BMC Medicine

RESEARCH ARTICLE

Open Access

Comparing the prognostic value of geriatric health indicators: a population-based study

Alberto Zucchelli[1,2*†] , Davide L. Vetrano[1,3†], Giulia Grande[1], Amaia Calderón-Larrañaga[1], Laura Fratiglioni[1,4], Alessandra Marengoni[1,2†] and Debora Rizzuto[1†]

Abstract

Background: The identification of individuals at increased risk of poor health-related outcomes is a priority. Geriatric research has proposed several indicators shown to be associated with these outcomes, but a head-to-head comparison of their predictive accuracy is still lacking. We therefore aimed to compare the accuracy of five geriatric health indicators in predicting different outcomes among older persons: frailty index (FI), frailty phenotype (FP), walking speed (WS), multimorbidity, and a summary score including clinical diagnoses, functioning, and disability (the Health Assessment Tool; HAT).

Methods: Data were retrieved from the Swedish National Study on Aging and Care in Kungsholmen, an ongoing longitudinal study including 3363 people aged 60+. To inspect the accuracy of geriatric health indicators, we employed areas under the receiver operating characteristic curve (AUC) for the prediction of 3-year and 5-year mortality, 1-year and 3-year unplanned hospitalizations (1+), and contacts with healthcare providers in the 6 months before and after baseline evaluation (2+).

Results: FI, WS, and HAT showed the best accuracy in the prediction of mortality [AUC(95%CI) for 3-year mortality 0.84 (0.82–0.86), 0.85 (0.83–0.87), 0.87 (0.85–0.88) and AUC(95%CI) for 5-year mortality 0.84 (0.82–0.86), 0.85 (0.83–0.86), 0.86 (0.85–0.88), respectively]. Unplanned hospitalizations were better predicted by the FI [AUC(95%CI) 1-year 0.73 (0.71–0.76); 3-year 0.72 (0.70–0.73)] and HAT [AUC(95%CI) 1-year 0.73 (0.71–0.75); 3-year 0.71 (0.69–0.73)]. The most accurate predictor of multiple contacts with healthcare providers was multimorbidity [AUC(95%CI) 0.67 (0.65–0.68)]. Predictions were generally less accurate among younger individuals (< 78 years old).

Conclusion: Specific geriatric health indicators predict clinical outcomes with different accuracy. Comprehensive indicators (HAT, FI, WS) perform better in predicting mortality and hospitalization. Multimorbidity exhibits the best accuracy in the prediction of multiple contacts with providers.

Keywords: Health indicators, Older persons, Prognosis

Background

The identification of individuals at increased risk of poor health-related outcomes is a clinical and public health priority. Indeed, risk stratification plays a pivotal role in medical decision-making, public resource allocation, and research [1, 2]. For example, unplanned hospitalizations, which are a major driver of healthcare costs, often lead to disability onset or progression [3, 4] and delirium [5, 6], preventing older adults from being discharged home. The identification of older persons at increased risk of unplanned hospital admissions could help to better target preventive strategies [7] (i.e. therapeutic review) toward specific groups of patients.

Accomplishing such a task is particularly critical among older persons. In fact, persons older than 60 are among the most strenuous users of healthcare resources [8, 9], and their number is expected to double worldwide by 2050 [10]. Indeed, a noteworthy variability is found among older persons, even of the same

* Correspondence: a.zucchelli001@unibs.it; alberto.zucchelli.1@ki.se
Alberto Zucchelli and Davide L Vetrano are first shared authorship.
Alessandra Marengoni and Debora Rizzuto are last shared authorship.
[1]Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, 11330 Stockholm, Sweden
[2]Department of Clinical and Experimental Sciences, University of Brescia, Viale Europa 11, 25121 Brescia, Italy
Full list of author information is available at the end of the article

age, in terms of functional and cognitive performance, number and severity of chronic diseases, quality of life, and prognosis [11, 12].

In the last decades, researchers in geriatrics have proposed several indicators shown to be strongly associated with the development of poor health-associated outcomes, such as death and unplanned hospitalizations. The co-occurrence of multiple chronic conditions in the same individual (multimorbidity), for example, has a strong impact on health, higher than that expected by simply summing diseases [13]. Frailty, a state of increased vulnerability to stressors due to poor resolution of homeostasis [14], is another concept that gained recent recognition because of its prognostic value, even beyond the borders of geriatric practice [15, 16]. Furthermore, simple functional measures, such as the evaluation of normal pace walking speed, have been shown to be strongly associated with survival [17]. Lastly, summary scores evaluating multiple domains have been shown to have high predictive accuracy [18, 19].

These indicators differ not only in their theoretical foundation, but also in their operationalization. For example, while a general consensus on the definition of frailty has been reached [20], several ways to assess it in clinical practice and research are in use [14]. Furthermore, while these indicators have been validated in various cohorts [21–23], a head-to-head comparison of their accuracy in the prediction of different outcomes is still lacking. Such studies are of particular interest, as they may allow clinicians (as well as researchers and policy makers) to choose the most suitable predictive tool according to aims, needs, and data availability.

Thus, the aim of this study is to compare the accuracy of five geriatric health indicators (the frailty index, the frailty phenotype, multimorbidity, walking speed, and a summary score—the Health Assessment Tool) in the prediction of mortality, unplanned hospitalizations, and multiple contacts with healthcare providers.

Methods
Study population
Data were gathered from the Swedish National Study on Aging and Care in Kungsholmen (SNAC-K). SNAC-K is an ongoing population-based study, started in 2001. Individuals aged 60+ living in the central area of Stockholm (Sweden), either at home or in institutions, were asked to participate in the study. A comprehensive assessment using standard questionnaires, medical examinations, and interviews was performed to retrieve demographic, clinical, and functional measures of the 3363 (response rate 73.3%) persons enrolled. Data from neuropsychological assessments and physical tests were also collected, as elsewhere described [24]. Every wave of the study was approved by the Regional Ethical Review

Board in Stockholm, Sweden. Written informed consent was obtained from each participant, or from a proxy, in case of cognitive impairment. The public or patients were not involved during the development of this study: anyhow, we plan to disseminate the findings of this research to participants of SNAC-K and to the public.

Geriatric health indicators
Frailty index (FI)
The frailty index is a commonly employed measure of frailty, firstly proposed by Rockwood et al. [25]. It is based on the ratio (range 0–1) between the number of deficits (i.e. signs, symptoms, diseases, biomarkers, functional status, physical performance indicators) exhibited by the individual and the total number of potential deficits taken into consideration by researchers. In SNAC-K, two geriatricians (DLV and AZ) selected 45 variables (Additional file 1: Table S1) and re-codified them, in accordance with the recommendations provided by Searle et al. [26]. For baseline description purposes, participants were considered frail if they exhibit a FI ≥ 0.25, robust with a FI ≤ 0.08, and pre-frail in between, as previously reported [27]. The frailty index was considered missing if two or more variables were not available ($N = 348$).

