Rockwood index outperforms patient age as a predictor of poor patient outcomes.

Alex Bokov 1,✉, Sara Espinoza 1, Chandana Tripathy 1, …, …, …, and Kathleen R. Stevens 1

2020-08-24

We have found that the electronic frailty index (EFI), a risk score developed using the Rockwood deficit-accumulation framework, is a strong predictor of falls, hospitalizations, hospital-acquired infections, and loss of independence without relying on any predictors including patient age. In fact, EFI outperformed patient age as a predictor. EFI’s accuracy was not negated by the fact that we used patients of all ages, not just older patients.

# Introduction

Frailty is the lifelong erosion of stress resistance and accumulation of impairments across multiple physiological systems. Among older community dwelling adults 32% have been classified as pre-frail and 24% have been classified as frail (Hoover *et al.*, 2013). Frailty predicts disability, injurious falls, and mortality (Pajewski *et al.*, 2019), emergency room visits and hospitalizations (Fried *et al.*, 2001), and long-term care admissions (Rockwood, 2005; Rockwood *et al.*, 2006; Pajewski *et al.*, 2019). The Fried phenotype (Fried *et al.*, 2001) and Rockwood deficit accumulation index (Mitnitski, Mogilner and Rockwood, 2001; Rockwood and Mitnitski, 2007), are the most commonly used methods for operationalizing frailty. There is reasonable convergence between these two approaches (Malmstrom, Miller and Morley, 2014; Li *et al.*, 2015) but the deficit accumulation approach does not require individual questionnaires nor physical assessments yet offers similar (Malmstrom, Miller and Morley, 2014) or possibly better (Kulminski *et al.*, 2008) predictive accuracy.

Possible points still left to cover:

* outcomes important to clinical care of older patients
* how care could be improved with better modeling of these outcomes
* information about our site

# Methods

## Population

A random 1% sample (N=14,844) was drawn from the deidentified patient records of a large academic health center and its teaching hospital partner. Visits during which patient age was less than 18 years old were excluded and then patients who had fewer than three visits in the remaining data were excluded. To avoid bias, for each patient an index visit date was chosen and only data recorded on or after that date was used in analysis. To avoid distorted results in patients with sparse visit histories, those who had fewer than two visit-dates after index visit assignment were removed from the sample as were patients whose EFI was never higher than 0. Finally, the patients were randomly assigned to a develpoment cohort (N=2,497 patients, 52,372 visit-days) or a testing cohort (N=3,220 patients, 56,320 visit-days). Sensitivity analysis was done to see the effect of leaving in data from all visits by adult patients and the overall direction of EFI’s effect was the same but [the performance improvement relative to patient age was inflated]. All decisions about data processing and statistical analysis were made using only the development cohort and blinded to the testing cohort. For publication, the same analysis scripts were run on the testing cohort and used to create all results reported here (the development version of each table and figure is available in the supplemental materials). The baseline characteristics of the testing cohort are shown in Table 1.

