SHORT REPORT

A classification tree to assist with routine scoring of the Clinical Frailty Scale

Olga Theou[1,2,3], Mario Ulises Pérez-Zepeda[2,3], Alexandra M. van der Valk[3], Samuel D. Searle[3], Susan E. Howlett[2,4], Kenneth Rockwood[2,3]

[1]School of Physiotherapy, Dalhousie University, Halifax, NS, Canada
[2]Geriatric Medicine, Dalhousie University, Halifax, NS, Canada
[3]Geriatric Medicine, Nova Scotia Health, Halifax, NS, Canada
[4]Pharmacology, Dalhousie University, Halifax, NS, Canada

Address correspondence to: Kenneth Rockwood, Division of Geriatric Medicine, Dalhousie University and Nova Scotia Health, Camp Hill Veterans' Memorial Bldg, 5955 Veterans' Memorial Lane, Halifax, NS B3H 2E1, Canada.
Tel: +1(902)473-8687; Fax: +1(902)473-1050. Email: Kenneth.Rockwood@dal.ca

Abstract

Background: the Clinical Frailty Scale (CFS) was originally developed to summarise a Comprehensive Geriatric Assessment and yield a care plan. Especially since COVID-19, the CFS is being used widely by health care professionals without training in frailty care as a resource allocation tool and for care rationing. CFS scoring by inexperienced raters might not always reflect expert judgement. For these raters, we developed a new classification tree to assist with routine CFS scoring. Here, we test that tree against clinical scoring.
Objective/Methods: we examined agreement between the CFS classification tree and CFS scoring by novice raters (clerks/residents), and the CFS classification tree and CFS scoring by experienced raters (geriatricians) in 115 older adults (mean age 78.0 ± 7.3; 47% females) from a single centre.
Results: the intraclass correlation coefficient (ICC) for the CFS classification tree was 0.833 (95% CI: 0.768–0.882) when compared with the geriatricians' CFS scoring. In 93%, the classification tree rating was the same or differed by at most one level with the expert geriatrician ratings. The ICC was 0.805 (0.685–0.883) when CFS scores from the classification tree were compared with the clerk/resident scores; 88.5% of the ratings were the same or ±1 level.
Conclusions: a classification tree for scoring the CFS can help with reliable scoring by relatively inexperienced raters. Though an incomplete remedy, a classification tree is a useful support to decision-making and could be used to aid routine scoring of the CFS.

Keywords: frailty, clinical frailty scale, classification tree, ageing, older people

Key points

- CFS scoring by inexperienced raters might not always be identical to expert judgement.
- The agreement was good for the CFS classification tree when compared with CFS ratings by experienced and inexperienced raters.
- The CFS classification tree can aid routine scoring of the CFS.

Introduction

Treatment decisions require information about an individual's goals of care, their severity of illness, level of frailty and other health indicators. With COVID-19, such decisions now can include how to allocate access to limited resources, particularly critical care admissions. For this, the Clinical Frailty Scale (CFS) is recommended [1–4].

Introduced in the second clinical examination of the Canadian Study of Health and Aging, the CFS summarises the level of fitness or frailty of an older adult after evaluation by a health care professional [5]. Now a 9-point scale from 1 ('very fit') to 9 ('terminally ill'), higher scores represent greater risk [2]. The CFS aims to reflect the baseline health state (2 weeks before); scoring it requires clinical judgement. Although still used to summarise a Comprehensive Geriatric Assessment [6], more widespread uptake requires that CFS scoring be undertaken by people new to frailty assessment [7–10].

In the UK, outside geriatrics and frailty services, few clinicians receive formal frailty identification training; many lack confidence in this, and desire more frailty education [11]. When used against a background of care rationing more than for the development of a traditional care plan, the results might not always reflect expert geriatrician judgement [12]. CFS inter-rater reliability is generally very good [13,14]. Even so, some evidence suggests that personal bias may play a role in judgement-based frailty assessment, especially with inexperienced raters [15–17]. In consequence, we developed a classification tree to improve CFS reliability when employed by inexperienced raters. Our objective was to compare the scoring of the CFS by the classification tree with the scoring of the CFS done by experienced (i.e. geriatricians) and inexperienced (i.e. trainee) CFS raters. As a secondary objective, we also compared inter-rater reliability of inexperienced versus experienced raters.

