

# Predictive performance of four frailty measures in an older Australian population

IMAINA S. WIDAGDO<sup>1</sup>, NICOLE PRATT<sup>1</sup>, MARY RUSSELL<sup>2</sup>, ELIZABETH E. ROUGHEAD<sup>1</sup>

<sup>1</sup>Quality Use of Medicines and Pharmacy Research Centre, Sansom Institute, School of Pharmacy and Medical Sciences, University of South Australia, Adelaide, SA, Australia

<sup>2</sup>Occupational Therapy Board, Australian Health Practitioner Regulation Agency, Melbourne, VIC, Australia

Address correspondence to: I. S. Widagdo. Tel: (+61) 8 83021249. Email: imaina.widagdo@mymail.unisa.edu.au

## Abstract

**Background:** there are several different frailty measures available for identifying the frail elderly. However, their predictive performance in an Australian population has not been examined.

**Objective:** to examine the predictive performance of four internationally validated frailty measures in an older Australian population.

**Methods:** a retrospective study in the Australian Longitudinal Study of Ageing (ALSA) with 2,087 participants. Frailty was measured at baseline using frailty phenotype (FP), simplified frailty phenotype (SFP), frailty index (FI) and prognostic frailty score (PFS). Odds ratios (OR) were calculated to measure the association between frailty and outcomes at Wave 3 including mortality, hospitalisation, nursing home admission, fall and a combination of all outcomes. Predictive performance was measured by assessing sensitivity, specificity, positive and negative predictive values (PPV and NPV) and likelihood ratio (LR). Area under the curve (AUC) of dichotomised and the multilevel or continuous model of the measures was examined.

**Results:** prevalence of frailty varied from 2% up to 49% between the measures. Frailty was significantly associated with an increased risk of any outcome, OR (95% confidence interval) for FP: 1.9 (1.4–2.8), SFP: 3.6 (1.5–8.8), FI: 3.4 (2.7–4.3) and PFS: 2.3 (1.8–2.8). PFS had high sensitivity across all outcomes (sensitivity: 55.2–77.1%). The PPV for any outcome was highest for SFP and FI (70.8 and 69.7%, respectively). Only FI had acceptable accuracy in predicting outcomes, AUC: 0.59–0.70.

**Conclusions:** being identified as frail by any of the four measures was associated with an increased risk of outcomes; however, their predictive accuracy varied.

**Keywords:** frail elderly, sensitivity, specificity, Australia, older people

## Introduction

Ageing is associated with declining health; therefore, the increase in the ageing population is likely to have significant impact on healthcare services [1]. Not all older people have the same rate of declining health, some are at higher risk of adverse health outcomes than the others, and this is often recognised as frailty [1–4].

Identifying frailty is becoming increasingly important in clinical decision making, particularly when considering the risks and benefits of treatment in an older population [5–8]. Many frailty measures have been developed to help identify frail older people, and the measures have shown that frailty is associated with an increased risk of adverse outcomes, including falls, nursing home admission, hospitalisation and death [9–11]. Several population-based studies have examined the predictive performance of frailty measures in predicting

adverse outcomes [9, 11–13]. However, no study has assessed the predictive performance of frailty measures in the Australian setting.

Our study aimed to compare the accuracy of four internationally validated frailty measures for predicting mortality, hospitalisation, nursing home admission and fall, and these outcomes combined in an older Australian population.

## Methods

### Study population

This study used data from the Australian Longitudinal Study of Ageing (ALSA) which is a population-based cohort study of 2,087 older people aged 70 and older, or their spouses aged 65 and older [14]. The ALSA data set contains 11 waves of data, collected between 1992 and 2010. Baseline

data (Wave 1) were used to measure frailty status. Outcome variables were obtained from Wave 3 follow-up data. The follow-up time from Wave 1 to Wave 3 was up to 3 years. Further details of the ALSA study can be found at: <http://www.flinders.edu.au/sabs/fcas/alsa/>.