Frailty phenotype (FP)
The frailty phenotype is a commonly used and validated operational definition of physical frailty, originally proposed by Fried et al. [28]. It evaluates five criteria: slow walking speed, low grip strength, unintentional weight loss, exhaustion, and low physical activity (the operationalization carried out in SNAC-K is available elsewhere [29]). For baseline description purposes, individuals meeting at least three criteria were considered frail, and those meeting one or two criteria were considered pre-frail, while the remaining were considered robust. Values were missing for 599 people in at least one criterion.

Multimorbidity
In SNAC-K, diseases were coded in accordance with the International Classification of Diseases 10th edition. Diagnoses were ascertained by physicians based on medical history, medical records, physical examinations, and instrumental and laboratory analyses. For baseline description purposes, we defined multimorbidity as the count of chronic conditions, based on 60 disease categories identified by Calderon-Larranaga et al. [30]. To examine the distribution of multimorbidity in our population, we used the cut-off of two or more chronic diseases.

Walking speed (WS)
In SNAC-K, a nurse noted the time needed for the participant to complete a 6-m straight path, walking at

usual pace. Participants were allowed to use walking aids but had to complete the path without help. In case of inability to complete the path, a walking speed of zero was recorded. For those who self-reported slow walking speed or in case of at-home assessment, a 2.4-m path was used. For baseline description purposes, a WS cut-off of < 0.8 m/s was used to identify slow walking speed in our study population, as previously suggested [17].

Health Assessment Tool (HAT)

Proposed by our group [18], HAT is a summary score evaluating five characteristics: walking speed, Mini-Mental State Examination (MMSE) score, limitations in instrumental activities of daily living, limitations in basic activities of daily living, and count of chronic diseases. HAT was built regressing these characteristics against the latent variable "health status" using a nominal response model (more details are available in the appendix of the original article [18]), obtaining a score ranging from 0 (poor health) to 10 (good health). It has been shown to be reliable over time and to adequately predict different adverse outcomes [18, 31]. For baseline description purposes, poor health status was considered for individuals with a HAT score ≤ 3.3, while good health was considered for those with a HAT score ≥ 6.6. Data were missing for eight people.

Outcomes

Vital status was retrieved within 3 and 5 years of follow-up using the Swedish Cause of Death Register [32]. The Stockholm County Council Register (as part of the National Patient Register [33, 34]) was used to gather data on hospitalizations and contacts with outpatient care providers (i.e. visits to both primary and specialist care), as previously described [18]. These registers contain information on the type of admission (i.e. planned or unplanned), among others. We defined "acute hospitalization" as experiencing at least one unplanned admission during the first year or the first 3 years after the baseline assessment. "Multiple provider contacts" was defined as having multiple outpatient visits in the 6 months prior and after the baseline assessment. We used the median number of planned outpatient visits (i.e. 2) as the cut-off.

Other measures

Education level was measured as the highest degree obtained. Cognitive status was assessed using the MMSE score (both as a continuous variable and using a cut-off of 24 [35]). Disability was defined as being impaired in at least one out of six basic activities of daily living [36].

Statistical analyses

To assess the accuracy of the different geriatric health indicators, we used the area under the receiver operating characteristic curve (AUC). In this paper, we employed the AUC as measure of predictive accuracy, since it allows to simultaneously consider the sensitivity and specificity of a continuous variable in the prediction of an outcome. The AUC was obtained using non-parametric ROC analysis [37], including the different indicators as continuous variables. The analyses were repeated stratifying by age, using a cut-off of 78 years, the median age of our study population. To compare the average scores of the different indicators across individuals of the same age, the raw scores were standardized into z-scores, using the baseline mean and standard deviation of the population. The analyses were conducted on 10 imputed datasets performing multiple imputation by chained equations. For those people for whom data on the health indicators were missing (28.4%), we created an indicator variable. This variable was equal to 1 if a given observation was missing in any health indicators and to 0 otherwise. We performed logistic regression with missing value as the outcome to test whether any of the other variables were associated with the probability to be missing (Additional file 2: Table S2). These variables were used in the imputation process. For the main analyses, pooled estimates were calculated according to Rubin's rule [38]. The same analyses were conducted in the complete case sample (71.6%), showing consistent results in terms of direction and magnitude (Additional file 3: Table S3). All analyses were performed using Stata 15 (Stata Corp, Texas, USA), with an alpha level of .05.

Results

The baseline characteristics of the study population are shown in Table 1: the mean age was 74.7 (standard deviation, SD 11.2) and 2182 (65%) participants were female. Older (i.e. ≥ 78 years, $N = 1581$) individuals were more likely to be female, less educated, and affected by disability, while younger participants were more likely to have better cognitive performance (all $p < 0.001$).

The scores for all indicators were worse among older individuals, as shown in Table 1 and Fig. 1a, with the exception of the count of chronic conditions, which exhibited a plateau and a subsequent slight decline after the age of 90 years. The proportion of individuals characterized by poor health according to HAT (≤ 3.3) and of those frail according to the FI (≥ 0.25) was similar across all ages (Fig. 1b). The proportion of persons with slow WS (< 0.8 m/s) and of those considered frail according to the FP steeply increased after the age of 80 years.

The mean follow-up time in our study was 4.41 years. Figure 2 (and Additional file 4: Table S4 and Additional