Methods

This is a prospective study of 115 patients aged 65+ years seen in the Emergency Department having been referred to Internal Medicine ($N = 43$), in clinic ($N = 40$), on the Geriatric Medicine inpatient consult service ($N = 21$), or at home ($N = 11$). The CFS scoring followed a Comprehensive Geriatric Assessment, both completed by clinical clerks, residents or geriatricians. To inform their decisions, they used any available multidisciplinary team assessments, and medication and diagnostic data from the patient's health record. This information was then integrated into a care plan. After reviewing the assessment and plan, and interviewing the patient, one of the two attending geriatricians (SDS; KR) assigned the patient a CFS score. For 52 patients, clerks/residents participated in the Comprehensive Geriatric Assessment and assigned a CFS score independently from the geriatrician; residents did so for 38 patients and clerks for 14. Clerks and residents were either entirely new to the

CFS or had limited experience. Local training is a 1-hour session, with on-service, case-by-case geriatrician review.

After completing the judgement-based CFS scoring, the CFS was scored separately using the classification tree. The classification tree (Figure 1) asks questions based on the descriptions of each of the CFS levels. We also developed a two-page questionnaire that could be used to collect the data needed to complete the classification tree. It asks about basic and instrumental activities of daily living, chronic conditions, self-rated health, energy level and physical activity (Supplementary Appendix A). As these items are often collected in clinical care and do not require specialist training, the questionnaire is not essential for the classification tree. Here, as a process check and for later database comparisons, the clerks/residents or geriatricians administered the questionnaire to all participants. They used the best information (e.g. validated diagnoses) typically with caregiver input. A research team member reviewed each participant's responses and used the classification tree to derive a CFS score. This project was undertaken as a Quality Assurance (QA) initiative and was approved by the local QA Committee.

Intraclass correlation coefficients (ICCs) and their 95% confidence intervals (CIs) were used to compare the classification tree CFS scores with the CFS scoring by experienced and inexperienced raters. ICC reflects both degree of correlation and agreement between measurements [18,19]. As this was a clinical study, it was not possible for the same set of raters to rate all patients, and as such, ICC estimates were based on a single measurement, absolute agreement and a one-way random effects model [20]. Recognising the arbitrariness of any cut-point in the early stages of research, a minimum reliability of 0.70 is sufficient to conclude good agreement [21]. As a sensitivity analysis, we repeated this analysis excluding CFS level 9 (being terminally ill). CFS level 9 focuses on the current health state, less their baseline health; this impacted only classification tree versus geriatrician score comparisons. Spearman correlation coefficients were used to describe the association between the CFS scores and age. Analyses were conducted using SPSS version 26. To detect an interrater reliability of at least 0.70, assuming a null hypothesis of nominal correlation of 0.20, and given a tolerance of ± 0.20, with $\alpha = 0.05$ and $\beta = 0.80$, we estimated a need for 40 ratings [22].

