Journal of Critical Care 43 (2018) 148–155

Contents lists available at ScienceDirect

Journal of Critical Care

journal homepage: www.jccjournal.org

The ability of intensive care unit physicians to estimate long-term prognosis in survivors of critical illness

Ivo W. Soliman, M.Sc., M.D. [a,*], Olaf L. Cremer, M.Sc., M.D., Ph.D. [a], Dylan W. de Lange, M.Sc., M.D., Ph.D. [a], Arjen J.C. Slooter, M.Sc., M.D., Ph.D. [a], Johannes (Hans) J.M. van Delden, M.Sc., M.D., Ph.D. [c], Diederik van Dijk, M.Sc., M.D., Ph.D. [a], Linda M. Peelen, M.Sc., Ph.D. [a,b]

[a] Department of Intensive Care Medicine, University Medical Center Utrecht, Utrecht University, The Netherlands
[b] Department of Epidemiology, Julius Center for Health Sciences and Primary Care University Medical Center Utrecht, Utrecht University, The Netherlands
[c] Department of Medical Humanities, Julius Center for Health Sciences and Primary Care University Medical Center Utrecht, Utrecht University, The Netherlands

ARTICLE INFO

Keywords:
Long-term outcomes
Intensive care
Critical care
Intuitive physician prognosis
Subjective prognosis
Health related quality of life

ABSTRACT

Purpose: To assess the reliability of physicians' prognoses for intensive care unit (ICU) survivors with respect to long-term survival and health related quality of life (HRQoL).
Methods: We performed an observational cohort-study in a single mixed tertiary ICU in The Netherlands. ICU survivors with a length of stay >48 h were included. At ICU discharge, one-year prognosis was estimated by physicians using the four-option Sabadell score to record their expectations. The outcome of interest was poor outcome, which was defined as dying within one-year follow-up, or surviving with an EuroQoL5D-3 L index <0.4.
Results: Among 1399 ICU survivors, 1068 (76%) subjects were expected to have a good outcome; 243 (18%) a poor long-term prognosis; 43 (3%) a poor short-term prognosis, and 45 (3%) to die in hospital (i.e. Sabadell score levels). Poor outcome was observed in 38%, 55%, 86%, and 100% of these groups respectively (concomitant c-index: 0.61). The expected prognosis did not match observed outcome in 365 (36%) patients. This was almost exclusively (99%) due to overoptimism. Physician experience did not affect results.
Conclusions: Prognoses estimated by physicians incorrectly predicted long-term survival and HRQoL in one-third of ICU survivors. Moreover, inaccurate prognoses were generally the result of overoptimistic expectations of outcome.

1. Introduction

The ICU physician is increasingly involved in decision making concerning follow up and post-ICU treatment of patients who have survived ICU care [1]. To do so, it is important to identify patients with an increased risk of poor outcome at the time of ICU discharge [2,3]. Currently prognoses at ICU discharge are largely based on the intuitive insight of the treating physicians. Based on their clinical expertise, they incorporate a patient's condition before ICU admission (medical history, functional status, quality of life and social environment) and the events during hospital and ICU stay into a holistic prognosis for the patient. This contrasts with the 'objective' multivariable prediction models typically used in ICU research, which incorporate a patient's vital status, age and pre-existing comorbidities at ICU admission. However, these models typically do not incorporate prior functional status or quality of life [4,5], and are mostly focused on prediction of short-term mortality rather than long-term functional outcomes.

Because of these omissions, researchers have tried to validate the ICU physicians' estimations of the risk of poor ICU outcomes [6,7]. And they directly compared the prognostic performance of physicians' prognoses to those of statistical models [8-10]. A systematic review of such studies showed that at ICU admission, physicians were more accurate in discriminating patients who would die in comparison to contemporary statistical models [10]. However, thus far, only in the domain of the neurologically critically ill were the studies focused on predicting functional status as outcome of interest [11]. Studies in the general ICU population focused on survival alone [8,12]. As a result, it is unknown whether ICU physicians are also accurate at predicting survival in conjunction with quality of life, at the moment a general ICU patient is discharged from the ICU.

Abbreviations: ICU, Intensive Care Unit; HRQoL, Health Related Quality of Life; IRB, Institutional Review Board; APACHE, Acute Physiology and Chronic Health Evaluation; SOFA, Sequential Organ Failure Assessment; EQ5D, EuroQoL 5D-3L™; c-index, concordance index; IQR, Interquartile Range; 95% CI, 95% Confidence Interval.

* Corresponding author.
E-mail addresses: I.W.Soliman@umcutrecht.nl (I.W. Soliman), O.L.Cremer@umcutrecht.nl (O.L. Cremer), D.W.deLange@umcutrecht.nl (D.W. de Lange), A.Slooter-3@umcutrecht.nl (A.J.C. Slooter), J.J.M.vanDelden@umcutrecht.nl (J.(H.)J.M. van Delden), D.vanDijk@umcutrecht.nl (D. van Dijk), L.M.Peelen@umcutrecht.nl (L.M. Peelen).

http://dx.doi.org/10.1016/j.jcrc.2017.09.007
0883-9441/

So, with the increasing attention for the long-term functional consequences of ICU care in all critically ill patients [13,14], prognostic estimations made at ICU discharge should accurately reflect both the probability of long-term survival and that of an adequate health related quality of life (HRQoL) [12,13]. Therefore, the aim of this study was to assess the ability of physicians to accurately prognosticate survivors of critical illness upon ICU discharge with respect to their long-term survival and HRQoL.

