Severity scores in ICU

DM seminar

Dr.Vamsi Krishna Mootha
Dept of Pulmonary medicine, PGIMER

Introduction

Severity of illness scoring systems are developed to evaluate delivery of care and provide prediction of outcome of groups of critically ill patients who are admitted to intensive care units

Why do we need severity scoring systems ?

Potential uses of severity scoring systems:

- In RCTs and clinical research
- To assess ICU performance
- Assess individual patient prognosis and guide care
- Administrative purposes

Ideal variables

- Objective
- Simple
- Well defined
- Reproducible
- Widely available
- Continuous variable

- Measurements or data collected routinely in the course of patient care

Appraisal of scoring systems

- Reliability : good inter-observer agreement
- Content validity: appropriate number of variables
- Methodological rigor:
  avoidance of local bias
  avoidance of unusual data
  consistency in data collection
  rules dealing with missing data
- Discrimination
- Calibration

Discrimination

- Ability of score to distinguish a survivor from non survivor

- Area under ROC:

  0.5 – chance performance

  1 perfect prediction

- 0.8 accepted cut-off

  randomly selected non-survivor has 90% chance of having severe score than survivor

Calibration

- Agreement between observed and predicted mortality in within severity strata
- Hosmer – Lameshow goodness of fit test
- Lower the X^2 value, better the calibration

Severity scores in Medical and Surgical ICU

1980-85:	1986-1990:	1990-95:	1996-2000:	2000- current
APACHE SAPS APACHE II	SAPS II MPM	APACHE III MODS MPM II ODIN	SOFA CIS	SAPS III APACHE IV

Numerous scoring systems for trauma, burns, cardiac surgery patients
Numerous scoring systems for pediatric patients

APACHE II

- Disease specific scoring system
- Validated in 5815 ICU patients
- APACHE II score (0-67) is sum of
  - Acute physiology score
  - GCS
  - Age and chronic health score
- Hospital mortality is predicted from
  - APACHE II score
  - Diagnostic category (choose 1 from 50)
  - Need for emergency surgery

APACHE II

- Predicted death rate:
  Logit = -3,517+(Apache II) 0.146

- Predicted Death Rate (adjusted):
  Logit = -3,517+(Apache II) 0.146 +diagnostic category weight

- Limitations:
  Failure to compensate for lead time bias
  Requirement to choose one disease
  Poor inter-observer reliability

Crit care med 1985;13:818-829

APACHE III

- To improve the risk prediction by reevaluating the selection and weighting of physiologic variables
- To examine how differences in patient selection for and timing of admission to ICUs related to outcome
- To clarify the distinction between risk of mortality within independently defined patient groups Vs individual risks
- The selection of patients and the timing of scoring

APACHE III

- Not applicable for patients aged <16 yrs and patients with burns/ MI
- APACHE III score (0-299) is sum of

  Acute physiology score (0-252)

  Age score (0-24)

  Chronic health score (0-23)

  Worst values in 24 hrs taken into account

- Predicted mortality calculated form

  APACHE III score

  Disease category (1 from 78) (reason for ICU admission)

  Patients prior location (ward, other ICU, ICU readmission)

APACHE III

- A 5 point increase in APACHE III score is associated with a statistically significant increase in the relative risk of hospital death

- With first-day APACHE III equation 95 percent of ICU admissions could be given a risk estimate for hospital death that was within 3% of observed

Limitations:

- Should be used to risk stratification rather than risk prediction
- Calculating risk at admission but not on subsequent days

NOT IN PUBLIC DOMAIN

APACHE II,III,IV

	ROC	Prediction at 50%probability	Calibration (Goodness of fitness statistic)
APACHE II	0.85	85.5	
APACHE III version (H)	0.90	88.2	48.7
APACHE III version (I)	Unpublished	Unpublished	24.2
APACHEIII (H) in 2003-04 cohort	Unpublished	Unpublished	276

APACHE IV

- Derived from data of 66,272 patients and validated in 44,288 patients
- Not applicable for <16yrs, burns patients and patients shifted from other ICU
- Worst values in 1st 24 hrs are used
- Separate scoring system for post CABG patients
- Disease specific score – includes 116 disease categories

