease, summarizing performance claims across diagnosis, prognosis, and risk stratification, and emphasizing the need for external validation and transparent reporting.
18	Li, X.; Dunn, J.; Salins, D.; et al. Digital Health: Tracking Physiomes and Activity Using Wearable Biosensors Reveals Useful Health-Related Information. *PLoS Biol.* 2017, *15*, e2001402. https://doi.org/10.1371/journal.pbio.2001402	The present analysis was among the first to demonstrate that multi-sensor digital biomarkers can signal physiological disturbances before clinical symptoms emerge. Combines continuous multi-sensor wearable data with clinical measures to define "digital physiomes", showing that biosensor-derived features can detect inflammatory and metabolic disturbances and distinguish insulin-resistant from insulin-sensitive individuals.
19	Mishra, T.; Wang, M.; Metwally, A.A.; et al. Pre-Symptomatic Detection of COVID-19 from Smartwatch Data. *Nat. Biomed. Eng.* 2020, *4*, 1208–1220. https://doi.org/10.1038/s41551-020-00640-6	Uses heart rate, steps, and sleep data from consumer smartwatches, combined with anomaly-detection algorithms, to identify COVID-19 infections, often days before symptom onset, illustrating AI-enabled early warning from passive physiological monitoring.
20	Radin, J.M.; Wineinger, N.E.; Topol, E.J.; Steinhubl, S.R. Harnessing Wearable Device Data to Improve Real-Time Surveillance of Influenza-Like Illness. *Lancet Digit. Health* 2020, *2*, e85–e93. https://doi.org/10.1016/S2589-7500(19)30222-5	Aggregates wearable step/heart-rate data to improve real-time influenza-like illness surveillance; demonstrates AI-ready population biosensing.
3. Digital phenotyping, behavior analytics and adaptive interventions		
21	Sano, A.; Picard, R.W. Stress Recognition Using Wearable Sensors and Mobile Phones. In *Proceedings of the IEEE International Conference on Affective Computing and Intelligent Interaction*	Multimodal wearable + smartphone sensing for stress recognition in free-living conditions; shows fusion improves stress classification accuracy.

	(ACII); IEEE: Geneva, Switzerland, 2013; pp. 671–676. https://doi.org/10.1109/ACII.2013.117	
22	Nahum-Shani, I.; Smith, S.N.; Spring, B.J.; Collins, L.M.; Witkiewitz, K.; Tewari, A.; et al. Just-in-Time Adaptive Interventions (JITAIs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Ann. Behav. Med.* 2017, *52*, 446–462. https://doi.org/10.1007/s12160-016-9830-8	Provides a conceptual and methodological framework for JITAIs, specifying how algorithmic decision rules (often implemented via AI/ML) can personalize the timing and content of behavioral health support using continuously collected context and sensor data.
23	Triantafyllidis, A.; Tsanas, A. Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature. *J. Med. Internet Res.* 2019, *21*, e12286. https://doi.org/10.2196/12286	Reviews real-world digital health interventions that embed ML (e.g., for depression management, stress reduction, smoking cessation), emphasizing how predictive models are integrated into feedback, coaching, and adaptive support.
24	Trifan, A.; Oliveira, M.; Oliveira, J.L. Passive Sensing of Health Outcomes through Smartphones: Systematic Review of Current Solutions and Possible Limitations. *JMIR Mhealth Uhealth* 2019, *7*, e12649. https://doi.org/10.2196/12649	Systematic review of smartphone-based passive sensing for health and well-being, summarizing sensors (GPS, accelerometer, app use), target outcomes (physical activity, mental health), and key methodological and engagement challenges.
25	Zhou, W.; Chan, Y.E.; Foo, C.S.; Zhang, J.; Teo, J.X.; Davila, S.; et al. High-Resolution Digital Phenotypes from Consumer Wearables and Their Applications in Machine Learning of Cardiometabolic Risk Markers: Cohort Study. *J. Med. Internet Res.* 2022, *24*, e34669. https://doi.org/10.2196/34669	Derives high-resolution heart-rate features from consumer wearables and applies ML models to estimate cardiometabolic markers and genetic risk scores. Demonstrates that high-resolution digital phenotypes outperform traditional summary metrics, underscoring the added value of fine-grained physiological data captured in free-living conditions.
26	Asare, K.O.; Moshe, I.; Terhorst, Y.; Vega, J.; Hosio, S.; Baumeister, H.; et al. Mood Ratings and Digital Biomarkers from Smartphone and Wearable Data Differentiate and Predict Depression Status: A Longitudinal Data Analysis. *Pervasive Mob. Comput.* 2022, *83*, 101621. https://doi.org/10.1016/j.pmcj.2022.101621	Uses multimodal passive data (smartphone usage, activity, sleep) plus mood ratings to derive digital biomarkers that distinguish and predict depression status over time, showing that joint behavioral and physiologic features improve prediction.
