Evidence based evaluation of immuno-coagulatory interventions in critical care

Systematic reviews with meta-analyses and trial sequential analyses of randomised clinical trials

Arash Afshari

This review has been accepted as a thesis together with 6 previously published papers by University of Copenhagen 19th of May 2011 and defended on 29th of June 2011.

Tutors: Ann M. Møller & Jørn Wetterslev

Official opponents: Jørgen B. Dahl, Else K. Tønnesen & Andrew F. Smith

Correspondence: Department, Department of Anaesthesiology, Juliane Marie center, dept. 4013-4014, Rigshospitalet, University of Copenhagen, Blegdamsvej 9, 2100 Copenhagen, Denmark

E-mail: arriba.a@gmail.com

Dan Med Bull 2011;58:(9) B4316

THE 6 ORIGINAL PAPERS ARE


Background

Clinicians are faced with the increasing challenge of ensuring the best possible care for patients while ensuring the most rational allocation of resources. Due to increasing documentation requirements of interventions impact on health care, evidence based medicine is receiving a growing attention in our society in order to provide guidance for clinical practice and health care decision making. This applies not least to treatment of critically ill patients where interventions are most often costly and the clinical conditions are associated with high mortality.

A. EVIDENCE BASED-MEDICINE (EBM)

EBM is defined as the conscientious use of the best research evidence in clinical decision making while the best research evidence is defined as having the smallest risk of systematic error (bias) and random error (play of chance) providing the most reliable results. However, definition of the best research evidence depends not only on the particular clinical questions but also on the study designs since some are superior to others. The hierarchy of evidence enables the clinicians to assess the benefits and harms of clinical interventions. Randomised clinical trials (RCTs) and systematic reviews are positioned at the top of the hierarchy based on their ability to minimise bias (appendix III). Additionally, assessment of the risk of systematic and random errors are considered the pillars of hierarchy of evidence.

B. THE COCHRANE COLLABORATION AND SYSTEMATIC REVIEWS

The Cochrane Collaboration is an international non-profit and independent organisation dedicated to making up-to-date, accurate information about the effects of health-care widely available. The Cochrane Collaboration produces systematic reviews of health-care interventions and was founded in 1993. It is named after the physician and epidemiologist Archie Cochrane. Systematic reviews are the cornerstone in the Cochrane Collaboration and aim to present information, rather than offer direct treatment guidelines. A systematic Cochrane review uses a predefined, explicit methodology specified in a peer-reviewed protocol. The methods used include meta-analysis and various approaches in order to minimise systematic and random error (bias) during the entire process. Systematic reviews with meta-analysis is an overall statistical approach in which there is a precision weighted combination of results from several trials that address a set of related research hypotheses. Meta-analysis produces a more powerful and precise estimate of an intervention effect and is placed at the top of the evidence hierarchy, due to the ability to increase power and precision.

The Cochrane Library 2010, Issue 12, included 6500 records; comprising 4515 systematic reviews and 1985 protocols for a systematic review from 52 review groups. The Cochrane Anaesthesia Review Group (CARG) has published 70 reviews and 60 protocols for a systematic review in the The Cochrane Library.
CARG reviews make up 1.55% of the Cochrane reviews, CARG protocols make up 3.22% of the Cochrane protocols, and 2.06% of the total number of reviews and protocols.7

C. CONDITIONS

C1. Sepsis and septic shock

Sepsis is diagnosed according to clinical signs, including hypotension, tachycardia, tachypnea, hypoperfusion, lactic acidosis, and an altered body temperature of >38.3°C or <36°C.10 There are several risk factors for the development of sepsis such as male sex, race, increasing age, co-morbid medical conditions, alcohol abuse, and lower socioeconomic status.10 Sepsis and septic shock occur among 15% of all patients admitted to an intensive care unit (ICU), and is associated with a high morbidity, mortality, substantial hospital cost and overall health-care expense.10-12 There is around 750,000 patients with severe sepsis annually in the US and EU with an estimated annual cost of more than $16.7 billion.10-12 Additionally, there is an increase in the incidence of severe sepsis by 1.5-5% per annum, probably due to the rising age of patients and growing population.10,14 Despite massive resource investments to modernise and improve care for critically ill patients, the mortality for sepsis still ranges from 20% to 50% while mortality in septic shock can reach as high as 70-87% in patients with multiple organ dysfunction syndrome (MODS).11,16-18 Severe sepsis is the second most frequent cause of death in ICU surpassed only by cardiovascular events.18 Sepsis as a primary cause of death has more than tripled between 1980 and 2000.19

C2. Disseminated intravascular coagulation (DIC)

Critical illness results in uncontrolled inflammation and vascular damage not only in the presence of infection but also in trauma, malignancy, complications of pregnancy, poisoning, allergic reactions, or liver failure. The inflammation associated with critical illness is characterized by a simultaneous increase in the activity of pro-inflammatory and pro-coagulant processes.17 DIC also known as consumptive coagulopathy is a pathological systemic activation of the coagulation mechanisms that happens in response to a variety of diseases, which at its worst results in simultaneous widespread microvascular thrombosis and profuse bleeding from various sites. These processes lead to capillary leakage, severe disturbance of the microcirculation, tissue damage, and eventually multiorgan failure and death.20 The onset of DIC can be fulminant, as in endotoxic shock or amniotic fluid embolism, or it may be insidious and chronic, as in cases of carcinomatosis. The most common cause of DIC is sepsis. There is a stepwise increase in the prevalence of DIC as the clinical conditions deteriorate from systemic inflammatory response syndrome (SIRS) to sepsis. Irrespective of infectious or non-infectious insults, the clinical combination of SIRS and DIC can synergistically result in MODS.21 The presence of DIC in patients with sepsis is an independent and strong predictor of mortality, probably even stronger than other risk factors.22,23

C3. Acute respiratory distress syndrome (ARDS) and acute lung injury (ALI)

ARDS is defined as acute non-cardiogenic pulmonary oedema, acute severe hypoxemia (PaO2/FiO2 < 200) irrespective of positive end expiratory pressure, bilateral infiltrates on chest radiography, and a pulmonary artery occlusion pressure ≤18 in any adult or child more than one month of age.24 ALI is defined by a hypoxia score between 200 to 300 mmHg in addition to the ARDS criteria. ARDS and ALI are characterized by an inflammatory process of the alveolar-capillary membrane that may arise from a primary lung disease or is secondary to a number of systemic disease processes.25 Hypoxemia in ARDS and ALI is mainly due to a ventilation-perfusion mismatch, resulting in increased intrapulmonary shunting due to pulmonary vasodilatation in non-ventilated lung regions and vasoconstriction in ventilated areas as well as pulmonary hypertension.26 Pulmonary hypertension is believed to be caused by mechanical obstruction of the pulmonary microcirculation by microthromboemboli and hypoxic pulmonary vasoconstriction due to alveolar and interstitial oedema triggered by inflammatory mediators.27 The incidence of ARDS and ALI is reported to be between 14 to 86 persons per 100,000 in a general adult population.28,29 But recent reports indicate a smaller yearly incidence of ALI and ARDS of 10.6 and 5 per 100,000, respectively.30 Mortality among adults with ALI and ARDS is 24% to 60% depending on age, the underlying health status, and clinical condition (e.g., worst when sepsis, multi-organ failure or immunocompromised).31-34 The yearly incidence of ARDS and ALI among children is around 13 per 100,000 with in-hospital mortality of 18% to 23%; with pneumonia, aspiration, and sepsis as the primary causes.35,36

C4. Coagulopathy and massive transfusion

Coagulopathy defined as a defect in the body's mechanisms for blood clotting, causing bleeding diathesis can be secondary to massive transfusion and uncontrolled bleeding. Coagulopathy leads to: defect clot firmness due to fibrinogen, coagulation factor, and platelet deficiency; decreased clot stability due to hyperfibrinolysis and factor XIII deficiency; and prolonged clot generation due to various coagulation factor deficiencies.37-40 Coagulopathy is frequently enhanced by hypothermia (body core temperature < 35°C); acidosis (pH < 7.1); hypovolaemia and low arterial blood pressure (< 70 mmHg); shock and tissue anoxia; low haematocrit (< 30%); low ionized calcium; extensive tissue trauma; coagulation factor and fibrinogen deficiency; and finally silent pre-existing bleeding disorders.41,42 Coagulopathy as an isolated entity is just one cause of bleeding. However, despite the ability of various test systems to identify coagulopathy, they are unable to predict bleeding in a reliable fashion. Often, surgical bleeding or arterial injuries are the dominant reasons for blood loss resulting in high transfusion requirement. Thus, identifying the cause of bleeding does not automatically resolve the problem. Massive transfusion is defined as the total replacement of a patient's blood volume in a period of 24 hours, or a transfusion of at least four red blood cell concentrates within one hour or the replacement of 50% of the total blood volume within three hours. Massive transfusion is an independent risk factor of death.43-45

D. SEPSIS INDUCED ACTIVATION OF INFLAMMATION AND COAGULATION

Sepsis is often regarded as a condition of pro-inflammation.10 Since most of the symptoms of sepsis and fulminant septic shock are considered to be caused by pro-inflammatory and coagulant mediators, extensive search for immuno-treatment has been undertaken with the aim of reducing the inflammatory response by acting on specific mediators.45 Interventions targeted at the inflammation and coagulation pathway are based on the rationale that severe septic patients develop DIC with microvascular dysfunction and impaired tissue oxygenation. In severe sepsis, mononuclear cells, stimulated by pro-inflammatory cytokines such as IL-1 or IL-6, express tissue factor, which seems to cause a
systemic activation of coagulation. Pro-inflammatory cytokines are potential biomarkers and independent predictors of an adverse outcome. Inflammatory activation in patients with severe infection is almost invariably related to activation of coagulation, which in turn may modulate the inflammatory response (Figure 1). Three major anticoagulant pathways are believed to regulate the activation of coagulation in sepsis: AT III, the protein C system, and tissue factor pathway inhibitor. During sepsis-induced activation of coagulation, the function of all three pathways can be impaired (Figure 2).