Table 1: Cohort demographics

|  | Nonfrail, < 0.1 (N=1876) | Prefrail, 0.1 - 0.19 (N=398) | Frail, > 0.2 (N=223) | Overall (N=2497) |
| --- | --- | --- | --- | --- |
| **Sex** |  |  |  |  |
| f | 1116 (59.5%) | 209 (52.5%) | 125 (56.1%) | 1450 (58.1%) |
| m | 759 (40.5%) | 189 (47.5%) | 98 (43.9%) | 1046 (41.9%) |
| u | 1 (0.1%) | 0 (0%) | 0 (0%) | 1 (0.0%) |
| **Race** |  |  |  |  |
| Asian | 37 (2.0%) | 11 (2.8%) | 5 (2.2%) | 53 (2.1%) |
| Black | 140 (7.5%) | 27 (6.8%) | 11 (4.9%) | 178 (7.1%) |
| Other | 135 (7.2%) | 10 (2.5%) | 7 (3.1%) | 152 (6.1%) |
| Unknown | 134 (7.1%) | 13 (3.3%) | 2 (0.9%) | 149 (6.0%) |
| White | 1425 (76.0%) | 337 (84.7%) | 198 (88.8%) | 1960 (78.5%) |
| Missing | 5 (0.3%) | 0 (0%) | 0 (0%) | 5 (0.2%) |
| **Hospitalization** |  |  |  |  |
| Yes | 206 (11.0%) | 44 (11.1%) | 88 (39.5%) | 338 (13.5%) |
| No | 1670 (89.0%) | 354 (88.9%) | 135 (60.5%) | 2159 (86.5%) |
| **Discharge to ICF or SNF** |  |  |  |  |
| Yes | 10 (0.5%) | 2 (0.5%) | 6 (2.7%) | 18 (0.7%) |
| No | 1866 (99.5%) | 396 (99.5%) | 217 (97.3%) | 2479 (99.3%) |
| **Deceased** |  |  |  |  |
| Yes | 46 (2.5%) | 9 (2.3%) | 6 (2.7%) | 61 (2.4%) |
| No | 1830 (97.5%) | 389 (97.7%) | 217 (97.3%) | 2436 (97.6%) |
| **Hospital Acquired Infections** |  |  |  |  |
| Yes | 70 (3.7%) | 44 (11.1%) | 48 (21.5%) | 162 (6.5%) |
| No | 1806 (96.3%) | 354 (88.9%) | 175 (78.5%) | 2335 (93.5%) |
| **Falls** |  |  |  |  |
| Yes | 49 (2.6%) | 19 (4.8%) | 26 (11.7%) | 94 (3.8%) |
| No | 1827 (97.4%) | 379 (95.2%) | 197 (88.3%) | 2403 (96.2%) |
| **Patient age (years)** |  |  |  |  |
| Mean (SD) | 51.0 (18.1) | 51.2 (16.0) | 56.2 (16.3) | 51.5 (17.7) |
| Median [Min, Max] | 48.6 [18.3, 88.5] | 51.4 [18.4, 87.7] | 58.6 [21.4, 87.8] | 50.3 [18.3, 88.5] |
| **Frailty** |  |  |  |  |
| Mean (SD) | 0.0194 (0.0275) | 0.140 (0.0298) | 0.286 (0.0704) | 0.0623 (0.0891) |
| Median [Min, Max] | 0 [0, 0.0833] | 0.146 [0.104, 0.188] | 0.271 [0.208, 0.563] | 0.0208 [0, 0.563] |
| **Number of visit-days** |  |  |  |  |
| Mean (SD) | 14.5 (34.4) | 26.3 (36.5) | 66.1 (119) | 21.0 (50.8) |
| Median [Min, Max] | 7.00 [2.00, 1170] | 13.0 [2.00, 318] | 31.0 [2.00, 1180] | 8.00 [2.00, 1180] |
| **Length of Stay** |  |  |  |  |
| Mean (SD) | 4.89 (3.91) | 3.47 (2.09) | 4.47 (3.17) | 4.59 (3.56) |
| Median [Min, Max] | 3.75 [1.00, 31.0] | 3.00 [1.00, 11.0] | 4.00 [1.00, 21.0] | 3.50 [1.00, 31.0] |
| Missing | 1670 (89.0%) | 354 (88.9%) | 135 (60.5%) | 2159 (86.5%) |

## Electronic Frailty Index

Clegg *et al.* (2016) developed an electronic frailty index (EFI) for UK health systems following the methodology of Mitnitsky, Rockwood, et al. (Mitnitski, Mogilner and Rockwood, 2001; Searle *et al.*, 2008; Song, Mitnitski and Rockwood, 2010). Recently, Pajewski *et al.* (2019) adapted the EFI to use ICD10 and ICD9. We further built on this work by also mapping laboratory tests to specific LOINC codes. For each patient visit, all distinct diagnoses and abnormal lab results over the preceding two-year window for that patient were aggregated into a single EFI, a numeric value that can range from 0 to 1 (but in practice seldom exceeded 0.6). As mentioned above, we omitted all visits prior to a randomly select index visit for each patient. This did not interfere with EFI calculation because those EFI values were calculated separately for every distinct patient-date in our health system, and then joined to the EHR data.

## Outcomes

The primary outcomes we predicted with EFI were falls (any ICD9 codes in E880-E888.9, E987.x, V00.x, or ICD10 codes in W00-W19), hospital admissions (EHR encounter type = ‘inpatient’), hospital-acquired infections (having any ICD10 code associated with hospital-acquired infections per Southern et al. Southern *et al.* (2017) ), discharge to intermediate care facilities (ICF) or skilled nursing facilities (SNF) (using discharge disposition codes, for patients whose admit-source for that encounter was ‘Home’), and all-cause mortality insofar as it can be ascertained from the EHR discharge disposition or vital status (which, in the absence of linkage to external death indexes are likely to under-report the actual mortality and so should be interpreted with caution).

