



Published in final edited form as:

*Ageing Res Rev.* 2016 March ; 26: 53–61. doi:10.1016/j.arr.2015.12.003.

## Frailty assessment instruments: Systematic characterization of the uses and contexts of highly-cited instruments

Brian J. Buta<sup>a,b</sup>, Jeremy D. Walston<sup>a,b</sup>, Job G. Godino<sup>c,d</sup>, Minsun Park<sup>d</sup>, Rita R. Kalyani<sup>a,b</sup>, Qian-Li Xue<sup>a,b</sup>, Karen Bandeen-Roche<sup>a,d</sup>, and Ravi Varadhan<sup>a,b,e,\*</sup>

<sup>a</sup>Johns Hopkins Older Americans Independence Center and the Center on Aging and Health, Johns Hopkins University, Baltimore, MD, United States

<sup>b</sup>School of Medicine, Johns Hopkins University, Baltimore, MD, United States

<sup>c</sup>Center for Wireless and Population Health Systems, University of CA, San Diego, United States

<sup>d</sup>Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

<sup>e</sup>Department of Oncology, Division of Biostatistics and Bioinformatics, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, United States

### Abstract

The medical syndrome of frailty is widely recognized, yet debate remains over how best to measure it in clinical and research settings. This study reviewed the frailty-related research literature by (a) comprehensively cataloging the wide array of instruments that have been utilized to measure frailty, and (b) systematically categorizing the different purposes and contexts of use for frailty instruments frequently cited in the research literature. We identified 67 frailty instruments total; of these, nine were highly-cited ( > 200 citations). We randomly sampled and reviewed 545 English-language articles citing at least one highly-cited instrument. We estimated the total number of uses, and classified use into eight categories: risk assessment for adverse health outcomes (31% of all uses); etiological studies of frailty (22%); methodology studies (14%); biomarker studies (12%); inclusion/exclusion criteria (10%); estimating prevalence as primary goal (5%); clinical decision-making (2%); and interventional targeting (2%). The most common assessment context was observational studies of older community-dwelling adults. Physical Frailty Phenotype was the most used frailty instrument in the research literature, followed by the Deficit Accumulation Index and the Vulnerable Elders Survey. This study provides an empirical evaluation of the current uses of frailty instruments, which may be important to consider when selecting instruments for clinical or research purposes. We recommend careful consideration in the selection of a frailty instrument based on the intended purpose, domains captured, and how the instrument has been used in the past. Continued efforts are needed to study the validity and feasibility of these instruments.

\*Corresponding author at: Division of Biostatistics and Bioinformatics, Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, 550 N. Broadway, Suite1111-E, Baltimore, MD, United States. Fax: +1 410 955 0859. ravi.varadhan@jhu.edu (R. Varadhan).

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.arr.2015.12.003>.

## Keywords

Frailty assessment; Instrument; Review; Operational definition

---

## 1. Introduction

As the population ages, a central focus of health care providers is to understand, and beneficially intervene upon, the factors that place older adults at elevated risk of precipitous declines in health and function. The syndrome of frailty has been hypothesized to represent such risk, in particular the increased vulnerability to stressors (e.g. infection, injury, changes in medication) that characterizes many older adults (Fried et al., 2001; Bandeen-Roche et al., 2006; Varadhan et al., 2008).

While frailty is widely recognized, there continues to be considerable debate over how best to assess it. Many operational definitions have been introduced to attempt to distinguish frail from non-frail older adults (Gobbens et al., 2010; Hogan et al., 2003). These definitions vary in their conceptual underpinnings, clinical practicality, domains, and assessment items (Bouillon et al., 2013; de Vries et al., 2011; Gobbens et al., 2010; Sternberg et al., 2011). There appears to be general agreement that operational definitions of frailty should be: multi-dimensional; exclusive of disability and, possibly, of comorbidity; dynamic; predictively valid for adverse outcomes; and feasible (Gobbens et al., 2010; Hogan et al., 2003). However, instrument variability has led to controversy over which frailty assessment instrument is appropriate in which context, and importantly, what is actually being assessed (for example, frailty versus disability) depending on the chosen instrument. Recent reviews of frailty instruments have highlighted the need for greater reliability and validity testing (Bouillon et al., 2013; de Vries et al., 2011). A systematic review by Sternberg et al., 2011 concluded that the needs and goals of the study or clinic may determine the most suitable frailty instrument, similar to the perspectives of Martin and Brighton (2008) and Cesari et al. (2014a).

A consensus-building effort by Rodríguez-Mañas et al. (2013) led to agreement on a conceptual framework for frailty, the inclusion of specific domains, and its distinction from disability, but no consensus on an overall operational definition of frailty was reached. A separate consensus effort by Morley et al. (2013) to define frailty reached agreement on four key points related to the assessment of physical frailty: (1) it is an important medical syndrome; (2) it can potentially be targeted and treated; (3) there are available screening tests; and (4) all persons 70 years and older should be screened. A published response to the second effort called for careful attention to the choice of instruments for frailty assessment and their validation and refinement (Xue and Varadhan, 2014).

Frailty research is evolving rapidly, with multiple frailty studies published every year despite the relative lack of validation studies and refinement efforts needed to maximize the clinical utility and reproducibility of frailty assessment. Questions such as “what is the best definition of frailty?” and “which instrument should be used to assess frailty?” are often posed, although no answer is readily available. A plausible reason as to why standardization and consensus efforts have been unfruitful is that they have not explicitly considered the

purpose and the context of frailty assessment. Though answering the above questions are important, the primary goals of this study were to gain insight into the spectrum of original studies, reviews and other types of articles that comprise the current frailty-related research literature, and to better understand whether high citation counts truly equated with wide use of an instrument, or if, perhaps, citations were more indicative of references in reviews or other types of papers. To accomplish this goal, we aimed to comprehensively catalog the wide array of instruments that have been utilized to measure frailty and provide an empirical foundation of the various purposes and contexts in which highly-cited instruments have been used. Trends of instrument uses, along with further considerations of theory, validity, and feasibility, can help to guide the development, selection and implementation of appropriate frailty instruments in the future, where “appropriate” means matching the assessment instrument to the purpose and context.

