Article:
Pugh, RJ, Battle, CE, Thorpe, C et al. (8 more authors) (2019) Reliability of frailty assessment in the critically ill: a multicentre prospective observational study. Anaesthesia, 74 (6). pp. 758-764. ISSN 0003-2409

https://doi.org/10.1111/anae.14596

Submitting author: Dr Richard Pugh

Department of Anaesthetics

Glan Clwyd Hospital

Sarn Lane, Bodelwyddan, Denbighshire, UK.

richard.pugh@wales.nhs.uk

MS 21673: Reliability of frailty assessment in the critically ill: a multi-centre prospective observational study

R.J. Pugh, [1] C.E. Battle, [2] C. Thorpe, [3] C. Lynch, [4] J.P. Williams, [5] A. Campbell, [6] C.P. Subbe, [7] R. Whitaker, [8] T. Szakmany, [9] A.P. Clegg, [10] N.I. Lone [11]

[1] Consultant in Anaesthetics and Intensive Care, Department of Anaesthetics, Glan Clwyd Hospital, Bodelwyddan, Denbighshire, UK. richard.pugh@wales.nhs.uk, @Richard05080183

[2] Consultant Physiotherapist, Ed Major Critical Care Unit, Morriston Hospital, UK. ceri.battle@wales.nhs.uk, @ceribattle

[3] Consultant in Anaesthetics and Intensive Care, Department of Anaesthetics, Ysbyty Gwynedd, UK. chris.thorpe@wales.nhs.uk, @christhorpe5

[4] Consultant in Anaesthetics and Intensive Care, Intensive Care Unit, Royal Glamorgan Hospital, Llantrisant, UK. ceri.lynch5@wales.nhs.uk, @ceri_lynch

[5] Consultant in Anaesthetics and Intensive Care, Critical Care Directorate, Royal Gwent Hospital, Newport, UK. james.p.williams@wales.nhs.uk, @williamsmbbch

[6] Consultant in Anaesthetics and Intensive Care, Department of Anaesthetics, Wrexham Maelor Hospital, Wrexham, UK. andy.campbell@doctors.org.uk, @andyfresher

[7] Senior Clinical Lecturer in Acute and Critical Care Medicine, School of Medical Sciences, Bangor University, Bangor, UK. christian.subbe@wales.nhs.uk, @csubbe

[8] Statistical and Methodological Consultant, Whitaker Research Ltd. rhiannon@whitres.co.uk, @rhiannonwhit

[9] Senior Lecturer in Intensive Care, Director of Critical Illness Research, Cardiff University, Cardiff, UK. szakmanyt1@cardiff.ac.uk, @iamyourgasman

[10] Clinical Senior Lecturer, Academic Unit of Elderly Care and Rehabilitation, University of Leeds, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, UK. a.p.clegg@leeds.ac.uk, @drandyclegg

[11] Clinical Senior Lecturer in Critical Care, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK. nazir.lone@ed.ac.uk, @ICULone

Correspondence: Dr Richard Pugh

Email: richard.pugh@wales.nhs.uk

Key words: ageing; frailty; reliability; proxy; critical care

Twitter: @Richard05080183, @wicsarg

Short title**:**

Frailty Assessment Reliability in ICU (FAR-ICU)

Summary

Demand for critical care among older patients is increasing in many countries. Assessment of frailty may inform discussions and decision-making, but acute illness and reliance on proxies for history-taking pose particular challenges in the critically ill. Our aim was to investigate the reliability of frailty assessment in critical care. We conducted a prospective multicentre study comparing assessments of frailty by staff from medical, nursing and physiotherapy backgrounds. Each assessment was made independently by two assessors using the Clinical Frailty Scale after review of clinical notes and interview with an individual who maintained close contact with the patient. Inter-rater reliability was assessed using linear weighted kappa. Factors associated with higher frailty scores were assessed and those which were statistically significant included in ordinal regression. We made 202 assessments in 101 critical care patients (median age 69 years, 59% male, median APACHE II score 19) at the critical care units of seven hospitals in Wales and Scotland. 53% of patients were able to participate in interview. 35% patients were considered frail. Linear weighted kappa was 0.74 (95% confidence interval 0.67 to 0.80) indicating good level of agreement. However, frailty assessment score differed by at least one category in 47% cases; factors independently associated with higher score on regression analysis included the patient being female and medical assessor background. We therefore identified good level agreement in assessments of frailty using the Clinical Frailty Scale, supporting its use in clinical care, but identified factors independently associated with higher scores which could indicate personal bias.

