

Willeke M Ravensbergen, Jeanet W Blom, Andrea WM Evers, Mattijs E Numans, Margot WM de Waal and Jacobijn Gussekloo

Measuring daily functioning in older persons using a frailty index:

a cohort study based on routine primary care data

Abstract

Background

Electronic health records (EHRs) are increasingly used for research; however, multicomponent outcome measures such as daily functioning cannot yet be readily extracted.

Aim

To evaluate whether an electronic frailty index based on routine primary care data can be used as a measure for daily functioning in research with community-dwelling older persons (aged ≥ 75 years).

Design and setting

Cohort study among participants of the Integrated Systemic Care for Older People (ISCOPE) trial (11 476 eligible; 7285 in observational cohort; 3141 in trial; over-representation of frail people).

Method

At baseline (T0) and after 12 months (T12), daily functioning was measured with the Groningen Activities Restriction Scale (GARS, range 18–72). Electronic frailty index scores (range 0–1) at T0 and T12 were computed from the EHRs. The electronic frailty index (electronic Frailty Index – Utrecht) was tested for responsiveness and compared with the GARS as a gold standard for daily functioning.

Results

In total, 1390 participants with complete EHR and follow-up data were selected (31.4% male; median age = 81 years, interquartile range = 78–85). The electronic frailty index increased with age, was higher for females, and lower for participants living with a partner. It was responsive after an acute major medical event; however, the correlation between the electronic frailty index and GARS at T0 and over time was limited.

Conclusion

Because the electronic frailty index does not reflect daily functioning, further research on new methods to measure daily functioning with routine care data (for example, other proxies) is needed before EHRs can be a useful data source for research with older persons.

Keywords

activities of daily living; aged; electronic health records; frail older persons; general practice.

INTRODUCTION

The use of routine care data such as electronic health records (EHRs) for research and population health management is increasing. These EHRs could be a valuable data source for research with older persons, which is often expensive and time-consuming. Some variables (for example, diagnoses, death, hospital admissions, polypharmacy, and multimorbidity) can be easily extracted from GPs' EHRs. However, often in research with older persons, complex, multicomponent outcome measures such as quality of life and functioning are used. These variables cannot be readily extracted from EHR data.^{1,2} Daily functioning, which is often used as an outcome measure in the older population, is such a variable.^{3–7} It is described in terms of basic Activities of Daily Living (ADL) and instrumental Activities of Daily Living (iADL). Both in research and clinical practice, these are currently assessed with questionnaires such as the Katz ADL scale, the Lawton iADL scale, or the Groningen Activities Restriction Scale (GARS [ADL and iADL]).^{8–10} As reflected in those questionnaires, daily functioning is the result of a patient's physical, psychological, cognitive, and social status.¹¹ A potential measure of daily functioning based on items of the EHR should therefore incorporate these different aspects.

The frailty index (FI), as outlined by Rockwood *et al*,¹² integrates the different aspects mentioned above (that is, physical, psychological, cognitive, and social functioning) into one measure.^{13,14} An FI consists of a comprehensive list of deficits and functional losses in different domains, from which a continuous score is calculated by dividing the number of deficits present in an individual by the total number of deficits from the list (score range 0–1).^{15,16} Most FIs are derived from questionnaires, but more recently FIs were developed that were derived from routine care data.^{12,17} Previous research has shown that the scores of the FI are stable across different versions of the FI and across different data sources used.^{15,16}

Some researchers have suggested that the integration of multiple domains of functioning into the FI make it a potentially useful evaluative measure for health status or functioning.^{13,15,18} However, other researchers state that a measure of functional decline should not only include the number of deficits, but also the severity and impact of each deficit, which would make the FI unfit as a measure of daily functioning.^{19,20} If an older person's daily functioning can be extracted from routine care data it opens new opportunities for research in large datasets, potentially saving costs and time in research. The FI is

WM Ravensbergen, MSc, PhD candidate;
JW Blom, MD, PhD, senior researcher;
ME Numans, MD, PhD, professor of general practice; **MWM de Waal**, PhD, senior researcher, Department of Public Health and Primary Care, Leiden University Medical Center, Leiden.
AWM Evers, PhD, professor of health psychology, Health, Medical and Neuropsychology Unit, Institute of Psychology, Leiden University, Leiden, the Netherlands. **J Gussekloo**, MD, PhD, professor of research in primary care, Department of Internal medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden.

Address for correspondence

Willeke M Ravensbergen, Department of Public Health and Primary Care, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, the Netherlands.

Email: W.M.Ravensbergen@lumc.nl

Submitted: 25 February 2020; **Editor's response:** 13 April 2020; **final acceptance:** 19 May 2020.

