Can wearable technologies contribute to an age-friendly walkability environment? First insights from a systematic review of the literature

Frida Milella, Michela Oltolini, Stefania Bandini
University of Milano-Bicocca, Department of Informatics, Systems and Communication (DISCo)
Milano, Italy

ABSTRACT

Research on wearable devices and sensors applied to older adults has witnessed significant growth, primarily focusing on health monitoring but also on researching real-world applications of urban walking, specifically, walkability. The use of pedestrians' bodily responses to the environment, known as the people-centric sensing strategy, has been demonstrated to be more suitable for detecting challenging environmental conditions and enhancing walkability. However, the extent to which such data may accurately pinpoint environmental distress in a particular demographic, such as older adults, has not been thoroughly investigated. The purpose of this in-progress systematic review of the literature is to outline the strategy employed and highlight some preliminary findings on the use of wearable sensors or sensor-based technologies in gathering the bodily responses of older adults and/or their family caregivers to stimuli originating from real urban walking scenarios. Our preliminary findings showed that the current research on using wearable devices to detect bodily responses in older adults (or informal caregivers) population in relation to walkability in outdoor environments tends to exhibit a uniform strategy, that is, collecting physiological and location data from older adults through controlled outdoor walking routes using wrist-wearable and GPS; training supervised classifiers to differentiate between physiological stress and non-stress signals to environmental conditions of external interaction; and finally, using hotspot analysis to group together individual physiological responses in areas with high-stress interactions with the external environment using GIS. Although the body of literature appears to be still in its early stages, using wearable sensors and a GIS-based approach could be a promising method for spatio-temporally capturing people's direct bodily responses to the environmental stressors, and this offer rooms for potential integration of simulation-based models into Digital Twins environments with GIS-based analysis to further enhance our understanding of older adults mobility in urban settings. However, our preliminary findings indicates a gap in the research regarding the detection of bodily responses to stimuli on outdoor walking paths using wearable sensors, specifically among informal carers.

PETRA '24, June 26–28, 2024, Crete, Greece

ACM ISBN 979-8-4007-1760-4/24/06
https://doi.org/10.1145/3652037.3652046

CCS CONCEPTS

• **Human-centered computing → Ambient intelligence**.

KEYWORDS

Wearable sensors, walkability, elderly, informal caregiver, outdoor environment, bodily responses, active learning

ACM Reference Format:
Frida Milella, Michela Oltolini, Stefania Bandini. 2024. Can wearable technologies contribute to an age-friendly walkability environment? First insights from a systematic review of the literature. In *The PErvasive Technologies Related to Assistive Environments (PETRA) conference (PETRA '24), June 26–28, 2024, Crete, Greece.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3652037.3652046

1 INTRODUCTION

Research on wearable devices and sensors applied to older adults has witnessed significant growth, primarily focusing on health monitoring but also on researching real-world applications of urban walking, specifically, walkability (e.g., [2, 3, 41]). This emerging field aims to improve the quality of life for older adults by monitoring their movements and collecting valuable data to gain a deeper understanding of their perception of the environment (e.g., [19]). Indeed, the increasing number of older adults in our society has brought attention to the issue of senior mobility [14]. Older adults mobility, referring to an individual's ability to achieve access to preferred people or locations [29, 34], is a crucial determinant of their quality of life [29]. As individuals age, their mobility tends to decline, leading to increased reliance on urban services that are easily accessible by walking [4, 31]. Therefore, an age-friendly environment should prioritise the provision of urban facilities that are conveniently located within a walkable distance [4]. Nevertheless, a crucial first step in promoting a walkable environment is to identify environmental issues that may limit a person's ability to access and utilise the outdoor environment [16, 44]. The use of pedestrians' bodily responses to the environment, known as the people-centric sensing strategy [16, 44], has been demonstrated to be more suitable for detecting challenging environmental conditions and enhancing walkability [26, 44]. Bodily responses refer to the natural and involuntary physiological, behavioural, and cognitive reactions that humans display when they interact with a stimuli in their environment [44]. Recently, there has been a growing body of research that has utilised wireless sensors worn on the wrist to monitor physiological signals of humans in real-world environments (e.g.,[11, 42]) [43]. However, the extent to which such data may accurately pinpoint environmental distress in a particular demographic, such as older adults, has not been thoroughly investigated [1].

The purpose of our systematic review is to investigate the use of wearable sensors or sensor-based technologies in gathering the bodily responses of older adults and/or their family caregivers to stimuli originating from real urban walking scenarios. A previous systematic literature review [35] found that sensor-based technologies are primarily used indoors as part of an intervention approach to support older adults in aging in place and family caregivers in their supervision duties. Moreover, a similar review [36] found that the use of advanced technological interventions to create age-friendly outdoor areas can provide further assistance to informal carers who support adults in need of care. Hence, incorporating informal carers, also known as unpaid or family caregivers, into our analysis's target population could provide valuable insights into the ongoing discourse on age-friendly environments. Accessibility to care is indeed currently being discussed in a user-centered context and is being integrated with a personal mobility dimension [30] that refers to measures of the age-friendliness of the urban environment based on the concept of walkability [4]. Therefore, learning about how the current research address the use of wearable sensors or sensor-based technology to collect physiological, behavioral, or cognitive responses of older adults and/or their family caregivers throughout outdoor walks may offer also valuable insights for enhancing the physical accessibility to care. Since this study is currently in the process of reviewing existing literature, it will outline the strategy employed for the systematic review and highlight some preliminary findings. The paper is structured as follows: Section 2 describes the methodology used in this study; Section 3 presents the preliminary results of the study, while Section 4 offers a concise summary of the main findings and some closing remarks. Section 5 describes future steps for the review study.

