Rockwood index as a predictor of patient safety events and readmissions.

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

2020-08-26

We have found that the electronic frailty index (EFI), a risk score developed using the Rockwood deficit-accumulation framework, is a strong predictor of ED visits, 30-day readmissions, healthcare-associated non-procedural trauma, cardiac complications, adverse drug events, venous thromboembolism, fluid management complications, as well as patient safety events in general and severe patient safety events.

# 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 left to cover:’

* frailty measures as an emerging best practice?
* Do we have any remarks about how quality improvement relates to implementation science?
* outcomes important for improving quality and reducing costs
* how care could be improved with better modeling of these outcomes
* information about our site

# Methods

## Population

Yes, I realize at the moment this is identical to the corresponding section in the aging paper. The phrasing will naturally diverge from the aging paper over subsequent edits. For now I left it all here because these facts are relevant to this paper as well and I didn’t want to lose any of them.

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 development 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.

[1] "

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%)

Emergency Department

Yes

339 (18.1%)

91 (22.9%)

77 (34.5%)

507 (20.3%)

No

1537 (81.9%)

307 (77.1%)

146 (65.5%)

1990 (79.7%)

Readmission Within 30 Days of a Previous Discharge

Yes

45 (2.4%)

6 (1.5%)

31 (13.9%)

82 (3.3%)

No

1831 (97.6%)

392 (98.5%)

192 (86.1%)

2415 (96.7%)

Hospital Acquired Trauma

Yes

72 (3.8%)

31 (7.8%)

42 (18.8%)

145 (5.8%)

No

1804 (96.2%)

367 (92.2%)

181 (81.2%)

2352 (94.2%)

Cardiac Complications

Yes

50 (2.7%)

17 (4.3%)

41 (18.4%)

108 (4.3%)

No

1826 (97.3%)

381 (95.7%)

182 (81.6%)

2389 (95.7%)

Drug Adverse Event

Yes

6 (0.3%)

8 (2.0%)

15 (6.7%)

29 (1.2%)

No

1870 (99.7%)

390 (98.0%)

208 (93.3%)

2468 (98.8%)

Venous Thromboembolism

Yes

3 (0.2%)

8 (2.0%)

9 (4.0%)

20 (0.8%)

No

1873 (99.8%)

390 (98.0%)

214 (96.0%)

2477 (99.2%)

Fluid Management Events

Yes

9 (0.5%)

4 (1.0%)

12 (5.4%)

25 (1.0%)

No

1867 (99.5%)

394 (99.0%)

211 (94.6%)

2472 (99.0%)

CNS Complications

Yes

14 (0.7%)

3 (0.8%)

8 (3.6%)

25 (1.0%)

No

1862 (99.3%)

395 (99.2%)

215 (96.4%)

2472 (99.0%)

GI Complications

Yes

5 (0.3%)

6 (1.5%)

7 (3.1%)

18 (0.7%)

No

1871 (99.7%)

392 (98.5%)

216 (96.9%)

2479 (99.3%)

Any Patient Safety Event

Yes

191 (10.2%)

99 (24.9%)

114 (51.1%)

404 (16.2%)

No

1685 (89.8%)

299 (75.1%)

109 (48.9%)

2093 (83.8%)

Severe Patient Safety Event

Yes

18 (1.0%)

5 (1.3%)

12 (5.4%)

35 (1.4%)

No

1858 (99.0%)

393 (98.7%)

211 (94.6%)

2462 (98.6%)

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 Visits

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%)

"

Table: Table 1: Cohort demographics

## 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).

Points that were omitted or briefly covered in aging paper and might be appropriate here (probably more diplomatically phrased though):

* EFI has been used in the UK for several years (Clegg *et al.*, 2016), and has only recently been adapted to a US EHR system (Pajewski *et al.*, 2019).
* Our version has fewer EHR-specific dependencies
* Ours is easier to maintain because it uses value-flags instead of manually curated threshold values
* Our EFI is longitudinal, compare our results to those of Stow, Matthews and Hanratty (2018)
* The importance of a “non-greedy” software license: not only lower costs but also greater transparency and makes it easier for stakeholders to contribute improvements.