### Frailty measures

The four frailty measures selected in this study were as follows: frailty phenotype (FP)—developed by Fried *et al.* [3], simplified frailty phenotype (SFP)—developed by Kiely *et al.* [15], frailty index (FI)—developed by Mitnitski *et al.* [16] and the prognostic frailty score (PFS)—developed by Ravaglia *et al.* [17].

### Frailty phenotype

The FP is a unidimensional measure which only assesses the physical characteristic of frailty. In the original study, it included five variables: unintentional weight loss, low grip strength, self-rated exhaustion—assessed using two questions from the Center of Epidemiologic Studies Depression (CES-D) Scale, low physical activity—assessed by calculating energy expenditure (Kcal/week) and slow walking time [3]. In our study, the FP was compiled as follows: body mass index (BMI) <20, low grip strength, self-rated exhaustion—assessed using two questions from the Center of Epidemiologic Studies Depression (CES-D) Scale, physical activity—assessed using a question on walking for exercise or recreation in the past 2 weeks and slow walking time. Frailty status was categorised into three levels: frail, pre-frail and non-frail. The presence of three or more variables was defined as frail, and pre-frail was defined by the presence of up to two variables. Non-frail was the absence of any variables. Participants with two or more missing variables were excluded from the analysis.

### Simplified frailty phenotype

The SFP is a short version of FP; it included only three variables of physical characteristics: unintentional weight loss, inability to rise from a chair 5 times without use of arms and low energy level (assessed using one question from the CED-S scale) [15]. In our study, only the unintentional weight loss variable was modified to BMI <20. This measure also categorised frailty status into three levels: participants with two or more variables present were classified as frail and the presence of one variable was categorised as pre-frail. Only participants with complete data in all of the three variables were included in the analysis.

### Frailty index

The FI is a multidimensional measure; it assesses four different characteristics of frailty: physical, medical, psychological and social contributors [16, 18]. In this study, we used frailty index with 39 variables (Supplementary data, Appendix S1, available in *Age and Ageing* online) that was previously used by Mitnitski *et al.* [16] in assessing frailty among ALSA population. Frailty was calculated by dividing the number of

deficits present with the total number of deficits assessed for each subject. The frailty score is a continuous score which ranges from 0 to 1; a greater score indicates increased frailty. The score was dichotomised to allow comparison with other frailty measures [19, 20]. Consistent with previous research that has dichotomised the frailty index, a frailty score of 0.25 or greater was the criteria for frailty; a score of <0.25 was classified as non-frail [19, 20]. Participants with 10 or more missing variables were excluded in the analysis.

### Prognostic frailty score

The PFS is also a multidimensional measure; it assesses four different characteristics of frailty: age, physical, medical and psychological contributions [17]. The score consists of nine variables: age  $\geq 80$  years, male gender, low physical activity defined as <4 h per week of moderate intensity activity, co-morbidity, sensory deficit, calf circumference <31 cm, instrumental activity of daily living (IADL) dependence, gait problem or health pessimism [17]. In this study, the nine variables were age  $\geq 80$  years old, male gender, low physical activity—assessed by a question on walking for exercise/recreation in the past 2 weeks, co-morbidity, sensory deficit, calf circumference <31 cm, IADL dependence, gait problem and health pessimism. Frailty was classified as the presence of three or more variables, and the presence of two or less variables was classified as non-frail [17]. Participants with complete data or those with only one missing variable of the nine variables were included in the analysis.

### Reference standard

In assessing the accuracy of a measure, a reference standard is needed to differentiate between persons with a target condition and those without [21]. There is no available consensus on the clinical or laboratory markers of frailty [4], which can be used as a reference standard. In this situation, the prognostic value of a measure can be used as a reference standard [21, 22]. Experts have agreed that a frailty measure ‘can help in identifying and stratifying older persons at high risk of disability and/or other adverse outcomes’ [4]; therefore, the presence of adverse outcomes has been commonly used as a reference standard in assessing the performance of a frailty measure [13]. There is no agreement yet of which adverse outcomes are important to assess and whether to assess them individually or as a combined variable [2].