2 METHODS

This section outlines the methodology used to carry out this study. A systematic review of the literature (SRL) was chosen as review approach to address the following research questions, as it is a commonly employed approach for identifying, evaluating, and combining all available evidence pertaining to a certain research question with explicit and systematic methods [17, 21]. The review protocol was defined following the guidelines by Moher et al [37]; as the review of the literature is still in progress, this study will only outline the methodological steps required to present the preliminary findings currently available.

2.1 Research questions

The following research questions were formulated to aid in conducting the literature review for this study:

(1) *What are the main studies on wearable sensors or sensor-based technology utilized for gathering data from older adults or their family caregivers in real urban walking scenarios?*;
(2) *What are the main stimuli and responses collected by wearable sensors or sensor-based technologies that are currently used to study the personal mobility of older adults or their informal caregivers?*

2.2 Search strategy

The SLR was conducted using the Scopus database that is currently the largest existing multidisciplinary database [10]. The question regarding the essential databases for retrieving all pertinent references in a systematic review remains unresolved [8]. Prior research has examined the additional benefit of multiple databases in review studies [8]: some studies have investigated the sufficiency of searching a single database, and have found that searching multiple databases does not impact the results (e.g., [48]), while alternative research has determined that depending solely on one database is insufficient for obtaining all the essential references for systematic reviews (e.g.,[7]). Developing a search strategy involves balancing comprehensiveness and relevancy, but the decision on the level of effort put into the search strategy also depends on the topic of the review [20]. Considering the cross-relevance of our topic in multiple field, employing a bibliographic database that encompasses multiple fields can be deemed suitable. Nonetheless, the Scopus database covers a wide range of highly-cited documents in the field of Engineering and Computer Science [33], making it suitable for researching wearable technologies in the context of human health monitoring [12]. The first reviewer (FM) conducted the search on the Scopus database using the final search string on December 26, 2023. The database search was conducted on the title, abstract, and keywords, with a restriction to publications from 2000 to 2023, including only journal articles and conference papers. The search was also restricted to articles published in the English language. A total of 4214 articles were found. Table 1 displays the search string employed for querying the digital source.

Table 1: Search query

Query	Database
((TITLE-ABS-KEY("wearable sens*") OR TITLE-ABS-KEY("wearable") OR TITLE-ABS-KEY("sens*"))) AND ((TITLE-ABS-KEY("physiolog*") OR TITLE-ABS-KEY("behav*") OR TITLE-ABS-KEY("cognitive") OR TITLE-ABS-KEY("physiolog* response*") OR TITLE-ABS-KEY("behav* response*") OR TITLE-ABS-KEY("cognitive response*") OR TITLE-ABS-KEY("bodily response*"))) AND ((TITLE-ABS-KEY("outdoor*") OR TITLE-ABS-KEY("outdoor environment") OR TITLE-ABS-KEY ("built environment") OR TITLE-ABS-KEY("environment") OR TITLE-ABS-KEY("outdoor mobility") OR TITLE-ABS-KEY("mobility") OR TITLE-ABS-KEY("walkability") OR TITLE-ABS-KEY("walk*"))) AND ((TITLE-ABS-KEY("older adult*") OR TITLE-ABS-KEY("older people") OR TITLE-ABS-KEY("elderly") OR TITLE-ABS-KEY("senior*") OR TITLE-ABS-KEY("informal caregiver*") OR TITLE-ABS-KEY("unpaid caregiver*")))	Scopus

2.3 Selection criteria

Criteria for inclusion and exclusion were chosen according to the research goals of the study. The studies included in the subsequent analysis were selected based on their adherence to the following inclusion criteria:

(1) The target population was older adults (aged over 65) or informal caregivers;

(2) The article used wearable-sensors or sensor-based technologies to gather bodily responses (ie physiological, behavioural or cognitive responses) of the target population against walkability in a real outdoor environment.

Conversely, the articles were excluded if they met any of the subsequent exclusion criteria:

(1) The targeted population was exclusively composed of people aged less than 65 or without any caregiving responsibility;
(2) The article exclusively employed non-wearable or non-sensor-based technologies;
(3) The article measured bodily responses which are not strictly responses to stimuli derived from the walkable outdoor environment or for the purpose of getting insights on walkable outdoor settings;
(4) The article measured the bodily responses of the target population against walkability not in a real outdoor environment;
(5) The article focused on the usability or qualitative appraisal of the use of wearable-sensors or sensor-based technologies even though in the context of interest;
(6) Articles were reviews, book proceedings, theoretical frameworks, editorial, letters.

2.4 Study selection

The first reviewer (FM) retrieved all the detected articles from the database and uploaded them to the web-based tool Rayyan [38] to facilitate the process of identifying and removal of duplicates. Duplicates were removed using the systematic auto resolver, an AI-based function integrated into Rayyan that automates the deduplication process according to the reviewer-defined criteria for resolving duplicates. The resolving criteria for this study included the title, authors, and DOI, which remained unchanged throughout the process. The minimal threshold for articles similarity was first set at 95%, and later raised to 100% to conduct two consecutive iterations. Eighteen remaining articles that did not meet the deletion criteria were manually reviewed for similarity.