## Outcomes

The primary outcomes we predicted with EFI were ED visits, readmission within 30 days of a prior discharge from an inpatient stay, non-procedural hospital trauma, cardiac complications, adverse drug events, venous thromboembolisms, fluid management events, CNS complications, and GI complications. All outcomes except ED visits and readmissions were defined as visits during which at least one of the ICD10 codes from the corresponding groups published by Southern *et al.* (2017) were recorded. Of the other 10 PSIs defined by Southern *et al.* (2017), 2 were excluded because pregnancies were outside the scope of our protocol and the remaining 8 had prevalences too low to reliably analyze in this sample size. However, all the diagnoses from Southern *et al.* (2017) except maternal and infant complications contributed to the ‘any patient safety event’ outcome. A subset of these were aggregated into the ‘severe patient safety’ event defined as “proximally threatening to life or to major vital organs” (Southern *et al.*, 2017).

## 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 a time-varying predictor. For each outcome we compared the predictive performance of EFI to that patient age at visit.

# 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: 1.8-fold for emergency department, 3-fold for readmission within 30 days of a previous discharge, 2.2-fold for hospital acquired trauma, 2.8-fold for cardiac complications, 2.6-fold for drug adverse event, 2.8-fold for venous thromboembolism, 3-fold for fluid management events, 2-fold for cns complications, 2.8-fold for gi complications, 2.4-fold for any patient safety event, and 2.4-fold for severe patient safety event. 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 (twelve outcomes reported in this study) using the Holm (1979) method and remain significant.

Older patients trend toward frailty and one might ask whether EFI merely reflect patient age. To test this we compared the predictive accuracy of the EFI to that of patient age. For each outcome, we fit an additional Cox proportional hazard model using age at visit as the predictor instead of EFI. We found that EFI predicts all outcomes better than patient age. Models using both patient age and EFI offered no significant improvement over EFI alone [not shown?].