©British Journal of General Practice

This is the full-length article (published online 3 Nov 2020) of an abridged version published in print. Cite this version as: **Br J Gen Pract 2020; DOI: <https://doi.org/10.3399/bjgp20X713453>**

How this fits in

Daily functioning is an often used outcome measure in the older population. If it could be extracted from routine care data it could save cost and time for both research and general practice. Although there are currently no established methods to measure daily functioning with routine primary care data, an electronic frailty index was suggested as a potentially useful evaluative measure for functioning. The electronic frailty index tested in this study (electronic Frailty Index — Utrecht) was responsive after an acute major medical event, but did not compare well with the gold standard for daily functioning (that is, the Groningen Activities Restriction Scale). Therefore, in its current state and context, the electronic frailty index cannot be used in research or general practice because of its limited ability to reflect daily functioning.

currently the only multicomponent outcome measure that can be extracted from EHRs, but it is still unclear whether it could serve as a proxy for daily functioning. The aim of this study was to test whether an electronic FI based on routine primary care data can be used as an evaluative measure for daily functioning in research with older persons.

METHOD

Design

This was a prospective cohort study embedded in the Integrated Systemic Care for Older People (ISCOPE) trial. Further details about the trial are described elsewhere.²¹

ISCOPE study

The ISCOPE study included 59 general practices from the Leiden region (the Netherlands). All patients aged ≥ 75 years enlisted in these practices were invited to participate. Exclusion criteria were:

- life expectancy < 3 months;
- nursing home resident;
- non-Dutch speaking; and
- considered to be too ill to participate by the GP.

Postal screening questionnaires together with an invitation to participate in the study were sent to 11 476 older persons. The ISCOPE screening questionnaire consisted of questions on four health domains (that is, functional; somatic; psychological; and social). Those who filled in and returned the ISCOPE screening questionnaire and

the informed consent form ($n = 7285$) were included in the study. Inclusion took place from September 2009 to September 2010. All participants gave informed consent.²¹

For the trial, a selection of the participants ($n = 3141$) were included for a 12-month follow-up. This sample consisted of all participants with problems on three or four domains of the ISCOPE screening questionnaire, a random sample of 60% of participants with problems on two domains, and a random sample of 15% of participants with problems on one or no domain. At baseline (T0) they were visited at home by a research nurse to collect extra information on sociodemographic characteristics and to administer additional questionnaires (that is, GARS and Mini-Mental State Examination [MMSE]; range = 0 to 30). After 12 months (T12) the measurements were administered again. In addition, data over a period of 5 years until 1 year after the first home visit were extracted from the participants' EHRs. The extracted data contained both diagnoses with *International Classification of Primary Care* (ICPC)-1-NL codes, prescriptions with Anatomical Therapeutic Chemical (ATC) codes, and free text. The EHR data were linked to the study data on a person-level using a personal identification number.

Participants

Inclusion criteria for this secondary analysis were a complete follow-up (T12), an available EHR, and at least one ICPC or ATC code registered in the EHR (that is, necessary to compute the electronic FI). Participants with missing values on either the GARS or the electronic FI were also excluded from the analyses ($n = 23$).

Measures

Electronic Frailty Index — Utrecht (eFI-U). In this study the electronic FI was used as developed by Drubbel *et al*²²⁻²⁶ (the eFI-U). This FI is generated from routine primary care data and consists of a list of 50 deficits (Supplementary Table S1). It includes physical, psychological, cognitive, and social deficits. Each deficit again consists of a list of ICPC and ATC codes related to that deficit. If one ICPC or ATC code was present in the previous 6 months or 5 years (depending on the code), the corresponding deficit scores positive (that is, one point). Diagnostic measurement data were not included in the eFI-U of this study, because these data were not extracted in the ISCOPE study.

