Bian et al. Gerontechnology (2020) 19:3
https://doi.org/10.4017/gt.2020.19.003.06

Technology for home-based frailty assessment and prediction: A systematic review

Chao Bian MS[a,*], Bing Ye MS[b,d], Charlene H. Chu RN GNC PhD[b,c], Katherine S. McGilton PhD RN FAAN[b,c], Alex Mihailidis PhD PEng[a,b,d]

[a]*Institute of Biomedical Engineering, University of Toronto, Toronto, Canada;* [b]*KITE Research, Toronto Rehabilitation Institute, University Health Network, Toronto, Canada;* [c]*Lawrence S. Bloomberg Faculty of Nursing, University of Toronto, Toronto, Canada;* [d]*Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, Canada; *Corresponding author: chao.bian@mail.utoronto.ca*

Abstract

Background: The current clinical frailty assessments are time-consuming and subjective which can lead to inaccurate results and delayed medical attention. Sensor technology and artificial intelligence enable home-based frailty assessment; however, there are no systematic reviews of existing technological methods for home-based frailty assessment and prediction.
Objective: To analyze and synthesize the frailty criteria, sensor technology, and the statistical or artificial intelligence methods used in home-based frailty assessment and prediction.
Methods: An exhaustive database search was performed. Three reviewers screened all studies by following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. The sensors and AI used for assessing frailty were synthesized with a particular focus on home-based technology. The Sackett's Level of Evidence Scale was also used to evaluate clinical evidence for the included studies.
Results: Body-worn sensors were the most commonly used (72%) technology in home-based frailty assessment. All of the body-worn sensors were accelerometer-based. 88% of the included studies measured physical activity for assessing frailty commonly defined by Fried's Frailty Index (75%). Heterogenous machine learning algorithms have been applied for classifying frailty. However, none of the AI methods were tested for the predictability of frailty. Only one longitudinal study followed up older participants for 10 years and revealed a high odds ratio for the development of frailty using physical activity.
Conclusion: The database search was limited to definitions of physical frailty and the English language. Various types of sensor technology with good accuracy are used to measure specific frailty criteria and functional tests. However, there is a lack of longitudinal studies for predicting frailty progression. To date, there is limited testing of the sensors using older populations with functional and cognitive comorbidities and because they are at higher risk of frailty they should be a priority moving forward.

Keywords: Frailty, assessment, technology, home, sensor, Artificial Intelligence

Introduction

The World Health Organization (WHO) projects a 56% growth in the global number of people aged sixty years or over between 2015 and 2030 (WHO, 2015). The majority of the aging population is expected to experience age-related issues such as declining functional capacity and increasing vulnerability to disease, disability, and death (Mitnitski et al., 2015). One of the most common age-related issues is frailty, widely considered as *"a condition in which the individual is in a vulnerable state at increased risk of adverse health outcomes and/or dying when exposed to a stressor"* (Morley et al., 2013). A recent systematic review on the prevalence of frailty in community-dwelling older adults found a 9.9% mean prevalence of physical frailty among people aged 65 years or older, with a higher prevalence of 13.6% when psychosocial frailty was also included (Collard et al., 2012).

Fortunately, researchers (Morley et al., 2013; Puts et al., 2016) have found that pre-frailty is a reversible status if frailty is identified and appropriate interventions are applied early. To assess frailty, clinicians use frailty scales typically in the form of standardized self-report questionnaires, or a battery of physical tests to evaluate the physical, physiological, cognitive or social aspects of patient health conducted by clinicians (Bruyère et al., 2017). However, these assessment methods have several limitations, such as subjective self-report, increased complexity to administer, and infrequent assessment and follow-up. Additionally, numerous frailty assessment scales exist due to a lack of consensus on the operational definition of frailty among clinicians. As such, these limitations may result in inaccurate or delayed identification of frailty.

There are new attempts by researchers to use novel technologies to continuously monitor and assess frailty in people's everyday living settings. In contrast to acute conditions, the onset and progress of frailty are slow and gradual, and its symptoms are usually reflected in individuals' ability to complete activities throughout their day. In community settings, frailty predicts future hospitalization, worsened quality of life, and loss of ability to carry out activities of daily living. Routinely identifying frailty offers opportunities for targeted care, including applying clinical practice guidelines and tools specific to frailty. The clinical reality is that frailty clinical assessments are completed periodically; and as a consequence, there is a high likelihood that one's deterioration may not be detected and the opportunity to initiate interventions early will be missed. The use of technologies including sensors, and artificial intelligence (AI) in home settings for frailty assessment can play an important role to overcome this practical challenge. Sensor-based devices, such as smart wearable watches, motion sensors, or pressure sensors, can be used in everyday living settings to capture physiological or behavior changes (Mainetti et al., 2017). The internet of things (loT) can connect these sensors remotely and enable data to be collected and analyzed anywhere in real-time. In the field of healthcare, AI can find patterns in sensor data usually using machine learning-based methods, and various IoT and AI-based technologies have been proposed to help solve healthcare problems by using sensor data (Azimi et al., 2017; Zouba et al., 2009). Technology can be used to enable early detection of frailty and early intervention for reversing or better managing frailty; however, despite recent increased attention given to technology for home-based frailty monitoring, no systematic reviews have been completed in this area.

This systematic review aims to synthesize the use of technologies to assess and predict frailty in home settings. The following research questions are addressed:
(1) What frailty criteria are being measured by the technologies in home settings?
(2) What types of technology are used for in-home frailty assessment and prediction?
(3) What statistical or artificial intelligence methods have been used for frailty assessment and prediction of frailty progression, and their assessment and prediction accuracy compared to clinical methods?

First, this review will synthesize technologies including hardware and software systems, AI for frailty assessment, and prediction in home-based systems. Then the studies will be evaluated by a clinical evidence scale to find the technologies with clinical evidence. The goal of this review is to compare and identify the state-of-art technology-based in-home frailty assessment and prediction methods, find potential opportunities and gaps in this field, and discuss future research directions.

Methods

Database search

An information specialist in Toronto Rehabilitation Institute Library conducted an extensive database search in the following electronic databases: Medline (including Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily), Cochrane Central Register of Controlled Trials, NHS Economic Evaluation Database, Cochrane Health Technology Assessment, Embase, PsycINFO, CINAHL, Scopus, Compendex, and Proquest Dissertations and Theses Global. The search was based on a combination of standardized database vocabulary in the following three areas: (1) frail, (2) home, and (3) technology. Variants of the keywords from these three areas were used in searches. The search was limited to English language only and papers published in 2000 and onwards due to the rapid advancement of technology. A full list of search strategies used across the entire databases can be found under Search Strategy section in Appendix. The first complete search was conducted on July 28, 2017, and two more update searches using the same search criteria as the first search were conducted on July 03, 2018 and May 10, 2019, respectively.