Table 1 Baseline characteristics of the study population, stratified by age

	< 78 years old N = 1782 (53.0%)	≥ 78 years old N = 1581 (47.0%)	p value	Total N = 3363
Age, mean (SD)	65.5 (4.8)	85.1 (6.1)	< 0.001	74.7 (11.2)
Sex (female), n (%)	1024 (57.4)	1158 (73.2)	< 0.001	2182 (64.9)
Education*				
Elementary, n (%)	154 (8.6)	451 (28.5)	< 0.001	590 (17.7)
High school, n (%)	817 (45.9)	848 (53.7)		1651 (49.6)
University, n (%)	811 (45.5)	282 (17.8)		1090 (32.7)
MMSE, mean (SD)	29.0 (2.1)	25.2 (7.4)	< 0.001	27.6 (4.9)
MMSE < 24, N (%)	25 (1.4)	348 (22.0)	< 0.001	462 (14.7)
≥ 1 impaired ADL, n (%)	26 (1.5)	303 (19.2)	< 0.001	283 (8.6)
Frailty index				
≤ 0.08, n (%)	1390 (78.0)	436 (27.6)	< 0.001	1826 (54.3)
0.08–0.25, n (%)	375 (21.0)	856 (54.1)		1231 (36.6)
≥ 0.25, n (%)	17 (1.0)	289 (18.3)		306 (9.1)
Median (IQR)	0.04 (0.05)	0.12 (0.14)	< 0.001	0.07 (0.09)
Frailty phenotype				
Robust, n (%)	972 (54.6)	319 (20.2)	< 0.001	1291 (38.4)
Pre-frail, n (%)	738 (41.4)	681 (43.0)		1419 (42.2)
Frail, n (%)	72 (4.0)	581 (36.8)		653 (19.4)
Health assessment tool				
≥ 6.6, n (%)	1672 (93.8)	792 (50.1)		2464 (73.3)
3.3–6.6, n (%)	94 (5.3)	506 (32.0)	< 0.001	600 (17.8)
≤ 3.3, n (%)	16 (0.9)	283 (17.9)		299 (8.9)
Median (IQR)	9 (0.8)	6.5 (4.4)	< 0.001	8.6 (2.6)
Multimorbidity				
0 chronic dis., n (%)	97 (5.4)	9 (0.6)	< 0.001	106 (3.1)
1 chronic dis., n (%)	289 (16.2)	37 (2.3)		326 (9.7)
2+ chronic dis., n (%)	1396 (78.4)	1535 (97.1)		2931 (87.1)
Median (IQR)	3 (2)	5 (4)	< 0.001	4 (3)
Walking speed				
< 1.0 m/s, n (%)	267 (15.0)	1112 (70.3)	< 0.001	1379 (50.0)
< 0.8 m/s, n (%)	141 (7.9)	890 (56.3)	< 0.001	1031 (30.7)
Median (IQR)	1.2 (0.5)	0.6 (0.6)	< 0.001	1 (0.6)

Abbreviations: n number, *SD* standard deviation, *IQR* interquartile range, *MMSE* Mini-Mental State Examination

file 5: Figure S1) depicts the predictive accuracy (AUC: area under the ROC curve) of the different indicators.
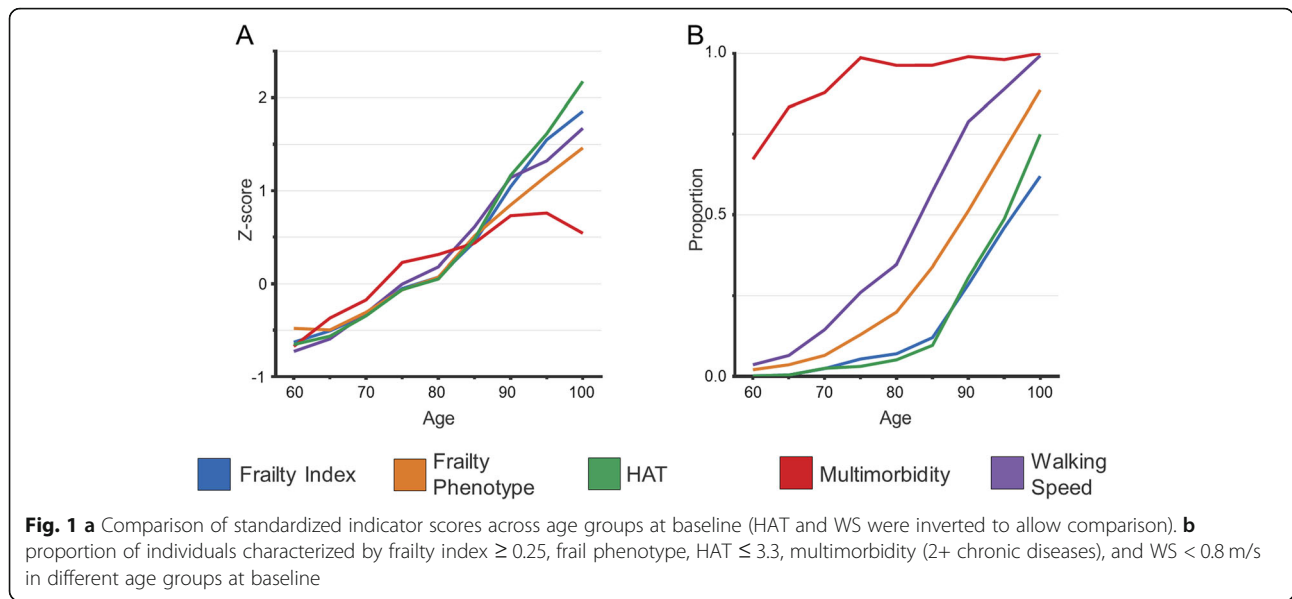
Mortality

In our study population, 477 participants (14.2%) died in the first 3 years of follow-up and another 291 in the subsequent 2 years (5-year mortality 22.8%). All indicators, with the exclusion of MM, predicted mortality with AUCs higher than 0.75: FP was the least performing indicator [3-year mortality AUC (95%CI) 0.80 (0.78–0.82); 5-year mortality AUC (95%CI) 0.79 (0.77–0.80)], while HAT showed the best AUCs [3-year mortality AUC

(95%CI) 0.87 (0.85–0.88); 5-year mortality AUC (95%CI) 0.86 (0.85–0.88)]. Mortality was predicted with similar AUCs by the FI [3-year mortality AUC (95%CI) 0.84 (0.82–0.86); 5-year mortality AUC (95%CI) 0.84 (0.82–0.86)] and WS [3-year mortality AUC (95%CI) 0.85 (0.83–0.87); 5-year mortality AUC (95%CI) 0.85 (0.83–0.86)]. MM showed the worst AUC overall [3-year mortality AUC (95%CI) 0.71 (0.68–0.73)].

Acute hospitalization

The 16.1% (N = 542) of our sample experienced at least one unplanned hospitalization in the first year of follow-

Fig. 1 a Comparison of standardized indicator scores across age groups at baseline (HAT and WS were inverted to allow comparison). **b** proportion of individuals characterized by frailty index ≥ 0.25, frail phenotype, HAT ≤ 3.3, multimorbidity (2+ chronic diseases), and WS < 0.8 m/s in different age groups at baseline

up, while 1134 participants (33.7%) had one or more unplanned hospitalizations in the first 3 years following baseline assessment. Indicators exhibited AUCs ranging from 0.66 (0.64–0.68) [AUC(95%CI) for FP in the prediction of 3-year unplanned hospitalization] to 0.73 (0.71–0.76) [AUC(95%CI) for FI in the prediction of 1-year unplanned hospitalization].

Multiple provider contacts

The number of individuals who had at least two contacts with care providers in the 6 months prior and after the baseline assessment was 1959 (58.2%). Among the outcomes considered, "multiple provider contacts" was predicted with the lowest AUCs. The best AUC (95% CI) was exhibited by MM 0.67 (0.65–0.68).