27	Stamatis, C.A.; Meyerhoff, J.; Meng, Y.; Lin, Z.C.C.; Cho, Y.M.; Liu, T.; et al. Differential Temporal Utility of Passively Sensed Smartphone Features for Depression and Anxiety Symptom Prediction: A Longitudinal Cohort Study. *npj Ment. Health Res.* 2024, *3*, 1. https://doi.org/10.1038/s44184-023-00041-y	Large-scale passive sensing study (n≈1000) evaluating how different smartphone features (location, communication, app use) predict depression and anxiety at varying time lags, clarifying when and how digital markers are informative.
28	Bufano, P.; Laurino, M.; Said, S.; Tognetti, A.; Menicucci, D. Digital Phenotyping for Monitoring Mental Disorders: Systematic Review. *J. Med. Internet Res.* 2023, *25*, e46778. https://doi.org/10.2196/46778	Synthesizes digital phenotyping studies for mental disorders, including depression, anxiety, bipolar disorder, and psychosis, summarizing sensor modalities, analytic pipelines, and effect sizes for symptom prediction and monitoring.
29	Zhou, D.; Keogh, J.W.L.; Ma, Y.; Tong, R.K.; Khan, A.R.; Jennings, N.R. Artificial Intelligence in Sport: A Narrative Review of Applications, Challenges and Future Trends. *J. Sports Sci.* 2025, 1–16. https://doi.org/10.1080/02640414.2025.2518694	Narrative review of AI in sport, covering performance monitoring, injury risk prediction, tactical analytics, and athlete profiling. Links sport-specific AI pipelines to broader precision-health and behavior analytics frameworks.
30	Jandoubi, B.; Akhloufi, M.A. Multimodal Artificial Intelligence in Medical Diagnostics. *Information* 2025, *16*, 591. https://doi.org/10.3390/info16070591	Comprehensive review of multimodal AI architectures for medical diagnostics, covering data harmonization, fusion strategies, and evaluation across imaging, EHR, and physiological signals. Emphasizes that multimodal models generally outperform unimodal baselines in diagnostic performance. The authors additionally

		address challenges in interpretability and robustness of multimodal architectures, which are critical for clinical adoption.

Supplementary Table S5. Summary of the 30-Study Corpus for the Decision layer (C) Factor

No	Reference	Description
Domain 1. Decision layer (C) for Health Value, HTA, and Priority Setting		
1	Thokala, P.; Devlin, N.; Marsh, K.; Baltussen, R.; Boysen, M.; Kaló, Z.; et al. Multiple Criteria Decision Analysis for Health Care Decision Making: An Introduction. *Value Health* 2016, *19*, 1–13. https://doi.org/10.1016/j.jval.2015.12.003	Foundational ISPOR Task Force introduction outlining MCDM principles, procedural steps, and its role as a structured alternative to informal deliberation in complex healthcare decision-making. Highlights core components, including criteria construction, scoring, weighting, and transparency.
2	Marsh, K.; Ijzerman, M.; Thokala, P.; Baltussen, R.; Boysen, M.; Kaló, Z.; et al. Multiple Criteria Decision Analysis for Health Care Decision Making: Emerging Good Practices. *Value Health* 2016, *19*, 125–137. https://doi.org/10.1016/j.jval.2015.12.016	Good-practice guidance summarizing methodological standards for MCDM in healthcare, covering criteria definition, evidence structuring, weighting approaches, uncertainty analysis, and reporting principles for robust implementation.
3	Angelis, A.; Kanavos, P. Multiple Criteria Decision Analysis for Evaluating New Medicines in Health Technology Assessment and Beyond: The Advance Value Framework. *Soc. Sci. Med.* 2017, *188*, 137–156. https://doi.org/10.1016/j.socscimed.2017.06.024	Introduces the Advanced Value Framework, an MCDM-based structure integrating clinical, economic, ethical, and societal dimensions to enhance transparency and consistency in HTA appraisals for new medicines.
4	Oliveira, M.D.; Mataloto, I.; Kanavos, P. Multi-Criteria Decision Analysis for Health Technology Assessment: Addressing Methodological Challenges to Improve the State of the Art. *Eur. J. Health Econ.* 2019, *20*, 891–902. https://doi.org/10.1007/s10198-019-01052-3	Methodological synthesis detailing how MCDM frameworks organize evidence inputs in HTA, examining challenges in criteria formalization, weighting logic, and evidence translation. Emphasizes the need for improved robustness, transparency, and reproducibility in contemporary MCDM applications.