APACHE IV

Limitations:

- Complexity – has 142 variables

  But web-based calculations can be done at *cerner.com*

- Developed and validated in ICUs of USA only

Crit Care Med 2006; 34:1297–1310

SAPS

- Designed to Simplify then existing APS or APACHE

- 14 variables scored from 0-4

  (Age, GCS,HR, SBP, temp., RR, UO, B.urea, hematocrit, TLC, serum glucose, sodium, potassium, bicarbonate)
- Worst values in 24 hrs used

- Developed in 679 patients

- APACHE and SAPS scores in same population showed comparable results

Crit care med 1984;12:975-977

SAPS

- Mortality increases (0 – 80%) with increasing SAPS score
- Cut-off of 14 had sensitivity and specificity of 0.56 and 0.82

SAPS score	Mortality
4	-
5-6	10.7±4.1%
7-8	13.3±3.9%
9-10	19.4±7.8%
11-12	24.7±4.1%
13-14	30.7±5.5%
15-16	32.1±5.1%
17-18	44.2±7.6%
19-20	50.2±9.4%
>20	81.1±5.4%

Crit care med 1984;12:975-977

SAPS II

- First scoring system to use statistical modeling techniques
- 13,152 patients (65% - developmental cohort, 35%- validation cohort)

- Not applicable for patients younger than 18 yrs, burns patients, coronary care and cardiac surgery patients

- 17 variables:

  12 physiological variables, age, type of admission, underlying AIDS, metastatic or hematological malignancy

  Worst values recoded in 1st 24 hrs are used

SAPS II

Parameter	Value (score)						
HR			<40 (11)	40-69 (2)	70-119 (0)	120-159 (4)	>160 (7)
SBP			<70 (13)	70-99 (5)	100-199 (0)	>200 (2)	
Temp					<39°C (0)	>39°C (3)	
PaO2/FiO2	<100(11)	100-199 (9)	>200 (6)				
UO (ml)		<500 (11)	>500(4)		>1000 (0)		
S.Urea					<28 (0)	28-83(6)	>84(10)
TLC(10^3/cc)				<1(12)	1-20 (0)	>20 (3)	
K*				<3(3)	3-4.9 (0)	>5(3)	
Na*				<125 (5)	125-144 (0)	>145 (1)	
Bicarb			<15(6)	15-19 (3)	>20 (0)		
bilirubin					< 4(0)	4-5.9 (4)	>6 (9)
GCS	<6 (26)	6-8 (13)	9-10 (7)	11-13 (5)	14-15 (0)		

Age - score
<40- 0
40-59- 7
60-69- 12
70-74- 15
75-79- 16
>80- 18

Chronic disease:
Metastatic cancer-9
Hemat.malig-10
AIDS-17

Type of admission:
Sched. Surgical- 0
Medical- 6
Emer.surgical-8

JAMA 1993;270(24):2957-2963

SAPS II

- Probability of death is given by the following equation:

  Logit = $\beta 0 + \beta 1$(SAPS II score)$+ \beta 2$ [ln (SAPS II score+1)]

- Area under ROC for SAPS was 0.8 where as SAPSII has a better value of 0.86

- Calibration- C-3.7 (p-0.883)

JAMA 1993;270:2957-2963

SAPS III

- 16784 patients
- Scores based on data collected within 1st hour of entry to ICU

- Allows predicting outcome before ICU intervention occurs
- Better evaluation of individual patient rather than an ICU
- Not effected by Boyd Grounds effect

- But less time for collecting data and can have greater missing information

Intensive Care Med 2005; 31:1345–1355

Sequential Organ Failure Assessment (SOFA)

- Organ dysfunction is a process rather than an event
- Time evaluation of MODS allows understanding of disease process or influence of therapy
- Designed for patients with sepsis and hence named initially as “Sepsis related organ failure assessment”