Results

The mean (SD) age was 78.0 (7.3) years (range 65–93); 61 patients (53%) were male. CFS scores of 5 or greater were assigned to 86.1% (95% CI: 78.4–91.8%) of the patients based on the classification tree, 80.9% (72.5–87.6%) based on geriatrician scoring and 71.2% (56.9–82.9%) based on the clerk/resident scoring (Supplementary Appendix B). The CFS scores based on the classification tree were not significantly correlated with age (rho = 0.125, $P = 0.183$) and the
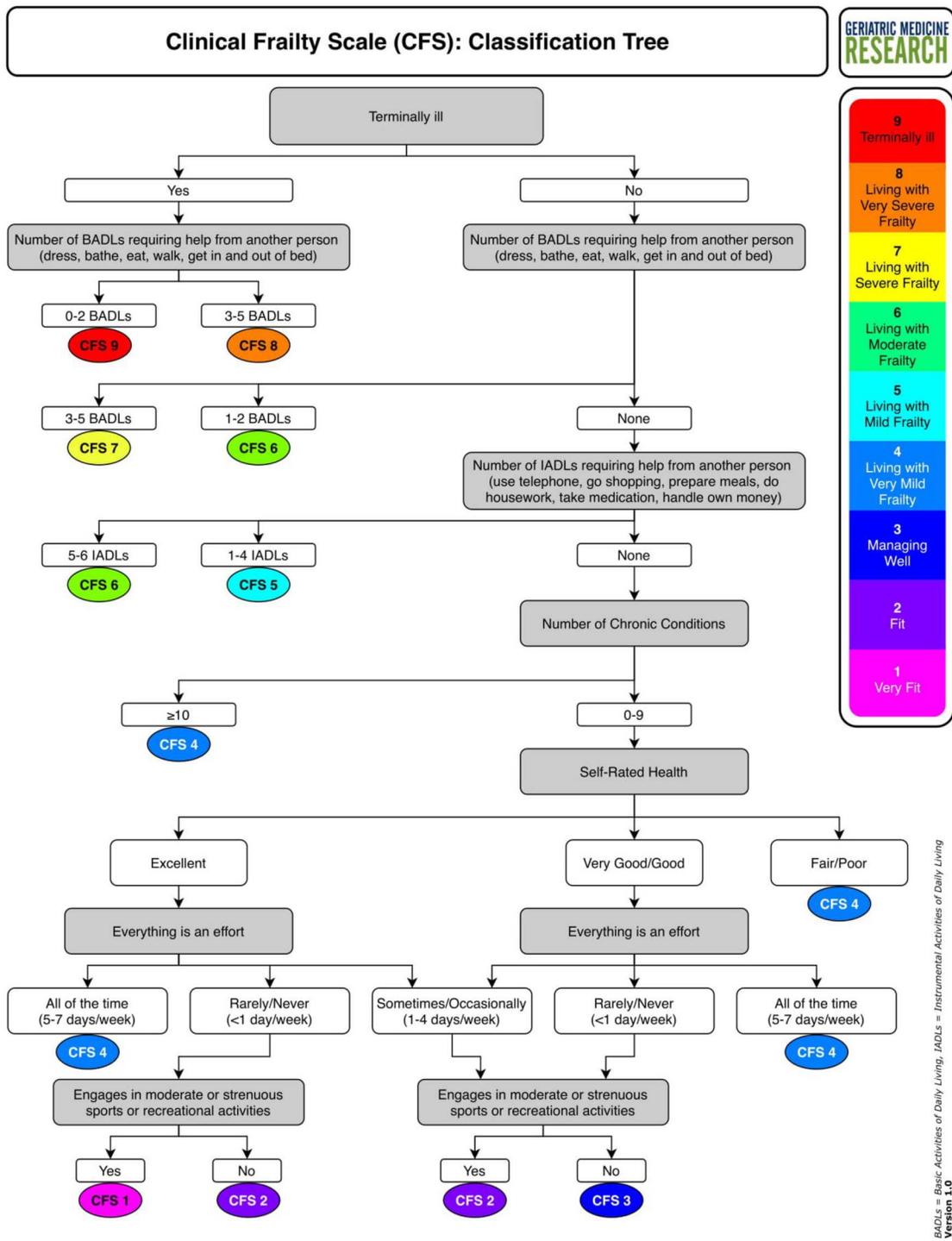
Figure 1. The Clinical Frailty Scale classification tree.

clerks/residents (rho = 0.212, P = 0.131), but the CFS scores based on the geriatricians were (rho = 0.191, P = 0.041).

The classification tree scores coincided with the geriatrician scores in 62.6% of cases; 93% received the same or ±1 score. The CFS rating between the classification tree and the clerks/residents were the same in 57.7% of cases, and within one score in 88.5%. Agreement was stronger when comparing the CFS rating between the geriatricians and the clerks/residents (same score in 76.9% of cases; the same or ±1 score in 96.2%) (Figure 2). The most common discrepancy in CFS scoring was for 14 patients in whom tree classification was 6 and the geriatrician's score was 5. Otherwise, there was no consistent pattern in re-assignment by the geriatrician or the clerks/residents (Supplementary Appendix C).
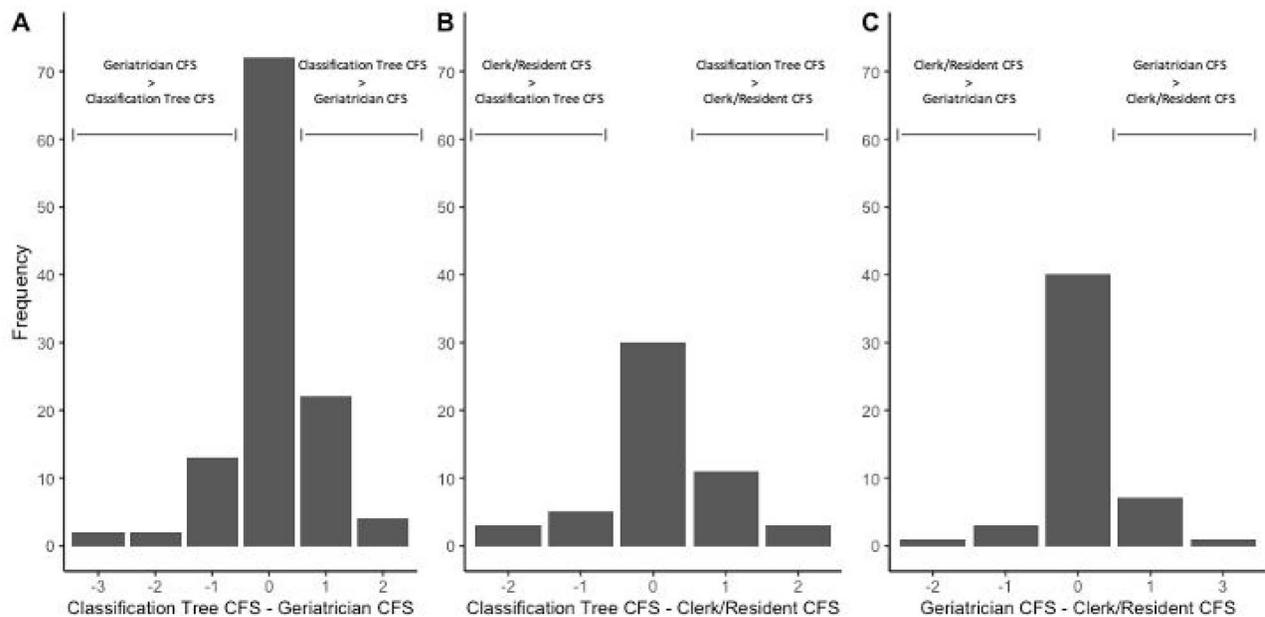
Figure 2. Agreement on CFS scoring between the classification tree and the two healthcare professional raters.

