

PATIENT - CENTRED
ACUTE CARE
TRAINING



AN ESICM MULTIDISCIPLINARY DISTANCE LEARNING PROGRAMME
FOR INTENSIVE CARE TRAINING

Clinical outcome

Professionalism

Update October 2010

Module Authors (Update 2010)

Maurizia CAPUZZO

Intensive Care Unit, Azienda Ospedaliero-Universitaria di
Ferrara, Ferrara, Italy

Rui MORENO

Intensive Care Unit, Hospital de St. António dos Capuchos,
Centro Hospitalar de Lisboa Central, E.P.E., Lisbon, Portugal

Module Authors (first edition)

Saxon Ridley

Intensive Care Unit, Norfolk & Norwich Hospital, Norwich,
United Kingdom

Rui Moreno

Intensive Care Unit, Hospital de St. António dos Capuchos,
Centro Hospitalar de Lisboa Central, E.P.E., Lisbon, Portugal

Module Reviewers

Brian Cuthbertson and Janice Zimmerman

Section Editor

Carl Waldmann



Clinical outcome

Update October 2010

Editor-in-Chief	Dermot Phelan, Intensive Care Dept, Mater Hospital/University College Dublin, Ireland
Deputy Editor-in-Chief	Francesca Rubulotta, Imperial College, Charing Cross Hospital, London, UK
Medical Copy-editor	Charles Hinds, Barts and The London School of Medicine and Dentistry
Editorial Manager	Kathleen Brown, Triwords Limited, Tayport, UK
Business Manager	Estelle Flament, ESICM, Brussels, Belgium
Chair of Education and Training Committee	Hans Flaatten, Bergen, Norway

PACT Editorial Board

Editor-in-Chief	Dermot Phelan
Deputy Editor-in-Chief	Francesca Rubulotta
Acute respiratory failure	Anders Larsson
Cardiovascular dynamics	Jan Poelaert/Marco Maggiorini
Neuro-intensive care and Emergency medicine	Giuseppe Citerio
HSRO/TAHI	Carl Waldmann
Environmental hazards	Janice Zimmerman
Systemic inflammation and Sepsis/Infection	Johan Groeneveld
Metabolism, endocrinology, nephrology, nutrition	Charles Hinds
Perioperative ICM and surgery	Position Vacant
ETC/Ethics	Gavin Lavery
Education and assessment	Lia Fluit
Consultant to the PACT Board	Graham Ramsay

LEARNING OBJECTIVES

After studying this module on Clinical outcome, you should be able to:

1. Define and appropriately select relevant clinical outcomes
2. Measure clinical outcomes
3. Understand the use of severity scores and general outcome prediction model
4. Understand the principles and methods behind the measurement of ICU (intensive care unit) performance

FACULTY DISCLOSURES

The authors of this module reported no disclosures.

DURATION

5 hours

Introduction	1
1/ How to define and appropriately select clinical outcomes.....	2
Patient oriented clinical outcomes	2
Short-term outcome measures.....	2
Long-term outcome measures	3
Complications and adverse events.....	4
Quality of life and functional outcome.....	5
ICU staff orientated outcomes.....	6
Short-term mortality	7
Long-term mortality.....	8
Society oriented outcomes.....	9
Health managers, economists and politicians	9
Views of society as a whole.....	10
2/ How to measure outcome.....	11
Survival	11
ICU survival.....	11
Hospital survival.....	12
Long-term survival.....	13
Beyond survival: health status and quality of life	15
Health status.....	17
Quality of life	19
3/ How to predict outcome	20
Severity evaluation and outcome prediction.....	20
Outcome prediction models	21
Acute Physiology and Chronic Health Evaluation system	22
Simplified Acute Physiology Score system	23
Mortality Probability Models system.....	25
Intensive Care National Audit & Research Centre model	26
Scores of evolution	27
Specific scores.....	28
Use of outcome prediction models.....	29
Patient selection	30
Collection of predictive variables and hospital outcome.....	31
Computation of the severity score	31
Transformation of the severity score to a probability of death	32
Evaluating the performance of a severity score	34
Discrimination.....	34
Calibration	35
Uniformity-of-fit.....	36
Poor predictive performance	37
4/ How to measure and compare performance of ICUs.....	39
Evaluating the performance of the ICU	39
Effectiveness of the ICU and the lack of demonstration of ICU efficacy	39
Comparing ICUs: how to adjust for case-mix.....	40
Standardised mortality ratio and ICU performance.....	40
Limitations of the SMR approach	42
A final comment on performance evaluation.....	43
Conclusion	45
Self-assessment Questions.....	46
Patient challenges.....	49

INTRODUCTION

Outcome refers to the result of a process or an event; clinical outcome refers to the results of any healthcare intervention, including the entire range of activities performed in an intensive care unit (ICU). Despite other outcomes being important for the intensivist e.g. effectiveness in the use of resources, clinical outcomes are the most important measure of critical care activity.

Clinical outcome is crucially important to patients and individuals working both in and outside critical care

There are many clinical outcomes that can be measured. Hospital mortality is the most frequently used outcome in clinical practice but more and more often the quality of life after discharge is valued by patients, families, health workers, policy makers and all of society.



Hinds CJ, Watson JD. *Intensive Care: A Concise Textbook*. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. pp. 15–16; Mortality, long-term survival and quality of life.

Clinical outcomes may be categorised on the basis of:

- **Time:** usually refers to short-term time frames and may be at ICU or hospital discharge or at a fixed time after ICU e.g. 28 days after ICU admission. Longer term time frames relate to time after the patient's discharge from the hospital e.g. at 3, 6 or 12 months.
- **Perspective:** refers to the perspective of the different stakeholders, including patients, caregivers, other staff and hospital administrators, economists and politicians.
- **Purpose:** refers to an endpoint for research and audit. It assesses the performance of the ICU as a whole to compare within a specific ICU over time or with other ICUs (benchmarking).

Link to **ESICM Flash Conference**: Philipp Metnitz, Vienna, Austria. Benchmarking ICUs. Berlin 2007

You will find the following reference helpful in understanding articles comparing clinical outcomes between ICUs.



Needham DM, Dowdy DW, Mendez-Tellez PA, Herridge MS, Pronovost PJ. Studying outcomes of intensive care unit survivors: measuring exposures and outcomes. *Intensive Care Med* 2005; 31(9): 1153–1160. Review. PMID 15909169. Full text (pdf)

Measuring clinical outcomes plays a pivotal role in influencing the way critical care is practised. Advances in both basic science and clinical research are more systematically applied as improvement in clinical outcome is used to drive changes in interventions or treatments. Ultimately as resources become more scarce in relation to the number of individuals requiring healthcare, clinical outcome will be used to allocate funding and to demonstrate efficiency.

Clinical outcome will determine the way all aspects of critical care are delivered

1/ HOW TO DEFINE AND APPROPRIATELY SELECT CLINICAL OUTCOMES

Clinical outcomes are the end result of any therapeutic interventions applied to patients. They may be immediate and short-term such as the effects of vasoactive drugs on the cardiovascular system or long-term when the effects of the whole treatment episode on the patient's well-being are evaluated. Clinical outcomes following intensive care can be assessed from at least four different perspectives:

- Patients and their relatives
- ICU staff
- Health managers, economists and politicians
- The population at large (i.e. society)

The clinical outcome of interest may differ depending on whose perspective is being considered.

Link to **ESICM Flash Conferences**: Ralf Kuhlen, Berlin, Germany. Quality indices for ICUs (European project). Berlin 2007

Rui Moreno, Cascais, Portugal. Risk adjustment: Current situation and future perspectives. Berlin 2007

Jukka Takala, Bern, Switzerland. Severity of illness assessment: Any impact on clinical practice? Berlin 2007

THINK: When considering any question concerning outcome, think about who is asking the question. This will enable a clear focus on the perspective of the questioner and the use of terminology that will be meaningful to the questioner.

Patient oriented clinical outcomes

Outcome evaluation after critical illness has been a rapidly growing area for research. Society used to place an emphasis on objective indices such as whether the patient was able to go back to work, but recently the emphasis has moved towards more subjective, as well as patient-centred outcome data.

Outcome measures may be in the form of mortality in ICU or on the ward afterwards, as well as in the first year or longer after intensive care treatment, or may involve physical, psychological and cognitive data. Outcome measures may be short or long term and may reflect side-effects or complications and adverse incidents arising from intensive care management.

Short-term outcome measures

Firstly, the patients and their relatives are interested in whether or not the patient is going to survive the acute phase of illness responsible for ICU admission. Nevertheless, ICU survival is just a first step, because the ultimate aim of any health system is

What are his chances, doctor?

to return the patient to his/her home and family. Therefore, hospital survival is usually the most important consideration for both patient and family and is the risk usually assessed.

Hospital outcome is primarily determined by age, underlying health status, severity of the acute illness, timing and reason(s) for ICU admission, diagnosis, presence and degree of acute physiological derangement, timing and application of treatments and patient response to treatments. Nowadays, age and underlying health status are responsible for about 50% of the prognostic capability of severity of illness scores, a proportion much greater than some years ago.



Moreno R, Jordan B, Metnitz P. The changing prognostic determinants in the critically ill patient. In: Vincent JL, editor. Yearbook of Intensive Care and Emergency Medicine 2007. Springer-Verlag; 2007. pp. 899-907. ISBN 978-3-540-49432-4

Long-term outcome measures

Patients will also be interested in the duration and quality of their survival. Both patients and their relatives will have a strong interest in knowing whether the patient will be capable of independent living, returning to work or will suffer permanent disability. Don't forget that neuropsychological aspects have an important influence on quality of life of the patient and their family circle.

NOTE To answer patient and family questions about long-term outcome, you must ask about, and appreciate the factors that the individual patient considers important (social environment, physical capability, work and leisure activities). As time passes, severity of illness and acute physiological derangement become less important, and underlying diagnosis and health status become more important for understanding and forecasting long-term outcomes.

Always obtain a thorough history, especially of the patient's prior activities

What defines 'adequate' life is a uniquely personal perception that varies with the individual's view of the value of life and their expectations.

Q. How are you going to obtain relevant background information regarding an unconscious patient?

A. This type of information can be obtained from:

- The relatives, but be aware that close relatives or carers can provide reliable information about physical aspects and social life, but not about emotional aspects and perceived quality of life.
- Co-morbidity and assessments of functional status can sometimes be obtained from the primary care physician.

Link to **ESICM Flash Conference**: Barbara McLean, Atlanta, USA. Patients' perception of the post-operative ICU experience. Berlin 2007

Complications and adverse events

Link to **ESICM Flash Conference**: Andreas Valentin, Vienna, Austria. Do intensivists bring safety or quality? Vienna 2009

Patients will also be interested in outcome measures relating to complications, adverse events and sequelae associated with critical illness.

Complications may reflect the standard of care delivered in the ICU and will also be of interest to clinical staff. Higher rates of nosocomial infection, for example, have been linked to higher mortality. Another aspect that is becoming more understood and relevant is the risks related to intra- and inter-hospital transport of the patient. More information about this association is contained in the following reference and the PACT module on Transportation.



Vincent J-L, Rello J, Marshall J, Silva E, Anzueto A, Martin CD, Moreno R, Lipman J, Gomersall C, Sakr Y, Reinhart K; EPIC II Group of Investigators. International study of the prevalence and outcomes of infection in intensive care units. *JAMA* 2009; 302: 2323–2329. PMID 19952319

Adverse events may be classified as

- Unpreventable e.g. allergic reaction in a patient who did not have previous reactions, or possibly weight loss, critical illness neuropathy and joint stiffness

Critically ill patients may suffer further morbidity in the ICU

- Preventable e.g. complications related to safety issues such as accidental tube or catheter removal, administration of the wrong drug or dosage.

Q. In what other ways might adverse events be classified?

A. Adverse incidents may be classified in several ways. They may be classified by what action precipitated the incident, by whom or how the incident was recognised or rectified. Adverse incidents could also be classified according to when they occurred e.g. during routine care, during a procedure or at another time.

Q. How common do you think adverse events are?

A. Unfortunately complications and adverse events are common in ICU (38.8 events per 100 patient days in one reported prevalence multinational study), and are detected mainly at the times of shift changes and handovers. The commonest complications involve intravascular and other catheters, drains and medication errors.

The Australian Incident Monitoring Study (AIMS) classified adverse events into five major types: Airway and ventilation; Drugs and therapeutics; Procedures and equipment systems; Patient management and environment and ICU management (see Beckmann reference, below).

The prevalence of adverse incidents in 205 ICUs worldwide and the associated ICU- and patient-related factors have been described by Valentin et al. in 2006
Details of the errors in administration of parenteral drugs in ICU can be found in Valentin et al. 2009.



Donchin Y, Gopher D, Olin M, Badihi Y, Biesky M, Sprung CL, et al. A look into the nature and causes of human errors in the intensive care unit. *Crit Care Med* 1995; 23(2): 294–300. PMID 7867355

Beckmann U, Baldwin I, Hart GK, Runciman WB. The Australian Incident Monitoring Study in Intensive Care: AIMS-ICU. An analysis of the first year of reporting. *Anaesth Intensive Care* 1996; 24(3): 320–329. PMID 8805886

Valentin A, Capuzzo M, Guidet B, Moreno RP, Dolanski L, Bauer P, et al; Research Group on Quality Improvement of European Society of Intensive Care Medicine; Sentinel Events Evaluation Study Investigators. Patient safety in intensive care: results from the multinational Sentinel Events Evaluation (SEE) study. *Intensive Care Med* 2006; 32(10): 1591–1598. PMID 16874492. Full text (pdf)

Valentin A, Capuzzo M, Guidet B, Moreno R, Metnitz B, Bauer P, et al; Research Group on Quality Improvement of the European Society of Intensive Care Medicine (ESICM); Sentinel Events Evaluation (SEE) Study Investigators. Errors in administration of parenteral drugs in intensive care units: multinational prospective study. *BMJ* 2009; 338: b814 doi: 10.1136/bmj.b814. PMID 19282436

Q. What can be done to prevent errors in the administration of parenteral drugs?

A. Organisational factors such as an existing critical incident reporting system, an established routine of checks at nurses' shift change and an adequate patient to nurse ratio have been all demonstrated to be associated with reduced risk for such errors.

Quality of life and functional outcome

Other areas of interest to the patient are quality of life and functional outcome.