Table 2 Differences between commonly used scoring systems and the SOFA score

Scoring systems	SOFA score
Evaluate risk of mortality	Evaluate morbidity
Aim = prediction	Aim = description
Often complex	Simple, easily calculated
Does not individualize the degree of dysfunction/failure of each organ usually obtained early after admission	Does individualize the degree of dysfunction/failure of each organ obtained daily

Intensive Care Med (1996) 22:707-710

SOFA

- Maximal SOFA score (during entire ICU stay) of >15 has predicted mortality of 90% and correct classification
- Mean SOFA score for first 10 days is significantly higher in non-survivors
- ΔSOFA :

  44% of non-survivors showed increase compared to only 20% in survivors

  33% of survivors showed decrease compared to 21% of non-survivors
- Mortality rate increases as number of organs with dysfunction increases

Crit Care Med 1998;26:1793-1800

- Maximal SOFA and ΔSOFA have been found good predictors of mortality

Intensive Care Med. 1999 Jul;25(7):686-96
Intensive Care Med. 2000 Aug;26(8):1023-4

- Found useful in cardiac surgery patients also

Chest 2003 ;123(4):1229-39

Multiple organ dysfunction score (MODS)

	0	1	2	3	4
Respiratory Po_2/Fio_2	>300	226-300	151-225	76-150	<75
Renal S. Creatinine (µmol/L)	<100	101-200	201-350	351-500	>500
Hepatic Serum bilirubin (µmol/L)	<20	21-60	61-120	121-240	>240
Cardiovascular (PAR)	<10	10.1-15	15.1-20	20.1-30	>30
Hematological Platelet count (100/ µL)	>120	120-80	80-50	50-20	<20
Neurological (GCS)	15	14-13	12-10	9-7	<7

MODS

- Objective scale to measure organ dysfunction in ICU
- MODS score correlates well with mortality and ICU stay in survivors
- 0% mortality for score 0
  25% mortality for score 9-12
  50% mortality for score 13-16
  75% mortality for score 17-20
  100% mortality for score>20
- Area under ROC – 0.936
- Greater the organ systems that have failed (score>3), higher the mortality
- ΔMODS also predicts mortality and to a greater extent than Admission MODS score

Crit care medicine 1995;23:1638-1652

Logistic Organ Dysfunction System (LODS)

System	Parameter	Value (Score)				
Neurological	GCS	14,15 (0)	13-9(1)	8-6(3)	5-3(5)	
CVS	HR	>140 (1)	140-30(0)	<30(5)		
	SBP	>270(3)	240-269(1)	70-89(1)	69-40 (3)	<40(5)
Hematological system	TLC (1000/cc)	<1 (3)	1-2.4(1)	2.4-50 (0)	>50 (1)	
	Platelet (10^3/cc)	<50 (1)	>50 (0)			
Respiratory system	PaO2	<150(3)	>150 (1)			
Hepatic system	Bilirubin (mg/dl)	<2 (0)	>2 (1)			
	PT	N+3 (0)	>N+3(1)			
Renal system	Urea (mg/dl)	>120 (5)	119-60(3)	59-35 (1)	<35 (0)	
	Creatinine (mg/dl)	>1.16(3)	1.59-1.2 (1)	<1.2 (0)		
	UO (L/24 hr)	>10 (3)	10-0.75 (0)	0.75-0.5 (3)	<0.5 (5)	

LODS

- Worst values in 1st 24 hrs of ICU stay
- Worst value in each organ system
- Total score ranges from 0-22
- Good calibration and discrimination (area under ROC 0.85)

JAMA 1996;276:802-810

Organ dysfunction and/or infection (ODIN)