## 2. Methods

### 2.1. Literature search and inclusion to identify frailty instruments

A search strategy was developed to identify frailty instruments using the following steps (see Appendix 1 for flow chart of the search strategy):

1. We first performed a PubMed database search using the “frail elderly” MESH term in combination with the term “instrument.” Relevant articles that included frailty assessment instruments were identified by reading the abstract and, when necessary, the full article. This search was performed from the start of the database through December 2013.
2. In our PubMed search, we identified three recent review papers that have examined the components and domains of frailty instruments (Bouillon et al., 2013; de Vries et al., 2011; Sternberg et al., 2011); two were found directly in the search results (Bouillon et al., 2013; **de Vries et al., 2011**) and the third article (Sternberg et al., 2011) was found by perusing the references of the review by Bouillon and colleagues.
3. We then screened the three identified review papers and found additional instruments not found in our PubMed search that met our inclusion criteria (described below).
4. Lastly, as we conducted the citation review described below, additional frailty assessment instruments were found in the literature that met our inclusion criteria.

For inclusion, we defined a frailty instrument as a specific and reproducible set of criteria for assessing frailty status. For each frailty instrument identified, we determined the instrument’s seed article(s) where the definition of the instrument for measuring frailty was first published. Generally, one article served as the seed article but in two cases (Deficit Accumulation Index and FRAIL Scale) the authors agreed that multiple articles were appropriate because each could serve as a reasonable seed for the instrument’s definition. Seed articles were identified from the searches above or by searching the references of the articles found in the searches above. The authors reached consensus on which article(s) qualified as the seed article(s) for each instrument.

## 2.2. Citation search to identify highly-cited frailty instruments

We performed citation searches of the agreed upon seed article(s) for each identified frailty instrument. Though PubMed is an outstanding search tool for keyword search (Linder et al., 2015), we selected the Web of Science (WoS) for our citation searches based on (a) its emphasis on top tier journals and high quality publications (<http://wokinfo.com/essays/journal-selection-process/>) and (b) its longer period of citation analysis coverage when compared to other databases (Falagas et al., 2008). Additionally PubMed includes less journals in its database than WoS (Falagas et al., 2008). We searched citations in WoS from the start of the year of the earliest identified seed articles, January 1991, through December 2013. Seed articles with at least 200 ( $\pm 5$ ) citations were used to denote the most highly-cited frailty instruments. For instruments with more than one seed article, the total number of de-duplicated citing articles had to be at least 200.

## 2.3. Random sampling of citing articles per instrument and characterizing the randomly-selected literature

In order to feasibly characterize the sizable number of citations, we selected a random sample of 20% of citations for eight of the highly-cited frailty instruments' seed article(s), all except for the PFP in which a 10% random sample was selected due to the comparatively large number of citations. In addition to feasibility, we chose random sampling because it helps to ensure that our sample is a reasonable depiction of trends among all citations. This type of random sampling approach has been performed in previous review studies (Annalingam et al., 2014; Burgess et al., 2006). Simple random sampling was performed using R statistical software (R Core Team, 2014, Vienna, Austria). Only English language articles were reviewed.

For each highly-cited frailty instrument, the randomly selected citations were assigned to two of four reviewers (BJB, JGG, MP, RV). Reviewers read and assessed the full articles for the following characteristics: type of article (original, review, commentary, or other); "actual use," which we define as whether frailty status was assessed using the given instrument (or some minor modification of it); and if yes, how the frailty variable was used (purpose); study design and setting (context), and results. A meeting of the four reviewers occurred to discuss the random sample of citations for each frailty instrument (multiple meetings took place for instruments with larger samples) and to reach 100% consensus on the characterization of each sampled article.

## 2.4. Estimating use of the frailty instruments in the literature from the random samples

To estimate total use of the frailty instruments in the literature, we calculated the per-instrument proportions of actual use found in each random sample and multiplied by the total number of citing articles per instrument (i.e., total estimated number of use = (number of use in the random sample/size of reviewed random sample)  $\times$  total number of citing articles). We then calculated the confidence intervals for the estimate using exact binomial confidence interval approach (using "binom.test" function in the R statistical software).

## 2.5. Systematic categorization of the uses and contexts of the frailty instruments

If an article included actual use of the frailty instrument, we determined a categorization of use through reviewer consensus. Actual uses of frailty instruments only occurred in original research studies; therefore this categorization of use only occurred in original studies. Using prior knowledge and based on the articles reviewed, we arrived at a list of possible uses (see below) and contexts (i.e., cohort studies, randomized controlled trials, or clinical assessment) for the frailty instruments. Articles may have used frailty instruments for more than one purpose, in which case we categorized a primary use for each, as well as additional uses where appropriate. We report primary uses which fell under one of the following eight categories and definitions:

- Frailty as a risk factor for adverse health outcomes: Research where frailty has been studied as a predictor of or a risk factor for various age-associated adverse health outcomes including disability and death;
- Risk factors for frailty: Studies where individual characteristics of subjects (e.g., age, socio-demographic variables, comorbid conditions) have been assessed as to whether they are associated with an increased risk of frailty;
- Methodology: Studies where the goal was to compare different frailty assessment instruments or to evaluate the clinimetric properties (e.g. reliability, predictive validity) of a particular frailty instrument;
- Biomarkers of frailty: Studies which assessed whether specific biomarkers (e.g. markers assayed in serum including inflammatory cytokines, hormones) were elevated or depressed in frailty;
- Frailty as inclusion/exclusion criteria: Studies where frailty status is among the criteria used to screen subjects into the study;
- Estimation of frailty prevalence: Studies where the main goal is to estimate the prevalence of frailty among older adults with certain characteristics (e.g., subjects with dementia) or in a particular context (e.g., living in nursing home);
- Frailty as a guide for clinical decision making: A broad category; examples include the use of frailty status as one of the criteria to evaluate the fitness of an older patient to undergo a surgical procedure; or to manage a particular disease differently in a frail older adult compared to a robust older adult;
- Frailty as a target for intervention: Studies where frailty status was assessed as an outcome measure to monitor the impact of an intervention.

To estimate total number of uses in the literature under each category, we used the same approach described above in §2.4, per category.

## 2.6. Validation

To validate our approach for estimating the extent of use of frailty instruments, we reviewed all articles that reference the Vulnerable Elders Survey (VES-13) and compared our findings to those based on random sampling. We chose the VES-13 because given its total number of

citations ( $n = 225$ ) it was one of the instruments that was feasible to review exhaustively. Our validation aimed to ensure that our estimates were reasonable.