## Statistical Analysis

For each outcome of interest, we used a Cox proportional hazard model to estimate the risk of the first occurrences of the outcome after the patients’ respective index visits using EFI as the predictor. Unlike earlier studies, we treated EFI as a time-varying numeric predictor with multiple followups per patient. Since only the first occurrence was being predicted, a recurring event model was not necessary. As a comparison, for each outcome a second analysis was done that was identical to the first except that patient age at visit (in days) was used as the predictor.

# Results

Table 2 shows the results of Cox proportional hazard models for each of the responses, with EFI as the predictor. For each 0.1 increase in EFI, we found at least a doubling of risk: 2.2-fold for hospitalization, 2.7-fold for discharge to ICF or SNF after having been admitted from home, 2.3-fold for deceased, 2.3-fold for hospital acquired infections, and 2.4-fold for falls. Furthermore, the probability of delayed discharge increased and thus longer stays were observed in patients with an EFI > 0.19. The p-values shown have been adjusted for multiple comparisons (six outcomes reported in one study) using the Holm (1979) method and in all cases are highly significant. It is well established in gerontology that an enormous and diverse range of poor health outcomes are correlated with an individual’s age, but chronological age provides no information about the progression of biological age in an individual. It has long been a priority in aging research to find a metric for aging that is subject to individual variation and intervention. Rockford et al. have pointed out {ref} that deficit accumulation has the hallmarks of such a measure.