Studies validating frailty measures have used a number of outcomes including, mortality and adverse health outcomes such as hospitalisation, fall or nursing home admissions [3, 9, 15–17]. In this study, the adverse outcomes chosen as the reference standard were death, nursing home admission in the last 12 months, hospitalisation in the past year, fall—at least one fall in the past year, and a combination of these outcomes. Death data were obtained from the Registry of Births, Deaths and Marriages [14], while other outcomes were obtained from self-reports at Wave 3. As interviews occurred at the subjects’ place of residence, nursing home admission is likely to be verified by the interviewer. Outcome variables

were coded as binary variables: the presence or absence of outcomes. Disability was not chosen as the reference standard in this study, because difficulties in daily activities that are used to define disability were also used in FI and PFS. Of the 1,679 participants at Wave 3, adverse outcomes data were missing in 81, 8 and 7 participants, for data on nursing home admission, hospitalisation and fall, respectively.

**Statistical analysis**

The proportion of participants with each adverse outcome was examined for each frailty status and compared with those no frailty status because of missing data. A logistic regression analysis was used to examine the association between frailty as identified by each of the measures and adverse outcomes. This analysis used a dichotomised frailty status (frail and non-frail); for those measures that had a pre-frail group, this was combined into non-frail. For each measure, the sensitivity and specificity, positive predictive value (PPV), negative predictive value (NPV) and likelihood ratio (LR) were calculated. The area under the curve (AUC) values were used to examine the ability of each measure to differentiate between the frail and non-frail participants. In this study, AUC values between 0.6 and 0.85 were considered as acceptable accuracy for a prognostic measure [23, 24]. The AUC values for the full model (i.e. multilevel or continuous scale) of the frailty measures were also assessed. Because the frailty index is a continuous measure, not a dichotomous measure, the predictive performance of the frailty index was assessed using different cut-off points to determine whether the predictive performance differed where different cut-off points were used (Supplementary data, Appendix S2, available in *Age and Ageing* online). To determine whether the predictive validity of the frailty measures using full population data was comparable, we also undertook a secondary analysis where the predictive performance of the four measures was compared on the sub-sample of the population that had complete data for all four measures (Supplementary data, Appendix S3, available in *Age and Ageing* online). No adjustment was made for age or gender, because the PFS includes

age and sex in the assessment of frailty. All analyses were undertaken using the SAS software (Version 9.3 of the SAS system for Windows. Copyright 2011 SAS Institute Inc., Cary, NC, USA).

**Results**

Baseline characteristics of the ALSA participants are presented in Table 1. Participants with missing data were older and more likely to be living in an institution compared with those in the analysis group. The majority of participants were living in the community. The prevalence of frailty varied between the measures; the SFP identified the lowest number of frail participants—2%, whereas the prognostic frailty scale identified the highest number of frail participants—49%.

Figure 1 shows the distribution of adverse outcomes by frailty status for all participants. The proportion of adverse outcomes across frailty status was different for all four measures, except for nursing home admission across frailty status identified by FP. Participants with missing data had a greater proportion of any adverse outcomes than their non-frail and pre-frail counterpart, but were often had less proportion of any adverse outcomes than those who were identified as frail.

In all four measures those identified as frail were more likely to have adverse outcomes than the pre-frail or non-frail (Table 2). The odds of death within 3 years was highest among the frail group identified by FP—odds ratio (OR): 3.6, 95% confidence interval (CI): 2.4–5.3. The SFP identified frail participants with the highest odds of hospitalisation and new nursing home admission. The odds of a fall were highest among the frail group identified by FI.