Introduction

Formal assessment of frailty as an aid to healthcare is relatively new to critical care [1]. As national populations age and the numbers of older patients referred to critical care services increase, evidence of the predictive validity of frailty assessment suggests its potential value for clinical decision-making, for development of clinical pathways, and for bench-marking processes [2,3]. However, critical illness presents particular challenges since a distinction between acute illness and frailty may be difficult, and there is typically reliance on others to provide history. If clinicians are to use frailty assessment to guide critical care decision-making, it must have proven reliability in this context [4].

A recent systematic review identified the judgement-based Clinical Frailty Scale (CFS) [5] as the most commonly reported frailty assessment tool in critical care [6], but found only limited evidence of inter-rater reliability in a small (n=30) single-centre study [7]. More recently, a dual centre Canadian study conducted with a small group of assessors also found no significant difference between paired assessments based on critical care chart review [8].

In practice, frailty assessment is typically made after discussions with family members and, where possible, with input from the patient themselves. Furthermore, in many centres, frailty assessment is made by a variety of clinicians, including medical staff, nurses and allied health professionals. The primary aim of this study was to assess the inter-rater reliability (IRR) of frailty assessment using CFS in a sample of patients from a United Kingdom (UK) critical care patient cohort when undertaken by members of a multidisciplinary critical care team.

Methods

The study was registered prospectively with the ISRCTN registry (ISRCTN17509500) and research ethics committee approval issued in Wales (REC 17/WA/0168) and Scotland (REC 17/SS/0121).

Written consent was obtained for each patient; for patients without capacity, consent was obtained from the nearest relative (Scotland) or consultee declaration from an individual close to the patient (Wales).

This was a prospective, observational multi-centre study. The study was carried out in the medical-surgical critical care units of six hospitals in Wales and one in Scotland. All participating units were sites where frailty assessment (using CFS) has been established as an element of routine practice and a training presentation which discussed frailty categorisation was provided to sites prior to study start date. Patient recruitment and initial data collection for all sites took place between December and April 2018, with completion of follow-up data collection July 2018.

The inclusion criteria for patient study entry were: age 60 years or more, receipt of active treatment and an expectation to remain in critical care for at least 24 hours. Lack of patient capacity did not preclude study entry. Assessors included study investigators working within the critical care team and members of clinical staff (medical, nursing and physiotherapy) caring for the patient and familiar with CFS application. The interviewee was required to be an individual who lived with the patient or a person who had known the patient for at least five years [9] and had spoken with him or her at least twice per month. The interview was expected to include the patient themselves if they had capacity to participate.

Our primary measure was assessment of frailty using a 9-category CFS reflecting a time-point approximately one month before hospital admission [5]; according to this assessment tool, CFS> 4 is considered "frail." Other variables included patient factors (age, sex, APACHE II score, surgical status [non-surgical, elective/ scheduled, urgent/ emergency], residence [home, nursing home or equivalent, non health-related institution, residential place of work or education, hospice or equivalent, no fixed abode or temporary abode], pre-hospital dependence [categorised by the clinical team as being able to live without assistance in daily activities, or minor, major or total

assistance with daily activities], use of mechanical ventilation or sedation in first 24 hours of critical care), Glasgow coma scale (GCS) in first 24 hours of critical care, assessor characteristics (clinical background and years of post-graduate experience) and interview descriptors (patient participation, and whether the interviewee was someone who ordinarily lived with the patient). Secondary measures were critical care length of stay and hospital mortality (censored at 30 days).

Data collected at time of interview included: assessor characteristics; usual contact between patient and main interviewee; availability of information with regards: residence, family support, external care, exercise tolerance and activities of daily living; and CFS rating (see Supplementary file: Data Collection Form).

Follow-up data were collected from each unit's Ward Watcher critical care database (Critical Care Audit Ltd, West Yorkshire, UK) one month after the end of site recruitment. Unit activity data included: numbers of admissions during the period of recruitment, age, sex, primary reason for admission, surgical status, APACHE II score, and receipt of mechanical ventilation in the first 24 hours of critical care. Additional patient-specific data for the recruited patients included: residence, pre-hospital dependence, surgical status, APACHE II score, use of sedation and mechanical ventilation in first 24 hours, GCS in first 24 hours, and critical care length of stay and hospital mortality.