Table 1. Definitions of organ dysfunctions

I. Respiratory dysfunction (presence of one or more of the following):
   A. $PaO_2 < 60$ mmHg on $FIO_2 = 0.21$
   B. Need for ventilatory support
II. Cardiovascular dysfunction (presence of one or more of the following, in the absence of hypovolemia[a]):
   A. Systolic arterial pressure < 90 mmHg with signs of peripheral hypoperfusion
   B. Continuous infusion of vasopressor or inotropic agents required to maintain systolic pressure > 90 mmHg
III. Renal dysfunction (presence of one or more of the following)[b]):
   A. Serum creatinine > 300 µmol/l
   B. Urine output < 500 ml/24 h or < 180 ml/8 h
   C. Need for hemodialysis or peritoneal dialysis
IV. Neurologic dysfunction (presence of one or more of the following):
   A. Glasgow coma scale ≤ 6 (in the absence of sedation at any one point in day)
   B. Sudden onset of confusion or psychosis
V. Hepatic dysfunction (presence of one or more of the following):
   A. Serum bilirubin > 100 µmol/l
   B. Alkaline phosphatase > 3× normal
VI. Hematologic failure (presence of one or more of the following):
   A. Hematocrit ≤ 20%
   B. White blood cell count $< 2000/mm^3$
   C. Platelet count $< 40000/mm^3$
VII. Infection (presence of one or more of the following associated with clinical evidence of infection):
   A. 2 or more positive blood cultures
   B. Presence of gros pus in a closed space
   C. Source of the infection determined during hospitalization, or at autopsy in case of death within the 24 h

[a] Excluding patients with a central venous pressure less than 5 mmHg
[b] Excluding patients on chronic dialysis before hospital admission

Table 4. Prediction of outcome using logistic regression analysis

Variables	Coefficient	Odds-ratio	*p*-value
Constant	−3.59		<0.0001
Cardiovascular dysfunction	1.19	3.28	<0.0001
Renal dysfunction	1.18	3.25	<0.0001
Respiratory dysfunction	1.09	2.97	<0.0001
Neurologic dysfunction	0.99	2.69	<0.0001
Hematologic dysfunction	0.86	2.36	0.011
Hepatic dysfunction	0.57	1.78	0.055
Infection	0.53	1.70	0.002

Intensive Care Med1993;19:137-144

ODIN

- Easily available data (within the first 24 h of admission), when precise diagnostic evaluation is not possible
- Less subjectivity
- Easy calculation

- Discrimination comparable to SAPS II & APACHE II

Mortality Probability Model (MPM)

- Developed in single ICU
- Not applicable for patients <14yrs, patients with burns, cardiac/ cardiac surgery patients
- Admission MPM (MPM_0) – 11 variables
  MPM at 24 Hrs (MPM_{24}) - 14 variables
  MPM at 48 Hrs (MPM_{48}) - 11 variables
  MPM over the time (MPM_{OT})- MPM_0
  (MPM_{24} - MPM_0)
  (MPM_{48} - MPM_{24})
- Probability is derived directly from these variables
- MPM_{OT} predicted better than MPM_0 for long term patients

Crit care med 1988;16:470-477

MPM_0

Variable	1	0	β
Level of consciousness	Coma / deep stupor	No coma/ deep stupor	2.89
Admission	Emergency	Elective	1.2671
Prior CPR	Yes	No	1.0137
Cancer	Present	Absent	0.94131
CRF	Present	Absent	0.64049
Infection	Probable	Not probable	0.047789
Previous ICU admission in 6mo	Yes	No	0.43946
Surgery before ICU admission	Yes	No	-0.37987
SBP			-0.04591
HR	10 beat/ min relative risk		0.00736
Age	10 years relative risk		0.047789
Constant			-2.9678

MPM II

Developed from 12,610 patients and validated in6514 patients

$\underline{\text{MPM II}_0}$

14 variables

Values within 1 hr of ICU admission

Discrimination: ROC 0.824

Calibration : C-6.21

Only score (apart from SAPS III) to calculate risk at admission

$\underline{\text{MPM II}_{24}}$:

13 variables

Worst values in 24 hrs

Discrimination: ROC- 0.844

Calibration: C-4.94

JAMA 1993;270: 2478-2486

Cellular injury score (CIS)

- Many of currently used Severity scores do not take Lead time bias into account
- CIS uses 3 biochemical parameters