Link to **ESICM Flash Conferences**: Didier Ledoux, Brussels, Belgium.
Assessment of survival and quality of life in octogenarian 1-year after cardiac surgery. Barcelona 2006

Despoina Koulenti, Athens, Greece. Survival, morbidity and quality of life outcome 18 months after ICU. Barcelona 2006

The terms quality of life and health-related quality of life are often used interchangeably

- Quality of life is a concept encompassing a broad range of physical and psychological characteristics that describe an individual's ability to function and derive satisfaction from doing so.

- Health-related quality of life has been defined as ‘a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’ by the World Health Organization.
- Functional outcome encompasses physical impairment, handicap or disability. Functional outcome is not synonymous with quality of life. Functional outcome concentrates on physical and mental capacity and does not measure well-being or sense of satisfaction.

There are many instruments to measure health-related quality of life

Instruments to measure health-related quality of life cover a broad range of physical and psychological attributes (called domains). Investigators need to select carefully the instrument to be used, according to validity and reliability already demonstrated in similar settings. The two most commonly used instruments and their respective domains are summarised in the table below. For a review of outcome measures, see the following references.



Capuzzo M, Metnitz PGH, Moreno R. Health related quality of life after ICU discharge. In: Kuhlen R, Moreno R, Ranieri M, Rhodes A, editors. 25 Years of Progress and Innovation in Intensive Care Medicine. Berlin: MVV Medizinisch Wissenschaftliche Verlagsgesellschaft; 2007. pp. 429–438. ISBN 9783939069478

Angus DC, Carlet J; 2002 Brussels Roundtable Participants. Surviving intensive care: a report from the 2002 Brussels Roundtable. Intensive Care Med 2003; 29(3): 368–377. PMID 12536269

Table – Two instruments to measure health-related quality of life

Short Form 36	EuroQoL
Physical functioning	Mobility
Role-physical	Self-care
Bodily pain	Usual activities
General health	Pain/discomfort
Vitality	Anxiety/depression
Social functioning	Overall health scoring (0-100)
Role-emotional	
Mental health	

ICU staff orientated outcomes

Staff orientated outcomes are diverse. Both short- and long-term patient mortality are naturally of importance to staff but so also are factors such as job satisfaction, burnout, communication and decisional latitude. These factors may affect risk-adjusted patient mortality. The impact of the organisation of the ICU on outcome has been extensively discussed and reported in the ESICM books

published in 2009 and 2010 which have been dedicated to the safety and quality aspects of intensive care as well as the principles of organisation and management of the ICU.



Hinds CJ, Watson JD. Intensive Care: A Concise Textbook. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. pp. 10–12; The influence of organization and resources on outcomes; Work-related stress

Reis Miranda D, Ryan DW, Schaufeli WB, Fidler V, editors. Organization and management of Intensive Care: a prospective study in 12 European countries. Berlin, Heidelberg: Springer-Verlag; 1997

Chiche JD, Moreno R, Putensen C, Rhodes A, editors. Patient Safety and Quality of Care in Intensive Care Medicine. Berlin: MVV Medizinisch Wissenschaftliche Verlagsgesellschaft; 2009. ISBN 978-3-941468-11-5

Flaatten H, Moreno R, Putensen C, Rhodes A, editors. Organization and management of intensive care. Berlin: MVV Medizinisch Wissenschaftliche Verlagsgesellschaft; 2010. ISBN 978-3-941468-27-6

Short-term mortality

The effects of disease and the consequences of medical interventions on mortality are important. New therapeutic options as well as groups of multifaceted interventions (called 'bundles of care' by the Institute for Healthcare Improvement) are frequently introduced with the expectation of reduced mortality. As for any other clinical intervention, their implementation requires proof that they are likely to have a beneficial effect on outcome.

Relatives view dying in or shortly after ICU as a failure of treatment

Taking the Surviving Sepsis Campaign as an example, an apparent decrease in hospital mortality of severe sepsis patients has been demonstrated after the implementation of the resuscitation bundle, both in ICU patients and in patients in the emergency department.



Ferrer R, Artigas A, Levy MM, Blanco J, González-Díaz G, Garnacho-Montero J, et al; Edusepsis Study Group. Improvement in process of care and outcome after a multicenter severe sepsis educational program in Spain. JAMA 2008; 299(19): 2294–2303. PMID 18492971

Levy MM, Dellinger RP, Townsend SR, Linde-Zwirble WT, Marshall JC, Bion J, et al. The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis. Intensive Care Med 2010; 36(2): 222–231. Review. PMID 20069275. Full text (pdf)

Q. Consider examples of categories of septic patients in which medicine is advancing so quickly that the indications for critical care admission require regular update and review?

A. Two examples are onco-haematological malignancies and some infectious diseases such as AIDS related infections.



Hinds CJ, Watson JD. Intensive Care: A Concise Textbook. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. pp. 350–351 and 354–356; Prognosis.

Lamia B, Hellot MF, Girault C, Tamion F, Dachraoui F, Lenain P, et al. Changes in severity and organ failure scores as prognostic factors in onco-hematological malignancy patients admitted to the ICU. Intensive Care Med 2006; 32(10): 1560–1568. PMID 16896863. Full text (pdf)

Vandijck DM, Benoit DD, Depuydt PO, Offner FC, Blot SI, Van Tilborgh AK, et al. Impact of recent intravenous chemotherapy on outcome in severe sepsis and septic shock patients with hematological malignancies Intensive Care Med 2008; 34(5): 847–855. PMID 18214437. Full text (pdf)

Lecuyer L, Chevret S, Thiery G, Darmon M, Schlemmer B, Azoulay E. The ICU trial: a new admission policy for cancer patients requiring mechanical ventilation. Crit Care Med 2007; 35(3): 808–814. PMID 17235261

Q. Apart from the primary pathology, what other clinical features or patient attributes determine ICU or hospital mortality?

A. Mortality attributable to critical illness is primarily determined by:

- Presence and degree of physiological dysfunction
- Age
- Pre-existing co-morbidity or chronic ill health
- Time and responsiveness to treatment
- Performance of the ICU to which the patient is admitted.

Long-term mortality

Long-term mortality is also of interest to intensive care staff. In the face of increasing ICU demand and the relative paucity of resources, ICU physicians have a responsibility to carefully triage patients in order to direct resources towards patients with the best chance of long-term survival. The decision whether or not to admit a patient depends upon many factors such as, the admission diagnosis, the severity of acute disease, age and previous health and health-related quality of life and importantly on the wishes of the patient. Moreover, the availability of an ICU bed may facilitate the admission of patients who are not so seriously ill, such as those requiring only monitoring. Accordingly, the most used criterion for triage is based on the risk–benefit ratio that the patient can expect from the ICU admission and stay.

Triage of patients for ICU admission requires realistic estimates of both short- and long-term mortality

You can obtain more examples and details of the long-term survival rates for different groups of critically ill patients in the references below.



- Roques S, Parrot A, Lavole A, Ancel PY, Gounant V, Djibre M, et al. Six-month prognosis of patients with lung cancer admitted to the intensive care unit. *Intensive Care Med* 2009; 35(12): 2044–2050. PMID 19768453. Full text (pdf)
- Broessner G, Helbok R, Lackner P, Mitterberger M, Beer R, Engelhardt K, et al. Survival and long-term functional outcome in 1,155 consecutive neurocritical care patients. *Crit Care Med* 2007; 35(9): 2025–2030. PMID 17855816
- Boumendil A, Somme D, Garrouste-Orgeas M, Guidet B. Should elderly patients be admitted to the intensive care unit? *Intensive Care Med* 2007; 33(7): 1252–1262. Review. PMID 17404703. Full text (pdf)
- Niskanen M, Kari A, Halonen P. Five-year survival after intensive care – comparison of 12,180 patients with the general population. Finnish ICU Study Group. *Crit Care Med* 1996; 24(12): 1962–1967. PMID 8968262

THINK about the short- and long-term mortality when considering triage of critically ill patients for admission to ICU. Both are important and should be balanced against the risks inherent to the ICU stay.



However, beware that prediction of outcome for individual patients is very difficult (all existing models are of a probabilistic nature) and may be unreliable.

Society oriented outcomes

Health budgets are not limitless and inevitably difficult decisions need to be made by those entrusted with the funding as to where to spend the budget and how much to spend. Health systems sometimes receive help from national bodies such as NICE (National Institute for Clinical Excellence) who weigh up the effectiveness of a treatment against cost; NICE generally is supportive of a treatment that costs <€40000 or £30000.

Health managers, economists and politicians

Health managers, economists and politicians are not as greatly influenced by the individual patient's prognosis or clinical problems as the healthcare professional working in the ICU. Their task is focused on distributive justice, in order to maximise the good for the whole of society. They are required to make rational decisions about healthcare resourcing based on accurate and reliable data. They must resolve conflicts concerning areas of healthcare, not only within the hospital but also in trying to establish the appropriate balance between primary and secondary healthcare. Their decision-making is best based upon rigorous economic evaluation using techniques of:

- Cost-effectiveness
- Cost-benefit

The perspective of this group may be different from the clinician involved in the care of an individual patient

- Cost-utility analyses
- Cost-minimisation

Countries around the world differ in their approach to distribution of resources for acute patient care. Nevertheless, clinicians and especially ICU directors need to be aware of the costs generated by the ICU to better inform allocation of resources.



Wunsch H, Angus DC, Harrison DA, Collange O, Fowler R, Hoste EA, et al. Variation in critical care services across North America and Western Europe. *Crit Care Med* 2008; 36(10): 2787–2793. e1–9 Review. PMID 18766102

See also the PACT module on Quality assurance and cost-effectiveness for further information.

The data required for such outcome measures will probably be of little interest to the critically ill patient presenting to ICU. However, for the professionals involved in intensive care, one of the pressing tasks is to produce accurate and reliable information to allow rational decision-making. Communication with health managers requires careful use of words, as their understanding of critical care processes is unlikely to be detailed. In this respect, the combination of clinical outcomes with economic outcomes is very important to understand the critical points of the ICU. Also, the Risk Profile of the ICU can give useful information about potential areas for improvement.



Rothen HU, Stricker K, Einfalt J, Bauer P, Metnitz PG, Moreno RP, et al. Variability in outcome and resource use in intensive care units. *Intensive Care Med* 2007; 33(8): 1329–1336. Epub 2007 Jun 1. PMID 17541552. Full text (pdf)

Nathanson BH, Higgins TL, Teres D, Copes WS, Kramer A, Stark M. A revised method to assess intensive care unit clinical performance and resource utilization. *Crit Care Med* 2007; 35(8): 1853–1862. PMID 17568328

Moreno RP, Hochrieser H, Metnitz B, Bauer P, Metnitz PG. Characterizing the risk profiles of intensive care units. *Intensive Care Med* 2010; 36(7): 1207–1212. PMID 20306015. Full text (pdf)

Views of society as a whole

The views of society as a whole may be expressed by politicians or patient societies or groups such as those offering support for patients suffering from a particular disease or groups campaigning for more resources to be directed at small groups of patients with special needs. The staff of ICU may become involved in discussions with such organisations e.g. organ transplant groups

2/ HOW TO MEASURE OUTCOME

The measures used to quantify the outcome of interest must be appropriate to the clinical perspective chosen. All outcome measures have limitations and it is important that clinicians are aware of such limitations so that results from outcome studies and data concerning outcome are properly interpreted. The most important clinical outcome measures are:

The choice of the outcome to be measured, as well as the choice of the instrument to do so is difficult

The patient's perspective is paramount. The general aim of healthcare staff and society is to allow patients to recover to their pre-critical illness health status. The specific aim of healthcare staff is to improve their patient's health status and outcomes.

- **Survival.** Long-term survival is probably the most important outcome measure for the patient and family. Moreover, the ICU staff is also interested in the short-term (hospital) mortality, because it is closer to their actions and is viewed as a more accurate reflection of the ICU performance.
- **Functional outcome.** Patients are interested in returning to an independent existence or at least to their previous level of activity and want to know what their physical and mental capabilities will be after they recover. Families and healthcare administrators as well as economists need to know the possible burden of care after hospital discharge.
- **Quality of life.** The patient's sense of well-being and satisfaction are uniquely personal perceptions. They can be measured with special instruments allowing comparison between health-related quality of life before and after intensive care but the assumptions of staff and family are not helpful in this respect.

THINK about issues other than survival when considering the care of patients in ICU and reflect on the perspectives of others.

Survival

Survival is most frequently reported at ICU or hospital discharge and then at various calendar periods up to many years following ICU admission. There are shortcomings with each of these reported survival periods.

ICU survival

ICU mortality may provide a global indication of ICU performance but is difficult to interpret in isolation because it is affected by many patient-related factors such as case-mix, severity of illness, co-morbidity and age. Moreover, other factors can influence ICU mortality; for example, an admission policy which is too strict (denying admission to patients with lower chances to recover) or a discharge policy whereby no

Varying admission and discharge criteria will inevitably affect ICU survival rates

terminal care is given in ICU with hopelessly ill patients being discharged to the general ward. Such practices will reduce the ICU mortality rate. Similarly, in some countries, ICU patients for whom death is expected to occur soon may be discharged home and managed there until death in accordance with local religious beliefs. In other cases, these patients will be offered the option to die in the ICU.



Try to establish in your own institution where the patients admitted to ICU come from (emergency department, operation room, ward or transfer from another hospital), and the reasons for ICU admission. Ascertain the mortality of these groups of patients both in ICU and in hospital after discharge from ICU. It will soon be apparent how very different these two outcome measures can be.

Hospital survival

Hospital mortality is the more frequently quoted outcome measure following critical illness. Hospital mortality, however, can be biased if the number of patients transferred to other acute hospitals is high, due to the lack of routine registration of vital status after transfer. For this reason, all developers of general outcome prediction models ask the user to follow the patient when they are transferred to another acute hospital until the final discharge to home or to a long-term facility. Moreover, hospital mortality is influenced by care beyond the ICU physician's control and reflects the institution's overall performance. For instance, an ICU in a hospital where emergency department patients are transferred to ICU after a long delay will have increased hospital length of stay and higher hospital mortality. Moreover, the in-hospital mortality rate following intensive care (called in some papers 'occult mortality') is surprisingly high. Figures of 20–30% have been reported, the majority of deaths occurring amongst the elderly and in patients with a high degree of residual organ dysfunction/failure or requiring extensive nursing care at ICU discharge.