Screening was performed by the study investigators. Two assessors were identified for each patient recruited; neither, one or both assessors could be a study investigator working within the critical care team; conversely, neither, one or both would be a member of clinical staff. Where one or both assessors were not study investigators, they were typically selected as a member of the clinical team who had pre-existing contact with the patient. Each assessor was asked to review the clinical records (including medical notes) available at time of assessment, and to interview the individual who lived with the patient or knew them well (see above), including the patient themselves too if they were

able to participate. The interview was unstructured, but the data collection form was used to collect information with regards assessor and interviewee background, the availability of data relevant to frailty assessment, and provided guidance for CFS rating (Supplementary file: Data Collection Form).

Representativeness of the included cohort was examined by comparison with characteristics of excluded but otherwise eligible patients (in terms of age, sex, illness severity, surgical status and need for mechanical ventilation). Attempt was made to reduce ascertainment bias by ensuring that the interview took place with an individual who lived with the patient or had known them for at least five years and spoken with them at least twice per month; ascertainment bias was further explored by recording the availability of information relevant to CFS assessment on the data collection form. Assessments of frailty were independent and blinded to the other rater.

On the basis of pilot study data [7], it was estimated that 100 participants would allow a kappa statistic of 0.80 to be estimated with a standard error of 0.05. After recruiting the first 60 patients (120 observations), CFS assessments were evaluated *a priori* to ensure that frail patients (with CFS classification above 4) were adequately represented, considered to reflect a minimum 16% of all observations, i.e. at least 20 observations of CFS> 4. At this point, there were 53 observations of CFS>4 , so study population enrichment was not required.

The distribution of continuous data was tested using Kolmogorov-Smirnov test, accepting non-normal distribution where the p value was 0.05 or less. Non-normally distributed data were summarised in terms of median and inter-quartile range (IQR). CFS rating and dependency were considered ordinal variables and presented as median value and percentage according to category. For purpose of analysis, assessors were designated "Rater 1" or "Rater 2" at time of data transcription on the basis of greater or lesser years of post-graduate experience, respectively.

The data were analysed using SPSS 23.0 (IBM, New York) and MedCalc 18.6 (MedCalc Software, Ostend). Inter-rater reliability was assessed using linear weighted kappa in order to minimise

influence of outlier ratings (0.41 - 0.6 reflecting moderate agreement; 0.61 - 0.80 representing good agreement; >0.80 representing very good agreement [10]). Chi-square test and chi-square test for trend were used to investigate the relationship between CFS rating and categorical and ordinal data, respectively. The Kruskal-Wallis test was used to examine the relationship between continuous (non-normally distributed) data and CFS rating. Ordinal regression was performed, with variable selection for the multivariable model based upon univariable associations at a 5% significance threshold. Assumption of proportional odds was assessed by a likelihood ratio test comparing the fit of the proportional odds model to a model with varying location parameters. Model fit was assessed by likelihood ratio test comparing the full model to intercept-only model.

Results

101 included critically ill patients each underwent two independent assessments of frailty at the seven participating sites (Figure 1). Approximately 18% of all eligible patients were included in the study. The main reason for non-enrolment was lack of researcher availability (96% excluded cases; Figure 1). Comparison of included and excluded patients demonstrated that included patients were significantly younger, sicker (according APACHE II score), more likely to be non-surgical and more likely to receive mechanical ventilation than excluded patients (Supplementary file: Table S2). Median recruitment per site was 12 patient (range 10 to 20. Supplementary file: Table S3).

Baseline characteristics of included patients are presented in Table 1. The median age was 69 years, 59% were male, and 75% patients were non-surgical. 62% patients received mechanical ventilation (including non-invasive ventilation) and 75% patients received sedation or had Glasgow Coma Score of less than 15 in the first 24 hours of critical care.

Most assessments were performed by medical staff (47%) or staff from a nursing background, including advanced critical care practitioners (44%), with a much smaller number by physiotherapists (9%). Nurse assessors had a significantly greater number of years experience (median 17 years, IQR

12-21) than medical assessors (median 13.5 years, IQR 3- 19; p=0.002); the number of physiotherapy assessors was too small to compare experience usefully with other clinical groups.