The Groningen Activities Restriction Scale (GARS) The GARS was used as a gold

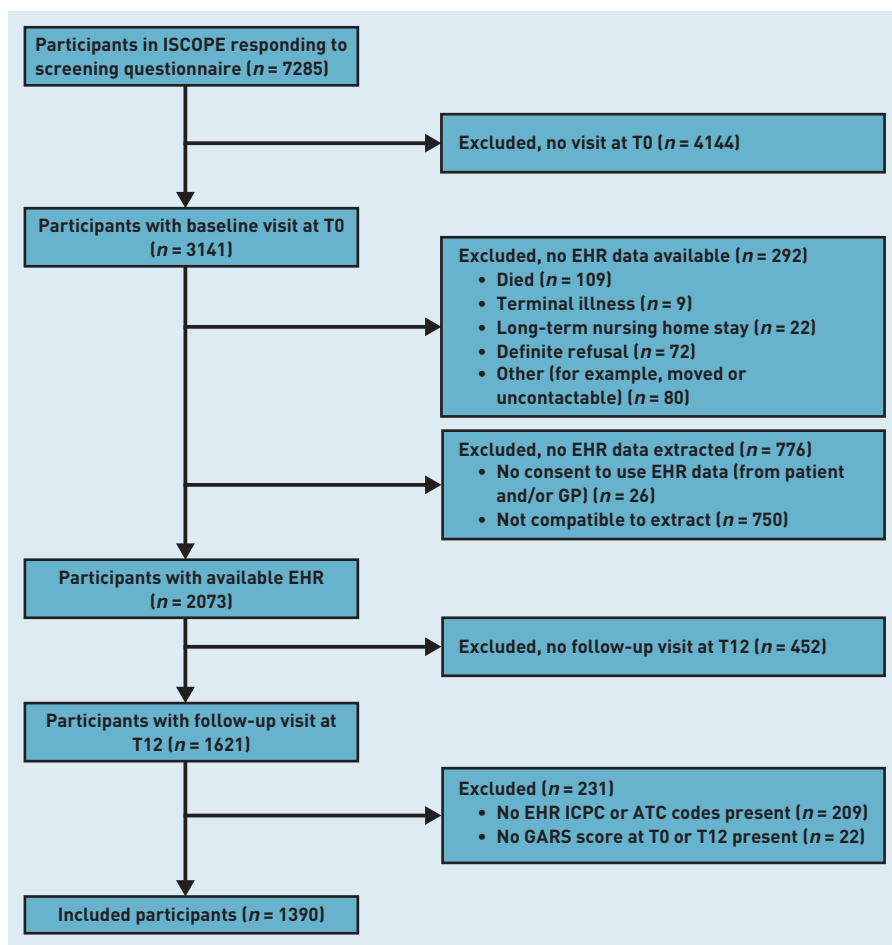


Figure 1. Participant inclusion flowchart.
 ATC = anatomical therapeutic chemical.
 EHR = electronic health record. GARS = Groningen Activities Restriction Scale. ICPC = International Classification of Primary Care. ISCOPE = Integrated Systemic Care for Older People. T0 = baseline. T12 = after 12 months.

standard for measuring daily functioning. The GARS is an 18-item questionnaire with 11 questions on basic ADL and seven questions on iADL. Each question has four answer categories:

- fully independent without problems;
- fully independent, but with some difficulty;
- fully independent, but with a lot of difficulty; and
- only with another person's help.

The total score ranges from 18 to 72 points, with a higher score indicating a lower level of daily functioning or more dependency.

Subgroups. Subgroups based on the occurrence of an acute major medical event during follow-up were compared. In this study, an acute major medical event was defined as a medical event with a sudden onset, which is likely to have a large impact on a person's daily functioning. In this study, hip fracture, myocardial infarction, and stroke were included as

acute major medical events. These events were considered to be present either if participants reported them in the follow-up questionnaire, or if corresponding ICPC codes were registered during the follow-up period. This was done to assure that all participants with an event during the follow-up period were identified. The ICPC codes included were L75 (femur fracture), K75 (acute myocardial infarction), K89 (transient cerebral ischaemia), and K90 (stroke).

Statistical analysis

Characteristics of the population at baseline were described. The construct validity of the eFI-U was assessed by comparing subgroups based on age, sex, and living status. Based on previous findings with the GARS, it was hypothesised that, if the baseline eFI-U measured daily functioning, average scores would increase with age, be higher for females compared with males, and be highest for those living in a residential care facility and lowest for those living independently with a partner. This was tested with Spearman's correlation (age), the Mann-Whitney *U* test (sex), and the Kruskal-Wallis test (living status: independently alone; independently with partner; or residential care facility).²⁷⁻³²

To test the eFI-U for floor and ceiling effects, a histogram of the eFI-U at baseline was created for visual inspection. Floor or ceiling effects were considered to be present if >15% of participants reached the lowest or highest possible score, which was also tested. The upper limit of the eFI-U was assessed by plotting the 99th percentiles of the baseline eFI-U in the cohort against age.

For both the eFI-U and the GARS the difference between the follow-up and baseline scores was calculated (delta = measurement at 12 months minus measurement at baseline). The delta scores were also corrected for the baseline scores, because the latter influence the potential change over time. The resulting relative deltas were calculated as the actual delta divided by the maximum delta possible for that patient (relative delta = [measurement at 12 months minus measurement at baseline] divided by [total score minus measurement at baseline plus 0.01]). An extra 0.01 was added to the denominator to avoid a value of zero.

To explore responsiveness, the occurrence of an acute major medical event during follow-up was used as an implicit external criterion of larger change. The delta and relative delta eFI-U scores of the groups with and without event were described

Table 1. Sociodemographic and functional characteristics of the total study population at T0 (baseline)