Age-stratified analyses

AUCs for mortality were lower among younger individuals than among older ones, as shown in Fig. 3 (and Additional file 6: Table S5), although most of the confidence intervals were overlapping. Among younger individuals, HAT, FI, and WS showed a trend of increased accuracy in predicting mortality and unplanned hospitalization. Multimorbidity and FI predicted provider contacts with similar accuracy among younger and older individuals.
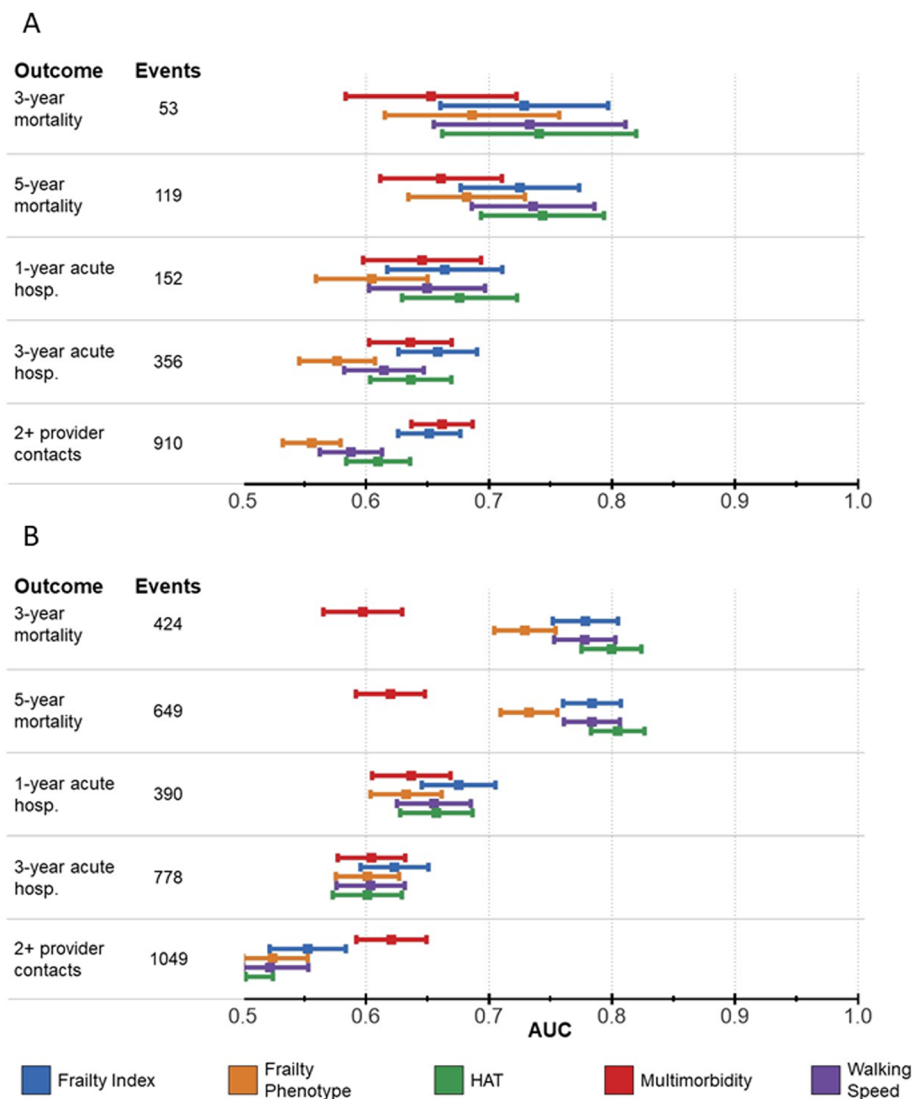

Fig. 2 Comparison between areas under the ROC curve (AUCs) of different indicators in the SNAC-K population (*n* = 3363). HAT: Health Assessment Tool
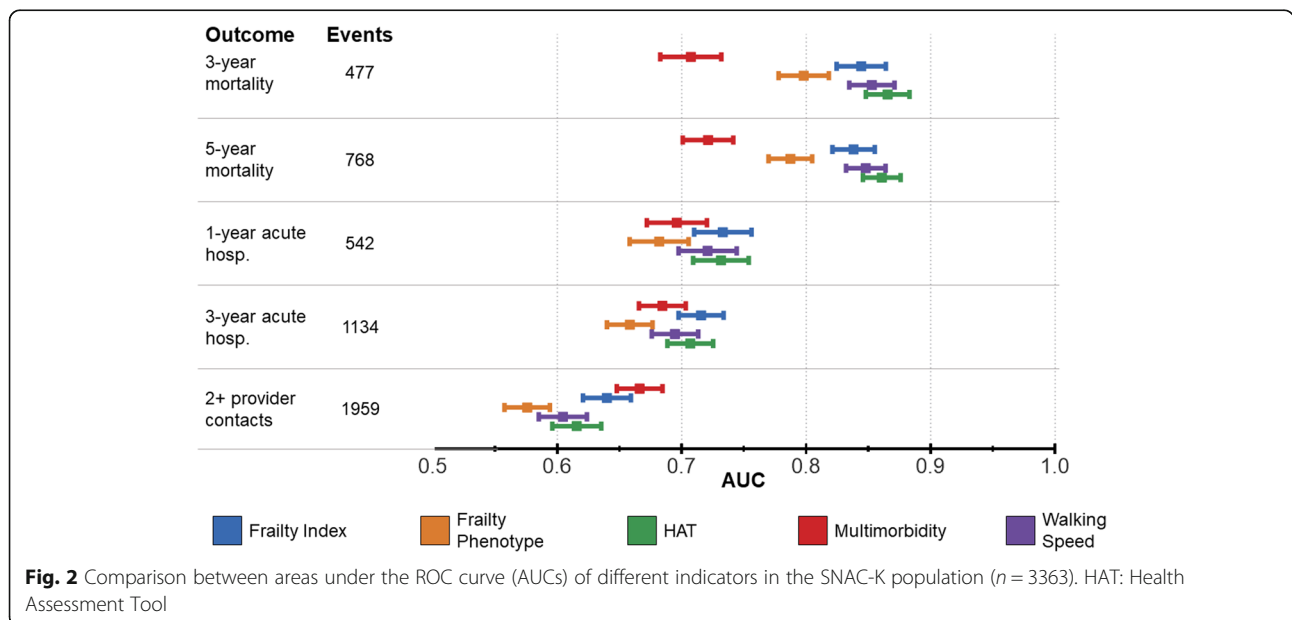
Fig. 3 Comparison between areas under the ROC curve (AUCs) of different indicators in **a** young older adults (< 78 years old) and **b** oldest old (≥ 78 years old). HAT: Health Assessment Tool

Sensitivity analyses conducted on the complete case dataset showed similar results in terms of magnitude and direction. Most indicators exhibited similar AUCs for the prediction of all outcomes, with the exception of FP and FI that showed a slightly lower predictive performance in the complete case analysis, compared to the main analysis (Additional file 3: Table S3).