Hospital mortality reflects institutional performance. Every effort should be made to optimise liaison between ICU clinicians and personnel on the ward to which the patient is to be transferred

You will find further information about mortality rates and ICU readmission in:



Chalfin DB, Trzeciak S, Likourezos A, Baumann BM, Dellinger RP; DELAY-ED study group. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. *Crit Care Med* 2007; 35(6): 1477–1483. PMID 17440421

de Rooij SE, Govers A, Korevaar JC, Abu-Hanna A, Levi M, de Jonge E. Short-term and long-term mortality in very elderly patients admitted to an intensive care unit. *Intensive Care Med* 2006; 32(7): 1039–1044. PMID 16791666. Full text (pdf)

Metnitz PG, Fieux F, Jordan B, Lang T, Moreno R, Le Gall JR. Critically ill patients readmitted to intensive care units – lessons to learn? *Intensive Care Med* 2003; 29(2): 241–8. PMID 12594586. Full text (pdf)

Fernandez R, Serrano JM, Umaran I, Abizanda R, Carrillo A, Lopez-Pueyo MJ, et al; Sabadell Score Study Group. Ward mortality after ICU discharge: a multicenter validation of the Sabadell score. *Intensive Care Med* 2010; 36(7): 1196–1201. PMID 20221748. Full text (pdf)



The differences in ICU case-mix and the standard of care outside the ICU (pre-admission or after discharge) make comparisons between the performance of different ICUs difficult.

Moreover, hospital mortality may be affected by the availability of discharge facilities such as intermediate care or step-down units which can provide a level of care which is higher than standard ward care to patients with residual organ dysfunction.

Whatever the destination, a transfer report prepared at ICU discharge constitutes the primary source of information for the receiving ward.

Post-ICU mortality may be related to a high level of nursing dependency following discharge. Nevertheless, it has been shown that higher dependency reflects unresolved organ dysfunction that the general wards may not be able to manage. More details can be found in Moreno et al while Smith et al illustrates the dangers of precipitate discharge from the ICU.



Moreno R, Miranda DR, Matos R, Fevereiro T. Mortality after discharge from intensive care: the impact of organ system failure and nursing workload use at discharge. *Intensive Care Med* 2001; 27(6): 999–1004. PMID 11497159. Full text (pdf)

Smith L, Orts CM, O'Neil I, Batchelor AM, Gascoigne AD, Baudouin SV. TISS and mortality after discharge from intensive care. *Intensive Care Med* 1999; 25(10): 1061–1065. PMID 10551959. Full text (pdf)



Building on the information obtained from the previous activity, compare patient characteristics and hospital mortality in two neighbouring ICUs with which you are familiar and gain an appreciation of the complexity of this task.

Long-term survival

Long-term survival is dependent upon a combination of the degradation of physiological reserve due to critical illness, the natural progression of the underlying pathology, and the gradual reduction of functional resources associated with increasing age. With appropriate follow-up, it is possible to draw a survival curve for a group of critically ill patients. However, interpretation of the patients' survival requires

Long-term follow-up of survivors may range from 30 or 90 days to decades

comparison with an appropriate control group. Two new problems then need careful consideration.

Which individuals should make up the control group? The ICU patient population is not a representative sample of the general population because ICU patients have generally been shown to have worse health conditions than those of the adjusted general population. Comparing their survival with that of an age and sex matched normal population may not be appropriate. See the figure 'Cumulative survival of a group of patients with a follow-up of 12 years' in the following reference.



Flaatten H, Kvåle R. Survival and quality of life 12 years after ICU. A comparison with the general Norwegian population. *Intensive Care Med* 2001; 27 (6): 1005–1011. PMID 11497132. Full text (pdf)

How long should patients be followed? ICU survivors should be followed until the gradient of their survival curve parallels that of the matched control group (if this phenomenon ever happens). Unfortunately, the length of time taken for the slopes of the survival curves to become parallel depends to a large extent on the population studied. It will take the longest time for patients with illnesses requiring long-term rehabilitation (for instance head trauma or stroke), and the shortest time for elective surgical patients, many having surgery for curable cancer. It has been reported that the survival curves for critically ill patients may not match those of a comparable population after a follow-up of as long as 15 years, but this phenomenon may depend on the start time of the follow-up, whether patients who died in hospital were included or excluded, and on the proportion of patients admitted with specific illness or conditions e.g. after cardiac surgery.



Niskanen M, Kari A, Halonen P. Five-year survival after intensive care – comparison of 12,180 patients with the general population. Finnish ICU Study Group. *Crit Care Med* 1996; 24(12): 1962–1967. PMID 8968262

Wright JC, Plenderleith L, Ridley SA. Long-term survival following intensive care: subgroup analysis and comparison with the general population. *Anaesthesia* 2003; 58(7): 637–642. PMID 12790812

Williams TA, Dobb GJ, Finn JC, Knuiman MW, Geelhoed E, Lee KY, et al. Determinants of long-term survival after intensive care. *Crit Care Med* 2008; 36(5): 1523–1530. PMID 18434893

Q. How would you estimate a patient's chances of long-term survival?

A. There are no validated systems for accurately predicting individual long-term survival. Moreover, the several reports of long-term survival in certain groups of critically ill patients likely reflect the continuous improvements in medical care and their probable influence on both short and long-term survival.

THINK: By trying to separate the prognosis of the pathological process that precipitated hospital admission from the effects of the acute physiological disturbance, try to work out how much of the acute physiological disturbance is actually reversible.

Beyond survival: health status and quality of life

The number of studies devoted to these topics has increased enormously in the last decade and we now have information about some of the consequences, both physical and mental, of life threatening critical illness. However, the matter is complex because health status and quality of life at least in part overlap. The following definitions will clarify the relationship between health and quality of life:

- **Impairment** concerns physiological aspects assessed objectively, for instance the FEV1 of a patient with chronic obstructive pulmonary disease.
- **Disability** refers to the symptoms associated with impairment such as dyspnoea for the patient with chronic obstructive pulmonary disease and a low FEV1, or the inability to manage one or two flights of stairs due to muscular atrophy.
- **Health-related quality of life** represents the effects of disability on individual satisfaction with life, for example not being able to garden or to play with the grandchild due to dyspnoea.

Link to **ESICM Flash Conferences**: Maurizia Capuzzo, Ferrara, Italy. Health related quality of life 90 days after ICU admission: A SAPS 3 substudy. Barcelona 2006

Els Bral, Leuven, Belgium. Long term health-related quality of life after cardiac surgery. Barcelona 2006



Curtis JR. Measuring health status after critical illness: Where are we and where do we go from here? In: Angus DC and Carlet J, editors. Surviving Intensive Care. Update in intensive care and emergency medicine. Berlin, Heidelberg, New York: Springer; 2003. pp. 181–196 ISBN 3540441492

Measurement of health status requires special tools

- Assessment of impairment needs measurement tools that have been designed for use in other areas of medicine or in different populations of patients.
- The results obtained will depend upon case-mix, patient selection and the confounding influences of socio-economic changes that affect the whole population.
- Cultural, social and religious factors can all play an important role in this perception.

Quality of life may be more significantly influenced by factors other than the direct effects of critical illness

- The rate of recovery of ICU survivors has been assessed in longitudinal studies comparing follow-up over time with baseline health-related quality of life. According to some of these studies, physical health is significantly reduced at one month and improves at nine months, but the variables related to ICU stay e.g. severity of illness generally do not influence health-related quality of life at six months. Health-related quality of life reaches pre-morbid values about 6–12 months after ICU admission but it may take longer in brain injured patients. Remember that major life events such as marriage, divorce, unemployment or inability to be with loved-ones may make the interpretation of changes in quality of life, related to intensive care, much more difficult.



- Graf J, Wagner J, Graf C, Koch KC, Janssens U. Five-year survival, quality of life, and individual costs of 303 consecutive medical intensive care patients – a cost-utility analysis. *Crit Care Med* 2005; 33(3): 547–555. PMID 15753746
- Graf J, Koch M, Dujardin R, Kersten A, Janssens U. Health-related quality of life before, 1 month after, and 9 months after intensive care in medical cardiovascular and pulmonary patients. *Crit Care Med* 2003; 31(8): 2163–2169. PMID 12973175
- Granja C, Teixeira-Pinto A, Costa-Pereira A. Quality of life after intensive care – evaluation with EQ-5D questionnaire. *Intensive Care Med* 2002; 28(7): 898–907. Epub 2002 Jun 15. PMID 12122528. Full text (pdf)
- Cuthbertson BH, Roughton S, Jenkinson D, MacLennan G, Vale L. Quality of life in the five years after intensive care: a cohort study. *Crit Care* 2010; 14(1): R6 Epub 2010 Jan 20. PMID 20089197

The pre-morbid levels of health-related quality of life of intensive care patients are generally lower than that of reference populations

- Any changes at follow-up can be measured only by comparison with baseline as ICU patients do not have the same distribution of pre-morbid physical and psychosocial attributes as the general population. The best approach is to assess pre-ICU health-related quality of life by direct patient interview. This is reliable particularly when the time frame involved does not exceed three months. Otherwise, the comparison between follow-up and baseline health-related quality of life has to be obtained from close relatives or partners who care for the patients, considering that surrogate answers are more reliable in the physical rather than the psychosocial domains.



- Scales DC, Tansey CM, Matte A, Herridge MS. Difference in reported pre-morbid health-related quality of life between ARDS survivors and their substitute decision makers. *Intensive Care Med* 2006; 32(11): 1826–1831. PMID 16957904. Full text (pdf)

Capuzzo M, Grasselli C, Carrer S, Gritti G, Alvisi R. Quality of life before intensive care admission: agreement between patient and relative assessment. *Intensive Care Med* 2000; 26(9): 1288–1295. PMID 11089755. Full text (pdf)

Diaz-Prieto A, Gorris MT, Badia X, Torrado H, Farrero E, Amador J, et al. Proxy-perceived prior health status and hospital outcome among the critically ill: is there any relationship? *Intensive Care Med* 1998; 24(7): 691–698. PMID 9722039. Full text (pdf)



Medical staff should always consider a patient's health-related quality of life. However, patients' perceptions are so uniquely personal that it is almost impossible for a third person to reliably describe their quality of life.

Health status

A number of measures have been used following intensive care. These may be broadly divided into five categories: physical impairment, functional status, mental function, neurological function and recovery.

Physical impairment will be common after severe trauma but the measurement of these parameters may not be appropriate for other causes of critical illness. Because respiratory failure is frequent in ICU patients and mechanical ventilation is a key therapeutic modality on ICU, measures of physical impairment have concentrated on the respiratory system and have included spirometry, diffusing capacity and bronchoscopy.

Functional status measures concern disability and may be broadly classified as either disease specific or general measures. The New York Heart Association classification, which grades physical activities, has been designed, validated and used in patients suffering with congestive cardiac failure. Generic measures of functional status include the activities of daily living: bathing, dressing, toileting, moving, continence and feeding. Most outcome measures examining functional status were developed for use in other fields and subsequently applied in ICU patients. For example, Katz's activities of daily living were originally developed from results obtained from elderly patients with femoral neck fracture. The application of the activities of daily living as a primary outcome measure to young critical care survivors may therefore be of limited value. For details, see Katz's activities of daily living in the reference below.



Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; 185: 914–919. No abstract available. PMID 14044222

Mental function measures may be divided into generic measures including the Hospital Anxiety and Depression Scale (HADS) and the Impact of Events Scale (IES). HADS was developed in non-psychiatric patients to measure mood disorders and depression; it contains 14 items asking about depression and anxiety. Disease specific mental function tests include the IES especially in the revised version developed to parallel the DSM-IV criteria for post-traumatic

stress disorder (PTSD). More recently the instrument developed by Stoll et al. has been modified and validated by Twigg et al. as a screening instrument for PTSD in ICU patients. For more information see the following references.



Weiss DS, Marmar CR. The impact of event scale – revised. In: Wilson JP, Keane TM, editors. *Assessing psychological trauma and PTSD*. New York, NY: The Guilford Press; 1997. pp. 399–411

Stoll C, Kapfhammer HP, Rothenhäusler HB, Haller M, Briegel J, Schmidt M, et al. Sensitivity and specificity of a screening test to document traumatic experiences and to diagnose post-traumatic stress disorder in ARDS patients after intensive care treatment. *Intensive Care Med* 1999; 25(7): 697–704. PMID 10470573. Full text (pdf)

Twigg E, Humphris G, Jones C, Bramwell R, Griffiths RD. Use of a screening questionnaire for post-traumatic stress disorder (PTSD) on a sample of UK ICU patients. *Acta Anaesthesiol Scand* 2008; 52(2): 202–208. PMID 18005373

Neurological function measurements assess cognitive function; examples include Trail-making Tests A and B, Wisconsin Card Sorting Test and Benton's test for visual retention. The tests are generally complex and used in a battery of tests assessing cognitive impairment; because the tests are complex, they usually require face-to-face interviews. They measure attention, perception, cognitive flexibility, information processing and visual memory. These tests have been used for patients with head injury or cardiac arrest, and also in paediatrics. Details of the tests and a review of the topic may be obtained from the references below.



Reitan RM. *Trail-making manual for administration, scoring and interpretation*. Indiana University Medical Center, Indianapolis, 1958

Heaton RK. *Wisconsin Card Sorting Test manual*. Psychological Assessment Resources, Odessa, 1981

Benton AL. *Revised Visual Retention Test: Clinical and Experimental Application*. Psychological Corporation, New York, 1984

Hopkins RO, Brett S. Chronic neurocognitive effects of critical illness. *Curr Opin Crit Care* 2005; 11(4): 369–375. Review. PMID 16015118

Recovery can be measured by multi-item scales such as the Glasgow Outcome Scale, which has five levels (good, moderate disability, severe disability, vegetative state, and death). The scale was originally intended to assess recovery six months after head injury. Simple single item scales such as returning to work or independent living have been used but their simplicity limits their usefulness as there are many grades of employment and varying levels of support available to maintain independence. For further information about the Glasgow Outcome Scale and its use in the assessment of head injury, read the PACT module on Traumatic brain injury

Quality of life

Specific instruments (questionnaires) to measure health-related quality of life have been widely used outside critical care, frequently in chronic diseases such as hypertension or rheumatoid arthritis. In the critical care setting, generic instruments are preferred because they allow us to study patients with different illnesses. More than one hundred instruments to measure health-related quality of life have been used, but only two of them have been recommended for use in critical care: the Short Form 36 and the EuroQol, that have been validated for use in almost all European countries and in different populations including patients after critical illness (see the table in Task 1).

Health-related quality of life can be described by a number of domains

When assessing critical care patients, investigators should report the psychometric properties of the instrument they use: **validity** (ability of the instrument to measure what it is intended to measure), **reliability** (individual items in a domain measure the same underlying concept and stability of the assessment over a short period of time) and **responsiveness** (sensitivity to detect clinically meaningful changes).

