Review Article

A systematic review of predictive accuracy via c-statistic of preoperative frailty tests for extended length of stay, post-operative complications, and mortality

ABSTRACT

Frailty, as an age-related syndrome of reduced physiological reserve, contributes significantly to post-operative outcomes. With the aging population, frailty poses a significant threat to patients and health systems. Since 2012, preoperative frailty assessment has been recommended, yet its implementation has been inhibited by the vast number of frailty tests and lack of consensus. Since the anesthesiologist is the best placed for perioperative care, an anesthesia-tailored preoperative frailty test must be simple, quick, universally applicable to all surgeries, accurate, and ideally available in an app or online form. This systematic review attempted to rank frailty tests by predictive accuracy using the c-statistic in the outcomes of extended length of stay, 3-month post-operative complications, and 3-month mortality, as well as feasibility outcomes including time to completion, equipment and training requirements, cost, and database compatibility. Presenting findings of all frailty tests as a future reference for anesthesiologists, Clinical Frailty Scale was found to have the best combination of accuracy and feasibility for mortality with speed of completion and phone app availability; Edmonton Frailty Scale had the best accuracy for post-operative complications with opportunity for self-reporting. Finally, extended length of stay had too little data for recommendation of a frailty test. This review also demonstrated the need for changing research emphasis from odds ratios to metrics that measure the accuracy of a test itself, such as the c-statistic.

Key words: Anesthesia, elderly, frailty test, post-operative complications, surgery

Introduction

Frailty is a syndrome of reduced physiological reserve that is present in 20% of patients undergoing emergency laparotomies aged 65 and above.[1,2] A recent systemic review has identified frailty as the strongest risk factor for developing post-operative morbidity in older patients.[3] Major stresses, such as surgery, temporarily decrease physiological reserves, meaning that the combination of frailty and surgery can result in significant mortality and morbidity. A diagnosis of frailty

Access this article online	
Website: https://journals.lww.com/sjan	**Quick Response Code**
DOI: 10.4103/sja.sja_358_23	

How to cite this article: Dunlop RA, Van Zundert A. A systematic review of predictive accuracy via c-statistic of preoperative frailty tests for extended length of stay, post-operative complications, and mortality. Saudi J Anaesth 2023;17:575-88.

Richard A. N. Dunlop, André Van Zundert

Department of Anaesthesia and Perioperative Medicine, Royal Brisbane and Women's Hospital and The University of Queensland, Brisbane, QLD, Australia

Address for correspondence: Prof. André Van Zundert, Lennard Travers Professor of Anaesthesia - Australian and New Zealand College of Anaesthetists, Honorary Fellow Royal College of Anaesthetists - London UK; Professor and Chairman Discipline of Anesthesiology, The University of Queensland, Faculty of Medicine and Biomedical Sciences, Brisbane, QLD, Australia; Chair, University of Queensland "Burns, Trauma and Critical Care Research Centre"; Chair, RBWH/University of Queensland "Centre for Excellence and Innovation in Anaesthesia"; Department of Anaesthesia and Perioperative Medicine, Royal Brisbane and Women's Hospital, Ned Hanlon Building L4 - Herston Campus-Brisbane, Queensland 4029, Australia. E-mail: a.vanzundert@uq.edu.au

Submitted: 01-May-2023, **Revised:** 03-May-2023, **Accepted:** 04-May-2023, **Published:** 18-Aug-2023

can increase 90-day post-operative mortality by a factor of 3.18.[2] Additionally, it is well-known that age is directly associated with the severity of frailty. Indeed, geriatric people may make up almost a quarter of the population by 2060 in the United States with more than 50% of this population requiring at least one surgery in their lives.[4,5] As such, frailty poses a significant threat to patients and the health systems of nations.

To combat this, surgical and anesthetic international societies have recommended preoperative frailty assessment since 2012.[6,7] However, their use in practice has been hampered by the sheer number of frailty tests available (a previous study found 35 alone) and the lack of census for which of these to use.[8,9] Currently, there exist three predominant models of frailty [Table 1]. There are two modalities of frailty assessment: clinical, where the assessor examines the patient in-person, and administrative, where hospital database information can be used to calculate a score. The closest to a gold standard for frailty assessment is the comprehensive geriatric assessment (CGA), a multidisciplinary process assessing the domains of multimorbidity, polypharmacy, nutrition, mobility, physiologic function/reserves, neurocognition, and psychological health to identify and manage these risk factors.[10] The CGA has already been employed in perioperative settings with great success, decreasing morbidity and mortality.[11,12] However, in the perioperative setting, the CGA can be an unwieldy program that is time-consuming and requires geriatrician expertise not commonly available in surgical teams.[13] As such, a frailty test needs to be tailored for the perioperative environment; tailored for the anesthesiologist who is best placed to accompany the patient throughout the entire perioperative journey. The ideal frailty test for the anesthesiologist needs to be feasible (able to be completed quickly with little extra training or equipment); universal (able to be applied to any surgical population); and accurate (able to correctly classify frail patients and predict post-operative outcomes). Finally, with the dawn of digital medicine, another desirable trait is digital interface of frailty tests, such as completion via an app on the phone, as well as easy online accessibility for physicians.

Table 1: The three most popular models of frailty according to the literature

Model	Definition	Archetypal Test
Phenotype of Frailty[14]	A disease-like syndrome consisting of energy depletion and inflammation, which exhibits itself as "weakness, decreased endurance, and slow performance."	Fried's Phenotype of Frailty
Accumulation of Deficits[15]	The accumulation of disabilities and conditions with emphasis on the number rather than the nature of the deficits.	Frailty Index
Multidimensional[16]	A dynamic state of loss affecting 1 or more areas of functioning such as the cognitive, physical, and social domains.	Comprehensive Geriatric Assessment

A current survey of the literature demonstrates an emphasis on feasibility because the use of odds ratios makes differentiation of predictive accuracy difficult. McIsaac *et al.*[3] commented that, despite only moderate agreement between frailty tests (Cohen's kappa = 0.1–0.8), many studies had found no difference in effect sizes for length of stay, post-operative complications and mortality. Indeed, odds ratios assess prevalence of an event in a population rather than the predictive accuracy of a test itself, the most commonly reported of such a metric being the c-statistic.[17] By appealing to feasibility, most reviews and guidelines have recommended the Clinical Frailty Scale for preoperative assessment.[18,19] In contrast, it is the aim of this systematic review to rank preoperative frailty tests according to their predictive accuracy, in the form of the c-statistic, as well as feasibility.

Materials and Methods

Search Strategy

Search terms were derived from initial scoping of previous systematic reviews covering preoperative frailty tests.[18,20-22] The search method was applied to Medline and EMBASE databases from inception to March 10, 2023. A summary of the search strategy has been included [Supplementary Table 1]. Reference lists of related systematic reviews and primary articles discovered in systematic search were also inquired for other studies not covered by the search method. No language restrictions were applied.

Study selection

Eligible studies were included if they: (1) studied a surgical population with a mean or median age greater than 60 years; (2) included a frailty instrument explicitly described or used according to its original publication and its result recorded before the surgery; (3) reported a predictive accuracy outcome in the form of the c-statistic for length of stay, 3-month or less post-operative complication or mortality.

Predictive accuracy of 3-month or less post-operative mortality was the primary outcome and 3-month or less post-operative complications, as defined by greater than or equal to grade 2 on the Clavien-Dindo classification model, and extended length of stay, as defined by a greater than 75th percentile length of stay, were secondaries.[23] Other secondary outcomes included feasibility parameters: completion time, equipment, training, database compatibility, and cost for frailty tests, which were recorded from original publications of frailty tests.

Studies were excluded if they: (1) included mixed populations with less than 50% of patients undergoing surgery; (2) included samples with greater than 50% of patients undergoing cardiac or major thoracic and abdominal vascular surgery (since frailty has a larger influence on post-operative outcome in these surgeries); (3) included samples with greater than 50% of patients with a cancer diagnosis or undergoing surgery specifically for cancer resection; (4) determined frailty by the CGA (since this is inappropriate for the perioperative environment); (5) determined frailty by a single laboratory or imaging technique (e.g., ultrasound scan for sarcopenia); (6) determined frailty using a score specific to a surgical subpopulation (e.g., Nottingham Hip Fracture Score). Conference abstracts or other grey literature were not included due to incomplete descriptions of methodology.

Data extraction and quality assessment

Screening of papers was conducted first by title and abstract and then by full text using Covidence. Removal of duplicate articles was done automatically by Covidence as well as manually by screeners. Data extracted included basic study and study population parameters and primary and secondary outcomes as above. Updated versions of frailty tests, such as modified frailty index 11-item and 5-item, were combined into one frailty test for analysis. Risk of bias was analyzed using the Quality in Prognosis Studies tool.[24]

Data analysis

The c-statistic is a measure of the discriminatory power of a predictive model calculated from the area under the receiver operating characteristic curve, which can be summarized as: "the proportion of all pairs of patients where one patient experienced the event of interest and the other patient did not experience the event, and the patient with the lower risk score was the one who did not experience the event."[25] For use in comparing predictive accuracy of frailty tests in this review, a c-statistic of 0.80 and above was defined as excellent predictive accuracy; 0.70 and above as good predictive accuracy; 0.60 and above as fair accuracy; 0.50 and above as poor accuracy.

The finding of the best frailty tests involved rounds of elimination based on desirable properties: (1) frailty tests must have data from at least three studies for predicting a single outcome (length of stay, post-operative complication or mortality); (2) total number of studies for predicting an outcome by a frailty test must not be composed by more than 50% of the same surgery type; (3) the frailty test must not take more than 5 minutes to complete. The remaining frailty tests were then ordered by their mean c-statistic in each outcome, and the 5 best were chosen for comparison.

Results

A total of 1772 records were screened after 590 duplicates were removed [Figure 1]. 304 full-text articles were assessed, and 35 studies were included from the systematic search. A further 17 studies were included after analyzing citations of reviews and primary articles. Thus, 52 studies in total have been included. Overall, included studies consist of 2,168,912 participants and were published between 2008 and 2023 with 2022 being the most common year of publication. A full summary of studies has been included [Table 2].