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

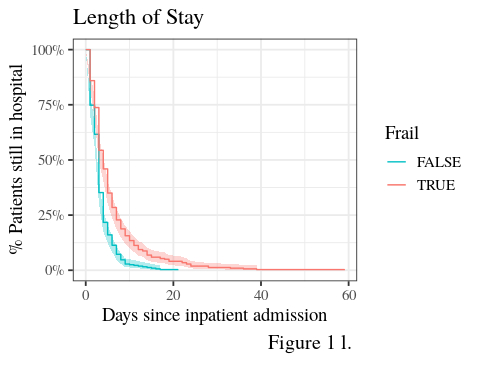
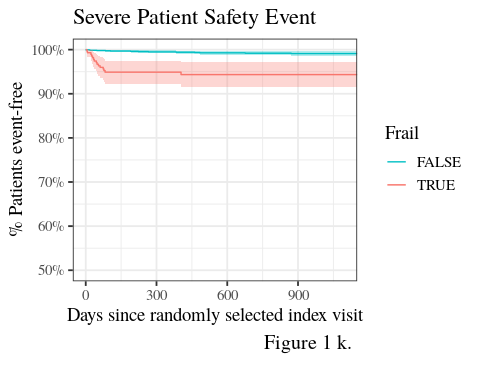
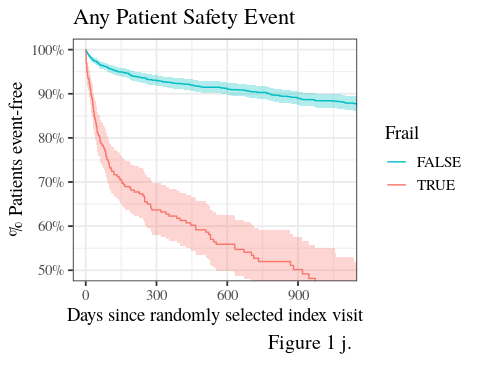
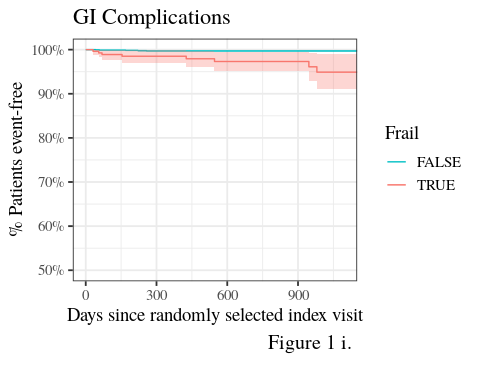
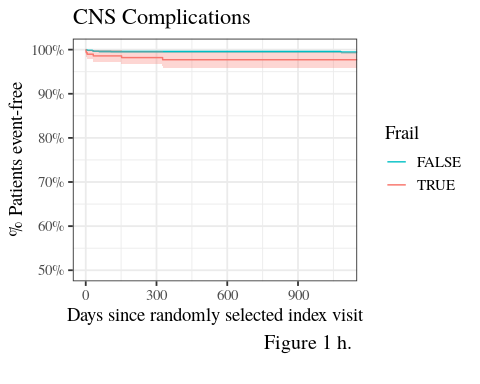
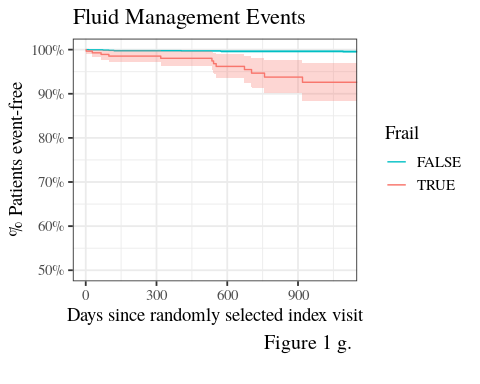
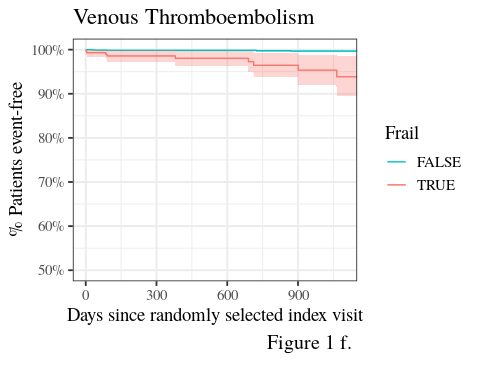
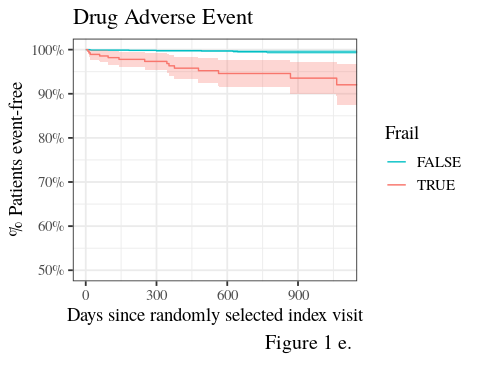
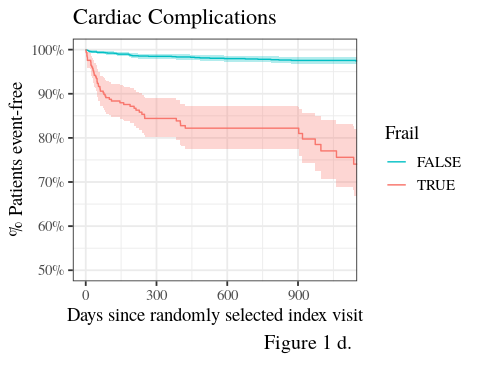
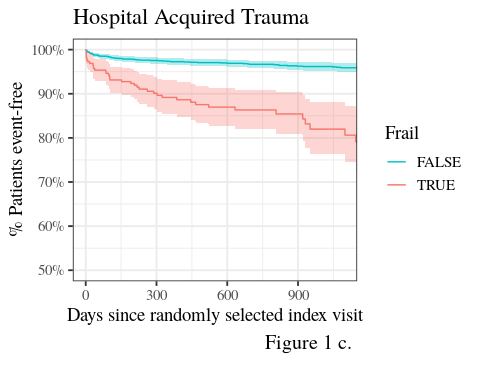
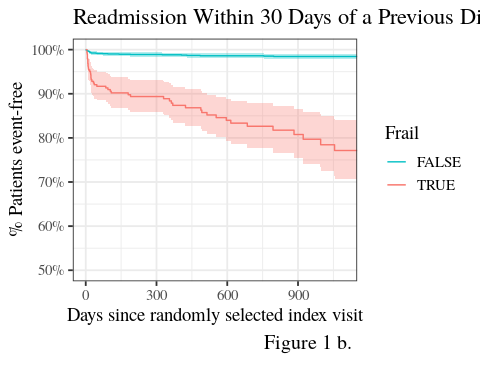
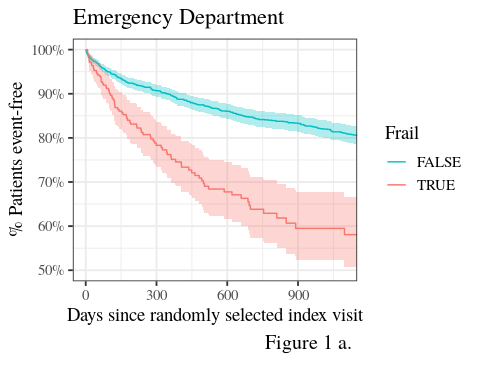
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Outcome | β^ (95% CI) | fold-change (95% CI) | SE | Z | P |
| Emergency Department | 0.59 (0.51, 0.66) | 1.80 (1.67, 1.93) | 0.0373 | 15.7 | 1.54e-54 |
| Readmission Within 30 Days of a Previous Discharge | 1.08 (0.94, 1.22) | 2.95 (2.56, 3.40) | 0.0725 | 14.9 | 2.26e-49 |