Study selection

The selection process followed the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines (Moher et al., 2009) which includes duplicate removal, title, and abstract screening, full-text screening, qualitative and quantitative analysis. After the database search, we first imported all studies into Endnote software for duplicate removal. The studies without duplicates were then imported into Covidence software for the title and abstract screening, and full-text screening. The inclusion and exclusion criteria used in the screening process were as follows:

Inclusion criteria:
- Population: individuals with pre-frail or frail conditions defined by any clinical frailty scale.
- Technology: use of information and communication technologies, sensors or AI.
- Frailty screening: frailty assessment, detection or monitoring.

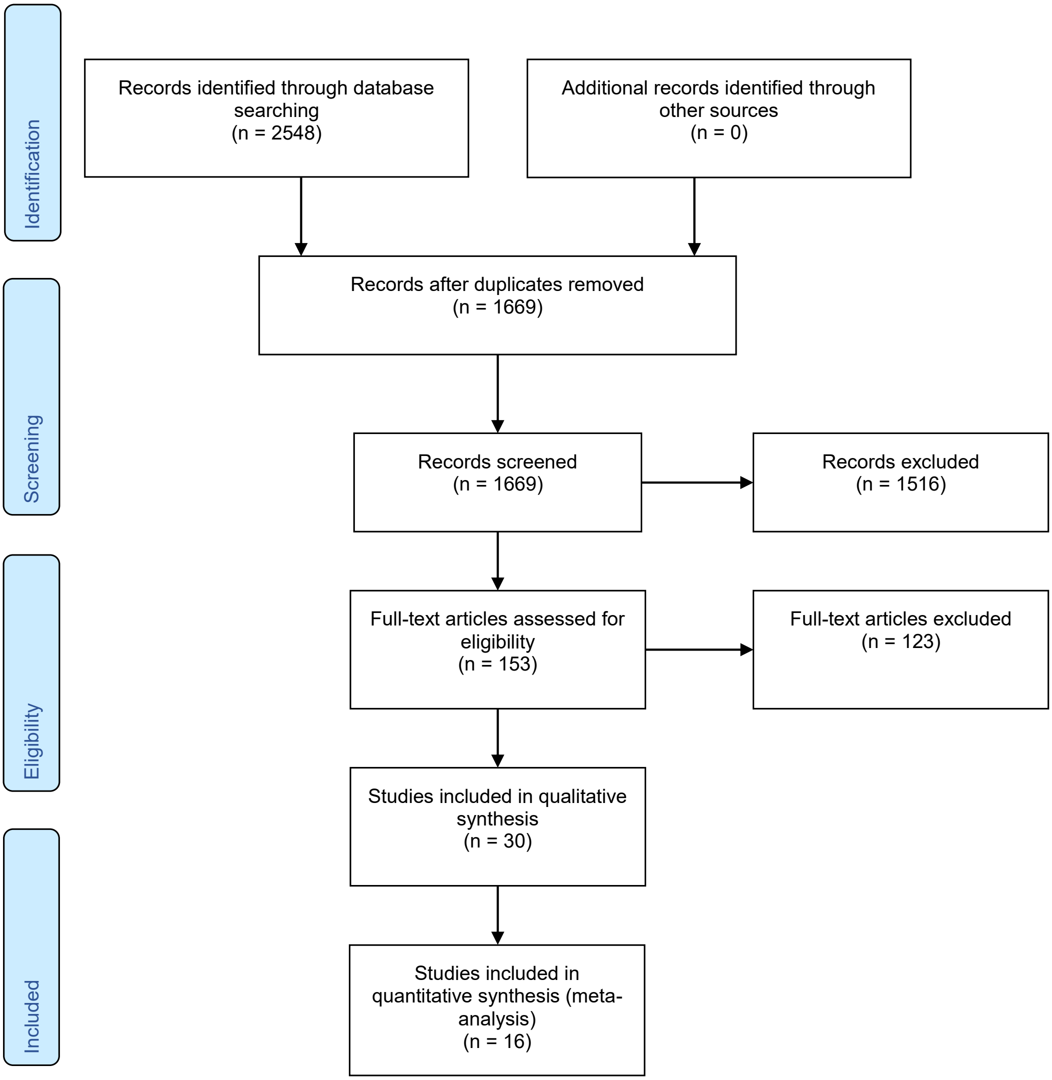

Figure 1. PRISMA Flow Diagram.

- Clinical frailty criteria: measurement of frailty-related criteria such as weight loss, gait speed, grip strength, defined by any clinical frailty index.
- Home-based: trials conducted in a home environment.

Exclusion criteria:
- Study subjects are not older adults.
- Frailty intervention studies that focus on applying any therapeutic method (such as exercise or medicine) to control or reverse frailty progress.
- Frailty prevalence studies.
- Qualitative, usability, non-peer reviewed, and non-academic studies.
- Systematic reviews, literature review, case reports, and letters.

The web-based Covidence systematic review software was used to manage the screening of the studies. Two researchers (C.B. and B.Y.) completed the title and abstract screening independently based on the inclusion and the exclusion criteria. In case of disagreements, the papers were discussed and reviewed with the third reviewer (C.C.). The remaining records after the title and abstract screening then underwent a full-text screening. One researcher (C.B.) completed the full-text screening using the above inclusion and exclusion criteria. A final list of studies after the full-text screening was then included in the review.

The data were extracted from the included studies according to the following categories: study objective, study type, author and year, demographic information including sample size and

Table 1. Study characteristic of studies with level 1 clinical evidence.

Author, Year	Study Design	Age	Sample Size	Frailty Measures	Test Period	Test Setting
Mohler et al., 2013	Cross-sectional	Not reported	20	PA	24 hours	Home/Uncontrolled
Schwenk et al., 2015	Cross-sectional	>=65	125 (44 non-frail, 60 pre-frail, 21 frail)	Gait parameters, balance, PA	24 hours for PA	Home/Controlled
Toosizadeh et al., 2015	Cross-sectional	>=65	117 (50 non-frail, 51 pre-frail, 16 frail)	Muscle strength	50 seconds	Home/Controlled
Parvaneh et al., 2017	Cross-sectional	78±8	120	PA	> 24 hours	Home/Uncontrolled
Geraedts et al., 2015	Cross-sectional	>=70	20	PA	1 week	Home/Controlled and uncontrolled
Chen et al., 2015	Cross-sectional	>65	1,527 (593 men, 934 women)	5 Fried criteria, social, health behavior, function	Mean 6.9 ±1.6 days	Lab and home/Controlled
Chang et al., 2013	Cross-sectional	>65	309 (131 men, 178 women)	Reaction time, TUG, Sit-to-Stand, weight, balance, PA	Not reported	Home/Uncontrolled
Alqahtani et al., 2017	Cross-sectional	87±6	29	Balance, muscle strength	Not reported	Long-term Care/Controlled
Jansen et al., 2015	Cross-sectional	73.6 ± 6.3	84	PA	7 days	Home/Uncontrolled

mean age, clinical frailty measures and frailty criteria, technology used, and study results.