Surgical and patient populations

Both abdominal and orthopedic surgical patients were the most studied populations (14 studies each [27%]), followed by mixed surgical patients (13 studies [25%]). Elective was the most studied surgical urgency population (24 studies [46%]), followed by emergency (20 studies [38%]) and mixed (8 studies [15%]). The average study population age ranged

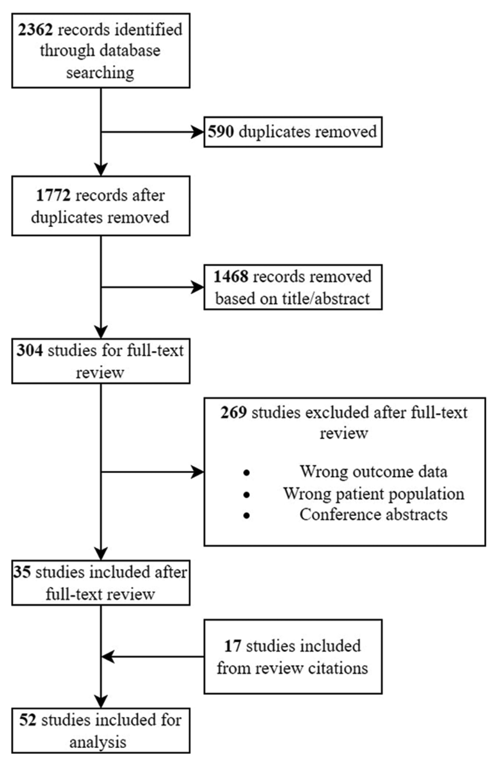


Figure 1: Preferred reporting items for systematic reviews and meta-analyses (PRISMA) flow diagram for inclusion and exclusion of papers

Table 2: Summary of 52 included studies for assessment of predictive accuracy of frailty tests. Where there were multiple frailty measurements, the lowest prevalence was used

Study	Study design	Study Size, *n*	Mean Age, Year	Sex, % Female	Frail patient, %	Surgical procedure	Procedure urgency	Frailty Measure
2008 Burgos[26]	P	232	85	85	-	Hip	EM	Barthel, CCI
2008 Dasgupta[27]	P	125	77.4	58	12.8	Non-cardiac	EL	EFS
2013 Robinson 1[28]	P	201	74	2	33.3	Colorectal	EL	Robinson
2013 Robinson 2[29]	P	98	74	4	26.5	Colorectal	EL	TUGT
2015 Kenig[30]	P	184	76.9	53.2	50	Abdominal	EM	BFI, G8, GFI, RFT, VES
2015 Revenig[31]	P	351	63	39	27.3	Abdominal	M	FP
2017 Hall[32]	P	1021	60.2	4.2	31	Mixed	EL	RAI-A, RAI-C, mFI-11
2017 Kapoor[33]	P	403	72	48.9	18	Mixed	EL	FP, LLFDI-F
2018 Gilbert[34]	R	1013590	84.1	57.4	57.6	Mixed	EM	HFRS
2018 Han[35]	P	176	69.5	53.4	23.7	Abdominal	EL	FP
2018 Kenig[36]	P	315	77	52.4	60.3	Abdominal	EM	G8
2018 Ondeck 1[37]	R	68580	65.1	55.8	3.9	THA	EL	CCI, Elixhauser, mFI-11
2018 Ondeck 2[38]	R	49738	82	72.3	5	Hip fracture	EM	CCI, Elixhauser, mFI-11
2018 Zattoni[39]	P	556	81	57.3	29	Abdominal	EM	CCI, TRST
2019 Al-Hamis[40]	R	295490	61	52	18	Colorectal	EL	mFI-5
2019 Amin[41]	R	158855	64.6	22.9	16.9	Urological	M	mFI-5, mFI-11, RAI-A
2019 Fu[42]	R	10527	69.2	56.4	2.5	Shoulder	EL	CCI, mFI-11
2019 Katlic[43]	P	513	80.5	63.7	47.4	Mixed	EL	CCI, FP
2019 Lima[44]	P	229	69	55	-	Mixed	EL	CFS
2020 Arya[45]	R	6856	60.7	3.6	19.9	Non-cardiac	EL	RAI-C, RAI-C Rev
2020 Barazanchi[46]	R	758	62	50.1	-	Laparotomy	EM	mFI-11
2020 Choi[47]	R	648	76.6	52.8	-	Mixed	EL	6-min walk, MFS
2020 He[48]	P	134	76.9	50	29.1	Abdominal	M	EFS
2020 Lu[49]	P	136	77.5	67	36.8	Hip fracture	EM	FI
2020 McIsaac[50]	P	645	74	49.8	36.6	Non-cardiac	EL	CFS, FI
2020 Rogozinski[51]	R	451	65.1	7.1	-	THA, TKA	EL	CCI, mFI-11
2020 Roopsawang[52]	P	200	72	78	43	Orthopedic	EL	Self-reported EFS
2021 Aguilar-Frasco[53]	P	140	72.7	47.1	35	Abdominal	EL	RFI
2021 Arteaga[54]	P	92	78.7	53.3	14.1	Abdominal	EM	FRAIL, FI, TRST, CFS
2021 Costa[55]	P	240	77.6	47.9	-	Abdominal	EM	EmSFI
2021 Lee[56]	R	4664	80	40.5	47.6	Mixed	EM	OFRS
2021 Pandit[57]	R	8681	76	32	24.5	LEA	EL	mFI-5, mFI-11
2021 Tse[58]	R	47197	66	1.1	-	LEA	M	RAI-A Rev
2021 Wu[59]	R	397	83.5	63	90	Hip fracture	EM	CCI, CFS, KPS
2021 Yi[60]	R	3893	68	77.6	24.8	Shoulder	EM	CCI, mFI-5
2021 Yin[61]	P	194	79	53.6	32.5	Abdominal	EL	CFS, FI, FRAIL
2022 Conlon[62]	R	6571	64	42.7	5.4	Spine	EM	mFI-5, RAI-A, RAI-A Rev
2022 Cotton[63]	R	298	67	0	65.8	LEA	EL	mFI-11, RAI-C
2022 Forssten[64]	R	2365	84	67.7	47	Hip	EM	CCI, mFI-5
2022 Ikram[65]	P	1577	83.6	56.4	44.3	Hip	EM	CFS
2022 Iwasaki[66]	R	476	92.4	64.3	32.5	Non-cardiac	M	ECOG-PS, mFI-5
2022 Kweh[67]	R	272	73.5	45.6	20.6	Spinal	M	mFI-5, mFI-11
2022 Le[68]	R	37186	67.9	51.4	20.2	Abdominal	M	FI, HFRS, mFI-5, RAI-A
2022 Lee[69]	R	1557	80.4	60.2	14.9	Mixed	EM	CCI, HFRS, OFRS
2022 Li[70]	R	923	73.5	37.6	24.4	GI	EL	CRI, mFI-11
2022 Palaniappan[71]	R	1434	65	51	10.6	Abdominal	EM	CFS
2022 Ruiz[72]	P	100	61.3	51	-	Abdominal	EM	UEF
2022 Wei[73]	R	4195	73.9	38.6	-	Abdominal	M	RAI-A Rev
2022 Yin[74]	P	194	77	53.6	37.6	Abdominal	EL	CFS, FI, FRAIL
2023 Darbyshire[75]	R	1508	66	54.1	-	Bowel	EM	HFRS

Contd...

Table 2: Contd...

Study	Study design	Study Size, *n*	Mean Age, Year	Sex, % Female	Frail patient, %	Surgical procedure	Procedure urgency	Frailty Measure
2023 McConaghy[76]	R	433311	>60	55.5	-	THA, TKA	EL	CCI, Elixhauser, mFI-5
2023 Sirisegaram[77]	R	535	72	40.2	21.1	Mixed	EL	EFS

P, prospective study; R, retrospective study; –, information not available; EL, Elective; EM, Emergency; M, Mixed; BFI, Balducci Frailty Index; CCI, Charlson Comorbidity Index; CFS, Clinical Frailty Scale; CRI, Composite Risk Index; ECOG-PS, Eastern Cooperative Oncology Group performance status; EFS, Edmonton Frailty Scale; EmSFI, Emergency Surgery Frailty Index; FI, Frailty Index; FP, Fried's Phenotype; G8, Geriatric 8; GFI, Groningen Frailty Indicator; HFRS, Hospital Frailty Risk Score; KPS, Karnofsky Performance Status; LEA, Lower Extremity Amputation; LLFDI, Late-Life Function and Disability Instrument; mFI-5, 5-Item Modified Frailty Index; mFI-11, 11-Item Modified Frailty Index; MFS, Multidimensional Frailty Score; OFRS, Operation Frailty Risk Score; RAI-A, Risk Analysis Index – Administrative; RAI-C, Risk Analysis Index – Clinical; RFT, Rockwood Frailty Test; THA, Total Hip Arthroplasty; TKA, Total Knee Arthroplasty; TRST, Triage Risk Screening Tool; TUGT, Timed-up-and-go test; UEF, Upper Extremity Function; VES, 13-Item Vulnerable Elders Survey; Rev, Revised

from 60.2 to 92.4 years, and the proportion of female patients ranged from 0% to 98% with the averages being 73.2 years and 48.7%, respectively. Frailty prevalence ranged from 2.5% to 90% with the average being 30.5%.

Frailty tests

Our review revealed twenty-nine unique frailty tests [Table 3]. The most prevalent was the Modified Frailty Index (30 data from 18 studies), followed by the Charlson Comorbidity Index (23 data from 12 studies), Frailty Index and Clinical Frailty Scale (9 data from 5 and 7 studies, respectively). The most common model of frailty used was the accumulation of deficits (13 frailty tests) followed by multidimensional (11 frailty tests) and phenotype of frailty (5 frailty tests). Seven frailty tests were database compatible, all of which used the accumulation of deficits model. Three frailty tests (Clinical Frailty Scale, Phenotype of Frailty and Upper Extremity Function) required training, two of which were due to the use of specialist equipment. Three multidimensional frailty tests (Composite Risk Index, Robinson Frailty Test and Multidimensional Frailty Score) required the use of imaging or blood tests. No tests used proprietary content.

Length of stay

A total of 24 data from ten frailty tests were found for predicting length of stay. The Modified Frailty Index was the frailty test with most data (8 values) followed by Charlson Comorbidity Index (7 values) and Elixhauser Comorbidity Index (3 values). Overall, predictive ability ranged from 0.50 to 0.88 with underreported frailty tests like Clinical Frailty Scale and FRAIL scale reporting excellent discrimination [Table 3.19: 0.88; Table 3.1: 0.81]. The three frailty tests with the most data all had fair discrimination [Table 3.12: 0.61; Table 3.14: 0.61; Table 3.15: 0.66]. Notably, these three tests are all of the accumulation of deficits frailty model and are database compatible. Although these three tests are automatically calculated and do not take time, Elixhauser Comorbidity Index required the most data followed by Charlson Comorbidity Index and Modified Frailty Index [Table 3.15: 31; Table 3.14: 17; Table 3.12: 5]. None of these tests have extra costs. The Modified Frailty Index has the best ratio of data required to accuracy for extended length of stay.

Post-operative complications

A total of 53 data from 25 frailty tests were found for predicting post-operative complications. The Modified Frailty Index was the frailty test with the most data (10 values) followed by the administrative Risk Analysis Index and Charlson Comorbidity Index (both 6 values), and the Phenotype of Frailty (4 values). Overall, predictive ability ranged from 0.52 to 0.88 with most underreported tests exhibiting low to fair predictive accuracy. The most reported tests had fair predictive abilities [Table 3.12: 0.65; Table 3.13: 0.61; Table 3.14: 0.60]. All frailty tests with ≥3 data were either automatically calculated via database with under 20 data items required or took less than 5 minutes to complete. Edmonton Frailty Scale was included among these with good discrimination [Table 3.24: 0.73]. The exceptions were: Frailty Index has an extensive item requirement of 40 but with good discrimination; Phenotype of Frailty has a time requirement of 15 minutes but with only fair discrimination [Table 3.17: 0.71; Table 3.5: 0.67]. Phenotype of Frailty also requires training. None of these tests have extra costs. Edmonton Frailty Scale has the best ratio of time required to accuracy for post-operative complications.