Agreement was good for the CFS classification tree: ICC = 0.833 (0.768–0.882) when compared with the geriatricians' CFS scoring and ICC = 0.805 (0.685–0.883) when compared with the clerk/resident CFS scoring. When comparing the CFS scoring between the geriatricians and the clerks/residents, the ICC was 0.879 (0.798–0.928). The sensitivity analysis estimated an ICC of 0.795 (0.715–0.854) for the CFS classification tree versus geriatrician comparison.

Discussion

A CFS classification tree afforded good agreement with an expert rating. Most classification tree scores were either the same or differed by one level compared with the other raters. Still, the classification tree classified people with at least a mild level of frailty slightly more often than the raters. In some cases, differing by at most a single CFS level might be acceptable. In other cases, a small difference could determine receiving or withholding care (e.g. where a predetermined CFS level represents a go/no go rule; [1]). For these cases, a stricter approach to grading risk is needed. Even so, whether the CFS classification tree could improve routine frailty scoring requires further study.

Employing the classification tree is not intended to replace the Comprehensive Geriatric Assessment. Inexperienced raters could use the classification tree to screen for frailty using routinely collected data or the questionnaire (Supplementary Appendix A). We acknowledge that because the raters of the Comprehensive Geriatric Assessment and the questionnaire used to calculate the CFS classification tree score were the same, bias could have been introduced. Future studies should explore this by having a rater complete the questionnaire independently from the Comprehensive Geriatric Assessment.

Discrepancies in CFS ratings were in more than one direction. Expert raters necessarily score the CFS based on information particular to the patient. Finely grained considerations, such as the pattern of illnesses and disabilities, or choice in performing some activity, do not readily translate into a classification tree. Less still does patient forbearance in coping with severe illness. This is what judgement seeks to do: provide an individual context for a given piece of information. Inexperienced raters will need to confirm whether their clinical judgement agrees with the classification tree CFS scoring; especially where they are material, differences will require adjudication.

Understanding the degree of a patient's frailty is important to prognostication [23–29]. A judgement-based, assessment-informed CFS aims to better understand likely challenges, such as the risk of common adverse outcomes [29,30]. Even so, frailty is just one factor that needs to be considered [2,7,9]. Outcomes of given CFS levels need to be addressed in relation to the severity of illness, and of course honouring patient preferences.

This is a single-site study, the sample was small and many raters contributed only a few ratings. Even so, we have power to detect a significant relationship between scoring sources, and the ICC is robust to using many raters [18–20]. One rater was an originator of the scale and in such settings agreement tends to be higher [18]. Whether using the CFS classification tree can help inexperienced raters with their CFS scoring is motivating further inquiry by our group.