Title and abstract screening phase of the study selection process was then executed via ASReview (version 1.4) [46], an AI tool-assisted screening system using active learning approach. Active learning is a machine learning approach where a model has the ability to select specific data points, such as records obtained through systematic search, for learning purposes [27]. While traditional machine learning approaches require previously labelled data, active learning actively questions the user to classify a desirable and relevant subset of the data, achieving comparable performance with less labelled data [40]. Active learning models are particularly suitable for the screening step in systematic reviews because of the availability of huge unlabeled datasets [18], as proven by their growing use in academic literature (e.g., [23, 24, 49, 51]). A recent research by Wang et al. [50] revealed that human reviewers have an error rate of 10.76% during abstract screening (equivalent to approximately one error in every nine abstracts), with variations ranging from 5.76% to 21.11% across different clinical areas and question types. Therefore, automating or semi-automating the process of title and abstract screening is a potentially beneficial step in the review process [5, 45] in light of the error-prone [46] and time-consuming nature of systematic reviewing [9, 46]. ASReview employs an active research-in-the-loop machine learning algorithm to prioritize articles based on their likelihood of meeting the inclusion criteria using text mining [47]. When used alongside well-selected stopping criteria, it can effectively decrease the number of papers that need to be reviewed, while also minimizing the risk of overlooking any important studies [52]. Therefore, implementing active learning in screening prioritizing enables substantial time savings [18] and enhancing the overall quality of the review screening [23], since the reviewer can choose to halt screening once a sufficient number of relevant publications have been identified [53].

Two reviewers worked independently on the title and abstract screening phase, which was organized into three steps. First, the first reviewer uploaded the set of articles with duplicates removed to EndNote (version 2.4) for the purpose of a cross-verification, that is a critical step prior to employing the AI tool for the filtering of titles and abstracts [47]. No further deduplication was required. However, an additional 46 records, which had already been classified as document types to be excluded in Scopus, were eliminated. This additional cross-check process against the exclusion criteria was necessary due to Rayyan's no option for selecting exclusion fields other than keywords. The resulting dataset articles was uploaded to ASreview.

Second, the algorithm was trained by the first reviewer using a set of labeled examples. This collection consisted of 7 papers that were classified as "relevant" and an additional set of 10 papers that were classified as "irrelevant". The anchoring documents were identified during the initial search conducted in Scholar on the topic, while the irrelevant papers were selected from the ones randomly suggested by ASReview based on the uploaded dataset. Once the prior knowledge was established for training, the first reviewer implemented the active learning model by selecting a feature extraction approach, a classifier, a query strategy, and a balance strategy. Considering the low computation time [46], the default settings were chosen. This includes using the frequency-inverse document frequency (TF-IDF) for the feature extraction strategy, the naive-Bayes as classifier, the certainty-based sampling (maximum in ASReview [46]) for the query strategy and the dynamic resampling to deal with the unbalancedness of the data. After being trained using labeled examples, ASReview generated an initial ranking of unlabelled articles based on their probability of relevance, with the highest-ranked articles being the most likely to be relevant [47]. The first reviewer evaluated and categorized the highest-ranked article according to the inclusion criteria (see Section 2.3). Hence, a new model was trained based on those additional information, subsequently resulting in a updated ranking and the identification of a new top-ranked article the first reviewer again evaluated and categorized for its pertinence. The process of AI generating rankings and the first reviewer making decisions was iterated until a specified data-driven halting condition of consecutive irrelevant articles was met. A user-specified optimal halting criterion has not been established in the literature; however, a considerable number of the articles retrieved on the subject employ values varying from 50 (e.g., [6, 23, 49]) to 100 consecutive irrelevant articles (e.g., [47]) or set the cut-off value between 5% (e.g., [51]) and 10% (e.g., [52]) of the dataset. In light of the substantial amount of data generated by the search string, a 10% threshold value (394 consecutive excluded papers) was deemed suitable for determining when to stop the process, as it might be an acceptable trade-off between screening efficiency and accuracy. The stopping criterion was met when 862 articles (22% of the total dataset) were labeled.

Third, to ensure the accuracy of the selection process, the second reviewer (MO) performed a cross-check of the selection process using 20% of the labeled articles (172 articles) that were extracted at random and examined for title and abstract, as suggested in other studies on the topic [23, 47]. Inter-reviewer agreement in study identification was assessed using Cohen's kappa statistic [13], which quantifies the level of agreement beyond what would be expected by chance alone [15]. This statistic is commonly employed when two raters are involved and there are two categories [32]. The inter-rater agreement was 0.988, with a Cohen's kappa coefficient of 0.894 (z-score = 11.8, p=0), indicating a very good level of agreement [13]. Hence, it can be inferred that the algorithm was properly trained by the first reviewer. The analysis was performed in RStudio (version 4.2.1).

The first reviewer and the second reviewer separately analyzed the whole text of the publications that could potentially contribute to the review to assess their pertinence against the inclusion criteria. Disagreements on the inclusion of studies were discussed until a consensus was reached. The full-text screening was performed in Rayyan to streamline the subsequent comparison of the selection process for relevant articles.