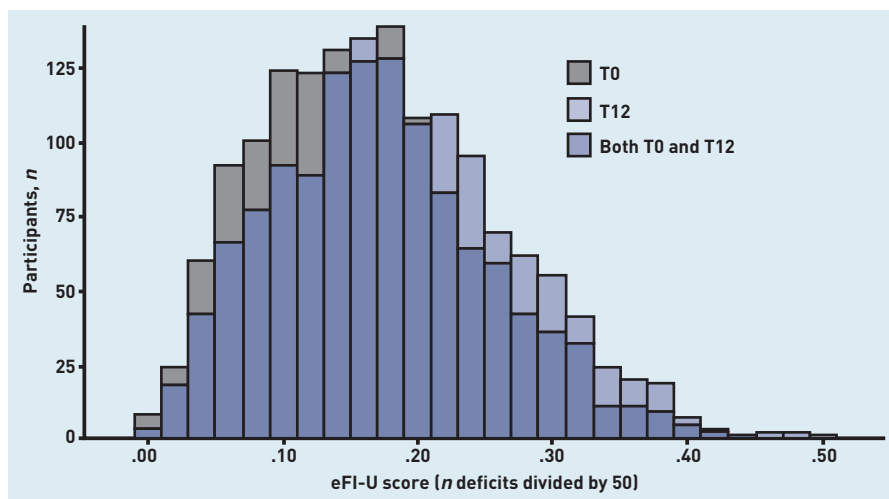
Characteristics	Total population (N = 1390)
Age, years ^a	81 (78 to 85)
Male, n (%)	436 (31.4)
Living situation, n (%)^b	
Independent alone	768 (55.3)
Independent together	507 (36.5)
Residential care facility	114 (8.2)
Education, n (%)^{b,c}	
Low	509 (36.6)
Medium	755 (54.3)
High	125 (9.0)
Low income, n (%)^{b,d}	206 (14.8)
MMSE^{a,b}	28 (26 to 29)
GARS^a	30 (24 to 38)
eFI-U^a	0.16 (0.10 to 0.22)

^aContinuous data are presented by median and interquartile range (25th and 75th percentile).

^bMissing data: living situation (n = 1, 0.1%); low education (n = 1, 0.1%); low income (n = 1, 0.1%); and MMSE (n = 11, 0.8%). ^cLow education defined as primary school only; medium education defined as secondary school only; high education defined as higher vocational education or university. ^dLow income defined as state pension only (no additional pension). eFI-U = electronic Frailty Index — Utrecht. GARS = Groningen Activities Restriction Scale. MMSE = Mini-Mental State Examination.

Figure 2. Distribution of the eFI-U scores at T0 and T12 of the total population (N = 1390).

eFI-U = electronic Frailty Index — Utrecht. T0 = baseline. T12 = after 12 months.



and compared with a Mann-Whitney *U* test. Standardised effect sizes (Cohen's *d*) were calculated for both the (relative) delta eFI-U and the (relative) delta GARS. The standardised effect sizes of the eFI-U and the GARS were expected to be similar and both were expected to be small to moderate.

Criterion validity of the eFI-U was assessed with Spearman's correlation between the baseline eFI-U and the baseline GARS. The association between changes (delta and relative delta) in the eFI-U and the GARS was also tested with Spearman's correlation. If the eFI-U measured daily functioning, the correlation coefficient was expected to be ≥ 0.70 in both cases.³²⁻³⁴

To get a better understanding of the relationship between the eFI-U and the GARS over time, participants were grouped in quartiles according to their delta GARS scores. All delta GARS quartiles were compared on delta eFI-U scores (Jonckheere-Terpstra test) and on the number of acute major medical events during follow up (χ^2 test for trend). In addition, the baseline GARS scores were compared between the delta GARS quartiles to check whether correction for baseline scores was needed. Because of significant differences between the quartiles in GARS score at baseline, the same analyses were repeated with quartiles based on the relative delta. The same analyses were also carried out with quartiles based on the (relative) delta eFI-U scores [Supplementary Tables S2 and S3].

RESULTS

A flowchart of the participants is presented in Figure 1. Table 1 displays the characteristics of the 1390 older persons included in the analyses. The delta eFI-U

was approximately normally distributed and ranged from -0.14 to +0.20 (Supplementary Table S4).

Construct validity

The baseline eFI-U scores were higher in the older participants, but the association with age was smaller than expected (Spearman's $\rho = 0.071$; $P = 0.008$). As expected, females on average had a higher eFI-U score at baseline compared with males (Mann-Whitney *U* test, $P < 0.001$; median females = 0.16, interquartile range [IQR] = 0.10 to 0.22 versus males = 0.14, IQR = 0.08 to 0.20). Furthermore, participants who lived in a residential care facility had the highest eFI-U score at baseline and those living independently with a partner the lowest (Kruskal-Wallis test, $P < 0.001$; median institutionalised = 0.18, IQR = 0.12 to 0.26; median independently alone = 0.16, IQR = 0.10 to 0.22; median independently with partner = 0.14, IQR = 0.10 to 0.20) (data not shown).