Arterial ketone body ratio	Ratio of acetoacetate and β hydroxy butyrate Marker of redox state in liver Indicates hypoperfusion or dysfynction of liver
Osmolar gap	Unmeasured solutes spilled over from injured cells
Blood lactate levelz	Surrogate for tissue oxygenation

Cellular injury score (CIS)

	0	1	2	3
AKBR	>0.71	0.7-0.41	0.4-0.26	<0.25
OG mOsm/kg	<2.2	2.3-10	10.1-20	>20.1
Blood Lactate mg/dl	<16	17-25	25-50	>51

CIS score 0-1 : 6.7%mortality rate
2-3 : 40%
4-5: 70%
6-7: 82%
8-9: 100%

J Trauma 1998;45(2):304-311

CIS

- Both SOFA score and CIS sequentially reflect the severity of MODS
- Comparable in diagnostic value as predictors of prognosis.
- These findings may indicate the possibility that MODS is a summation of effects of cellular injury

Intensive Care Med. 2000 Dec;26(12):1786-93

Three day recalibrated ICU outcome score (TRIOS)

	Parameter estimate	Odds ratio 95 % CI	*P* value (Wald)	Odds ratio 95 % CI (Bootstrap)
Intercept	–4.44		0.0001	
Transfer from ward (0/1)	0.5543	1.74 (1.25–2.42)	0.001	(1.253–2.453)
LOD at admission	0.1536	1.16 (1.085–1.253)	< 0.0001	(1.093–1.276)
SAPS II admission	0.0388	1.04 (1.027–1.053)	< 0.0001	(1.026–1.053)
Chronic illness (0/1)	0.8507	2.34 (1.677–3.269)	< 0.0001	(1.622–3.296)
SAPS2-SAPS3 alteration	0.4161	1.516 (1.04–2.22)	0.032	(1.055–2.373)
LOD2-LOD3 alteration	0.6940	2.00 (1.29–3.11)	0.0002	(1.292–3.019)

To compute the probability of hospital mortality:
(1) compute the logit:
Logit = (–4.44) + 0.5543 (Transfer) + 0.1536 (LOD) + 0.0388 (SAPS II) + 0.8507 (Chronic illness) + 0.4161 (SAPS2-SAPS3 alteration) + 0.6940 (LOD2-LOD3 alteration);
(2) calculate the probability of hospital mortality [*P*(death)]:
$P(\text{death}) = (e^{\text{Logit}})/(1+e^{\text{Logit}})$ where $e = 2.7182818$ (the base of the natural logarithm).

Intensive care med 2001;27:1012-21

Dynamic monitoring

Misclassifications in scores calculated at admission are due to:

- Exclusion of factors that cannot be measured at ICU admission
- Exclusion of complications occurring during OCU stay
- Exclusion of treatment effects

Dynamic monitoring circumvents some of these:

- MPM 48, MPM 72 calculated at 48 and 72 hrs of admission with same 13 variables with different constant terms have been used

  Critical care med 1994;22:1351-58

- Serial APACHE III score can be used to calculate daily risk

  Critical care med 1994;22:1359-68

- Mean and highest SOFA scores during first

  Crit Care Med 1998;26:1793-1800

Customized probability models

- SAPS III provides customized models for various regions

- SAPS II and MPM II_{24} have been customized for sepsis patients to improve discrimination and calibration

JAMA1995;273:644-650

Other scores

Scores for surgical patients:

Thoracoscore (thoracic surgery)
Lung Resection Score (thoracic surgery)
EUROSCORE (cardiac surgery)
ONTARIO (cardiac surgery)
Parsonnet score (cardiac surgery)
System 97 score (cardiac surgery)
QMMI score (coronary surgery)
Early mortality risk in redocoronary artery surgery
MPM for cancer patients

Scores for Pediatric patients:

PRISM (Pediatric RISk of Mortality)
P-MODS (Pediatric MODS)
DORA (Dynamic Objective Risk Assesment)
PELOD (Pediatric Logistic Organ Dysfunction)
PIM II (Paediatric Index of Mortality II)
PIM (Paediatric Index of Mortality)