### 3. Results

#### 3.1. Identification of frailty instruments

The PubMed search using the “frail elderly” MESH term and keyword “instrument” resulted in 132 articles. Among these, we identified 28 unique frailty assessment instruments. We identified another 23 unique, non-duplicative instruments from existing systematic review papers on the components and domains of frailty instruments (Bouillon et al., 2013; de Vries et al., 2011; Sternberg et al., 2011). As we conducted the following review, 16 additional frailty instruments were found in the literature. In total, 67 frailty instruments were identified (see website: <https://jhpeppercenter.jhmi.edu/FrailtyTool/InstrumentSummaryList.aspx>). Of note, the majority of the nine highly-cited instruments discussed in the next section were identified through our initial PubMed search; the remaining two highly-cited instruments were found in the existing review papers.

#### 3.2. Identification of highly-cited frailty instruments

We identified nine instruments with seed article(s) that are highly cited: (1) Physical Frailty Phenotype (PFP, also called CHS frailty phenotype) (Fried et al., 2001); (2) Deficit Accumulation Index (DAI, also called Frailty Index) (Mitnitski et al., 2001, 2004; Rockwood et al., 2006, 2007; Rockwood and Mitnitski, 2007); (3) Gill Frailty Measure (Gill et al., 2002); (4) Frailty/Vigor Assessment (Speechley and Tinetti, 1991); (5) Clinical Frailty Scale (Rockwood et al., 2005); (6) Brief Frailty Instrument (Rockwood et al., 1999); (7) Vulnerable Elders Survey (VES-13) (Saliba et al., 2001); (8) FRAIL Scale (Abellan van Kan et al., 2008a,b); and (9) Winograd Screening Instrument (Winograd et al., 1991). See Table 1. Each highly-cited instrument included some measure of physical function, though the inclusion of other domains (e.g., physical activity, cognition) varied; see Table 2. Appendix 2 includes additional information on the nine highly-cited instruments (including domains, items, and scoring).

Fig. 1 displays the number of citing articles (on a logarithmic scale) over time for each of the highly-cited frailty instruments. The three most-cited instruments were: the Physical Frailty Phenotype (PFP; 1891 citing articles); the Deficit Accumulation Index (DAI; 401 citing articles); and the Gill Frailty Measure (254 citing articles).

#### 3.3. Randomly sampled citations per instrument, and characteristics of the sample

In total, 591 articles were randomly sampled; of the sampled articles, 46 were unavailable or written in a language other than English. The remaining 545 randomly-selected articles were reviewed and characterized. On average, 63% of the randomly selected articles were original research studies. Details of the randomly sampled articles are provided in Table 1.

#### 3.4. Estimated uses of the frailty instruments in the literature

Actual uses of frailty instruments occurred in original research studies only. The percentage of citations in each random sample that included frailty assessment using the given

instrument ranged from 2.1% (using the Frailty/Vigor Assessment) to 29.9% (using the Deficit Accumulation Index). Using these percentages, we estimated the total number of original research study articles that have measured frailty using each of the highly-cited instruments (see Table 1). Based on this approach, the Physical Frailty Phenotype (PFP) – 440 total uses estimated among 1891 citing articles – was the most utilized frailty instrument in the research literature, followed by the Deficit Accumulation Index (DAI), with 116 total uses estimated among 401 citing articles, and the Vulnerable Elders Survey (VES-13), with 49 total uses estimated among 225 citing articles.

### 3.5. Systematic categorization of the uses and contexts of the frailty instruments

As shown in Table 3, per category of use, the most commonly identified purpose was to assess frailty as a risk factor for adverse health outcomes (31.3% of all estimated instrument uses), which was employed in seven out of the nine highly-cited instruments reviewed and most frequently with the PFP. Identifying risk factors for frailty was the next most frequent type of use (22.3%). This etiologic research was performed only in studies that used the PFP (121 estimated uses in this category) or the DAI (37 estimated uses in this category). The third most frequent type of use was methodology (13.9%), followed by biomarkers of frailty (12.4%) and then frailty as inclusion/exclusion criteria (9.9%). There was a relative paucity of uses for frailty as a guide for clinical decision-making (2.3%) and frailty as a target for intervention (2.3%). Fig. 2 summarizes the estimated uses per category.

Regarding context of use, the study designs and settings (e.g., clinical vs. epidemiological) varied among the highly-cited frailty instruments. Instruments were used among community-dwelling and hospitalized/institutionalized older adults; in observational studies, interventions, and randomized trials; in studies with sample sizes ranging from 15 participants to over 60,000; and in more than 20 countries. The PFP was used in epidemiological, clinic-based and intervention studies. The DAI was used in several large database studies, and in epidemiological and clinical studies. For the Gill Frailty Measure and FRAIL Scale, frailty assessment was solely among community-based cohorts. For the Clinical Frailty Scale, Brief Frailty Instrument, VES-13, and Winograd screening instrument, assessment took place solely in a hospital or clinical context. Use of the Frailty/Vigor Assessment occurred in a longterm care setting. Additional information on usage patterns and contexts per instrument are provided in Appendix 3.

### 3.6. Validation results

To validate our random sampling methodology, we reviewed all articles that reference the VES-13 seed article (Saliba et al., 2001). We found 65 articles in which the VES-13 was used to assess frailty. This was higher than our estimate of 49 uses, though well within the confidence interval (95% CI: 24, 85) of our estimate.

## 4. Discussion

Our goal was to comprehensively identify frailty assessment instruments and to categorize the use of highly-cited frailty instruments in the literature. We aimed to provide an initial foundation for the selection of frailty instruments given an intended purpose and context.

We identified 67 frailty instruments, 9 of which were highly-cited. Physical Frailty Phenotype (PFP), Deficit Accumulation Index (DAI), and the Gill Frailty Measure were the most cited. Citations for the Clinical Frailty Scale and the FRAIL Scale (published in 2005 and 2008, respectively) appeared to increase at rates similar to the PFP and DAI (see Fig. 1), though whether this growth rate will continue is to be determined. Among highly-cited frailty instruments, the PFP, DAI and Vulnerable Elders Survey (VES-13) were the most used instruments. Frailty was most commonly assessed as a risk factor for adverse health outcomes such as death, institutionalization, and falls, which aligns with prior reviews (de Vries et al., 2011; Sternberg et al., 2011). We identified other common uses (etiology, methodology, biomarkers) and found a paucity of clinical and interventional uses. Observational cohort studies that assess frailty among community-dwelling older adults were most common, though some instruments have been solely used in clinical settings.