Validity and reliability are important properties of any quality of life questionnaire

Investigators should be encouraged to select the tools that are specifically designed for the area of outcome in which they are interested: physical impairment, functional status, mental or neuropsychological function, recovery or health-related quality of life. The importance of selecting the correct measurement tool is discussed in the references below.



Rowan KM, Jenkinson C, Black NA. Health-related quality of life. In: Angus DC and Carlet J, editors. *Surviving Intensive Care. Update in intensive care and emergency medicine*. Berlin, Heidelberg, New York: Springer; 2003. pp. 36–50 ISBN 3540441492

Curtis JR. Measuring health status after critical illness: Where are we and where do we go from here? In: Angus DC and Carlet J, editors. *Surviving Intensive Care. Update in intensive care and emergency medicine*. Berlin, Heidelberg, New York: Springer; 2003. pp. 181–196 ISBN 3540441492

Hayes JA, Black NA, Jenkinson C, Young JD, Rowan KM, Daly K, Ridley S. Outcome measures for adult critical care: a systematic review. *Health Technol Assess* 2000; 4(24): 1–111. Review. PMID 11074394

3/ HOW TO PREDICT OUTCOME

Severity evaluation and outcome prediction

The evaluation of severity of illness in the critically ill patient is made through the use of general severity scores and general outcome prognostic models. Severity scores are instruments that aim at stratifying patients based on their illness severity, having assigned additional points to each patient as his/her severity of illness increases.

Severity of illness scoring systems can be used for patient stratification and outcome prediction for groups of patients

General outcome prognostic models, apart from their ability to stratify patients according to their severity of illness, aim at predicting a certain outcome, usually the vital status at hospital discharge based on a given set of prognostic variables and a particular modelling equation. Other outcomes, both short-term and long-term can also be considered, but most are more prone to bias and manipulation or are of little interest for the patients, their families and the healthcare providers.

Link to **ESICM Flash Conferences**: Edwin Koterba. A prospective, multicentric comparison between SAPS 3 and APACHE II scores. Barcelona 2006

Didier Ledoux, Brussels, Belgium. Assessment of survival and quality of life in octogenarian 1-year after cardiac surgery. Barcelona 2006

Isabelle Runge, Orleans, France. Can 'at risk of deterioration' patients be identified in the emergency department (ED)? Barcelona 2006

Didier Ledoux, Brussels, Belgium. Does SAPS 3 score perform better than the old APACHE II and SAPS 2 scores? Barcelona 2006

Christian Bosert, Feldkirch, Austria. Enhancing the performance of the score system SAPS 2 by an analytic tool. Barcelona 2006

Severity scores can be classified according to their aim, so they can look like:

- A picture of the patient's clinical status at ICU admission. This picture allows measurement of illness severity on the basis of assigned points and may be used for comparison of groups of patients e.g. those enrolled in clinical trials. Such scores provide a numerical estimate of the probability of hospital mortality for a group of similar patients. They may also be used to assess an individual ICU's performance over time or for comparison with other ICUs. These kinds of scores are presented in the section dealing with outcome prediction models.
- A movie describing the evolution of the patient's clinical status during their ICU stay. Such scores follow the clinical evolution of an individual patient and compares groups of patients enrolled in

clinical trials some days after ICU admission. These kinds of scores are described in the section dedicated to scores of evolution.

Like all methods of clinical measurement, severity of illness scoring systems are less than optimal in clinical practice and are not well suited to individual risk prediction. We should keep in mind that they have a false positive and false negative error rate, so that individual predictions may be wrong. Given their probabilistic nature, they cannot be used to predict outcome for a particular individual.



Hinds CJ, Watson JD. *Intensive Care: A Concise Textbook*. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. pp. 27 and 30; Predicting death in individual patients

In fact, the application of different models to the same patient can result in very different outcome predictions, as demonstrated in the figure 'Relationship between the APACHE (Acute Physiology and Chronic Health Evaluation) II predicted risk of death and the SAPS II predicted risk of death in a Portuguese multicentre study referenced below. Similar results have been obtained comparing the hospital mortality predicted by SAPS II and SAPS 3 Admission Score in data used in a more recent study – see figure in the second reference. Additional details on the statistics of scoring systems can be found in the other references.

NOTE

Current outcome prediction instruments are not valid in individual patients.



Moreno R, Morais P. Outcome prediction in intensive care: results of a prospective, multicentre, Portuguese study. *Intensive Care Med* 1997; 23(2): 177–186. PMID 9069003. Full text (pdf)

Capuzzo M, Scaramuzza A, Vaccarini B, Gilli G, Zannoli S, Farabegoli L, et al. Validation of SAPS 3 Admission Score and comparison with SAPS II. *Acta Anaesthesiol Scand* 2009; 53(5): 589–594. PMID 19419351

Cook DA. Methods to assess performance of models estimating risk of death in intensive care patients: a review. *Anaesth Intensive Care* 2006; 34(2): 164–175. Review. PMID 16617636

Peek N, Arts DG, Bosman RJ, van der Voort PH, de Keizer NF. External validation of prognostic models for critically ill patients required substantial sample sizes. *J Clin Epidemiol* 2007; 60(5): 491–501. PMID 17419960

Outcome prediction models

Outcome prediction in intensive care is made using general outcome prediction models. Such models aim to predict survival or death at hospital discharge (or less frequently at 28 days after ICU admission), based on a given set of variables evaluated at

Outcome prediction is based on general outcome prediction models

ICU admission or within 24 hours of ICU admission. All models were developed in large multicentre databases, which have been selected as being representative of the general critically ill patient population.

The rationale behind their construction is that derangement of homeostasis has an adverse effect on mortality and that the magnitude of change from normal for physiological and laboratory variables is proportional to their effects on outcome. Using logistic regression equations usually (nowadays almost always controlling for the ICU effect by using logistic regression with random effects), these instruments predict outcome according to the patients' particular past medical history, the acute clinical condition (defined by the values of the predictive variables) presuming the patients are being treated in a (theoretical) reference ICU.

The severity of illness/ risk stratification models most frequently used for outcome prediction in adult, general, critically ill patients can be classified in the following groups:

Several general outcome prediction models are available for use in the ICU

- The Acute Physiology and Chronic Health Evaluation (APACHE) system – last version available being APACHE IV. The most commonly used version worldwide is the second (APACHE II).
- The New Simplified Acute Physiology Score (SAPS) system – second (SAPS II) and third (SAPS 3 Admission Score) versions. Other variations of the method have been described e.g. the expanded SAPS II.
- The Mortality Probability Models system – second or third version (MPM II or III).
- The Intensive Care National Audit & Research Centre (ICNARC) model (specific to the population and the practices in the United Kingdom).

THINK: Relatively poor prognostic performance of all these systems became evident after their development. This may have been due to differences in the case-mix of the study population compared with that where the model was developed and/or medical progress and/or differences among ICUs in their quality of care.



Hinds CJ, Watson JD. Intensive Care: A Concise Textbook. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. pp. 21–27; Severity-scoring systems

Acute Physiology and Chronic Health Evaluation system

The Acute Physiology and Chronic Health Evaluation (APACHE) II is the most commonly used outcome prediction model; it was based on data collected from 1979 to 1982 in 13 hospitals in North America. A panel of experts chose the variables and their weights based on clinical judgement and on the published relationships between the degree of physiological derangement and outcome. The model uses the worst values

The APACHE II model uses information on physiological variables, age, surgical status and previous health status to assign a score to the patient

during the first 24 hours of critical illness for 12 physiological variables (weighted from 0 to 4 points according to the degree of change from normal values), age, and chronic health status. The APACHE II score is the sum of the scores attributed to the previous variables and varies between 0 and 71 points. A formula including the APACHE II score, and values assigned according to the main diagnostic category at ICU admission (chosen from a list of around 50 diagnoses) and the presence of emergency surgery, allows the user to calculate the probability of hospital death.

You can find a calculator for APACHE II at <http://www.sfar.org/scores2/apache22.html>

The latest version, APACHE IV, was developed between 2002 and 2003 from a sample of 110,558 patients admitted to ICUs of 45 North American hospitals and is quite complex, because there are 142 variables in the mortality equation, although most (115) are disease groups. A predictive equation provides the user with an estimate of the probability of in-hospital death. However, the most recent versions of APACHE, but not APACHE II, are a commercial product; the equations are not in the public domain and may be purchased from Cerner Corporation (www.cerner.com). This is the main reason why APACHE II, although old, is still more commonly used than the new versions. More details about the APACHE systems may be found in the online appendix and the following references:

The APACHE system needs information on diagnosis to estimate the probability of death



Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985; 13(10): 818–829. PMID 3928249

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): 1297–1310. PMID 16540951

Zimmerman JE, Kramer AA. Outcome prediction in critical care: the Acute Physiology and Chronic Health Evaluation models. *Curr Opin Crit Care* 2008; 14(5): 491–497. Review. PMID 18787439

Simplified Acute Physiology Score system

The Simplified Acute Physiology Score (SAPS) II was described in 1993 (see reference below) following a combined European and North American study. It was developed and validated in a large sample of patient data from 110 European and 27 North American hospitals. This model comprises 17 variables (12 physiological variables, age, type of admission (non-operative, emergency surgery and elective surgery), and three primary diagnoses (AIDS, metastatic cancer and haematological cancer). The SAPS II score varies between 0 and 163 points. All the physiological variables are registered as the worst values during

The SAPS II model was developed from a large multicentre database, mainly from Europe

the first 24 hours in the ICU. Calculation of the risk of death does not require selection of a primary admission diagnosis nor further information on chronic health status. See the electronic supplementary material and the reference of the original paper.



Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. *JAMA* 1993; 270(24): 2957–2963 Erratum in: *JAMA* 1994; 271(17): 1321. PMID 8254858

You can find a calculator for SAPS II at <http://www.sfar.org/scores/igs2.php>

The SAPS 3 admission model was described in 2005 on a database where the SAPS II showed poor performance (see references below). It was developed on 16,784 patients admitted to 303 ICUs worldwide. This model takes into account three groups of variables, named boxes. Box I (five variables) concerns patient characteristics and treatments before ICU admission; box II (five variables) on circumstances related to ICU admission and box III (ten physiological parameters) collected the hour before–after ICU admission. The SAPS 3 Admission Score model allows the probability of hospital mortality to be computed by an ad hoc logit, but users can also adopt a logit specific for a geographical area (customisation). The following references may be useful.

The SAPS 3 Admission Score was developed with the objective of evaluating not only individual patient outcome but also the effectiveness of ICU practices



Metnitz PG, Moreno RP, Almeida E, Jordan B, Bauer P, Campos RA, et al; SAPS 3 Investigators. SAPS 3 – From evaluation of the patient to evaluation of the intensive care unit. Part 1: Objectives, methods and cohort description. *Intensive Care Med* 2005; 31(10): 1336–1344. PMID 16132893. Full text (pdf)

Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al; SAPS 3 Investigators. 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): 1345–1355. Erratum in: *Intensive Care Med* 2006; 32(5): 796. PMID 16132892. Full text (pdf)

Capuzzo M, Moreno RP, Le Gall JR. Outcome prediction in critical care: the Simplified Acute Physiology Score models. *Curr Opin Crit Care* 2008; 14(5): 485–490. Review. PMID 18787438

A similar model targeting survival during the first 28 days after ICU admission was also described, based on the same database.

Moreno R, Metnitz P, Jordan B, Einfalt JPB. SAPS 3 28 days score: a prognostic model to estimate patient survival during the first 28 days in the ICU [Abstract]. *Intensive Care Med* 2006; 32: S203.

You can download a calculator for SAPS 3 Admission Score at <http://www.saps3.org>

Mortality Probability Models system

The Mortality Probability Models (MPM) II were described in 1993 to 1994 and were based on the same data as used for the development of SAPS II with additional data from six other ICUs in North America. In these models, the final result is expressed only as a probability of death and not as a score. The most recent MPM model (MPM 0-III) was published in 2009 (see reference).

The MPM II models were developed to be applied at admission, and after 24, 48 and 72 hours in the ICU



Higgins TL, Kramer AA, Nathanson BH, Copes W, Stark M, Teres D. Prospective validation of the intensive care unit admission Mortality Probability Model (MPM0-III). Crit Care Med 2009; 37(5):1619–1623. PMID 19325480

The MPM II models are defined according to the time frame considered:

- The MPM II admission model (MPM II0), computed within one hour of ICU admission. This model contains 15 variables. It is the only general model that is independent of treatment provided in the ICU and therefore can be used for patient stratification at the time of admission to the ICU.
- The MPM II 24 hour model (MPM II24), computed after 24 hours in the ICU. This model comprises 13 variables.
- The MPM II 48 hour model (MPM II48), computed after 48 hours in the ICU.
- The MPM II 72 hour model (MPM II72), computed after 72 hours in the ICU.

The MPM II48 and MPM II72 models use the same variables of the MPM II24 model, with different weights for risk of death calculation. Both are based on the worst values presented by the patient during the previous 24 hours.

The MPM 0-III (Higgins et al) is computed at ICU admission and contains 17 variables. The development study has shown that the previous model significantly over predicted mortality.

Further details on the MPM models may be found in the online appendix and the following references:



Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients. JAMA 1993; 270(20): 2478–2486. PMID 8230626

Lemeshow S, Klar J, Teres D, Avrunin JS, Gehlbach SH, Rapoport J, et al. Mortality probability models for patients in the intensive care unit for 48 or 72 hours: a prospective, multicenter study. *Crit Care Med* 1994; 22(9): 1351–1358 PMID 8062556

Higgins TL, Teres D, Copes WS, Nathanson BH, Stark M, Kramer AA. Assessing contemporary intensive care unit outcome: an updated Mortality Probability Admission Model (MPMO–III). *Crit Care Med* 2007; 35(3): 827–835. PMID 17255863

Higgins TL, Teres D, Nathanson B. Outcome prediction in critical care: the Mortality Probability Models. *Curr Opin Crit Care* 2008; 14(5): 498–505. Review. PMID 18787440

You can find a calculator for MPM scores at <http://www.sfar.org/article/315/scores>

Most ICUs use at least one outcome prediction model. The most used systems in Europe are the SAPS II and APACHE II models. All of these systems allow an assessment of performance.

Q. Give arguments as to why the use of performance measurement/a general outcome prediction model in the ICU might be mandatory.

A. Performance measurement:

- (1) Is an indicator of quality
- (2) Is the first step in quality improvement programmes
- (3) Used constructively, is a means of identifying the consequences of deficiencies in critical care delivery.



