

Inter-Rater Reliability of the Retrospectively Assigned Clinical Frailty Scale Score in a Geriatric Outreach Population



Jasmine Davies, MD¹, Jennifer Whitlock, RN¹, Iris Gutmanis, PhD², Sheri-Lynn Kane, MD, FRCPC³

¹Specialized Geriatric Services, St. Joseph's Health Care London, Parkwood Institute, London, ON; ²Lawson Health Research Institute, London, ON; ³Division of Geriatric Medicine, St. Joseph's Health Care London, Parkwood Institute, London, ON, Canada

DOI:<https://doi.org/10.5770/cgj.21.263>

ABSTRACT

Background

Frailty, a common clinical syndrome in older adults associated with increased risk of poor health outcomes, has been retrospectively calculated in previous publications; however, the reliability of retrospectively assigned frailty scores has not been established. The aim of this study was to see if frailty scores, based on chart review data, agreed with clinician-determined scores based on a comprehensive geriatric assessment.

Methods

Per standard practice, all patients seen by one nurse clinician (JW) from the Southwestern Ontario Regional Geriatric Program, a tertiary care-based outreach service, between August 15, 2013 and December 31, 2015 received a comprehensive geriatric assessment which included the assignment of an interview-based Clinical Frailty Scale score (CFS-I). Subsequently, a medical student researcher (JD), blinded to the CFS-I, assigned each consenting patient a frailty score based on chart review data (CFS-C). The inter-rater reliability of the CFS-I and CFS-C was then determined.

Results

Of the 41 consented patients, 39 had both a CFS-I and CFS-C score. The median CFS score was 6, indicating patients were moderately frail and required assistance for some basic activities of daily living. Cohen's kappa coefficient was 0.64, indicating substantial agreement.

Conclusion

CFS scores can be reliably assigned retrospectively, thereby strengthening the utility of this measure.

Key words: Clinical Frailty Scale, retrospective, frailty, reliability

INTRODUCTION

Due to its robust nature, frailty has recently become a commonly researched topic in many medical specialties.^(1,2) Although multiple frailty measures have been proposed, there is currently no one widely accepted tool.⁽³⁾ Some initially proposed measures have been criticized either for being too time-consuming or for requiring measures that are not routinely captured in clinical practice.⁽⁴⁾ In contrast, the Clinical Frailty Scale (CFS), a nine-level (scores range from 1: very fit, to 9: terminally ill) global frailty rating scale based on clinical judgement, allows health-care providers to assign a score based solely on a standard clinical interview.⁽⁵⁾ Scores are assigned based on: a) a patient's level of fitness (i.e., regularly active, occasionally active, and not regularly active); b) presence/absence of active disease symptoms (i.e., no active disease symptoms, controlled symptoms which do not limit activity, and symptoms limiting activity); c) dependency on others for daily help with activities of daily living (ADLs) (i.e., help needed with both high-order instrumental ADLs (IADLs) and some basic ADLs (BADLs), and completely dependent for all ADLs); and d) cognition.

As CFS scores are based on data routinely recorded during a comprehensive geriatric assessment, researchers conducting retrospective chart reviews have used this efficient frailty measure rather than other tools which require measurements, such as grip strength and walking speed.⁽⁶⁾ However, it is unknown if CFS scores can be reliably assigned retrospectively, based on information in client charts.^(7,8,9) Thus, the purpose of this study was to determine if CFS scores, based on chart review (CFS-chart or CFS-C), resembled CFS scores assigned after in-person patient interviews (CFS-interview or CFS-I).

METHODS

All patients who had been referred to tertiary care-based Specialized Geriatric Services (SGS), and then seen for an initial outpatient consultation by one SGS outreach nurse clinician (JW) were invited to participate. As part of the SGS outreach service, older community-dwelling adults who are unable to attend an outpatient visit receive a comprehensive geriatric assessment in their homes by one of the team's nurse clinicians, who then provides recommendations on appropriate treatment and services and links them with other team members (e.g., geriatrician, physiotherapist, occupational therapist, and social worker) as needed.