The included studies were then categorized based on their level of evidence (LOE) according to the Centre for Evidence-Based Medicine (CEBM) (CEBM Levels of Evidence Working Group, 2009) in order to identify the level of clinical evidence to support the use of technology in frailty assessment and prediction. The CEBM guideline tool evolved from Sackett's LOE (Sackett, 1989), which is one of the earliest and widely recognized scales for evaluating the level of clinical evidence. Levels 1, 2, and 3 of clinical evidence indicate that a study is a validating cohort study with good reference standards, an exploratory cohort study with good reference standards, and a study without consistently applied reference standards, respectively. Only studies with Level 1 clinical evidence (i.e. studies that used a reference standard) will be analyzed in detail. As clinicians have not reached a consensus on which frailty operational scale should be used in routine care (Morley et al., 2013), any clinical frailty scale was eligible as a reference standard when determining the level of clinical evidence in this review.

HEADING 1>RESULTS

The electronic database search generated a total of 2,548 studies. After the duplicates were removed, the remaining 1669 records underwent the two screening stages (title and abstract screening; full-text screening). After the title and abstract screening, a total of 1516 studies were removed and the remaining 153 full-text studies were screened. Of these, 123 were removed and 30 studies were included in the final analysis. The PRISMA flow diagram is shown in *Figure 1*.

Although we defined the year range in the search to be post-2000 to present (2019), 29 of 30 eligible papers were published post-2010 and only one paper was published in 2009 (Zouba et al., 2009). Furthermore, among all papers that published post-2010, 90% were published between 2013 - 2019.

Study characteristics

The age criteria for study participants in the clinical evidence studies were 65 years old or

Table 1. Study characteristic of studies with level 1 clinical evidence (cont.)

Author, Year	**Study Design**	**Mean Age**	**Sample Size**	**Frailty Measures**	**Test Period**	***Test Settings***
Razjouyan et al., 2018	Cross-sectional	75 ± 10	153	PA	48 hours	Home/Uncontrolled
Bastone Carvalho et al., 2015	Cross-sectional	66-86	26	Aerobic fitness, PA	1 week	Home/Uncontrolled
Goonawardene et al.,	Longitudinal	60 - 91	46 (19 men, 27 women)	39 health deficits defined in Clinical Frailty Index	30 days (1-year follow-up)	Home/Uncontrolled
Okubo et al., 2018	Cross-sectional	71.7	185	Gait speed, strength, endurance, malnutrition and protein intake	A 30-minute session and 1 week (PA)	Outside and Home/Controlled and uncontrolled
Rainoldi et al., 2018	Cross-sectional	71 ± 6	25 (60% women)	Daily grade of performed PA	1 week	Home/Uncontrolled
Tegou et al., 2019	Cross-sectional	76.8 ± 5.2 (men)	271 (102 men, 169	PA: room transitions	1-7 days	Home/Uncontrolled
Yuki et al., 2019	Longitudinal	71.1± 4.3	401	PA: steps, activity time	7 days (10 years follow-up)	Home/Uncontrolled

older. The highest mean age reported was 87±6 (Alqahtani et al., 2017). The sample size in the included studies varied from 20 to 1527. The study characteristics can be found in *Table 1* and in *Appendix I*. The most commonly used inclusion criteria were *"no gait or mobility disorders"* as gait tests were the most used assessments to measure clinical frailty using Fried's frailty phenotype scale. The most common exclusion criteria included cognitive impairment indicated by a low MMSE score (<18-24 depending on studies) and the presence of a terminal illness.

16 of the 30 (53%) studies have a Level 1 clinical evidence as identified by the CEBM rating scale (*Table 2*). Specifically, these manuscripts of novel technologies compared their technology to clinical gold standards for frailty assessment with promising results to be used as an alternative method for frailty assessment. 12 of the 16 studies with Level 1 clinical evidence used Fried's frailty phenotype scale (consisting of five frailty criteria: slow walking speed, weak handgrip strength, weight loss in the last year, exhaustion and low physical activities) as the clinical reference gold standard for the proposed technologies making it the most common clinical reference for technology. The other 4 studies used Tilburg Frailty Indicator (TFI) (Gianaria et al., 2016), Groningen Frailty Indicator (GFI) (Geraedts et al., 2015), *"Cumulative Index"* model of frailty by Kenneth Rockwood (Goonawardene et al., 2018) and Identification Seniors At Risk - Hospitalized Patients' questionnaire (ISAR-HP) (Jansen et al., 2015). The choice of the operational scale could depend on the context. Fried's frailty scale is usually used for persons living in the community. Other frailty scales are more commonly used in hospitals.

The research design of 14 (out of 16) studies were cross-sectional in which researchers used proposed technological tools, such as a wearable sensor to collect frailty-related data from study participants within a defined time. The outcome of the study, which is a frailty class calculated from the data collected by the proposed technology was achieved right after the data collection. The data collection periods were short, for example, less than a day or shorter durations like a one-time assessment of the required measurements. 2 studies were longitudinal study with longer monitoring period (30 days and 7 days, respectively) (Goonawardene et al., 2018; Yuki et al, 2019). There was a marked lack of longitudinal studies that monitored the progression of frailty over many years in the home.

Table 2. Frailty criteria measurements in the studies with level 1 clinical evidence.