| Hospital Acquired Trauma | 0.77 (0.65, 0.88) | 2.15 (1.92, 2.42) | 0.0596 | 12.9 | 5.08e-37 |
| Cardiac Complications | 1.01 (0.89, 1.14) | 2.75 (2.43, 3.13) | 0.0649 | 15.6 | 6.76e-54 |
| Drug Adverse Event | 0.95 (0.72, 1.18) | 2.58 (2.06, 3.24) | 0.116 | 8.18 | 1.69e-15 |
| Venous Thromboembolism | 1.03 (0.76, 1.31) | 2.81 (2.14, 3.69) | 0.139 | 7.46 | 2.59e-13 |
| Fluid Management Events | 1.08 (0.84, 1.33) | 2.95 (2.31, 3.78) | 0.126 | 8.63 | 4.48e-17 |
| CNS Complications | 0.71 (0.42, 1.01) | 2.04 (1.52, 2.74) | 0.15 | 4.75 | 2.06e-06 |
| GI Complications | 1.01 (0.71, 1.31) | 2.75 (2.04, 3.70) | 0.152 | 6.67 | 5.21e-11 |
| Any Patient Safety Event | 0.87 (0.80, 0.94) | 2.38 (2.22, 2.56) | 0.0368 | 23.6 | 6.85e-122 |
| Severe Patient Safety Event | 0.88 (0.65, 1.11) | 2.40 (1.91, 3.02) | 0.117 | 7.51 | 2.39e-13 |
| Length of Stay | -0.25 (-0.31, -0.19) | 0.78 (0.73, 0.83) | 0.0321 | -7.73 | 5.23e-14 |