Table 2: Cox-proportional hazards with EFI as a predictor

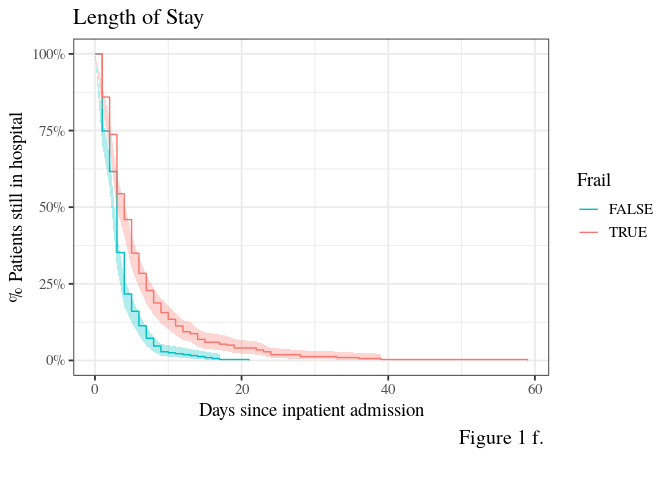
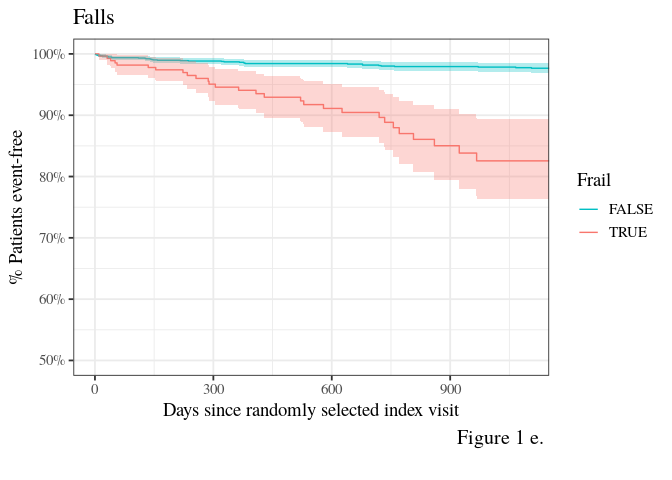
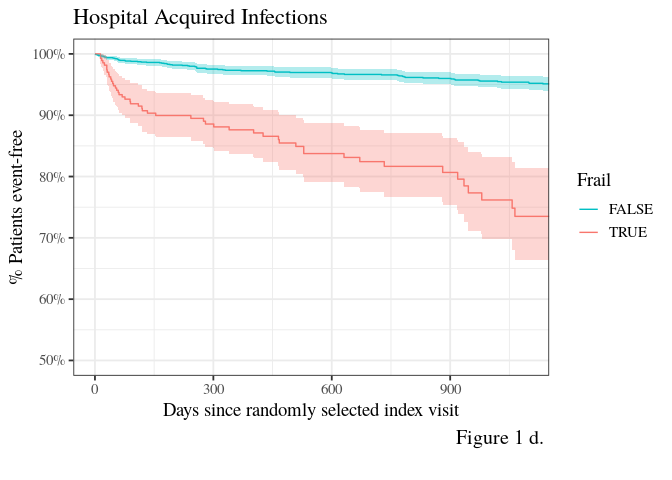
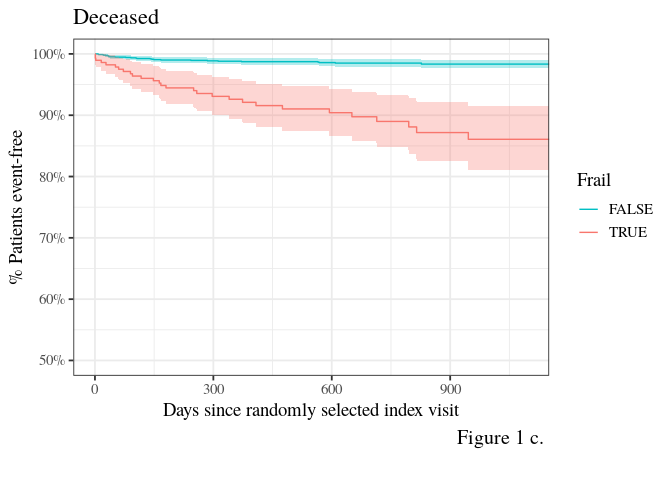
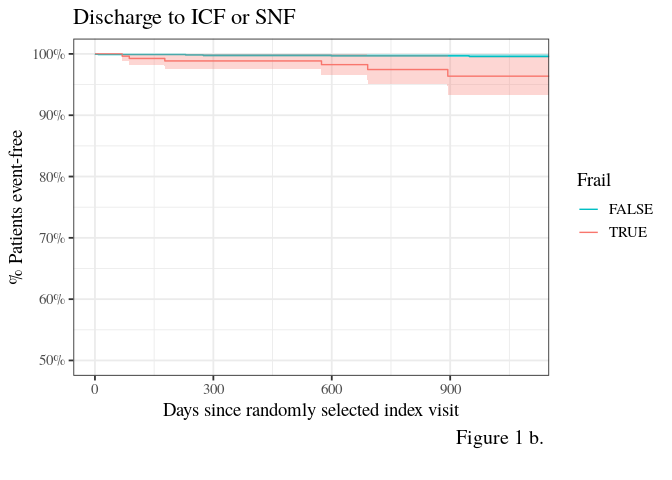
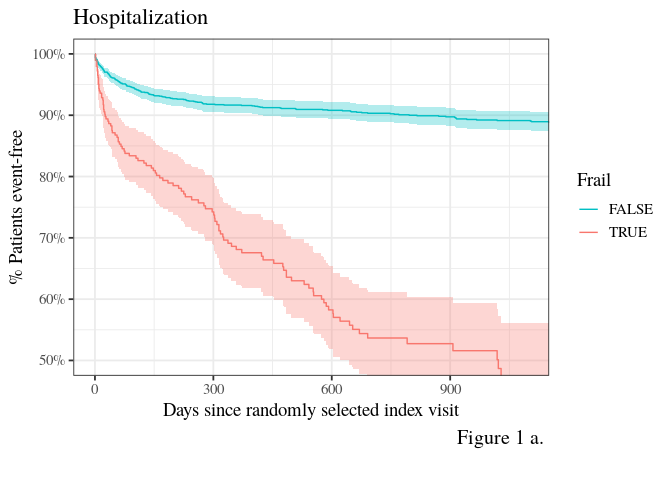
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Outcome | β^ (95% CI) | Fold-change (95% CI) | SE | Z | P |
| Hospitalization | 0.80 (0.72, 0.88) | 2.23 (2.06, 2.42) | 0.0412 | 19.5 | 8.93e-84 |
| Discharge to ICF or SNF | 0.99 (0.72, 1.27) | 2.70 (2.05, 3.57) | 0.142 | 7 | 2.52e-12 |
| Deceased | 0.85 (0.68, 1.02) | 2.34 (1.98, 2.77) | 0.0854 | 9.99 | 5.32e-23 |
| Hospital Infections | 0.85 (0.74, 0.95) | 2.34 (2.10, 2.60) | 0.0537 | 15.8 | 1.49e-55 |
| Falls | 0.88 (0.74, 1.01) | 2.40 (2.10, 2.75) | 0.0679 | 12.9 | 1.52e-37 |
| Length of Stay | -0.25 (-0.31, -0.19) | 0.78 (0.73, 0.83) | 0.0321 | -7.73 | 2.09e-14 |

Therefore, we compared the predictive accuracy of the EFI to that of patient age. For each outcome, we fit an additional Cox proportional hazard model, this time using age at visit as the predictor variable. It appears, at least in this population, that not only can EFI accurately predict mortality and poor patient outcomes without any knowledge of patient age but it does so better than patient age.