Determine the reasoning behind the choice of system and to what extent the system is suited to the spectrum of patients in your ICU.

Apart from the general outcome prediction models described above, other instruments are used in particular contexts. Examples include organ failure scores aiming at quantifying and describing organ failure and there are also a series of specific scores developed for use in particular situations – see ‘specific scores’ below.

Intensive Care National Audit & Research Centre model

The main purpose of the Intensive Care National Audit & Research Centre (ICNARC) model was to support the Case-Mix Program developed in the UK, for comparative audit of risk-adjusted outcomes from adult general critical care units in the UK. The setting for this development was 163 adult, general critical care units in England, Wales, and Northern Ireland, during the period 1995–2003. The risk prediction model considers a set of physiological variables combined with information relating to the critical care admission, namely age, diagnostic category, source of admission, and cardiopulmonary resuscitation before admission.



Seek a regional or national group working on the evaluation of ICU performance in your country. If you do not find any, think of joining a large international group.



Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K. A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model. *Crit Care Med* 2007; 35(4): 1091–1098. PMID 17334248

Harrison DA, Rowan KM. Outcome prediction in critical care: the ICNARC model. *Curr Opin Crit Care* 2008; 14(5): 506–512. Review. PMID 18787441

Scores of evolution

Scores related to clinical evolution were developed to quantify organ failures in critically ill patients, and are now used to measure organ failures or dysfunction in a patient over time. Among them are the Multiple Organ Dysfunction Score (MODS), the Logistic Organ Dysfunction (LOD) Score and the Sequential Organ Failure Assessment (SOFA) Score. The SOFA score was initially developed for sepsis and named Sepsis-related Organ Failure Assessment. Subsequently it was applied in many settings with success; it is now the most commonly used score and its name has been changed accordingly. These scores describe evolving dysfunction of individual organs, measured as a continuum ranging from normality to complete failure. What makes the SOFA score original is that it takes into account the level of organ dysfunction or failure, in the light of the effect of treatments such as mechanical ventilation and infusion of vasoactive agents.

The scores are designed to be applied serially (usually daily) throughout the ICU stay and aim at patient description rather than outcome prediction. This is the case for all the systems except the LOD Score, which incorporates an equation to predict hospital mortality based on the degree of organ dysfunction after 24 hours in the ICU. Also, for maximum SOFA, an equation has been described to predict patient ICU survival.

Organ failure scores aim in general at describing and following the evolution of patients and not at outcome prediction

The various organ dysfunction/failure scoring systems are described in more detail in the online appendix and in the following references:



Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995; 23(10): 1638–1652. Review. PMID 7587228

Le Gall JR, Klar J, Lemeshow S, Saulnier F, Alberti C, Artigas A, et al. The Logistic Organ Dysfunction system. A new way to assess organ

dysfunction in the intensive care unit. ICU Scoring Group. JAMA 1996; 276(10): 802–810. PMID 8769590

Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça 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): 707–710. No abstract available. PMID 8844239. Full text (pdf)

Moreno R, Vincent JL, Matos R, Mendonça A, Cantraine F, Thijs L, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. Working Group on Sepsis related Problems of the ESICM. Intensive Care Med 1999; 25(7): 686–696. PMID 10470572. Full text (pdf)

You can find a calculator for SOFA at <http://www.sfar.org/scores/sofa.php>

Specific scores

Although most instruments are designed to be applied to a general population of ICU patients, alternative scores, developed for particular pathologies or patient groups, are available. Examples include severity scores for children, trauma victims and patients with myocardial infarction, after cardiac surgery, and acute pancreatitis. Although intuitively attractive since they incorporate the best predictive variables in a particular context, these models have not been proven superior to the general outcome prediction models in ICU, except for trauma. In this case a combination of anatomical and physiological methodology can increase accuracy.

Special scores are applicable in particular conditions

NOTE

In paediatrics, specific scores must be used because of differences in the normal range of physiological variables. The following references show the Pediatric Risk of Mortality (PRISM) and an example of its application:



Pollack MM, Ruttimann UE, Getson PR. Pediatric risk of mortality (PRISM) score. Crit Care Med 1988; 16(11): 1110–1116. PMID 3048900

Cogo PE, Poole D, Codazzi D, Boniotti C, Capretta A, Langer M, et al. Outcome of children admitted to adult intensive care units in Italy between 2003 and 2007. Intensive Care Med 2010; 36(8): 1403–1409. PMID 20490781. Full text (pdf)

Harrison DA, D'Amico G, Singer M. The Pancreatitis Outcome Prediction (POP) Score: a new prognostic index for patients with severe acute pancreatitis. Crit Care Med 2007; 35(7): 1703–1708. PMID 17522578

Q. If you are treating a heterogeneous group of patients in your ICU, how might you deal with this case-mix?

A. The use of general severity scores will be appropriate if you deal with groups of patients with a variety of illnesses, but specific scoring systems are appropriate if your ICU population is specialised.

Link to **ESICM Flash Conference**: Jean-Roger Le Gall, Paris, France.
 Development of severity of illness systems: Historical perspective. Berlin 2007

Q. Name organ or system specific scoring systems that can be used on the general ICU.

A. Examples of such scores are the Injury Severity Score (ISS) and the (Revised) Trauma Score/Injury Severity Score (TRISS) (for trauma), the Ranson, Imrie or Pancreatitis Outcome Prediction (POP) scores (for acute pancreatitis), the Child–Pugh or Model of End stage Liver Disease (MELD), and the Glasgow Outcome Scale or the FOUR score (for neurotrauma). You should evaluate if the number of patients in specific groups is significant in your ICU. For instance, if your ICU is a specialised liver unit, you could use a specific score such as the MELD score. Similarly if the proportion of patients in a specific group is high, you should assess whether specific scores are likely to perform better than general outcome models.

Links to

- PACT module on Traumatic brain injury
- PACT module on Pancreatitis
- PACT module on Sepsis and MODS
- PACT module on Acute hepatic failure

References for some specific scoring systems:



Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method Trauma Score and the Injury Severity Score. *J Trauma* 1987; 27(4): 370–378. PMID 3106646

Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated Pediatric Risk of Mortality score. *Crit Care Med* 1996; 24(5): 743–752. PMID 8706448

Nashef SA, Roques F, Michel P, Gauducheau E, Limeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 1999; 16(1): 9–13. PMID 10456395

Wijdicks EF, Bamlet WR, Maramattom BV, Manno EM, McClelland RL. Validation of a new coma scale: The FOUR score. *Ann Neurol* 2005; 58(4): 585–593. PMID 16178024

Harrison DA, D’Amico G, Singer M. The Pancreatitis Outcome Prediction (POP) Score: a new prognostic index for patients with severe acute pancreatitis. *Crit Care Med* 2007; 35(7): 1703-1708. PMID 17522578

Use of outcome prediction models

Several important steps when using any of these models to predict outcome:

- Selection of the patient group suitable for the analysis
- Collection of the predictive variables and the outcome data
- Computation of the severity score
- Transformation of the severity score into a probability of death.

Several steps are involved in computing the probability of death in hospital

THINK: about the steps needed for computation of the probability of death in your patients. If you separate out and analyse all the steps involved, you will be able to optimise them and improve the performance of the model.

Patient selection

Although many are termed 'general', none of these outcome prediction models can be applied to all categories of patient. Specific patient groups were excluded in their development and should therefore not be included in subsequent analyses. There is variation among the scores with regard to specific exclusions, but for most, the following groups should be excluded:

None of the outcome prediction models are applicable to all patients in the ICU

- Readmissions (only evaluate first ICU admission), to avoid counting more than one hospital outcome for the same patient
- Paediatric patients (less than 16 or 18 years old), require purpose designed scores
- Patients with a very short length of stay in the ICU (less than 12 or 24 hours) for all scores except those of the MPM 0 and SAPS 3 admission models where just one hour in the ICU is sufficient
- Burn patients, patients following coronary artery by-pass surgery or those admitted to rule out myocardial infarction. The most recent SAPS 3 admission model and the APACHE III and IV models include coronary care and cardiac surgery patients.



If the proportion of a specific type of patient in your unit is too high, the mortality estimate given by the model will be based on a minority of patients similar to yours. Mortality prediction for your case-mix will therefore be inaccurate. It is also necessary to follow the definitions and apply rules strictly as described by the original authors.



Rowan K. The reliability of case mix measurements in intensive care. *Curr Opin Crit Care* 1996; 2: 209-13.



Don't forget that the patient transferred to another acute hospital is usually counted as discharged by the hospital administration. You should collect information about the vital status of this type of patient at their discharge from the final acute care hospital.



Kahn JM, Kramer AA, Rubenfeld GD. Transferring critically ill patients out of hospital improves the standardized mortality ratio: a simulation study. *Chest* 2007; 131(1): 68–75. PMID 17218558

Q. What proportion of the patients admitted to your ICU can be evaluated using a general outcome prediction model?

A. The proportion of patients that can be evaluated using general outcome models varies substantially from ICU to ICU. Each unit should empirically test the figures in their own setting, since if the number of excluded patients is too high comparability with other ICUs will be compromised. With newer models, this proportion is greater than 95% in general ICUs.

Collection of predictive variables and hospital outcome

Great care should be applied during data collection. An outcome prediction can only be reliably computed if the definitions and time frames used for data collection follow the original description. The errors most frequently seen include:

- Definitions of the variables
- Correct choice of a variable's worst value (between highest and lowest)
- Time frames for data collection
- Frequency of evaluation of the variables (especially important if you are using automated data collection software)
- Exclusion criteria
- Data handling before analysis (what does it mean?).

It is important to have a clear idea of why this data is being collected and how it will be used

THINK: Are you using the models according to their original descriptions? Are you using manual or automatic data collection of the physiological variables?



Strand K, Strand LI, Flaatten H. The interrater reliability of SAPS II and SAPS 3. *Intensive Care Med* 2010; 36(5): 850–853. Epub 2010 Feb 4. PMID 20130828. Full text (pdf)

Polderman KH, Jorna EM, Girbes AR. Inter-observer variability in APACHE II scoring: effect of strict guidelines and training. *Intensive Care Med* 2001; 27(8): 1365–1369. PMID 11511950. Full text (pdf)

Suistomaa M, Kari A, Ruokonen E, Takala J. Sampling rate causes bias in APACHE II and SAPS II scores. *Intensive Care Med* 2000; 26(12): 1773–1778. PMID 11271084. Full text (pdf)

Computation of the severity score

The next step is the computation of the severity score. Using the relevant original tables or a computer programme, points are assigned to each variable according to the degree of variation from normal values provided by the chosen score. The impact on mortality of the variations can be uni-directional (e.g. low values for the Glasgow Coma Score or for the PaO₂/FiO₂ ratio) or bi-directional (e.g. low and high values for leucocytes, heart rate or blood pressure).

An aggregated score is then computed by summing the values for all the variables. However, this step is not necessary with the MPM II and III models.

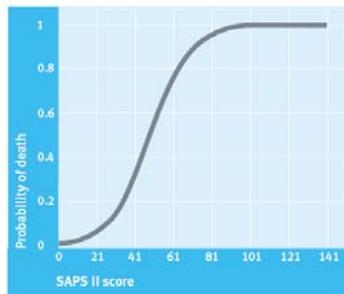
Transformation of the severity score to a probability of death

The last step is transforming the severity score (or directly the variables as in the case of the MPM systems) to a probability of death at hospital discharge. This transformation is made using a logistic regression equation for the SAPS or APACHE scores. The following figure relates to SAPS II, but the shape is the same for SAPS 3 Admission Score.

If you use the SAPS 3 Admission Score, you can adopt either the general equation, or the equation customised for your geographical area, or the customisation most suitable for your own ICU.

Link to **ESICM Flash Conferences**: Zepeda Monares, México, Mexico. SAPS 3 vs SOFA vs APACHE II in the evaluation and prognosis of critical care patients. Barcelona 2006

Juan Carlos Sotillo Diaz, Madrid, Spain. Serum transthyretin as predictor of death in critically ill patients. Barcelona 2006



The sigmoid shape of the logistic regression equation curve has important implications for scores in the middle range of severity evaluation

Legend: Relationship between the SAPS II score and the probability of death in hospital. Adapted from Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA. 1993; 270(24): 2957-2963. PMID 8254858

Q. Given the shape of the curve above, at the extremes of the score (very low or very high values), the associated changes in the probabilities of death are small. How would you interpret the situation for intermediate values?

A. For intermediate values even small differences in the score are associated with large changes in the probability of death.



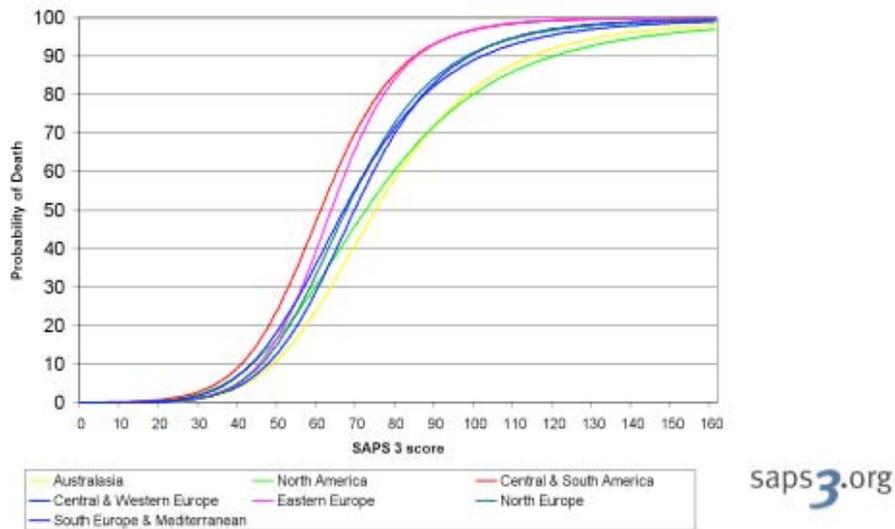
From the above figure, it is possible to estimate the probability of death of a patient with a SAPS II score of 30. However there are likely to be other important prognostic factors that are not captured in the SAPS II score. Consider what these might be and why population based prediction models are not yet appropriate for individual patients.

In figure 2 in the following reference, you will see the relationship between the SAPS 3 Admission Score and the respective probabilities of hospital mortality.

Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al; SAPS 3 Investigators. 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): 1345–1355. Erratum in: *Intensive Care Med* 2006; 32(5): 796. PMID 16132892. Full text (pdf)

Please note that the relationship between the score and the associated probability of mortality changes according to several factors including the geographical area.