Frailty Criteria	**Sensor measurement**	**Gold standard measurement**	**LOE**	***Reference***
PA, 5-Chair-Sit-to-Stand	Percent time lying on side, standing, walking, number of walking episodes, number of steps	Fried	Level 1	(Mohler et al., 2013)
Gait, balance, PA	Gait: speed, stride time, stride length, double support, gait variability defined as coefficient of variation of stride velocity Balance: sway of ankle, hip, and center of mass PA: postural transitions, movements such as walking, standing, or sitting	Fried	Level 1	(Schwenk et al., 2015)
Slowness, muscle strength, exhaustion	Kinematics and kinetics of elbow flexion (speed, flexibility, power, rise time, moment, jerkiness, and speed reduction)	Fried	Level 1	(Toosizadeh et al., 2015)
PA	Total number of transitions, stand-to-walk, and walk-to-stand	Fried	Level 1	(Parvaneh et al., 2017)
PA	Time-on-legs, mobility postures (active period, sit to stand transfer, walking and lying)	GFI	Level 1	(Geraedts et al., 2015)
PA	Energy expenditure of physical activity quantified as kilocalories/kg	Fried	Level 1	(Chen et al., 2015)
Reaction time, slowness, weakness, weight, PA	Reaction time, TUG, Sit-to-Stand, weight, balance, PA	Fried	Level 1	(Y. C. Chang et al., 2013)
Balance, muscle strength	Standing balance test (body sway), knee extension, knee flexion, hip abduction, hip flexion, ankle plantar flexion and dorsiflexion	Fried	Level 1	(Alqahtani et al., 2017)
PA	Activity intensity, GPS location, active transportation distances	ISAR-HP	Level 1	(Jansen et al., 2015)

Frailty criteria and measurements

Technologies in the reviewed studies measured frailty in two ways: (1) directly measured frailty criteria such as physical activity, gait speed and muscle strength which are defined in clinical frailty scales such as the Fried's frailty phenotypes; and (2) measured the performances in standardized functional tests and studied the correlation between the functional performance and frailty measured as a baseline by researchers using clinical frailty scales. The proposed technologies and the frailty criteria are shown in *Table 2* and *Appendix I*. The most commonly used criteria are various aspects of physical activities (PA) which have been used in 14 of the 16 studies (88%). The Timed up and Go (TUG) and Sit to Stand test are the two most common standardized functional tests that technologies assessed to measure frailty, followed by measuring balance (e.g. sway of ankle, hip, and center of mass), gait (e.g. walking speed, stride time, gait variability) and muscle/grip strength.

The specific parameters for measuring physical activity in the review studies include activity intensity of light to vigorous levels (Razjouyan et al., 2018; Yuki et al., 2019), duration of activities (Bastone et al., 2015; Geraedts et al., 2015; Mohler et al., 2013), number of transitions between different postures such as standing, walking and lying (Geraedts et al., 2015; Parvaneh et al., 2017; Schwenk et al., 2015), daily step counts (Bastone et al., 2015; Mohler et al., 2013) and energy expenditure (kcal/day) (Chen et al., 2015). For instance, Mohler et al. used sensors to continuously monitor participants throughout the day and measured their percentage of time spent in a day either lying down on their side, standing, and walking; a number of walking episodes, and a number of steps using an accelerometer motion sensor (Mohler et al., 2013). In contrast, Parvaneh et al. used sensors to measure the total number of transitions during stand-to-walk and walk-to-stand tests (Parvaneh et al., 2017). *Table 2* lists the sensor measurements and their corresponding clinical gold standard references.

Balance was also assessed by measuring specific parameters like body sway and center of mass. Specifically, Schwenk et al. captured data related to the sway of the ankle, hip, and center of mass in medial-lateral and anterior-posterior direction during their experiments (Schwenk et al., 2015). Using technology to capture balance parameters during a standing balance test was another way

Table 2. Frailty criteria measurements in the studies with level 1 clinical evidence (cont.)

Frailty Criteria	**Sensor measurement**	**Gold standard measurement**	**LOE**	***Reference***
PA	Activity intensity, postural parameters, sleep quantity parameters	Fried	Level 1	(Razjouyan et al., 2018)
Aerobic fitness, PA	Incremental Shuttle Walk Test including parameters oxygen consumption, respiratory exchange ratio, and heart rate; and physical activity variables - daily step count, daily counts, daily duration of sedentary activity, light activity, moderate activity, daily energy expenditure	Fried	Level 1	(Bastone et al., 2015)
Daily living patterns (PA)	Away duration and frequency, sleep duration, location occupancy and frequency, intensity of sensor firings, in-home transitions	Clinical Frailty Scale (Rockwood)	Level 1	(Goonawardene et al., 2018)
Gait speed, strength, endurance, PA,	Gait speed, grip strength, Sit-to-Stand, PA, nutrition, 6-minute walk	Fried	Level 1	(Okubo et al., 2018)
PA	Daily grade of performed physical activity, as: Very Low, Low, Medium, High, and Very High mobility	TFI	Level 1	(Rainoldi et al., 2018)
PA	Room transitions	Fried	Level 1	(Tegou et al., 2019)
PA	Steps, activity intensity	Fried	Level 1	(Yuki et al., 2019)

to assess balance (Alqahtani et al., 2017).

The specific gait-related measurements include speed, stride, double support, walking time, and distance (Schwenk et al., 2015). Such data can be collected in a home-based 15-feet walk test (Schwenk et al., 2015), or along a 6-m course with 1-m acceleration/deceleration space at each end (Okubo et al., 2018).

Muscle strength was measured in both upper extremity motion and lower extremity motion. Toosizadeh et al. measured kinematics and kinetics of elbow flexion (speed, flexibility, power, rise time, moment, jerkiness, and speed reduction) in an upper extremity motion test (Toosizadeh et al., 2015), whereas Alqahtani et al. chose to measure lower extremity muscle strength by measuring knee extension, knee flexion, hip abduction, hip flexion, ankle plantar flexion and dorsiflexion (Alqahtani et al., 2017).

Technology was also used to measure standardized functional tests, including TUG and Sit to Stand, where sensors were able to discern many different features of movement parameters. For example, 44 sensor-derived features were extracted from a TUG test (Greene, Doheny, O'Halloran, et al., 2014) and were grouped into four categories: temporal gait parameters, spatial gait parameters, tri-axial angular velocity parameters and turn parameters. Similarly, another study extracted 52 TUG features which also include temporal, spatial, turning, and rotational characteristics (Greene, Doheny, Kenny, et al., 2014). The same study also extracted 82 movement parameters during a Five Times Sit to Stand (FTSS) test.

Although Fried's frailty phenotype was the most commonly used clinical reference (75%), two of the five frailty criteria defined in Fried's frailty index - weight loss and exhaustion – have not been extensively measured by the technologies in the included studies. For those that measured these criteria, data were collected with a modified home chair (weight) (Chang et al., 2013) and an elbow flexion test (exhaustion) (Toosizadeh et al., 2015).

heading 2>Technology

Sensors

Table 3 in the Appendix outlines all the sensor technologies used for frailty assessment and prediction. Body-worn technology was in 13 (72%) of 16 Level 1 clinically evident studies; 5 studies used ambient technology. All the 13 studies that involved body-worn technology used inertial measurement unit (IMU) which consists of ac-

Table 3. Sensor technologies for frailty assessment and prediction in the studies with level 1 clinical evidence.