Post-operative Mortality

A total of 71 data over 23 frailty tests were found for predicting post-operative mortality. The Modified Frailty Index was the frailty test with the most data (12 values) followed by the Charlson Comorbidity Index (10 values), and the administrative Risk Analysis Index and Clinical Frailty Scale (both 7 values). Overall, predictive ability ranged from 0.52 to 0.98 with most underreported tests exhibiting poor to good predictive accuracy; some exceptions such as Eastern Cooperative Oncology Group Performance score and Karnofsky's Index of Performance Status had excellent discrimination [Table 3.20: 0.98; Table 3.21: 0.82]. The top three most accurate frailty tests with ≥3 data were Frailty Index, Clinical Frailty Scale, and administrative Risk Analysis Index [Table 3.17: 0.80; Table 3.19: 0.77; Table 3.13: 0.76].

Table 3: A summary of 29 frailty tests and their usefulness in preoperative screening. Accuracy of predicting post-operative length of stay, complication, and mortality is represented by the mean c-statistic and its range. Mortality and complication outcomes are measured within 3 months or less of surgery. Electronic database compatibility is defined as being based on International Classification of Diseases diagnosis codes or National Surgical Quality Improvement Program variables. Electronic database scores were not recorded as having any time since they are not used manually. Cost is defined as using any proprietary tests or the use of blood tests or imaging

Test Number	Test Name	Duration (min)	Items, *n*	Extended Length of Stay			Increased Post-Operative Complications			Increased Mortality			Database use	Training	Equipment	Cost
				Mean	Range	*n*	Mean	Range	*n*	Mean	Range	*n*				
Phenotype of Frailty																
1	FRAIL[78]	<5	5	0.81			0.68			0.80	(0.72-0.87)	2				
2	TUGT[79]	5	1	-			0.78			-						
3	UEF[80]	5	6	-			-			0.74				X	X	
4	6MWT[81]	10	1	-			0.67			-						
5	PF[14]	15	5	0.74			0.67	(0.60-0.76)	4	0.75	(0.68-0.81)	2		X	X	
n = 5 tests																
Accumulation of deficits																
6	Bal[82]	<5	4	-			0.55			0.56						
7	TRSI[83]	<5	5	-			0.61	(0.52-0.69)	2	0.65	(0.52-0.72)	3				
8	RAI-C[32]	<5	14	-			0.64			0.74	(0.70-0.82)	4				
9	ESFI[55]	5	9	-			0.54			0.76						
10	Bar[84]	5	10	-			0.67			0.69						
11	VES[85]	5	13	-			0.68			0.64						
12	MFI[86]	-	5	0.61	(0.53-0.76)	8	0.65	(0.52-0.88)	10	0.74	(0.59-0.96)	12	X			
13	RAI-A[32]	-	14	0.67			0.61	(0.57-0.66)	6	0.76	(0.60-0.98)	6	X			
14	CCI[87]	-	17	0.61	(0.50-0.72)	7	0.60	(0.52-0.71)	6	0.70	(0.59-0.82)	7	X			
15	ECI[88]	-	31	0.66	(0.56-0.74)	3	0.57			0.62	(0.62-0.62)	2	X			
16	FI[89]	-	40	0.85			0.71	(0.64-0.75)	3	0.80	(0.73-0.85)	5	X			
17	HFRS[34]	-	109	0.68			0.71			0.67	(0.55-0.82)	4	X			
18	OFRS[56]	-	111	-			-			0.72	(0.62-0.81)	2	X			
n = 13 tests																
Multidimensional																
19	CFS[90]	<5	1	0.88			0.75			0.77	(0.63-0.91)	7		X		
20	ECOG[91]	<5	1	-			-			0.98						
21	KIPS[92]	<5	1	-			-			0.82						
22	RFI[93]	<5	1	-			0.52			0.57						
23	G8[94]	<5	8	-			0.70	(0.56-0.83)	2	0.75	(0.57-0.92)	2				
24	EFS[95]	<5	11	-			0.73	(0.69-0.81)	3	0.75						
25	LLFDI[96]	<5	40	-			0.67			-						
26	GFI[97]	5	15	-			0.60			0.58						
27	RFS[28]	15	7	-			0.70			-				X	X	X
28	CRI[70]	20	8	-			0.65			-				X	X	X
29	MFS[98]	20	9	-			0.75			-				X	X	X
n = 11 tests																

–, information not available; 6MWT, 6-minute walk test; Bal, Balducci frailty test; Bar, Barthel frailty test; CCI, Charlson Comorbidity Index; CFS, Clinical Frailty Scale; CRI, Composite Risk Index; ECOG-PS, Eastern Cooperative Oncology Group performance status; EFS, Edmonton Frailty Scale; EmSFI, Emergency Surgery Frailty Index; FI, Frailty Index; FP, Fried's Phenotype; G8, Geriatric 8; GFI, Groningen Frailty Indicator; HFRS, Hospital Frailty Risk Score; KPS, Karnofsky Performance Status; LEA, Lower Extremity Amputation; LLFDI, Late-Life Function and Disability Instrument; mFI-5, 5-Item Modified Frailty Index; mFI-11, 11-Item Modified Frailty Index; MFS, Multidimensional Frailty Score; OFRS, Operation Frailty Risk Score; RAI-A, Risk Analysis Index – Administrative; RAI-C, Risk Analysis Index – Clinical; RFT, Rockwood Frailty Test; TRST, Triage Risk Screening Tool; TUGT, Timed-up-and-go test; UEF, Upper Extremity Function; VES, Vulnerable Elders Survey

Again, Frailty Index suffers from requiring 40 items compared to the others which take less than 5 mins or have fewer than 20 items. However, Clinical Frailty Scale also requires training. None of these tests have extra costs. Clinical Frailty Scale proved to have the best ratios of time to accuracy for mortality.

Assessment of best frailty tests

Of a total of 29 frailty tests, only six were included in the final assessment after excluding undesirable findings and properties [Figure 2]. 18 of the tests were excluded due to insufficient data for predicting either length of stay, post-operative complication or mortality. Four frailty tests were also rejected due to overrepresentation of specific surgical populations. Notably, Charlson and Elixhauser Comorbidity indices overrepresented orthopedics. Finally, the Phenotype of Frailty was excluded due to it taking ≥5 minutes to complete.

Of the remaining six tests, five were included in the final ranking [Table 4]. Forms used to complete these frailty tests have also been included [Supplementary Table 2.1-5] The excluded frailty test, Hospital Frailty Risk Scale, had a predictive ability for mortality far below the remaining 5 (0.67) and not enough data to predict length of stay or post-operative complication. The top five include three clinical frailty tests (Clinical Frailty Scale, clinical Risk Assessment Index, and Edmonton Frailty Scale) and two administrative frailty tests (Modified Frailty Index and administrative Risk Assessment Index). Overall, the best frailty test for predicting increased mortality and increased post-operative complications was the Clinical Frailty Scale and Edmonton Frailty Scale, respectively. Both had good predictive ability (0.77 and 0.73). Only the Modified Frailty Scale had enough data for predicting extended length of stay, the ability of which was found to be poor (0.61).

All of the top five frailty tests are very quick to perform either being automatically calculated from a database using software or taking ≤5 minutes in preoperative clinic. However, database tests with numbers of items greater than ten, such as the administrative Risk Analysis Index, can struggle with missing data. Only the Edmonton Frailty Scale has been validated for self-reporting and only the Clinical Frailty Scale has an app available for tablet or phone. Although the frailty test recommends training, which is freely available and short, the Clinical Frailty Scale combines the best predictive ability with the best feasibility of fastest time to complete and phone app availability.

Risk of bias analysis

Risk of bias results according to the Quality in Prognosis Studies tool has been reported [Supplementary Table 3].

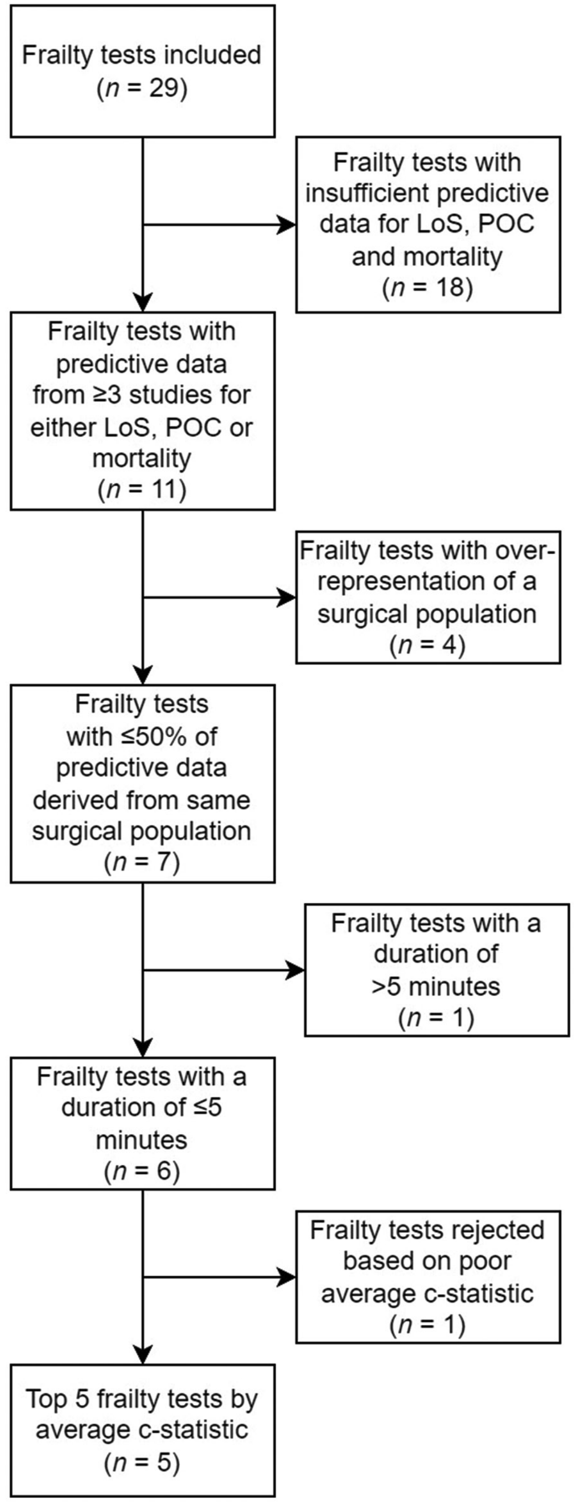


Figure 2: Flowchart of process for selection of top five frailty tests. Process has been outlined in methods. LoS, length of stay; POC, post-operative complications

A major contributor to high risk of bias was not reporting confounding factors like duration and stress of surgery and method of anesthesia and not accounting for these confounding factors in analysis. Poor reporting of missing

Table 4: Summary of top five frailty tests with average c-statistic. Average c-statistics with ranges are derived from Table 3. All tests take no longer than 5 minutes to complete or are derived automatically from administrative data. All tests have been assessed in diverse surgical populations. Assessor accessibility is a measure of the ability for populations to be able to complete the test

Ranking	Test Name	Extended Length of Stay	Increased Post-operative Complications	Increased Mortality	Assessor Accessibility			Advantages	Disadvantages
					Doctor	Nurse	Patient		
Clinical Frailty Tests									
1	CFS	-	-	0.77 (0.63-0.91)	X	X		Very quick to complete (44 s on average).[50] Visual chart available for reference. Online training module available that takes <30 minutes.[99] App available for phone or tablet.[100]	Concerns over inter-rater reliability (0.78-0.79).[101,102] Training recommended. Lacking self-reporting data. Likely difficult due to training recommended.
2	RAI-C	-	-	0.74 (0.70-0.82)	X	X		Form available for completion. Recent revision completed.[45]	Lacking self-reporting data.
3	EFS	-	0.73 (0.69-0.81)	-	X	X	X	Form available for completion. Validated for self-reporting.[77] Includes nutrition and cognitive components.	
Administrative Frailty Tests									
1	MFI	0.61 (0.53-0.76)	0.65 (0.52-0.88)	0.74 (0.59-0.96)	X	X		Automatically generated from hospital database using NSQIP codes.	Requires hospital database.
2	RAI-A	-	0.61 (0.57-0.66)	0.76 (0.60-0.98)	X	X		Automatically generated from hospital database using NSQIP codes. Recent revision completed.[45]	Requires hospital database. Requires more items for calculation. More likely to have missing data.

