Outcome prediction with Physiological and Operative Severity Score for the enumeration of Mortality and Morbidity score system in elderly patients submitted to elective surgery

ABSTRACT
Context: Elderly patients have a higher risk of complications and 30-day mortality than younger patients. Population is aging and this is an emergent preoccupation.
Aims: The aim of this study was to evaluate the performance of Physiological and Operative Severity Score for the enumeration of Mortality and Morbidity (POSSUM) system on 30-day mortality in elderly patients submitted to elective surgery. Additionally, the correlation of WHODAS 2.0 and Clinical Frailty Score (CFS) with mortality was evaluated.
Settings and Design: An observational prospective study was conducted between May and July 2017.
Methods and Material: Patients submitted to elective orthopedic, gynecologic, urologic, vascular, plastic, and general surgery were included. Exclusion criteria were as follows: age <60 years old; inability to give informed consent; emergency/urgency surgery, inability to understand Portuguese; patients admitted in the ICU after surgery. POSSUM was used to estimate postoperative mortality risk. WHODAS 2.0 and CFS were used to assess quality of life and health status. Mortality was evaluated during hospital stay and 30 days after surgery. area under the receiver operating characteristic (AUROC) was analyzed to test the discrimination of P-POSSUM, WHODAS 2.0 and CFS scale.
Statistical Analysis Used: Statistical analysis was done using the SPSS Software (version 24.0).
Results: POSSUM-predicted mortality was 3.0% with a standardized mortality ratio = 0.87; 95% CI 0.62–0.93; and a good calibration (H–L: P = 0.646); however, the AUROC was poor (0.563). We identified an association between mortality and a higher CFS grade (P = 0.000 and AUROC = 0.859) and a higher WHODAS 2.0 score (P = 0.000 and AUROC = 0.808).
Conclusions: WHODAS and CFS appear to be a better assessment tolls for predicting postoperative mortality with a good discrimination comparing with P-POSSUM system.
Key words: Clinical Frailty Score; elderly; Physiological and Operative Severity Score for the enumeration of Mortality and Morbidity; surgery; World Health Organization Disability Assessment Schedule 2.0
Introduction

According to World Health Organization, the population is aging and the rise in life expectancy allied with an increase in quality of life are important contributes.\cite{1,2,3,11} Over 40% of surgical procedures are performed in those with over 65 years old, with higher risk of complications and 30-day mortality (with rates 5–10%) than younger patients.\cite{4,5,6,7}

Our aim was to evaluate Physiological and Operative Severity Score for the enumeration of Mortality and Morbidity (POSSUM) on predicting 30-day mortality in elderly patients submitted to elective surgery. Additionally, correlation of World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) and Clinical Frailty Score (CFS) with mortality was evaluated.

Materials and Methods

Ethical approval for this study was provided from Institutional Ethics Committee. A prospective longitudinal study was conducted at a university hospital.

We selected patients over 60 years old who submitted to general, regional, or combined anesthesia for surgical interventions in general surgery, urology, gynecology, plastic, orthopedic, or maxillofacial surgery between May of 2017 and July of 2017. Patients were excluded if they were submitted to emergency surgery, were unable to understanding Portuguese or give the informed consent, had life-threatening condition, had a cognitive impairment, or were admitted to an intensive care unit. Two to 24 h before surgery, the following data were collected: demographic characteristics of patients, diagnosis, type of surgery, date of admission in the hospital, comorbidities, usual medication and physical state of the American Society of Anesthesiologists (ASA), WHODAS 2.0 self-administered questionnaire, and the CFS. A patient’s comorbidities were used to calculate the age-adjusted Charlson Comorbidity Index that includes 19 diagnoses and ranges 0–43.\cite{8}

Perioperative parameters for POSSUM score system were collected postoperatively from anesthetic chart and discharge data from the hospital were also collected.