Floor or ceiling effects

The histogram of the baseline eFI-U showed a slight right-skewed distribution, approaching a gamma distribution (Figure 2). The baseline eFI-U score in the total group ranged from 0.00 to 0.46. The 15% highest score was ≥ 0.25 and the 15% lowest score was ≤ 0.08 , suggesting that there was no floor or ceiling effect. No common maximum of the eFI-U at every age was observed, which again suggested that there was no ceiling effect.¹⁶

Responsiveness of the Electronic Frailty Index — Utrecht (acute major medical events)

During follow-up, 193 participants (13.9%) experienced an acute major medical event (that is, hip fracture, myocardial infarction, and/or stroke) (Table 2). Of those 193 participants, 185 had one type of event and eight had two different types of events during follow-up. In total, 22 (1.6%) participants had a hip fracture, 64 (4.6%) a myocardial infarction, and 115 (8.3%) a stroke (data not shown). Characteristics of the participants with and without an acute major medical event during follow-up are described in Table 2.

There was a significant difference in (relative) delta eFI-U between participants with and without an acute major medical event during follow-up (mean absolute delta = 0.039, standard deviation [SD] 0.052 versus 0.020, SD 0.043; $P < 0.001$; relative delta = 0.047, SD 0.064 versus 0.023, SD 0.051; $P < 0.001$) (Table 2). The

Table 2. Characteristics of subgroups based on the presence of an acute major medical event during follow-up

Characteristics	Acute major medical event		P-value
	Yes (N = 193)	No (N = 1197)	
Age, years ^a	82 (78 to 86)	81 (78 to 85)	
Male, N(%)	76 (39.4)	360 (30.1)	
MMSE at T0 ^{a,b}	28 (27 to 29)	28 (26 to 29)	
Living situation, n(%)^b			
Independent alone	101 (52.3)	667 (55.7)	
Independent together	78 (40.4)	429 (35.8)	
Residential care facility	14 (7.3)	100 (8.4)	
eFI-U^a			
Baseline	0.20 (0.13 to 0.26)	0.14 (0.10 to 0.20)	
Follow-up	0.24 (0.16 to 0.30)	0.18 (0.12 to 0.24)	
Delta ^c	0.0392 (0.05239)	0.0202 (0.04284)	<0.001
Relative delta ^c	0.0472 (0.06421)	0.0229 (0.05141)	<0.001
GARS^a			
Baseline	34 (27 to 43)	29 (23 to 38)	
Follow-up	39 (30 to 48)	32 (25 to 41)	
Delta ^c	3.7927 (6.88889)	2.4436 (6.31102)	0.012
Relative delta ^c	0.1082 (0.22952)	0.0547 (0.23711)	0.003

^aContinuous data are presented by median and interquartile range (25th and 75th percentile). ^bMissing data: living situation (n = 1, 0.1%, for the 'No' response), and MMSE (n = 11, 0.8%). ^cApproximately normal distribution (visual). eFI-U = electronic Frailty Index — Utrecht. GARS = Groningen activities restriction scale. MMSE = Mini-Mental State Examination. T0 = baseline.

standardised effect sizes were 0.42 (delta) and 0.45 (relative delta), which can both be considered small but present (data not shown). The difference in delta and relative delta GARS between participants with and without an acute major medical event during follow-up was also significant. The standardised effect size was 0.21 for the delta GARS and 0.23 for the relative delta GARS, which can both be considered small but present, just like the standardised effect

sizes of the (relative) delta eFI-U (data not shown).

Criterion validity

At baseline Spearman's ρ between the eFI-U and the GARS was 0.374 ($P < 0.001$). Figure 3 is a graphic representation of the relationship between the delta eFI-U and the delta GARS. The correlation coefficient between the delta eFI-U and the delta GARS was 0.088 and the correlation coefficient of the relative deltas was 0.097 (both $P \leq 0.001$). No regression analysis was done because of the low correlation between the delta GARS and the delta eFI-U.

Comparison GARS quartiles

More in-depth, the median delta eFI-U across the quartiles of the delta GARS was 0.02 (IQR = 0.00 to 0.04) for the first quartile, 0.02 (IQR = 0.00 to 0.04) for the second quartile, 0.02 (IQR = 0.00 to 0.06) for the third quartile, and 0.02 (IQR = 0.00 to 0.04) for the fourth quartile ($P = 0.003$) (Table 3). By contrast, there was a large and significant difference in median delta GARS over the delta GARS quartiles, as expected ($P < 0.001$). Furthermore, the incidence of acute major medical events during follow-up increased over the quartiles (13.0% in the lowest quartile compared with 20.5% in the highest quartile; $P = 0.005$). The baseline GARS was highest for the participants in the lowest delta GARS quartile ($P = 0.029$). These differences in GARS at baseline suggest that the low change of the GARS during follow-up in the lowest quartile might be partly due to a high baseline GARS (that is, participants are not able to get much higher). Therefore, the same analyses were repeated with quartiles based on the relative delta GARS. Apart from the baseline GARS score the findings did not change much (Supplementary Table S5).