Our study is the first to categorize the different purposes for which highly-cited frailty instruments have been employed. Frailty assessment was commonly used as a risk factor for adverse health outcomes, generally and among most highly-cited instruments. Only the PFP and the DAI were found to assess risk factors for frailty; and biomarker studies were found among a limited number of instruments. Also, frailty assessment as used for clinical decision-making or as an interventional target was found to be scarce. For clinical decision-making, the small number of uses estimated may be due to the evolving focus on the clinical applicability of frailty (Rodríguez-Mañas and Fried, 2015; Sourial et al., 2013); given this, more clinically-focused uses will likely emerge in the coming years. Also, some recently recommended instruments for clinical frailty screening (Pialoux et al., 2012; Turner and Clegg, 2014; Vellas et al., 2013) are not among the highest cited and therefore were not categorized in our study. For frailty as an interventional target, the small number of uses estimated in our study aligns with other research. A recent systematic review found only 11 interventional studies with frailty as an outcome (Lee et al., 2012); among these studies, 4 used the PFP (or a modified version) and 7 used less common frailty indicators. Also, two recent interventional studies measured frailty outcomes using the PFP (Fairhall et al., 2015; Cesari et al., 2014b). Overall, few studies have focused on interventions to ameliorate frailty syndrome, hence this points to a critical need in the frailty literature.

The highly-cited instruments reviewed here stem from different underlying theories of frailty, and each has a unique set of measures. The notion of frailty as vulnerability to poor outcomes is a commonality and declining or poor physical function is a common domain. But as shown in Table 2, operationalization varies across instruments. The distinct components and varied uses of these differing instruments highlight the need for careful reflection on what is being measured in frailty assessment and the appropriateness of the measures for a given study or clinic. This is especially true in distinguishing between assessments for risk stratification versus etiological investigation of frailty (Xue and Varadhan, 2014). It is also true when considering instruments that include measures of disability and comorbidity. Over half of the highly-cited instruments include measures of disability or comorbidity (Table 2; Appendix 2). Disability, especially, has been agreed to be a distinct entity from frailty (Rodríguez-Mañas et al., 2013). Others have argued that “frailty is distinct from, but overlapping with, both comorbidity and disability” (Fried et al., 2004),



and that definitions of frailty should “not include disease, comorbidity, or disability” (Gobbens et al., 2010).

The wide use of an instrument for a specific type of assessment alone is not sufficient to recommend it as the most appropriate. In addition to the patterns of use of frailty instruments explored in our study, consideration for instrument selection should be given to: (1) intended purpose; (2) theoretical basis and validity of the constructs included in the instrument for the intended purpose; and (3) feasibility, given the intended purpose and context. Theoretical views of frailty provide frameworks for operationalizing frailty assessment (Bortz, 2002; Buchner and Wagner, 1992; Ferrucci et al., 2005; Fried et al., 2001; Rockwood and Mitnitski, 2007; Varadhan et al., 2008). For example, the ‘cycle of frailty’ provided a conceptual theory for proposing the clinical manifestation of frailty using five criteria (Fried et al., 2001). As science advances and theories of frailty are refined, the choice of measures should be representative of the underlying constructs, and selected instruments should be validated. Predictive validity is important, and it has been shown among most of the widely-used instruments (de Vries et al., 2011; Bouillon et al., 2013). Further evidence of discriminant validity, construct validity and reliability are needed (Xue and Varadhan, 2014); we should determine if frailty instruments are measuring the frailty-related constructs they were designed to measure, and doing so consistently.

The instrument selected should also be feasible in the given research or clinical context. Frailty instruments in this review range from an assessment with two performance measures (Gill Frailty Measure), to five self-reported questions (FRAIL Scale), to as many as 92 signs, symptoms and health attributes (DAI). Restraints related to time, funding, resources, and space are often present, especially in busy clinics. Certain instruments may better suit specific clinic or research needs, such as DAI (or a similar type index) for risk stratification among mobility-reduced inpatients, or PFP for intervention studies that aim to improve malleable measures of frailty. It has been suggested that some frailty instruments are less feasible in a primary care setting (Pialoux et al., 2012), but there are examples of measures once argued to be infeasible or difficult to collect that are now routinely employed in primary care (e.g., C-reactive protein as a predictive inflammatory marker for cardiovascular disease) (Pepys and Hirschfield, 2003). Such shifts occur when measures are demonstrated to have clinical value or when technological advances or policy changes facilitate or mandate the implementation.

The main limitation of our study is that although we endeavored to perform a comprehensive review, we may have missed uses of certain instruments and certain high quality studies that were not represented in the random samples of articles chosen. The choice of the random sample selection allowed us to maintain methodological validity while keeping the review feasible. We examined the validity of our random sampling methodology by conducting a validation study, where we comprehensively reviewed one of the nine highly-cited instruments, viz. VES-13. The results presented above reinforce the validity of our random sampling methodology. Another limitation is that this review excludes frailty instruments introduced after December 2013; it also does not represent the uses of existing instruments from 2014 to present. This study focuses on a specific time period; as such temporal bias is inherent, especially in a rapidly evolving field such as frailty research. We

recommend future studies that investigate temporal trends in the uses of frailty instruments. Additionally, it is possible that the use of these instruments in clinics may differ slightly from the uses found in the research literature.

## 5. Conclusion

Our study systematically identified frailty instruments and categorized the purposes and contexts of use for highly-cited frailty instruments. We found that frailty was most commonly assessed to evaluate its predictive value as a risk factor for adverse health outcomes. We found scarcity of frailty assessment for the purpose of clinical decision-making and as an interventional target. The most common context was observational cohort studies that assess frailty among community-dwelling older adults. The Physical Frailty Phenotype was the most widely used instrument for assessing frailty, but wide use alone does not suffice for an unconditional recommendation. The results of this study can inform the development and implementation of appropriate frailty instruments in the future, where the instrument for assessing frailty may need to be optimally matched to the purpose and the context of its use. When selecting an appropriate instrument to measure frailty, the intended purpose is a key consideration: if risk prediction is the goal, then whichever instrument has been shown to be most predictive for the outcome of interest may be appropriate. However, if the aim is to investigate the construct of frailty as a medical syndrome or to understand its etiology, then it is recommended that the selected instrument not include measures of disease, disability and comorbidity. Our study provides empirical evidence on the range and frequency of uses of frailty instruments. We recommend careful consideration in the selection of a frailty instrument based on the intended purpose, domains captured, and how the instrument has been used in the past. Continued efforts are needed to study the validity and feasibility of these instruments.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors would like to acknowledge and thank Ms. Carrie Price and Ms. Claire Twose from the Welch Medical Library at Johns Hopkins University for their invaluable assistance with the Web of Science citation searches.