In 63% cases, interviews took place with an individual who lived with the patient; in 53% cases, the patient was also able to participate. In all but one case, the medical notes were available to support assessment; information regarding residence (98%), family support (92%), external care (75%), exercise tolerance (93%), and activities of daily living (94%) was available from notes and from interview to a variable degree.

The frequency distributions of frailty assessments are presented in Figure 2. The median CFS was 3.5 (inter-quartile range 2 to 5). Using a CFS cut-off >4, patients were considered frail by 35% assessors. There was perfect agreement in frailty assessment in 54 out of 101 cases (53%). Overall, there was good agreement between assessments (linear weighted kappa 0.74, 95% confidence interval 0.67 to 0.80; Table 2).

However, there was a difference of one CFS category in 40 (40%) cases, a difference of two in 5 (5%) cases, and difference of 3 categories in 2 (2%) cases. In 9 out of 101 cases, there was a difference in assessment such that one assessor considered a patient "frail" and one considered them "non-frail" (according to a cut-off of CFS>4). Among different staff pairings, the lowest level of agreement was found for the sub-group of patients for whom one assessor was from medical and one from nursing background (n=28; linear weighted kappa 0.59; 95% confidence interval 0.44 to 0.75; Table 2).

Patient factors associated with a significantly higher CFS rating were: increasing age and APACHE II score, female sex, increasing pre-hospital dependency, sedation in first 24 hours, and mechanical ventilation in first 24 hours (Table 3). Regarding the clinical background of assessor, those from a medical background recorded the highest CFS ratings.

Cumulative odds ordinal logistic regression with proportional odds was performed to further study the association between these factors and CFS rating. The assumption of proportional odds was met ($\chi^2(60) = 26.909$ P =1.000) and the final model predicted the dependent variable over and above the intercept-only model ($\chi^2(10) = 145.864$, P<0.001), indicating a good fit. Female patient sex , medical assessor background, higher category of pre-hospital dependence and higher APACHE II score were independently associated with a higher CFS rating (Table 5).

There were no statistically significant differences in mortality or critical care length of stay (for survivors) between CFS categories (Supplementary file: Table S4).

Discussion

We found a good level of agreement between ratings of CFS in critically ill patients aged 60 years and over according to linear weighted kappa. However, we also identified a difference in frailty scores in 47% patients. We identified factors independently associated with a higher CFS: female sex of the patient, medical background of the assessor, higher category of pre-hospital dependence and higher APACHE II score.

Considering our study's limitations, a large proportion of eligible patients (79%) were excluded due to lack of researcher availability; consequently, there were differences in characteristics between included and excluded patients, with excluded patients more likely to have been surgical, to have had lower APACHE II score and less likely to have received mechanical ventilation. This probably indicates a degree of selection bias excluding patients with shorter critical care stay, but suggests that we have evaluated performance of CFS in a cohort of particular interest, i.e. those who were particularly sick and less likely to be able to participate in interview. Indeed, the age, sex, illness severity score, surgical status and proportion of ventilated patients among those included were broadly similar to those reported in two recent systematic reviews [3,6]. Moreover, the proportion of assessments which could not involve patients (because of lack of capacity) was also similar [6].

We did not explicitly collect data relating to co-morbidity, which we anticipate will have influenced rater CFS assessments [11,12] , though severe comorbidity will have contributed to APACHE II calculation. The data collected on assessors was limited to clinical background and number of years of post-graduate experience; we did not capture experience of clinical staff in using CFS which may have varied between clinical backgrounds. However, a strength of our study is that we recruited patients from multiple critical care units and utilised clinical assessors with representative backgrounds and a range of clinical experience. As such we believe our results are broadly generalisable.

Our study demonstrated a good overall level of agreement between CFS assessments. This is consistent with findings from our earlier pilot work [7] and with a recently published Canadian study at two units comparing assessments made using chart review by a smaller group of assessors. [8] It is an important finding for a setting in which there is a high degree of reliance of proxy input for frailty assessment [8,13,14] and where retrospective assessment introduces risk of recall bias, particularly in the context of acute or sub-acute illness [2,15].