Table 3: Comparing EFI and age as predictors

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Outcome | Predictor | Concordance | Log Likelihood | AIC |
| Hospitalization | Frailty | 0.7219 | -2337 | 4675 |
|  | Patient Age | 0.5501 | -2476 | 4954 |
| Discharge to ICF or SNF | Frailty | 0.86 | -107.8 | 217.6 |
|  | Patient Age | 0.7855 | -117.1 | 236.2 |
| Deceased | Frailty | 0.7841 | -414.1 | 830.2 |
|  | Patient Age | 0.7408 | -432.8 | 867.5 |
| Hospital Acquired Infections | Frailty | 0.7801 | -1072 | 2146 |
|  | Patient Age | 0.5515 | -1162 | 2326 |
| Falls | Frailty | 0.8 | -608.7 | 1219 |
|  | Patient Age | 0.5816 | -664.6 | 1331 |
| Length of Stay | Frailty | 0.6001 | -3456 | 6914 |
|  | Patient Age | 0.5341 | -3486 | 6975 |

In Table 3 we compare the performance of EFI and age as predictors for each of the six outcomes. In all cases, the EFI models has a robustly lower AIC and log-likelihood than the respective patient-age models, as well as a higher concordance (in all cases greater than 0.7). Cox models incorporating EFI together with age, with or without an interaction term, did not have a significantly better fit than the univariate EFI model (not shown, available in supplement).



In figures 1 a-f we show Kaplan-Meier plots for each of the outcomes stratified by whether EFI is greater than 0.19 (Frail=TRUE) or less (Frail=FALSE) (Stow *et al.*, 2018).

# Discussion

Our data shares the fundamental limitation of the EHR system from which it was obtained: like all EHR systems, it only has information that providers and coders put into it. Events taking place outside the health system or at un-connected health systems are not visible to our analysis. On the other hand, providers who rely on EHR systems at point of care are also working under these limitations on. The data we used is representative of this scenario, and despite the limitations EFI provides accurate predictions of poor patient outcomes. Because our implementation of the Rockwood index real EHR data, it is more directly transferable to clinical use than implementations based on curated registries. This suggests that EFI is most accurate for patients who have accumulated a reasonable in-system visit-history. Further work is needed to find a more precise relationship between the length of a patient’s visit-history and the accuracy of EFI and better distinguish genuinely non-frail patients from those who get most of their care outside the researchers’ health system.

Vital status is often out of date in EHR systems if the patient did not die in the hospital or shortly after their visit, so our data may under-represent the true mortality rates. Nevertheless, the relationship between EFI and mortality risk observed here agrees with previous EFI studies. Unlike previous studies we sampled all adults rather than just older adults so we could assess the performance of EFI across the lifespan and include individuals who might have early-onset frailty.

There is no gold standard method to assess Frailty in clinical practice. Currently available frailty assessment tools used in geriatric practice have good validity ( example Linda Fried’s phenotype method) but these are time intensive and often difficult to implement in a busy general practice. Assessing frailty helps clinicians identify high risk patients and tailor interventions to prevent health decline and poor outcomes. Because of the simplicity of the Rockwood Frailty Index, it is more likely to be adopted by clinicians in a busy practice.

Other possible points to cover:

* What we learned by not restricting the sample to older adults (or move to methods/theory paper?)
* What agrees and what disagrees with previous work?
* What is new?
  + Per-visit EFI allows plotting an EFI trajectory over time
  + Strictly limiting data elements to those we can count on being available in any EHR system, at any site
* future work
* theoretical and practical implications

# Conclusions

TBD

# Acknowledgments

This work was supported by NIH/NCATS UL1TR001120 (AFB, KRS), the Long School of Medicine KL2 Award (AFB), …

# References

Clegg, A. *et al.* (2016) ‘Development and validation of an electronic frailty index using routine primary care electronic health record data’, *Age and Ageing*, 45(3), pp. 353–360. doi: [10.1093/ageing/afw039](https://doi.org/10.1093/ageing/afw039).