THE SAPS 3 ADMISSION SCORE - AREAS



Q. Why is it important to validate the outcome prediction model at national and regional level?

A. The identification of the system that best applies in your country (or in ICUs with similar case-mix) is important, since it will probably provide more accurate outcome prediction in your setting.

Q. Why is it important that risk-adjusted mortality is compared between ICUs with similar characteristics?

A. Case-mix adjustment will be most reliable when the characteristics of the ICUs being compared are similar. Comparisons may be unreliable when specialised ICUs (e.g. neurosurgical, liver or cardiothoracic ICUs) are compared with general medical, surgical or combined ICUs.

The comparison of major case-mix characteristics of your patients with similar ICUs is important, since it can give you insight as to your performance compared to other ICUs.



Peek N, Arts DG, Bosman RJ, van der Voort PH, de Keizer NF. External validation of prognostic models for critically ill patients required substantial sample sizes. *J Clin Epidemiol* 2007; 60(5): 491–501. PMID 17419960

Capuzzo M, Scaramuzza A, Vaccarini B, Gilli G, Zannoli S, Farabegoli L, et al. Validation of SAPS 3 Admission Score and comparison with SAPS II. *Acta Anaesthesiol Scand* 2009; 53(5): 589–594. PMID 19419351

Evaluating the performance of a severity score

All statistical models need validation, that is evidence that they perform effectively and reproducibly (the probability predicted by the model equals the actual mortality). In order to check how the model used performs in predicting mortality in a sample of patients which is independent of the population sample in which it was developed, three fundamental features should always be evaluated – discrimination, calibration and uniformity-of-fit. This assessment should be performed at regional or national level, and should consist of the following measures:

All outcome prediction models need validation

- Discrimination.
- Calibration.
- Uniformity-of-fit.

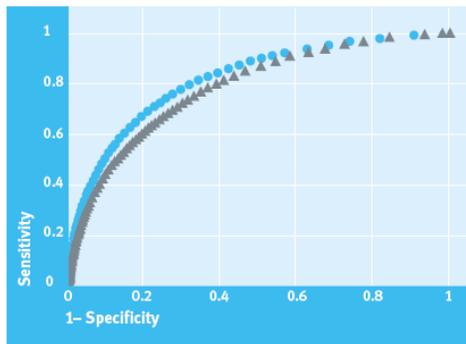
Discrimination

Discrimination is the ability of the model to distinguish survivors from non-survivors. It is usually quantified using the Receiver Operator Characteristic (ROC) curve (figure below) which plots the proportion of correctly classified non-survivors (sensitivity) against the proportion of correctly classified survivors (1 minus specificity) at different cut-off values of the score. The interpretation of discrimination is easy: a perfect model will have an area under the ROC curve of 1.0 and a model whose discriminative capability is no greater than chance has an area of 0.5. For most models, this value should be greater than 0.80.

Discrimination, the ability of the model to distinguish survivors from non-survivors, is usually evaluated using the area under the receiver operating characteristic curve



Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143(1): 29–36. PMID 7063747



The greater the area under the ROC curve, the better is the discrimination of the model

Receiver operating characteristic curve for SAPS II and MPM II in the EURICUS I study. The curves plot the sensitivity and 1 minus the specificity for all possible values of the two scores. Adapted from Moreno R, Miranda DR, Fidler V, Van Schilfgaarde R. Evaluation of two outcome prediction models on an independent database. Crit Care Med 1998; 26 (1): 50-61. PMID 9428543

Q. Determine the area under the curves in the above illustration. Does this indicate satisfactory discrimination?

A. The areas under the two curves are 0.785 and 0.822. In statistical terms, an area under a ROC curve of 0.9 indicates excellent discrimination but few diagnostic tests achieve this. Most well performing tests have areas under the ROC curve of around 0.8. A value of 0.5 means that the test performs no better than chance.

For further details about ROC curves see:



Faraggi D, Reiser B. Estimation of the area under the ROC curve. Stat Med 2002; 21(20): 3093–3106. PMID 12369084

Calibration

Calibration quantifies the degree of correspondence between the predicted probabilities of death and the observed mortality over classes of risk. It is usually evaluated by two statistical tests for goodness-of-fit proposed by Hosmer and Lemeshow (the C test and the H test) which divide the population into deciles of risk (C test refers to groups of equal numbers of patients, while H test refers to groups of patients at equal intervals of predicted mortality), and compare the expected with the actual number of survivors (and non-survivors) in each decile. However, there are other tests that can be used to assess calibration, such as the Cox test (see references). Please note that the Hosmer–Lemeshow test is influenced by a variety of factors, particularly sample size; other methods of assessing calibration (such as Cox-plots) should be used for analysis of large databases.

Calibration is the degree of correspondence between the probabilities assigned by the model and the observed mortality over classes of risk



Hosmer DW, Lemeshow S. Confidence interval estimates of an index of quality performance based on logistic regression models. Stat Med 1995; 14(19): 2161–2172. PMID 8552894

Seillier-Moiseiwitsch F. Predictive diagnostics for logistic models. *Stat Med* 1996; 15(20): 2149–2160. PMID 8910960

Cox D. Two further applications of a model for a method of binary regression. *Biometrika* 1958; 45: 562–565

A more intuitive evaluation, although less formal, can be made by the use of the so-called calibration curve, that is plotting the points identifying predicted and observed mortality for each category of risk. The diagonal line reported for comparison represents the perfect identity between observed and predicted mortality (see figure 1 in the Soares reference below).

The greater the agreement between observed and predicted mortality, the better the calibration of the model

THINK: of the possible explanations for the finding in the Soares illustration.



Soares M, Salluh JI. Validation of the SAPS 3 admission prognostic model in patients with cancer in need of intensive care. *Intensive Care Med* 2006; 32(11): 1839–1844. PMID 16977481. Full text (pdf)

Hosmer DW, Lemeshow S. *Applied logistic regression*. New York, (NY): John Wiley & Sons Inc.; 1989. ISBN 0471615536

Harrison DA, Brady AR, Parry GJ, Carpenter JR, Rowan K. Recalibration of risk prediction models in a large multicenter cohort of admissions to adult, general critical care units in the United Kingdom. *Crit Care Med* 2006; 34(5): 1378–1388. PMID 16557153

Glance LG, Dick AW, Osler TM, Mukamel D. Using hierarchical modeling to measure ICU quality. *Intensive Care Med* 2003; 29(12): 2223–2229. PMID 14534777. Full text (pdf)

Poole D, Rossi C, Anghileri A, Giardino M, Latronico N, Radrizzani D, et al. External validation of the Simplified Acute Physiology Score (SAPS) 3 in a cohort of 28,357 patients from 147 Italian intensive care units. *Intensive Care Med* 2009; 35(11): 1916–1924. PMID 19685038. Full text (pdf)

Bertolini G, D'Amico R, Nardi D, Tinazzi A, Apolone G. One model, several results: the paradox of the Hosmer–Lemeshow goodness-of-fit test for the logistic regression model. *J Epidemiol Biostat* 2000; 5(4): 251–253. PMID 11055275

Poole D, Bertolini G. Outcome-based benchmarking in the ICU Part I: Statistical tools for the creation and validation of severity scores. In: Chiche J-D, Moreno R, Putensen C, Rhodes A, editors. *Patient Safety and Quality of Care in Intensive Care Medicine*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2009. pp. 141–150. ISBN 978-3-941468-11-5

Poole D, Bertolini G. Outcome-based benchmarking in the ICU Part I: Use and limitations of severity scores in critical care. In: Chiche J-D, Moreno R, Putensen C, Rhodes A, editors. *Patient Safety and Quality of Care in Intensive Care Medicine*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2009. pp. 151–160. ISBN 978-3-941468-11-5

Uniformity-of-fit

Uniformity-of-fit reflects the model's performance in sub-groups of patients. The rationale for this evaluation is to explore specific groups of patients that can have an important impact on

the performance of the model. No consensus exists about the sub-groups suitable for this analysis or the best technique for identifying them. The following factors are those commonly considered to most influence the case-mix and are consequently used to test the model performance:

Sub-groups of patients in your ICU can reduce the performance of the model and introduce bias into your results

- Location in the hospital before ICU admission
- Patient type (non-operative, emergency surgery, scheduled surgery)
- Degree of physiological dysfunction
- Physiological reserve (age, chronic diagnosis)
- Acute diagnosis.

This evaluation is more important when studying highly specialised ICUs with specific patient characteristics. However, it can also be important in general ICUs. For further details on the above issues you will find the following reference helpful:



Moreno R, Apolone G, Miranda DR. Evaluation of the uniformity of fit of general outcome prediction models. *Intensive Care Med* 1998; 24(1): 40–47. PMID 9503221. Full text (pdf)

Poor predictive performance

The discrimination of these prognostic models in new samples of patients is generally acceptable or good. Calibration is frequently not adequate, with mortality underestimated in low-risk patients and overestimated in high-risk patients. To solve the problem of poor performance, models can be modified to describe the new population of patients more accurately. This process is called customisation and usually is performed by the following two techniques:

Two techniques are available for customisation: first- and second-level customisation

- **First-level customisation** involves modification of the original formula with maintenance of the same variables with the same weights as in the original model.
- **Second-level customisation** means re-evaluation of the components of the model, which may have weights changed and/or new variables added.

These techniques need large samples and considerable expertise, so they are performed only in the context of collaborative groups, since the sample sizes needed for their correct application is usually beyond those available at a single centre. The major drawback is that customisation results in a loss of comparability between the new population and the reference population. For instance, if a customised model is developed and adopted in a given country, then only ICUs of that country can be compared while no comparison can be made with ICUs of other countries. Also, it should be noted that customisation does not have any impact (first-level customisation) or only a small impact (second-level customisation) on calibration.

Q. What is the value of having customised models at national and regional level?

A. In some European countries, such as the United Kingdom, Italy, Switzerland, Austria and Spain, general outcome models have been adapted for the specific characteristics of their population. For national comparisons, these models probably provide more accurate prognostic information.

References on first and second-level customisation:



Moreno R, Apolone G. Impact of different customization strategies in the performance of a general severity score. *Crit Care Med* 1997; 25(12): 2001–2008. PMID 9403750

Metnitz PG, Valentin A, Vesely H, Alberti C, Lang T, Lenz K, et al. Prognostic performance and customization of the SAPS II: results of a multicenter Austrian study. *Simplified Acute Physiology Score. Intensive Care Med* 1999; 25(2): 192–197. PMID 10193547. Full text (pdf)

Aegerter P, Boumendil A, Retbi A, Minvielle E, Dervaux B, Guidet B. SAPS II revisited. *Intensive Care Med* 2005; 31(3): 416–423. PMID 15678308. Full text (pdf)

Metnitz B, Schaden E, Moreno R, Le Gall JR, Bauer P, Metnitz PG; ASDI Study Group. Austrian validation and customization of the SAPS 3 Admission Score. *Intensive Care Med* 2009; 35(4): 616–622. PMID 18846365. Full text (pdf)

You can find a calculator for the expanded version of SAPS II at http://www.sfar.org/scores/igs2_expanded.php

4/ HOW TO MEASURE AND COMPARE PERFORMANCE OF ICUS

Evaluating the performance of the ICU

Link to [ESICM Flash Conference Philipp Metnitz, Vienna, Austria. Benchmarking ICUs. Berlin 2007](#)

As medical practitioners we have a responsibility to audit the quality of care we provide to our patients. An important aspect of this audit is evaluation of performance, or, in other words, how well we are treating our patients. In some contexts, regulatory agencies will also ask for relevant information.



List five reasons, of importance in your local context, why you might compare the performance of your unit with that of other ICUs. Discuss the list with colleagues.

Effectiveness of the ICU and the lack of demonstration of ICU efficacy

Drugs as well as other health care interventions must be shown to have efficacy before their introduction in the market or their widespread implementation. When considering overall ICU care it would be useful to have a measure of the effect on patient outcome (efficacy). Unfortunately, we have to admit that the efficacy of ICU care has not been demonstrated using rigorous scientific methods such as randomised control trials. Such a study would be unethical because it is inappropriate to withhold advanced organ support to patients who might benefit. Therefore, we can only rely on an evaluation of effectiveness, that is on the assessment of efficacy relative to the real world clinical practice around us.

The comparison of an individual ICU with reference ICUs participating in patient data collection for the development of a given outcome prediction model allows evaluation of the relative performance of the ICU.

A large group of ICUs collecting patient data for the development of a given outcome prediction model is taken as a reference. Consequently, we can compare the effectiveness of our individual ICU with the reference ICUs, by applying the prediction model to our patients. In other words, the question is whether the outcome of the patients treated in our ICU is in accordance with the expected outcome i.e. a relative evaluation of ICU performance or a 'benchmarking' process.

With this approach, the actual outcome in the population under analysis is compared to the outcome in a reference population while controlling for case-mix by using general outcome prediction models. We replace the question 'How good are our results?' with 'Are our results better (or worse) than others?'

THINK: of the practical reasons that make the evaluation of ICU performance a difficult task. Also think of clinical situations where the arguments developed in paragraph one above (in relation to absolute effectiveness) might not be unethical?

Comparing ICUs: how to adjust for case-mix

There are large differences between ICUs in terms of patient diagnoses, previous health status, demographic factors and degree of physiological dysfunction. Comparison of crude mortality rates may therefore be misleading and specific instruments are needed to adjust crude mortality rates for the severity of illness of the patients being treated in each ICU. The instruments available are the outcome prediction models.

There are important differences in case-mix between ICUs that must be taken into account in order to ensure comparability



Niskanen M, Reinikainen M, Pettilä V. Case-mix-adjusted length of stay and mortality in 23 Finnish ICUs. *Intensive Care Med* 2009; 35(6): 1060–1067. PMID 19125233. Full text (pdf)

Standardised mortality ratio and ICU performance

Several investigators proposed the use of the ratio between observed and predicted deaths (standardised mortality ratio, SMR) as an indicator of the effectiveness of care in the ICU. The assumption is that although ICUs may admit very heterogeneous groups of patients in terms of age, co-morbidities or acute health status, outcome prediction models can account for most of these confounders.

To evaluate the performance of a given ICU we can sum the expected probabilities of death for all the patients in that ICU to generate an overall expected mortality. The expected mortality is then compared with the observed mortality to generate the SMR. Additional calculations will produce the confidence interval of the ratio (usually using non-parametric estimations).