There were differences in rater assessments in nearly half the cases, and given the judgement-based nature of CFS assessment [2] we explored this further by evaluating factors associated with higher CFS rating. That age[16-18], illness severity [16,18] and pre-hospital dependence [13,19] were associated with higher CFS scores was expected from earlier critical care literature. Positive [18] and negative [20] associations between mechanical ventilation and frailty score have been previously reported, and we would suggest that the higher frailty ratings observed among non-ventilated patients in our study are likely to reflect care processes at participating centres and a lower tendency to progress to invasive ventilation among those perceived as frail.

However, of particular interest given our intent to study the psychometric properties of a measurement tool, we identified that assessors from a medical background rated frailty significantly

more highly than assessors from nursing and physiotherapy backgrounds (though years of experience had no apparent influence); consequently, we found the lowest level of inter-rater reliability was between assessors from a medical and from a nursing background. We also identified significantly higher CFS ratings for female patients, independent of age, prior dependency and APACHE II score.

An association between frailty and a subject's sex has been identified in previous critical care[16] and non-critical care studies [21], with sarcopoenia, longevity and progressive rather than sudden decline suggested as explanation for a higher incidence of frailty in women [21]. However, particularly given current uncertainties over the influence of a patient's sex on access to critical care [22,23], our findings indicate the possibility of personal bias in the application of a judgement-based frailty assessment [2]. There is evidently a need to further explore how frailty is perceived by a multidisciplinary critical care team and the potential influence that such understanding may have on clinical decision-making. Despite the ease of CFS application in critical care and its predictive validity [3,4], the relative advantages and disadvantages of CFS versus other frailty assessment tools (e.g. frailty index generated from routinely collected data [18,24,25]) are under-explored in this setting.

Conclusion

We have demonstrated in a multi-centre study that assessment of frailty in the critically ill using CFS has a good level of inter-rater reliability, according to linear weighted kappa. This is encouraging for critical care clinicians wishing to reference frailty in discussions and clinical decision-making, and for researchers looking to explore variations in critical care access, processes of care, resource requirements and outcomes in the context of frailty. However, we found evidence to suggest that personal bias may influence the application of CFS, and there is a need to evaluate further the relative performance of CFS versus other frailty assessment tools in this setting.

Acknowledgements

We are extremely grateful to members of the Welsh Intensive Care Society Audit and Research Group (WICSARG) and to Scottish critical care colleagues for facilitating participation at each of the sites and in particular wish to thank to Rose Dean, Una Gunter, Emma Hall and Tracy Owen for their help with data collection. We are also grateful to Rossela Roberts for her input to study design.

Competing interests

The authors declare that they have no competing interests

Figure 1. Patient flowchart

Figure 2. Distribution of frailty assessments

References

1. McDermid RC, Stelfox HT, Bagshaw SM Frailty in the critically ill: a novel concept. *Critical Care* 2011; **15:** 301.

2. Hogan DB, Maxwell CJ, Afilalo J, et al. A Scoping Review of Frailty and Acute Care in Middle-Aged and Older Individuals with Recommendations for Future Research. *Canadian Geriatrics Journal* 2017; **20:** 22-37.

3. Muscedere J, Waters B, Varambally A, et al. The impact of frailty on intensive care unit outcomes: a systematic review and meta-analysis. *Intensive Care Medicine* 2017; **43:** 1105-22.

4. Flaatten H, Clegg A Frailty: we need valid and reliable tools in critical care. *Intensive Care Medicine* 2018. Published online 11th October 2018. https://doi.org/10.1007/s00134-018-5404-5.

5. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *Canadian Medical Association Journal* 2005; **173:** 489-95.

6. Pugh RJ, Ellison A, Pye K, et al. Feasibility and reliability of frailty assessment in the critically ill: a systematic review. *Critical Care* 2018; **22:** 49.

7. Pugh RJ, Thorpe CM, Subbe CP A critical age: can we reliably measure frailty in critical care? *Critical Care* 2017; **21:** 121.

8. Shears M, Takaoka A, Rochwerg B, et al. Assessing frailty in the intensive care unit: A reliability and validity study. *Journal of Critical Care* 2018; **45:** 197-203.

9. Ahasic AM, Van Ness PH, Murphy TE, Araujo KL, Pisani MA Functional status after critical illness: agreement between patient and proxy assessments. *Age Ageing* 2015; **44:** 506-10.

10. Altman DG. *Practical statistics for medical research* London, Chapman and Hall, 1991.

11. Theou O, Rockwood MR, Mitnitski A, Rockwood K Disability and co-morbidity in relation to frailty: how much do they overlap? *Archives of Gerontology and Geriatrics* 2012; **55:** e1-8.