Technology	Sensor Type	Statistical Analysis	Machine Learning	Accuracy/Significance	Reference
3-axial accelerometer motion sensor (BioSensics)	Body	Not reported	N/A	p = 0.001-0.005 for postural and PA discriminators	(Mohler et al., 2013)
Gait assessment system, balance assessment system, motion sensor in T-shirt (BioSensics)	Body	Logistic regression	N/A	AUC > 0.76: non-frail vs. pre-frail; AUC > 0.8: non-frail, pre-frail, frail	(Schwenk et al., 2015)
Two tri-axial wearable gyroscope sensors (BioSensics)	Body	ANOVA, tukey honestly significant difference tests, multivariate logistic regression model	N/A	Sensitivity = 94%, specificity = 98% for **classifying** frailty and prefrailty	(Toosizadeh et al., 2015)
Shirt-embedded sensor (BioSensics)	Body	General linear model tests, logistic regression	N/A	Walk-to-stand (OR = 0.997 p = 0.013), quick sitting (OR = 1.036, p = 0.05), age (OR = 1.073, p = 0.016)	(Parvaneh et al., 2017)
Necklace-worn IMU sensor (Philips Research)	Body	Independent sample t-test	N/A	No significant difference between non-frail and frail	(Geraedts et al., 2015)
3-axial accelerometer (Omron)	Body	Univariate, multivariate logistic analyses	N/A	OR > 1: non-frail vs. pre-frail, non-frail vs. frail	(Chen et al., 2015)
e-furniture (eScale, eChair, ePad, eReach, eQuestionnaire)	Ambient	N/A	ANN	Sensitivity: 79.71%, Specificity: 86.25%, 3%	(Chang et al., 2013)
Uniaxial load cell, dual-axis accelerometer (Analog Devices)	Body, Ambient	Spearman rank correlation	N/A	p> 0.05, Spearman ρ > 0.40 for body sway, knee flexion, ankle dorsiflexion, hip abduction strength	(Alqahtani et al., 2017)
Accelerometer (ActiGraph), GPS device (QStarz)	Body	Bootstrapped linear regression	N/A	no significant differences in PA between frailty groups	(Jansen et al., 2015)

celerometers, gyroscopes, and sometimes also magnetometers. They were used to measure various frailty criteria (physical activity, muscle strength, gait, etc.). The form factor of the IMU sensors is different. The IMU sensor can be embedded in a shirt (Parvaneh et al., 2017; Schwenk et al., 2015) or a standalone wearable device that can be worn in the neck (Geraedts et al., 2015), or chest (Mohler et al., 2013).

Unlike body-worn sensors which are homogeneously IMU-based, ambient sensors are heterogeneous in type. The ambient technologies include furniture embedded sensors (Chang et al., 2013), motion sensors and door sensors (Goonawardene et al., 2018), balance pad (Alqahtani et al., 2017), and Bluetooth beacons (Tegou et al., 2019). The furniture embedded sensors were a suite of technologies that consists of eScale, eChair, ePad, eReach (Y. C. Chang et al., 2013). Collectively, these were used to measure reaction time and slowness, weight, Sit to Stand, TUG, and functional reach.

Two studies took a sensor fusion approach that used a combination of body-worn and ambient sensors to measure multiple frailty related criteria (Alqahtani et al., 2017; Okubo et al., 2018). Alqahtani et al. used a uniaxial load cell (ambient) and a dual-axis accelerometer (body-worn). The load cell was made into a lower extremity muscle strength device to measure knee extension, knee flexion, hip abduction, hip flexion, Ankle plantar flexion, and dorsiflexion. The Airex Pad and the accelerometer were used to conduct a standing balance test.

Four studies were conducted in participant's home but with the researchers on site instructing the tests (Alqahtani et al., 2017; Chen et al., 2015; Schwenk et al., 2015; Toosizadeh et al., 2015) and two studies were conducted using both controlled and uncontrolled protocols in home settings (Geraedts et al., 2015; Okubo et al., 2018). 10 of the 16 studies were conducted in a *"free-living uncontrolled environment"*, meaning that older adults are living freely within their environments according to their own pace and schedule without following any predefined protocol created by re-

Table 3. Sensor technologies for frailty assessment and prediction in the studies with level 1 clinical evidence (cont.).

Technology	**Sensor Type**	**Statistical Analysis**	**Machine Learning**	**Accuracy/Significance**	***Reference***
Tri-axial accelerometer-based pendant sensor (BioSensics)	Body	N/A	Regression, decision trees	91.8% (Sensitivity), 81.4% (specificity), 84.7% (accuracy), 0.88 (AUC) (91.8%, 81.4%, 84.7%, and 0.88, respectively)	(Razjouyan et al., 2018)
Accelerometer (ActiGraph), portable metabolic measurement system	Body	Unpaired t-test, logistic regression	N/A	OR 0.96 – 1.02 for distance, steps, sedentary, light activity, energy	(Bastone et al., 2015)
Passive infrared motion sensor, door contact sensor	Ambient	N/A	Logistic regression, linear discriminant analysis, naïve bayes	AUC = 0.98 (highest, Logistic regression)	(Goonawardene et al., 2018)
Hand-grip strength meter, 3-axis accelerometer (Omron)	Body, Ambient	Spearman's rank correlation coefficient	N/A	$p < 0.05$ for global scores (including physical and nutritional criteria) and physical functionality	(Okubo et al., 2018)
Smart watch	Body	Linear regression	N/A	$p = 0.031$ for daily grade of performed physical activity	(Rainoldi et al., 2018)
Localization system (beacons and smart-phones)	Ambient	N/A	Naïve bayes, 10-nearest neighbour, neural networks, decision tree, random forests	Highest accuracy 82.33% (Random forests)	(Tegou et al., 2019)
Uniaxial accelerometry sensor (Suzuken)	Body	Generalized estimate equation	N/A	OR 1.72 – 3.19 for walking < 5000 steps, < 7.5 minutes/d of moderate to vigorous PA	(Yuki et al., 2019)

searchers and without the presence of researchers.

Depending on the frailty criteria and measurements, the data collection period ranged from 50 seconds (Toosizadeh et al., 2015) to 30 days (Goonawardene et al., 2018). The studies that used IMU-based body-worn devices to measure physical activity collected data for at least 24 hours to a week.

Technologies from companies like BioSensics, ActiGraph, Analog Device, Omeron have been used by a number of studies in this review. IMU-based wearable devices such as LEGSys and PAMSys trackers from BioSensics, IMU from Omron were used in a similar way to monitor motions during physical activities (Chen et al., 2015; Mohler et al., 2013; Parvaneh et al., 2017; Schwenk et al., 2015). Besides the dominant number of IMU-based technology, a portable metabolic measurement system by COSMED was used to study the association between frailty and aerobic fitness (Bastone et al., 2015). The COSMED K4 system consists of a face mask, HR chest strip, battery and transmitting unit (containing the O_2 and CO_2 gas analyzers), and a receiving unit. Data including oxygen consumption, respiratory exchange ratio, and heart rate was collected during an incremental shuttle walk test.