Sample size tables indicated that approximately 40 patients were required (two evaluators, null hypothesis: $\rho = 0$, alternative hypothesis: $\rho = 0.6$; $\alpha = 0.05$, $\beta = 0.20$).⁽¹⁰⁾ Recruitment began on August 15, 2013. The initial clinical assessment and assignment of the CFS-I score was done by one nurse clinician (JW) with over 15 years of experience working with older adults. Captured by all outreach clinicians during the initial assessment, CFS-I scores are only recorded on the team's intake assessment form and are not part of the electronic patient health record. Subsequently, the senior medical student researcher (JD), who had no extra training aside from some elective rotations in Geriatric Medicine and who was blinded to the CFS-I score, determined the CFS-C score based on the nurse-generated consultation note that included past medical history, medications, social history, history of presenting illness, investigations, and physical examination, per standard practice.

Patient characteristics, including living setting, access to formal and informal supports, ability to mobilize with/without a gait aid, cognitive status, comorbidities, and ability to perform BADLs and IADLs, were also abstracted from the electronic patient record. The Charlson Co-morbidity Index (CCI) score^(11,12) was calculated using past medical history.

Once CFS-C scores were assigned, the CFS-I scores, as well as information on the number of emergency department (ED) visits and falls in the prior six and prior three months, respectively, to the initial visit as captured on the intake form, were entered into the study database. Frequencies and measures of central tendency were calculated to describe the patient population. Cohen's kappa coefficient was calculated using SPSS v. 21 for the 39 patients with recorded CFS-I and CFS-C scores. This study was approved by Western University's Research Ethics Board.

RESULTS

Recruitment ran from August 15, 2013 to December 31, 2015 when 41 patients or their substitute decision-makers had provided informed consent. There were no changes in the usual practice of the outreach team during this time. Although 41 patients were consented, only 39 had documented CFS-I scores.

Twenty-nine (70.7%) patients were female and 24 (58.5%) were living with family (Table 1). While 19.5% of patients had not visited the ED in the six months prior to the consultation, 34.1% had visited the ED once, and 45.9% had attended the ED more than once. While 58.5% of study patients did not report a fall in the three months prior to their consultation, one patient reported 11 falls. Twenty-four people were receiving formal support as offered mainly through the Community Care Access Centre. Six patients were completely dependent for all BADLs and IADLs, and 21 (51.2%) used a walker for ambulation. Although only 2.4% of patients had been formally diagnosed with dementia, 53.7% cited undiagnosed memory concerns as part of the reason for referral. While CCI scores ranged from 0–7, over 60% had scores of 2 or below.

CFS-I scores varied from 4 to 7 (5 people [12.9%] were assigned a score of 4 [vulnerable], 19 [74.4%] had a score of 5 or 6 [mildly or moderately frailty], and 5 [12.8%] had a score of 7 [severely frail]), while CFS-C scores varied from 1 to 7 (7 [18.0%] scored between 0 and 4; 26 [66.6%] were mildly or moderately frail and 6 [15.4%] were severely frail). No one was classified as either very severely frail or terminally ill. Mean scores were not statistically significantly different (mean CFS-I score: 5.56 [95% confidence interval (CI): 5.28-5.74] vs. mean CFS-C score: 5.41 [95% CI: 5.03-5.79]), and the median for both CFS-I and CFS-C scores was 6.0. Cohen's kappa coefficient was 0.64, indicating substantial agreement.⁽¹³⁾ Twenty-nine (74.4%) of 39 scores agreed perfectly, while nine scores were only one level apart (Figure 1). Among the 32 people with frailty (CFS score of five or more), 26 (81.3%) scores agreed perfectly and six differed by one point. Only once did the scores differ by four levels; the CFS-I was 5 (mildly frail), while the CFS-C was 1 (very fit).

DISCUSSION

Study findings indicate that overall CFS scores determined from a retrospective chart review agree with those assigned immediately following clinical assessment (Cohen's kappa: 0.64), suggesting that retrospectively assigned CFS scores are a reliable measure of clinical frailty. However, the validity of retrospectively assigned scores warrants future study.

Increased understanding of the psychometric properties of frailty measures is needed, as frailty has become an increasingly common predictor variable in medical studies. For example, Masud and colleagues⁽⁸⁾ noticed that the variation in elderly patients' recovery from burns could not be simply explained by chronological age. Suspecting that frailty was a better outcome predictor, they assigned CFS scores based on retrospective chart review and found that those with lower frailty scores had improved survival rates. However, as the psychometric properties of a retrospectively determined CFS were unknown, they had to cite this as a study limitation.

The retrospective assignment of CFS scores may be dependent on the assessor's clinical abilities and impression.