2. Methods and materials

2.1. Study design, setting and participants

This study was designed as a cohort study using data prospectively collected for the purpose of benchmarking and follow-up data for quality of care evaluation. The study was performed at the ICU of the University Medical Center Utrecht. This ICU's population is a mix of medical-surgical critically ill patients, including those after major cardiac, neurological (trauma, vascular and oncology related), gastro-intestinal and transplant surgery, and most types of medical patients. We included ICU survivors admitted to the ICU between March 2012 and December 2014, with a ICU length of stay of over 48 h. If patients were readmitted to the ICU during this period, they were only included once (the first ICU admission). Patients under the age of 18 and those who declined to participate in any medical research (through a hospital-wide opt-out procedure) were excluded. The Institutional Review Board (IRB) of the University Medical Center Utrecht waived the need for informed consent when working with anonymised patient and follow-up data after approval of the follow-up protocol (UMC Utrecht IRB protocol number 10/006).

2.2. Data collection and follow-up

The data extracted for this study included patients' demographics, chronic comorbidities, admission diagnosis, severity of disease at admission (acute physiology and chronic health evaluation (APACHE) IV predicted mortality [4]), disease progression during admission (sequential organ failure assessment (SOFA) scores [15]), length of stay, and active treatment restrictions at ICU discharge. All these variables were recorded during ICU admission for every patient as part of regular care. Consequently the data were validated by trained research personnel as part of the National Intensive Care Evaluation collection of ICU benchmarking data.

One year after ICU discharge the municipal registry was checked for the survival status of all patients. If a patient had died, the date of death was extracted. If a patient had survived, the address was extracted from the registry and the patient was sent a questionnaire concerning HRQoL; non-responders were resent the questionnaire and received a reminder per telephone call [16].

2.3. Physician prognosis at intensive care unit discharge

The ICU physician's prognosis at ICU discharge was measured using the Sabadell score [17,18]. The Sabadell score consists of four prognostic categories. The treating ICU physician is supposed to select the one most applicable to the patient's expected outcome: "0 - good prognosis", "1 - long-term poor prognosis (>6 months) with unlimited ICU readmission", "2 - short-term poor prognosis (<6 months); ICU readmission debatable", or "3 - Death expected during hospitalisation, ICU readmission not recommended" [17]. The Sabadell score was translated into Dutch while maintaining similar wording. Assigning a Sabadell score was mandatory for the ICU physician when discharging a patient from the ICU. The score was saved into the electronic patient file, but was not made available during post-ICU treatment. The physicians were not informed about the use of this score for the current study during data collection. Next to Sabadell score, we assigned the physician completing the Sabadell score into one of three levels of experience and recorded the level. At level 1 were first year registrars (around 3 years of experience; US: residents), at level 2 were sixth year registrars (around 5 years of experience; US: fellows) and at level 3 were consulting intensivists (at least 8 years of experience; US: board certified/attending).

2.4. Study outcomes

The main study outcomes of interest were survival up to one year and long-term poor outcome. We defined poor outcome on the composite of survival and HRQoL. This was operationalised as dying within a year after ICU discharge or surviving follow-up with a low HRQoL [16]. For this study the EuroQoL 5D-3L™ (EQ5D) was used to assess HRQoL [19] [see Additional file 1]. A low HRQoL was defined as an EQ5D index below 0.4. Any patients with a low HRQoL according to this threshold, reported at least one severe EQ5D disability [16,19], and would fall below the average HRQoL of patients with moderate to severe physical [20], cognitive [21], or psychiatric disabilities [22].

2.5. Missing data

Missing data were likely to occur in one-year survivors not responding to the HRQoL questionnaire. As ignoring possible selective missingness typically leads to biased results [23,24], multiple imputation techniques were used to replace the missing EQ5D responses per dimension of the questionnaire for non-responding one-year survivors [23-27]. See Additional file 1 for more details on missing rates and imputation techniques.

2.6. Statistical analyses

Before analysing outcomes, the association between patient characteristics and the prognosis assigned at ICU discharge was studied by comparing characteristics across Sabadell score groups. This was done using a Chi-square test for comparisons of categorical variables and the Kruskal-Wallis test for comparisons of continuous variables. One-year survival was studied using Kaplan-Meier survival curves for the different ICU physician prognosis groups. The survival curves were compared using the log-rank test. Further, the physician's prognosis' predictive performance for one-year survival and for poor outcome were studied by the discrimination, as measured by the concordance index (c-index) (obtained by rank correlation for censored data when studying survival outcomes separately).

Second, patient characteristics were assessed for having an association with a prognosis at ICU discharge which did not match the observed long-term outcome. To this end, patients were grouped into overpessimistic and overoptimistic prognosis groups. The overpessimistic group contained patients where death was expected during hospital stay but which did not occur, and patients who were assigned a short-term poor prognosis while those patients survived follow-up with a high HRQoL. The overoptimistic group contained patients where a good prognosis was expected at ICU discharge but the patient either died within a year or survived with a low HRQoL, and patients assigned a long-term poor prognosis while the patient died within a year of follow-up. The remaining patients made up the correct prognosis group. Patient characteristics were compared between the overoptimistic and correct prognosis group, and between the overpessimistic and correct prognosis group separately.

Finally, to study the effects of physician experience we stratified our data along three levels of experience in a post-hoc analysis. Within these three subgroups agreement between estimated prognosis and observed outcome was tabulated, and the c-index for discrimination was calculated.

For all statistical analyses pooling across imputation datasets was performed using Rubin's rules for the appropriate test statistic. A p-value of 0.05 was considered statistically significant. Analyses were

performed using R version 3.2.2 (R Foundation for Statistical Computing, 2015).

3. Results

3.1. Study population

Among 1676 unique patients enrolled during the inclusion period having a length of stay over 48 h, 1419 (84.7%) survived their ICU stay. Of these eligible patients, three were excluded because of a missing Sabadell score. Seventeen patients were considered lost to follow-up because they were not registered in the Dutch municipal registry, or because there was no available address. This left 1399 patients to be included in this study (fig. 1).