5	Góngora-Salazar, P.; Rocks, S.; Fahr, P.; Rivero-Arias, O.; Tsiachristas, A. The Use of Multi-Criteria Decision Analysis to Support Decision-Making in Healthcare: An Updated Systematic Literature Review. *Value Health* 2023, *26*, 780–790. https://doi.org/10.1016/j.jval.2022.11.007	Updated systematic review mapping MCDM applications across healthcare decision contexts. Summarizes dominant methodological families (e.g., AHP, VIKOR), types of decisions supported, and the extent of stakeholder involvement.
6	Dai, Z.; Xu, S.; Wu, X.; Hu, R.; Li, H.; He, H.; et al. Knowledge Mapping of Multicriteria Decision Analysis in Healthcare: A Bibliometric Analysis. *Front. Public Health* 2022, *10*, 895552. https://doi.org/10.3389/fpubh.2022.895552	Bibliometric study charting the evolution of MCDM research in healthcare. Identifies key thematic clusters (HTA, priority setting, resource allocation) and illustrates structural patterns and research fronts shaping methodological development.
7	Marsh, K.; Lanitis, T.; Neasham, D.; Orfanos, P.; Caro, J. Assessing the Value of Healthcare Interventions Using Multi-Criteria Decision Analysis: A Review of the Literature. *Pharmacoeconomics* 2014, *32*, 345–365. https://doi.org/10.1007/s40273-014-0135-0	Systematic review synthesizing early MCDM applications for valuing healthcare interventions, detailing criteria architectures, evidence-integration approaches, and practical implementation considerations that informed later developments in HTA-focused MCDM models.
8	Baltussen, R.; Niessen, L. Priority Setting of Health Interventions: The Need for Multi-Criteria Decision Analysis. *Cost Eff. Resour. Alloc.* 2006, *4*, 14. https://doi.org/10.1186/1478-7547-4-14	Conceptual paper articulating the rationale for applying MCDM in health-intervention priority setting. Highlights the limitations of conventional approaches and argues for transparent, structured, criteria-based evaluation frameworks.
9	Sussex, J.; Rollet, P.; Garau, M.; Schmitt, C.; Kent, A.; Hutchings, A. A Pilot Study of Multicriteria	Pilot MCDM application for orphan drug appraisal incorporating rarity, unmet need, uncertainty and other

	Decision Analysis for Valuing Orphan Medicines. *Value Health* 2013, *16*, 1163–1169. https://doi.org/10.1016/j.jval.2013.10.002	context-specific value dimensions into a structured assessment framework.
10	Iskrov, G.; Miteva-Katrandzhieva, T.; Stefanov, R. Multi-Criteria Decision Analysis for Assessment and Appraisal of Orphan Drugs. *Front. Public Health* 2016, *4*, 214. https://doi.org/10.3389/fpubh.2016.00214	Develops and applies an MCDM model tailored to orphan drug appraisal, demonstrating how broader value components can be systematically structured, weighted and compared in transparent decision processes.
Domain 2. MCDM for clinical technologies, treatments and system-level interventions		
11	Elezbawy, B.; Fasseeh, A.N.; Németh, B.; Gamal, M.; Eldebeiky, M.; Refaat, R.; et al. A Multicriteria Decision Analysis (MCDA) Tool to Purchase Implantable Medical Devices in Egypt. *BMC Med. Inform. Decis. Mak.* 2022, *22*, 289. https://doi.org/10.1186/s12911-022-02025-y	Develops an MCDM tool for national-level procurement of implantable medical devices, integrating clinical benefit, safety, cost, and organizational considerations into a structured evaluation process.
12	Pimenta, J.F.; Vieira, A.C.L. Selection and Prioritization of Medical Devices for HTA Evaluation: A Systematic Review of Existing Approaches. *Appl. Health Econ. Health Policy* 2025, *23*, 1005–1025. https://doi.org/10.1007/s40258-025-00981-w	Systematic review mapping approaches for selecting and prioritizing medical devices for HTA. Highlights widespread application of MCDM, heterogeneity in criteria, and persistent methodological gaps in prioritization frameworks.