Table 3: Comparing EFI and age as predictors

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Outcome | Predictor | Concordance | Log Likelihood | AIC |
| Emergency Department | Frailty | 0.681 | -3496 | 6994 |
|  | Patient Age | 0.6308 | -3539 | 7081 |
| Readmission Within 30 Days of a Previous Discharge | Frailty | 0.8546 | -524 | 1050 |
|  | Patient Age | 0.5713 | -608.4 | 1219 |
| Hospital Acquired Trauma | Frailty | 0.7629 | -980.2 | 1962 |
|  | Patient Age | 0.4912 | -1041 | 2085 |
| Cardiac Complications | Frailty | 0.8293 | -698.8 | 1400 |
|  | Patient Age | 0.7222 | -763.5 | 1529 |
| Drug Adverse Event | Frailty | 0.8503 | -182.2 | 366.4 |
|  | Patient Age | 0.6226 | -205.8 | 413.6 |
| Venous Thromboembolism | Frailty | 0.8536 | -123 | 248 |
|  | Patient Age | 0.6227 | -142.3 | 286.7 |
| Fluid Management Events | Frailty | 0.8193 | -146.1 | 294.3 |
|  | Patient Age | 0.5993 | -174.2 | 350.4 |
| CNS Complications | Frailty | 0.7017 | -170.6 | 343.3 |
|  | Patient Age | 0.5701 | -178.1 | 358.2 |
| GI Complications | Frailty | 0.7792 | -113.3 | 228.7 |
|  | Patient Age | 0.5049 | -130 | 262 |
| Any Patient Safety Event | Frailty | 0.7841 | -2705 | 5412 |
|  | Patient Age | 0.5806 | -2900 | 5803 |
| Severe Patient Safety Event | Frailty | 0.7781 | -237.3 | 476.6 |
|  | Patient Age | 0.6937 | -252.9 | 507.8 |
| 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 twelve 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).

In figures 1 a-k 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). [Let’s talk about whether to put length of stay here or aging paper]



# 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.

Assessing frailty helps clinicians identify high risk patients and tailor interventions to prevent health decline and poor outcomes. Currently available frailty assessment tools used in geriatric practice have good validity (for example Fried *et al.* (2001)) but these are time intensive and often difficult to implement in a busy general practice. Because EFI can be calculated automatically it is more likely to be adopted by clinicians in a busy practice. Also, EFI can be calculated retroactively on historic records of patients who never received in-person assessments, giving a clearer picture of a health system’s overall performance. Furthermore, our version of the EFI algorithm can be made vendor-independent since it relies only on diagnoses, medications, lab codes, and vitals which are data in modern EHR systems. This, together with the fact that this algorithm is publicly available under an open-source license, is intended to facilitate adoption and enhancement by diverse health systems.

Other possible points to cover:

* Tie-in to implementation science
* What agrees and what disagrees with previous work but from a more quantitative point of view than in the aging paper: distributions and prevalences compared to Southern *et al.* (2017), Clegg *et al.* (2016), Stow *et al.* (2018), and Stow, Matthews and Hanratty (2018) (for trajectories, maybe).
* Possibly examples from other industries of how open source software encourages collaboration and innovation, and that healthcare field lags in understanding how to leverage open source.
* 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).

Stow, D., Matthews, F. E. and Hanratty, B. (2018) ‘Frailty trajectories to identify end of life: A longitudinal population-based study’, *BMC Medicine*, 16(1), p. 171. doi: [10.1186/s12916-018-1148-x](https://doi.org/10.1186/s12916-018-1148-x).