Scores for trauma patients:
Trauma Score
Revised Trauma Score
Trauma and injury Severity score (TRISS)
A Severity Characterization of trauma (ASCOT)

Comparison of scoring scales

Study	Year	Country	APACHE II		APACHE III		SAPS II		MPM II	
			chi²	ROC	*chi²*	ROC	*chi²*	ROC	*chi²*	ROC
Sirio	1992	Japan		0.78						
Oh	1993	Hong kong		0.89						
Rowan	1993	UK	79	0.83						
Rowan	1994	UK	81	0.83					251	0.74
Wong	1995	Canada		0.86						
Castella	1995	Europe		0.85		0.86		0.85		0.81
Apolone	1996	Italy					71	0.8		
Bastos	1996	Brazil			400	0.82				
Moreno	1997	Portugal	33	0.79			29	0.82		
Beck	1997	UK	99	0.8	130	0.84				

Comparison of scoring scales

Study	Year	Country	APACHE II		APACHE III		SAPS II		MPM II	
			chi²	ROC	*chi²*	ROC	*chi²*	ROC	*chi²*	ROC
Moreno	1998	Europe					218	0.822	437	0.785
Goldhill	1998	UK	181							
Zimmerman	1998	US			48	0.89				
Sirio	1999	US			2407	0.9				
pappachan	1999	UK			312	0.89				
Markgraf	2000	Germany		0.832		0.846		0.846		
Capuzzo	2000			0.8				0.8		
Arabi Y	2000	S.Arabia		0.83				0.79		0.85
Livigston	2000	UK	366	0.763	67	0.795	142	0.784	452	0.741
Beck	2003	UK	232	0.835	443	0.867	257	0.852		
Kim	2005	Korea			2.58	0.981	4.37	0.978		0.941
Geater A	2007	Thailand	66	0.91			54	0.88		

Limitations of current scoring scales

- Good discrimination
- Poor calibration:

  Patients who had a probability of mortality between 10% and 40%, there was a significantly higher number of observed deaths compared with predicted deaths

  Zimmerman JE et al. Crit Care Med 1998;26:1317-1326

Diagnostic categories and case mix

Lead time bias and source of referral

Discharge practices

Accuracy in data collection, analysis

Indian perspective

- APACHE II, SAPS II, MPM II_0,MPM II_{24} had modest discrimination (area under ROC 0.66–0.78) and poor calibration
- Tendency to under predict hospital death in patients with lower mortality probability estimates
- No differences between the models with regard to discrimination and calibration

AN Aggarwal et al. Respirology 2006; 11: 196–204

- The scoring system also showed a poor calibration as well as discrimination.
- Suggested lowering down the cut-off value in allotment of age points and by awarding the score to factor like co-existing immunocompromised state

Indian J Med Res. 2004;119(6):273-82

Why do we need severity scoring systems ?

Potential uses of severity scoring systems:

- In RCTs and clinical research
- To assess ICU performance
- Assess individual patient prognosis and guide care
- Administrative purposes

Assessing ICU performance

- Confounded by the fact that ICU that admit sicker patients will have higher than predicted mortality
- But can evaluate any given ICU over time
- Assessment of individual intensivist cannot be done as ICU care is a team management

Care of an individual patient

- Scores cannot be used to triage patients

  As all of them have been developed in patients who received ICU care

- Patient selection for therapeutic interventions

  eg: indications for XIGRIS in severe sepsis

- Decision of withdrawal of support cannot be based in current scoring systems as their area under ROC (discrimination) is far less than 0.99

Which score to use?

- APACHE, SAPS, MPM – only of historic significance
- APACHE II – most widely used and quoted in USA
- SAPS II – commonly used in Europe
- APACHE III – not in public domain
- SAPS III, APACHE IV – better design
- MPM, MODS, LODS – uncommonly used

Conclusions

- APACHE II, APACHE III, SAPS II, MPM II give comparable results
- Good discrimination but poor calibration
- Scores need customization before use, at the cost of loss of comparability
- Can be used to compare study population in RCTs, assess ICU
- Not to be used for individual patient management