In the total study population, the median ICU length of stay was 5 days (interquartile range (IQR): 3.1–10.1 days). Twenty-three percent of patients were admitted after elective surgery. Admissions after coronary artery bypass graft or cardiac valve surgery surmised 14% of the total population. The APACHE IV predicted probability of mortality, showing disease severity at admission, had a median of 0.14 (IQR: 0.05–0.38). Further patient details at and during admission are presented in Table 1.

Sabadell scores of 0, 1, 2 and 3 were found in 1068 (76%), 243 (18%), 43 (3%), 45 (3%) patients, respectively. Higher Sabadell scores (corresponding to poorer physician estimates of prognosis) were associated with increased ICU length of stay, a decreased proportion of elective surgery patients, an increased proportion of patients admitted with a metastasized neoplasm, confirmed infection or stroke, and an increased APACHE IV predicted mortality at ICU admission (table 1).

3.2. Predictive performance of physicians at ICU discharge

Overall, 322 (23%) of included ICU survivors died during the subsequent year. The four prognostic groups showed significantly distinct survival curves (log-rank $p < 0.001$; fig. 2). The mortality rates at hospital discharge and at one year are shown in Additional file 2 Table E. Based on the prognosis groups the c-index for discrimination of one year survival was 0.69 (95%CI: 0.66–0.72).

Low HRQoL was observed in 287 (27%) of the 1077 patients who survived beyond the first year. Consequently, poor outcome (as defined by the composite of death or low HRQoL), occurred in 609 (44%) of all patients included in this study. When grouped according to estimated physician prognosis, poor outcome occurred in 393 (38%), 134 (55%), 37 (86%), and 45 (100%) patients with Sabadell scores of 0, 1, 2 and 3, respectively (see Additional file 2 Table E). The resulting c-index for discrimination of poor outcome based on physician prognosis was 0.61 (95%CI: 0.60–0.66).

3.3. Overpessimistic and overoptimistic prognosis groups

An overpessimistic prognosis, i.e. assigning a worse prognosis at ICU discharge than was observed at the end of one year follow-up, occurred in 0.4% ($n = 6$) of patients. The overoptimistic group consisted of 36%

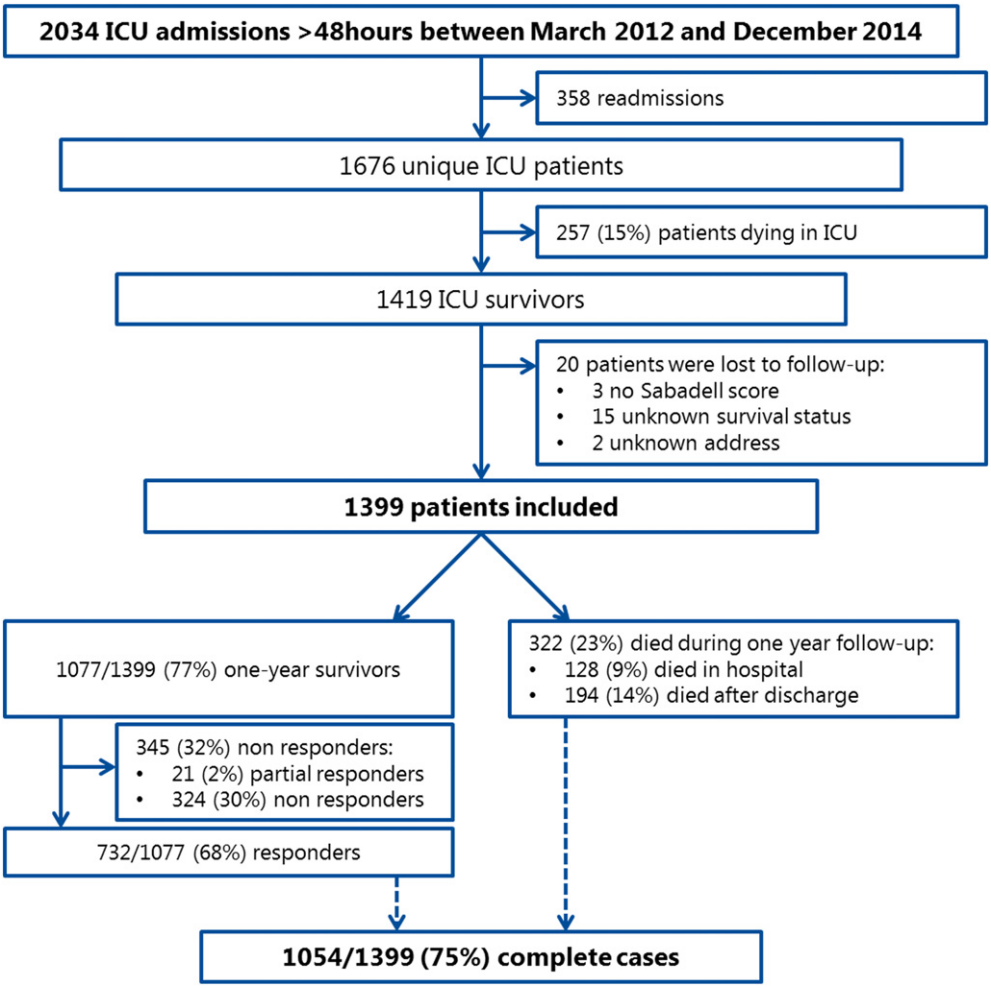

Fig. 1. Flowchart of inclusion. ICU, intensive care unit.

Table 1
Patient characteristics.