AT III predominately blunts activation of many inflammatory mediators by inhibiting factor Xa and thrombin but also has inhibitory properties toward tissue factor–factor VIIa (TF-FVIIa) and FXa. Apart from its anticoagulant activities, AT III is believed to possess direct anti-inflammatory properties many of which are mediated by its actions in the coagulation cascade. Other anti-inflammatory actions of AT III are mediated by direct interaction with leukocytes and lymphocytes and by inducing prostacyclin release from endothelial cells. AT III directly hinders leukocyte migration and adhesion to endothelial cells, which in turn may alter the severity of capillary leakage and subsequent organ damage.

Protein C is an endogenous anticoagulant and anti-inflammatory protein that is activated by binding to the thrombin–thrombomodulin complex on the endothelium. Lower levels of protein C and higher levels of circulating thrombomodulin indicate a procoagulant state. However, low levels of protein C can be caused by several other conditions such as liver insufficiency, vitamin K deficiency, and DIC. Protein C together with its cofactor protein S degrade the coagulation cofactors Va and VIIa, thus acting as an effective anticoagulant. In patients with severe inflammation, the protein C system is malfunctioning at virtually all levels (e.g., falling plasma levels, diminished activation, inadequate function, down-regulation of endothelial receptors, increased systemic resistance).

Tissue factor pathway inhibitor (TFPI) is the third inhibitory mechanism of thrombin generation and is the main inhibitor of the tissue factor–factor VIIa complex. The role of TFPI in the regulation of inflammation-induced coagulation activation is not completely clear.

Proinflammatory cytokines, tumour necrosis factor-α, interleukin-1β, and interleukin-6, contribute to the development of disseminated intravascular coagulation (DIC) and neutrophil-endothelial activation, thus leading to microvascular thrombosis, endothelial injury, and multiple organ dysfunction syndrome (MODS). The bidirectional interplay between coagulation and inflammation through protease-activated receptors (PAR) also plays a role in the development of MODS. TF, tissue factor; TFPI, tissue factor pathway inhibitor; AT, antithrombin; PC, protein C; TM, thrombomodulin; SIRS, systemic inflammatory response syndrome; TNF, tumour necrosis factor; IL, interleukin.

E. INFLAMMATION AND COAGULATION IN ARDS AND ALI

Although patients with ARDS or ALI are a heterogeneous population, they are all characterized by having local and systemic inflammation that causes lung damage and fluid leakage across the alveolar-capillary barrier. NO and inhaled prostacyclins are used because of the potential benefit of modifying the process of inflammation, preserving or restoring oxygen delivery and decreasing mortality in ALI and ARDS patients. In ARDS or ALI, there is an activation of inflammation and derangement of the coagulation and fibrinolytic pathways. The haemostatic systems appear intimately involved in the development and progression of lung failure similar to the coagulation abnormalities observed in sepsis, MODS, and DIC. There is a tissue factor-mediated activation of coagulation, initiated by inflammatory cells and endothelial cell damage followed by breakdown of regulatory anticoagulant mechanisms, especially of the AT III and protein C pathways, inhibition of fibrinolysis as measured by elevated levels of plasminogen activator inhibitors (PAI), and activation of thrombin-activatable fibrinolysis inhibitor.

These processes lead to formation and deposition of fibrin in the microcirculation and the alveoli, obstruction of the microcirculation, disturbance of gas exchange, and ultimately organ failure. Some authors speculate that sepsis and ALI are often not distinguishable, and the observed abnormalities are caused by concomitant sepsis or inflammation rather than by ALI itself. Intravascular thrombin, fibrin, and neutrophils interact synergistically to increase lung endothelial permeability to protein. DIC associated microvascular thrombosis, together with neutrophil-endothelial activation, and secondary endothelial injury contribute to the initiation, course, and prognosis of ARDS or ALI. Higher levels of biomarkers of inflammation and the proinflammatory cytokines such as IL-6, IL-8, intercellular adhe-

Fig.1. Effects of proinflammatory cytokines on the regulator of thrombosis such as platelets, neutrophils, monocytes/macrophages, and endothelium. TNF, tumor necrosis factor-α; IL-1, interleukin-1; IL-6, interleukin-6; ICAM-1, intracellular adhesion molecule-1; VCAM-1, vascular cell adhesion molecule-1; PAF, platelet-activating factor; NO, nitric oxide; EPOR, endothelial protein C receptor; t-PA, tissue-type plasminogen activator; PAI-1, plasminogen activator inhibitor-1; S-HT, S-hydroxytryptamine; VWF, von Willebrand factor; TFPI, tissue factor pathway inhibitor; MMP, matrix metalloprotease; DIC, disseminated intravascular coagulation; MODS, multiple organ dysfunction syndrome. Reprint with permission from publisher.

Fig.2. Proinflammatory cytokines, tumour necrosis factor-α, interleukin-1β, and interleukin-6, contribute to the development of disseminated intravascular coagulation (DIC) and neutrophil-endothelial activation, thus leading to microvascular thrombosis, endothelial injury, and multiple organ dysfunction syndrome (MODS). The bidirectional interplay between coagulation and inflammation through protease-activated receptors (PAR) also plays a role in the development of MODS. TF, tissue factor; TFPI, tissue factor pathway inhibitor; AT, antithrombin; PC, protein C; TM, thrombomodulin; SIRS, systemic inflammatory response syndrome; TNF, tumour necrosis factor; IL, interleukin. Reprint with permission from publisher.
sion molecule 1 (ICAM-1) and plasminogen activator inhibitor 1 (PAI-1) appear to predict worse outcomes, increased length of ventilation and raised mortality.\textsuperscript{51,63-66} Lower levels of protein C in pulmonary edema fluid as well as lower plasma levels have also been associated with increased mortality in ARDS or ALI regardless of the presence or absence of sepsis. However, a recent RCT of activated protein C administration in patients with ARDS or ALI failed to show beneficial effect and was stopped early because of lack of efficacy (Michael Matthay, unpublished data). Additionally, another recent published RCT among patients with ALI in the absence of severe sepsis and high disease severity failed to show a beneficial effect on length of ventilation and mortality.\textsuperscript{67}

Injurious high tidal volumes and positive-pressure ventilation alone can cause activation of inflammatory cells, derangements in coagulation and fibrinolysis, induce tissue remodelling, and finally activate inflammatory cells, derangements in fibrinolysis contributing to bronchoalveolar fibrin turnover, or whether this is a consequence of a systemic disease state.

F. MICROVASCULAR THROMBOSIS

Various disorders in critically ill patients lead to disseminated microvascular thrombosis. There are two distinct major syndromes associated with microvascular thrombosis and MODS in critically ill patients, which share pathologic features: thrombotic microangiopathies (TMA) and DIC (Figure 3).\textsuperscript{31} TMA is often seen in disorders such as thrombotic thrombocytopenic purpura (TTP)/haemolytic uremic syndrome (HUS), antiphospholipid antibody syndrome, and the haemolytic anaemia, elevated liver enzyme, and low platelets syndrome. DIC is characterized by the widespread activation of tissue-factor dependent coagulation, insufficient control of coagulation by physiologic anticoagulation pathways, and plasminogen activator inhibitor-1–mediated attenuation of fibrinolysis.\textsuperscript{21} There is a widespread formation of fibrin clots, microvascular occlusion, and reduced oxygen delivery to cells and tissues which may lead to MODS. Regardless of the pre-existing condition, activation of coagulation begins as an adaptive host response, with the aim of preventing spread of microorganisms into the systemic circulation, limit exsanguination, and/or promote wound healing. DIC is an exaggeration of this response leading to a consumptive coagulopathy and hemorrhagic diathesis due to the consumption of platelets and coagulation proteins. DIC can be subdivided into three types:

1) Thrombotic phenotype: associated with sepsis, ARDS or ALI, and a late-stage of trauma manifested by microvascular thrombosis and organ dysfunction.\textsuperscript{75}

2) Fibrinolytic (haemorrhagic) phenotype: associated with excessive haemorrhage, often seen in hematologic malignancies, and associated with tissue-type plasminogen activator-induced fibrinolysis and consumptive coagulopathy.\textsuperscript{76,77}

3) Fibrinolytic phenotype associated with early trauma: the presence of excessive fibrinolysis, hypothermia, acidosis, and dilutional coagulopathy may lead to an accentuation of blood loss.\textsuperscript{21}

The interactions of pro-inflammatory cytokines with regulators of thrombosis, such as platelets, leukocytes, and endothelium is complicated and differs depending on the etiology. However, the pro-inflammatory cytokines (e.g., TNF-α, IL-1β, and IL-6) appear to be the most important inflammatory mediators that regulate the process of microvascular thrombosis (Figure 1). Additionally, leukocyte–endothelial cell interactions play essential roles in the endothelial damage resulting from inflammatory responses.\textsuperscript{78}

G. MASSIVE TRANSFUSION AND COMPLICATIONS

Patients requiring massive transfusions suffer from increased risk of death, not only related to the trauma, surgery or underlying disease, but also directly related to transfusions. Blood transfusion per se has several potential adverse effects such as increased risk of disease transmission as well as non-infectious serious hazards such as coagulopathy, immune system compromising effects, ARDS or ALI, and circulatory overload.\textsuperscript{82} A recent systematic review suggests that in adult ICU, trauma, and surgical patients, RBC transfusions are associated with increased morbidity and mortality.\textsuperscript{82} There are many possible explanations for the increased risk of infection and harm associated with transfusion and there is ever more increasing evidence to support that it is beneficial to reduce the amount of blood products transfused in major injury, among critically ill, and during surgery.