3 RESULTS

A total of 4214 articles was retrieved through the search in Scopus database. From this, 225 duplicate articles were excluded using the AI-based auto resolver integrated into Rayyan and additional 46 articles were eliminated through the cross-check process against the exclusion criteria in EndNote. A total of 28 articles were deemed eligible for the inclusion using ASReview. After full text screening, a total of 7 articles were included in the analysis. Figure 1 shows the flow chart of the study selection process.

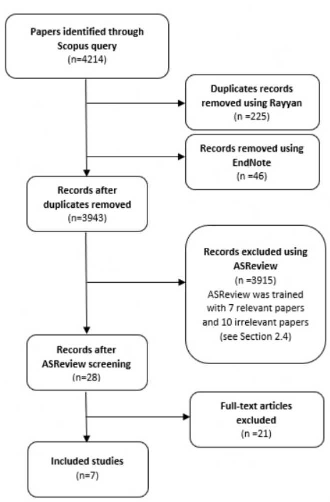

Figure 1: The flow chart shows the selection procedure of the included papers.

A preliminary analysis of the included papers was conducted to gain initial insights into the two research questions (see Section 2.1). First, our analysis revealed a significant similarity in the scope of the studies. The largest proportion of the included articles (5 out of the 7) mostly pertained to the older adults stressful interaction with their physical environment [29, 43], as well as the use of wearable-based collective sensing to identify the environmental barriers encountered by older adults [28] or the environmental constraints associated with the mobility of older people [1, 42]. However, two of the included studies expanded the rationale of the research including the effects of outdoor exposure also on subjective well-being and cognitive functioning [39] or behavioural responses [44]. From an overall perspective, the publication timeline spans only four years, from 2019 to 2022. In 2022, 43.9% of the chosen articles were published, while there was a significant peak in 2020 with 28.6% of the total publications. Furthermore, it is noteworthy that most of the articles share the same first authors [28, 29, 42–44], with only two articles having different first authors [1, 39].

Second, the main data collected by wearable sensors used in studying the personal mobility of target users consists of physiological response data, such as electrodermal activity (EDA) [1, 28, 29, 43, 44], gait patterns [1], photoplethysmogram (PPG) [29, 42–44], and electroencephalography (EEG) [44]. However, most of the research (4 out of 7) combine visual sensing data [1], location [29, 42, 43] and environmental data [43] to explore the potential of using multi-modal data to capture and evaluate environmental distress related to the mobility of older persons in urban areas (e.g., [1, 29]) or the influence of visuospatial configurations of urban space on older adults' physiological stress (e.g., [42]). One of the included study integrates physiologic health data with real-time air quality and noise data to analyse the activation of stress levels in older individuals while walking in neighbourhoods with varying spatial and environmental conditions [39]. In the majority of the articles (5 out of 7) physiological signals, including PPG and EDA, were collected using wrist-worn biometric sensors. The sensors used for this purpose included the Empatica E4 [1, 42–44] and an in-house built app android smartwatch [39] with EDA data sampled at 4Hz [1, 29, 44], and PPG data at 64 Hz [29, 42] or at 100Hz [39]. EEG data were recorded during experimental outdoor environmental walks using the wearable EMOTIV EPOC+ headset at a frequency of 128 Hz [44]. Additionally, data on foot plantar pressure distribution, contact forces, and three-axis acceleration were collected using the Moticon OpenGo insole sensor at a frequency of 50 Hz [44], while the GPS data were collected using a smartphone [1, 39] or wore a belt-clip-type GPS sensor Qstarz BT-Q1000XT [42–44] at 1Hz [42, 43]. However, two out of the seven articles did not specify the smart wristband used to monitor the physiological responses from older adults and the sensor used to gather GPS data [28, 29].

Third, four out of seven articles employed hot-spot analysis, a Geographic information system (GIS)-based mapping clustering tool, to pinpoint high-demanding locations along the path where multiple older adults' physiological responses were recorded during outdoor walks using their respective GPS coordinates [29, 44]; to spatially correlate hot spots and cold spots with older adults' perceived responses, in order to identify stressful person-environment interactions caused by spatial factors [42]; to identify locations with high-risk clusters of stressful interactions, thus identifying risk stress hotspots for older adults through simulation-based risk hotspot analysis [43]. Training of machine learning classifiers was also integrated to predict individual stress levels based on physiological signals [29], as well as location and environmental data such as temperature and humidity [43], by including the informativeness of the bodily responses [44]. One of the seven chosen articles examines the impact of spaziovisual indicators on stress and non-stress physiological responses in older adults using Self-Organizing Maps (SOM) [42].

Fourth, none of the chosen articles specifically focus on informal carers who provide support to older adults, as the study samples solely consisted of older adults. The study's sample sizes are typically small, averaging around 10 participants [28, 42–44]. The lower bound is 9 [1], while the upper bound is 11 [39], with the exception of the work by [29] that recruited 30 participants. Three articles focus on the elderly population aged over 65 [42–44], with one study [39] reporting an average age of 65 years. Only three of the included articles do not provide specific age information for older adults participants [1, 28, 29].

4 DISCUSSION AND CONCLUSIONS

This section examines some specific issues that have arisen in the ongoing literature study and draws the first conclusions.