DISCUSSION

Summary

This study explored whether an electronic FI based on routine primary care data can be used as an evaluative measure for daily functioning in research with older persons. As the electronic FI tested in this study (eFI-U) changed over time and did not have floor or ceiling effects, it might be useful as an evaluative measure; however, there was a moderate overlap between the eFI-U and the GARS. Furthermore, the eFI-U was responsive after an acute major medical event, just like the GARS, but it was barely responsive over time in the population as a whole, which was different from the GARS. These findings suggest that

Figure 3. Delta eFI-U scores against delta GARS scores for those with and without an acute major medical event during follow-up. With event (red) (n = 193); without event (blue) (n = 1197). eFI-U = electronic Frailty Index — Utrecht. GARS = Groningen Activities Restriction Scale.

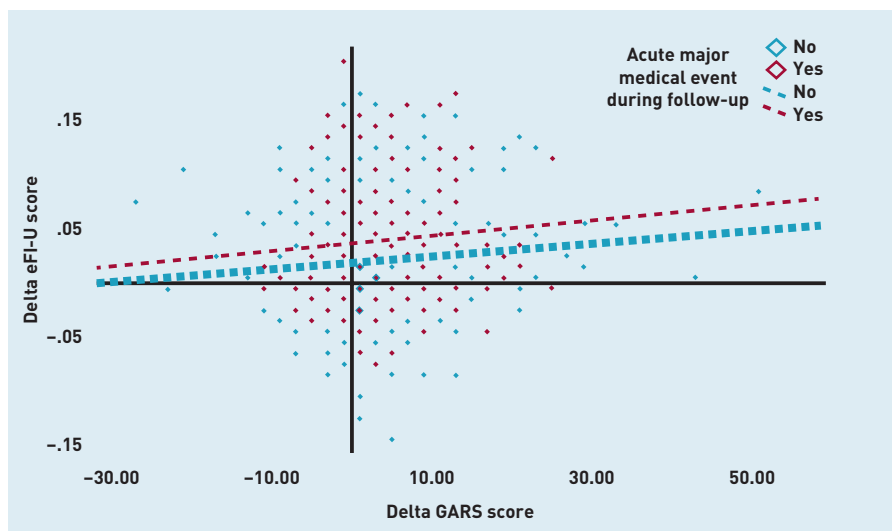


Table 3. Comparison between lowest and highest delta Groningen Activities Restriction Scale (GARS) quartiles

	Delta GARS			
	Quartile 1 (0%–<25%)	Quartile 2 (25%–<50%)	Quartile 3 (50%–<75%)	Quartile 4 (75%–100%)
GARS, median (IQR)^a				
Baseline ^b	33 [27 to 43]	25 [20 to 34]	30 [23 to 37]	30 [25 to 38]
Follow-up	29 [23 to 38]	26 [21 to 35]	33 [26 to 41]	41 [35 to 51]
Delta ^b	-3 [-5 to -2]	0 [0 to 1]	3 [2 to 4]	9 [7 to 12]
eFI-U, median (IQR)				
Baseline	0.16 [0.10 to 0.22]	0.14 [0.08 to 0.18]	0.16 [0.12 to 0.22]	0.18 [0.10 to 0.22]
Follow-up	0.18 [0.12 to 0.24]	0.16 [0.10 to 0.20]	0.18 [0.12 to 0.24]	0.18 [0.14 to 0.26]
Delta ^b	0.02 [0.00 to 0.04]	0.02 [0.00 to 0.04]	0.02 [0.00 to 0.06]	0.02 [0.00 to 0.04]
Major medical events, n (%)^b	50 (13.0)	29 (10.0)	42 (11.6)	72 (20.5)

^aThe GARS and eFI-U were not normally distributed for any of the variables. ^bP-value for trend: baseline GARS P = 0.029; delta GARS P < 0.001; delta eFI-U P = 0.003; and major medical events P = 0.005. eFI-U = electronic Frailty Index – Utrecht. GARS = Groningen Activities Restriction Scale.

the eFI-U does not reflect daily functioning in older persons.

Strengths and limitations

The main strength of this study is the high generalisability of the results due to the data and the instrument used. Previous studies already showed that the FI, because of the underlying concept of deficit accumulation, is a flexible instrument that can be based on different deficits and data sources, and still give the same results.^{15,16} The data used in this study (EHRs from Dutch general practices) are similar to many other routine care data in that they contain both codes and free text, which increase the generalisability of the results of this study. Another strength is the availability of a combination of routine care data and standardised questionnaires from the same community-dwelling older population and time period. Combining these data sources allows for a direct comparison of the EHR-derived instrument with a gold standard for daily functioning (that is, GARS). Furthermore, because of the availability of extensive prospective data, the authors were able to assess responsiveness by looking both over time (gradual decline in ADL/iADL) and after an event (sudden change in ADL/iADL).

This study also has some limitations. First, part of the lack of correlation in the study might be explained by the EHR data on which the electronic FI was based. Quality and completeness of coded routine care data fully rely on the ability and willingness of the primary care team to code and prioritise their findings in routine healthcare systems. Second, quite a few patients had to be excluded because they

were not selected for follow-up or were lost to follow-up in the ISCOPE trial. This drop-out is likely to be associated with poor daily functioning and/or a higher level of deficits. The attrition and complete case analysis in this study, therefore, might have skewed the responses and weakened the effects found. Some patients were also excluded because of missing or unavailable EHRs; however, most of these missing EHRs are expected to be completely random as they were missing at practice level because of software problems. Thus, the influence on the results is expected to be limited.