The PFS had the highest sensitivity values (58.6–77.1%); however, its specificity values were the lowest (54.7–59.5%) across all outcomes. The other three measures had <50% sensitivity but had high specificity (>80%). In all four measures, the PPV was highest when assessed for a combined outcome (59.1–70.8%) than for any individual outcome. The NPV was higher in all four measures for individual outcomes than for combined outcomes. PFS had low LR values (1.22–1.70) in all

**Table 1.** Baseline characteristics and prevalence of frailty across four measures

| Characteristics           | Frailty phenotype       |                      |         | Simplified frailty phenotype |                      |         | Frailty index<br>Analysis<br>(n = 2,087) | Prognostic frailty score |                      |         |
|---------------------------|-------------------------|----------------------|---------|------------------------------|----------------------|---------|--|--------------------------|----------------------|---------|
|                           | Analysis<br>(n = 1,566) | Missing<br>(n = 521) | P-value | Analysis<br>(n = 1,173)      | Missing<br>(n = 914) | P-value |  | Analysis<br>(n = 1,485)  | Missing<br>(n = 602) | P-value |
| Age, mean (SD)            | 77.4 (6.4)              | 80.5 (7.01)          | <0.0001 | 76.7 (6.0)                   | 80.1 (6.99)          | <0.0001 | 78.2 (6.7)                               | 77.3 (6.3)               | 80.4 (7.03)          | <0.0001 |
| Gender, n (%)             |                         |                      |         |                              |                      |         |  |                          |                      |         |
| Male                      | 816 (52.1)              | 240 (46.1)           | 0.0169  | 606 (51.7)                   | 450 (49.2)           | 0.2710  | 1,056 (50.6)                             | 766 (51.6)               | 290 (48.2)           | 0.1581  |
| Female                    | 750 (47.9)              | 281 (53.9)           |         | 567 (48.3)                   | 464 (50.8)           |         | 1,031 (49.4)                             | 719 (48.4)               | 312 (51.8)           |         |
| Living arrangement, n (%) |                         |                      |         |                              |                      |         |  |                          |                      |         |
| Community living          | 1,505 (96.1)            | 456 (87.5)           | <0.0001 | 1,148 (97.9)                 | 813 (88.9)           | <0.0001 | 1,961 (94)                               | 1,428 (96.2)             | 533 (88.5)           | <0.0001 |
| Institution               | 61 (3.9)                | 65 (12.5)            |         | 25 (2.1)                     | 101 (11.1)           |         | 126 (6)                                  | 57 (3.8)                 | 69 (11.5)            |         |
| Frailty prevalence, n (%) |                         |                      |         |                              |                      |         |  |                          |                      |         |
| Frail                     | 137 (8.8)               | n/a                  | n/a     | 24 (2)                       | n/a                  | n/a     | 366 (17.5)                               | 733 (49.4)               | n/a                  | n/a     |
| Pre-frail                 | 904 (57.7)              |                      |         | 172 (14.8)                   |                      |         | –  | –                        |                      |         |
| Non-frail                 | 525 (33.5)              |                      |         | 976 (83.2)                   |                      |         | 1,721 (82.5)                             | 752 (50.6)               |                      |         |