Firstly, our initial findings indicate that the current research on using wearable devices to detect bodily responses in older adults (or informal caregivers) population in relation to walkability in outdoor environments is still in its early stages. The selected scientific literature on the topic appear to be limited, with only 7 identified contributions, even though there has been an observable increase in research output, particularly from 2022 onwards. Most of the articles analysed in this ongoing review exhibit a minimal variation in their research goals, which mainly focus on the use of wearable sensors or sensor-based technologies to identify stressful person-environment interactions that may restrict the mobility of older adults in the built environment [1, 28, 29, 42, 43]. Furthermore, the analyzed articles exhibit a uniform strategy, that is, collecting physiological and location data from older adults through controlled outdoor walking routes using wrist-wearable and GPS; training supervised classifiers to differentiate between physiological stress and non-stress signals to environmental conditions of external interaction; and finally, using hotspot analysis to group together individual physiological responses in areas with high-stress interactions with the external environment using GIS [29, 42–44]. As an example, in their study, Torku et al. [43] conducted a field experiment to measure the stress levels of older adults while interacting with the built environment. They gathered data from participants' physiological responses, location, and environment using wearable sensors and applying a multimodal information fusion technique (specifically, feature level fusion of EDA signal, PPG signal, location and environmental data using parametric or non-parametric machine learning algorithms) [43] to reduce errors in data collection and provide additional insights into how humans interact with their surroundings in real-life situations. Several machine learning algorithms were trained and tested to detect the older adults' stressful interactions from physiological signals, location and environmental data, with Ensemble bagged tree achieving the best performance (98.25% accuracy). Using the spatial relative risk (SRR) function to identify clusters of high-stress hotspots that present an increased risk to older adults, a simulation-based risk hotspot analysis was employed to identify environmental barriers within the study area that pose a high risk to older adults [43]. The utilisation of a simulation-based methodology exhibits encouraging outcomes in producing physiological point-level data that accurately represents the entire study area. Researchers and urban planners will benefit from the ability to identify risk hotspot locations with high statistical power in order to identify actual urban stress hotspots that present a greater risk to ageing adults and to comprehend the relationship between the built environment and stress [43]. Furthermore, Torku et al [42] specifically examined how the visuospatial arrangements of urban environments impact the physiological stress levels of older adults. The experimental study gathered physiological and perceived stress responses of senior adults as they walked through an urban environment. By employing spatial clustering and hot spot analysis, areas exhibiting concentrations of physiological responses induced by spatial factors were identified and classified as stress-free or stress-inducing, according to the perception of stress among the participants. Isovist analysis was employed to extract the perceived visual elements of the urban environment, whereas machine learning algorithms, principal component analysis, and self-organizing maps were utilised to comprehend the relationship between these elements. The findings of the study indicated that older adults exhibited a non-stress physiological response when the environment provided a layout that allowed for a wider field of vision (such as increased maximum visibility length, perimeter and isovist area). However, two of the studies included in the analysis frame their proposal into the Digital Twins paradigm (DT), proposing to combine wearable biosensors and geographical information systems (GIS)-based hotspot analysis to identify areas where older adults experience stress when interacting with the built environment [29], or using multiple urban sensing data to potentially identify less demanding routes for older adults based on the environmental factors that affect their mobility in a spatial context [1]. To this respect, Ahn and his colleagues [1] conducted a field experiment utilising an alternative processing method to analyse physiological data in order to identify built environment elements that induce distress to pedestrians. The study primarily examined physiological signals, including EDA and gait patterns, which were collected through wearable devices worn by older individuals while walking at their preferred speeds on a predetermined path that contained various environmental stressors at specific locations (POIs) (e.g., a partially broken wall, a dumpster, an uneven sidewalk, and overhanging dead branches and leaves). The physiological response data were divided into segments using a bottom-up segmentation approach, and a physiological saliency cue (PSC) was introduced and utilised to calculate the uniqueness of physiological responses within the segment of interest compared to others, and to compute a collective PSC of each POI across all participants [25]. Visual sensing data (i.e., images) were also gathered at a few locations, and ranked using a pairwise comparison in which participants scored the perceptual distress of each image by comparing a pair of street-level images. The findings demonstrated that the combination of physiological response data and crowdsourced visual sensing data effectively capture local environmental stressors, and when incorporating into a 3D virtual city model with geospatial localization, older pedestrians can potentially exploited the distress detected from both data sources to simulate their convenient daily trips using a virtual simulation environment. Similarly, in a pilot study carried out by Lee et al [29], wristband-type biosensors were employed to collect EDA and PPG signals from older adults as they walked along a controlled route and in their daily trips that included challenging environmental obstacles. Through the controlled route data collection, a total of 22 features were extracted from EDA and PPG and the best elderly individual stress classifier was selected by comparing the validation accuracy of several machine learning algorithms (ie Decision Tree, Gaussian Support Vector Machine, bagging tree and ensamble subspace KNN), resulting in an accuracy rate ranging from 79% to 92.5%. A bagging tree-based classifier was used to categorise the EDA and PPG signals of senior subjects collected during their daily trip into stress and no stress categories. The stress data samples were then geocoded based on GPS data simultaneously collected with physiological signals and distributed among grid cells for hotspot detection. Using the GIS-based kernel density estimation (KDE) technique, the significance of the density stress samples on grid-cells was tested based on the Monte-Carlo method, resulting in the identification of 40 grid cells (0.22% of the tested grid cells) as stress hotspots (with a significance level higher than the predetermined threshold of 0.01). A post hotspot investigation in the pilot study revealed that the identified hotspots align with older subjects' stressful experiences in the built environment. This information can enhance the Digital Twins' analytics platform, allowing for simulations of different scenarios and interventions to gain a deeper understanding of how older adults interact with the built environment in stressful ways. Interestingly, this paves the way for future research on integrating simulation-based models into Digital Twins environments with GIS-based analysis to further enhance our understanding of older adults personal mobility in urban settings.