12. Yarnall AJ, Sayer AA, Clegg A, Rockwood K, Parker S, Hindle JV New horizons in multimorbidity in older adults. *Age and Ageing* 2017; **46:** 882-8.

13. Le Maguet P, Roquilly A, Lasocki S, et al. Prevalence and impact of frailty on mortality in elderly ICU patients: a prospective, multicenter, observational study. *Intensive Care Medicine* 2014; **40:** 674-82.

14. Hope AA, Hsieh SJ, Petti A, Hurtado-Sbordoni M, Verghese J, Gong MN Assessing the Usefulness and Validity of Frailty Markers in Critically Ill Adults. *Annals of the American Thoracic Society* 2017; **14:** 952-9.

15. Iwashyna TJ Trajectories of recovery and dysfunction after acute illness, with implications for clinical trial design. *American journal of respiratory and critical care medicine* 2012; **186:** 302-4.

16. Bagshaw SM, Stelfox HT, McDermid RC, et al. Association between frailty and short- and long-term outcomes among critically ill patients: a multicentre prospective cohort study. *Canadian Medical Association Journal* 2014; **186:** E95-102.

17. Fisher C, Karalapillai DK, Bailey M, Glassford NG, Bellomo R, Jones D Predicting intensive care and hospital outcome with the Dalhousie Clinical Frailty Scale: a pilot assessment. *Anaesthesia and Intensive Care* 2015; **43:** 361-8.

18. Zampieri FG, Iwashyna TJ, Viglianti EM, et al. Association of frailty with short-term outcomes, organ support and resource use in critically ill patients. *Intensive Care Medicine* 2018; **44:** 1512-20.

19. Bagshaw SM, Stelfox HT, Johnson JA, et al. Long-Term Association Between Frailty and Health-Related Quality of Life Among Survivors of Critical Illness: A Prospective Multicenter Cohort Study. *Critical Care Medicine* 2015; **43:** 973-82.

20. Brummel NE, Bell SP, Girard TD, et al. Frailty and Subsequent Disability and Mortality among Patients with Critical Illness. *American journal of respiratory and critical care medicine* 2017; **196:** 64-72.

21. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC Prevalence of frailty in community-dwelling older persons: a systematic review. *Journal of the American Geriatrics Society* 2012; **60:** 1487-92.

22. Fowler RA, Sabur N, Li P, et al. Sex-and age-based differences in the delivery and outcomes of critical care. *Canadian Medical Association Journal* 2007; **177:** 1513-9.

23. Garland A, Olafson K, Ramsey C, Yogendranc M, Fransoo R Reassessing access to intensive care using an estimate of the population incidence of critical illness. *Critical Care* 2018; **22:** 208.

24. Hope AA, Gong MN, Guerra C, Wunsch H Frailty Before Critical Illness and Mortality for Elderly Medicare Beneficiaries. *Journal of the American Geriatrics Society* 2015; **63:** 1121-8.

25. Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age and Ageing* 2016; **45:** 353-60.

Table 1. Characteristics of patients undergoing frailty assessment. Values are presented as median (IQR [range]) or number (proportion).

Patient characteristic		**All N=101**
Age; years		69 (65- 75 [60-80])
APACHE II score		19 (15- 23.0 [7-33])
Sex	Male	59 (58%)
	Female	42 (42%)
Residence	Home	99 (98%)
	Nursing home or equivalent	1 (1%)
	Non health-related institution	1 (1%)
Dependence	Able to live without assistance in daily activities	66 (65%)
	Minor assistance with some daily activities	25 (25%)
	Major assistance with majority of/ all daily activities	7 (7%)
	Total assistance with all daily activities	3 (3%)
Pathology	Respiratory	35 (35%)
	Gastrointestinal	27 (27%)
	Cardiovascular	16 (16%)
	Renal	10 (10%)
	Neurological	5 (5%)
	Polytrauma	3 (3%)
	Metabolic	2 (2%)
	Skin, soft tissue, isolated musculo-skeletal	2 (2%)
	Toxin	1 (1%)
Surgical status	Non-surgical	75 (74%)
	Surgical	26 (26%)
	-Elective/ scheduled	9 (9%)
	- Urgent/ emergency	17 (17%)
Mechanical ventilation in first 24 hours		63 (62%)
Sedation in first 24 hours		61 (60%)
GCS <15*		16(39%)
Patient unable to participate in interview		49 (49%)
Hospital mortality †		25 (25%)
ICU Length of stay; days ‡		6.5 (4.5- 15.9 [1-30])

*of those not receiving sedation in the first 24 hours (n=41).