WHODAS 2.0 12-item Portuguese version, self-administered questionnaire was used with five possible answers for each item: none, mild, moderate, severe, and extreme or cannot do. With these results, a score was calculated that varies from 0 (no disability) to 100 (total disability). This score evaluates limitations over the last 30 days in six domains including cognition, mobility, self-care, getting along, life activities, and participation. The Portuguese version was shown to be valid, reliable, easily applied, understood, and equivalent to the original version.\cite{9,10}

In Rockwood et al., CFS was used with a scale from 1 to 9. The CFS-1 “very fit” corresponds to people who are robust, active, energetic, and motivated; they usually exercise regularly and they are among the fittest to their age. CFS-2 “well” represents people with no active disease symptoms but is less fit than CFS-1, often or occasionally they exercise. CFS-3 “managing well” corresponds to people whose medical problems are controlled but are not regularly active beyond their routine walking. CFS-4 “vulnerable” includes those who are not dependent on others for daily help but often symptoms limit activities and they commonly complain of being “slowed up” or tired during the day. CFS-5 “mildly frail” corresponds to people who have more slowing and need help in high-order daily activities; they need supervision for walking outside alone and taking their medication. CFS-6 “moderately frail” represents people who need help with all outside activities and with keeping house, they often have problems with stairs and need help with bathing and might need minimal assistance (cuing, standby) with dressing. In CFS-7, “severely frail” people are completely dependent for personal care, from physical or cognitive cause; they seem stable and not at high risk of dying (within ~6 months). In CFS-8, “very severely frail” people are completely dependent, approaching the end of life; they could not recover even from a minor illness. In the last category CFS-9, “terminally ill” people are approaching the end of life; their life expectancy is less than 6 months.\cite{11} Both scores data were collected 2–24 h before the surgery.

According to Copeland et al., POSSUM system includes a physiological score (PS) with 12 preoperative variables and range from 12 to 88 and an operative score (OS) that includes 6 variables and range between 6 and 48.\cite{12} The PS included the following variables: age, cardiac signs, respiratory signs, systolic blood pressure, pulse rate, Glasgow Coma Score, serum urea, serum sodium, serum potassium, hemoglobin level, white blood cell count and electrocardiogram signs. The OS was based on operative magnitude, number of operations within 30 days, blood loss, peritoneal contamination, presence of malignancy, and timing of operation. Each variable had a 4-grade classification with an exponentially increasing score (1, 2, 4, 8), and if data are not available, the score allocated is 1.\cite{12} P-POSSUM was calculated for all patients with the following formulae (where $R$ is the risk of mortality): P-POSSUM, $\ln[R(1-R)] = -9.065 + (0.1692 \times PS) + (0.155 \times OS)$. 


Descriptive statistics are presented as numbers and percentages for categorical variables and continuous variables as mean and standard deviation or as median and range, depending if there is a normal or skewed distribution for what Kolmogorov–Smirnov test for normality was performed.

POSSUM system score was assessed using calibration fit models and observed over expected mortality ratio using the Hosmer–Lemeshow test (H–L T) and standardized mortality ratio (SMR).

POSSUM system score was assessed using calibration fit models and observed over expected mortality ratio using the Hosmer–Lemeshow test (H–L T) and standardized mortality ratio (SMR).

POSSUM system score was assessed using calibration fit models and observed over expected mortality ratio using the Hosmer–Lemeshow test (H–L T) and standardized mortality ratio (SMR).

Calibration was considered poor when the P value was <0.05 and the Chi-square value was large. [13,14] The area under the receiver operating characteristic (AUROC) curve was used to discriminate between patients who died in the postoperative period and those who did not. [15] Additionally, Mann–Whitney U test was used.

For all data collection and statistical analyses, SPSS 24.0 version was used.

Results

There were 229 patients, 103 men (45%) and 126 women (55%) with a median age of 69 years (range 60–91 years).

Ninety-four patients (41%) were submitted to general surgery, 53 (23.1%) to urology, 11 (4.8%) to gynecology, 15 (6.1%) to plastic, 35 (15.3%) to orthopedic, and 21 (9.2%) to vascular surgery.

The median Charlson Comorbidity Index was 6 (range 1–27). The mean hospital stay was 10.55 days with a maximum of 87 days and the mean stay after surgery 7.04 days with a maximum stay of 66 days. Concerning ASA classification, 4.4% (n = 10) patients were class I, 55% (n = 126) were class II, 36.2% (n = 83) were class III, and 4.4% (n = 10) were class IV.