There are indications that a restrictive and conservative approach to red blood cell (RBC) transfusion is at least as efficient as a liberal transfusion strategy and that most critically ill patients can tolerate haemoglobin levels as low as 7 g/dl.\textsuperscript{84} RBC transfusion is believed to increase the risk of infection in adult and paediatric settings, with the increased age of blood products as an independent risk factor.\textsuperscript{84,85} But there is also increasing evidence that...
platelet transfusion in the perioperative period might increase the risk of serious adverse events with indications of a persistent negative risk-adjusted effect of RBC and platelet transfusion on quality of life extending well beyond hospitalisation.\textsuperscript{98,99} Application of an aggressive transfusion treatment of coagulopathy in clinical settings such as trauma is often advocated by many investigators despite no real evidence to support such routines.\textsuperscript{90} This practice is mainly based on the limitation of standard laboratory tests and the clinical reality in which the clinicians are often faced with little time in their decision-making. However, in trauma, this early and aggressive transfusion strategy might be an independent predictor of ARDS.\textsuperscript{91} Transfusion-related acute lung injury (TRALI) and transfusion-associated circulatory overload are two feared complications to transfusion in critically ill patients.\textsuperscript{92} Despite controversy regarding the true incidence of TRALI, it appears to be more common than previously considered and independently associated with decreased long-term survival.\textsuperscript{93,94}

H. THE INFLAMMATION-COAGULATION AXIS

Critically ill patients represent a heterogeneous population characterized by systemic inflammation regardless of the cause of their illness. Both infectious (sepsis) and non-infectious insults (trauma) can produce systemic inflammatory response syndrome (SIRS), characterized by systemic pro-inflammatory cytokine release and generalized activation of leukocytes and endothelium, leading to damage of tissues and organs and ultimately multi organ dysfunction syndrome (MODS).\textsuperscript{95} Patients with SIRS often have coagulation abnormalities ranging from subtle activation of coagulation to more robust coagulation activation often evident by a small decrease in platelet count and prolongation of global clotting times, to fulminant DIC characterized by simultaneous widespread microvascular thrombosis and profuse universal bleeding.\textsuperscript{96} There is increasing evidence to support the extensive bidirectional influence between the systems of inflammation and coagulation, whereby inflammation leads to systemic activation of coagulation, and coagulation also markedly affects inflammatory activity.\textsuperscript{96} Activation of coagulation in the critically ill is mediated by inflammatory activity primarily via pro-inflammatory cytokines contributing directly to disease. The coagulation-driven modulation of inflammatory activity is driven by specific cell receptors on inflammatory cells and endothelial cells. This simultaneous activation of coagulation and inflammation can lead to microvascular thrombosis and MODS in patients with severe sepsis.\textsuperscript{97} Apart from the latter, coagulation abnormalities have other adverse effects such as thrombocytopenia and low levels of coagulation factors in sepsis, increasing the risk of bleeding and mortality.\textsuperscript{98-100} The fibrinolytic system is also depressed leading to impaired fibrin removal and enhanced fibrin formation.

The endothelium appears to play a central role in all major pathways involved in the bidirectional co-play between inflammation and coagulation. Endothelial cells may be a source of tissue factor and therefore involved in the initiation of coagulation activation. Additionally, all physiologic anticoagulant systems and various adhesive molecules that modulate both inflammation and coagulation are connected to the endothelium and thus have direct effect on other functions including maintaining vascular barrier function, nitric oxide-mediated vasodilatation, and antioxidant functions.\textsuperscript{98,99} Finally, in the setting of inflammation-induced activation of coagulation, platelets can be activated directly by endotoxin or by pro-inflammatory mediators (e.g., platelet activator) or via thrombin. This activation of platelets accelerates and enhances fibrin formation.\textsuperscript{94}

I. INTERVENTIONS

II. Antithrombin III (AT III)

AT III is primarily a potent anticoagulant with independent anti-inflammatory properties. AT III irreversibly inhibits serine proteases (e.g., activated factor X and thrombin) in a one-to-one ratio, with the generation of protease-AT III complexes. Heparin prevents AT III from interacting with the endothelial cell surface by: binding to sites on the AT III molecule; competing for the AT III binding site; and reducing AT III ability to interact with its cellular receptor. AT III’s anti-coagulant effect is thus greatly accelerated (by a factor of 1000) by heparin; heparin reduces AT III’s anti-inflammatory properties, weakens vascular protection, and increases bleeding events.\textsuperscript{101,102} The blood concentration of AT III falls by 20% to 40% in septic patients and these levels correlate with disease severity and clinical outcome.\textsuperscript{101,103} This reduction in concentration is due to the combined effect of: reduced production of AT III in the liver; inactivation by the enzyme elastase, which is increased during inflammation; and loss of AT III from the circulation into tissues through inflamed and leaking capillary blood vessels. These processes reduce the half-life of AT III from a mean of 55 hours to 20 hours.\textsuperscript{104} The main mechanism of AT III depletion in severe sepsis is linked to consumption of the molecule. It is this depletion of AT III and its presumed ability to modify inflammation that has prompted research into the potential benefits of replenishing AT III levels.

II. Inhaled nitric oxide (INO)

Nitric oxide (NO) is a potent endogenous vasodilator that can be administered via inhalation. A recent survey from Canada found that 39% of specialists used INO for intervention against ARDS.\textsuperscript{105} It is synthesized by the conversion of the terminal guanidine nitrogen atom of L-arginine via endothelial cell calcium-dependent enzyme NO synthetase and then diffuses across the cell membrane to activate the enzyme guanylate cyclase. This enzyme enhances the synthesis of cyclic guanosine monophosphate (cGMP), causing relaxation of vascular and bronchial smooth muscle and vasodilatation of blood vessels.\textsuperscript{106,107} INO provides selective pulmonary vasodilatation in well-ventilated lung units, improves ventilation-perfusion mismatch, and subsequently reduces the elevated pulmonary vascular resistance and pulmonary hypertension seen in ARDS.\textsuperscript{108} INO also increases right ventricular ejection fraction and decreases right end-systolic volume and thus prevents the decapsulation of acute cor pulmonale.\textsuperscript{109} INO has a half life of three to five seconds, binds to haemoglobin, with high affinity to form methaemoglobin and then diffuses from the alveoli to vascular smooth muscle cells adjacent to the alveoli causing vasodilatation.\textsuperscript{110} NO is involved in both the production of and protection from oxidative injury, regulates both immune and inflammatory responses, decreases neutrophil sequestration in the lung, decreases edema formation and regulates its own production.\textsuperscript{111,112} NO is rapidly converted to active intermediates, including nitrogen dioxide, peroxynitrite, and nitro-tyrosine in the presence of superoxide. However, systemic exposure to NO, which is a cytotoxic free radical, or accumulation of its degradation products could result in deleterious effects through formation of other free radicals causing further lung tissue damage, impaired surfactant function, or aggravated circulatory failure.\textsuperscript{113,114} NO alters the
immune function by modifying the release of cytokines and other components of the inflammatory cascade from alveolar macrophages and inhibits the active adhesion molecules and the neutrophil oxidative burst involved in neutrophil migration.\textsuperscript{117-119} Adenosine diphosphate (ADP) and collagen-induced platelet aggregation is significantly inhibited by INO due to an increase in intra-platelet cGMP during the passage of platelets through the lung, and bleeding time is significantly prolonged in a non-dose related manner during inhalation.\textsuperscript{120-122}

I.3. Inhaled prostacyclins
Prostaglandins are lipid mediators, which are synthesized from essential fatty acids by cellular enzymes and have strong physiological properties. They have important effects on endothelium, platelet, uterine and mast cells and are found in virtually all tissues and organs.\textsuperscript{15} Prostacyclin is produced in endothelial cells from prostaglandin H2 (PGH2) by the action of the enzyme prostacyclin synthase. Although prostacyclin is considered an independent mediator, it is called PG5 (prostacyclin I2) and is a member of the family of lipid molecules known as eicosanoids, and a member of the prostanooids (together with the prostaglandins and thromboxane). PG5 is a naturally occurring prostaglandin which is synthesized by vascular endothelial and smooth muscle cells within the lung with a half-life of three to six minutes.\textsuperscript{15} PG5 is a potent vascular smooth muscle relaxant causing vasodilatation of the systemic and pulmonary vasculature resulting in re-distribution of the systemic and pulmonary vasculature. It can be administered as intravenous for pulmonary hypertension, and inhalable preparation for ALI and ARDS.\textsuperscript{123} PG5 has anti-inflammatory properties since it appears to inhibit platelet aggregation and neutrophil adhesion.

Iloprost is a stable, synthetic analogue of PG5 with a plasma half-life of 20 to 30 minutes and similar pulmonary and hemodynamic properties as PG5. It can be administered as an intravenous or inhalable solution.\textsuperscript{124}

Prostaglandin E1 (PGE1) is a naturally occurring prostaglandin with anti-inflammatory capabilities. PGE1 is an arterial vasodilator, a platelet aggregation inhibitor and stimulates intestinal and uterine smooth muscle.\textsuperscript{125} It is mainly used to treat sexual dysfunction or as an intravenous treatment for neonates with congenital heart defects, in order to maintain patency of ductus arteriosus until surgery. Its half-life is 5 to 10 minutes and is primarily removed by the pulmonary vascular bed. It can also be used as an inhalable solution for severe hypoxemia.\textsuperscript{125} Aerosolized PGE1 and PG2 are potent vasodilators, which seem to reduce pulmonary arterial hypertension by lowering pulmonary vascular resistance, improve right-heart function, redistribute pulmonary blood flow to ventilated segments of the lung with matching improvements in ventilation and perfusion to result in better oxygenation. Inhaled prostacyclins cause minimal systemic vasodilatation and have minor transfer to the vascular system. Prostacyclins seem to reduce obstruction of pulmonary microcirculation in ALI and ARDS and modulate the underlying inflammation due to their ability to reduce leukocyte adhesion and their antithrombotic and platelet disaggregating properties.\textsuperscript{126} Additionally, they block neutrophil tethering to blood vessels, and decreases endothelial cell production of various cytokines and chemokines.\textsuperscript{123} However, the principle action of aerosolized prostacyclins is their property of selective vasodilatation to reduce hypoxemia.