Secondly, some articles discuss the environmental stressors that are used to evaluate age-friendly walkability environments in the field experiments (e.g., [1, 29]) or use the features affecting the functionality, safety, or appearance of the walking path to grade between high and low demanding environment [44]. Only two of the selected articles examine the physiological responses to spatial stimulus (ie the spatial configuration) [42] and to temporal stimuli (ie weather and noise level) [39], while none of the included articles consider individual stimuli (ie health conditions or previous experience [42]). None of the articles provide a comprehensive description of older adults participating in the experimental sessions. They only briefly mention their socio-economic status (specifically low income) [39] and assess their cognitive function using the Mini-Mental State Examination (MMSE) screening tool (e.g. [42–44]). However, our current review indicates a gap in the research regarding the detection of bodily responses to stimuli on outdoor walking paths using wearable sensors, specifically among informal carers. Future research should focus on this target population, as they play a significant role in providing health and social care services to elderly or disabled individuals [22].

From an overall perspective, based on our ongoing literature review, it would appear that using wearable sensors and a GIS-based approach could be a promising method for spatio-temporally capturing people's direct bodily responses to the environmental stressors [29]. As an example, the accuracy performance achieved by the trained bagging tree algorithms as reported in two of the included studies (ie accuracy ranging from 92.5% to 98.25% [29, 43]), may effectively suggest the applicability of the wearable biosensors in accurately identifying stress levels in elderly individuals as they engage with the built environment (e.g.,[29]). Therefore, future research ought to further investigate this people-centric sensing approach [44] and expand its application to examine older pedestrian-friendly environments in districts of different sizes and characteristics.

5 NEXT STEPS

Once the initial findings have been presented, the next step of action will involve delving deeper into the analysis of the full paper obtained through the chosen methodology. This will aim to offer a more comprehensive response to the research questions and can aid in better understanding how such devices can collect meaningful data during the daily activities of older adults and their related carers, as well as their reliability and suitability for an age-friendly urban environment.

ACKNOWLEDGMENTS

This publication was produced with the co-funding of European Union – Next Generation EU, in the context of the National Recovery and Resilience Plan, PE8 "Conseguenze e sfide dell'invecchiamento", Project Age-It (AGE - IT - A Novel Public-private Alliance to Generate Socioeconomic, Biomedical and Technological Solutions for an Inclusive Italian Ageing Society- Ageing Well in an Ageing Society) - AGE-IT- PEO0000015 - CUP: H43C22000840006.

REFERENCES

[1] Changbum Ahn, Youngjib Ham, Jinwoo Kim, and Jaeyoon Kim. 2020. A digital twin city model for age-friendly communities: Capturing environmental distress from multimodal sensory data. (2020).

[2] Stefania Bandini and Francesca Gasparini. 2021. Social and active inclusion of the elderly in the city through affective walkability. *The Review of Socionetwork Strategies* 15 (2021), 557–573.

[3] Stefania Bandini, Francesca Gasparini, and Flavio S Correa da Silva. 2022. Positive Artificial Intelligence Meets Affective Walkability. In *Living Beyond Data: Toward Sustainable Value Creation*. Springer, 161–178.

[4] Rümeysa Bayar and Merve Yılmaz. 2023. Measuring age-friendliness based on the walkability indices of older people to urban facilities. *Urban Design International* 28, 1 (2023), 35–51.

[5] Elaine Beller, Justin Clark, Guy Tsafnat, Clive Adams, Heinz Diehl, Hans Lund, Mourad Ouzzani, Kristina Thayer, James Thomas, Tari Turner, et al. 2018. Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). *Systematic reviews* 7 (2018), 1–7.

[6] Hidde Bleijendaal, Philip M Croon, Marinka D Oudkerk Pool, Arjan Malekzadeh, Simona Aufiero, Ahmad S Amin, Aeilko H Zwinderman, Yigal M Pinto, Arthur AM Wilde, and Michiel M Winter. 2022. Clinical applicability of artificial intelligence for patients with an inherited heart disease: a scoping review. *Trends in Cardiovascular Medicine* (2022).

[7] Wichor M Bramer, Dean Giustini, and Bianca MR Kramer. 2016. Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study. *Systematic reviews* 5 (2016), 1–7.

[8] Wichor M Bramer, Melissa L Rethlefsen, Jos Kleijnen, and Oscar H Franco. 2017. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic reviews* 6 (2017), 1–12.

[9] Niamh Carey, Marie Harte, and Laura Mc Cullagh. 2022. A text-mining tool generated title-abstract screening workload savings: performance evaluation versus single-human screening. *Journal of Clinical Epidemiology* 149 (2022), 53–59.

[10] Angela Carrera-Rivera, William Ochoa, Felix Larrinaga, and Ganix Lasa. 2022. How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX* 9 (2022), 101895.

[11] Benjamin W Chrisinger and Abby C King. 2018. Stress experiences in neighborhood and social environments (SENSE): a pilot study to integrate the quantified self with citizen science to improve the built environment and health. *International journal of health geographics* 17, 1 (2018), 1–13.