The Standardised Mortality Ratio is a summary measure of ICU performance, controlling for case-mix

$$\text{SMR} = \frac{\text{Observed mortality}}{\text{Predicted mortality}}$$

If the observed mortality in the ICU is higher than the expected (SMR > 1.0) its performance may be poor, while if the observed mortality is lower than the predicted (SMR < 1.0) performance may be better. The 95% confidence interval will give strength to the results, when it does not include the number, 1. An example of this methodology can be found in figure 1 in the Le Gall reference below. The figure shows standardised mortality ratios (SMR) with 95% confidence intervals for 16 French ICUs participating in the EURICUS I study. Predicted mortality was based on the SAPS II model using either the original, the expanded, or the customised SAPS II. A–P indicate different intensive care units.

Note the variation among ICUs with the original model, the customised model



Le Gall JR, Neumann A, Hemery F, Bleriot JP, Fulgencio JP, Garrigues B, Gouzes C, Lepage E, Moine P, Villers D. Mortality prediction using SAPS II: an update for French intensive care units. *Crit Care* 2005; 9 (6): R645-52. PMID: 16280063. Available from <http://ccforum.com/content/9/6/R645>



If your ICU were to appear in the upper half of the illustration on the previous screen (i.e. $SMR > 1.0$), prepare an explanation for this finding that might be intelligible to a lay person.

Assuming that the errors resulting from data collection are small and randomly distributed and that the model is well-calibrated, the differences between predicted and observed mortality may be attributed to local variations in the quality of care. Nevertheless, SMR methodology assumes that the relationship between severity of illness and performance is constant, but an ICU may, for instance, perform well when treating low-risk patients but perform less well when treating high-risk patients. Therefore, visualising the individual risk profile of each ICU over the whole spectrum of the expected hospital mortality should allow any given ICU to understand the quality of care given in each severity band. Accordingly, each ICU can see where care needs to be improved.

Figure 1 in the Moreno reference below shows the association between SAPS II score and predicted mortality in the reference cohort of 77 Austrian ICUs and a single ICU. In this example, the ICU has a poor performance in low-risk patients (mortality predicted in the reference ICUs = 10%; observed in the ICU in question = 17%).



Moreno RP, Hochrieser H, Metnitz B, Bauer P, Metnitz PG. Characterizing the risk profiles of intensive care units. *Intensive Care Med* 2010; 36(7): 1207–1212. PMID 20306015. Full text (pdf)

Q. What is the value of recording SMR on a continuous basis? What assumptions are made in interpreting SMR data?

A. Evaluation of the standardised mortality ratio over time is a good way of tracking changes in your unit's performance, assuming that your case-mix has not altered significantly change in recent years e.g. due to a major organisational change. Great care should be used in the collection and analysis of data to ensure comparability.

Q. If the SMR in your ICU has changed significantly in recent years what significance might be attached to this finding?

A. The pattern of changes in standardised mortality ratio can provide useful insights into performance. Account must be taken, however, of random variability; computing the confidence intervals for the ratio may be helpful in this regard.

NOTE In comparing ICU performance, do not forget that unmeasured case-mix variation can have an important impact on the findings. If the reference population is very different from your population, the uncritical observer might be misled.



Liu J, Louis TA, Pan W, Ma JZ, Collins AJ. Methods for Estimating and Interpreting Provider-Specific Standardized Mortality Ratios. *Health Serv Outcomes Res Methodol* 2003; 4(3): 135–149. PMID 19606272

Moreno RP, Hochrieser H, Metnitz B, Bauer P, Metnitz PG. Characterizing the risk profiles of intensive care units. *Intensive Care Med* 2010; 36(7): 1207–1212. PMID 20306015. Full text (pdf)

Limitations of the SMR approach

Almost all the assumptions behind the use of the SMR have been challenged in recent years. Differing definitions of the variables used, inadequate intra- and inter-observer reliability of collected data and poor calibration of the models used have all been shown to affect mortality prediction. In addition, our ability to control for case-mix variation is increasingly questioned. Even the choice of the outcome (usually hospital mortality) is subject to debate since it may be prone to bias and manipulation.

Furthermore, the use of hospital outcome to evaluate the performance of an ICU means that poor outcome in the general wards following ICU discharge will be attributed to ICU performance. Unfortunately, a significant number of deaths do occur in hospital after ICU discharge - the so-called post-ICU mortality or occult mortality. You can find information on post-ICU mortality in:

Hospital mortality is influenced by care beyond the ICU physician's control



Fernandez R, Serrano JM, Umaran I, Abizanda R, Carrillo A, Lopez-Pueyo MJ, et al; Sabadell Score Study Group. Ward mortality after ICU discharge: a multicenter validation of the Sabadell score. *Intensive Care Med* 2010; 36(7): 1196–1201. PMID 20221748. Full text (pdf)

Black N. Assessing the quality of hospitals. *BMJ* 2010; 340: c2066 doi: 10.1136/bmj.c2066. PMID 20406863

Outcome prediction models, by design, do not provide information regarding the process of ICU care. In other words, a significantly high (or low) risk-adjusted mortality does not provide any insights into the underlying reasons for the observation. Therefore, the standardised mortality ratio cannot be used in isolation, but should be complementary to a series of quality indicators, and eventually be supplemented by audits examining specific issues.

There are several weaknesses in using standardised mortality ratio to analyse and improve the quality of ICU care



Hinds CJ, Watson JD. *Intensive Care: A Concise Textbook*. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. pp. 30–31; Criticisms of severity scoring

THINK: about how performance evaluation can help you improve the quality of care in your unit.



Rothen HU, Stricker K, Einfalt J, Bauer P, Metnitz PG, Moreno RP, et al. Variability in outcome and resource use in intensive care units. *Intensive Care Med* 2007; 33(8): 1329–1336. PMID 17541552. Full text (pdf)

Kahn JM, Brake H, Steinberg KP. Intensivist physician staffing and the process of care in academic medical centres. *Qual Saf Health Care* 2007; 16(5): 329–333. PMID 17913772

A final comment on performance evaluation

To improve the quality of our clinical practice, we need to evaluate ICU performance taking into account not only mortality but also other outcomes e.g. quality of life, quality of death and the appropriate use of resources. Also, our performance evaluation should include the health system as a whole, beginning with the general practitioner and finishing with the post-hospital long-term rehabilitation and after care. The ICU as part of this continuum of care can and should be evaluated, not only in terms of the short-term results but also in terms of its interactions with the global healthcare system.

To improve quality of care we need multiple measures of the structure, the effectiveness and the process of ICU care

We must address organisational issues of our ICU rather than focusing exclusively on clinical, patient-related, factors. The non-clinical organisation and management of ICUs undoubtedly influence the clinical outcome of intensive care patients. This implies that these factors should be taken into account in our evaluation process.

ICU performance is significantly influenced by non-clinical factors

THINK: of the way performance can be influenced by non-clinical factors. Try to indicate the most important in your setting.



Reinikainen M, Karlsson S, Varpula T, Parviainen I, Ruokonen E, Varpula M, et al. Are small hospitals with small intensive care units able to treat patients with severe sepsis? *Intensive Care Med* 2010; 36(4): 673–679. PMID 20143222. Full text (pdf)

Najjar-Pellet J, Jonquet O, Jambou P, Fabry J. Quality assessment in intensive care units: proposal for a scoring system in terms of structure and process.

Intensive Care Med 2008; 34(2): 278–285. PMID 17932651. Full text (pdf)

Kahn JM, Fuchs BD. Identifying and implementing quality improvement measures in the intensive care unit. Curr Opin Crit Care 2007; 13(6): 709–713. Review. PMID 17975395

Since providing the clinician with risk-adjusted mortality figures does not have any detectable effect on outcomes, we should replace our current benchmarking process with more comprehensive measures that simultaneously assess several dimensions of performance. These can then be used to improve our practices.

We need to replace the present benchmarking process with more comprehensive measures

See also the PACT modules on

- Organisation and management
- Quality assurance and cost-effectiveness; How to assess the performance of your ICU? Vertical versus horizontal comparison

CONCLUSION

In the first Task we reviewed the different aspects of outcome after critical illness and emphasised the importance of selecting the appropriate outcome measure(s). In Task 2 we discussed how outcome can be measured and stressed the importance of selecting the correct tool to measure the outcome of interest. In the third Task we discussed various outcome prediction models and how to use them. The importance of evaluating the performance of severity scores was stressed. In the final Task we considered approaches to performance evaluation and in particular the use and limitations of the standardised mortality ratio. Overall the valid and reliable measurement of clinical outcomes is vitally important if critical care is to advance through research and continue to improve the quality of care delivered to our patients.

SELF-ASSESSMENT QUESTIONS

EDIC-style Type K

1. Long-term survival after intensive care is dependent on

- A. Co-morbidity
- B. Maximum organ failure during ICU admission
- C. Age
- D. Admission group

2. Hospital survival of ICU patients

- A. Is influenced by inter-hospital transport
- B. Is reported to be the same as the ICU survival rate
- C. Is reported not to be dependent on the standard of post-ICU care
- D. Is dependent on organ dysfunctions at ICU discharge

3. ICU patients compared to the general population

- A. Are older
- B. Prior to admission have similar health-related quality of life (HRQOL)
- C. Prior to admission have more co-morbidities
- D. After 2–3 years post-ICU discharge have the same life-expectancy

4. Quality of life after ICU discharge

- A. Declines from 3 to 12 months
- B. Is equivalent to the pre-admission level after 12 months
- C. Is not related to other life events in the first 12 months after ICU discharge
- D. Is not related to the severity of illness (during ICU stay) at 12 months post-ICU discharge

5. Common tools to investigate Health-related quality of life (HRQOL) include

- A. Glasgow outcome score
- B. Short form 36
- C. EuroQual 5-D
- D. Karnofsky score

6. Disease specific scoring in ICU patients

- A. Includes the Ranson score in acute pancreatitis
- B. Uses the Imrie score for acute renal failure
- C. The Child–Pugh score is used in acute liver failure
- D. Are superior to general severity scoring systems

7. The standardised mortality ratio (SMR) is frequently used in evaluation of the performance in an ICU

- A. The value should be as close to 1.0 as possible
- B. Lower values mean better performance
- C. Is the predicted: observed mortality
- D. Is the observed: predicted mortality

EDIC-style Type A

8. Mortality attributable to critical illness is primarily dependent on the following EXCEPT:

- A. Degree of physiologic dysfunction
- B. Age of the patient
- C. Performance of the ICU
- D. Gender
- E. Co-morbidity

9. Frequently reported mental disturbances after ICU discharge include the following EXCEPT:

- A. Anxiety
- B. Depression
- C. Asperger syndrome
- D. Post-traumatic stress disorder
- E. Cognitive dysfunctions

10. The chances of survival in an individual critically ill patient can best be estimated by

- A. SAPS III score
- B. Maximum SOFA score
- C. APACHE IV score
- D. Trained clinical judgement
- E. Glasgow coma score at admission

11. Which patient group should NOT be included in SAPS II and APACHE II score?

- A. Acute pancreatitis
- B. Severe sepsis
- C. Acute myocardial infarction
- D. Multiple trauma with head injury
- E. Exacerbation of COPD

12. Regarding the SAPS II and the SAPS III scoring systems for severity of acute disease: all of the following are true EXCEPT:

- A. SAPS II depends more on acute physiologic derangements
- B. SAPS III includes time in hospital before ICU admission
- C. SAPS III includes effect of gender
- D. SAPS II includes presence of AIDS
- E. SAPS III includes a shorter time frame for collecting data

13. The ability of a severity based outcome model to distinguish survivors from non-survivors is called:

- A. Calibration
- B. Uniformity-of-fit
- C. Limitation
- D. Discrimination
- E. Dissertation

Self-assessment answers**Type K**

Q1.	Q2.	Q3.	Q4.	Q5.	Q6.	Q7.
A. T	A. T	A. T	A. F	A. F	A. T	A. F
B. F	B. F	B. F	B. F	B. T	B. F	B. T
C. T	C. F	C. T	C. F	C. T	C. T	C. F
D. T	D. T	D. T	D. T	D. F	D. F	D. T

Type A

- 8. Answer D is correct
- 9. Answer C is correct
- 10. Answer D is correct
- 11. Answer C is correct
- 12. Answer C is correct
- 13. Answer D is correct

PATIENT CHALLENGES

The regional office responsible for the ICU in your hospital has announced the release of extra financial resources to improve services for critically ill patients. In your region there is a network of several ICUs. All are adult general ICUs and include neurosurgical services but not hepatic, cardiac or other specialty tertiary centres. Your ICU network has two teaching hospitals and other, not necessarily smaller, non-university affiliated hospitals.

Within your network, there is pressure on the critical care services resulting in transfers (for reasons other than an upgrade in medical care) and cancellation of elective surgery because of lack of nursing and medical staff. Management has decided to release extra funding but wants to direct the money to those ICUs with the best performances. Clearly relevant data need to be supplied.

LEARNING ISSUES

Disbursement of healthcare funds
Methods of categorising ICUs



Department of Health. Comprehensive Critical Care. A review of adult critical care services. London; Department of Health, 2000:9–10

<http://www.dh.gov.uk/Publications>

LEARNING ISSUES

Regional intensive care services
Transportation

Link to PACT module on Transportation

Link to PACT module on Traumatic brain injury Task 5

Pressure on critical care services

Q. What categories of data reflect performance?

A. Data can be broadly categorised into those evaluating the economic performance of the ICU e.g. activity, length of stay and costs and those that assess clinical performance e.g. risk-adjusted mortality, safety and quality indicators.

Activity data describe throughput and include information such as the number of patients treated, occupancy rate and ICU length of stay. ICUs may have similar activity but different workloads. Workloads reflect the intensity of work and can be measured by parameters such as severity of illness, diagnosis and case-mix, dependency of patients and therapeutic interventions provided. It also allows the

classification of the ICUs into different levels of care and they can be evaluated for efficacy by the work utilisation ratio. Cost is the measurement of resource consumption and allows comparisons between ICUs; it is the first step in the assessment of cost-effectiveness.

LEARNING ISSUES

Data collection
Severity of illness
Case-mix
Level of intensive care
Degree of patient dependency



Moreno R, Reis Miranda D. Nursing staff in intensive care in Europe: the mismatch between planning and practice. *Chest* 1998; 113(3): 752–758. PMID 9515853

Q. Can you think of examples of systems that collect data about activity, workload and cost?

A. Various models exist, the characteristics of which are influenced in part by local factors. In the UK, the Augmented Care Period data set is a minimum amount of data that is collected on each critically ill patient. Similar systems exist in many other countries, such as Austria (ASDI), Italy (GiViTi) and Australia-New Zealand (ANZICS database).