I4. Thromboelastography (TEG) or thromboelastometry (ROTEM)
Transfusion of blood products can be guided by clinical judgment or standard laboratory tests (SLTs), thromboelastography (TEG) or thromboelastometry, or a combination of these, in a more or less fixed transfusion algorithm. Generally SLTs include aPTT (activated partial thromboplastin time), PT (prothrombin time), international normalized ratio (INR), platelet count, and plasma-fibrinogen. However, none of these tests were developed to predict bleeding or to guide coagulation management in the surgical setting and they are of limited use in diagnosis, risk assessment, and in relation to algorithms used to guide the administration of blood products for surgical or critically ill patients.\textsuperscript{125} The limitations of these tests include: lack of real-time monitoring; inability to identify singular or multiple coagulation factor deficiencies; no measurement of the effects of hypothermia on haemostasis; no rapid assessment of fibrinolysis, platelet dysfunction, or haemostatic response to injury or surgery.\textsuperscript{127} Additionally, all these tests are performed in plasma at 37°C without the presence of platelets or other blood cells and seem unable to predict the role of the components measured in the context of haemostasis as a whole. Thus, none of these tests can estimate the risk of bleeding but may guide therapy in the presence of clinical bleeding.

TEG is a whole blood coagulation analyser invented by Hartert imitating sluggish venous flow.\textsuperscript{128} It provides an evaluation of the kinetics of all stages of clot initiation, formation, stability, strength and dissolution in whole blood or plasma.\textsuperscript{129,129} In the conventional TEG, a blood sample is placed into a cup, which is then rotated gently with a sensor shaft in the cup while in the rotary thromboelastometry (ROTEM) the sensor shaft rotates, rather than the cup. A clot forms between the cup and the sensor with the speed and patterns of changes in strength and elasticity in the clot measured in various ways by a computer and depicted as a graph.\textsuperscript{130,131} TEG or ROTEM are point of care assays in the perioperative and emergency setting, which produce rapid graphical and numerical results of the haemostatic status with the ability to detect the anti-coagulant effect of acidosis, hypo- or hyperthermia as it can be performed between 22°C and 42°C. TEG and ROTEM are able to detect and quantify the underlying cause of coagulopathy such as thrombocytopenia, factor deficiency, heparin effect, hypofibrinogenemia, and hyperfibrinolysis.\textsuperscript{132}

Treatment for such disorders may involve the transfusion of red blood cells, blood products or specific drugs such as antifibrinolytics or factor substitution and their effect can be evaluated in vitro.\textsuperscript{134} Additionally, it may enable a distinction between a surgical cause of bleeding and coagulopathy, potentially saving time and effort for clinicians, provide a balanced transfusion, reduce incidents of coagulopathies related to massive transfusion and reduce morbidity and mortality by avoiding interventional hesitation and reducing the amount of blood products transfused.\textsuperscript{88,87,92,94}

Much time and many resources have been invested for treatment of critically ill patients. As previously described, often in sepsis, ARDS and ALI the initial trigger appears to be overstimulation of the inflammatory cascade resulting in coagulation dysfunction. Thus various interventions and agents targeting specific mediators of the inflammatory response and the coagulation cascade have been investigated. I also acknowledge that there are multiple additional interventions of interest such as avoidance of injurious tidal volume ventilation, early aggressive resuscitation and early appropriate antibiotics in severe sepsis.\textsuperscript{132,134} However, this is beyond the scope of this dissertation. Nevertheless, I will dis-

DANISH MEDICAL BULLETIN 6
cuss the evidence of additional interventions such as tight glycaemic control, activated protein C and corticosteroids for sepsis in the section summary of evidence.

RISK OF SYSTEMATIC ERRORS (BIAS), RANDOM ERRORS (PLAY OF CHANCE) AND DESIGN ERRORS

Systematic reviews of randomised clinical trials (RCTs) are considered the most comprehensive and most valid way of examining the benefits and harms of interventions. However, the methodological quality of RCTs included in a systematic review can have a substantial impact on estimates of intervention effects, which may alter the validity of the conclusions of a systematic review. RCTs with inadequate methods are associated with bias and tend to exaggerate intervention effects. Bias in clinical trials may be described as systematic errors or deviation from the truth that encourage one intervention over others. Bias can ultimately mislead at all levels of health-care decision making, from a patient perspective to a national public health point of view. The evaluation of the validity of the included trials is an essential component of a Cochrane review influencing the analysis, interpretation, and conclusions of the review.

Critical appraisal of the trial quality is a multidimensional concept, which is only possible if the design, conduct, and analysis of the trials are accurately described. However, the reporting of RCTs is often incomplete due to poor methodology and other mechanisms. Thus, it is essential for the authors of systematic reviews to apply careful appraisal of the methodological characteristics of the RCTs in order to identify strength and weakness of the existing evidence. This will enable the authors to formulate recommendations for treatment strategies and improvement of the conduct and value of future research.

The risk of random error (play of chance) may affect the estimation of intervention effect by overestimation or underestimation of intervention effect. When random error is combined with human wish-bias, it may have the same directional influence on the estimation of the intervention effect as systematic errors, that is overestimation of benefit and underestimation of harms. The risk of design errors may be caused by many factors, such as selection of patients, selection of both experimental interventions, selection of dose controlled interventions, type of outcome measure, timing of outcome measure assessment, patients included in the analyses, etc.

Quality and validity assessment of trials in a systematic review is considered to have two general dimensions: Internal validity and external validity.

1. INTERNAL VALIDITY (SYSTEMATIC AND RANDOM ERRORS)

Internal validity describes the quality of the trial design and execution in order to prevent systematic errors or bias. Bias can result from flaws in the design, conduct, analysis, interpretation, or reporting of a study. In RCTs, systematic errors have been classified into five general categories: selection, performance, detection, attrition and reporting bias. Internal validity implies that the differences observed between groups of patients allocated to different interventions may, apart from random error, be attributed to the interventions under investigation. Quality, or risk of bias, assessment in systematic reviews has been a highly debated and controversial issue. The use of scales or checklists for assessing quality or risk of bias is explicitly discouraged in Cochrane reviews due to lack of empirical evidence. Despite their simplicity (e.g., calculating a summary score), their validity assessment has often shown to be unreliable and they are less likely to be transparent to users of the review. Due to lack of a gold standard to assess the validity of RCTs, the Cochrane Collaboration has introduced a risk of bias tool to assess the internal validity of RCTs.

The risk of bias tool for randomized trials is based on six domains with empirical evidence supporting an association with systematic error: sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other sources of bias (e.g., vested interests). For each domain, the authors are required to conduct critical assessments on the risk of bias (high, low, unclear) separately on the basis of the trial report as well as additional documents (e.g., trial protocol). Differences in risk of bias are found both between the different levels of evidence and within each level of evidence. RCTs should have an overall assessment based on the appraisal of the individual domains. Trials with one or more systematic error components assessed as inadequate or unclear are considered to be of high-risk of bias, while trials with all quality components assessed as adequate are considered to be of low-risk of bias. Differences in risks of bias can help explain variation in the results of the included trials in a systematic review (i.e., heterogeneity). RCTs classified as high or unclear risk of bias trials have larger effect estimates of benefits of interventions than RCTs with a low risk of bias which are more likely to estimate an intervention effect close to the “truth.”

<table>
<thead>
<tr>
<th>Type of bias</th>
<th>Description</th>
<th>Relevant domains in the Cochrane Collaboration’s “risk of bias” tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection bias</td>
<td>Systematic differences between baseline characteristics of the groups that are compared</td>
<td>Sequence generation, allocation concealment</td>
</tr>
<tr>
<td>Perform-ance bias</td>
<td>Systematic differences between groups in the care that is provided, or in exposure to factors other than the intervention of interest</td>
<td>Blinding of participants, personnel &amp; outcomes assessors</td>
</tr>
<tr>
<td>Attrition bias</td>
<td>Systematic differences between groups in withdrawals from a study</td>
<td>Incomplete outcome data, Blinding of participants, personnel &amp; outcomes assessors</td>
</tr>
<tr>
<td>Detection bias</td>
<td>Systematic differences between groups in how outcomes are determined</td>
<td>Blinding of participants, personnel &amp; outcomes assessors</td>
</tr>
<tr>
<td>Reporting bias</td>
<td>Systematic differences between reported and unreported findings</td>
<td>Selective outcome reporting</td>
</tr>
</tbody>
</table>

Table 1. Cochrane classification scheme for bias

2. SYSTEMATIC ERROR CATEGORIES

2.1 Selection bias

Selection bias occurs when prognostic factors are unevenly distributed between the intervention group and the control group, often resulting in less favourable outcomes for patients in the control group. The aim of randomization is to reduce such bias by creating comparable groups for any known or unknown potential confounding factors due ideally to the elimination of confounding by indication. Confounding by indication is a bias frequently encountered in observational epidemiologic studies but may also impact the results of randomised trials. Confounding by indication during randomisation appears when patients are allocated to the intervention or control group on the basis of patient and investigator preferences patient characteristics, and clinical his-
tory, a manipulation that may take place in case of inadequate or unclear allocation sequence generation or inadequate or unclear allocation concealment.139