Patient characteristics	Total population ($N = 1399$)	Sabadell score groups[a]				
		0 ($n = 1068$)	1 ($n = 243$)	2 ($n = 43$)	3 ($n = 45$)	p-Value[b]
Sex (female)	497 (35.5%)	380 (35.6%)	84 (34.6%)	16 (37.2%)	17 (37.8%)	0.969
Age, years	62 [50–71]	61 [49–70]	63 [50–74]	65 [56–75]	66 [58–73]	<0.001
Hospital LoS, days	21.5 [12.7–38.7]	20.6 [12.3–36]	29.7 [15.9–47.8]	28.9 [14.7–50.4]	11.5 [7.5–14.6]	<0.001
ICU LoS, days	5.4 [3.1–10.1]	4.9 [3–9]	7.6 [3.9–14.6]	9.6 [4.4–17]	6.4 [3.8–10.8]	<0.001
pre-ICU hospital LoS, days	0.4 [0–1.6]	0.4 [0–1.4]	0.4 [0–2.3]	0.2 [0–0.8]	0.1 [0–1]	0.019
post-ICU hospital LoS, days	12.2 [5.9–22.8]	12.1 [6–21.9]	17 [7.1–27]	18 [4.4–25.6]	1.4 [0.6–4.3]	<0.001
Admission type						<0.001
Elective Surgical	324 (23.2%)	286 (26.8%)	37 (15.2%)	1 (2.3%)	0 (0%)	
Urgent Surgical	411 (29.4%)	327 (30.6%)	56 (23%)	10 (23.3%)	18 (40%)	
Medical	664 (47.5%)	455 (42.6%)	150 (61.7%)	32 (74.4%)	27 (60%)	
Admission diagnosis						<0.001
Cardiac surgery (CABG or valve)	199 (14.2%)	184 (17.2%)	15 (6.2%)	0 (0%)	0 (0%)	
Sepsis	218 (15.6%)	149 (14%)	55 (22.6%)	11 (25.6%)	3 (6.7%)	
Subarachnoid haemorrhage	53 (3.8%)	33 (3.1%)	10 (4.1%)	6 (14%)	4 (8.9%)	
Traumatic brain injury	124 (8.9%)	94 (8.8%)	16 (6.6%)	7 (16.3%)	7 (15.6%)	
Cardiac, non-surgical	122 (8.7%)	88 (8.2%)	23 (9.5%)	2 (4.7%)	9 (20%)	
Other	683 (48.8%)	520 (48.7%)	124 (51%)	17 (39.5%)	22 (48.9%)	
Comorbidities						
AIDS/HIV	12 (0.9%)	9 (0.8%)	3 (1.2%)	0 (0%)	0 (0%)	0.760
Chronic cardiovascular (NYHA IV)	167 (11.9%)	134 (12.5%)	31 (12.8%)	2 (4.7%)	0 (0%)	0.032
Chronic dialysis	24 (1.7%)	16 (1.5%)	8 (3.3%)	0 (0%)	0 (0%)	0.144
Chronic kidney insufficiency	102 (7.3%)	78 (7.3%)	19 (7.8%)	4 (9.3%)	1 (2.2%)	0.558
Chronic respiratory	135 (9.6%)	93 (8.7%)	33 (13.6%)	7 (16.3%)	2 (4.4%)	0.030
Cirrhosis	17 (1.2%)	12 (1.1%)	3 (1.2%)	1 (2.3%)	1 (2.2%)	0.826
COPD	149 (10.7%)	114 (10.7%)	27 (11.1%)	4 (9.3%)	4 (8.9%)	0.963
Diabetes	205 (14.7%)	163 (15.3%)	33 (13.6%)	7 (16.3%)	2 (4.4%)	0.223
Immunological insufficiency	159 (11.4%)	113 (10.6%)	34 (14%)	6 (14%)	6 (13.3%)	0.427
Metastasized neoplasm	30 (2.1%)	14 (1.3%)	8 (3.3%)	4 (9.3%)	4 (8.9%)	<0.001
Acute complications on the first day of ICU admission						
Acute kidney injury	117 (8.4%)	81 (7.6%)	27 (11.1%)	6 (14%)	3 (6.7%)	0.160
Cerebrovascular incident (stroke)	195 (13.9%)	121 (11.3%)	39 (16%)	17 (39.5%)	18 (40%)	<0.001
CPR	298 (21.3%)	216 (20.2%)	61 (25.1%)	14 (32.6%)	7 (15.6%)	0.073
Dysrhythmia	119 (8.5%)	83 (7.8%)	20 (8.2%)	4 (9.3%)	12 (26.7%)	<0.001
Gastro-intestinal bleed	142 (10.2%)	102 (9.6%)	28 (11.5%)	5 (11.6%)	7 (15.6%)	0.481
Cranial mass effect	20 (1.4%)	15 (1.4%)	4 (1.6%)	0 (0%)	1 (2.2%)	0.823
Confirmed infection	193 (13.8%)	117 (11%)	43 (17.7%)	14 (32.6%)	19 (42.2%)	<0.001
Mechanical ventilation at ICU admission	1237 (88.4%)	963 (90.2%)	202 (83.1%)	35 (81.4%)	37 (82.2%)	0.004
Mechanical ventilation within 24 h	1313 (93.9%)	1009 (94.5%)	224 (92.2%)	41 (95.3%)	39 (86.7%)	0.107
APACHE IV predicted mortality	0.14 [0.05–0.38]	0.12 [0.05–31]	0.22 [0.08–0.48]	0.43 [0.22–62]	0.52 [0.33–0.72]	<0.001
Sum of SOFA scores during ICU stay	33 [20–65]	31 [19–57]	45 [24–96]	57 [30 − 102]	43 [24–80]	<0.001
Totalmax SOFA score	10 [7–13]	10 [7–13]	10 [7–13]	10 [9–12]	10 [8–12]	0.450
Treatment restrictions at ICU discharge	100 (7.1%)	38 (3.6%)	30 (12.3%)	10 (23.3%)	22 (48.9%)	<0.001

Values are expressed as number (percentage) for categorical variables and as median [interquartile range] for continuous variables. LoS, length of stay; ICU, intensive care unit; CABG, coronary artery bypass grafting; AIDS, acquired immune deficiency syndrome; HIV, human immunodeficiency virus; NYHA IV, New York Heart Association classification four; COPD, chronic obstructive pulmonary disease; CPR, cardiopulmonary resuscitation; APACHE IV, acute physiology and chronic health evaluation version four; SOFA, sequential organ failure assessment.