The cost of commercial technologies varies. Some examples include the LEGSys or PAMSys IMU from BioSensics (≈ USD 4000), Active style Pro of Omron (¥20,000 ≈ USD 184), and Qstarz BT-Q1000x GPS receiver (≈ USD 100). The most expensive device in this review is the COSMED K4 portable metabolic measurement system (≈ USD 9,900).

Statistics and AI

Table 3 shows the literature summary on statistical and AI methods. The studies in the review used either statistical or AI-based methods to analyze the data collected by the proposed technologies. 12 of 16 (75%) studies used traditional statistical methods and the remaining 4 studies implemented AI-based methods.

AI-based frailty classification and prediction

The studies that adopted AI-based methods trained machine learning models to classify older adults into different frailty levels such as frail, pre-frail, or non-frail. Heterogeneous machine learning algorithms such as logistic regression, naïve bayes, and neural networks (Chang et al., 2013; Goonawardene et al., 2018; Tegou et al., 2019) were used to build a model for classifying frailty levels. Goonaward et al. reported that the logistic regression yielded the highest AUC (0.98) in assessing if an elderly is frail or robust using 7 features of in-home living patterns. In contrast, Tegou et al. reported the random forest method outperformed other methods with an 82.33% accuracy in classifying frailty status. Chang et al. trained their own artificial neural networks with 11 neurons (sensor features) in the input layers and 2 neurons (pre-frail and normal) as output. However, none of these machine learning models were used to predict when a person would progress into frailty from a pre-frail status in the future.

Statistical methods
Statistical methods were used for testing the statistically significant associations of different frailty related parameters with clinical frailty. For instance, Analysis of Variation and t-tests were used to test differences in each sensor-derived clinical parameter between different classes of frailty (Bastone et al., 2015; Toosizadeh et al., 2015). Similarly, general linear model tests were used to compare postural transitions parameters between frailty groups. Regression is the most commonly used method to determine independent predictors of frailty among sensor-derived parameters. Multiple studies used logistic regression (Parvaneh et al., 2017; Schwenk et al., 2015; Toosizadeh et al., 2015). For instance, Schwenk et al. used multinomial logistic regression to analyze temporal-spatial gait parameters, postural balance parameters, and physical activity parameters extracted from data collected by 5 inertial sensors from 125 participants (Schwenk et al., 2015). They found that stride length (AUC=0.86), double support (AUC=0.84), and walking bout duration variability (AUC=0.82) are the most sensitive parameters for discriminating between frail, pre-frail, and non-frail levels. In contrast, Chen et al. used univariate and multivariate logistic analyses to calculate odds ratios (OR) for each frailty related factor and revealed high ORs for multiple factors such as age increment, alcohol consumption, and social engagement (Chen et al., 2015). Notably, the only longitudinal study that reported predictability of frailty used generalized estimate equation to analyze steps and activity intensity data (Yuki et al., 2019). These data were collected from a cohort of community-dwelling Japanese older adults with 10 years follow-up. They reported ORs for development of frailty with an increase in 1000 steps (OR = 0.92), an increase in 10 minutes of light physical activity (OR = 0.96), and moderate to vigorous physical activity (OR = 0.83).

Discussion

To our knowledge, this is the first systematic review summarizing and discussing home-based sensor and AI technology for frailty assessment and prediction and assessing these methods using a clinical evidence scale.

One notable body-worn sensor technology in this review is the IMU sensor. The IMU sensor was the most common sensor technology in our reviewed studies, which provides evidence for its validation and good efficacy of the body-worn IMU-based technology in assessing frailty. This homogeneity in the studies was notable since the IMU sensor can only measure gait and activity-related parameters. Technologies measuring new criteria of physical frailty such as muscle, weight, and exhaustion, or measuring the same gait criteria but used a different way than extant literature, could be a future research area for assessing physical frailty. For example, the exhaustion criteria of the Fried's frailty index were not assessed with the use of technology, and the activity criteria were mostly measured by IMU sensor. Perhaps a more natural and interactive method powered by technologies such as natural language processing (NLP) could be used to measure these criteria. The NLP technology could ask the individual about their level of exhaustion or daily activity. These new technologies are still yet to be tested and validated.

In addition, studies that tested the body-worn IMU sensors, muscle strength meter and portable metabolic measurement system were conducted in supervised settings and therefore have two major limitations: (1) it requires training or clinician's supervision; (2) accuracy of the measurement for frailty depends on the high level of user compliance to wearing the technology. In other words, few of the evaluated technologies can be used independently with zero effort. Considering the application in home settings, the limitations of the technology creates considerable barriers for older adults themselves to use at home if training is not adequate and sustainable or a clinician is not able to be present to supervise.

The ecological automatic measurement of grip strength by home-based technology is found to be understudied and not validated as of yet. Only one study (Okubo et al., 2018) in the reviewed studies used an unidentified hand-grip strength device as part of a functional assessment. It was not clear in the paper if the hand-grip strength assessment was administered with supervision or automatically. In addition, standardization of hand-grip strength measurement was not discussed either. More research is needed for investigating the feasibility and standardization of ecological automatic grip strength measurement by home-based technology which is why it is not proposed for this study.

Furthermore, the exclusion of older adults with more severe cognitive impairment and mobility issues from the samples was an interesting finding with implications to the clinical relevance of this body of work. As the severity of cognitive impairment is associated with the physical decline in older adults (Laurin D et al., 2001) and at high risk of frailty, a wider sample that reflects the range of cognitive and mobility impairments in this population is required. Similarly, eight studies also excluded older adults with mobility issues including unable to ambulate for a certain period of time or distance, musculoskeletal injuries, severe handicaps, and severe balance impairments. The reason for excluding them was related to the specific mobility test that the technology required for assessing frailty. However, older adults commonly have mobility impairments which increase their risk of being physically frail. This prompts us as to whether technologies that are not relying on mobility tests could be used to assess frailty especially for people with mobility issues.

Although two studies (Goonawardene et al., 2018; Yuki et al., 2019) followed up participants for 1 year and 10 years respectively, other studies being reviewed were cross-sectional and unable to capture the progression of frailty over time, therefore the capability of AI to predict frailty is still to be determined.