Another limitation concerns the combination of the electronic FI tested in this study with the type of data from which it is derived. The eFI-U is a cumulative score based on EHRs of general practices. As a result, those patients who have been registered with their GP for a long time and those who visit more often are more likely to accumulate recorded deficits and thus have a higher eFI-U score compared with other patients. For any instrument based on EHR data, the influence of consultation frequency and registration period, among other factors, should be taken into account.

Comparison with existing literature

In previous literature some researchers suggested that an FI could serve as an evaluative measure for daily functioning because of its multicomponent nature.^{13,15,18} However, other researchers stated that this was not possible because frailty and disability are different constructs, and because no information on severity and impact is included in an FI.^{19,20,35,36}

The results of this study using the eFI-U are in line with studies that showed a limited association between frailty and daily functioning.^{35,36} The authors of these studies propose that frailty and disability are overlapping but distinct concepts. Thus, an instrument that is designed to measure frailty will not be able to measure disability and vice versa. The findings of this study showed that an FI based on EHR data also does not reflect measurements of (daily) functioning. Furthermore, these findings are in line with studies on the relationship between the number of diseases or deficits and functional decline.^{19,20} As was already concluded by those studies, functioning or daily functioning is not only a matter of the number of deficits (which is the approach of an FI), but also of the severity and impact of each deficit. The current study shows that this is also the case when routine primary care data are used to count deficits. An electronic FI could be enriched

with information on severity, and more importantly impact, through the use of new techniques such as plain-text mining and other advanced reading techniques, which are a proven approach to increase the quality of algorithms like an electronic FI. However, it is doubtful whether EHRs contain enough information on severity and impact.

Implications for research and practice

An evaluative measure for daily functioning that can be obtained from routine care data could be useful both for research (to replace time-consuming questionnaires) and clinical purposes (to monitor patients). In research, such a measure may save costs and time for both the researcher and the clinician. Furthermore, it may allow for more efficient and faster research, which might in the end improve patient

outcomes and day-to-day general practice management. This study showed that the FI (with a deficit-counting approach), in its current state and context, has a limited ability to reflect daily functioning. As the electronic FI does not measure the aimed construct it cannot be used as an evaluative measure of daily functioning for research. The lack of precision and congruence of the eFI-U with the GARS means that it is even further away from use in clinic to monitor individual patients' daily functioning. Further research could focus on other approaches (that is, other proxies or adjusted versions of the electronic FI) to measure daily functioning with routine care data. It must be noted that previous research has shown that the eFI-U can be used in population health management as a frailty identification instrument on a population level.^{23,25}

Funding

The Integrated Systemic Care for Older People (ISCOPE) trial (the Netherlands trial register: NTR1946) was funded by the Netherlands Organisation for Health Research and Development (Zon-MW) (project number: 311060201). The work of Willeke M Ravensbergen was funded by the Leiden University research profile area: Health, Prevention and the Human Life Cycle. Both funders had no role in the design and conduct of the study; collection, management, analysis, or interpretation of the data; or preparation, review, or approval of the manuscript. The researchers were independent from the funders.

Ethical approval

The ISCOPE trial was approved by the Medical Ethical Committee of the Leiden University Medical Center (reference number: P09.096). All participants gave written informed consent.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors thank the members of the Advisory Board of Older Persons 'Care and Wellbeing' South-Holland North for their involvement in the design and conduct of the study.