2.2 Performance bias
Performance bias occurs when there are systematic differences between groups in the care that is provided (e.g., additional treatment interventions provided preferentially to one group) or exposure to factors other than the intervention.140 Adequate blinding of patients and care providers prevents performance bias and also safeguards against differences in placebo responses between the groups.141

2.3 Attrition bias
Attrition bias refers to systematic differences between groups in withdrawals from a study.7 Protocol deviations (e.g., violation of eligibility criteria, non-adherence to treatments) and loss to follow-up often lead to the exclusion of patients after their allocation to intervention groups, which may introduce attrition bias.143 Loss to follow up refers to patients becoming unavailable for examinations at some stage during the study period due to various reasons such as refusal of further participation (drop outs), unable to contact patients, or clinical decisions to stop the assigned intervention. In case of missing data, intention to treat analysis (ITT) is recommended. ITT includes all patients, whereas per-protocol analyses exclude data from patients with protocol deviations from the analyses.155 Suggested strategies to curb the consequences of dropout include “last observation carried forward” and multiple imputation calculating the most likely outcome based on the data and outcome of other patients and the known data from patients with missing data. However, “last observation carried forward”, being a single imputation conferring too much weight to the imputed values and hence an overestimated precision to the effect estimate, may introduce bias and is not recommended by Cochrane Collaboration.7

If patients with missing data are mainly outliers, per-protocol analyses may increase homogeneity and precision.139 However, if losses to follow-up are related to prognostic factors, adverse events, or lack of treatment response, per-protocol analyses may overestimate the intervention effects.159 In general, the ITT analysis is the most reliable method for analyzing data from RCTs. Despite the argument that a few missing outcomes will not cause a problem, in half of trials more than 10% of randomised patients may have missing outcomes.156 Thus, complete case analysis will lead to lose of power by reducing the sample size, and bias may be introduced if being lost to follow-up is related to a patient’s response to treatment.157 Fundamentally, breaking the ITT principle may lead to violation of the effects of randomization and a non-ITT analysis can be considered just a variant of an analysis of an observational study as the balance of prognostic factors in the intervention groups may eventually be skewed.

2.3 Detection bias
Detection bias arises if the knowledge of patient assignment leads to systematic differences between groups in how outcomes are assessed.7 This is avoided by the blinding of those assessing the outcomes.146

2.4 Reporting bias
Reporting biases arise when the dissemination of research findings is influenced by the nature and direction of results. This is a due to the fact that only a proportion of research projects are ultimately published.7 Table 2 summarizes various types of reporting biases and I will briefly discuss two of these in more detail.

2.5.1 Publication bias and time lag bias
Studies with statistically significant results are more likely to be published than studies with non-significant results with significantly shorter times to publication.157 Journals are not solely to blame since publication bias may reflect a reluctance to submit reports on negative trials for publication.139 Selective or delayed publication of the findings of trials with unwanted results seems to be a widespread problem.7 Selective publication of the findings of trials with positive results, time-lag bias, and outcome-reporting bias may lead to overestimation and false-positive conclusions about treatment effects. Incompletely reported outcomes are almost twice as likely to be statistically non-significant as fully reported outcomes.139 In meta-analyses, smaller trials often show different and larger intervention effects. One possible explanation is publication bias and occurs when the chance of a smaller study being published is increased in case of a stronger effect.158 This leads to a biased interpretation of the results of meta-analyses and systematic reviews. Bias may be visualised by the application of funnel plot, which plots each trial’s effect size against a measure of its variability.159-162 Asymmetry in funnel plots is an indication of bias. Heterogeneity is an additional source of asymmetry in funnel plots and occurs when smaller trials may select patients who are more likely to benefit from the intervention. Effects like these have been referred to as small study effects.160-162

<table>
<thead>
<tr>
<th>PUBLICATION BIAS</th>
<th>The publication or non-publication of research findings, depending on the nature and direction of results</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME LAG BIAS</td>
<td>The rapid or delayed publication of research findings, depending on the nature and direction of the results</td>
</tr>
<tr>
<td>MULTIPLE (DUPLICATE) PUBLICATION BIAS</td>
<td>THE MULTIPLE OR SINGULAR PUBLICATION OF RESEARCH FINDINGS, DEPENDING ON THE NATURE AND DIRECTION OF THE RESULTS</td>
</tr>
<tr>
<td>LOCATION BIAS</td>
<td>THE PUBLICATION OF RESEARCH FINDINGS IN JOURNALS WITH DIFFERENT EASE OF ACCESS OF LEVELS OF INDEXING IN STANDARD DATABASES, DEPENDING ON THE NATURE AND DIRECTION OF RESULTS</td>
</tr>
<tr>
<td>CITATION BIAS</td>
<td>The citation or non-citation of research findings, depending on the nature and direction of the results</td>
</tr>
<tr>
<td>LANGUAGE BIAS</td>
<td>THE PUBLICATION OF RESEARCH FINDINGS IN A PARTICULAR LANGUAGE, DEPENDING ON THE NATURE AND DIRECTION OF THE RESULTS</td>
</tr>
<tr>
<td>OUTCOME REPORTING BIAS</td>
<td>THE SELECTIVE REPORTING OF SOME OUTCOMES BUT NOT OTHERS, DEPENDING ON THE NATURE AND DIRECTION OF THE RESULTS</td>
</tr>
</tbody>
</table>

Table 2. Definitions of some of the types of reporting biases according to Cochrane handbook

2.5.2 Outcome reporting bias (selective reporting bias)
Selective reporting bias is defined as the selection for publication of a subset of the original recorded outcomes based on the results. It may occur in relation to outcome analyses, subgroup analyses, and per protocol analyses, rather than ITT analyses.163
Selective reporting of outcomes may arise in several ways, affecting not only the interpretation and validity of the trial data but also the corresponding meta-analytic estimates and conclusions in systematic reviews. There is strong evidence that trials reporting positive or statistically significant results are more likely to be published, and outcomes that are statistically significant have higher odds of being fully reported. Additionally, 40-62% of trials change, introduce, or omit at least one primary outcome compared to their protocols in the final publication. The decision to omit outcomes from publications seems to be made by investigators based on a combination of journal space restrictions, the importance of the outcome, and the statistical results. Industry funded researchers appear to be less willing or able to offer data from their studies. Regardless of the funding source, authors may have a reluctance to reveal biased practices and thus be non-responsive or inaccurate in their response once approached by systematic review authors.

3. RISK OF BIAS DOMAINS FOR SYSTEMATIC ERRORS

3.1 Randomisation (sequence generation and allocation concealment)
With randomization, each patient’s treatment is assigned according to the play of chance. At the same time, randomization ensures that unknown and unmeasured differences as well as those that are known and measured are controlled. Adequate randomization involves both generation of an unpredictable allocation sequence and concealment of allocation. Sequence generation refers to a rule for allocating interventions to participants, based on some chance (random) process. Allocation concealment is defined as procedures, which ensure strict implementation of the schedule of random assignments by preventing knowledge of forthcoming allocations by study participants or by those recruiting them to the trial. Irrespective of how the allocation sequence is generated, bias may be introduced if the allocation of the next patient is known before the randomization takes place, e.g., by increasing the risk of excluding the potential participant from the trial. If properly implemented, randomization prevents selection bias in allocating interventions to participants. Additionally, random assignment permits the use of probability theory to express the likelihood that any difference in outcome between intervention groups merely reflects chance.

Finally, random allocation may facilitate blinding the identity of interventions to the investigators, participants, and evaluators, possibly by use of a placebo, thus reducing performance bias after assignment of interventions. However, the most important advantage of adequate randomisation is reduction of selection bias at trial entry. Despite the widespread acceptance of the Consolided Reporting of Trials (CONSORT) statement and the central role of allocation, reporting of the methods used for allocation of participants to interventions is very inadequate in the published papers and many are not truly randomised. On average, non-randomised trials and trials with inadequate allocation sequence generation, inadequate allocation concealment, and inadequate blinding lead to overestimation of intervention effect and may contribute to discrepancies between the results of large and small RCTs in meta-analyses. However, it is not generally possible to predict the magnitude, or even the direction, of possible selection biases and consequent distortions of intervention effects. Trials with inadequate or unclear allocation concealment result in 23-31% larger estimates of effect than trials with adequate allocation concealment and trials that are not blinded yield 25% larger estimates.

3.2. Blinding
Blinding (masking) is defined as keeping participants, health-care providers, data collectors, outcome assessors, data analysts or authors of manuscripts unaware of the assigned intervention. The purpose of blinding is to prevent bias associated with patients’ and investigators’ expectations. Blinding can be especially important for assessment of subjective outcomes, such as degree of postoperative pain. However, blinding may also be important for objective outcomes in trials where enthusiasm for participation or follow-up may be influenced by group allocation. If interventions are compared with no intervention, an identical placebo may be used which should be identical in taste, smell, appearance, and mode of administration since any difference may destroy the blinding. Nevertheless, some interventions are difficult to blind. For instance it is usually not possible to blind people to whether or not major surgery has been undertaken.

Lack of blinding may introduce bias if knowledge of intervention groups affected the care received or the assessment of outcomes. This could happen independently of a possible selection bias due to inadequate allocation concealment. Interpretations, definitions and reporting of single, double, and triple blinding varies greatly between physicians, textbooks and published papers. There is a strong association between adequate allocation concealment and blinding since about 3/4 of trials with adequate concealment are classified as blinded.

3.3 Incomplete outcome data
Incomplete (missing) outcome data in clinical trials is due to either exclusions or attrition. Exclusions refer to situations in which some participants are omitted from reports of analyses, despite outcome data being available to the trialists. Attrition refers to situations in which outcome data are not available. As previously described, incomplete reporting is a key obstacle in the assessment of risk of bias or quality and raises the possibility that effect estimates are biased. When available, published protocols are a particularly valuable source of information.