Discussion

All geriatric health indicators showed an AUC ≥ 0.70 in the prediction of mortality, while they were less accurate in predicting unplanned hospitalization and contact with multiple providers. Besides, important differences were observed in the prediction of one same clinical outcome by the different indicators.

AUCs were lower among younger old persons for all indicators, with the exception of multimorbidity. HAT, WS, and FI were the most accurate predictors of mortality and unplanned hospitalization, while multimorbidity showed the highest AUCs in the prediction of contact with multiple healthcare providers.

Our findings are in line with the literature that reports AUCs ≥ 0.80 for the prediction of mortality using the FI [27, 39, 40]. Previous studies showed a prognostic accuracy for the FP ranging between 0.70 [40, 41] and 0.75 [42], although a significant variability in the assessment of the five phenotypical criteria is present. Ritt et al. [42] reported an AUC of 0.50 in the prediction of unplanned hospitalizations using the FP: the fact that the assessment was conducted in routine clinical practice

conditions and the short follow-up (i.e. 6 months) might explain the difference with our findings. Several different multi-domain scores have been proposed in the previous years: despite the noteworthy variability in the variables included, reported AUCs for the prediction of unplanned hospitalization were generally higher than 0.70 [43].

Our results confirm the ability of physical function to accurately predict poor health outcomes among older individuals [17, 44–47]. Several studies suggest that disability and functional measures are strongly associated with poor health-related outcomes among older adults [17, 48, 49]. The combination of physical function and other domains, such as cognition [50, 51] or the severity of a pre-defined number of chronic conditions [19], has already been shown to help better stratify older individuals with poor prognosis. In our study, comprehensive indicators (FI and HAT) exhibited a minor but significantly higher AUCs for mortality and hospitalization, when compared to a single functional measure (WS). Different studies compared the accuracy in the prediction of mortality of physical functional indicators, such as the FP, and more comprehensive ones, such as the FI, showing different results. Our findings confirm the results of Ritt et al. [39] and Wigadgo et al. [52], who found that FP exhibited a lower discriminative performance than FI in hospitalized and community-dwelling adults. Anyhow, Li et al. [53] found similar AUCs for these two indicators. The differences with our results might be explained by the fact that in this last study, all phenotypical criteria were derived from the questions of the Short Form Survey (SF-36) and not by directly assessing walking speed or grip strength. Probably, comprehensive indicators benefit from the diversity of the information taken into account, with the inclusion of measures corresponding to different domains [51].

Interestingly, our results showed that WS alone exhibited higher AUCs for every outcome when compared with FP, despite the inclusion of walking speed among its criteria. Walking speed has been shown to be a reliable proxy of physical frailty [54, 55]: this might suggest that gait speed already provides a consistent part of the details captured by this operationalization of frailty. On the other hand, walking speed cut-offs employed for frailty phenotypical criteria (lowest quintile, adjusted by sex and height [28]) are particularly strict. While this seems to improve the specificity of FP, it might negatively affect its sensitivity [52] and, thereby, its AUC.

Furthermore, our study confirms that the simple count of chronic diseases is the most accurate indicator in predicting the use of healthcare resources, but is not as reliable in the prediction of mortality, as already described by previous studies [56]. Indeed, diagnoses—more than frailty and mobility impairment—seem to trigger clinical consultations. Previous

studies already showed that increased mortality risk among persons affected by multimorbidity is probably due to a limited number of index diseases, rather than to the accumulation of chronic conditions [57]. Specific clusters of multimorbidity and the speed of accumulation—rather than the simple number—of chronic diseases have been shown to be reliably associated with several negative outcomes [58–61].

Having multiple contacts with care providers was poorly predicted by the studied indicators compared to other outcomes. Several factors might influence the number of contacts with providers, beyond people's healthcare needs: behavioural and psychological traits, distance from the provider's office, as well as social support, economical, and economical and insurance statuses, among others [62–65]. The studied indicators do not evaluate these aspects. Our findings highlight the need for more accurate tools to predict outpatient healthcare use.

Finally, our findings show a general trend of lower predictive accuracy for mortality when the indicators were applied to younger persons. It is likely that a higher functional resilience among younger individuals might explain the inability of currently used indicators to accurately predict poor outcomes among this subset of individuals. These results strengthen the need for a reliable tool, able to capture vulnerability to poor outcomes even among younger old individuals.

The results of the present study should be read in light of some limitations. All indicators were assessed at baseline: change of status during the follow-up might have affected the estimation of the predictive accuracy. Furthermore, minor differences with the original operationalization of some indicators exist and are related to data availability in SNAC-K. In addition, as previously described [11], the SNAC-K population is highly educated and wealthy: this might limit the generalizability of our findings. Anyhow, this issue might play a minor role because our main aim was to investigate the accuracy of different health indicators, which are based on participants' clinical and functional characteristics. Furthermore, we found the prevalence of MM, WS, and FP to be similar to those described in previous studies [55, 66, 67]. Our study has also several major strengths. Firstly, we developed all indicators using variables derived from an in-depth and comprehensive assessment, conducted by physicians and nurses [24]. Furthermore, outcomes were retrieved from national registers, minimizing the risk of loss of information. Lastly, all indicators were built using the same data, allowing therefore a direct comparison of their predictive accuracy. Indeed, to the best of our knowledge, this is the first study directly comparing the accuracy of several indicators

commonly used in geriatric research and practice for the prediction of different clinical outcomes.

Implications

Physicians might employ indicators exhibiting a high prognostic value to better tailor diagnostic and therapeutic decisions. For example, older persons with low life expectancy benefit from therapeutic revisions aimed to control symptoms and improve quality of life [68, 69] and from the avoidance of screening tests that might lead to overdiagnosis [70]. Furthermore, high accuracy indicators might also help to prompt discussion between physicians and patients about preferences in late life [71]. The identification of older persons at increased risk of unplanned hospitalizations might be used in the clinic to plan interventions proven to lower such risk, such as more strict follow-ups [72, 73].