[a] Sabadell score groups: 0, good prognosis; 1, long-term poor prognosis (>6 months) with unlimited ICU readmission; 2, short-term poor prognosis (<6 months); ICU readmission debatable; 3, death expected during hospitalisation, ICU readmission not recommended.

[b] p-Value for difference across Sabadell score groups.

($n = 498$) of all patients. This left 64% ($n = 895$) of patients in the correct prognosis group (see fig. 3).

Table 2 presents descriptive statistics on these three patient groups. Because of the small size of the overpessimistic group, no statistical tests were performed on this group. When comparing the overoptimistic group with the correct prognosis group, a patient with an overoptimistic prognosis was less likely to have been admitted after elective surgery, and more likely to have been admitted with sepsis. Additionally, overoptimism was associated with patients with more frequent chronic kidney insufficiency or diabetes as comorbidities at ICU admission, and with more patients requiring cardiopulmonary resuscitation on the first day of ICU admission.

3.4. Post-hoc analysis: physician experience level

First year registrars, sixth year registrars and consulting intensivists respectively provided the prognosis for 578 (41.3%), 414 (29.6%) and 407 (29.1%) of all patients. The proportions of correct prognoses for these three groups were 62%, 64% and 67% (see supplementary material 2 Table G). The concomitant c-indices for the discrimination of poor outcome were 0.58 (95%CI: 0.57–0.65), 0.62 (95%CI: 0.61–0.66) and 0.65 (95%CI: 0.63–0.67) respectively. Neither the proportions of correct prognoses, nor the c-indices differed significantly from each other.

4. Discussion

We investigated to what extent the physician's estimation of a patient's prognosis at ICU discharge was in accordance with observed long-term outcomes, and found that ICU physicians performed only moderately. Moreover, when studying the predictive performance for survival and HRQoL combined, the discriminatory performance was poor. One third of all ICU survivors experienced an outcome which was particularly worse than what was expected by the ICU physician.

Interestingly, in patients whom the physicians overoptimistically estimated the prognosis, certain pre-existing comorbidities (like chronic kidney insufficiency, chronic obstructive pulmonary disease or diabetes), admission types (medical), admission diagnoses (sepsis) were seen and acute events early during ICU stay occurred more often. Yet,

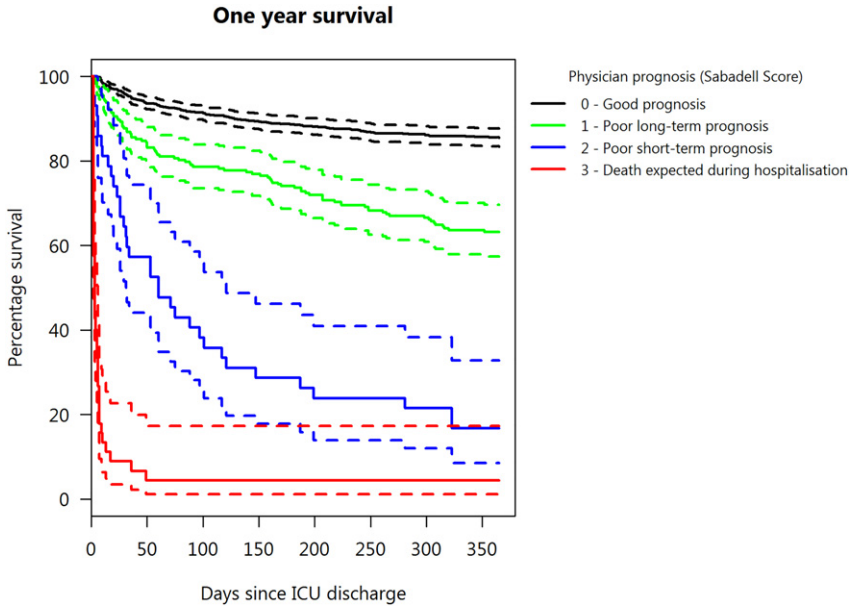

Fig. 2. Kaplan Meier survival curve per prognosis group. ICU, intensive care unit. Dotted lines are 95% confidence intervals. Log rank test for difference in survival distributions *p*-value <0.001.

this overoptimism was not associated with disease severity at admission or organ failure (as measured by the maximum SOFA-score) during ICU admission. These findings make it seem that patient characteristics available at ICU admission can be used to identify patients in which estimating the prognosis may be difficult. Especially when the physician initially feels optimistic about such a patient's outlook, overoptimistic estimations are a potential risk. In clinical practice this should urge the ICU physician to be extra cautious when estimating the individual prognosis for patients with these characteristics. However, we do need to acknowledge these are secondary findings in an explorative part of this study. As such, these findings do not necessarily imply causal links between factors known early after ICU admission and the eventual correctness of a physicians' estimate of long-term outcome.

With regard to previous research on this topic, it was already confirmed that patients with a pessimistic ICU physician prognosis at ICU discharge would have poorer long-term survival than patients where the physician was optimistic would have [17,18,28]. However, there were differences between our study and similar prior ones. For example, in external validations of the Spanish prognostic estimation score (the so-called Sabadell score used in this study), the survival rates differed considerably from ours. The in-hospital mortality of the entire cohorts in these studies was lower than in our cohort (5.3–6.7% versus 9.1%). This mortality difference remained across all four Sabadell score groups: in patients with a good prognosis (1.3–1.5% versus 4.8%), in patients with long-term poor prognosis (3.2–9% versus 9.5%), in patients with short-term poor prognosis (16–23% versus 39.5%), and in patients with expected death in hospital (49–64% versus 82.2%) [18,28]. Further, in an external validation study in the United Kingdom, long-term mortality was also lower than in our study [28] (exact data not available), while the physicians more often seemed to expect poor prognoses than in our study. These differences are likely to be caused by the studies' case-mixes and settings. In particular the exclusion of patients with