3.4 Selective outcome reporting
Selective outcome reporting is defined as the selection of subset of original variables recorded, on the basis of the results, for inclusion in publication of trials. The particular concern is that statistically non-significant results might be selectively withheld from publication. Selective outcome reporting may occur in several ways according to The Cochrane Handbook.

1. Selective omission of outcomes from reports occurs when only some of the analysed outcomes are included in the published report, for instance outcomes with statistical significance
2. Selective choice of data for an outcome may occur when there is different time points at which the outcome has been measured, or when different instruments are used to measure the outcome at the same time point (e.g. different scales, or different assessors).
3. Selective reporting of analyses using the same data refers to settings where there are different ways in which an outcome can be analysed, e.g. blood pressure reduction as a continuous or dichotomous variable, with the further possibility of selecting from multiple cut-points. Another approach would be switching from an intended comparison of final values to a comparison of changes from baseline because of observed baseline imbalances. The latter might actually introduce bias in itself rather than remove it.
4. Selective reporting of subsets of the data may occur if outcomes can be subdivided (e.g. selecting sub-scales of a full measurement scale or a subset of events).

5. Selective under-reporting of data occurs when some outcomes are reported with inadequate detail of the data to be included in a meta-analysis. For instance the authors just state "not significant" or "P>0.05" instead of the actual exact P-value.

3.5 Risk of bias from other sources

Beyond previously described domains, further issues may raise concern about the possibility of bias. This domain is defined as risk of bias from "other sources" and relate to sources of diversity (heterogeneity) or measures of research quality that are unrelated to bias. I will describe three of the most important ones.

3.5.1 Baseline imbalance

Bias can occur in the estimation of the intervention effect when there is baseline imbalance in factors that are strongly associated to outcome measures. This can happen through chance alone or through unconcealed allocation of interventions. Additionally trial authors can cause imbalance in participant characteristics in the different intervention groups when some randomised individuals are excluded from the analyses.7 Sequence generation, lack of allocation concealment or exclusion of participants should each be addressed using the specific entries for these in the tool. If further inexplicable baseline imbalance is observed after assessment of sequence generation and allocation concealment, these should be noted, since they may lead to important exaggeration of effect estimates. However, base line imbalance is just as likely to occur as result of random error in a RCT if there are multiple baseline variables. Tests of baseline imbalance are not generally recommended, but very large differences could suggest bias in the allocation of the intervention.7

3.5.2 Early stopping

Trials that are stopped earlier than planned are more likely to show extreme intervention effects than those that continue to the end, particularly if they have very few events.175 Trials can stop earlier than planned when: result of an interim-analysis shows larger than expected benefit or harm on the experimental intervention; when investigators find evidence of no important difference between experimental and control interventions (stopping for futility); or when the trial becomes unviable (e.g., lack of funding, no access to eligible patients or trial interventions, results of other trials make the research question irrelevant).175 RCTs stopped early for benefit may systematically overestimate treatment effects.7 This difference is independent of the presence of statistical stopping rules and is greatest in smaller studies.176 However, when studies are stopped early for reasons apparently independent of trial findings, or when they reach their planned termination, they are unlikely to introduce bias by stopping.146,176,177 Thus, the impact of early stopping in a study depends primarily on the definition of the stopping rules and the level of statistical significance of the interim analysis.144 Thus, it may be quite reasonable to halt a trial if a large efficacy difference is observed, acknowledging that the true difference is likely to be smaller than what is observed.7 Conversely, trial should be allowed to continue when the magnitude of benefit and perhaps of safety is essential to know with precision.178 Early stopping for efficacy, in case of proper implementation, analysis and reporting may have a relatively small effect on efficacy estimates when compared to letting a trial continue. The best protection against inappropriate stopping of a trial is an efficient and wise data monitoring committee, guided but not ruled by statistical stopping guidelines.178,180

3.5.3 Competing (vested) interests

Trials associated with industry funding generally have higher quality than trials without external funding.181 However, external funding and financial interests may bias the interpretation of trial results due to violation of the uncertainty principle, publication bias, and biased interpretation of trial results.181,182 The uncertainty principle is defined as enrolment of patients in a trial, only when there is substantial uncertainty about which of the treatments in the trial is most appropriate for the patient.183 Hence, trials may be considered unethical if patients allocated to the control group are not offered a known effective intervention. Selective sponsoring of trials with known beneficial effects may be related to violation of the uncertainty principle.181

4. RISK OF RANDOM ERROR ("PLAY OF CHANCE")

Measurement errors are composed of two components: random error and systematic error. Bias refers to systematic error, meaning that multiple replications of the same study would reach the wrong answer on average. Whereas, systematic errors are predictable, the concept of random error is closely related to the concept of precision. Imprecision refers to random error, meaning that multiple replications of the same study will produce unpredictable fluctuations in the effect estimates because of sampling variation even if they yield the right answer on average.7 The word random indicates that these errors are inherently unpredictable, and have null expected value, namely, the random errors are scattered about the true value, and tend to have null arithmetic mean when a measurement is repeated several times with the same instrument.184 All measurements are prone to random error. The results of smaller trials are subject to greater sampling variation and hence are less precise. The higher the precision of the effect estimate, the smaller the variability, reflected in the confidence interval around the intervention effect estimate from each trial and in the weight given to the results of each trial in a meta-analysis.7 Random error refers to the risk of drawing a false conclusion based on sparse data defined as either a type I error (false rejection of the null hypothesis) or type II error (false acceptance of the null hypothesis).144 Meta-analyses are typically defined as ‘positive’ or ‘negative’ based on statistical tests (test statistics), with a corresponding P-value or confidence interval. A P-value reflects the probability of obtaining a difference in outcome between two interventions, assuming that the null hypothesis is true. However, the p-values of intervention effect estimates are not suitable for comparison of the risk of random error between different trials since they do not sufficiently reflect the true risk of random error especially during accumulation of data and sequential testing.144,175 Standard error (SE) as a measure of uncertainty and the degree of variation in the population and the sample size can be applied for evaluation of risk of random error.144 In RCTs, the influence of the risk of random error appears to be much larger than previously perceived and may be one explanation for the early stopping of trials at interim analyses when benefit or harm appear to be significant.175,185,186 Additionally, random error may play a role in the repeated analyses of accumulating data in both trials and meta-analyses.145,187,188 As systematic reviews are updated with new trials (cumulative meta-analysis) as recommended by The Cochrane Collaboration, more statistical tests may be applied which increases the likelihood of observing a false positive or
false negative result. This phenomenon is commonly referred to as 'multiplicity due to repeated significance testing'. A typical scenario for multiplicity occurs when data on the primary outcome, on which the sample size calculation is based, does not show statistical significance, while another outcome measure, for which no separate sample size calculation is performed, indicates statistical significance. In meta-analysis as well as RCTs it is essential to minimize the risk of false positive or false negative conclusions. When a meta-analysis includes only a small number of trials and a small number of patients, random errors can cause spurious findings. On the contrary, in case of large number of patients, and in case of several trials confirming the findings of previous trials, test statistics and intervention effect estimates will typically be closer to the 'truth'. Risk of random error or imprecision only causes problems if statistical tests (and intervention effect estimation) are employed at stages where the magnitude of the random error or imprecision is 'extreme enough' to result in spurious statistical findings as illustrated in figure 4.

Figure 4. Examples of false positive and false negative statistical test results over time in two different randomised clinical trials A and B. Figure 1A: Significance testing at X1 and X2 would have resulted in a type I error (false positive) while X3 and X4 would have resulted in a type II error (false negative). X5 and X6 would not.

Systematic reviews with meta-analyses of several randomised clinical trials are considered to be the best available evidence. However, 'the best available evidence' may not be synonymous with 'sufficient evidence' or 'strong evidence' or 'even the best obtainable evidence'. About 25% of conventional meta-analyses with small number of events and patients, may falsely estimate a statistically significant intervention effect. Empirical evidence also indicate that as more evidence is accumulated, large pooled intervention effects observed in early positive meta-analyses tend to disappear. As in RCTs, repeated significance testing on accumulating data (updating the meta-analyses) is known to increase the overall risk of type I error. To avoid random errors, a meta-analysis should include a sample size at least as large as that of an adequately powered single RCT. Additionally, in a meta-analysis, there is likely heterogeneity across included trial populations, interventions and methods. Thus, the meta-analysis sample size considerations should be adjusted (increased) in order to allow for the heterogeneity induced variance. The strength of evidence should be measured based on the accrued number of patients, observed number of events in the included trials, and the impact of multiplicity.

Trial sequential analysis (TSA) is an approach with the potential to provide the required meta-analytic sample size, that is the required information size for a meta-analysis, and minimise the risk of type I error. Meta-analyses not reaching the TSA calculated sample size are analysed with trial sequential monitoring boundaries. This is a more restrictive analysis, analogous to an interim monitoring boundary in a single trial. Trial sequential monitoring boundaries protect against the increased risk of random error due to repeated and early significance testing, and adjust the required significance level for obtaining statistical significance according to the number of events and participants in a meta-analysis. The fewer events and participants, the more restrictive the monitoring boundaries are, and a lower P value is required to obtain statistical significance (more detailed description of TSA in Appendix II).