CFS, Clinical Frailty Scale; EFS, Edmonton Frailty Scale; MFS, Modified Frailty Index; RAI-A, Risk Assessment Index – Administrative; RAI-C, Risk Assessment Index – Clinical preoperative frailty test data also commonly contributed to high risk of bias. The removal of high risk of bias studies from the ranking analysis [see Supplementary Table 3 in red] did not change the ranking of preoperative frailty tests.

Discussion

In this systematic review that compared predictive accuracy of preoperative frailty tests using c-statistic values from 52 studies, we found that: (a) Clinical Frailty Scale combined the best predictive accuracy for mortality with the best feasibility thanks to speed of completion and phone app availability; (b) Edmonton Frailty Scale was the best predictor for post-operative complications and potentially has similarly excellent feasibility due to its validated ability for self-reporting; (c) at this time, we would not recommend any frailty test for predicting extended length of stay due to poor accuracy of tests and lacking data; (d) Modified Frailty Index was found to be the best administrative frailty test overall but administrative Risk Analysis Index outperforms it in mortality prediction though it requires far more data items, which predisposes it to missing data.

Previously, the choice of frailty test had been obscured using odds ratios, from which no statistical difference could be detected between frailty tests despite poor-to-modest agreement between them.[3] In our analysis of predictive accuracy, we were able to rank tests by their average c-statistic, but not able to calculate statistical differences due to the small number of studies. Moreover, the small differences in average c-statistic are unlikely to translate into any observable difference in predictive ability in practice. As such, despite the use of the c-statistic, we have found similar results to other reviews such as Aucoin *et al.*,[18] who found that, whilst the Clinical Frailty Scale has the highest odds ratio of 4.89 for mortality, it was not statistically different from other frailty tests. In head-to-head cohort studies comparing predictive ability of different frailty tests, the Clinical Frailty Scale is also consistently better than other frailty tests, but only by differences in c-statistic of 0.01–0.02, which have little observable clinical effect.[103] Indeed, despite the use of the c-statistic, our ranking of frailty tests is still based on feasibility, for which the Clinical Frailty Scale excels at. However, if more studies were available presenting c-statistics and their confidence intervals, statistical differences may be able to be calculated. Especially now that the effects of frailty on post-operative outcomes are well established via odds ratios, the emphasis of research needs to be on predictive accuracy of frailty tests so that such statistical tests can be done in future reviews.

Another large contributing factor to difficulty choosing frailty tests for preoperative screening is the explosion of frailty tests designed for different surgical populations. Tests such as the Nottingham Hip Fracture Score and Addenbrookes Vascular Frailty Score, which include significant frailty components, have been designed for predicting post-operative complications in orthopedic and vascular populations.[104,105] Despite the increased accuracy that these may provide, we believe they are not feasible for the anesthesiologist working mixed caseloads. The widespread use of the American Society of Anesthesiologists physical status score can be partly attributed to its use in all surgery types.[106] Indeed, rather than having very many frailty tests for each surgical population, it is more prudent to have a single frailty test, which has a component that quantifies the risks of different surgeries. We recommend surgery-specific frailty tests be reserved for surgeons.

If the Clinical Frailty Scale is found to be the most accurate frailty test, the issue of bias due to its judgment component needs to be addressed. Although the scale's training program emphasizes that physicians look out for leniency and central tendency bias effects, two separate studies have reported good inter-rater reliability for the Clinical Frailty Scale, finding kappa values between 0.74 and 0.85 after standardized training.[101,102] In context, the inter-rater reliability of the American Society of Anesthesiologists physical status score has been found to be 0.40, 0.61, and 0.21–0.4 in different studies, suggesting that the Clinical Frailty Scale may be more reliable than one of the most widely used preoperative risk scores.[107-109]

The future of frailty tests should take advantage of the multidimensional nature of frailty, as exemplified by the many categories of the CGA, so that the accuracy of the Clinical Frailty Scale is further improved. Some tools found in this systematic review have already tried to combine multiple techniques to increase risk prediction but lacked sufficient data to be properly analyzed [Table 2.27-29]. While we need to balance feasibility, it is possible and desirable for a frailty test to include other dimensions like nutrition and cognition, both of which in poor condition can increase post-operative mortality by 3.86 in hip fracture surgery and 1.6 in any elective surgery, respectively.[110,111] Edmonton Frailty Scale already includes the clock drawing test for cognition and a basic screening question for nutrition, but such tests need to be further refined, especially for difficult to detect conditions like mild cognitive impairment, which contributes significantly to post-operative complications like post-operative delirium.[112] The ideal preoperative assessment would be a "mini-CGA," which is able to combine all dimensions into a single score that can predict surgical risk.

Other possible additions to this mini-CGA could be emerging imaging and blood biomarkers for frailty. Sarcopenia, the loss of muscle mass related to age, which commonly coincides with frailty, can be measured with computed tomography or ultrasound scans.[113] The measurement of quadriceps depth with ultrasound, which suffers from requiring a trained sonographer, can accurately predict post-operative delirium with a c-statistic of 0.89.[114] Turning to blood biomarkers, serum albumin, which is very commonly measured, is a composite of nutrition and liver and kidney condition and its preoperative to post-operative change has been found to predict post-operative complication.[115] Newer frailty biomarkers, like interleukin-6 and alpha-1-acid glycoprotein, may have better accuracy, a recent trial finding a 0.781 c-statistic for both predicting morbidity, but may not be available in most pathology labs.[116] Other experimental neurological biomarkers such as neurofilament light chain and glial fibrillary acidic protein may be more useful for anticipating post-operative delirium and cognitive dysfunction.[117,118] All such additions would significantly improve post-operative risk prediction but may have feasibility issues.

The problems of feasibility associated with further additions to different dimensions to the mini-CGA could be attenuated by allowing patients to rate their own frailty, which could be accommodated using digital apps. The Edmonton Frailty Scale was recently validated as a self-reporting tool if the timed-up-and-go and clock drawing tests are removed.[77] This missing data could be collected by phone apps: walking pace, which has strong agreement with the timed-up-and-go, can be tracked via accelerometer; and many apps exist to test cognition such as the Cogstate Brief Battery, which has better sensitivity and specificity for detecting mild cognitive impairment than the clock drawing test.[119-121] Instruction to use these apps could be provided before preoperative consultation so that the data may be used for frailty assessment. Current barriers include the proprietary nature of cognitive testing phone apps and technological illiteracy in elderly populations.[122] Nevertheless, such additions could completely revolutionize feasibility and accuracy.

Finally, an area of preoperative assessment that needs significant accuracy improvement is prediction of extended length of stay. In this review, extended length of stay had less than half the data than post-operative complications and mortality and had poor accuracy in frailty tests with enough data for assessment. A common cause of extended length of stay is low severity post-operative complications.[123] According to Mah *et al.*,[124] the Modified Frailty Index may be very good at predicting Clavien-Dindo grade 3–5 complications (0.92) but inclusion of grade 2 reduces accuracy significantly (0.74). A possible explanation for this is that frailty may have poor association with minor post-operative complications. However, since some of the best frailty tests, like Clinical Frailty Scale, have little to no extended length of stay data, this cannot be completely verified. As such, more research should be focused on assessing preoperative frailty screening tools, especially clinical frailty tests, for predicting extended length of stay.

Strengths and Limitations

This systematic review is the first of its kind to assess predictive accuracy of post-operative outcomes via the c-statistic as its primary outcome. This allows a quantification of accuracy of specific tests without the interference of outcome prevalence in a sample, which is the problem with odds ratios.[17] To avoid interference via confounding factors, this review had well-defined inclusion criteria, which excluded surgeries where frailty is more likely to influence post-operative outcomes such as cardiac and major thoracic and abdominal vascular surgery. We also assessed feasibility and displayed important test properties for the reference of anesthesiologists. Where possible, we followed the systematic review best practice by including risk of bias assessment, PRISMA and displaying of our complete search methodology.

While we did attempt to measure predictive accuracy, other insightful statistical metrics exist such as predictive values, likelihood ratios and calibration.[125] Indeed, similar c-statistics may not necessarily have comparable positive predictive values. Lack of available data made use of these metrics unfeasible for review at this time. While we can rank accuracy based on the average c-statistic, many tests differed by small values, which may have little effect in practice. Finally, since the average was used, there was no analysis of heterogeneity, and all studies were weighted the same. It can be predicted that, due to the wide variety of surgical types and different cutoffs for frailty tests, heterogeneity may be high.

Additionally, feasibility of frailty tests in the emergency surgery setting was not explored, which poses unique challenges with timing and communication with patients of reduced consciousness. A systematic review of feasibility in acute trauma suggests that the Clinical Frailty and FRAIL scales can be completed in between 71% and 100% and 62% and 100% of cases, respectively.[126] Problems may arise in longer frailty tests such as the Frailty Index, which was found to have a completion rate of only 31.9%.[127] Additionally, institutionalization, a metric of patient post-operative quality of life, was not assessed. Despite a finding of 22% of elderly patients being institutionalized after abdominal surgery, this metric had poor data availability during initial scoping research and requires further exploration.[128]

Conclusions

This is the first systematic review to rank preoperative frailty tests according to a metric of predictive accuracy in addition to their feasibility. Clinical Frailty Scale was found to be the best for predicting mortality; this, alongside its standout time efficiency and phone app availability, made it the definitive preoperative frailty test. Another notable test was Edmonton Frailty Scale, which had the best predictive ability for post-operative complications and represents future opportunities for feasibility via self-reporting. Research emphasis must continue to move away from odds ratios to predictive accuracy metrics like the c-statistic, especially for extended length of stay.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References

1. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013;381:752-62. doi: 10.1016/S0140-673662167-9. Erratum in: Lancet 2013;382:1328.
2. Parmar KL, Law J, Carter B, Hewitt J, Boyle JM, Casey P, *et al*. Frailty in older patients undergoing emergency laparotomy: Results from the UK observational emergency laparotomy and frailty (ELF) study. Ann Surg 2021;273:709-18.
3. McIsaac DI, MacDonald DB, Aucoin SD. Frailty for perioperative clinicians: A narrative review. Anesth Analg 2020;130:1450-60.
4. Colby SL, Ortman JM. Projections of the Size and Composition of the U.S. population: 2014 to 2060. Washington DC (US): U.S. Census Bureau; 2015. Available from: https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf.
5. Yang R, Wolfson M, Lewis MC. Unique aspects of the elderly surgical population: An anesthesiologist's perspective. Geriatr Orthop Surg Rehabil 2011;2:56-64.
6. Chow WB, Rosenthal RA, Merkow RP, Ko CY, Esnaola NF, American College of Surgeons National Surgical Quality Improvement Program, *et al*. Optimal preoperative assessment of the geriatric surgical patient: A best practices guideline from the American College of Surgeons National Surgical Quality Improvement Program and the American Geriatrics Society. J Am Coll Surg 2012;215:453-66.
7. Griffiths R, Beech F, Brown A, Dhesi J, Foo I, Goodall J, *et al*. Peri-operative care of the elderly 2014: Association of Anaesthetists of Great Britain and Ireland. Anaesthesia 2014;69(Suppl 1):81-98.
8. Aguayo GA, Donneau AF, Vaillant MT, Schritz A, Franco OH, Stranges S, *et al*. Agreement between 35 published frailty scores in the general population. Am J Epidemiol 2017;186:420-34.
9. Rodríguez-Mañas L, Féart C, Mann G, Viña J, Chatterji S, Chodzko-Zajko W, *et al*. Searching for an operational definition of frailty: A Delphi method based consensus statement: The frailty operative definition-consensus conference project. J Gerontol A Biol Sci Med Sci 2013;68:62-7.
10. Parker SG, McCue P, Phelps K, McCleod A, Arora S, Nockels K, *et al*. What is comprehensive geriatric assessment (CGA)? An umbrella review. Age Ageing 2018;47:149-55.
11. Partridge JS, Harari D, Martin FC, Dhesi JK. The impact of pre-operative comprehensive geriatric assessment on postoperative outcomes in older patients undergoing scheduled surgery: A systematic review. Anaesthesia 2014;69(Suppl 1):8-16.
12. Kim KI, Park KH, Koo KH, Han HS, Kim CH. Comprehensive geriatric assessment can predict postoperative morbidity and mortality in elderly patients undergoing elective surgery. Arch Gerontol Geriatr 2013;56:507-12.
13. Kocman D, Regen E, Phelps K, Martin G, Parker S, Gilbert T, *et al*. Can comprehensive geriatric assessment be delivered without the need for geriatricians? A formative evaluation in two perioperative surgical settings. Age Ageing 2019;48:644-9.
14. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, *et al*. Frailty in older adults: Evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56:M146-56.
15. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal 2001;1:323-36.
16. Gobbens RJ, Luijkx KG, Wijnen-Sponselee MT, Schols JM. In search of an integral conceptual definition of frailty: Opinions of experts. J Am Med Dir Assoc 2010;11:338-43.
17. Grund B, Sabin C. Analysis of biomarker data: Logs, odds ratios, and receiver operating characteristic curves. Curr Opin HIV AIDS 2010;5:473-9.
18. Aucoin SD, Hao M, Sohi R, Shaw J, Bentov I, Walker D, *et al*. Accuracy and feasibility of clinically applied frailty instruments before surgery: A systematic review and meta-analysis. Anesthesiology 2020;133:78-95.
19. Frailty Guideline Working Group. Guideline for Perioperative Care for People Living with Frailty Undergoing Elective and Emergency Surgery. [Internet]. London (UK): Centre for Perioperative Care; 2021. Available from: https://www.cpoc.org.uk/sites/cpoc/files/documents/2021-09/CPOC-BGS-Frailty-Guideline-2021.pdf.
20. Ko FC. Preoperative Frailty Evaluation: A Promising Risk-stratification Tool in Older Adults Undergoing General Surgery. Clin Ther 2019;41:387-99.
21. Gené C, Senti S, Parrales M, Troya J, Fernández-Llamazares J, Julián JF, *et al*. Preoperative assessment of geriatric surgical patients: Update on clinical scales used for elective general and digestive surgery. Surg Laparosc Endosc Percutan Tech 2021;31:368-75.
22. Darvall JN, Gregorevic KJ, Story DA, Hubbard RE, Lim WK. Frailty indexes in perioperative and critical care: A systematic review. Arch Gerontol Geriatr 2018;79:88-96.
23. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: A new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004;240:205-13.
24. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med 2013;158:280-6.
25. Caetano SJ, Sonpavde G, Pond GR. C-statistic: A brief explanation of its construction, interpretation and limitations. Eur J Cancer 2018;90:130-2.
26. Burgos E, Gómez-Arnau JI, Díez R, Muñoz L, Fernández-Guisasola J, Garcia del Valle S. Predictive value of six risk scores for outcome after surgical repair of hip fracture in elderly patients. Acta Anaesthesiol Scand 2008;52:125-31.
27. Dasgupta M, Rolfson DB, Stolee P, Borrie MJ, Speechley M. Frailty is associated with postoperative complications in older adults with medical problems. Arch Gerontol Geriatr 2009;48:78-83.
28. Robinson TN, Wu DS, Pointer L, Dunn CL, Cleveland JC Jr, Moss M. Simple frailty score predicts postoperative complications across surgical specialties. Am J Surg 2013;206:544-50.

29. Robinson TN, Wu DS, Sauaia A, Dunn CL, Stevens-Lapsley JE, Moss M, *et al*. Slower walking speed forecasts increased postoperative morbidity and 1-year mortality across surgical specialties. Ann Surg 2013;258:582-8; discussion 588-90.
30. Kenig J, Zychiewicz B, Olszewska U, Barczynski M, Nowak W. Six screening instruments for frailty in older patients qualified for emergency abdominal surgery. Arch Gerontol Geriatr 2015;61:437-42.
31. Revenig LM, Canter DJ, Kim S, Liu Y, Sweeney JF, Sarmiento JM, *et al*. Report of a simplified frailty score predictive of short-term postoperative morbidity and mortality. J Am Coll Surg 2015;220:904-11.e1.
32. Hall DE, Arya S, Schmid KK, Blaser C, Carlson MA, Bailey TL, *et al*. Development and initial validation of the risk analysis index for measuring frailty in surgical populations. JAMA Surg 2017;152:175-82.
33. Kapoor A, Matheos T, Walz M, McDonough C, Maheswaran A, Ruppell E, *et al*. Self-reported function more informative than frailty phenotype in predicting adverse postoperative course in older adults. J Am Geriatr Soc 2017;65:2522-8.
34. Gilbert T, Neuburger J, Kraindler J, Keeble E, Smith P, Ariti C, *et al*. Development and validation of a hospital frailty risk score focusing on older people in acute care settings using electronic hospital records: An observational study. Lancet 2018;391:1775-82.
35. Han B, Wang Y, Chen X. Predictive value of frailty on postoperative complications in elderly patients with major abdominal surgery. Biomed Res 2018;29:1308-15.
36. Kenig J, Mastalerz K, Lukasiewicz K, Mitus-Kenig M, Skorus U. The Surgical Apgar score predicts outcomes of emergency abdominal surgeries both in fit and frail older patients. Arch Gerontol Geriatr 2018;76:54-9.
37. Ondeck NT, Bohl DD, Bovonratwet P, McLynn RP, Cui JJ, Grauer JN. Discriminative ability of Elixhauser's comorbidity measure is superior to other comorbidity scores for inpatient adverse outcomes after total hip arthroplasty. J Arthroplasty 2018;33:250-7.
38. Ondeck NT, Bovonratwet P, Ibe IK, Bohl DD, McLynn RP, Cui JJ, *et al*. Discriminative ability for adverse outcomes after surgical management of hip fractures: A comparison of the charlson comorbidity index, elixhauser comorbidity measure, and modified frailty index. J Orthop Trauma 2018;32:231-7.
39. Zattoni D, Montroni I, Saur NM, Garutti A, Bacchi Reggiani ML, Galetti C, *et al*. A simple screening tool to predict outcomes in older adults undergoing emergency general surgery. J Am Geriatr Soc 2019;67:309-16.
40. Al-Khamis A, Warner C, Park J, Marecik S, Davis N, Mellgren A, *et al*. Modified frailty index predicts early outcomes after colorectal surgery: An ACS-NSQIP study. Colorectal Dis 2019;21:1192-205.
41. Amin KA, Lee UJ, Jin C, Boscardin J, Medendorp AR, Anger JT, *et al*. A national study demonstrating the need for improved frailty indices for preoperative risk assessment of common urologic procedures. Urology 2019;132:87-93.
42. Fu MC, Ondeck NT, Nwachukwu BU, Garcia GH, Gulotta LV, Verma NN, *et al*. What associations exist between comorbidity indices and postoperative adverse events after total shoulder arthroplasty? Clin Orthop Relat Res 2019;477:881-90.
43. Katlic MR, Coleman J, Khan K, Wozniak SE, Abraham JH. Sinai abbreviated geriatric evaluation: Development and validation of a practical test. Ann Surg 2019;269:177-83.
44. Lima DFT, Cristelo D, Reis P, Abelha F, Mourão J. Outcome prediction with physiological and operative severity score for the enumeration of mortality and morbidity score system in elderly patients submitted to elective surgery. Saudi J Anaesth 2019;13:46-51.
45. Arya S, Varley P, Youk A, Borrebach JD, Perez S, Massarweh NN, *et al*. Recalibration and external validation of the risk analysis index: A surgical frailty assessment tool. Ann Surg 2020;272:996-1005.
46. Barazanchi A, Bhat S, Palmer-Neels K, Macfater WS, Xia W, Zeng I, *et al*. Evaluating and improving current risk prediction tools in emergency laparotomy. J Trauma Acute Care Surg 2020;89:382-7.
47. Choi JY, Kim KI, Choi Y, Ahn SH, Kang E, Oh HK, *et al*. Comparison of multidimensional frailty score, grip strength, and gait speed in older surgical patients. J Cachexia Sarcopenia Muscle 2020;11:432-40.
48. He Y, Li LW, Hao Y, Sim EY, Ng KL, Lee R, *et al*. Assessment of predictive validity and feasibility of Edmonton Frail Scale in identifying postoperative complications among elderly patients: A prospective observational study. Sci Rep 2020;10:14682.
49. Lu W, Dai L, Wu G, Hu R. Comparison of two frailty indexes in hip fractures. J Orthop Surg (Hong Kong) 2020;28:2309499020901891. doi: 10.1177/2309499020901891.
50. McIsaac DI, Taljaard M, Bryson GL, Beaulé PE, Gagné S, Hamilton G, *et al*. Frailty as a predictor of death or new disability after surgery: A prospective cohort study. Ann Surg 2020;271:283-9.
51. Rogozinski J, Kiskaddon E, Flanigan T, Spitz H, Froehle A, Chen R, *et al*. The utility of the Charlson Comorbidity Index and modified Frailty Index as quality indicators in total joint arthroplasty: A retrospective cohort review. Curr Orthop Pract 2020;31:543-8.
52. Roopsawang I, Thompson H, Zaslavsky O, Belza B. Predicting hospital outcomes with the reported edmonton frail scale-Thai version in orthopaedic older patients. J Clin Nurs 2020;29:4708-19.
53. Aguilar-Frasco JL, Rodríguez-Quintero JH, Moctezuma-Velázquez P, Morales-Maza J, Moctezuma-Velázquez C, Pastor-Sifuentes F, *et al*. Frailty index as a predictive preoperative tool in the elder population undergoing major abdominal surgery: A prospective analysis of clinical utility. Langenbecks Arch Surg 2021;406:1189-98.
54. Arteaga AS, Aguilar LT, González JT, Boza AS, Muñoz-Cruzado VD, Ciuró FP, *et al*. Impact of frailty in surgical emergencies. A comparison of four frailty scales. Eur J Trauma Emerg Surg 2021;47:1613-9.
55. Costa G, Bersigotti L, Massa G, Lepre L, Fransvea P, Lucarini A, *et al*. The Emergency Surgery Frailty Index (EmSFI): Development and internal validation of a novel simple bedside risk score for elderly patients undergoing emergency surgery. Aging Clin Exp Res 2021;33:2191-201.
56. Lee SW, Nam JS, Kim YJ, Kim MJ, Choi JH, Lee EH, *et al*. Predictive model for the assessment of preoperative frailty risk in the elderly. J Clin Med 2021;10:4612. doi: 10.3390/jcm10194612.
57. Pandit V, Tan TW, Kempe K, Chitwood J, Kim H, Horst V, *et al*. Frailty syndrome in patients with lower extremity amputation: Simplifying how we calculate frailty. J Surg Res 2021;263:230-5.
58. Tse W, Dittman JM, Lavingia K, Wolfe L, Amendola MF. Frailty Index associated with postoperative complications and mortality after lower extremity amputation in a national veteran cohort. J Vasc Surg 2021;74:963-71.
59. Wu HHL, Van Mierlo R, McLauchlan G, Challen K, Mitra S, Dhaygude AP, *et al*. Prognostic performance of clinical assessment tools following hip fracture in patients with chronic kidney disease. Int Urol Nephrol 2021;53:2359-67.
60. Yi BC, Gowd AK, Agarwalla A, Chen E, Amin NH, Nicholson GP, *et al*. Efficacy of the modified Frailty Index and the modified Charlson Comorbidity Index in predicting complications in patients undergoing operative management of proximal humerus fracture. J Shoulder Elbow Surg 2021;30:658-67.
61. Yin Y, Jiang L, Xue L. Comparison of three frailty measures for 90-day outcomes of elderly patients undergoing elective abdominal surgery. ANZ J Surg 2021;91:335-40.
62. Conlon M, Thommen R, Kazim SF, Dicpinigaitis AJ, Schmidt MH, McKee RG, *et al*. Risk analysis index and its recalibrated version predict postoperative outcomes better than 5-factor modified frailty index in traumatic spinal injury. Neurospine 2022;19:1039-48.
63. Cotton J, Cabot J, Buckner J, Field A, Pounds L, Quint C. Increased frailty associated with higher long-term mortality after major lower extremity amputation. Ann Vasc Surg 2022;86:295-304.
64. Forssten MP, Cao Y, Trivedi DJ, Ekestubbe L, Borg T, Bass GA, *et al*.