Observed outcome Physician prognosis at ICU discharge	Survived with high HRQoL	Survived with low HRQoL	Died during one year of post ICU follow-up
0 - Good prognosis	659 (61.7%)	254 (23.8%)	155 (14.5%)
1 - Long-term poor prognosis (>6 months); with unlimited ICU readmission	110 (45.3%)	44 (18.1%)	89 (36.6%)
2 - Short-term poor prognosis (<6 months); ICU readmission debatable	6 (14.0%)	2 (4.7%)	35 (81.4%)
3 - Death expected during hospitalisation, ICU readmission not recommended	0 (0%)	2 (4.4%)	43 (95.6%)

Fig. 3. Agreement between physician prognosis and observed outcome Values are expressed as number (percentage of row total). ICU, intensive care unit. HRQoL, health related quality of life. Green cells indicate the correct prognosis group, orange cells indicate the overpessimistic group and blue cells indicate the overoptimistic group.

Table 2
Overpessimistic and overoptimistic versus correct prognosis groups.

	Overpessimistic prognosis ($n = 6$)	Correct prognosis ($n = 895$)	Overoptimistic ($n = 498$)	p-Value[a]
Sex (female)	4 (66.7%)	317 (35.4%)	176 (35.3%)	1
Age	65 [55–74]	61 [49–70]	62 [50–73]	0.372
Hospital LoS, days	27.7 [23.1–32.9]	21.5 [13.3–37.9]	21.5 [11.3–40.3]	0.590
ICU LoS, days	14 [9.6–18.7]	5.3 [3.2–10]	5.4 [3–10.1]	0.781
Pre-ICU hospital LoS, days	0.1 [0–0.2]	0.4 [0.1–1.4]	0.4 [0–1.9]	0.592
Post-ICU hospital LoS, days	17.9 [6.7–19.3]	12.9 [5.9–22.8]	12 [5.3–22.7]	0.377
Admission type				0.004
Elective Surgical	0 (0%)	231 (25.8%)	93 (18.7%)	
Urgent Surgical	4 (66.7%)	263 (29.4%)	144 (28.9%)	
Medical	2 (33.3%)	401 (44.8%)	261 (52.4%)	
Admission diagnosis				<0.001
Cardiac surgery	0 (0%)	150 (16.8%)	49 (9.8%)	
Sepsis	1 (16.7%)	118 (13.2%)	99 (19.9%)	
Subarachnoid haemorrhage	1 (16.7%)	39 (4.4%)	13 (2.6%)	
Traumatic brain injury	2 (33.3%)	75 (8.4%)	47 (9.4%)	
Cardiac, non-surgical	0 (0%)	89 (9.9%)	33 (6.6%)	
Other	2 (33.3%)	424 (47.4%)	257 (51.6%)	
Comorbidities				
AIDS/HIV	0 (0%)	6 (0.7%)	6 (1.2%)	0.464
Chronic cardiovascular (NYHA IV)	1 (16.7%)	117 (13.1%)	49 (9.8%)	0.089
Chronic dialysis	0 (0%)	11 (1.2%)	13 (2.6%)	0.092
Chronic renal insufficiency	1 (16.7%)	49 (5.5%)	52 (10.4%)	0.001
Chronic respiratory	1 (16.7%)	93 (10.4%)	41 (8.2%)	0.225
Cirrhosis	0 (0%)	10 (1.1%)	7 (1.4%)	0.830
COPD	1 (16.7%)	80 (8.9%)	68 (13.7%)	0.008
Diabetes	0 (0%)	113 (12.6%)	92 (18.5%)	0.004
Immunological insufficiency	1 (16.7%)	93 (10.4%)	65 (13.1%)	0.158
Metastasized neoplasm	0 (0%)	19 (2.1%)	11 (2.2%)	1
Acute complications on the first day of ICU admission				
Acute kidney injury	1 (16.7%)	76 (8.5%)	40 (8%)	0.844
Cerebrovascular incident (stroke)	2 (33.3%)	124 (13.9%)	69 (13.9%)	1
CPR	2 (33.3%)	158 (17.7%)	138 (27.7%)	<0.001
Dysrhythmia	0 (0%)	92 (10.3%)	27 (5.4%)	0.003
Gastro-intestinal bleed	1 (16.7%)	101 (11.3%)	40 (8%)	0.066
Cranial mass	0 (0%)	14 (1.6%)	6 (1.2%)	0.760
Confirmed infection	3 (50%)	118 (13.2%)	72 (14.5%)	0.560
Mechanical ventilation at ICU admission	5 (83.3%)	799 (89.3%)	433 (86.9%)	0.225
Mechanical ventilation within 24 h	5 (83.3%)	848 (94.7%)	460 (92.4%)	0.097
APACHE IV predicted mortality	0.39 [0.22–0.61]	0.14 [0.05–0.38]	0.15 [0.06–0.37]	0.143
Sum of SOFA scores during ICU stay	90 [68–157]	34 [20–64]	33 [19–65]	0.423
Totalmax SOFA score	12 [11–14]	10 [7–13]	10 [7–13]	0.250
Treatment restrictions at ICU discharge	0 (0%)	61 (6.8%)	39 (7.8%)	0.551

Values are expressed as number (percentage) for categorical variables and as median [interquartile range] for continuous variables. LoS, length of stay; ICU, intensive care unit; CABG, coronary artery bypass grafting; AIDS, acquired immune deficiency syndrome; HIV, human immunodeficiency virus; NYHA IV, New York Heart Association classification four; COPD, chronic obstructive pulmonary disease; CPR, cardiopulmonary resuscitation; APACHE IV, acute physiology and chronic health evaluation version four; SOFA, sequential organ failure assessment.
[a] p-Value for comparison between correct and overoptimistic prognosis groups.

a length of stay <48 h in our study sets it aside from these prior studies [17,18]. Adding the patients with a short ICU stay to our cohort would indeed result in a lower hospital and one-year mortality in all prognosis groups, more alike that of previous studies (see Additional file 2 Table F). However, adding these short stay patients to the study population would not affect the predictive performance as measured by the discrimination for survival (c-index 0.68; 95%CI 0.67–0.70) or for poor outcome including survival and HRQoL (c-index 0.63; 0.62–0.66).