Healthcare policy makers could employ information regarding patients' risk of poor health-related outcomes (such as death and hospitalizations) to better allocate resources. For example, accurately identifying individuals with decreased life expectancy is important for the integration of palliative care in modern healthcare systems [74]. Moreover, several interventions have been shown to decrease the number of hospitalizations [75, 76]: better defining the share of the population at risk of such events might enhance the effectiveness of these strategies. Furthermore, our findings showed that the count of chronic diseases could be used to predict an increased number of outpatient visits.

The indicators considered in our study might be employed according to data availability. For example, WS has already been proposed as a simple measure to be evaluated in clinical practice [77, 78], while the FI might be easily calculated from electronic clinical records [79]. HAT is based on measures easily available in clinical settings [30].

Conclusions

Despite their different theoretical background and practical construction, HAT, WS, and FI were the most accurate predictors of mortality and unplanned hospitalizations in a population of older adults. On the other hand, multimorbidity was the most accurate predictor of contact with multiple providers. The accuracy of the considered indicators was generally lower among younger old persons compared to older ones. Different assessment tools can be used in different circumstances to support physicians during their decision-making process. Some of these tools may also be used to forecast future use of healthcare resources, including both hospital-based and outpatient services.

Additional files

Additional file 1: **Table S1.** Deficits included in the frailty index. (DOCX 13 kb)

Additional file 2: **Table S2.** Variables associated with indicator variable of missing values used during multiple imputation process. (DOCX 13 kb)

Additional file 3: **Table S3.** Areas under ROC curves for different indicators – complete dataset analyses. (DOCX 13 kb)

Additional file 4: **Table S4.** Areas under ROC curves for different indicators – imputed dataset analyses. (DOCX 13 kb)

Additional file 5: **Figure S1.** ROC curves comparison. (JPG 241 kb)

Additional file 6: **Table S5.** Areas under ROC curves for different indicators – analyses stratified by age. (DOCX 15 kb)

Abbreviations
AUC: Area under the receiver operating characteristic curve; FI: Frailty index; FP: Frailty phenotype; HAT: Health Assessment Tool; SD: Standard deviation; WS: Walking speed

Transparency statement
The lead authors (AZ and DLV) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported, that no important aspects of the study have been omitted, and that any discrepancies from the study as originally planned have been explained.

Authors' contributions
AZ, DLV, LF, AM, and DR contributed to the conception and design of the work. AZ contributed to the data analysis. All coauthors contributed to the interpretation of the results. AZ and DLV contributed to the drafting the article. All coauthors contributed to the critical revision of the manuscript. All coauthors contributed to the final approval of the manuscript. All the authors fulfil the ICMJE criteria for authorship.

Availability of data and materials
Data are from the SNAC-K project, a population-based study on ageing and dementia (http://www.snac-k.se/). Access to these original data is available to the research community upon approval by the SNAC-K data management and maintenance committee. Applications for accessing these data can be submitted to Maria Wahlberg (Maria.Wahlberg@ki.se) at the Aging Research Center, Karolinska Institutet.

Ethics approval and consent to participate
Every wave of the study was approved by the Regional Ethical Review Board in Stockholm, Sweden. Written informed consent was obtained from each participant, or from a proxy, in case of cognitive impairment. Public or patients were not involved during the development of this study: anyhow, we plan to disseminate the findings of this research to participants of SNAC-K and to the public.

Competing interests
The authors declare that they have no competing interests.

Zucchelli *et al. BMC Medicine*   (2019) 17:185

Page 9 of 10

Author details
[1]Aging Research Center, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet and Stockholm University, 11330 Stockholm, Sweden. [2]Department of Clinical and Experimental Sciences, University of Brescia, Viale Europa 11, 25121 Brescia, Italy. [3]Centro di Medicina dell'Invecchiamento, IRCCS Fondazione Policlinico "A. Gemelli" and Catholic University of Rome, 00168 Rome, Italy. [4]Stockholm Gerontology Research Center, 11330 Stockholm, Sweden.