To the best of our knowledge, there is no review that has synthesized the statistical and AI methods used for assessing and predicting frailty. From the review results, we suggest more heterogeneous machine learning algorithms can be investigated for predicting frailty progression. In the meantime, more longitudinal data that is collected for many years are required to study the transition between healthy and frail conditions. The technology to assess and predict frailty in homes can also be used to assess frailty in other settings like acute care once it has been accurately tested in homes. This is an area for future research. In addition, another frailty test may be better suited for acute care as mentioned above.

Strengths and limitations

Our review has several strengths:

(1) Extensive database searches.
(2) The first systematic review that focuses on home-based technology for frailty assessment and prediction.
(3) The only systematic review on technology and frailty that has reviewed the use of AI and statistical methods for frailty data analysis.
(4) The only systematic review on technology and frailty that evaluated the level of evidence of reviewed studies. Our review also has limitations. We focused our search strategy on technology specifically for assessing frailty (as a general term) as well as physical phenotypes of frailty. As frailty could also present in the form of cognitive frail or social frail, this review did not search the phenotypes that could contribute to the detection of cognitive or social frailty, unless those articles also contain the term frailty which could be included in our search strategy. We also only included research conducted in English.

Conclusions

This systematic review focused on the use of technology for frailty assessment and prediction with particular attention to home-based technology and AI. Our findings highlight several gaps in this field for future work. First, more different types of technology other than IMU-based wearable sensors should be studied for evaluating the efficacy in assessing frailty as existing studies concentrated too much on IMU-based technology. Second, there are opportunities for technology to measure frailty criteria that have not been extensively measured by existing research such as exhaustion, muscle strength, weight loss in the gold standard Fried' frailty index. For any new technology that measures frailty, it is recommended to validate the technology assessment with clinical standard reference in order to prove its clinical evidence. Third, few longitudinal studies were found in the review, however, this type of study should be a future direction because longitudinal data is essential for AI to track an individual's progress in frailty, which the cross-sectional study wouldn't be able to do. Home-based continuous frailty monitoring is ideal to capture frailty progress as frailty phenotypes usually express in daily life.

References

Alqahtani, B. A., Ferchak, M. A., Huppert, T. J., Sejdic, E., Perera, S., Greenspan, S. L., & Sparto, P. J. (2017). Standing balance and strength measurements in older adults living in residential care communities. Aging Clinical and Experimental Research, 29(5), 1021–1030.

Azimi, I., Rahmani, A. M., Liljeberg, P., & Tenhunen, H. (2017). Internet of things for remote elderly monitoring: a study from user-centered perspective. Journal of Ambient Intelligence and Humanized Computing, 8(2), 273–289. https://doi.org/10.1007/s12652-016-0387-y

Bastone, A. de C., Ferriolli, E., Teixeira Patricia, C., Dias Marcos, J., & Dias Corrêa, R. (2015). Aerobic Fitness and Habitual Physical Activity in Frail and Nonfrail Community-Dwelling Elderly. Journal of Physical Activity & Health, 12(9), 1304.

Bruyère, O., Buckinx, F., Beaudart, C., Reginster, J. Y., Bauer, J., Cederholm, T., Cesari, M. (2017). How clinical practitioners assess frailty in their daily practice: an international survey. Aging Clinical and Experimental Research, 29(5), 905–912. https://doi.org/10.1007/s40520-017-0806-8

CEBM Levels of Evidence Working Group. (2009). The Oxford Levels of Evidence 1. Oxford Centre for Evidence-Based Medicine. Retrieved from https://www.cebm.net/index.aspx?o=5653

Chang, Y. C., Lin, C. C., Lin, P. H., Chen, C. C., Lee, R. G., Huang, J. S., & Tsai, T. H. (2013). eFurniture for home-based frailty detection using artificial neural networks and wireless sensors. Medical Engineering & Physics, 35(2), 263–268.

Chen, S., Honda, T., Chen, T., Narazaki, K., Haeuchi, Y., Supartini, A., & Kumagai, S. (2015). Screening for frailty phenotype with objectively-measured physical activity in a west Japanese suburban community: evidence from the Sasaguri Genkimon Study. BMC Geriatrics, 15(1), 36.

Collard, R. M., Boter, H., Schoevers, R. A., & Oude Voshaar, R. C. (2012). Prevalence of frailty in community-dwelling older persons: A systematic review. Journal of the American Geriatrics Society, 60(8), 1487–1492. https://doi.org/10.1111/j.1532-5415.2012.04054.x

Geraedts, H. A. E., Zijlstra, W., Van Keeken, H. G., Zhang, W., & Stevens, M. (2015). Validation and user evaluation of a sensor-based method for detecting mobility-related activities in older adults. PLoS ONE, 10(9), 1–11.

Gianaria, E., Grangetto, M., Roppolo, M., Mulasso, A., & Rabaglietti, E. (2016). Kinect-based gait analysis for automatic frailty syndrome assessment. Proceedings - International Conference on Image Processing, ICIP, 2016-Augus, 1314–1318. https://doi.org/10.1109/ICIP.2016.7532571

Goonawardene, N., Tan, H., & Tan, L. B. (2018). Unobtrusive Detection of Frailty in Older Adults. Human Aspects of IT for the Aged Population. Applications in Health, Assistance, and Entertainment, 290–302. Springer International Publishing.

Greene, B. R., Doheny, E. P., Kenny, R. A., & Caulfield, B. (2014). Classification of frailty and falls history using a combination of sensor-based mobility assessments. Physiological Measurement, 35(10), 2053–2066. https://doi.org/10.1088/0967-3334/35/10/2053

Greene, B. R., Doheny, E. P., O'Halloran, A., & Kenny, R. A. (2014). Frailty status can be accurately assessed using inertial sensors and the TUG test. Age and Ageing, 43(3), 406–411. https://doi.org/10.1093/ageing/aft176

Jansen, F. M., Prins, R. G., Etman, A., Van Der Ploeg, H. P., De Vries, S. I., Van Lenthe, F. J., & Pierik, F. H. (2015). Physical activity in non-frail and frail older adults. PLoS ONE, 10(4), 1–15.

Laurin D, Verreault R, Lindsay J, MacPherson K, & Rockwood K. (2001). Physical activity and risk of cognitive impairment and dementia in elderly persons. Archives of Neurology, 58(3), 498–504. https://doi.org/10.1001/archneur.58.3.498

Mainetti, L., Patrono, L., Secco, A., & Sergi, I. (2017). An IoT-Aware AAL system to capture behavioral changes of elderly people. Journal of Communications Software and Systems, 13(2), 68–76.

Mitnitski, A., Collerton, J., Martin-Ruiz, C., Jagger, C., Von Zglinicki, T., Rockwood, K., & Kirkwood, T. B. L. (2015). Age-related frailty and its association with biological markers of ageing. BMC Medicine.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Medicine, 6(7).