The median CFS was 3 with a range 1–7; no patient was registered as “very severely frail” (CFS-8) or “terminally ill” (CFS-9). The mean WHODAS score was 20.3 (median, 12.5; range 0–81.25). The mean POSSUM physiology score was 19.02 (median, 19; range 12–40) and the mean OS was 8.03 (median, 8; range 6–13).

The overall hospital mortality rate was 2.62% (n = 6) and 30-day mortality rate was 3.93% (n = 9). There was no significant difference in mortality rate among the different surgical specialties. The differences between the patients that survived and the others are described in Table 1.

POSSUM-Predicted 30-day mortality was 3.0%. The observed-to-expected ratio (O: E) or SMR showed that there was no significant difference on predicted mortality facing the observed 30-days number of deaths for POSSUM (SMR = 0.87; 95% CI 0.93–0.062); however, the AUROC was poor (0.563) with 95% confidence interval (CI) of 0.375–0.748. We identified significant correlations between mortality and a higher CFS grade (Spearman’s ρ = −0.247; P = 0.000) and a higher WHODAS 2.0 score (Spearman’s ρ = −0.208; P = 0.000). The AUROC and 95% CI for CFS and WHODAS were 0.859 (0.750–0.968) and 0.808 (0.638–0.979), respectively. A AUROC comparison is illustrated in Figure 1.

Discussions

It became necessary to accurate the risk–benefit before the surgery and make a pondered choice of whom benefits the most with the surgical procedures according with each patient’s expectations. [5,18] Several scores systems have been developed with a main purpose of prediction and an outcome of an individual patient. The perfect system should be simple, reproducible, objective, and available to all patients. [19] Additionally, it would be cheap and preferably based on preoperative risk factors instead of intra and/or postoperative data. [19] The main goal is to classify the patient’s

Table 1: Differences between the patients with the outcome (death) and without the outcome (alive)

|                      | Alive (n=220) | Death (n=9) | P   |
|----------------------|--------------|------------|-----|
| Age (years), median (range) | 69 (59-91)   | 79 (61-85) | 0.471 |
| CFS, median (range)    | 3 (1-7)      | 6 (3-7)    | 0.005 |
| WHODAS 2.0, median (range) | 11.46 (0-79.17) | 54.17 (2.08-81.25) | 0.029 |
| POSSUM physiology score, mean (standard deviation) | 18.97 (4.10) | 20.33 (7.63) | 0.350 |
| POSSUM operative score, mean (standard deviation) | 8.00 (1.83) | 8.67 (2.60) | 0.297 |
| Charlson index, median (range) | 6 (1-24) | 14 (8-27) | 0.002 |
| ASA, median (range)    | 2 (1-4)      | 3 (2-4)    | 0.008 |
| Hospital stay, median (range) | 5 (1-87) | 25 (8-82) | 0.005 |

ASA: Physical state of the American Society of Anesthesiologists; CFS: Clinical Frailty Score; POSSUM: Portsmouth Physiological and Operative Severity Score for the enumeration of mortality and morbidity; WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0
In this study, P-POSSUM score predicted the mortality with good calibration (H–L test; \(P = 0.646\)); however, it showed to be a poor discriminator of outcomes (0.6–0.7 AUROC). Observed-to-expected mortality ratio was no different from predicted, with O: E mortality 0.87 (0.93–0.062). No differences between the different types of surgery were found. Additionally, this study confirms a correlation between mortality and disability evaluated by WHODAS (\(P = 0.000\)) and the degree of frailty, represented by CFS (\(P = 0.000\)).

POSSUM is one of the most widely validated score systems.\[^{19}\] In the last years, several studies have been assessing the applicability of POSSUM models in various surgical specialties with no consensus outcomes. Some studies have confirmed the validity of POSSUM models,\[^{20,21}\] whereas others have found no advantages in using it.\[^{15}\] Our results support the fact of P-POSSUM that should not be used to predict mortality for one particular patient and that can continue to be used for audits due to its poor discrimination and good calibration.\[^{22}\] The fact that the model variables are mainly dichotomy and do not represent the continuum of the disease can explain this lack on representing a specific individual outcome.\[^{15}\]

Although all the utilities of P-POSSUM system (such as being quickly calculated without special examinations, being used to make a decision beside operative methods, and being a tool for evaluating the operative skills among institutes), there are others important parameters that should be taken in account and that are not included in this score system.\[^{22}\] Another limitation of POSSUM system is the operative parameters that are collected after the patient been submitted to the surgery and this limits the application of POSSUM as a predicting preoperative instrument.