### Role of funding sources

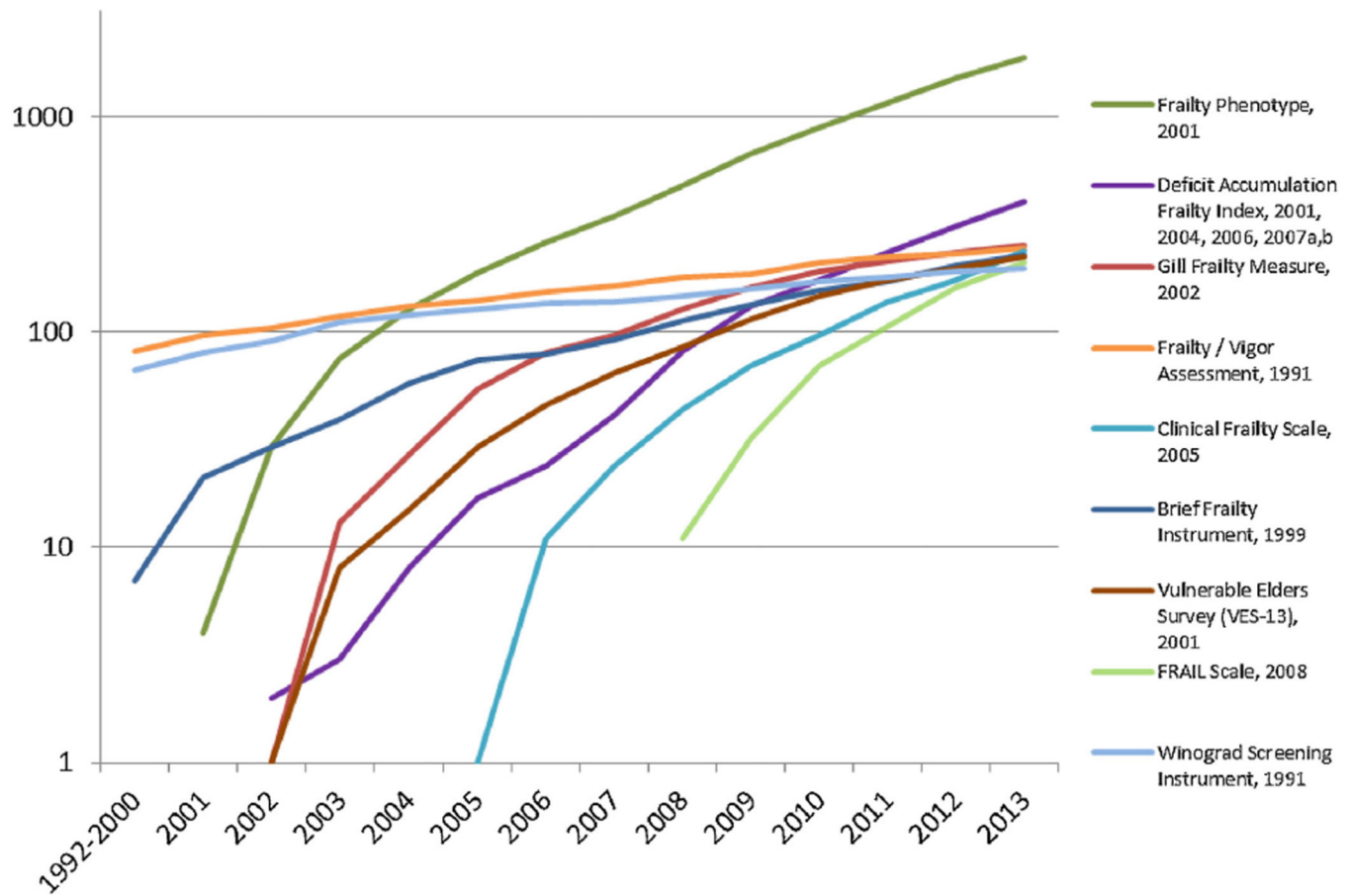
This work was supported by the Johns Hopkins University Claude D. Pepper Older Americans Independence Center of the U.S. National Institute on Aging (NIA) under award number P30AG021334, and by the Johns Hopkins Epidemiology and Bio-statistics of Aging Training Program of the U.S National Institute on Aging (NIA) under award number T32AG000247. NIA center grant P30AG021334 provided support for Mr. Buta, Dr. Walston, Dr. Xue, Dr. Bandeen-Roche, and Dr. Varadhan. Dr. Godino was a postdoctoral fellow supported by NIA training grant T32AG000247 during his involvement on this project and manuscript. Dr. Kalyani was supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) career development award K23DK093583. The study sponsors had no role in the study's design, collection, analysis, interpretation of data, or in the writing of the report or in the decision to submit the paper for publication.