Discuss this article

Contribute and read comments about this article: bjgp.org/letters

REFERENCES

1. Thygesen LC, Ersboll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. *Eur J Epidemiol* 2014; **29**(8): 551–558.
2. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health* 2016; **37**: 61–81.
3. Haywood KL, Griffin XL, Achten J, Costa ML. Developing a core outcome set for hip fracture trials. *Bone Joint J* 2014; **96-B**(8): 1016–1023.
4. Hutchings L, Fox R, Chesser T. Proximal femoral fractures in the elderly: how are we measuring outcome? *Injury* 2011; **42**(11): 1205–1213.
5. Bowers TM, Parker MJ. Assessment of outcome after hip fracture: development of a universal assessment system for hip fractures. *SICOT J* 2016; **2**: 27.
6. Parsons N, Griffin XL, Achten J, Costa ML. Outcome assessment after hip fracture: is EQ-5D the answer? *Bone Joint Res* 2014; **3**(3): 69–75.
7. Akpan A, Roberts C, Bandeen-Roche K, et al. Standard set of health outcome measures for older persons. *BMC Geriatr* 2018; **18**(1): 36.
8. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; **9**(3): 179–186.
9. Kempen GI, Suurmeijer TP. The development of a hierarchical polychotomous ADL-IADL scale for noninstitutionalized elders. *Gerontologist* 1990; **30**(4): 497–502.
10. Katz S, Ford AB, Moskowitz RW, et al. Studies of illness in the aged. The index of ADL: a standardised measure of biological and psychosocial function. *JAMA* 1963; **185**: 914–919.
11. World Health Organization. *International classification of functioning, disability and health*. 2001. <https://www.who.int/classifications/icf/en> [accessed 25 Sep 2020].
12. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005; **173**(5): 489–495.
13. Sieliwonczyk E, Perkisas S, Vandewoude M. Frailty indexes, screening instruments and their application in Belgian primary care. *Acta Clin Belg* 2014; **69**(4): 233–239.
14. Faller JW, Pereira DDN, de Souza S, et al. Instruments for the detection of frailty syndrome in older adults: a systematic review. *PLoS One* 2019; **14**(4): e0216166.
15. Cesari M, Gambassi G, van Kan GA, Vellas B. The frailty phenotype and the frailty index: different instruments for different purposes. *Age Ageing* 2014; **43**(1): 10–12.
16. Searle SD, Mitnitski A, Gahbauer EA, et al. A standard procedure for creating a frailty index. *BMC Geriatr* 2008; **8**: 24.
17. Rockwood K, Mitnitski AB, MacKnight C. Some mathematical models of frailty and their clinical implications. *Rev Clin Gerontol* 2002; **12**(2): 109–117.
18. Rockwood K, Theou O, Mitnitski A. What are frailty instruments for? *Age Ageing* 2015; **44**(4): 545–547.
19. Boeckxstaens P, Vaes B, Legrand D, et al. The relationship of multimorbidity with disability and frailty in the oldest patients: a cross-sectional analysis of three measures of multimorbidity in the BELFRAIL cohort. *Eur J Gen Pract* 2015; **21**(1): 39–44.
20. Fried LP, Ferrucci L, Darer J, et al. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci* 2004; **59**(3): 255–263.
21. Blom J, den Elzen W, van Houwelingen AH, et al. Effectiveness and cost-effectiveness of a proactive, goal-oriented, integrated care model in general practice for older people. A cluster randomised controlled trial: Integrated Systematic Care for Older People — the ISCOPE study. *Age Ageing* 2016; **45**(1): 30–41.
22. Bleijenberg N, Drubbel I, Ten Dam VH, et al. Proactive and integrated primary care for frail older people: design and methodological challenges of the Utrecht primary care PROactive Frailty Intervention Trial (U-PROFIT). *BMC Geriatr* 2012; **12**: 16.
23. Drubbel I, de Wit NJ, Bleijenberg N, et al. Prediction of adverse health outcomes in older people using a frailty index based on routine primary care data. *J Gerontol A Biol Sci Med Sci* 2013; **68**(3): 301–308.
24. Drubbel I. *Frailty screening in older patients in primary care using routine care data*. Thesis. Utrecht: Utrecht University, 2014.
25. Bleijenberg N, Drubbel I, Schuurmans MJ, et al. Effectiveness of a proactive primary care program on preserving daily functioning of older people: a cluster randomized controlled trial. *J Am Geriatr Soc* 2016; **64**(9): 1779–1788.
26. Drubbel I, Bleijenberg N, Kranenburg G, et al. Identifying frailty: do the Frailty Index and Groningen Frailty Indicator cover different clinical perspectives? A cross-sectional study. *BMC Fam Pract* 2013; **14**: 64.
27. Kempen GI, Miedema I, Ormel J, Molenaar W. The assessment of disability with the Groningen Activity Restriction Scale. Conceptual framework and psychometric properties. *Soc Sci Med* 1996; **43**(11): 1601–1610.
28. Suurmeijer TP, Doeglas DM, Moum T, et al. The Groningen Activity Restriction Scale for measuring disability: its utility in international comparisons. *Am J Public Health* 1994; **84**(8): 1270–1273.
29. Moerman S, Mathijssen NM, Tuinebreijer WE, et al. Less than one-third of hip fracture patients return to their prefracture level of instrumental activities of daily living in a prospective cohort study of 480 patients. *Geriatr Gerontol Int* 2018; **18**(8): 1244–1248.
30. Rodgers V, Neville S, La Grow S. Health, functional ability and life satisfaction among older people 65 years and over: a cross-sectional study. *Contemp Nurse* 2017; **53**(3): 284–292.
31. Lutomski JE, Krabbe PF, den Elzen WP, et al. Rasch analysis reveals comparative analyses of activities of daily living/instrumental activities of daily living summary scores from different residential settings is inappropriate. *J Clin Epidemiol* 2016; **74**: 207–217.
32. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018; **27**(5): 1147–1157.
33. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; **60**(1): 34–42.
34. Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saude* 2017; **26**(3): 649–659.
35. Fried LP, Tangen CM, Walston J, et al. Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001; **56**(3): M146–M156.
36. Rodriguez-Manas L, Feart C, Mann G, et al. 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**(1): 62–67.