Developing and validating a scoring system for measuring frailty in patients with hip fracture: A novel model for predicting short-term postoperative mortality. Trauma Surg Acute Care Open 2022;7:e000962. doi: 10.1136/tsaco-2022-000962.
65. Ikram A, Norrish AR, Marson BA, Craxford S, Gladman JRF, Ollivere BJ. Can the clinical frailty scale on admission predict 30-day survival, postoperative complications, and institutionalization in patients with fragility hip fracture? A cohort study of 1,255 patients. Bone Joint J 2022;104-B: 980-6.
66. Iwasaki M, Ishikawa M, Namizato D, Sakamoto A. Worse ECOG-PS is associated with increased 30-day mortality among adults older than 90 years undergoing non-cardiac surgery: A single-center retrospective study. J Nippon Med Sch 2022;89:295-300.
67. Kweh BTS, Lee HQ, Tan T, Liew S, Hunn M, Wee Tee J. Posterior instrumented spinal surgery outcomes in the elderly: A comparison of the 5-item and 11-item modified frailty indices. Global Spine J 2022. doi: 10.1177/21925682221117139.
68. Le ST, Liu VX, Kipnis P, Zhang J, Peng PD, Cespedes Feliciano EM. Comparison of electronic frailty metrics for prediction of adverse outcomes of abdominal surgery. JAMA Surg 2022;157:e220172. doi: 10.1001/jamasurg. 2022.0172.
69. Lee SW, Kim KS, Park SW, Kim J, Choi JH, Lee S, *et al*. Application of the new preoperative frailty risk score in elderly patients undergoing emergency surgery. Gerontology 2022;68:1276-84.
70. Li CQ, Zhang C, Yu F, Li XY, Wang DX. The composite risk index based on frailty predicts postoperative complications in older patients recovering from elective digestive tract surgery: A retrospective cohort study. BMC Anesthesiol 2022;22:7.
71. Palaniappan S, Soiza RL, Duffy S, Moug SJ, Myint PK, Older People's Surgical Outcomes Collaborative (OPSOC), *et al*. Comparison of the clinical frailty score (CFS) to the National Emergency Laparotomy Audit (NELA) risk calculator in all patients undergoing emergency laparotomy. Colorectal Dis 2022;24:782-9.
72. Ruiz M, Peña M, Cohen A, Ehsani H, Joseph B, Fain M, *et al*. Physical and cognitive function assessment to predict postoperative outcomes of abdominal surgery. J Surg Res 2021;267:495-505.
73. Wei B, Zong Y, Xu M, Wang X, Guo X. The revised-risk analysis index as a predictor of major morbidity and mortality in older patients after abdominal surgery: A retrospective cohort study. BMC Anesthesiol 2022;22:301.
74. Yin Y, Jiang L, Xue L. Which frailty evaluation method can better improve the predictive ability of the SASA for postoperative complications of patients undergoing elective abdominal surgery? Ther Clin Risk Manag 2022;18:541-50.
75. Darbyshire AR, Kostakis I, Meredith P, Toh SKC, Prytherch D, Briggs J. Novel predictors of mortality in emergency bowel surgery: A single-centre cohort study. Anaesthesia 2023;78:561-70.
76. McConaghy KM, Orr MN, Emara AK, Sinclair ST, Klika AK, Piuzzi NS. Can extant comorbidity indices identify patients who experience poor outcomes following total joint arthroplasty? Arch Orthop Trauma Surg 2023;143:1253-63.
77. Sirisegaram L, Owodunni OP, Ehrlich A, Qin CX, Bettick D, Gearhart SL. Validation of the self-reported domains of the edmonton frail scale in patients 65 years of age and older. BMC Geriatr 2023;23:15.
78. Abellan van Kan G, Rolland YM, Morley JE, Vellas B. Frailty: Toward a clinical definition. J Am Med Dir Assoc 2008;9:71-2.
79. Podsiadlo D, Richardson S. The timed "Up & Go": A test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142-8.
80. Toosizadeh N, Wendel C, Hsu CH, Zamrini E, Mohler J. Frailty assessment in older adults using upper-extremity function: Index development. BMC Geriatr 2017;17:117.
81. Guyatt GH, Sullivan MJ, Thompson PJ, Fallen EL, Pugsley SO, Taylor DW, *et al*. The 6-minute walk: A new measure of exercise capacity in patients with chronic heart failure. Can Med Assoc J 1985;132:919-23.
82. Balducci L, Beghe C. The application of the principles of geriatrics to the management of the older person with cancer. Crit Rev Oncol Hematol 2000;35:147-54.
83. Meldon SW, Mion LC, Palmer RM, Drew BL, Connor JT, Lewicki LJ, *et al*. A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department. Acad Emerg Med 2003;10:224-32.
84. Mahoney FI, Barthel DW. Functional evaluation: The Barthel Index. Md State Med J 1965;14:61-5.
85. Saliba D, Elliott M, Rubenstein LZ, Solomon DH, Young RT, Kamberg CJ, *et al*. The Vulnerable Elders Survey: A tool for identifying vulnerable older people in the community. J Am Geriatr Soc 2001;49:1691-9.
86. Subramaniam S, Aalberg JJ, Soriano RP, Divino CM. New 5-factor modified frailty index using American College of Surgeons NSQIP data. J Am Coll Surg 2018;226:173-81.e8.
87. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. J Chronic Dis 1987;40:373-83.
88. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998;36:8-27.
89. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr 2008;8:24. doi: 10.1186/1471-2318-8-24.
90. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, *et al*. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005;173:489-95.
91. Zubrod CG, Schneiderman M, Frei III E, Brindley C, Gold GL, Shnider B, *et al*. Appraisal of methods for study of chemotherapy of cancer in man: Comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. J Chronic Dis 1960;11:7-33.
92. Karnofsky DA, Burchenal JH. The Clinical Evaluation of Chemotherapeutic Agents in Cancer. New York (US): Columbia University Press; 1949.
93. Rockwood K, Stadnyk K, MacKnight C, McDowell I, Hébert R, Hogan DB. A brief clinical instrument to classify frailty in elderly people. Lancet 1999;353:205-6.
94. Bellera CA, Rainfray M, Mathoulin-Pélissier S, Mertens C, Delva F, Fonck M, *et al*. Screening older cancer patients: First evaluation of the G-8 geriatric screening tool. Ann Oncol 2012;23:2166-72.
95. Rolfson DB, Majumdar SR, Tsuyuki RT, Tahir A, Rockwood K. Validity and reliability of the Edmonton Frail scale. Age Ageing 2006;35:526-9.
96. McDonough CM, Tian F, Ni P, Kopits IM, Moed R, Pardasaney PK, *et al*. Development of the computer-adaptive version of the late-life function and disability instrument. J Gerontol A Biol Sci Med Sci 2012;67:1427-38.
97. Steverink N. Measuring frailty: Developing and testing the GFI (Groningen Frailty Indicator). Gerontologist 2001;41:236-7.
98. Kim SW, Han HS, Jung HW, Kim KI, Hwang DW, Kang SB, *et al*. Multidimensional frailty score for the prediction of postoperative mortality risk. JAMA Surg 2014;149:633-40.
99. The Ottawa Hospital. Clinical Frailty Scale (CFS) Training Module – Overview. Ottawa (CA): The Ottawa Hospital; 2019. Available from: https://rise.articulate.com/share/deb4rT02lvONbq4AfcMNRUudcd6QMts3#/.
100. Acute Frailty Network, National Health Service Elect. Clinical Frailty Scale (CFS). California (US): Apple; 2020. Available from: https://apps.apple.com/us/app/clinical-frailty-scale-cfs/id1508556286.
101. Mirabelli LG, Cosker RM, Kraiss LW, Griffin CL, Smith BK, Sarfati MR, *et al*. Rapid methods for routine frailty assessment during vascular surgery clinic visits. Ann Vasc Surg 2018;46:134-41.
102. Fornaess KM, Nome PL, Aakre EK, Hegvik TA, Jammer I. Clinical frailty scale: Inter-rater reliability of retrospective scoring in emergency abdominal surgery. Acta Anaesthesiol Scand 2022;66:25-9.