5. EXTERNAL VALIDITY: THE RISK OF DESIGN ERROR ('WRONG DESIGN TO ANSWER THE POSED QUESTION')

External validity is defined as the extent to which results of trials provide a correct basis for generalisation and application to other circumstances. External validity describes whether the trial is asking an appropriate research question with regard to specified external conditions, such as patient populations (e.g., age, sex, severity of disease and risk factors, co-morbidity); intervention regimens (e.g., dosage, timing, route of administration, type of intervention within a class of interventions/treatments, concomitant treatments); settings (e.g., level of care, experience and specialisation of care providers) and finally modalities of outcomes (e.g., type or definitions of outcomes and duration of follow-up). In case of sufficient internal validity, i.e., low risks of systematic errors and random errors, it becomes relevant to consider the risks of design errors (external validity) in answering a posed question. Among the many potential variables to be considered, outcome measures have a major relevance and play a central role for clinical research. Outcome measures can be divided into three categories according to the GRADE classifications (Figure 5). Primary outcome measures are central in deciding the use of one intervention over another. Large estimate differences in the primary outcome measure between groups in a clinical trial may result in early stopping. Primary outcome measures should be chosen according to GRADE classification of “critical for decision-making”.

Secondary outcome measures are supplementary outcomes and often surrogates. If the effect estimate of the secondary outcomes is positively influenced by an intervention, the intervention may only be recommended if it provides a beneficial effect on the primary outcome or if no clinically relevant and statistically significant effect exists on the primary outcomes. The secondary outcomes should be chosen according to the second and third GRADE categories of important, but not critical for decision making and “not important for decision making”. The GRADE approach provides a nine point scale to judge importance with the upper ratings of 7 to 9 identifying outcomes of critical importance for decision making, the middle ratings of 4 to 6 representing important but not critical outcomes and the lower ratings of 1 to 3 defined as items of limited importance. Depending on the outcomes, this scale should sometimes be considered nominal and in other situations considered functional. GRADE provides a definition for the quality of evidence that reflects the extent to which confidence in an estimate of the effect is adequate to support recommendations. Since systematic reviews do not—or preferably should not—make recommendations, the quality of evidence reflects the extent of confidence that an estimate of effect is correct. The GRADE approach involves making separate ratings for quality of evidence according to patients’ perspective for each important outcome and identifies five factors that can lower the quality of the evidence. These factors can downgrade the quality of the studies and are defined as:
1) Study limitations (design): inadequate randomisation, lack of blinding, large losses to follow-up; lack of intention-to-treat; stopping early for benefit; outcome reporting bias.

2) Inconsistency (heterogeneity or variability in results): Widely differing estimates of intervention effect across studies indicate true differences in underlying treatment effect. Variability may arise from differences in populations (e.g. drugs may have different effect estimates in sicker populations), interventions (e.g., larger effect estimates with higher drug doses), or outcomes (e.g. diminishing treatment effect with time). Heterogeneity without a plausible explanation decreases the quality of evidence.

3) Indirectness: First type of indirectness of evidence refers to indirect comparisons of the magnitude of effect estimates of various interventions that are not necessarily comparable. The second type includes differences between the population, intervention, comparator to the intervention, and outcome of interest, and those included in the relevant studies.

4) Imprecision: Small studies with few patients and few events with wide confidence intervals.

5) Publication bias.

Grading of outcome measures may vary according to the clinical question with the possibility of outcomes becoming interchangeable within a category (i.e., critical, important, or not important). However, it is hard to argue that outcomes between categories (i.e., critical, important, or not important) are interchangeable (e.g., mortality is always more important than length of stay in hospital).^144^ Figure 5. Hierarchy of outcomes according to importance to patients, GRADE classification

6. QUALITY OF REPORTS OF RCTS

CONSORT statement was first published in 1996 and then revised in 2001 with latest revision in 2010.\[^{137,147,202}\] It provides recommendations for authors regarding how to prepare reports of trial findings. The CONSORT checklist contains details such as sample size calculations, primary outcomes, random sequence generation, allocation concealment, and handling of attrition.\[^{203}\] Despite improvement in the reporting of several important aspects of trial methods since the introduction of CONSORT statement, there are still inadequate reporting in more than half of the publications assessed.\[^{204,205}\] Additionally, reporting is not only often incomplete but also sometimes inaccurate, more commonly in the specialty journals and journals published in languages other than English.\[^{138,143,170,204,206}\] A significant association between inadequate or unclear bias protection and overestimation of beneficial effects and underreporting of adverse effects is still present in RCTs.\[^{7,139,138,154}\] There is still ample room for improvement in reporting and conduct of trials and meta-analyses and I will discuss different aspects and the importance of this issue further in the discussion section.

7. QUALITY OF REPORTS OF SYSTEMATIC REVIEWS

Recent estimates indicate that at least 2,500 new systematic reviews reported in English are indexed in MEDLINE annually.\[^{216}\] However, there is considerable evidence that key information is often poorly reported in systematic reviews, thus diminishing their potential validity and usefulness.\[^{207,212}\] Systematic reviewers sometimes draw too optimistic conclusions or do not consider the harms equally and as carefully as the benefits.\[^{214}\] Systematic reviews should be subject to full and transparent reporting to allow readers to assess the strengths and weaknesses of the systematic reviews as well as the trials they include.\[^{215}\] Poor reporting of systematic reviews diminishes their value to clinicians, policy makers, and other users. Since the publication of the QUOROM statement (QUality Of Reporting Of Meta-analysis) as a reporting guideline for authors of meta-analyses of RCTs in 1999, there have been several conceptual, methodological, and practical advances regarding the conduct and reporting of systematic reviews and meta-analyses.\[^{216}\] However, despite these advances, the quality of the conduct and reporting of systematic reviews remains far from adequate.\[^{207-213}\] As a direct consequence, the PRISMA statement (Preferred Reporting Items for Systematic reviews and Meta-Analyses) was conceived as an update of QUOROM statement in order to ensure clear presentation of what was planned, done, and found in a systematic review.\[^{208}\] The PRISMA statement consists of a 27-item checklist and a four-phase flow diagram based on empirical evidence. It focuses on ways in which authors can ensure the transparent and complete reporting of systematic reviews since recent data indicate that this reporting guidance is much needed.\[^{207}\]

The number of Cochrane reviews published each year has become relatively stable over the last 7 years.\[^{208}\] However, the number of non-Cochrane systematic reviews is increasing at a rapid rate and in 2008 accounted for more than 75% of the systematic reviews indexed in MEDLINE (Figure 6).\[^{208}\] This relative increase in non-Cochrane systematic reviews may be related to the strict methodological requirements of the Cochrane Collaboration in order to increase the quality of the systematic reviews. In principal, this increasing number of systematic reviews should improve the basis for clinical decision making as systematic reviews are considered essential sources of evidence for clinical practice and guideline development.\[^{217}\] Unfortunately, empirical evidence suggests that the methodological quality of the majority of the published systematic reviews in various specialties seem to have serious methodological flaws leading to a high risk of bias and thereby erroneous conclusions and raising concerns about validity.\[^{207,209-213}\]

While Cochrane systematic reviews on average appear to be of higher methodological quality than systematic reviews in regular journals, some of them also have methodological shortcomings.\[^{212}\] Cochrane Collaboration is the largest provider of systematic reviews with currently more than 4,500 systematic reviews included in the Cochrane Library. Cochrane systematic reviews use a more rigorous methodology and have a higher methodological quality, on average, than systematic reviews published in regular journals.\[^{212,218-221}\] The lower methodological quality of the so-called systematic reviews in regular journals could be...
explained by insufficient reporting of the used methods due to space restrictions in regular journals.\textsuperscript{224}

However, comparisons of Cochrane systematic reviews published in regular journals with Cochrane systematic reviews published in the Cochrane Library have shown no true differences in terms of methodological quality\textsuperscript{219,220} The shortcomings of systematic reviews are substantial. For instance, a recent papers reported that in more than one third of systematic reviews in regular journals, searching for trials was limited to MEDLINE, which is incomprehensive and may fail to identify all relevant trials.\textsuperscript{221} Also, the identified studies may not be representative of all relevant trials, which can lead to possibly false conclusions based on various types of bias such as publication bias (only published trials included) or language bias (only trials reports written in English) as previously discussed.

A challenging group among these publications is the industry-supported meta-analyses, which appear to lack scientific rigour, are less transparent, have more biased conclusions and have poorer methodological quality on average than meta-analyses based on various types of bias such as publication bias (only published trials included) or language bias (only trials reports written in English) as previously discussed.

Aromatase inhibitors are indicated to promote quality standards for performing systematic reviews among the authors, editors and readers of the literature. Additionally, transparency is essential for readers to make their own judgment about medical interventions guided by the results of meta-analyses, which I will discuss further in the discussion section.

**OBJECTIVES**

The objectives were to assess the beneficial and harmful effects of the following interventions by performing systematic reviews and meta-analyses, if appropriate on:

1) **Antithrombin III (AT III) in critically ill patients**
2) **Inhaled nitric oxide for acute respiratory distress syndrome (ARDS) or acute lung injury (ALI)**
3) **Aerosolized prostacyclin for ARDS or ALI**
4) **Thrombelastography (TEG) or thromboelastometry (ROTEM) to monitor haemotherapy versus usual care in patients with massive transfusion requirement**

In addition, I planned to assess the available evidence for activated protein C, steroids and tight glycaemic control for critically ill patients in this dissertation.