Other studies into the prognostic ability of ICU physicians have focused on the different levels of physician experience [6,29]. In these studies, attending intensivists consistently outperformed ICU fellows with regard to the predictive value of their outcome predictions. In our post-hoc analysis, even the most experienced ICU physicians were incorrect(ly overoptimistic) in over 30% of prognoses, discriminating those with a poor outcome only slightly better than physicians with the least experience. We believe this result may be explained by the timing of estimating the prognosis in our study. Both aforementioned papers studied (daily) prognostication on the first day(s) of ICU admission. When studying prognostication at ICU discharge however, it is more likely that the entire team of physicians has discussed the patient's course of disease and the patient's probable prognosis. This could have positively influenced the prognostic accuracy of the less experienced physicians, bringing the results of the three subgroups closer together.

This study has several strengths. First, patients and clinicians alike will find prognostic performance which includes long-term HRQoL next to long-term survival more relevant for practical use, than prognostic performance for (short-term) survival alone [13,30]. Second, when studying HRQoL as an outcome in an observational study, patients lost to follow-up due to death or non-response need to be taken into account. In this study the composite outcome of survival and HRQoL minimizes any bias which could exist when studying HRQoL in survivors only, while the use of multiple imputation techniques takes into account the possible selective loss to follow-up due to non-response within survivors.

Next to these strengths, this study has limitations which need to be addressed. First, this was a single-centre study. Therefore the results presented here may not generalize to ICUs worldwide, or even other ICUs in the Netherlands, even though similar results were seen in the few prior studies on the Sabadell score. Second, due to its limited options the Sabadell score might not be a tool sensitive enough to capture the prognosis made by the ICU physicians. Third, this study was not set up to verify interrater differences or whether specific ICU physicians influenced concordance between intuitive prognosis and observed outcome. Consequently, we did not have multiple physicians score each individual patient. However, the interrater agreement of two physicians scoring the relatively simple Sabadell score for the same patient has

been reported as moderate at best (interrater Kappa: 0.68) [31]. One might argue that consistently overoptimistic or overpessimistic ICU physicians might have biased our results. Future studies using physician estimates of prognosis could opt to have the patient's prognosis performed by at least two physicians, independently. This would enable the study of specific physicians completing the score (and the association between experience or 'prognostics' training, and prognostication). Lastly, no correction for multiple testing was performed, making the statistical inferences in this study prone to high false positive rates. Corrections were not performed because the aim of our study was to compare the subgroups on a general level, without drawing causal conclusions about specific differences.

To our knowledge, studies which used other tools to record the physicians' prognosis at ICU discharge, such as assigning the probability of outcome directly, are non-existent. Moreover, risk prediction models which could provide statistically modelled prognoses specifically at ICU discharge have been not reported, or are not yet completely developed [32]. Consequently, this is the first study specifically describing the predictive performance of (physician's) prognoses made at ICU discharge for long-term clinically relevant outcomes.

5. Conclusions

The subjective prognosis estimated by ICU physicians incorrectly predicted long-term survival and HRQoL in one out of three ICU patients, regardless of physician experience. This suggests that ICU physicians are currently unable to perform sufficiently reliable risk stratifications in survivors of critical illness with respect to long-term patient-centered outcomes.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jcrc.2017.09.007.

Ethical Approval and Consent to participate

The institutional review board (IRB) of the University Medical Center Utrecht approved the study protocol and waived the need for informed consent when working with anonymised patient and follow-up data (UMC Utrecht IRB protocol number 10/006).

Consent for publication

Not applicable.

Availability of supporting data

A minimal version of this study's research datasets are available from the corresponding author on reasonable request, which also takes into account Dutch Law and good scientific practice for sharing anonymised biomedical patient data.

Competing interests statement

This study was supported by the NutsOhra Foundation, project nr 1404–013, entitled "Prognostics and decision making in prolonged intensive care treatment". Additionally, all authors were appointed as researchers and/or medical doctors by the University Medical Center Utrecht, the academic hospital where this study was performed. Neither the NutsOhra foundation, nor the hospital had influence on any part of the conducting or writing of this study. On behalf of all authors, the corresponding author states there was no conflicts of interest.

Funding

This study was supported by the NutsOhra Foundation, project nr 1404-013, entitled "Prognostics and decision making in prolonged intensive care treatment".

Authors' contributions

IWS contributed to data collection, carried out the data analysis and drafted the manuscript. IWS and OC conceived the study. OC contributed to data collection. DL participated in the study's design and coordination and helped to draft the manuscript. AS contributed to data collection. JD participated in the drafting of the manuscript. DD participated in the drafting of the manuscript and contributed to data collection. LP participated in the data analysis, advised on the methodological design of the study and helped draft the manuscript. All authors contributed to writing of the final manuscript.

Acknowledgements

Not applicable.