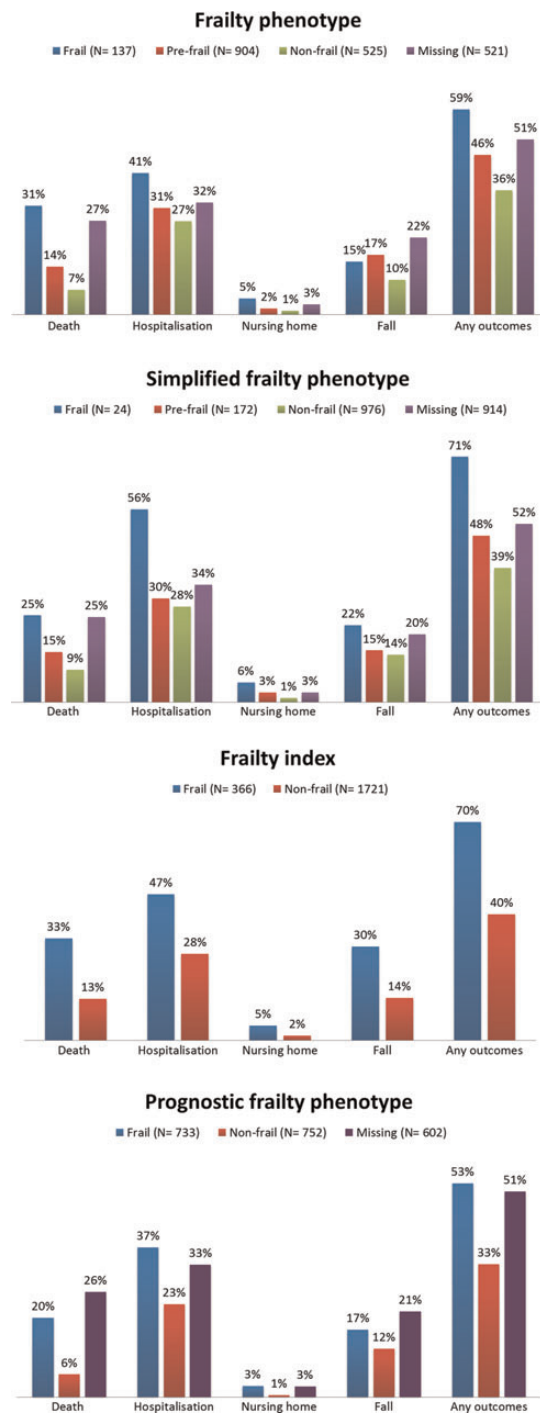


Figure 1. Adverse outcomes by frailty status.

of the adverse outcomes assessed, compared with a range of 1.99–2.76 from another multidimensional measure—FI.

The AUC values of the dichotomised model of frailty measures were comparable to those of the multilevel or continuous model (Table 2). Of the 2,087 participants, only 1,113 had complete data on the variables of all four measures; the predictive performance of all four measures was reduced when applied in this group (Supplementary data, Appendix S3, available in *Age and Ageing* online).

## Discussion

This is the first study in an older Australian population that examined four internationally validated frailty measures for their accuracy in predicting adverse outcomes. Our findings showed that being identified as frail by each of the four measures was associated with a range of adverse health outcomes. The PFS was found to be the most sensitive measure in our population; this means that the PFS correctly identified people who were at higher risk of adverse outcomes as frail more often than the other three measures. However, the low PPV values indicate that only a small number of those who were identified as frail had the adverse outcomes. The probability of having any of the adverse outcomes was highest among the frail older people identified by the SFP and the FI (PPV = 70%). Therefore, SFP and FI may be more useful in clinical practice, as recognising the high probability of adverse outcomes can provide better information for clinicians when considering the risks and benefits of a treatment for their frail older patients.

FI was the only measure that was able to be used to assess frailty in all participants without any loss of participants due to missing data. Between 25% up and 44% of participants were unable to be assessed by the other three measures due to missing data. Participants who were excluded because of missing data had a higher proportion of adverse outcomes than the pre-frail or the non-frail group. Given the purpose of frailty measure is to identify individuals with higher risk of adverse outcomes [4], the large proportion of adverse outcomes among the missing group may have influenced the measure's performance. This may have contributed to the lower accuracy for the FP, SFP and PFS. Several modifications to the variables used in the assessment of frailty were needed in this study, due to lack of data availability. These modifications may have some effects on the predictive performance of the four measures. A similar population-based study of the Italian older population examined the predictive accuracy of PFS (without modifications) and SFP (with a different modification from our study) [17]. The Italian study used adverse outcomes including death, hospitalisation, nursing home admission and fall, with a 4-year follow-up. It reported comparable sensitivity values for the PFS 48–78%; however, the AUC values were slightly better: 0.65–0.83 [9]. For SFP, the Italian study reported higher sensitivity values 11–36% compared with 2.8–6.7% in our study [9].

A limitation of our study was the reliance on self-report data for some of the reference standards including hospitalisation and fall. We cannot rule out recall bias; further, the frail elderly may have been more likely to have had cognitive impairment that led to poorer recall. Nursing home admission was less likely to suffer from self-report bias as it could have been validated by the interviewer at the time of the interview at the subjects' residence. While these limitations exist, the results of the self-reported outcomes were concordant with the outcome for death.