[12] Yolanda-María de-la Fuente-Robles, Adrián-Jesús Ricoy-Cano, Antonio-Pedro Albín-Rodríguez, José Luis López-Ruiz, and Macarena Espinilla-Estévez. 2022. Past, Present and Future of Research on Wearable Technologies for Healthcare: A Bibliometric Analysis Using Scopus. *Sensors* 22, 22 (2022), 8599.

[13] Joseph R Dettori and Daniel C Norvell. 2020. Kappa and beyond: is there agreement? *Global Spine Journal* 10, 4 (2020), 499–501.

[14] N Distefano, G Pulvirenti, and S Leonardi. 2021. Neighbourhood walkability: Elderly's priorities. *Research in transportation business & management* 40 (2021), 100547.

[15] JA Doust, E Pietrzak, S Sanders, and PP Glasziou. 2005. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *Journal of clinical epidemiology* 58, 5 (2005), 444–449.

[16] Kate Duchowny, Philippa Clarke, Nancy Ambrose Gallagher, Robert Adams, Andrea L Rosso, and Neil B Alexander. 2019. Using mobile, wearable, technology to understand the role of built environment demand for outdoor mobility. *Environment and Behavior* 51, 6 (2019), 671–688.

[17] Hannah Ewald, Irma Klerings, Gernot Wagner, Thomas L Heise, Jan M Stratil, Stefan K Lhachimi, Lars G Hemkens, Gerald Gartlehner, Susan Armijo-Olivo, and Barbara Nussbaumer-Streit. 2022. Searching two or more databases decreased the risk of missing relevant studies: a metaresearch study. *Journal of Clinical Epidemiology* 149 (2022), 154–164.

[18] Gerbrich Ferdinands, Raoul Schram, Jonathan de Bruin, Ayoub Bagheri, Daniel L Oberski, Lars Tummers, Jelle Jasper Teijema, and Rens van de Schoot. 2023. Performance of active learning models for screening prioritization in systematic reviews: a simulation study into the Average Time to Discover relevant records. *Systematic Reviews* 12, 1 (2023), 100.

[19] Francesca Gasparini, Alessandra Grossi, Katsuhiro Nishinari, and Stefania Bandini. 2020. Age-related walkability assessment: A preliminary study based on the emg. In *International Conference of the Italian Association for Artificial Intelligence.* Springer, 423–438.

[20] Michael Gusenbauer and Neal R Haddaway. 2020. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research synthesis methods* 11, 2 (2020), 181–217.

[21] Julian PT Higgins, Sally Green, et al. 2008. Cochrane handbook for systematic reviews of interventions. (2008).

[22] Frédérique Hoffmann and Ricardo Rodrigues. 2010. *Informal carers: who takes care of them.* European Centre for Social Welfare Policy and Research Vienna.

[23] Jacoba Huizenga, Aukelien Scheffelaar, Agnetha Fruijtier, Jean Pierre Wilken, Nienke Bleijenberg, and Tine Van Regenmortel. 2022. Everyday experiences of people living with mild cognitive impairment or dementia: A scoping review. *International Journal of Environmental Research and Public Health* 19, 17 (2022), 10828.

[24] Haitham Jahrami, Khaled Trabelsi, Omar Boukhris, Jumana Hasan Hussain, Ahmad F Alenezi, Ali Humood, Zahra Saif, Seithikurippu R Pandi-Perumal, and Mary V Seeman. 2022. The prevalence of mild, moderate, and severe nomophobia symptoms: a systematic review, meta-analysis, and meta-regression. *Behavioral Sciences* 13, 1 (2022), 35.

[25] Jinwoo Kim, Megha Yadav, Changbum R Ahn, and Theodora Chaspari. 2019. Saliency detection analysis of pedestrians' physiological responses to assess adverse built environment features. In *ASCE International Conference on Computing in Civil Engineering 2019.* American Society of Civil Engineers Reston, VA, 90–97.

[26] Jinwoo Kim, Megha Yadav, Theodora Chaspari, and Changbum R Ahn. 2020. Saliency detection analysis of collective physiological responses of pedestrians to evaluate neighborhood built environments. *Advanced Engineering Informatics* 43 (2020), 101035.

[27] Jan Kremer, Kim Steenstrup Pedersen, and Christian Igel. 2014. Active learning with support vector machines. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4, 4 (2014), 313–326.

[28] G Lee, B Choi, CR Ahn, and S Lee. 2019. Feasibility of wearable-based collective sensing to detect environmental barriers for facilitating the elderly's mobility. In *International Conference on Smart Infrastructure and Construction 2019 (ICSIC) Driving data-informed decision-making.* ICE Publishing, 143–149.

[29] Gaang Lee, Byungjoo Choi, Changbum Ryan Ahn, and SangHyun Lee. 2020. Wearable biosensor and hotspot analysis–based framework to detect stress hotspots for advancing elderly's mobility. *Journal of management in engineering* 36, 3 (2020), 04020010.

[30] Jean-Frederic Levesque, Mark F Harris, and Grant Russell. 2013. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. *International journal for equity in health* 12 (2013), 1–9.

[31] Kevin M Leyden. 2003. Social capital and the built environment: the importance of walkable neighborhoods. *American journal of public health* 93, 9 (2003), 1546–1551.

[32] Ming Li, Qian Gao, and Tianfei Yu. 2023. Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. *BMC cancer* 23, 1 (2023), 799.