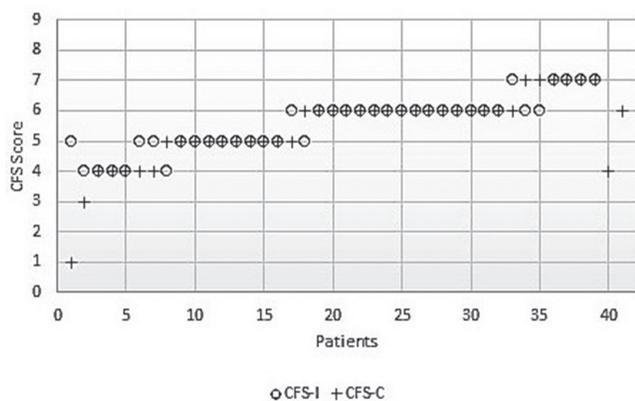
TABLE 1.
Study population

| <i>Continuous Variables</i> | | |
|---|-----------------------------------|-------------------------|
| <i>Measure</i> | <i>Mean (SD)</i> | <i>Range</i> |
| Age | 83.29 (8.04) | 65-98 |
| Emergency department visits (last 6 months) | 1.74 (1.81) | 0-7 |
| Falls (last 3 months) | 1.13 (2.15) | 0-11 |
| Charlson Co-morbidity Index | 1.48 (1.57) | 0-7 |
| <i>Categorical Variables</i> | | |
| <i>Variable</i> | <i>Categories</i> | <i>Number (Percent)</i> |
| Sex | • Male | 12 (29.3) |
| | • Female | 29 (70.7) |
| Formal supports | • Present | 24 (58.5) |
| | • Absent | 17 (41.5) |
| Living arrangement | • Alone | 12 (29.3) |
| | • With family | 24 (58.5) |
| | • Retirement home | 5 (12.2) |
| Informal supports | • Present | 31 (75.6) |
| | • Absent | 9 (22.0) |
| | • Missing data | 1 (2.4) |
| Exercise | • Regularly throughout the year | 4 (9.8) |
| | • Not regularly active | 32 (78.0) |
| | • Missing data | 5 (12.2) |
| Basic Activities of Daily Living | • Independent | 18 (43.9) |
| | • Dependent | 23 (56.1) |
| Instrumental Activities of Daily Living | • Independent | 8 (19.5) |
| | • Dependent | 33 (80.5) |
| Gait aid indoors | • None | 12 (29.3) |
| | • Cane | 4 (9.8) |
| | • Walker | 21 (51.2) |
| | • Wheelchair | 4 (9.8) |
| | • Scooter | 0 (0) |
| Gait aid outdoors | • None | 10 (24.4) |
| | • Cane | 5 (12.2) |
| | • Walker | 20 (48.8) |
| | • Wheelchair | 4 (9.8) |
| | • Scooter | 2 (4.9) |
| Memory concerns | • None | 6 (14.6) |
| | • Undiagnosed but concerns raised | 22 (53.7) |
| | • Mild Cognitive Impairment | 4 (9.8) |
| | • Dementia | 8 (19.5) |
| | • Missing data | 1 (2.4%) |

SD = standard deviation.

In this study, the initial assessment was done by one nurse clinician with considerable clinical experience with older patients. Therefore, it is likely that information required for the determination of a retrospectively assigned CFS score was captured in the consultation note. It is unknown if the same result would have been obtained if the consultation note

had been done by someone with little geriatric experience. Further, as the medical student researcher did not have much additional geriatric medicine training, our results are likely generalizable to other researchers without a strong geriatrics background, assuming they have access to similarly comprehensive documentation. Additionally, while a consultation



CFS-I = Clinical Frailty Scale score assigned by nurse clinician (JW); CFS-C = Clinical Frailty Scale score assigned by medical student researcher (JD). Overlapping shapes (circle with cross inside) represents 100% agreement between the two raters.

FIGURE 1. Comparison of CFS-I and CFS-C scores

note can capture the information required to assign a retrospective CFS score, it is possible that notes do not fully capture both clinical impression and all of the information needed to retrospectively assign a CFS score. This may explain the one score that differed by four points.

To limit variability only two raters were used. Consequently, it took more than two years to consent the required number of patients. However, it is unlikely that time was a significant confounder as there were no changes to outreach team practice during this time. Additionally, as with all retrospective studies, we were unable to control which data were captured in the patient record and intake forms and, as a result, some patient characteristics and two CFS-I scores were unavailable. Also, this study focused on community-dwelling patients living in one location who were referred to tertiary care-based SGS and were unable to attend an outpatient appointment. Thus, it is likely that study patients were frailer than a general community sample. Future multisite studies with larger sample sizes are needed to determine broader generalizability.