†(censored at 30 days).

‡ of those surviving to critical care discharge

Table 2. Agreement and inter-rater reliability of frailty assessments according to paired backgrounds of assessors. Values presented as numbers (proportion), linear weighted kappa and 95% confidence intervals (CI). There were no assessments made by a doctor and a physiotherapist pair, or by two physiotherapists.

Paired background	Number	Perfect agreement	Linear weighted kappa	95% CI
All pairs	101	54 (53%)	0.74	0.67 to 0.80
Doctor-Doctor	32	20 (69%)	0.70	0.56 to 0.84
Doctor-Nurse	28	10 (36%)	0.59	0.44 to 0.75
Nurse-Nurse	22	12 (55%)	0.63	0.45 to 0.82
Nurse-Physiotherapist	19	14 (74%)	0.88	0.80 to 0.96

Table 3. Association between Clinical Frailty Scale ratings and patient, interview and rater characteristics. Data are presented (where applicable) as median (IQR [range]). P values are presented according to chi-square test and chi-square test-for-trend for categorical variables and Kruskal-Wallis test for continuous variables.

Characteristic (categorical variables)			**CFS, median (IQR [range])**	**P Value**
Interview	Patient participation	Yes	4 (2- 5 [1-9])	0.526
		No	3 (2- 5 [1-7])	
	Interviewee lives with patient	Yes	4 (2-5 [1-7])	0.286
		No	3 (2-5 [1-9])	
Rater	Clinical background	Medical	5 (3- 6 [1-9])	<0.001
		Nursing	3 (2- 4 [1-7])	
		Physiotherapy	2 (1- 5 [1-6])	
Patient	Sex	Male	3 (2- 5 [1-9])	<0.001
		Female	4 (3- 6 [1-7])	
	Ventilation	Yes	3 (2- 5 [1-9])	0.003
		No	4 (3- 6 [1-7])	
	Sedation	Yes	3 (2- 5 [1-9])	0.001
		No	4 (3- 6 [1-7])	
	Surgical status	Surgical	3 (1.25- 5 [1-7])	0.122
		Non-surgical	4 (2- 5 [1-9])	
	Dependency	Independent	3 (2- 4 [1-6])	<0.001
		Minor	5 (4- 6 [1-9])	
		Major	6 (4.75- 6 [4-7])	
		Total	6.5 (6- 7 [6-7])	
	Residence	Home	3 (2- 5 [1-9])	0.195

		Nursing Home	5 [5]	
		Institution - non-health related	5.5 (5- 6 [5-6])	
Characteristic (continuous variables)				**P Value**
Rater	Experience; years			0.759
Patient	Age; years			<0.001
	APACHE II score			<0.001

Table 4. Ordinal regression: association between Clinical Frailty Scale rating and patient and rater characteristics. The results of cumulative odds ordinal logistic regression - with proportional odds performed to identify variables with statistically significant effect on the dependent variable (Clinical Frailty Scale) - presented as odds ratio with 95% confidence intervals and associated p values.

Factor			**Odds Ratio**	**95% confidence interval**		**P Value**
				Lower	**Upper**	
Rater	Clinical background	Medical	7.320	2.744	19.528	<0.001
		Nursing	2.341	0.908	6.037	0.078
		Physiotherapy	1			
Patient	Ventilation	Yes	1			
		No	2.993	0.985	9.097	0.053
	Sedation	Yes	1			
		No	0.609	0.202	1.839	0.379
	Sex	Male	1			
		Female	2.791	1.573	4.952	<0.001
	Dependency	Independent	0.008	0.001	0.049	<0.001
		Minor	0.124	0.022	0.706	0.019
		Major	0.155	0.022	1.114	0.064
		Total	1			
	Age		0.975	0.935	1.017	0.240
	APACHE II		1.055	1.108	1.084	0.037

Supplementary File: Data collection form

Supplementary File: Table S1. Comparison of included and excluded patients

Supplementary File: Table S2. Number of patients recruited per site

Supplementary File: Table S3: Mortality and length of stay in relation to frailty