The Augmented Care Period data set contains 12 simple data items (a numerical identifier, data about where the patient was referred from, who looked after the patient, how long the patient was in ICU, whether the patient required intensive or intermediate care, patient survival and to where the patient was discharged). The APACHE system, developed in the USA, partially reflects workload as well as outcome. In these countries and elsewhere methods are being developed for national comparative audit of severity of illness adjusted outcome and standard cost collection in ICUs.

LEARNING ISSUES

Cost issues



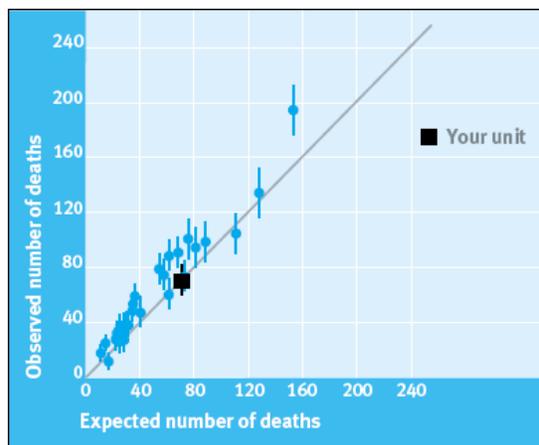
Hinds CJ, Watson JD. *Intensive Care: A Concise Textbook*. 3rd edition. Saunders Ltd; 2008. ISBN: 978-0-7020259-6-9. p. 16; Costs



Edbrooke D, Hibbert C, Ridley S, Long T, Dickie H. The development of a method for comparative costing of individual intensive care units. *The Intensive Care Working Group on Costing. Anaesthesia* 1999; 54(2): 110–120. PMID 10215705

Fortunately the ICUs in your region have been collecting severity of illness data for a number of years and so you are able to show the performance of your unit in comparison with others using the standardised mortality ratio (SMR).

Below is a scatter diagram of the SMRs in some ICUs from your country together with their 95% confidence limits. Your unit is shown as a square while all the others are shown as circles.



You are pleased to report that your ICU has a SMR of one and from the graph is clearly one of the best units because, in relation to others, your unit has fewer expected deaths.

LEARNING ISSUES

SMRs

Scoring systems (critical care)

Limitations of scoring systems

Measurement of performance



Rotondi AJ, Angus DC, Sirio CA, Pinsky MR. Assessing intensive care unit performance: a new conceptual frame-work. *Current Opin Crit Care* 2000; 6:155-157.

Q. This performance measure is based on the APACHE system. Are there any problems applying the APACHE system in this way?

Prompt: The APACHE score is one of the scoring systems used internationally. Another widely used scoring system is SAPS II.

A. The APACHE system is well refined as it is now in its fourth revision but it was developed using exclusively patients from ICUs in the USA. International differences in case-mix may mean that its prediction of hospital mortality is not as reliable when used in Europe or in other continents. The scatter diagram above shows that the majority of units in Europe have a higher than expected proportion of fatalities.

Q. How would you interpret these findings?

A. This means that either these ICUs genuinely have a high adjusted death rate or that the scoring system is not properly calibrated for the population under study.

Q. Is the SMR sufficiently robust to be used as a performance measure? Mention two practical issues that may have an important influence on the score.

A. The SMR is crucially dependent on the calculated risk of hospital mortality which itself is generated from the physiological data collected for the APACHE (or SAPS) score. Small changes (increases or decreases of two points) in APACHE score can make big differences to the SMR. Also, small variations in post-ICU discharge mortality can have a striking effect on the SMR.

This important observation is addressed in greater detail in:



Goldhill DR, Withington PS. Mortality predicted by APACHE II. The effect of changes in physiological values and post-ICU hospital mortality. *Anaesthesia* 1996; 51(8): 719–723. PMID 8795311

The regional office accepts the information you provide about your SMR and expresses satisfaction with your ICU's performance. However the office feels that more information is needed before it can identify the best performing units with confidence and so determine which ICU should receive the available extra resources. You decide to summarise the data concerning the number of patients you have treated and their length of stay in the ICU.

Q. When summarising length of stay, which measures of central tendency and dispersion are most appropriate?

Prompt: You can find a definition of these terms in the Weissman reference below.

A. Measures of central tendency include the mode, mean and median while measures of dispersion include the standard deviation, the interquartile range and range. As the distribution of length of ICU stay is always positively skewed with most patients

staying a short period of time, the summary statistics should not be disproportionately influenced by values at the extreme end. Therefore length of ICU admission should be summarised by the median and either the interquartile range or range.



Weissman C. Analyzing intensive care unit length of stay data: problems and possible solutions. *Crit Care Med* 1997; 25(9): 1594–1600. Review. PMID 9295838

LEARNING ISSUES

Relevant statistical analysis



Altman DG. *Practical Statistics for Medical Research*. London: Chapman and Hall; 1991. ISBN 0412276305. pp. 21-38

Unfortunately the regional office is still not satisfied and asks you for information demonstrating the effectiveness of intensive care in general, as it has to allocate funds between several different healthcare programmes.

LEARNING ISSUES

Effectiveness of intensive care



Kuhlen R, Moreno R, Ranieri M, Rhodes A, eds. *25 Years of Progress and Innovation in Intensive Care Medicine*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft, 2007:399-408. ISBN 9783939069478

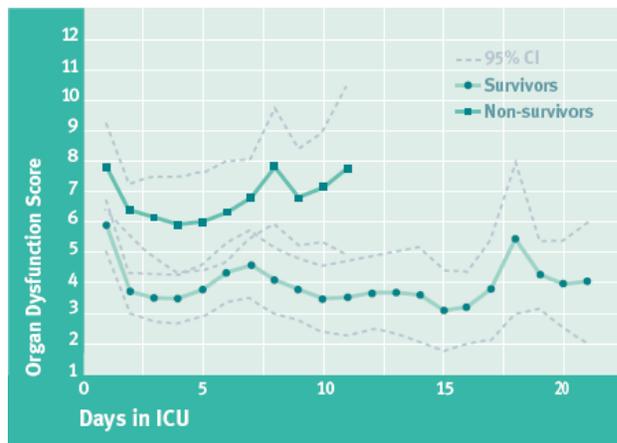
Alberti C, Durand-Zaleski I. ICUs: from performance appraisal to executive dashboard? *Intensive Care Med* 2007; 33(8): 1313–1315. No abstract available. PMID 17541547

Vincent J-L, Rello J, Marshall J, Silva E, Anzueto A, Martin CD, Moreno R, Lipman J, Gomersall C, Sakr Y, Reinhart K; EPIC II Group of Investigators. International study of the prevalence and outcomes of infection in intensive care units. *JAMA* 2009; 302: 2323–2329. PMID 19952319

Q. How you are going to demonstrate that intensive care is effective?

A. Randomised controlled trials comparing intensive care with alternative care elsewhere in hospital are essentially not possible to organise – not least for ethical reasons. However intensive care support may be divided into specific therapy directed at the underlying pathology and general supportive care. Organ dysfunction scores can be used as a global measure of physiological status.

If used over time, organ dysfunction scores may reflect changes in physiology secondary to the effect of intensive care support. Studies have shown that over the first seven days of intensive care, organ failure scores of survivors are lower than non-survivors. In patients staying longer than three days, sequential organ failure scoring has shown that irrespective of the eventual outcome, patients' scores decrease (figure below). The value of the delta (maximum minus admission score) can also be used as an easy way to quantify changes after ICU admission and has been demonstrated to have an excellent correlation with outcome.



This represents an improvement in the physiological status, which could demonstrate the effectiveness of general supportive care delivered in your ICU when compared with the reference population from which the model has been developed.

LEARNING ISSUES

Randomised controlled trials
Organ failure scoring



- Vincent JL, de Mendonça A, Cantraine F, Moreno R, Takala J, Suter PM, et al. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Working group on "sepsis-related problems" of the European Society of Intensive Care Medicine. *Crit Care Med* 1998; 26(11): 1793–1800. PMID 9824069
- Hutchinson C, Craig S, Ridley S. Sequential organ scoring as a measure of effectiveness of critical care. *Anaesthesia* 2000; 55(12): 1149–1154. PMID 11121921
- Moreno R, Vincent JL, Matos R, Mendonça A, Cantraine F, Thijs L, et al. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care. Results of a prospective, multicentre study. Working Group on Sepsis related Problems of the ESICM. *Intensive Care Med* 1999; 25(7): 686–696. PMID 10470572. Full text (pdf)

You report to the regional office saying that in your ICU, the physiological status of 80% of patients staying for more than three days can be maintained or improved by the supportive care of your ICU. The regional office accepts your figures about achieving physiological stability but goes on to ask you to provide information about staff costs, consumables and clinical support services (the three largest cost components in ICUs).

Q. How are you going to cost intensive care so that your figures will be amenable to comparison with your neighbours?

A. Fortunately a valid costing method has been described. The method divides the resources used in intensive care into six cost blocks. The cost blocks are:

- Current cost of equipment
- Estates/Property
- Non-clinical support services
- Clinical support services
- Consumables
- Staff

It has been found that with this method it is difficult to measure the first three cost blocks.

LEARNING ISSUES

Costing methods



Edbrooke D, Hibbert C, Ridley S, Long T, Dickie H. The development of a method for comparative costing of individual intensive care units. *The Intensive Care Working Group on Costing. Anaesthesia* 1999; 54(2): 110–120. PMID 10215705

Hibbert CL, Corcoran M, Dean J, Edbrooke D, Coates E, Jacobs PD. How much does it cost to survive in the intensive care unit? *Rev Port Med Intensiva* 2001;10:179-81.

Csomós A, Janecskó M, Edbrooke D. Comparative costing analysis of intensive care services between Hungary and United Kingdom. *Intensive Care Med* 2005; 31(9): 1280–1283. PMID 15959758. Full text (pdf)

LEARNING ISSUES

Implications of resource constraint Management accounting/budgeting



Edbrooke DL, Hibbert CL. Cost determinants and economic assessment in the critical care setting. *Current Opin Critical Care* 1999; 5: 316-320

LEARNING ISSUES

Presentation of data



Altman DG. Practical Statistics for Medical Research. London: Chapman and Hall; 1991. ISBN 0412276305. pp. 19-47

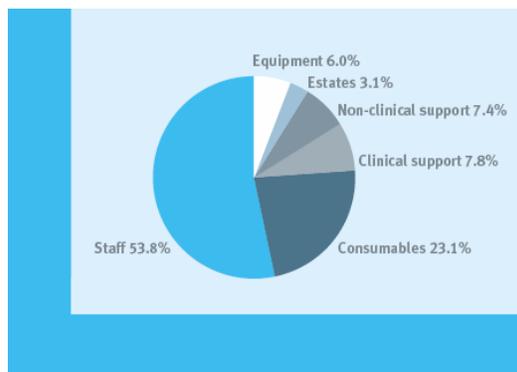
LEARNING ISSUES

Society's expectations in relation to outcome

Q. In the pie chart below the relative contribution of the different cost blocks is identified for a number of ICUs. Clinical support, consumables and staff contribute 85% to the total costs. What are the implications of these figures?

Prompt: Consider the principles of resourcing critical care

A. If costing studies are undertaken in ICU, it is important that as much information as possible about these three cost blocks is obtained.



Unknown to you, the regional office has also asked your neighbouring ICUs for the same information and has produced three league tables with data from all ICUs. The first league table ranks the number of patients treated per ICU bed (activity); the second ranks the SMRs (risk-adjusted mortality) and the third ranks the costs per ICU bed, per day and per patient.

Q. What are the problems associated with league tables that may limit their usefulness?

Prompt: Strengths and weaknesses of league tables are important to understand.

A. League tables should have outcomes that can be generalised to local clinical practice. They should compare units with similar facilities, therapeutic options and healthcare resources. League tables generate a false sense of precision and encourage unwarranted conclusions.



Goldstein H, Spiegelhalter DJ. League tables and their limitations: statistical issues in comparisons of institutional performance. *J R Stat Soc A* 1996;159:385-443.

Poole D, Bertolini G. Outcome-based benchmarking in the ICU Part I: Use and limitations of severity scores in critical care. In: Chiche J-D, Moreno R, Putensen C, Rhodes A, eds. *Patient Safety and Quality of Care in Intensive Care Medicine*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft, 2009:151-60. ISBN 9783941468115

Poole D, Bertolini G. Outcome-based benchmarking in the ICU Part I: Statistical tools for the creation and validation of severity scores. In: Chiche J-D, Moreno R, Putensen C, Rhodes A, eds. *Patient Safety and Quality of Care in Intensive Care Medicine*. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft, 2009:vol 141-150). ISBN 9783941468115

Finally the regional office announces that in the future funding will be distributed on the basis of the ICUs' position in these three league tables.

Q. Is it possible that the regional office has made a mistake?

A. The regional office has made a brave attempt to direct resources in a reasonable manner but may have overlooked:

- The fact that the league tables do not contain important outcome measures that are of key interest to the patient
- That ICU is a support service within the hospital and cannot be considered in isolation. Underfunding intensive care will have repercussions on other services in the hospital, for example cancellation of major elective surgery. Such consequences may not be captured by league tables that focus on one aspect of outcome. League tables cannot accurately quantify quality and humanity of care.

Q. Might it have been possible to consider an alternative approach?

A. Regular consultation between healthcare professionals and the managers is one approach. The conduct of external audits is another.

A further option might be to establish and develop an integrated approach to patient care in conjunction with healthcare professionals elsewhere in the hospital rather than adopting the more traditional compartmentalised (silo) approach.

LEARNING ISSUES

Communication

Continuum of patient care



Ridley SA. Intermediate care, possibilities, requirements and solutions. *Anaesthesia* 1998; 53(7): 654–664. Review. PMID 9771174

PACT module on Communication

Q. How might society be able to evaluate the effectiveness of decision-making by management?

A. The effectiveness of critical care is determined by clinical decisions; however, society can evaluate decision-making by requesting measures that reflect the efficiency with which critical care services are run. Such measures might include the number of patients refused admission or transferred because of lack of facilities and the frequency of adverse events attributable to failures of the organisation or system.

On reflection – Clinical outcome measurement is key to ensuring the quality of critical care, as well as in other branches of medicine. Properly performed, using the tools outlined in this module, measurement of outcome allows meaningful analysis of ICU performance over time (relative to itself) and in benchmarking (relative to other ICUs). Regardless of the healthcare system, resource allocation is a difficult problem, especially as demand is increasing and difficult decisions will need to be based on reliable data concerning the effectiveness of (critical) care delivery.