## References

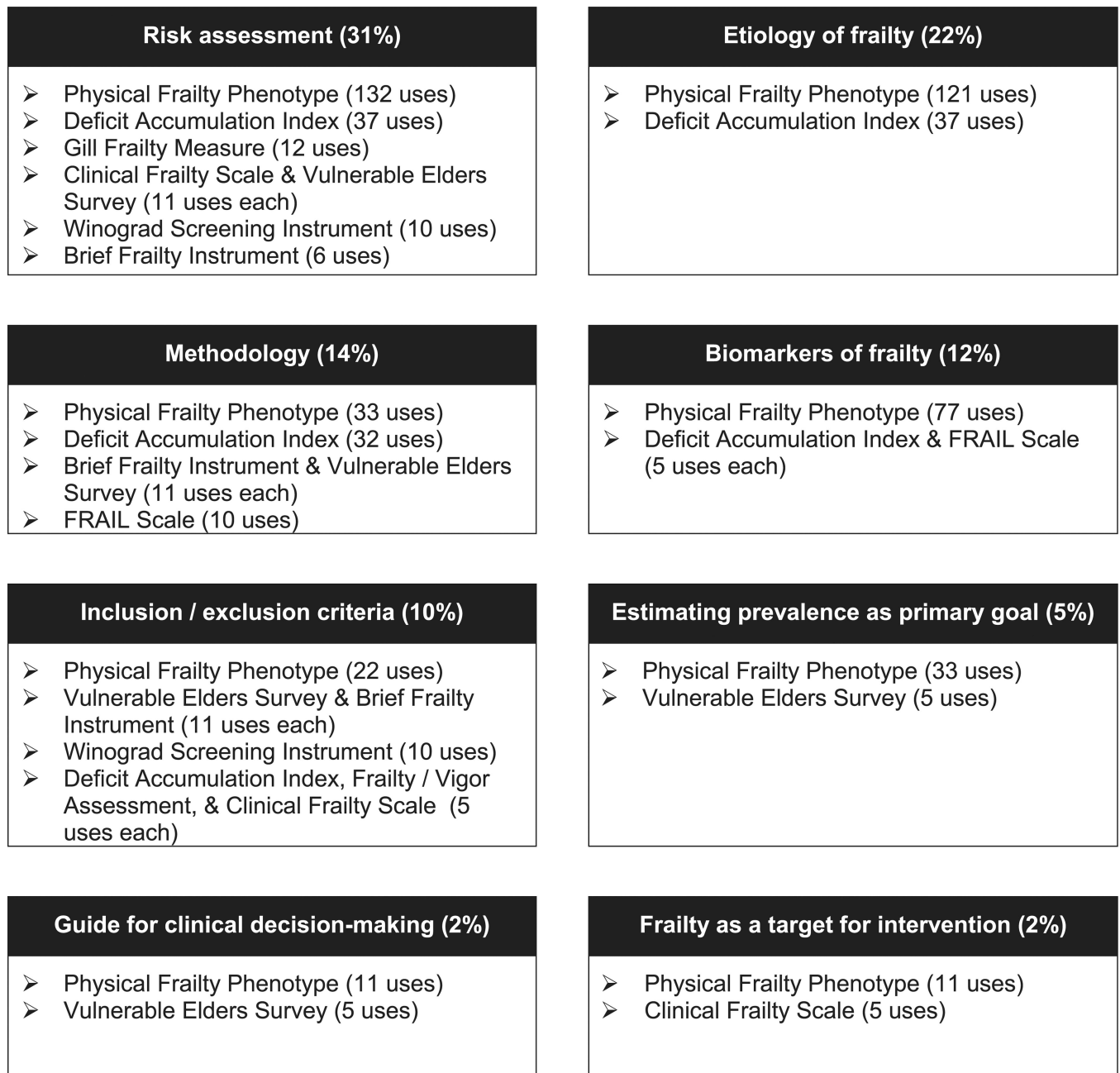
- Abellan van Kan G, Rolland Y, Bergman H, Morley JE, Kritchevsky SB, Vellas B. On behalf of the geriatric advisory panel. The I. A. N. A Task Force on frailty assessment of older people in clinical practice. *J. Nutr. Health Aging*. 2008a; 12:29–37. [PubMed: 18165842]
- Abellan van Kan G, Rolland YM, Morley JE, Vellas B. Frailty: toward a clinical definition. *J. Am. Med. Dir. Assoc.* 2008b; 9:71–72. [PubMed: 18261696]
- Annalingam A, Damayanthi H, Jayawardena R, Ranasinghe P. Determinants of the citation rate of medical research publications from a developing country. *Springerplus*. 2014; 14:140. [PubMed: 25674441]
- Bandeem-Roche K, Xue QL, Ferrucci L, Walston J, Guralnik JM, Chaves P, Zeger SL, Fried LP. Phenotype of frailty: characterization in the women's health and aging studies. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2006; 61:262–266. [PubMed: 16567375]
- Bortz WM. A conceptual framework of frailty: a review. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2002; 57:M283–M288. [PubMed: 11983721]
- Bouillon K, Kivimaki M, Hamer M, Sabia S, Fransson EI, Singh-Manoux A, Gale CR, Batty GD. Measures of frailty in population-based studies: an overview. *BMC Geriatr.* 2013; 13:64. [PubMed: 23786540]
- Buchner DM, Wagner EH. Preventing frail health. *Clin. Geriatr. Med.* 1992; 8:1–17. [PubMed: 1576567]
- Burgess K, Singh PJ, Koroglu R. International J operations & production management, suppl. *Supply Chain Manag. Theory Practice*. 2006; 26:703–729.
- Cesari M, Gambassi G, Abellan van Kan G, Vellas B. The frailty phenotype and the frailty index: different instruments for different purposes. *Age Ageing*. 2014a; 43:10–12. [PubMed: 24132852]
- Cesari M, Vellas B, Hsu FC, Newman AB, Doss H, King AC, Manini TM, Church T, Gill TM, Miller ME, Pahor M. A physical activity intervention to treat the frailty syndrome in older persons—results from the LIFE-P study. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2014b; 70:216–222. [PubMed: 25387728]
- de Vries NM, Staal JB, van Ravensberg CD, Hobbelen JSM, Olde Rikkert MGM, Nijhuis-van der Sandenn MWG. Outcome instruments to measure frailty: a systematic review. *Ageing Res. Rev.* 2011; 10:104–114. [PubMed: 20850567]
- Fairhall N, Kurrle SE, Sherrington C, Lord SR, Lockwood K, John B, Monaghan N, Howard K, Cameron ID. Effectiveness of a multifactorial intervention on preventing development of frailty in pre-frail older people: study protocol for a randomised controlled trial. *BMJ Open*. 2015; 5:e007091.
- Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, web of science, and Google Scholar: strengths and weaknesses. *FASEB J.* 2008; 22:338–342. [PubMed: 17884971]
- Ferrucci L, Windham G, Fried LP. Frailty in older persons. *Genus*. 2005; 61:39–53.
- Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA. For the cardiovascular health study collaborative research group. frailty in older adults: evidence for a phenotype. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2001; 56:M146–M157. [PubMed: 11253156]
- Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2004; 59:255–263. [PubMed: 15031310]
- Gill TM, Baker DI, Gottschalk M, Peduzzi PN, Allore H, Byers A. A program to prevent functional decline in physically frail, elderly persons who live at home. *N. Engl. J. Med.* 2002; 347:1068–1074. [PubMed: 12362007]
- Gobbens RJ, Luijckx KG, Wijnen-Sponselee MT, Schols JM. Toward a conceptual definition of frail community dwelling older people. *Nurs. Outlook*. 2010; 58:76–86. [PubMed: 20362776]
- Hogan DB, MacKnight C, Bergman H. Steering Committee Canadian Initiative on Frailty and Aging. Models, definitions and criteria of frailty. *Ageing Clin. Exp. Res.* 2003; 15:1–29. [PubMed: 14580013]