Previous studies had demonstrated that WHODAS 2.0 is good at predicting postoperative disability and recovery.\[^{23,24}\] Pedro-Cuesta et al. demonstrated that an increase in WHODAS 2.0 score is an independent predictor for a high risk of death in nonhospitalized patients with chronic obstructive pulmonary disease, chronic heart failure, and stroke.\[^{25}\] In our study, a positive correlation was demonstrated between the six major life domains included in this score and the impact in the patients’ outcome after surgery. With this result, applicability of WHODAS could be expanded and used as a preoperative mortality predicting score. One of major advantages of this score is asking about the limitations in the last 30 days and showing how was the mean state of the patient for that time.

The CFS is a measure of frailty based on clinical judgment that takes into account cognition, mobility, function, and comorbidities.\[^{26}\] Several studies had confirmed an association between an high CFS grade and mortality.\[^{27,28}\] Despite being a semiquantitative, subjective scale and with a predisposition for an interobserver variability, CSF does not appear to reduce its capacity to predict outcomes. This result is like other studies that had demonstrated that frailty is an indicator of patient’s vulnerability that is high associated with adverse outcomes such as mortality.\[^{27,28}\] Despite the main disadvantage of being a subjective scale, it is a more realistic reflection of routine clinical practice, which often requires a clinical impression. Frailty was identified to be a major predictor of postoperative complications and death after scheduled or unscheduled surgery.

The absence of mortality difference between surgical specialties may prove that the type of surgery itself is no more important than the functional state of the patient.
There are several limitations that must be considered. The fact of selecting an elderly population makes a high-risk group for itself. However, it is important to study this specific population group because it is one of the groups most submitted to surgery and with more doubts on surgery’s decision. Another limitation is that this study has been conducted at a single center, including 229 patients that were distributed by different areas with few patients representing each area. Different surgery specialties were evaluated and this could be a limitation by itself. The reduced number of deaths in this study was too small and may be a limitation on data evaluation. Expanding this study to other institutions would improve the findings of the study.

These results support that some old people have a good physical and cognitive status and are active, showing that age is no more synonymous of adverse surgical outcomes; however, some patients suffer from multiple disabilities. That condition appears to be of major importance when surgical procedures are considered. Frailty is an important variable in this study, showing a good correlation with outcome, and that could be a preoperative parameter used more often in clinical practice. Surprisingly, the most subjective scales had the best performance predicting mortality. It is expected that clinical disability scores will be used as a predicting tool more often.