CONCLUSION

The substantial agreement between the CFS-I and the CFS-C provides evidence that the CFS can be reliably used as a measure of frailty in retrospective chart reviews if the charts contain all the elements required to assign a CFS score.

CONFLICT OF INTEREST DISCLOSURES

The authors declare that no conflicts of interest exist.

REFERENCES

1. Makary MA., Segev DL, Pronovost PJ, et al. Frailty as a predictor of surgical outcomes in older patients. *J Am Coll Surg.*

2010;210(6):901–908. Accessed 2016 Jan. Available from: <http://dx.doi.org/10.1016/j.jamcollsurg.2010.01.028>

2. Wallis SJ, Wall J, Biram RWS, et al. Association of the Clinical Frailty Scale (CFS) with hospital outcomes. *QJM.* 2015;108(12):943–949. Accessed 2016 Jan. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/25778109>

3. Bouillon K, Kivimaki M, Hamer M, et al. Measures of frailty in population-based studies: an overview. *BMC Geriatr.* 2013;13(1):64. Accessed cited 2016 Jan. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3710231&tool=pmcentrez&rendertype=abstract>

4. Singh I, Gallacher J, Davis K, et al. Predictors of adverse outcomes on an acute geriatric rehabilitation ward. *Age Ageing.* 2012;41(2):242–246. Accessed 2016 Jan. Available from: <http://ageing.oxfordjournals.org/content/41/2/242.long>

5. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ.* 2005;173(5):489–495. Accessed 2016 Jan. Available from: <http://www.cmaj.ca/content/173/5/489.long>

6. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults : evidence for a phenotype. *J Gerontol A Biol Sci Med Sci.* 2001;56(3):146–156. Accessed 2016 Jan. Available from: <http://biomedgerontology.oxfordjournals.org/content/56/3/M146.abstract>

7. Parmar KR, Xiu PY, Chowdhury MR, et al. In-hospital treatment and outcomes of heart failure in specialist and non-specialist services: a retrospective cohort study in the elderly. *Open Heart.* 2015;2(1):e000095. Accessed 2016 Jan. Available from: <http://openheart.bmj.com/cgi/doi/10.1136/openhrt-2014-000095>

8. Masud D, Norton S, Smailes S, et al. The use of a frailty scoring system for burns in the elderly. *Burns.* 2013;39(1):30–36. Accessed 2016 Jun. Available from: <http://dx.doi.org/10.1016/j.burns.2012.03.002>

9. Ormerod JOM, Ramcharitar S. Does specific interventional risk scoring better predict mortality than comorbidity in nonagenarians undergoing coronary angioplasty? *Cardiovasc Revasc Med.* 2014;15(4):258–260. Accessed 2016 Jun. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/?term=Does+specific+interventional+risk+scoring+better+predict+mortality+than+comorbidity+in+nonagenarians+undergoing+coronary+angioplasty%E2%80%AF%3F>

10. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med.* 1998;17(1):101–110. Accessed 2016 Jun. Available from: [http://onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-0258\(19980115\)17:1%3C101::AID-SIM727%3E3.0.CO;2-E/abstract](http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0258(19980115)17:1%3C101::AID-SIM727%3E3.0.CO;2-E/abstract)

11. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic in longitudinal studies: development and validation. *J Chron Dis.* 1987;40(5):373–383. Accessed 2016 Jun. Available from: <http://www.sciencedirect.com/science/article/pii/0021968187901718>

12. Quan H, Li B, Couris CM, et al. Updating and validating the Charlson Comorbidity Index and score for risk adjustment in hospital discharge abstracts using data from 6 countries. *Am J*

Epidemiol. 2011;173(6):676–682. Accessed 2016 Jun. Available from: <http://aje.oxfordjournals.org/content/173/6/676.long>

13. Viera AJ, Garrett JM. Understanding interobserver agreement : the Kappa Statistic. *Fam Med.* 2005;37(5):360–363. Accessed 2016 Jun. Available from: <http://www.stfm.org/fmhub/fm2005/May/Anthony360.pdf>

Correspondence to: Jasmine Davies, MD, Specialized Geriatric Services, St. Joseph's Health Care London, Parkwood Institute, Main Building, 550 Wellington Rd., London, ON, N6C 0A7 Canada

E-mail: Jdavies2016@gmail.com