[33] Alberto Martín-Martín, Enrique Orduna-Malea, and Emilio Delgado López-Cózar. 2018. Coverage of highly-cited documents in Google Scholar, Web of Science, and Scopus: a multidisciplinary comparison. *Scientometrics* 116, 3 (2018), 2175–2188.

[34] David H Metz. 2000. Mobility of older people and their quality of life. *Transport policy* 7, 2 (2000), 149–152.

[35] Frida Milella, Davide Donato Russo, and Stefania Bandini. [n. d.]. How Artificial Intelligence can support informal caregivers in their caring duties to elderly? A systematic review of the literature. *In: AIXAS2023 Italian Workshop on Artifcial Intelligence for an Ageing Society, co-located with the 22nd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2023), November 6-9, 2023, Rome, Italy (forthcoming)* ([n. d.]).

[36] Frida Milella, Davide Donato Russo, and Stefania Bandini. 2023. AI-Powered Solutions to Support Informal Caregivers in Their Decision-Making: A Systematic Review of the Literature 1. *OBM Geriatrics* 7, 4 (2023), 1–11.

[37] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. 2010. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International journal of surgery* 8, 5 (2010), 336–341.

[38] Mourad Ouzzani, Hossam Hammady, Zbys Fedorowicz, and Ahmed Elmagarmid. 2016. Rayyan—a web and mobile app for systematic reviews. *Systematic reviews* 5 (2016), 1–10.

[39] Jenny Roe, Andrew Mondschein, Chris Neale, Laura Barnes, Medhi Boukhechba, and Stephanie Lopez. 2020. The urban built environment, walking and mental health outcomes among older adults: a pilot study. *Frontiers in public health* (2020), 528.

[40] Burr Settles. 2009. Active learning literature survey. (2009).

[41] Jeff Speck. 2013. *Walkable city: How downtown can save America, one step at a time.* macmillan.

[42] Alex Torku, Albert PC Chan, Esther HK Yung, and JoonOh Seo. 2021. The influence of urban visuospatial configuration on older adults' stress: A wearable physiological-perceived stress sensing and data mining based-approach. *Building and Environment* 206 (2021), 108298.

[43] Alex Torku, Albert PC Chan, Esther HK Yung, and JoonOh Seo. 2022. Detecting stressful older adults-environment interactions to improve neighbourhood mobility: A multimodal physiological sensing, machine learning, and risk hotspot analysis-based approach. *Building and Environment* 224 (2022), 109533.

[44] Alex Torku, Albert PC Chan, Esther HK Yung, JoonOh Seo, and Maxwell F Antwi-Afari. 2022. Wearable sensing and mining of the informativeness of older adults' physiological, behavioral, and cognitive responses to detect demanding environmental conditions. *Environment and Behavior* 54, 6 (2022), 1005–1057.

[45] Guy Tsafnat, Paul Glasziou, Miew Keen Choong, Adam Dunn, Filippo Galgani, and Enrico Coiera. 2014. Systematic review automation technologies. *Systematic reviews* 3 (2014), 1–15.

[46] Rens Van De Schoot, Jonathan De Bruin, Raoul Schram, Parisa Zahedi, Jan De Boer, Felix Weijdema, Bianca Kramer, Martijn Huijts, Maarten Hoogerwerf, Gerbrich Ferdinands, et al. 2021. An open source machine learning framework for efficient and transparent systematic reviews. *Nature machine intelligence* 3, 2 (2021), 125–133.

[47] Sanne HB van Dijk, Marjolein GJ Brusse-Keizer, Charlotte C Bucsán, Job van der Palen, Carine JM Doggen, and Anke Lenferink. 2023. Artificial intelligence in systematic reviews: promising when appropriately used. *BMJ open* 13, 7 (2023).

[48] Wynanda A van Enst, Rob JPM Scholten, Penny Whiting, Aeilko H Zwinderman, and Lotty Hooft. 2014. Meta-epidemiologic analysis indicates that MEDLINE searches are sufficient for diagnostic test accuracy systematic reviews. *Journal of clinical epidemiology* 67, 11 (2014), 1192–1199.

[49] F van Ommen, P Coenen, A Malekzadeh, AGEM de Boer, MA Greidanus, and SFA Duijts. 2023. Interventions for work participation of unemployed or work-disabled cancer survivors: a systematic review. *Acta Oncologica* (2023), 1–12.

[50] Zhen Wang, Tarek Nayfeh, Jennifer Tetzlaff, Peter O'Blenis, and Mohammad Hassan Murad. 2020. Error rates of human reviewers during abstract screening in systematic reviews. *PloS one* 15, 1 (2020), e0227742.

[51] Samuel L Warren and Ahmed A Moustafa. 2023. Functional magnetic resonance imaging, deep learning, and Alzheimer's disease: A systematic review. *Journal of Neuroimaging* 33, 1 (2023), 5–18.

[52] Michel Wijkstra, Katja Rogers, Regan L Mandryk, Remco C Veltkamp, and Julian Frommel. 2023. Help, My Game Is Toxic! First Insights from a Systematic Literature Review on Intervention Systems for Toxic Behaviors in Online Video Games. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play.* 3–9.

[53] Zhe Yu and Tim Menzies. 2019. FAST2: An intelligent assistant for finding relevant papers. *Expert Systems with Applications* 120 (2019), 57–71.