Overall, our analysis showed that all four frailty measures had low discriminating ability with LR values that were <5 [25, 26]. However, based on the AUC values, the FI had

Table 2. Predictive performance of four frailty measures in predicting adverse outcomes

| Frailty measure              | Prevalence of outcome | % outcome (frail/non-frail <sup>a</sup> ) | OR (95% CI)    | Sensitivity | Specificity | PPV  | NPV  | LR   | AUC (dichotomised model) | AUC (full model) <sup>b</sup> |
|------------------------------|-----------------------|---|----------------|-------------|-------------|------|------|------|--------------------------|-------------------------------|
| .....                        |                       |   |                |             |             |      |      |      |                          |                               |
| Death at Wave 3              |                       |   |                |             |             |      |      |      |                          |                               |
| Frailty phenotype            | 205 (13.1%)           | 31.4/11.3                                 | 3.6 (2.4–5.3)  | 20.9        | 93.1        | 31.4 | 88.7 | 3.03 | 0.57                     | 0.63                          |
| Simplified frailty phenotype | 122 (10.4%)           | 25.0/10.1                                 | 2.9 (1.2–7.6)  | 4.9         | 98.3        | 25.0 | 89.9 | 2.88 | 0.52                     | 0.55                          |
| Frailty index                | 346 (16.6%)           | 32.5/13.2                                 | 3.2 (2.4–4.1)  | 34.4        | 85.8        | 32.5 | 86.8 | 2.42 | 0.60                     | 0.66                          |
| Prognostic frailty score     | 188 (12.7%)           | 19.8/5.7                                  | 4.1 (2.8–5.8)  | 77.1        | 54.7        | 19.8 | 94.3 | 1.70 | 0.66                     | 0.66                          |
| Hospitalisation              |                       |   |                |             |             |      |      |      |                          |                               |
| Frailty phenotype            | 404 (30.1%)           | 40.8/29.3                                 | 1.7 (1.1–2.5)  | 9.9         | 93.8        | 40.8 | 70.7 | 1.59 | 0.52                     | 0.54                          |
| Simplified frailty phenotype | 292 (28.4%)           | 55.6/27.9                                 | 3.2 (1.3–8.3)  | 3.4         | 98.9        | 55.6 | 72.1 | 3.09 | 0.51                     | 0.52                          |
| Frailty index                | 513 (30.6%)           | 46.7/27.6                                 | 2.3 (1.7–3.0)  | 23.8        | 88.1        | 46.7 | 72.4 | 1.99 | 0.56                     | 0.59                          |
| Prognostic frailty score     | 379 (29.8%)           | 37.3/23.2                                 | 1.9 (1.6–2.5)  | 58.6        | 58.3        | 37.3 | 76.8 | 1.40 | 0.58                     | 0.58                          |
| Nursing home admission       |                       |   |                |             |             |      |      |      |                          |                               |
| Frailty phenotype            | 22 (1.7%)             | 4.6/1.5                                   | 3.2 (1.0–9.5)  | 18.2        | 93.4        | 4.6  | 98.5 | 2.76 | 0.56                     | 0.59                          |
| Simplified frailty phenotype | 15 (1.5%)             | 5.6/1.4                                   | 4.1 (0.5–33.1) | 6.7         | 98.3        | 5.6  | 98.6 | 3.94 | 0.53                     | 0.59                          |
| Frailty index                | 31 (1.9%)             | 4.7/1.5                                   | 3.3 (1.6–7.0)  | 35.5        | 85.8        | 4.7  | 98.5 | 2.5  | 0.61                     | 0.7                           |
| Prognostic frailty score     | 21 (1.7%)             | 2.8/0.8                                   | 3.9 (1.4–10.6) | 76.2        | 54.8        | 2.8  | 99.3 | 1.69 | 0.66                     | 0.66                          |
| Fall                         |                       |   |                |             |             |      |      |      |                          |                               |
| Frailty phenotype            | 195 (14.5%)           | 15.3/14.5                                 | 1.1 (0.6–1.9)  | 7.7         | 92.8        | 15.3 | 85.5 | 1.07 | 0.50                     | 0.57                          |
| Simplified frailty phenotype | 143 (13.9%)           | 22.2/13.8                                 | 1.8 (0.6–5.5)  | 2.8         | 98.4        | 22.2 | 86.2 | 1.75 | 0.50                     | 0.51                          |
| Frailty index                | 270 (16.1%)           | 29.9/13.5                                 | 2.7 (2.0–3.7)  | 28.9        | 87.0        | 29.9 | 86.5 | 2.22 | 0.58                     | 0.61                          |
| Prognostic frailty score     | 183 (14.4%)           | 16.9/12.1                                 | 1.5 (1.1–2.0)  | 55.2        | 54.7        | 16.9 | 87.9 | 1.22 | 0.55                     | 0.55                          |
| Any outcomes                 |                       |   |                |             |             |      |      |      |                          |                               |
| Frailty phenotype            | 686 (43.8%)           | 59.1/42.3                                 | 1.9 (1.4–2.8)  | 11.8        | 93.6        | 59.1 | 57.7 | 1.84 | 0.53                     | 0.57                          |
| Simplified frailty phenotype | 478 (40.8%)           | 70.8/40.1                                 | 3.6 (1.5–8.8)  | 3.6         | 98.9        | 70.8 | 59.9 | 3.27 | 0.51                     | 0.54                          |
| Frailty index                | 949 (45.5%)           | 69.7/40.3                                 | 3.4 (2.7–4.3)  | 26.8        | 90.3        | 69.7 | 59.7 | 2.76 | 0.59                     | 0.64                          |
| Prognostic frailty score     | 640 (43.1%)           | 53.3/33.1                                 | 2.3 (1.8–2.8)  | 61.1        | 59.5        | 53.3 | 66.9 | 1.51 | 0.60                     | 0.60                          |