References

[1] Schandl A, Bottai M, Holdar U, Hellgren E, Sackey P. Early prediction of new-onset physical disability after intensive care unit stay: a preliminary instrument. Crit Care 2014;18(4):455.
[2] van der Schaaf M, Beelen A, Dongelmans DA, Vroom MB, Nollet F. Functional status after intensive care: a challenge for rehabilitation professionals to improve outcome. J Rehabil Med 2009;41(5):360–6.
[3] Dowdy DW, Eid MP, Sedrakyan A, Mendez-Tellez PA, Pronovost PJ, Herridge MS, et al. Quality of life in adult survivors of critical illness: a systematic review of the literature. Intensive Care Med 2005;31(5):611–20.
[4] Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute physiology and chronic health evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med 2006;34(5):1297–310.
[5] Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3—from evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med 2005;31(10):1345–55.
[6] Barrera R, Nygard S, Sogoloff H, Groeger J, Wilson R. Accuracy of predictions of survival at admission to the intensive care unit. J Crit Care 2001;16(1):32–5.
[7] Meadow W, Weis K, Kim A, Patel S, Pohlman A, Hall JB. Foreseeing the future: predicting 6 month outcomes from events in the medical intensive care unit (MICU). Am J Resp Crit Care 2011;183.
[8] Litton E, Ho KM, Webb SA. Comparison of physician prediction with 2 prognostic scoring systems in predicting 2-year mortality after intensive care admission: a linked-data cohort study. J Crit Care 2012;27(4):423 (e9-15).
[9] Scholz N, Basler K, Saur P, Burchardi H, Felder S. Outcome prediction in critical care: physicians' prognoses vs. scoring systems. Eur J Anaesthesiol 2004;21(8):606–11.
[10] Sinuff T, Adhikari NK, Cook DJ, Schunemann HJ, Griffith LE, Rocker G, et al. Mortality predictions in the intensive care unit: comparing physicians with scoring systems. Crit Care Med 2006;34(3):878–85.
[11] Finley Caulfield A, Gabler L, Lansberg MG, Eyngorn I, Mlynash M, Buckwalter MS, et al. Outcome prediction in mechanically ventilated neurologic patients by junior neurointensivists. Neurology 2010;74(14):1096–101.
[12] Afessa B, Keegan MT. Predicting mortality in intensive care unit survivors using a subjective scoring system. Crit Care 2007;11(1):109.
[13] Oeyen SG, Vandijck DM, Benoit DD, Annemans L, Decruyenaere JM. Quality of life after intensive care: a systematic review of the literature. Crit Care Med 2010;38(12):2386–400.
[14] Kerckhoffs MC, Soliman IW, Wolters AE, Kok L, van der Schaaf M, van Dijk D. Long-term outcomes of ICU treatment. Ned Tijdschr Geneeskd 2016;160:A9653.
[15] Vincent JL, Moreno R, Takala J, Willatts S, De Mendonca A, Bruining H, et al. The SOFA (sepsis-related organ failure assessment) score to describe organ dysfunction/failure. On behalf of the working group on sepsis-related problems of the European Society of Intensive Care Medicine. Intensive Care Med 1996;22(7):707–10.
[16] Soliman IW, Frencken JF, Peelen LM, Slooter AJ, Cremer OL, van Delden JJ, et al. The predictive value of early acute kidney injury for long-term survival and quality of life of critically ill patients. Crit Care 2016;20(1):242.
[17] Fernandez R, Baigorri F, Navarro G, Artigas A. A modified McCabe score for stratification of patients after intensive care unit discharge: the Sabadell score. Crit Care 2006;10(6):R179.
[18] Fernandez R, Serrano JM, Umaran I, Abizanda R, Carrillo A, Lopez-Pueyo MJ, et al. Ward mortality after ICU discharge: a multicenter validation of the Sabadell score. Intensive Care Med 2010;36(7):1196–201.
[19] Szende A, Oppe M, De Charro F. Comparative review of time-trade-off value sets. In: Szende A, Oppe M, Devlin N, editors. EQ-5D value sets: inventory, comparative review and user guide. Dordrecht: Springer Netherlands; 2007. p. 21–7.
[20] Green C, Kiebert G, Murphy C, Mitchell JD, O'Brien M, Burrell A, et al. Patients' health-related quality-of-life and health state values for motor neurone disease/amyotrophic lateral sclerosis. Qual Life Res 2003;12(5):565–74.
[21] Winter Y, Korchounov A, Zhukova TV, Bertschi NE. Depression in elderly patients with Alzheimer dementia or vascular dementia and its influence on their quality of life. J Neurosci Rural Pract 2011;2(1):27–32.

[22] Hayhurst H, Palmer S, Abbott R, Johnson T, Scott J. Measuring health-related quality of life in bipolar disorder: relationship of the EuroQol (EQ-5D) to condition-specific measures. Qual Life Res 2006;15(7):1271–80.
[23] Rubin DB. Multiple imputation for non response in surveys. New York: Wiley; 1987.
[24] Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;59(10):1087–91.
[25] Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. JAMA 2015;314(18):1966–7.
[26] Groenwold RH, Donders AR, Roes KC, Harrell Jr FE, Moons KG. Dealing with missing outcome data in randomized trials and observational studies. Am J Epidemiol 2012; 175(3):210–7.
[27] van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1–67.
[28] Dempsey G, Astles T. Use of the sabadell score to predict hospital and 6-month outcome following critical care admission in a UK centre. Intensive Care Med 2013; 39:S505-S.
[29] Neville TH, Wiley JF, Holmboe ES, Tseng CH, Vespa P, Kleerup EC, et al. Differences between attendings' and fellows' perceptions of futile treatment in the intensive care unit at one academic health center: implications for training. Acad Med 2015; 90(3):324–30.
[30] Turnbull AE, Davis WE, Needham DM, White DB, Eakin MN. Intensivist-reported facilitators and barriers to discussing post-discharge outcomes with intensive care unit surrogates. A qualitative study. Ann Am Thorac Soc 2016;13(9):1546–52.
[31] Herrejon EP, Rico-Cepeda P, Diaz-Diaz D, de la Gandara AM, Martinez MV, Garcia BB. Validation of the sabadell score. Intensive Care Med 2009;35:177.
[32] Schandl A, Bottai M, Hellgren E, Sundin O, Sackey PV. Developing an early screening instrument for predicting psychological morbidity after critical illness. Crit Care 2013;17(5):R210.