Fried, L. P. *et al.* (2001) ‘Frailty in older adults: Evidence for a phenotype’, *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3), pp. M146–M157. doi: [10.1093/gerona/56.3.M146](https://doi.org/10.1093/gerona/56.3.M146).

Holm, S. (1979) ‘A simple sequentially rejective multiple test procedure.’, *Scandinavian Journal of Statistics*, 6, pp. 65–70.

Hoover, M. *et al.* (2013) ‘Validation of an index to estimate the prevalence of frailty among community-dwelling seniors’, *Health Reports*, 24(9), pp. 10–17.

Kulminski, A. M. *et al.* (2008) ‘Cumulative Deficits Better Characterize Susceptibility to Death in Elderly People than Phenotypic Frailty: Lessons from the Cardiovascular Health Study: FRAILTY, CUMULATIVE DEFICITS, AND SURVIVAL’, *Journal of the American Geriatrics Society*, 56(5), pp. 898–903. doi: [10.1111/j.1532-5415.2008.01656.x](https://doi.org/10.1111/j.1532-5415.2008.01656.x).

Li, G. *et al.* (2015) ‘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*. Edited by A. M. Chamberlain, 10(3), p. e0120144. doi: [10.1371/journal.pone.0120144](https://doi.org/10.1371/journal.pone.0120144).

Malmstrom, T. K., Miller, D. K. and Morley, J. E. (2014) ‘A Comparison of Four Frailty Models’, *Journal of the American Geriatrics Society*, 62(4), pp. 721–726. doi: [10.1111/jgs.12735](https://doi.org/10.1111/jgs.12735).

Mitnitski, A. B., Mogilner, A. J. and Rockwood, K. (2001) ‘Accumulation of Deficits as a Proxy Measure of Aging’, *The Scientific World JOURNAL*, 1, pp. 323–336. doi: [10.1100/tsw.2001.58](https://doi.org/10.1100/tsw.2001.58).

Pajewski, N. M. *et al.* (2019) ‘Frailty Screening Using the Electronic Health Record Within a Medicare Accountable Care Organization’, *The Journals of Gerontology: Series A*. Edited by A. Newman, 74(11), pp. 1771–1777. doi: [10.1093/gerona/glz017](https://doi.org/10.1093/gerona/glz017).

Rockwood, K. (2005) ‘A global clinical measure of fitness and frailty in elderly people’, *Canadian Medical Association Journal*, 173(5), pp. 489–495. doi: [10.1503/cmaj.050051](https://doi.org/10.1503/cmaj.050051).

Rockwood, K. and Mitnitski, A. (2007) ‘Frailty in Relation to the Accumulation of Deficits’, *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 62(7), pp. 722–727. doi: [10.1093/gerona/62.7.722](https://doi.org/10.1093/gerona/62.7.722).

Rockwood, K. *et al.* (2006) ‘Long-Term Risks of Death and Institutionalization of Elderly People in Relation to Deficit Accumulation at Age 70: LONG-TERM RISK OF DEATH DEFINED BY AGE 70’, *Journal of the American Geriatrics Society*, 54(6), pp. 975–979. doi: [10.1111/j.1532-5415.2006.00738.x](https://doi.org/10.1111/j.1532-5415.2006.00738.x).

Searle, S. D. *et al.* (2008) ‘A standard procedure for creating a frailty index’, *BMC Geriatrics*, 8(1). doi: [10.1186/1471-2318-8-24](https://doi.org/10.1186/1471-2318-8-24).

Song, X., Mitnitski, A. and Rockwood, K. (2010) ‘Prevalence and 10-year outcomes of frailty in older adults in relation to deficit accumulation’, *Journal of the American Geriatrics Society*, 58(4), pp. 681–687. doi: [10.1111/j.1532-5415.2010.02764.x](https://doi.org/10.1111/j.1532-5415.2010.02764.x).

Southern, D. A. *et al.* (2017) ‘Deriving ICD-10 Codes for Patient Safety Indicators for Large-scale Surveillance Using Administrative Hospital Data:’ *Medical Care*, 55(3), pp. 252–260. doi: [10.1097/MLR.0000000000000649](https://doi.org/10.1097/MLR.0000000000000649).

Stow, D. *et al.* (2018) ‘Evaluating frailty scores to predict mortality in older adults using data from population based electronic health records: Case control study’, *Age and Ageing*, 47(4), pp. 564–569. doi: [10.1093/ageing/afy022](https://doi.org/10.1093/ageing/afy022).