Mohler, J., Huyvetter, K. D., Schwenk, M., Wendel, C., Grewal, G., Fain, M. J., Najafi, B. (2013). In-Home Activity Monitoring in Frail Elders: A New Measure of Function. Journal of the American Geriatrics Society, 61, S48-49.

Morley, J. E., Vellas, B., Abellan van Kan, G., Anker, S. D., Bauer, J. M., Bernabei, R., Walston, J. (2013). Frailty consensus: A call to action. Journal of the American Medical Directors Association, 14(6), 392–397. https://doi.org/10.1016/j.jamda.2013.03.022

Okubo, Y., Nemoto, M., Osuka, Y., Jung, S., Seino, S., Figueroa, R., Tanaka, K. (2018). Development of the Nutrition and Functionality Assessment (NFA) Among Older Adults in Japan. Journal of Frailty and Aging, 7(3), 187–192.

Parvaneh, S., Mohler, J., Toosizadeh, N., Grewal, G. S., & Najafi, B. (2017). Postural Transitions during Activities of Daily Living Could Identify Frailty Status: Application of Wearable Technology to Identify Frailty during Unsupervised Condition. Gerontology, 63(5), 479–487.

Puts, M. T. E., Toubasi, S., Atkinson, E., Ayala, A. P., Andrew, M., Ashe, M. C., McGilton, K. S. (2016). Interventions to prevent or reduce the level of frailty in community-dwelling older adults: a protocol for a scoping review of the literature and international policies. BMJ Open, 6(3), e010959. https://doi.org/10.1136/bmjopen-2015-010959

Rainoldi, A., Donini, L. M., Feletti, L. C., Zia, G., & Signore, S. Del. (2018). Continuous ICT enabled recording vs . standard functional assessment can reliably indicate physical frailty status in community dwelling older adults. 11th International Conference on Cachexia, Sarcopenia and Muscle Wasting. Maastricht,The Netherlands.

Razjouyan, J., Naik, A. D., Horstman, M. J., Kunik, M. E., Amirmazaheri, M., Zhou, H., Najafi, B. (2018). Wearable sensors and the assessment of frailty among vulnerable older adults: An observational cohort study. Sensors (Switzerland), 18(5), 1–17. https://doi.org/10.3390/s18051336

Sackett, D. L. (1989). Rules of evidence and clinical recommendations on the use of antithrombotic agents. Chest, 95(2 Suppl), 2S-4S. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2914516

Schwenk, M., Mohler, J., Wendel, C., D'Huyvetter, K., Fain, M., Taylor-Piliae, R., & Najafi, B. (2015). Wearable sensor-based in-home assessment of gait, balance, and physical activity for discrimination of frailty status: Baseline results of the Arizona frailty cohort study. Gerontology, 61(3), 258–267.

Tegou, T., Kalamaras, I., Tsipouras, M., Giannakeas, N., & Votis, K. (2019). A Low-Cost Indoor Activity Monitoring System for Detecting Frailty in Older Adults.

Toosizadeh, N., Mohler, J., & Najafi, B. (2015). Assessing upper extremity motion: An innovative method to identify frailty. Journal of the American Geriatrics Society, 63(6), 1181–1186.

WHO. (2015). World Report on Ageing and Health.

Yuki, A., Otsuka, R., Tange, C., & Nishita, Y. (2019). Daily Physical Activity Predicts Frailty Development Among Community-Dwelling Older Japanese Adults. Journal of the American Medical Directors Association.

Zouba, N., Bremond, F., Thonnat, M., Anfosso, A., Pascual, É., Malléa, P., Guerin, O. (2009). A computer system to monitor older adults at home: Preliminary results. Gerontechnology, 8(3), 129–139.

Appendix I

Search strategy

Initial search terms were compiled and iteratively refined by content experts in the fields of library science, geriatrics, and technology. The following search strategy was used in the Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present> database: 1 Frail Elderly/ (9135), 2 frail*.mp. (18933), 3 prefrail*.mp. (261), 4 or/1-3 (18944), 5 exp Community Health Services/ (281911), 6 home*.mp. (506872), 7 house*.mp. (162253), 8 community*.mp. (472943), 9 residential*.mp. (31372), 10 "ag???? at home".mp. (80), 11 "ag???? in place".mp. (541), 12 assisted living.mp. (2302), 13 or/5-12 (1219972), 14 (smart adj2 (home? or house?)).mp. (349), 15 Signal Processing, Computer-Assisted/ (42735), 16 exp Accelerometry/ (5497), 17 (acceleromet* or actigraph*).mp. (17012), 18 wearable?.mp. (5474), 19 (sensor or sensors or sensing).mp. (161535), 20 exp Artificial Intelligence/ (72528), 21 artificial intelligence/ (21628), 22 machine learn*.mp. (12840), 23 wrist*.mp. (40720), 24 ((worn or wear) adj2 (body or wrist)).mp. (1109), 25 Wireless Technology/ (2505), 26 wireless*.mp. (10983), 27 exp Video Recording/ (36894), 28 (video* or videocamera* or camera*).mp. (162924), 29 (pressure adj2 mat).mp. (175), 30 Technology/ (8526), 31 exp Biomedical Technology/ (11446), 32 (gerontech* or tech or technolog*).mp. (435624), 33 gyro*.mp. (3706), 34 detector*.mp. (63113), 35 gauge*1.mp. (26357), 36 ??????meter*.mp. (27398), 37 exp Telemedicine/ (22258), 38 exp Telemetry/ (11554), 39 telehealth*.mp. (3245), 40 telemed*.mp. (20681), 41 telecar*.mp.(942), 42 telemonitor*.mp. (1214), 43 telesurveillanc*.mp. (22), 44 "internet of things".mp. (473), 45 IOT.mp. (408), 46 monitoring, physiologic/ (51408), 47 monitoring, ambulatory/ (7358), 48 (monitor* adj2 (tele* or health* or ambulator*)).mp. (27212), 49 or/14-48 (1079564), 50 4 and 13 and 49 (383), 51 50 not (exp animals/ not exp humans/) (383), 52 limit 51 to english language (361), 53 limit 52 to yr="2000 -Current" (315), 54 remove duplicates from 53 (298). Search strategies applied in the other databases were derived from the MEDLINE search. Reference lists of relevant articles were subsequently hand-searched to identify additional papers.

Abbreviation

ANOVA – Analysis of Variance
AUC – Area Under Curve
ISAR-HP - Identification Seniors at Risk - Hospitalized Patients' questionnaire
GFI - Groningen Frailty Index
OR – Odds ratio
PA - Physical Activity
SVM – Support Vector Machine
TFI - Tilburg Frailty Index
TUG - Timed Up and Go