- Lee PH, Lee YS, Chan DC. Interventions targeting geriatric frailty: a systemic review. *J. Clin. Gerontol. Geriatr.* 2012; 3:47–52.
- Linder SK, Kamath GR, Pratt GF, Saraykar SS, Volk RJ. Citation searches are more sensitive than keyword searches to identify studies using specific measurement instruments. *J. Clin. Epidemiol.* 2015; 68:412–417. [PubMed: 25554521]
- Martin FC, Brighton P. Frailty: different tools for different purposes? *Age Ageing.* 2008; 37:129–131. [PubMed: 18349007]
- Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. *Sci. World J.* 2001; 1:323–336.
- Mitnitski AB, Song X, Rockwood K. The estimation of relative fitness and frailty in community-dwelling older adults using self-report data. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2004; 59:M627–M632. [PubMed: 15215283]
- Morley JE, Vellas B, Abellan van Kan G, Anker SD, Bauer JM, Bernabei R, Cesari M, Chumlea WC, Doehner W, Evans J, Fried LP, Guralnik JM, Katz PR, Malmstrom TK, McCarter RJ, Gutierrez Robledo LM, Rockwood K, von Haehling S, Vandewoude MF, Walston J. Frailty consensus: a call to action. *J. Am. Med. Dir. Assoc.* 2013; 14:392–397. [PubMed: 23764209]
- Pepys MB, Hirschfield GM. C-reactive protein: a critical update. *J. Clin. Invest.* 2003; 111:1805–1812. [PubMed: 12813013]
- Pialoux T, Goyard J, Lesourd B. Screening tools for frailty in primary health care: a systematic review. *Geriatr. Gerontol. Int.* 2012; 12:189–197. [PubMed: 22233158]
- R Core Team. R Foundation for Statistical Computing. Vienna, Austria: 2014. R: A language and environment for statistical computing. URL: <http://www.R-project.org/>
- Rockwood K, Andrew M, Mitnitski A. A comparison of two approaches to measuring frailty in elderly people. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2007; 62:738–743. [PubMed: 17634321]
- Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2007; 62:722–727. [PubMed: 17634318]
- Rockwood K, Mitnitski A, Song X, Steen B, Skoog I. Long-term risks of death and institutionalization of elderly people in relation to de.cit accumulation at age 70. *J. Am. Geriatr. Soc.* 2006; 54:975–979. [PubMed: 16776795]
- Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, Mitnitski A. A global clinical measure of fitness and frailty in elderly people. *CMAJ.* 2005; 173:489–495. [PubMed: 16129869]
- Rockwood K, Stadnyk K, MacKnight C, McDowell I, Hébert R, Hogan DB. A brief clinical instrument to classify frailty in elderly people. *Lancet.* 1999; 353:205–206. [PubMed: 9923878]
- Rodríguez-Mañas L, Fried LP. Frailty in the clinical scenario. *Lancet.* 2015; 14:e7–e9. [PubMed: 25468154]
- Rodríguez-Mañas L, Féart C, Mann G, Viña J, Chatterji S, Chodzko-Zajko W, Gonzalez-Colaço Harmand M, Bergman H, Carcaillon L, Nicholson C, Scuteri A, Sinclair A, Pelaez M, Van der Cammen T, Beland F, Bickenbach J, Delamarche P, Ferrucci L, Fried LP, Gutiérrez-Robledo LM, Rockwood K, Rodríguez Artalejo F, Serviddio G, Vega E. FOD-CC group Searching for an operational definition of frailty: a Delphi method based consensus statement: the frailty operative definition-consensus conference project. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2013; 68:62–67. [PubMed: 22511289]
- Saliba D, Elliott M, Rubenstein LZ, Solomon DH, Young RT, Kamberg CJ, Roth C, MacLean CH, Shekelle PG, Sloss EM, Wenger NS. The Vulnerable Elders Survey: a tool for identifying vulnerable older people in the community. *J. Am. Geriatr. Soc.* 2001; 49:1691–1699. [PubMed: 11844005]
- Sourial N, Bergman H, Karunanathan S, Wolfson C, Payette H, Gutierrez-Robledo LM, Béland F, Fletcher JD, Guralnik J. Implementing frailty into clinical practice: a cautionary tale. *J. Gerontol. A. Biol. Sci. Med. Sci.* 2013; 68:1505–1511. [PubMed: 23640761]
- Speechley M, Tinetti M. Falls and injuries in frail and vigorous community elderly persons. *J. Am. Geriatr. Soc.* 1991; 39:46–52. [PubMed: 1987256]

- Sternberg SA, Wershof Schwartz A, Karunanathan S, Bergman H, Clarfield AM. The identification of frailty: a systematic literature review. *J. Am. Geriatr. Soc.* 2011; 59:2129–2138. [PubMed: 22091630]
- Turner G, Clegg A. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. *Age Ageing.* 2014; 43:744–747. [PubMed: 25336440]
- Varadhan R, Seplaki CL, Xue QL, Bandeen-Roche K, Fried LP. Stimulus-response paradigm for characterizing the loss of resilience in homeostatic regulation associated with frailty. *Mech. Ageing Dev.* 2008; 129:666–670. [PubMed: 18938195]
- Vellas B, Balardy L, Gillette-Guyonnet S, Abellan Van Kan G, Ghisolfi-Marque A, Subra J, Bismuth S, Oustric S, Cesari M. Looking for frailty in community-dwelling older persons: the Gérontopôle Frailty Screening Tool (GFST). *J. Nutr. Health Aging.* 2013; 17:629–631. [PubMed: 23933875]
- Winograd CH, Gerety MB, Chung M, Goldstein MK, Dominguez FJr, Vallone R. Screening for frailty: criteria and predictors of outcomes. *J. Am. Geriatr. Soc.* 1991; 39:778–784. [PubMed: 1906492]
- Xue QL, Varadhan R. What is missing in the validation of frailty instruments? *J. Am. Med. Dir. Assoc.* 2014; 15:141–142. [PubMed: 24405640]



**Fig. 1.** Cumulative number of articles that cite the Most-Referenced Frailty Instruments. This figure displays the cumulative number of citations (on the Y-Axis) per year (on the X-Axis) for nine highly-cited frailty instruments. The Y-axis is on a logarithmic scale. The Physical Frailty Phenotype is the most cited frailty instrument (1891); Deficit Accumulation Index is the second most cited (401).



**Fig. 2.** Frailty Instruments and Categories of Use, Ordered by Number of Estimated Uses per Category. This figure displays the percentage of use per category, and the number of estimated uses of the highly-cited frailty instruments per each category.

Table 1

Summary of the reviewed random samples of frailty literature and estimated total uses of frailty instruments in the literature.