<sup>a</sup>Non-frail group from frailty phenotype and simplified frailty phenotype analysis was combined with pre-frail group.

<sup>b</sup>Full model includes multilevel or continuous version of the frailty measure; % outcome, proportion of outcome among frail and non-frail group; CI, confidence interval; sensitivity, probability of persons with outcome who are classified as frail; specificity, probability of persons without outcome who are classified as non-frail; PPV, probability that a person has the outcome if he/she was identified as frail; NPV, probability that a person has no outcome if he/she was identified as non-frail; likelihood ratio (LR): [sensitivity/100 – specificity]—how much more likely is a person with the outcome to be classified as frail than those without the outcome; AUC, area under the curve.

acceptable accuracy for a prognostic measure with AUC between 0.59 and 0.7 [23, 24].

To improve the validity of frailty measures, future research could focus on assessing other aspects of validity such as construct validity.

### Conclusions

Our study showed that being identified as frail by any of the four measures was significantly associated with an increased risk of adverse outcomes. However, only the FI had acceptable predictive accuracy.

### Key points

- Being identified as frail by each of the four measures was associated with an increased risk of adverse health outcomes.
- Frailty measures had poor sensitivity in predicting adverse outcomes and had low ability in discriminating between the frail and the non-frail group.
- Participants identified as frail by SFP and FI were more likely to have the combined outcome (PPV = 70%).

### Acknowledgements

We thank the participants in ALSA, who have given their time over many years, and gratefully acknowledge the work of the project team at the Flinders Centre for Ageing Studies, Flinders University, Adelaide, who carried out the ALSA and provided data for this paper.

### Conflicts of interest

None declared.

### Funding

The Australian Longitudinal Study of Ageing (ALSA) was initially funded by a grant from the US National Institute of Health (AG 08523-02), with additional funding from the South Australian government, Flinders University and other NGOs. Subsequent funding has been provided by the Australian Research Council (ARC-LP 0669272, ARC-LP 100200413, ARC-DP 0879152 and ARC-DP 130100428) and the National Health and Medical Research Council (NH&MRC 179839 and 229922). This research was funded by a PhD scholarship provided by the Australian Government Department of Veterans' Affairs.