13	Beny, K.; Dubromel, A.; du Sartz de Vigneulles, B.; Gay, V.; Carrouel, F.; Negrier, C.; et al. Multiple Criteria Decision Analysis for Therapeutic Innovations in a Hemophilia Care Center: A Pilot Study of the Organizational Impact of Innovation in Hemophilia Care Management. *PLoS ONE* 2022, *17*, e0273775. https://doi.org/10.1371/journal.pone.0273775	Applies MCDM to assess organizational impacts of therapeutic innovations in a hemophilia care center, extending evaluation beyond clinical effectiveness to workflow and resource considerations.
14	Sir, E.; Batur Sir, G.D. Evaluating Treatment Modalities in Chronic Pain Treatment by the Multi-Criteria Decision Making Procedure. *BMC Med. Inform. Decis. Mak.* 2019, *19*, 191. https://doi.org/10.1186/s12911-019-0925-6	Uses MCDM to compare chronic pain treatment options across multiple clinical and patient-centred criteria, highlighting trade-offs between efficacy, safety, and convenience.
15	Mishra, V.; Singh, J. Health Technology Assessment of Telemedicine Interventions in Diabetes Management: Evidence from UAE. *FIIB Bus. Rev.* 2022. https://doi.org/10.1177/23197145221130651	Uses an MCDM approach (AHP-Entropy) for HTA of telemedicine interventions in diabetes management, comparing in-person care, hybrid care and pure telecare across cost-effectiveness, clinical effectiveness, safety and ease-of-use criteria.
16	Alamoodi, A.H.; Zaidan, B.B.; Albahri, O.S.; Garfan, S.; Ahmaro, I.Y.Y.; Mohammed, R.T.; et al. Systematic Review of MCDM Approach Applied to Medical Case Studies of COVID-19. *Complex Intell. Syst.* 2023, *9*, 4705–4731. https://doi.org/10.1007/s40747-023-00972-1	Reviews MCDM applications in COVID-19 case studies, including triage, treatment selection, hospital logistics and telemedicine decision-making during pandemic conditions.
17	Baltussen, R.; Jansen, M.P.M.; Bijlmakers, L.; Grutters, J.; Kluytmans, A.; Reuzel, R.P.; van der Wilt, G.J. Value Assessment Frameworks for HTA Agencies: The Organization of Evidence-Informed Deliberative Processes. *Value Health* 2017, *20*, 256–260. https://doi.org/10.1016/j.jval.2016.11.019	Describes evidence-informed deliberative processes to integrate multiple value dimensions in HTA; operationalises MCDA principles in institutions.
18	Gaievskyi, S.; Delfrate, N.; Ragazzoni, L.; Bahattab, A. Use of Multi-Criteria Decision Analysis to Support Decision-Making during Health Emergencies: A Scoping Review. *Front. Public*	Scoping review of MCDM applications in health emergencies and disaster response, mapping decision types and methodological patterns across crises.

	Health 2025, *13*, 1584026. https://doi.org/10.3389/fpubh.2025.1584026	
19	Băcescu Ene, G.-V.; Stoia, M.-A.; Cojocaru, C.; Todea, D.A. SMART Multi-Criteria Decision Analysis (MCDA)—One of the Keys to Future Pandemic Strategies. *J. Clin. Med.* 2025, *14*, 1943. https://doi.org/10.3390/jcm14061943	Proposes a SMART-based MCDM framework for pandemic strategy development, integrating epidemiological and socio-economic criteria to prioritize interventions and enhance system resilience.
20	Wang, G.; Shao, Q. Design of a Smart Medical Service Quality Evaluation System Based on a Hybrid Multi-Criteria Decision Model. *Sci. Rep.* 2024, *14*, 26407. https://doi.org/10.1038/s41598-024-71224-6	Proposes a hybrid MCDM model to evaluate innovative medical services across quality, safety, and efficiency criteria, thereby supporting the improvement of digital and telehealth systems.
Domain 3. MCDM for digital health, self-care, remote monitoring, and AI-linked prioritization		
21	Kapun, M.M.; Hlebec, V.; Zajc, M.; Kramberger, T. A Multi-Criteria Decision Model for Assessing the Health and Self-Care of Patients with Chronic Diseases in the Home Environment. *Cent. Eur. J. Oper. Res.* 2023, *31*, 1147–1176. https://doi.org/10.1007/s10100-022-00823-3	Develops a DEX-based MCDM model that integrates symptoms, functional status, and self-care behaviors to classify chronic patients' home health status and support timely self-management decisions.
22	Silva-Aravena, F.; Gutiérrez-Bahamondes, J.H.; Núñez Delafuente, H.; Toledo-Molina, R.M. An Intelligent System for Patients' Well-Being: A Multi-Criteria Decision-Making Approach. *Mathematics* 2022, *10*, 3956. https://doi.org/10.3390/math10213956	Applies an MCDM framework to compute patient well-being and vulnerability scores based on clinical and contextual indicators, enabling prioritization of chronic patients on follow-up waiting lists.