References

1. National Institute for Health and Care Excellence (NICE). COVID-19 rapid guideline: critical care in adults, NICE guideline [NG159]. In: NICE Guide–COVID-19, 2020.
2. Rockwood K, Theou O. Using the Clinical Frailty Scale in allocating scarce health care resources. Can Geriatr J CGJ 2020; 23: 210–5.
3. Aw D, Woodrow L, Ogliari G et al. Association of frailty with mortality in older inpatients with Covid-19: a cohort study. Age Ageing 2020; 49: 915–22.
4. Owen RK, Conroy SP, Taub N et al. Comparing associations between frailty and mortality in hospitalised older adults with or without COVID-19 infection: a retrospective observational study using electronic health records. Age Ageing 2020; afaa167. doi: 10.1093/ageing/afaa167.
5. Rockwood K, Song X, MacKnight C et al. A global clinical measure of fitness and frailty in elderly people. Can Med Assoc J 2005; 173: 489–95.
6. Pulok MH, Theou O, van der Valk AM et al. The role of illness acuity on the association between frailty and mortality in emergency department patients referred to internal medicine. Age Ageing 2020; 49: 1071–9.
7. Hubbard RE, Maier AB, Hilmer SN et al. Frailty in the face of COVID-19. Age Ageing 2020; 49: 499–500.
8. Lewis EG, Breckons M, Lee RP et al. Rationing care by frailty during the COVID-19 pandemic. Age Ageing 2021; 50: 7–10.
9. Rockwood K. Rationing care in COVID-19: if we must do it, can we do better? Age Ageing 2021; 50: 3–6.
10. Montero-Odasso M, Hogan DB, Lam R et al. Age alone is not adequate to determine healthcare resource allocation during the COVID-19 pandemic. Can Geriatr J 2020; 23: 152–4.
11. Taylor JK, Fox J, Shah P et al. Barriers to the identification of frailty in hospital: a survey of UK clinicians. Future Healthc J 2017; 4: 207–12.
12. Surkan M, Rajabali N, Bagshaw SM et al. Interrater reliability of the Clinical Frailty Scale by geriatrician and intensivist in patients admitted to the intensive care unit. Can Geriatr J CGJ 2020; 23: 235–41.
13. Ringer T, Thompson C, McLeod S et al. Inter-rater agreement between self-rated and staff-rated Clinical Frailty Scale scores in older emergency department patients: a prospective observational study. Acad Emerg Med 2020; 27: 419–22.
14. Lo AX, Heinemann AW, Gray E et al. Inter-rater reliability of clinical frailty scores for older patients in the emergency department. Acad Emerg Med 2020. doi: 10.1111/acem.13953.
15. Young RL, Smithard DG. The Clinical Frailty Scale: do staff agree? Geriatrics 2020; 5: 40.
16. Shrier W, Dewar C, Parrella P et al. Agreement and predictive value of the Rockwood Clinical Frailty Scale at emergency department triage. Emerg Med J 2020. doi: 10.1136/emermed-2019-208633.
17. Pugh RJ, Battle CE, Thorpe C et al. Reliability of frailty assessment in the critically ill: a multicentre prospective observational study. Anaesthesia 2019; 74: 758–64.
18. Streiner DL, Norman GR, Cairney J. Health Measurement Scales. Oxford: Oxford University Press, 2015.
19. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012; 8: 23–34.
20. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996; 1: 30–46.
21. Nunnally JC. Psychometric Theory. vol. 2. New York: McGraw-Hill, 1978.
22. Kraemer HC, Thiemann S. How Many Subjects? Statistical Power Analysis in Research. Newbury Park: Sage Publications, 1987.
23. Hewitt J, Carter B, Vilches-Moraga A et al. The effect of frailty on survival in patients with COVID-19 (COPE): a multicentre, European, observational cohort study. Lancet Public Health 2020; 5: e444–51.
24. Darvall JN, Bellomo R, Bailey M et al. Frailty and outcomes from pneumonia in critical illness: a population-based cohort study. Br J Anaesth 2020; 125: 730–8.
25. Utino Taniguchi L, Ibrahim Q, Azevedo LCP et al. Comparison of two frailty identification tools for critically ill patients: a post-hoc analysis of a multicenter prospective cohort study. J Crit Care 2020; 59: 143–8.
26. Fernando SM, McIsaac DI, Rochwerg B et al. Frailty and invasive mechanical ventilation: association with outcomes, extubation failure, and tracheostomy. Intensive Care Med 2019; 45: 1742–52.
27. Montgomery CL, Zuege DJ, Rolfson DB et al. Implementation of population-level screening for frailty among patients admitted to adult intensive care in Alberta, Canada. Can J Anesth 2019; 66: 1310–9.

28. Shears M, Takaoka A, Rochwerg B *et al.* Assessing frailty in the intensive care unit: a reliability and validity study. J Crit Care 2018; 45: 197–203.

29. Rickard F, Ibitoye S, Deakin H *et al.* The Clinical Frailty Scale predicts adverse outcome in older people admitted to a UK major trauma Centre. Age Ageing 2020; afaa180. doi: 10.1093/ageing/afaa180.

30. Church S, Rogers E, Rockwood K *et al.* A scoping review of the Clinical Frailty Scale. BMC Geriatr 2020; 20: 393.