Taken together, these results highlight the importance of a careful decision in patients with high POSSUM, CFS, and WHODAS grade. This decision should be made with an assessment of patients’ expectations, life expectancy, and the probability of functional recovery.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References
1. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel JP, et al. The World report on ageing and health: A policy framework for healthy ageing. Lancet 2016;387:2145-54.
2. Brinson Z, Tang VL, Finlayson E. Postoperative functional outcomes in older adults. Curr Surg Rep 2016;4:21.
3. Vaiserman A, Lushchak O. Implementation of longevity-promoting supplements and medications in public health practice: Achievements, challenges and future perspectives. J Transl Med 2017;15:160.
4. Amrock LG, Neuman MD, Lin HM, Deiner S. Can routine preoperative data predict adverse outcomes in the elderly? Development and validation of a simple risk model incorporating a chart-derived frailty score. J Am Coll Surg 2014;219:684-94.
5. Christmas C, Makary MA, Burton JR. Medical considerations in older surgical patients. J Am Coll Surg 2006;203:746-51.
6. Mizumoto K, Morita E. Evaluation of the Physiological and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM) scoring system in elderly patients with pressure sores undergoing fasciocutaneous flap-reconstruction. J Dermatol 2009;36:30-4.
7. Green D, Bidd H, Rashid H. Multimodal intraoperative monitoring: An observational case series in high risk patients undergoing major peripheral vascular surgery. Int J Surg 2014;12:231-6.
8. de Groot V, Beckerman H, Lankhorst GJ, Boutier LM. How to measure comorbidity: A critical review of available methods. J Clin Epidemiol 2003;56:221-9.
9. Moreira A, Alvarelhão J, Silva AG, Costa R, Queirós A. Tradução e validação para português do WHODAS 2.0-12 itens em pessoas com 55 ou mais anos [Validation of a Portuguese version of WHODAS 2.0-12 items in people aged 55 or more]. Rev Portuguesa de Saúde Públ 2015;33:179-82.
10. Silva C, Coleta I, Silva AG, Amaro A, Alvarelhão J, Queirós A, et al. Adaptation and validation of WHODAS 2.0 in patients with musculoskeletal pain. Rev Saude Publica 2013;47:752-8.
11. 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.
12. Copeland GP, Jones D, Walters M. POSSUM: A scoring system for surgical audit. Br J Surg 1991;78:355-60.
13. Lai D, Hardy RJ, Tsai SP. Statistical analysis of the standardized mortality ratio and life expectancy. Am J Epidemiol 1996;143:832-40.
14. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol 1982;115:92-106.
15. Scott S, Lund JN, Gold S, Elliott R, Vater M, Chakrabarty MP, et al. An evaluation of POSSUM and P-POSSUM scoring in predicting post-operative mortality in a level 1 critical care setting. BMC Anesthesiol 2014;14:104.
16. Bewick V, Cheek L, Ball J. Statistics review 13: Receiver operating characteristic curves. Crit Care 2004;8:508-12.
17. Cook NR. Statistical evaluation of prognostic versus diagnostic models: Beyond the ROC curve. Clin Chem 2008;54:17-23.
18. Preston SD, Southall AR, Nel M, Das SK. Geriatric surgery is about disease, not age. J R Soc Med 2008;101:409-15.
19. Barnett S, Moonesinghe SR. Clinical risk scores to guide perioperative management. Postgrad Med J 2011;87:535-41.
20. Sharrock AE, McLachlan J, Chambers R, Bailey IS, Kirkby-Bott J. Emergency abdominal surgery in the elderly: Can we predict mortality? World J Surg 2017;41:402-9.
21. Ng K, Yii M. Possum-A model for surgical outcome audit in quality care. Med J Malaysia 2003;58:516-21.
22. Merad F, Baron G, Pasquet B, Hennet H, Kohlmann G, Warfin F, et al. Prospective evaluation of in-hospital mortality with the P-POSSUM scoring system in patients undergoing major digestive surgery. World J Surg 2012;36:2320-7.
23. Shulman MA, Myles PS, Chan MT, McIlroy DR, Wallace S, Ponsford J. Measurement of disability-free survival after surgery. Anesthesiology 2015;122:524-36.
24. Ida M, Naito Y, Tanaka Y, Matsunari Y, Inoue S, Kawaguchi M. Feasibility, reliability, and validity of the Japanese version of the 12-item world health organization disability assessment schedule-2 in preoperative patients. J Anesth 2017;31:539-44.
25. de Pedro-Cuesta J, Garcia-Sagredo P, Alcalde-Cabero E, Alberquilla A, Damián J, Bosca G, et al. Disability transitions after 30 months in three community-dwelling diagnostic groups in Spain. PLoS One 2013;8:e77482.
26. Wallis S, Wall J, Biram R, Romero-Ortuno R. Association of the clinical frailty scale with hospital outcomes. QJM 2015;108:943-9.
aortic valve replacement. Circulation 2017;135:2013-24.

28. Ritt M, Ritt JI, Sieber CC, Gassmann KG. Comparing the predictive accuracy of frailty, comorbidity, and disability for mortality: A 1-year follow-up in patients hospitalized in geriatric wards.

29. Alfaadhel TA, Soroka SD, Kiberd BA, Landry D, Moorhouse P, Tennankore KK. Frailty and mortality in dialysis: Evaluation of a clinical frailty scale. Clin J Am Soc Nephrol 2015;10:832-40.