23	Almahdi, E.M.; Zaidan, A.A.; Zaidan, B.B.; Alsalem, M.A.; Albahri, O.S.; Albahri, A.S. Mobile-Based Patient Monitoring Systems: A Prioritization Framework Using Multi-Criteria Decision-Making Techniques. *J. Med. Syst.* 2019, *43*, 219. https://doi.org/10.1007/s10916-019-1339-9	Presents a hybrid MCDM framework that evaluates and ranks mobile patient-monitoring systems using clinical, technical, and usability criteria to guide telemedicine deployment decisions.
24	Mohammed, K.I.; Jaafar, J.; Zaidan, A.A.; Albahri, O.S.; Zaidan, B.B.; Abdulkareem, K.H.; et al. A Uniform Intelligent Prioritization for Solving Diverse and Big Data Generated from Multiple Chronic Diseases Patients Based on Hybrid Decision-Making and Voting Method. *IEEE Access* 2020, *8*, 91521–91530. https://doi.org/10.1109/ACCESS.2020.2994746	Proposes a hybrid decision-making and voting model (HDMVM) that aggregates outputs from multiple MCDM methods to generate stable, uniform prioritization of multi-morbidity patients under big-data conditions.
25	Mohammed, K.I.; Zaidan, A.A.; Zaidan, B.B.; Albahri, O.S.; Albahri, A.S.; Jasim, A.N. A Novel Technique for Reorganisation of Opinion Order to Interval Levels for Solving Several Instances Representing Prioritization in Patients with Multiple Chronic Diseases. *Comput. Methods Programs Biomed.* 2020, *185*, 105151. https://doi.org/10.1016/j.cmpb.2019.105151	Introduces the TROOIL method, which converts expert opinion orders into interval levels and applies weighting and ranking procedures to improve stability and accuracy of prioritization for multi-morbidity patients.
26	Ortiz-Barrios, M.; Cleland, I.; Donnelly, M.P.; Gul, M.; Yucesan, M.; Jiménez-Delgado, G.I.; et al. Integrated Approach Using Intuitionistic Fuzzy Multicriteria Decision-Making to Support Classifier Selection for Technology Adoption in Patients with Parkinson Disease: Algorithm Development and Validation. *JMIR Rehabil. Assist. Technol.* 2024, *11*, e57940. https://doi.org/10.2196/57940	Uses intuitionistic fuzzy MCDM to evaluate and select machine-learning classifiers for digital Parkinson's biomarkers, integrating algorithmic performance with clinical relevance and usability criteria.

27	Haktanır, E.; Kahraman, C. A Novel Picture Fuzzy CRITIC and REGIME Methodology: Wearable Health Technology Application. *Eng. Appl. Artif. Intell.* 2022, *113*, 104942. https://doi.org/10.1016/j.engappai.2022.104942	Develops a picture-fuzzy CRITIC–REGIME MCDM approach to assess wearable health technologies by combining indicators of performance accuracy, comfort, usability, and cost.
28	Yeh, T.-M.; Lu, H.-Y.; Pai, F.-Y. Applying Multi-Criteria Decision Analysis Methods to Explore the Key Factors in Using Interactive Intelligent Health Promotion Equipment. *SAGE Open* 2025, *15*, 21582440251327474. https://doi.org/10.1177/21582440251327474	Applies DEMATEL to identify relationships and weights among adoption-related factors for intelligent health-promotion equipment, highlighting direct and indirect influences on user acceptance.
29	Sánchez-Martínez, F.I.; Abellán-Perpiñán, J.M.; Martínez-Pérez, J.E.; Gómez-Torres, J.L. Design of a Multiple Criteria Decision Analysis Framework for Prioritizing High-Impact Health Technologies in a Regional Health Service. *Int. J. Technol. Assess. Health Care* 2024, *40*, e21. https://doi.org/10.1017/S0266462324000205	Develops/validates MCDA framework for prioritising high-impact technologies in a regional health service with explicit criteria and elicitation procedures.
30	Li, T.; Sun, J.; Fei, L. Application of Multiple-Criteria Decision-Making Technology in Emergency Decision-Making: Uncertainty, Heterogeneity, Dynamicity, and Interaction. *Mathematics* 2025, *13*, 731. https://doi.org/10.3390/math13050731	Reviews MCDM applications in emergency decision-making, analyzing how uncertainty, heterogeneity, dynamic change, and criterion interactions are handled, and outlining methodological gaps.