Frailty Instrument, Publication Year (# of citations through Dec 2013)	# of articles reviewed from random sample <sup>a</sup>	# of original research articles (% of sample)	# of review articles (% of sample)	# of commentary (% of sample)	# of other articles (% of sample)	% of use of frailty instrument in the random sample (uses/reviewed articles)	Estimated total number of uses of the frailty instrument in the literature [95% CI] <sup>b</sup>
Physical Frailty Phenotype, 2001 (1891)	172	107 (62.2%)	45 (26.2%)	14 (8.1%)	6 (3.5%)	23.3% (40/172 = .233)	440 [324,573]
Deficit Accumulation Index, 2001 (401)	76	54 (71.1%)	14 (18.4%)	8 (10.5%)	0	29.9% (22/76 = .299)	116 [77,162]
Gill Frailty Measure, 2002 (254)	44	30 (68.2%)	9 (20.5%)	1 (2.3%)	4 (9%)	4.6% (2/44 = .046)	12 [1,39]
Frailty/Vigor Assessment, 1991 (246)	47	37 (78.7%)	9 (19.1%)	1 (2.1%)	0	2.1% (1/47 = .021)	5 [1,28]
Clinical Frailty Scale, 2005 (239)	45	28 (62.2%)	11 (24.4%)	3 (6.7%)	3 (6.7%)	8.9% (4/45 = 0.89)	21 [6,51]
Brief Frailty Instrument, 1999 (226)	41	26 (63.4%)	14 (34.2%)	1 (2.4%)	0	12.2% (5/41 = .122)	28 [9,59]
Vulnerable Elders Survey, 2001 (225)	41	21 (51.2%)	17 (41.5%)	1 (2.4%)	2 (4.9%)	22.0% (9/41 = .220)	49 [24,85]
FRAIL Scale, 2008 (211)	41	17 (41.5%)	12 (29.3%)	11 (26.8%)	1 (2.4%)	7.3% (3/41 = .073)	15 [3,42]
Winograd Screening Instrument, 1991 (198)	38	26 (68.4%)	10 (26.3%)	2 (5.3%)	0	10.5% (4/38 = .105)	21 [6,49]

<sup>a</sup>We sampled 20% of all citing articles per instrument, except for the Physical Frailty Phenotype where 10% of the citing articles were sampled. We reviewed all available articles from each random sample, though some were not available or in a non-English language.

<sup>b</sup>Estimated number of individual articles that use the instrument in research studies through December 2013. Estimates rounded to the nearest positive integer.



**Table 2**

Domains included in highly-cited frailty instruments.

Highly-cited frailty instrument	Physical function? (includes disability?)	Physical activity?	Cognition?	Comorbidity?	Weight loss?	Other (social, sensory, demographic, etc)?
Physical Frailty Phenotype	Yes (No)	Yes	No	No	Yes	No
Deficit Accumulation Index	Yes (Yes)	No	Yes	Yes	No	Yes
Gill Frailty Measure	Yes (No)	No	No	No	No	No
Frailty/Vigor Assessment	Yes (Yes)	Yes	Yes	No	No	Yes
Clinical Frailty Scale	Yes (Yes)	Yes	No	Yes	No	Yes
Brief Frailty Instrument	Yes (Yes)	No	Yes	Yes	No	No
Vulnerable Elders Survey	Yes (Yes)	No	No	No	No	Yes
FRAIL Scale	Yes (No)	No	No	Yes	Yes	No
Winograd Screening Instrument	Yes (Yes)	No	Yes	Yes	No	Yes
Total out of nine instruments	9 (6)	3	4	5	2	5

**Table 3**

Estimated primary uses of frailty instruments per category in the literature<sup>a</sup>.

Frailty instrument (# citing articles)	Risk factor for adverse health outcomes	Risk factors for frailty	Methodology	Biomarkers of frailty	Inclusion or exclusion criteria for research	Estimation of frailty prevalence as primary goal	Guide for clinical decision-making	Frailty as target for intervention
Physical Frailty Phenotype (1891)	132 [69,224]	121 [61,211]	33 [7,95]	77 [31,155]	22 [3,78]	33 [7,95]	11 [1,61]	11 [1,61]
Deficit Accumulation Index (401)	37 [15,72]	37 [15,72]	32 [7,66]	5 [1,29]	5 [1,29]	0 [0,19]	0 [0,19]	0 [0,19]
Gill Frailty Measure (254)	12 [1,39]	0 [0,20]	0 [0,20]	0 [0,20]	0 [0,20]	0 [0,20]	0 [0,20]	0 [0,20]
Frailty Vigor Assessment (246)	0 [0,19]	0 [0,19]	0 [0,19]	0 [0,19]	5 [1,28]	0 [0,19]	0 [0,19]	0 [0,19]
Clinical Frailty Scale (239)	11 [1,36]	0 [0,19]	0 [0,19]	0 [0,19]	5 [1,28]	0 [0,19]	0 [0,19]	5 [1,28]
Brief Frailty Instrument (226)	6 [1,29]	0 [0,19]	11 [1,37]	0 [0,19]	11 [1,37]	0 [0,19]	0 [0,19]	0 [0,19]
Vulnerable Elders Survey <sup>b</sup> (225)	11 [1,37]	0 [0,19]	11 [1,37]	0 [0,19]	11 [1,37]	5 [1,28]	5 [1,28]	0 [0,19]
FRAIL Scale (211)	0 [0,18]	0 [0,18]	10 [1,35]	5 [1,27]	0 [0,18]	0 [0,18]	0 [0,18]	0 [0,18]
Winograd Screening Instrument (198)	10 [1,35]	0 [0,18]	0 [0,18]	0 [0,18]	10 [1,35]	0 [0,18]	0 [0,18]	0 [0,18]
Sum (%) of estimated uses for all instruments	219 (31.3%)	158 (22.3%)	97 (13.9%)	87 (12.4%)	58 (9.9%)	38 (5.4%)	16 (2.3%)	16 (2.3%)

<sup>a</sup> Estimated number (and confidence interval) of individual articles that use the highly-cited frailty instrument in research studies per category through December 2013. Estimates rounded to the nearest positive integer.

<sup>b</sup> One article (Ganz DA, Wenger NS, Roth CP, et al. The Effect of a Quality Improvement Initiative on the Quality of Other Aspects of Health Care: The Law of Unintended Consequences? Medical Care 2007; 45(1): 8–18) from the VES citing articles was categorized as “frailty as a descriptive measure” and is not included in the represented categories or sum.