## Supplementary data

Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

## References

- Shamliyan T, Talley KMC, Ramakrishnan R, Kane RL. Association of frailty with survival: A systematic literature review. *Ageing Res Rev* 2013; 12: 719–36.
- 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: 301–8.
- Fried L, Tangen C, Walston J *et al.* Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci* 2001; 56: M146–56.
- Rodríguez-Mañas L, Féart C, Mann G *et al.* 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: 62–7.
- Clegg A, Young J, Iliffe S *et al.* Frailty in elderly people. *Lancet* 2013; 381: 752–62.
- Pialoux T, Goyard J, Lesourd B. Screening tools for frailty in primary health care: a systematic review. *Geriatr Gerontol Int* 2012; 12: 189–97.
- Sternberg SA, Schwartz AW, Karunanathan S *et al.* The identification of frailty: a systematic literature review. *J Am Geriatr Soc* 2011; 59: 2129–38.
- Theou O, Rockwood K. Should frailty status always be considered when treating the elderly patient? *Aging Health* 2012; 8: 261–71.
- Forti P, Rietti E, Pisacane N *et al.* A comparison of frailty indexes for prediction of adverse health outcomes in an elderly cohort. *Arch Gerontol Geriatr* 2012; 54: 16–20.
- Ensrud KE, Ewing SK, Taylor BC *et al.* Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women. *Arch Intern Med* 2008; 168: 382–9.
- Woo J, Leung J, Morley JE. Comparison of frailty indicators based on clinical phenotype and the multiple deficit approach in predicting mortality and physical limitation. *J Am Geriatr Soc* 2012; 60: 1478–86.
- Daniels R, van Rossum E, Beurskens A *et al.* The predictive validity of three self-report screening instruments for identifying frail older people in the community. *BMC Public Health* 2012; 12: 69.
- Pijpers E, Ferreira I, Stehouwer CDA, Nieuwenhuijzen Kruseman AC. The frailty dilemma. Review of the predictive accuracy of major frailty scores. *Eur J Int Med* 2012; 23: 118–23.
- Flinders Centre for Ageing Studies. ALSA: The Australian Longitudinal Study on Ageing, 2011. [http://flinders.edu.au/sabs/fcas/alsa/alsa\\_home.cfm](http://flinders.edu.au/sabs/fcas/alsa/alsa_home.cfm) (7 April 2011, date last accessed).
- Kiely DK, Cupples LA, Lipsitz LA. Validation and comparison of two frailty indexes: the MOBILIZE Boston Study. *J Am Geriatr Soc* 2009; 57: 1532–9.
- Mitnitski A, Song X, Skoog I *et al.* Relative fitness and frailty of elderly men and women in developed countries and their relationship with mortality. *J Am Geriatr Soc* 2005; 53: 2184–9.
- Ravaglia G, Forti P, Lucicesare A *et al.* Development of an easy prognostic score for frailty outcomes in the aged. *Age Ageing* 2008; 37: 161–6.
- Mitnitski A, Mogilner A, Rockwood K. Accumulation of deficits as a proxy measure of aging. *Sci World* 2001; 1: 323–36.
- Rockwood K, Melissa A, Mitnitski A. A comparison of two approaches to measuring frailty in elderly people. *J Gerontol* 2007; 62A: 738.
- Theou O, Rockwood MRH, Mitnitski A, Rockwood K. Disability and co-morbidity in relation to frailty: how much do they overlap? *Arch Gerontol Geriatr* 2012; 55: e1–8.
- Bossuyt PM, Reitsma JB, E Bruns D *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Clin Chem Lab Med* 2003; 41: 68–73.
- Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. The Evidence Base of Clinical Diagnosis. West Sussex, UK: Wiley-Blackwell, 2009; 42–62.
- Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. *BMJ* 2009; 338: b604.
- Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009; 338: b606.
- Akobeng AK. Understanding diagnostic tests 2: likelihood ratios, pre- and post-test probabilities and their use in clinical practice. *Acta Paediatrica* 2007; 96: 487–91.
- Cronbach LJ, Meehl PE. Construct validity in psychological test. *Psychol Bull* 1955; 52: 281–302.

Received 20 August 2014; accepted in revised form 13 August 2015