References

1. Powers BW, Chaguturu SK, Ferris TG. Optimizing high-risk care management. JAMA. 2015;313(8):795–6.
2. Friedman DJ, Starfield B. Models of population health: their value for US public health practice, policy, and research. Am J Public Health. 2003;93(3):366–9.
3. Covinsky KE, Pierluissi E, Johnston CB. Hospitalization-associated disability: "she was probably able to ambulate, but I'm not sure". JAMA. 2011;306(16):1782–93.
4. Gill TM, Allore HG, Gahbauer EA, Murphy TE. Change in disability after hospitalization or restricted activity in older persons. Jama. 2010;304(17):1919–28.
5. Marcantonio ER. Delirium in hospitalized older adults. N Engl J Med. 2017;377(15):1456–66.
6. Inouye SK, Rushing JT, Foreman MD, Palmer RM, Pompei P. Does delirium contribute to poor hospital outcomes? A three-site epidemiologic study. J Gen Intern Med. 1998;13(4):234–42.
7. Ettinger WH. Can hospitalization-associated disability be prevented? Jama. 2011;306(16):1800–1.
8. Yoon J, Zulman D, Scott JY, Maciejewski ML. Costs associated with multimorbidity among VA patients. Med Care. 2014;52(Suppl 3):S31–6.
9. Geue C, Briggs A, Lewsey J, Lorgelly P. Population ageing and healthcare expenditure projections: new evidence from a time to death approach. Eur J Health Econ. 2014;15(8):885–96.
10. World Health Organization. World report on ageing and health. Geneva: World Health Organization; 2015.
11. Santoni G, Angleman S, Welmer AK, Mangialasche F, Marengoni A, Fratiglioni L. Age-related variation in health status after age 60. PLoS One. 2015;10(3):e0120077.
12. Lowsky DJ, Olshansky SJ, Bhattacharya J, Goldman DP. Heterogeneity in healthy aging. J Gerontol A Biol Sci Med Sci. 2014;69(6):640–9.
13. Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, Meinow B, Fratiglioni L. Aging with multimorbidity: a systematic review of the literature. Ageing Res Rev. 2011;10(4):430–9.
14. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet. 2013;381(9868):752–62.
15. Marengoni A, Vetrano DL, Manes-Gravina E, Bernabei R, Onder G, Palmer K. The relationship between COPD and frailty: a systematic review and meta-analysis of observational studies. Chest. 2018;154(1):21–40.
16. Howlett SE, Rockwood K. Ageing: Develop models of frailty. Nature. 2014;512:253.
17. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, Brach J, Chandler J, Cawthon P, Connor EB, et al. Gait speed and survival in older adults. JAMA. 2011;305(1):50–8.
18. Santoni G, Marengoni A, Calderón-Larrañaga A, Angleman S, Rizzuto D, Welmer AK, Mangialasche F, Orsini N, Fratiglioni L. Defining health trajectories in older adults with five clinical indicators. J Gerontol A Biol Sci Med Sci. 2017;72(8):1123–9.
19. Yourman LC, Lee SJ, Schonberg MA, Widera EW, Smith AK. Prognostic indices for older adults: a systematic review. JAMA. 2012;307(2):182–92.
20. Morley JE, Vellas B, van Kan GA, Anker SD, Bauer JM, Bernabei R, Cesari M, Chumlea WC, Doehner W, Evans J, et al. Frailty consensus: a call to action. J Am Med Dir Assoc. 2013;14(6):392–7.
21. Kojima G, Iliffe S, Walters K. Frailty index as a predictor of mortality: a systematic review and meta-analysis. Age Ageing. 2018;47(2):193–200.
22. Nunes BP, Flores TR, Mielke GI, Thume E, Facchini LA. Multimorbidity and mortality in older adults: a systematic review and meta-analysis. Arch Gerontol Geriatr. 2016;67:130–8.
23. Chang SF, Lin PL. Frail phenotype and mortality prediction: a systematic review and meta-analysis of prospective cohort studies. Int J Nurs Stud. 2015;52(8):1362–74.
24. Lagergren M, Fratiglioni L, Hallberg IR, Berglund J, Elmståhl S, Hagberg B, Holst G, Rennemark M, Sjölund BM, Thorslund M, et al. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). Aging Clin Exp Res. 2004;16(2):158–68.
25. 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(5):489–95.
26. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr. 2008;8:24.
27. Song X, Mitnitski A, Rockwood K. Prevalence and 10-year outcomes of frailty in older adults in relation to deficit accumulation. J Am Geriatr Soc. 2010;58(4):681–7.
28. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56(3):M146–56.
29. Zucchelli A, Vetrano DL, Marengoni A, Grande G, Romanelli G, Calderón-Larrañaga A, Fratiglioni L, Rizzuto D. Frailty predicts short-term survival even in older adults without multimorbidity. Eur J Intern Med. 2018;56:53–6.
30. Calderón-Larrañaga A, Vetrano DL, Onder G, Gimeno-Feliu LA, Coscollar-Santaliestra C, Carfì A, Pisciotta MS, Angleman S, Melis RJF, Santoni G, et al. Assessing and measuring chronic multimorbidity in the older population: a proposal for its operationalization. J Gerontol A Biol Sci Med Sci. 2017;72(10):1417–23.
31. Santoni G et al: Geriatric health charts for individual assessment and prediction of care needs: a population-based prospective study. - PubMed - NCBI. 2019.
32. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, Feychting M, Ljung R. The Swedish cause of death register. Eur J Epidemiol. 2017;32(9):765–73.
33. Ludvigsson JF, Andersson E, Ekbom A, Feychting M, Kim JL, Reuterwall C, Heurgren M, Olausson PO. External review and validation of the Swedish national inpatient register. BMC Public Health. 2011;11:450.
34. Socialstyrelsen. The National Patient Register. In: statistics and data, https://www.socialstyrelsen.se/en/statistics-and-data/registers/alla-register/the-national-patient-register/, vol. 2019: Socialstyrelsen; 2019. Accessed 16 May 2019.
35. Lezak MD. Neuropsychological assessment. 4th ed. Oxford: Oxford University Press; 2004.
36. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc. 1983;31(12):721–7.
37. Wieand S, Gail MH, James BR, James KL. A family of nonparametric statistics for comparing diagnostic markers with paired or unpaired data. Biometrika. 1989;76(3):8.
38. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol. 2009;9:57.
39. Ritt M, Ritt JI, Sieber CC, Gassmann KG. Comparing the predictive accuracy of frailty, comorbidity, and disability for mortality: a 1-year follow-up in patients hospitalized in geriatric wards. Clin Interv Aging. 2017;12:293–304.
40. Theou O, Brothers TD, Mitnitski A, Rockwood K. Operationalization of frailty using eight commonly used scales and comparison of their ability to predict all-cause mortality. J Am Geriatr Soc. 2013;61(9):1537–51.
41. Bongue B, Buisson A, Dupre C, Beland F, Gonthier R, Crawford-Achour E. Predictive performance of four frailty screening tools in community-dwelling elderly. BMC Geriatr. 2017;17(1):262.
42. Ritt M, Schwarz C, Kronawitter V, Delinic A, Bollheimer LC, Gassmann KG, Sieber CC. Analysis of Rockwood et Al's clinical frailty scale and Fried et Al's frailty phenotype as predictors of mortality and other clinical outcomes in older patients who were admitted to a geriatric Ward. J Nutr Health Aging. 2015;19(10):1043–8.
43. Alonso-Moran E, Nuno-Solinis R, Onder G, Tonnara G. Multimorbidity in risk stratification tools to predict negative outcomes in adult population. Eur J Intern Med. 2015;26(3):182–9.
44. Landi F, Calvani R, Tosato M, Martone AM, Bernabei R, Onder G, Marzetti E. Impact of physical function impairment and multimorbidity on mortality among community-living older persons with sarcopaenia: results from the ilSIRENTE prospective cohort study. BMJ Open. 2016;6(7):e008281.
45. Callahan KE, Lovato L, Miller ME, Marsh AP, Fielding RA, Gill TM, Groessl EJ, Guralnik J, King AC, Kritchevsky SB, et al. Self-reported physical function as a

predictor of hospitalization in the lifestyle interventions and Independence for elders study. J Am Geriatr Soc. 2018;66(10):1927–33.

46. Majer IM, Nusselder WJ, Mackenbach JP, Klijs B, van Baal PH. Mortality risk associated with disability: a population-based record linkage study. Am J Public Health. 2011;101(12):e9–15.

47. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. J Aging Phys Act. 2015;23(2):314–22.

48. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal. 2001;1:323–36.

49. Ferrucci L, Levine ME, Kuo PL, Simonsick EM. Time and the metrics of aging. Circ Res. 2018;123(7):740–4.

50. Thinggaard M, McGue M, Jeune B, Osler M, Vaupel JW, Christensen K. Survival prognosis in very old adults. J Am Geriatr Soc. 2016;64(1):81–8.

51. Marengoni A, Bandinelli S, Maietti E, Guralnik J, Zuliani G, Ferrucci L, Volpato S. Combining gait speed and recall memory to predict survival in late life: population-based study. J Am Geriatr Soc. 2017;65(3):614–8.

52. Widagdo IS, Pratt N, Russell M, Roughead EE. Predictive performance of four frailty measures in an older Australian population. Age Ageing. 2015;44(6): 967–72.

53. Li G, Thabane L, Ioannidis G, Kennedy C, Papaioannou A, Adachi JD. Comparison between frailty index of deficit accumulation and phenotypic model to predict risk of falls: data from the global longitudinal study of osteoporosis in women (GLOW) Hamilton cohort. PLoS One. 2015;10(3): e0120144.