103. McIsaac DI, Harris EP, Hladkowicz E, Moloo H, Lalu MM, Bryson GL, *et al*. Prospective comparison of preoperative predictive performance between 3 leading frailty instruments. Anesth Analg 2020;131:263-72.
104. Maxwell MJ, Moran CG, Moppett IK. Development and validation of a preoperative scoring system to predict 30-day mortality in patients undergoing hip fracture surgery. Br J Anaesth 2008;101:511-7.
105. Ambler GK, Brooks DE, Al Zuhir N, Ali A, Gohel MS, Hayes PD, *et al*. Effect of frailty on short- and mid-term outcomes in vascular surgical patients. Br J Surg 2015;102:638-45.
106. Daabiss M. American Society of Anaesthesiologists physical status classification. Indian J Anaesth 2011;55:111-5.
107. Riley R, Holman C, Fletcher D. Inter-rater reliability of the ASA physical status classification in a sample of anaesthetists in Western Australia. Anaesth Intensive Care 2014;42:614-8.
108. Sankar A, Johnson SR, Beattie WS, Tait G, Wijeysundera DN. Reliability of the American Society of Anesthesiologists physical status scale in clinical practice. Br J Anaesth 2014;113:424-32.
109. Mak PH, Campbell RC, Irwin MG, American Society of Anesthesiologists. The ASA Physical Status Classification: Inter-observer consistency. American Society of Anesthesiologists. Anaesth Intensive Care 2002;30:633-40.
110. van Wissen J, van Stijn MF, Doodeman HJ, Houdijk AP. Mini nutritional assessment and mortality after hip fracture surgery in the elderly. J Nutr Health Aging 2016;20:964-8.
111. Chen D, Chen J, Yang H, Liang X, Xie Y, Li S, *et al*. Mini-Cog to predict postoperative mortality in geriatric elective surgical patients under general anesthesia: A prospective cohort study. Minerva Anestesiol 2019;85:1193-200.
112. Racine AM, Fong TG, Gou Y, Travison TG, Tommet D, Erickson K, *et al*. Clinical outcomes in older surgical patients with mild cognitive impairment. Alzheimers Dement 2018;14:590-600.
113. Gingrich A, Volkert D, Kiesswetter E, Thomanek M, Bach S, Sieber CC, *et al*. Prevalence and overlap of sarcopenia, frailty, cachexia and malnutrition in older medical inpatients. BMC Geriatr 2019;19:120.
114. Canales C, Mazor E, Coy H, Grogan TR, Duval V, Raman S, *et al*. Preoperative point-of-care ultrasound to identify frailty and predict postoperative outcomes: A diagnostic accuracy study. Anesthesiology 2022;136:268-78.
115. Joliat GR, Schoor A, Schäfer M, Demartines N, Hübner M, Labgaa I. Postoperative decrease of albumin (ΔAlb) as early predictor of complications after gastrointestinal surgery: A systematic review. Perioper Med (Lond) 2022;11:7.
116. González-Martínez S, Olona Tabueña N, Martín Baranera M, Martí-Saurí I, Moll JL, Morales García MÁ, *et al*. Inflammatory markers as predictors of postoperative adverse outcome in octogenarian surgical patients: An observational prospective study. Cir Esp 2019;93:166-73.
117. Fong TG, Vasunilashorn SM, Ngo L, Libermann TA, Dillon ST, Schmitt EM, *et al*. Association of plasma neurofilament light with postoperative delirium. Ann Neurol 2020;88:984-94.
118. Sharipova V, Alimov A, Valihanov A. Interleukin-6 and glial fibrillary acidic protein in prediction of early postoperative cognitive dysfunction after orthopedic surgery. Clin Med Diagn 2020;10:38-42.
119. Zhong R, Rau PP. A mobile phone-based gait assessment app for the elderly: Development and evaluation. JMIR Mhealth Uhealth 2020;8:e14453. doi: 10.2196/14453.
120. Viccaro LJ, Perera S, Studenski SA. Is timed up and go better than gait speed in predicting health, function, and falls in older adults? J Am Geriatr Soc 2011;59:887-92.
121. De Roeck EE, De Deyn PP, Dierckx E, Engelborghs S. Brief cognitive screening instruments for early detection of Alzheimer’s disease: A systematic review. Alzheimers Res Ther 2019;11:21.
122. Kaustov L, Fleet A, Brenna CTA, Orser BA, Choi S. Perioperative neurocognitive screening tools for at-risk surgical patients. Neurol Clin Pract 2022;12:76-84.
123. Collins TC, Daley J, Henderson WH, Khuri SF. Risk factors for prolonged length of stay after major elective surgery. Ann Surg 1999;230:251-9.
124. Mah SJ, Anpalagan T, Marcucci M, Eiriksson L, Reade CJ, Jimenez W, *et al*. The five-factor modified frailty index predicts adverse postoperative and chemotherapy outcomes in gynecologic oncology. Gynecol Oncol 2022;166:154-61.
125. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007;115:928-35.
126. Cords CI, Spronk I, Mattace-Raso FUS, Verhofstad MHJ, van der Vlies CH, van Baar ME. The feasibility and reliability of frailty assessment tools applicable in acute in-hospital trauma patients: A systematic review. J Trauma Acute Care Surg 2022;92:615-26.
127. Jarman H, Crouch R, Baxter M, Wang C, Peck G, Sivapathasuntharam D, *et al*. Feasibility and accuracy of ED frailty identification in older trauma patients: A prospective multi-centre study. Scand J Trauma Resusc Emerg Med 2021;29:54.
128. Davis PJ, Bailey JG, Molinari M, Hayden J, Johnson PM. The impact of nonelective abdominal surgery on the residential status of older adult patients. Ann Surg 2016;263:274-9.

Supplementary Table 1: Search Terms Used for (1) Medline and (2) Embase Databases

Supplementary Table 1.1: Search terms used for Ovid Medline (R). Search period from 1946 to March 10, 2023

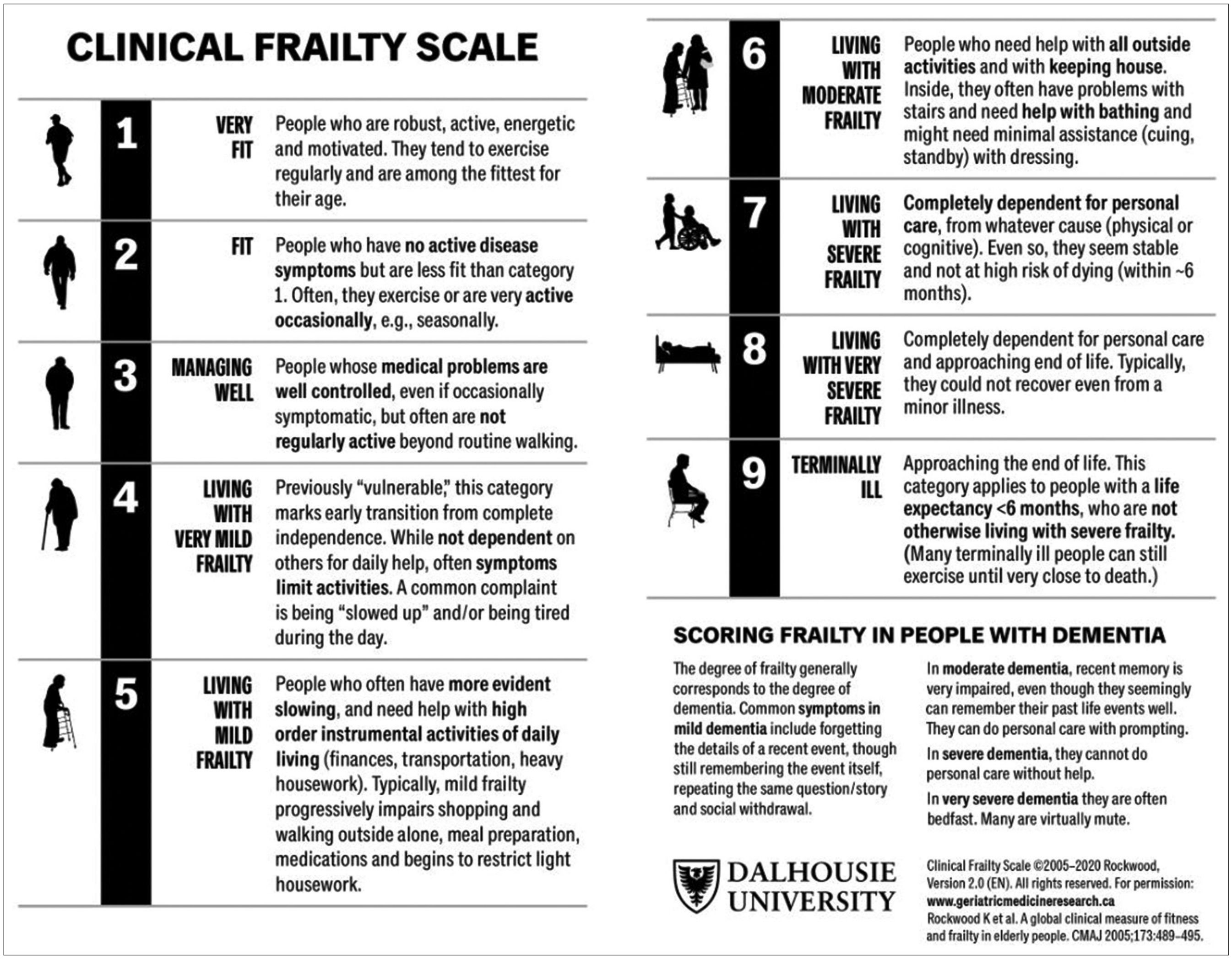
Search Step	Search Terms	Total Studies, *n*
1	(Preoperative Care or preoperative period).sh.	74390
2	(preoperat* or pre-operat*).tw.	385505
3	1 or 2	418662
4	(frailty or frail elderly).sh.	19638
5	frail*.tw.	32568
6	4 or 5	37347
7	geriatric assessment.sh.	32057
8	(test* or screen* or assess* or index* or indicator* or rule* or measur* or tool* or instrument* or scale* or score* or metri* or rating or resignation or phenotype).tw.	11619843
9	7 or 8	11627309
10	(mortality or death or morbidity or complication* or adverse event* or length of stay).tw.	2842226
11	3 and 6 and 9 and 10	1080

Supplementary Table 1.2: Search terms used for Embase. Search period from 1947 to March 10, 2023

Search Step	Search Terms	Total Studies, *n*
1	(preoperat* OR "pre-operat*")	547321
2	frail*	51617
3	(old OR elderly OR geriatric OR aged)	5217295
4	(test* OR screen* OR assess* OR index* OR indicator* OR rule* OR measur* OR tool* OR instrument* OR scale* OR score* OR metri* OR rating OR resignation OR phenotype)	14682871
5	(complication* OR "adverse event") AND ("post-operative")	65120
6	(mortality OR "length of stay")	1725801
7	5 OR 6	1773169
8	1 AND 2 AND 3 AND 4 AND 7	1282

Supplementary Table 2 : Resources for Completion of Top 5 Frailty Tests

Supplementary Table 2.1: Clinical Frailty Scale



CLINICAL FRAILTY SCALE

1	VERY FIT	People who are robust, active, energetic and motivated. They tend to exercise regularly and are among the fittest for their age.
2	FIT	People who have **no active disease symptoms** but are less fit than category 1. Often, they exercise or are very **active occasionally**, e.g., seasonally.
3	MANAGING WELL	People whose **medical problems are well controlled**, even if occasionally symptomatic, but often are **not regularly active** beyond routine walking.
4	LIVING WITH VERY MILD FRAILTY	Previously "vulnerable," this category marks early transition from complete independence. While **not dependent** on others for daily help, often **symptoms limit activities.** A common complaint is being "slowed up" and/or being tired during the day.
5	LIVING WITH MILD FRAILTY	People who often have **more evident slowing**, and need help with **high order instrumental activities of daily living** (finances, transportation, heavy housework). Typically, mild frailty progressively impairs shopping and walking outside alone, meal preparation, medications and begins to restrict light housework.
6	LIVING WITH MODERATE FRAILTY	People who need help with **all outside activities** and with **keeping house.** Inside, they often have problems with stairs and need **help with bathing** and might need minimal assistance (cuing, standby) with dressing.
7	LIVING WITH SEVERE FRAILTY	**Completely dependent for personal care**, from whatever cause (physical or cognitive). Even so, they seem stable and not at high risk of dying (within ~6 months).
8	LIVING WITH VERY SEVERE FRAILTY	Completely dependent for personal care and approaching end of life. Typically, they could not recover even from a minor illness.
9	TERMINALLY ILL	Approaching the end of life. This category applies to people with a **life expectancy <6 months**, who are **not otherwise living with severe frailty.** (Many terminally ill people can still exercise until very close to death.)

SCORING FRAILTY IN PEOPLE WITH DEMENTIA

The degree of frailty generally corresponds to the degree of dementia. Common **symptoms in mild dementia** include forgetting the details of a recent event, though still remembering the event itself, repeating the same question/story and social withdrawal.

In **moderate dementia**, recent memory is very impaired, even though they seemingly can remember their past life events well. They can do personal care with prompting.

In **severe dementia**, they cannot do personal care without help.

In **very severe dementia** they are often bedfast. Many are virtually mute.

DALHOUSIE UNIVERSITY

Rockwood K et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005;173:489–495.

Supplementary Table 2.2: Clinical risk analysis index

Risk Analysis Index (RAI)

Last Name:______________________ Last Four: _________

Date Form is Completed:___________ Date & Type of Anticipated Surgery:__________________

A. Age, Sex & Cancer

Age	Score *without* Cancer	Score *with* Cancer
< 69	2	20
70-74	3	19
75-79	4	18
80-84	5	17
85-89	6	16
90-94	7	15
95-99	8	14
100+	9	13

1. Sex Female= 0 Male= 5 ___________
2. Age ___________
3. Does the patient have cancer?
   (Excluding skin cancer,except for melanoma)
   If no, score *without* cancer ___________
   or
   If yes, score *with* cancer ___________

B. Medical Co-Morbidities

4. Have you had unintentional weight loss in the past 3 months (>10 lbs)? No= 0 Yes= 5 _______
5. Renal failure? No= 0 Yes= 6 _______
6. Chronic/congestive heart failure? No= 0 Yes= 4 _______
7. Poor appetite? No= 0 Yes= 4 _______
8. Shortness of breath (at rest)? No= 0 Yes= 8 _______

C. Cognition, Residence & Activity of Daily Living

9. Do you reside in a setting other than independent living?
   If yes, check answer: Skilled nursing facility □ Assisted living □ Nursing home □
   No= 0 Yes= 8 _______
   If yes, were you admitted within the past 3 months? No □ Yes □

D. Activities of Daily Living & Cognitive Decline (*Circle score for each ADL*)

10. Mobility/Locomotion	11. Eating	12. Toilet Use	13. Personal Hygiene
0. Independent	0. Independent	0. Independent	0. Independent
1. Supervised	1. Supervised	1. Supervised	1. Supervised
2. Limited assistance	2. Limited assistance	2. Limited assistance	2. Limited assistance
3. Extensive assistance	3. Extensive assistance	3. Extensive assistance	3. Extensive assistance
4. Total Dependence	4. Total Dependence	4. Total Dependence	4. Total Dependence

14. Have your cognitive skills or status deteriorated over the past 3 months? No □ Yes □ (*see score chart*)

ADL Score *without* Cognitive Decline (*Sum of ADL Scores*)	ADL Score *with* Cognitive Decline
0	ADL Score -2
1,2	ADL Score -1
3,4	ADL Score 0
5-7	ADL Score +1
8,9	ADL Score +2
10,11	ADL Score +3
12,13	ADL Score +4
14-16	ADL Score +5

Score *without* cognitive decline _____________ (0 to 16)
or
Score *with* cognitive decline _____________(-2 to 21)

Total RAI Score: ___________

Scoring Instructions: To calculate the RAI-C score, first look at the age/cancer table to determine the single value between 2 and 20 that corresponds to the patient's age and cancer status. Record this single value in the appropriate line for item 3. Next look at the ADL table and sum the scores (0–4) for the four ADLs queried in items 10–13. This sum is the ADL Score and should range between 0 and 16. Next look at the ADL/Cognitive-Decline table to determine the single value between -2 and 21 that corresponds to the patient's ADL Score and cognitive decline. Record the value in the appropriate line for item 14. Finally, sum the values for items 1,3-9, and 14 to yield a final RAI-C score between 0 and 81

Supplementary Table 2.3: Edmonton frail scale

The Edmonton Frail Scale

NAME : ________________________________

d.o.b. : ______________________ **DATE :** __________________

Frailty domain	Item	0 point	1 point	2 points
Cognition	Please imagine that this pre-drawn circle is a clock. I would like you to place the numbers in the correct positions then place the hands to indicate a time of 'ten after eleven'	No errors	Minor spacing errors	Other errors
General health status	In the past year, how many times have you been admitted to a hospital?	0	1–2	≥2
	In general, how would you describe your health?	'Excellent', 'Very good', 'Good'	'Fair'	'Poor'
Functional independence	With how many of the following activities do you require help? (meal preparation, shopping, transportation, telephone, housekeeping, laundry, managing money, taking medications)	0–1	2–4	5–8
Social support	When you need help, can you count on someone who is willing and able to meet your needs?	Always	Sometimes	Never
Medication use	Do you use five or more different prescription medications on a regular basis?	No	Yes	
	At times, do you forget to take your prescription medications?	No	Yes	
Nutrition	Have you recently lost weight such that your clothing has become looser?	No	Yes	
Mood	Do you often feel sad or depressed?	No	Yes	
Continence	Do you have a problem with losing control of urine when you don't want to?	No	Yes	
Functional performance	I would like you to sit in this chair with your back and arms resting. Then, when I say 'GO', please stand up and walk at a safe and comfortable pace to the mark on the floor (approximately 3 m away), return to the chair and sit down'	0–10 s	11–20 s	One of : >20 s , or patient unwilling , or requires assistance
Totals	Final score is the sum of column totals			

Scoring :

0 - 5 = Not Frail

6 - 7 = Vulnerable

8 - 9 = Mild Frailty

10-11 = Moderate Frailty

12-17 = Severe Frailty

TOTAL **/17**

Administered by : ______________________

Supplementary Table 2.4: Modified Frailty Index

Modified Frailty Index	
Item	**Score**
Functionally dependent	1
History of diabetes	1
Chronic obstructive pulmonary disease	1
Congestive heart failure	1
Hypertension	1
Total:	

A score of >2 designates a frail person.

Supplementary Table 2.5: Administrative risk analysis index

RAI Variable	MDS Variable	VASQIP Variable	RAI-A Scoring System*
1. Sex	Sex	SEX	+5 if "male"
2. Age	Age	AGE	Continuous (Scored as interaction with Cancer Diagnoses as per table in eFigure 1)
3. Cancer (excluding skin cancer, except for melanoma)	Cancer diagnosis with or without metastasis	DISCANCER or RADIO or CHEMO	1= any of the 3 variables "yes" 0 = all of the 3 variables "no"
4. Weight Loss ("Have you had unintentional weight loss in the past 3 months >10 pounds?"	Weight loss	WTLOSS	+5 if "yes"
5. Renal Failure	Renal failure	RENALFAIL or DIALYSIS	+6 if either variable "yes"
6. Chronic/Congestive Heart Failure	Chronic heart failure	HXCHF	+4 if "yes"
7. Poor Appetite	Poor appetite	WTLOSS	+4 if "yes"
8. Shortness of Breath at Rest	Shortness of breath	DYSPNEA	+8 if "yes"
9. Residence other than Independent Living	Recent admission to nursing home	TRANST	+8 if transferred to the hospital for the index operation from a nursing home, chronic care facility, spinal cord injury unit or intermediate care unit
10. Cognitive Deterioration ("Have your cognitive skills or status deteriorated over the last 3 months?")	Cognitive Deterioration	IMPSENS or COMA or CVANEURO	"yes" if any of 3 variables "yes" "no" if all of 3 variables "no" (Scored as interaction with Activities of Daily Living as per table below)
11. Activities of Daily Living (Mobility, Eating, Toileting, Personal Hygiene)	Short-Form ADL Scale in 4 dimensions	FNSTATUS	*Without Cognitive Decline* +16 = totally dependent +8 = partially dependent +0 = independent *With Cognitive Decline* +21 = totally dependent +10 = partially dependent -2 = independent

The RAI-A score is calculated the same way as the RAI-C, and both scores range between 0 and 81

Supplementary Table 3: Risk of bias assessment of 30 studies involved in top 5 ranking analysis according to Quality in Prognosis Studies tool. Green, yellow, red, and white represent low, moderate, and high and unclear risk of bias, respectively

Study	Study Participation	Study Attrition	Prognostic Factor Measurement	Outcome Measurement	Study Confounding	Statistical Analysis and Reporting	Overall Risk of Bias
Al-Hamis 2019	Green	Green	Green	Green	Yellow	Green	Green
Amin 2019	Green	Green	Red	Green	Red	Green	Red
Arteaga 2021	Yellow	Green	Yellow	Green	Red	Green	Red
Arya 2020	Green	Green	Yellow	Green	Yellow	Green	Yellow
Barazanchi 2020	Green	Green	Green	Green	Red	Green	Yellow
Conlon 2022	Green	Green	Green	Green	Yellow	Green	Green
Cotton 2022	Green	Green	White	Green	Green	Green	Green
Dasgupta 2008	Yellow	Green	Green	White	Yellow	Green	Yellow
Forssten 2022	Green	Green	Yellow	Green	Red	Green	Yellow
Fu 2019	Green	Green	Green	Green	Red	Yellow	Yellow
Hall 2017	White	Green	Yellow	Green	Red	Green	Yellow
He 2020	Green	Yellow	White	Green	Green	Green	Green
Ikram 2022	Yellow	White	Yellow	Red	Red	Yellow	Red
Iwasaki 2022	Green	Green	Green	White	Yellow	Green	Green
Kweh 2022	Yellow	Green	White	Green	Yellow	Yellow	Yellow
Le 2022	Green	Green	Green	Green	Red	Green	Yellow
Li 2022	Green	Green	Red	Green	Green	Green	Yellow
Lima 2019	Green	Yellow	Yellow	Green	Red	Green	Red
McConaghy 2023	Green	Green	Green	Green	Yellow	Green	Green
McIsaac 2020	Green	Yellow	Green	Green	Yellow	Green	Green
Ondeck 2018	Green	Green	Yellow	Green	Red	Green	Yellow
Ondeck 2018 2	Green	Green	Yellow	Green	Red	Green	Yellow
Palaniappan 2022	Green	Green	Yellow	Green	Red	Green	Yellow
Pandit 2021	Green	Green	Green	Green	Green	Green	Green
Rogozinski 2020	Green	Green	Yellow	Green	Red	Green	Yellow
Roopsawang 2020	Green	Yellow	Green	Green	Green	Green	Green
Tse 2021	Green	Green	Green	Green	Yellow	Green	Green
Wu 2021	Green	White	Yellow	Green	Green	Green	Yellow
Yi 2021	Green	Green	Yellow	Green	Yellow	Green	Green
Yin 2021	Yellow	Yellow	Green	Green	Green	Green	Green