54. Lee L, Patel T, Costa A, Bryce E, Hillier LM, Slonim K, Hunter SW, Heckman G, Molnar F. Screening for frailty in primary care: accuracy of gait speed and hand-grip strength. Can Fam Physician. 2017;63(1):e51–7.

55. Castell MV, Sanchez M, Julian R, Queipo R, Martin S, Otero A. Frailty prevalence and slow walking speed in persons age 65 and older: implications for primary care. BMC Fam Pract. 2013;14:86.

56. Bahler C, Huber CA, Brungger B, Reich O. Multimorbidity, health care utilization and costs in an elderly community-dwelling population: a claims data based observational study. BMC Health Serv Res. 2015;15:23.

57. Rizzuto D, Melis RJF, Angleman S, Qiu C, Marengoni A. Effect of chronic diseases and multimorbidity on survival and functioning in elderly adults. J Am Geriatr Soc. 2017;65(5):1056–60.

58. Vetrano DL, Calderón-Larrañaga A, Marengoni A, Onder G, Bauer JM, Cesari M, Ferrucci L, Fratiglioni L. An international perspective on chronic multimorbidity: approaching the elephant in the room. J Gerontol A Biol Sci Med Sci. 2018;73(10):1350–6.

59. Vetrano DL, Rizzuto D, Calderon-Larranaga A, Onder G, Welmer AK, Bernabei R, Marengoni A, Fratiglioni L. Trajectories of functional decline in older adults with neuropsychiatric and cardiovascular multimorbidity: a Swedish cohort study. PLoS Med. 2018;15(3):e1002503.

60. Calderon-Larranaga A, Fratiglioni L. Multimorbidity research at the crossroads: developing the scientific evidence for clinical practice and health policy. J Intern Med. 2019;285(3):251–4.

61. Calderon-Larranaga A, Santoni G, Wang HX, Welmer AK, Rizzuto D, Vetrano DL, Marengoni A, Fratiglioni L. Rapidly developing multimorbidity and disability in older adults: does social background matter? J Intern Med. 2018;283(5):489–99.

62. Kannan VD, Veazie PJ. Predictors of avoiding medical care and reasons for avoidance behavior. Med Care. 2014;52(4):336–45.

63. Broadhead WE, Gehlbach SH, deGruy FV, Kaplan BH. Functional versus structural social support and health care utilization in a family medicine outpatient practice. Med Care. 1989;27(3):221–33.

64. Hsu WC, Hsu YP. Patterns of outpatient care utilization by seniors under the National Health Insurance in Taiwan. J Formos Med Assoc. 2016;115(5):325–34.

65. Zayas CE, He Z, Yuan J, Maldonado-Molina M, Hogan W, Modave F, Guo Y, Bian J. Examining healthcare utilization patterns of elderly middle-aged adults in the United States. Proc Int Fla AI Res Soc Conf. 2016;2016:361–6.

66. Thompson MQ, Theou O, Karnon J, Adams RJ, Visvanathan R. Frailty prevalence in Australia: findings from four pooled Australian cohort studies. Australas J Ageing. 2018;37(2):155–8.

67. Excoffier S, Herzig L, N'Goran AA, Deruaz-Luyet A, Haller DM. Prevalence of multimorbidity in general practice: a cross-sectional study within the Swiss Sentinel Surveillance System (Sentinella). BMJ Open. 2018;8(3):e019616.

68. Morin L, Vetrano DL, Rizzuto D, Calderon-Larranaga A, Fastbom J, Johnell K. Choosing wisely? Measuring the burden of medications in older adults near the end of life: Nationwide, Longitudinal Cohort Study. Am J Med. 2017; 130(8):927–936.e929.

69. Holmes HM, Hayley DC, Alexander GC, Sachs GA. Reconsidering medication appropriateness for patients late in life. Arch Intern Med. 2006;166(6):605–9.

70. Croft P, Altman DG, Deeks JJ, Dunn KM, Hay AD, Hemingway H, LeResche L, Peat G, Perel P, Petersen SE, et al. The science of clinical practice: disease diagnosis or patient prognosis? Evidence about "what is likely to happen" should shape clinical practice. BMC Med. 2015;13:20.

71. Covinsky KE, Fuller JD, Yaffe K, Johnston CB, Hamel MB, Lynn J, Teno JM, Phillips RS. Communication and decision-making in seriously ill patients: findings of the SUPPORT project. The study to understand prognoses and preferences for outcomes and risks of treatments. J Am Geriatr Soc. 2000; 48(5 Suppl):S187–93.

72. McAlister FA, Youngson E, Kaul P, Ezekowitz JA. Early follow-up after a heart failure exacerbation: the importance of continuity. Circ Heart Fail. 2016;9(9): e003194. https://www.ahajournals.org/doi/full/10.1161/ CIRCHEARTFAILURE.116.003194?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3 Acrossref.org&rfr_dat=cr_pub%3Dpubmed.

73. Irewall AL, Ögren J, Bergström L, Laurell K, Söderström L, Mooe T. Nurse-led, telephone-based, secondary preventive follow-up after stroke or transient ischemic attack improves blood pressure and LDL cholesterol: results from the first 12 months of the randomized, controlled NAILED Stroke Risk Factor Trial. PLoS One. 2015;10:e0139997.

74. World Health Organization. Integrating palliative care and symptom relief into primary health care: a WHO guide for planners, implementers and managers: World Health Organization; 2018. Available at: https://apps.who. int/iris/handle/10665/274559. License: CC BY-NC-SA 3.0 IGO

75. Levine S, Steinman BA, Attaway K, Jung T, Enguidanos S. Home care program for patients at high risk of hospitalization. Am J Manag Care. 2012; 18(8):e269–76.

76. Kripalani S, Theobald CN, Anctil B, Vasilevskis EE. Reducing hospital readmission rates: current strategies and future directions. Annu Rev Med. 2014;65:471–85.

77. Studenski S, Perera S, Wallace D, Chandler JM, Duncan PW, Rooney E, Fox M, Guralnik JM. Physical performance measures in the clinical setting. J Am Geriatr Soc. 2003;51(3):314–22.

78. Cesari M, Kritchevsky SB, Newman AB, Simonsick EM, Harris TB, Penninx BW, Brach JS, Tylavsky FA, Satterfield S, Bauer DC, et al. Added value of physical performance measures in predicting adverse health-related events: results from the health, aging and body composition study. J Am Geriatr Soc. 2009; 57(2):251–9.

79. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, Mohammed MA, Parry J, Marshall T. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age Ageing. 2016; 45(3):353–60.

Publisher's Note