**SUMMARY OF FINDINGS**

In this section, I will briefly highlight the main findings of my published papers that are included in this dissertation. Further, I will discuss in more details the evidence of some additional immuno-modulatory interventions. Since this field is very broad and due to space limitations, I have only chosen three of the most widely discussed and controversial interventions: glucocorticosteroids for sepsis and septic shock, tight glycaemic control for critically ill patients, and activated protein C for sepsis. I will discuss the rationale and the controversies surrounding these three interventions and the level of evidence. In case of my own papers, I will kindly refer the readers to my own publications and appendices for more in-depth detail regarding the background, methodological aspects, results and analyses. In general, this chapter has been divided in to four different subsections:

1) **Anti-inflammatory and anti-coagulant strategies in sepsis: AT III and activated protein C**
2) **Anti-inflammatory and anti-coagulant strategies in ARDS or ALI: inhaled nitric oxide and inhaled prostacyclins**
3) **Endocrine-mediated systemic immune suppression: glucocorticosteroids and tight glycaemic control**
4) **The impact of point of care assays on transfusion requirements, and their potential influence on coagulation and inflammation: TEG or ROTEM**

**1. ANTI-INFLAMMATORY AND ANTI-COAGULANT STRATEGIES IN SEPSIS**

**1.1 Antithrombin III (AT III) in critically ill patients**

In this systematic review, we included 20 randomized trials with 3458 patients. Patients had sepsis (n=13) or were from paediatric (n=3), obstetric (n=2), and trauma (n=2) specialties.\textsuperscript{222-228} Combination of data from all 20 trials showed no significant effect of AT III on mortality, with 667 (39.1%) deaths in the intervention group compared with 699 (39.9%) in the control group (RR 0.96, 95% CI 0.89 to 1.03, I\textsuperscript{2}=0.3%, (Figure 7). AT III significantly increased the risk of bleeding events (1.52, 1.30 to 1.78, I\textsuperscript{2}=0.3%). A total of 32 subgroup and sensitivity analyses were carried out. We found no significant effect on mortality or other outcome measures (P>0.05) in all analyses of different subgroup populations (sepsis, paediatrics, obstetrics, and trauma). Additionally, subgroup analyses of trials with short and long duration of treatment, short or long follow-up, and high or low risk of bias showed no significant effect on the examined outcome measures. In the subgroup of patients who received AT III without adjuvant heparin, Anti-thrombin III was associated with a significant effect (0.87, 0.75 to 0.99). However, when we used a random effects model in the meta-analysis, the statistically significant effect was no longer present (0.87, 0.77 to 1.02). This might or might not support the previously generated hypothesis that AT III is beneficial in patients who do not receive adjuvant heparin.\textsuperscript{229} However, the use of heparin was not a stratification variable in the dominant trial which was split in this analysis conferring substantial uncertainty to the result from the fixed-effect model. AT III showed no significant effect in patients with adjuvant heparin (0.99, 0.90 to 1.09). Compared with no intervention or placebo, the analyses did not demonstrate that AT III had a statistically significant influence on
the proportion of respiratory failure, duration of mechanical ventilation, need for surgical intervention, and length of stay in hospital or in intensive care unit (ICU), or quality of life. In summary, treatment of critically ill patients with AT III does not significantly affect mortality and length of stay in hospital or in ICU, but is associated with a significantly increased risk of bleeding events. In patients who do not receive heparin there might be a benefit, though this should be explored in further randomised trials. Its use in critically ill patients cannot be recommended based on the available evidence even in patients without adjuvant heparin, but it may be relevant to explore this further in future trials.

In order to demonstrate or reject a beneficial effect on mortality in a single trial, assuming a RRR of 5% (an absolute risk reduction of 2.3%, from 48.5% to 46.2%) at least 14,294 patients should be randomized. As 3458 patients have already been included in the meta-analysis, without becoming statistically significant, a RRR of 10% or more on mortality is unlikely (Figure 1 in appendix IV). To my knowledge noRCT has been carried out on AT III in sepsis or other critically ill patients as defined in our systematic review.

1.2 Activated protein C (APC)

APC is an endogenous protein, which acts as an important modulator of the inflammatory and coagulatory responses associated with severe sepsis (Figure 8). In severe sepsis, low baseline levels of circulating protein C and early changes in protein C levels are associated with increased morbidity and mortality.230,231 APC exerts its anti-inflammatory effects by inhibiting the formation of TNF, IL-6, and IL-8 and by inhibiting neutrophil chemotaxis. APC also attenuates inflammation induced by thrombin receptor-mediated platelet activation, and nuclear factor-kB activation in blood and endothelial cells. Additionally, APC promotes fibrinolysis by binding and inhibiting plasminogen activator inhibitor-1, a potent antifibrinolytic factor. APC is also believed to posses antiapoptotic properties that may be neuroprotective.14

In 2001, the PROWESS trial of recombinant human APC (rhAPC) for severe sepsis, also referred to as drotrecogin alfa activated (DAA), reported a 6.1% absolute reduction in mortality and a 19.4% relative risk reduction of 28 days mortality.15 Given the high cost of this intervention (about £7000 per course of treatment) and the potentially large eligible patient group, post-hoc subgroup analyses of the PROWESS dataset was carried out in order to target the subgroups of patients who may most benefit from DAA. In November 2001, the US Food and Drug Administration (FDA) approved the use of DAA for septic patients with an APACHE II score ≥25. In 2002, the European Medicines Evaluation Agency (Formally known as EMEA, now EMA) approved its use for those with multiple organ failure. Subsequently FDA mandated further trials to evaluate the efficacy and safety of DAA in other patient groups such as those with severe sepsis and a low risk of death, paediatric populations, and an investigation of medium term survival of those enrolled in the PROWESS trial. The subsequent trial, ADDRESS which included severe septic patients with a low risk of death, as defined by an APACHE II score <25 or single organ failure was terminated prematurely due to low likelihood of efficacy.23 The following trial, RESOLVE conducted in paediatric setting failed to demonstrate efficacy of DAA in children with severe sepsis while there was more instances of central nervous system bleeding in the DAA group, particularly in children less than 60 days old.236

Figure 7. Forest plot of mortality, antithrombin III (AT III) vs control (subgroup analyses on risk of bias)

In order to demonstrate or reject a beneficial effect on mortality in a single trial, assuming a RRR of 5% (an absolute risk reduction of 2.3%, from 48.5% to 46.2%) at least 14,294 patients should be randomized (with 80% power and alpha 0.05, assuming a double-sided type I risk of 5% and a type II risk of 20%). However, solid evidence may be obtained with a lesser number of patients depending on the findings of future trials and their overall quality. On the other hand to demonstrate or reject an a priori anticipated intervention effect of a RRR of 10%, 3317 patients should be randomized. As 3458 patients have already been included in the meta-analysis, without becoming statistically significant, a RRR of 10% or more on mortality is unlikely (Figure 1 in appendix IV). To my knowledge no RCT has been carried out on AT III among critically ill since the publication of this paper and there are currently no registered RCT on AT III in sepsis or other critically ill patients as defined in our systematic review.


An additional trial failed to show any effect of DAA on long-term survival among the same population as those originally enrolled in the PROWESS trial. A subsequent post-hoc analysis indicated that patients with an APACHE II score ≥25 who were treated with DAA had a better overall survival at 3 months, 6 months, 1 year and 2.5 years.237 For patients treated with DAA and with an APACHE II score <25, there was a significant decrease in survival at 1 year (p=0.04), but not at other time-points.237 However, subsequent to the PROWESS trial, there has been a continuous debate over the efficacy and safety of this expensive treatment, since none of the other RCTs have been able to reproduce the same efficacy as the PROWESS trial. There has also been concern over protocol amendments for the original PROWESS trial, since an external evaluation committee advised a protocol change after inclusion of 720 patients since the committee considered a large number of recruited patients at high risk of death from causes other than sepsis. This led to exclusion of these patients. Subsequently, there was a marked reduction in mortality associated with treatment with DAA in this analysis potentially violating the ITT-principle. In 2009, a Cochrane systematic review based on the
above trials concluded that there was no evidence to support the use of DAA for treating patients with severe sepsis or septic shock.\textsuperscript{338} Additionally, the authors found an increased risk of bleeding and no association between the effectiveness of APC and the degree of severity of sepsis based on APACHE score.

### Table 3. Changes in 28-days mortality in PROWESS study before and after the protocol amendment

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Number</th>
<th>Placebo Group Mortality</th>
<th>Treatment Group Mortality</th>
<th>Relative Risk of Death</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Amendment</td>
<td>720</td>
<td>30</td>
<td>28</td>
<td>0.94</td>
<td>0.57</td>
</tr>
<tr>
<td>After Amendment</td>
<td>970</td>
<td>31</td>
<td>22</td>
<td>0.71</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Since the publication of this systematic review, results of an additional RCT on the extended use (96-h infusion) of DAA in patients with severe sepsis and vasopressor-dependent hypotension has been published.\textsuperscript{339} This trial included 193 participants, but DAA failed to reduce 28-day or in-hospital all-cause mortality, improve organ failure or result in a more rapid resolution of vasopressor-dependent hypotension compared with placebo. By updating the existing systematic review on the use of APC in septic setting, I found no statistically beneficial effect on 28-day all cause mortality (RR 0.97, 95% CI 0.91 to 1.22, \( I^2 = 58\%\), random effects model) or on mortality at the longest follow-up (RR 0.99, 95% CI 0.88 to 1.12, \( I^2 = 4\%\), Appendix V, figure 1&2). However, APC appears to significantly increase the risk of bleeding (RR 1.54, 95% CI 1.14 to 2.08, \( I^2 = 0\%\), Appendix V, figure 3)

There are two additional published RCTs on the use of DAA. Liu et al examined the efficacy and safety of APC among patients with ALI.\textsuperscript{240} This trial included 75 participants with ALI but failed to show any statistically significant difference in 60-days mortality, ventilator-free days or number of organ failure-free days between the two groups. XPRESS trial (Eli Lilly sponsored) examined the effects of concomitant administration of DAA and Heparin vs. DAA and placebo among 1935 patients with severe sepsis.\textsuperscript{241} All patients received DAA according to local hospital guidelines. Heparin co-administration was found to be associated with a non-significant reduction in 28-day mortality in patients randomized to treatment with heparin. The authors found no negative co-interaction between DAA and Heparin.

APC treatment remains approved only for the small subset of septic patients with a low risk for haemorrhage. This is a critical limitation in the therapeutic use of APC, considering that most sepsis patients are already at an increased risk of internal haemorrhage. Furthermore, data from recent non-industry-supported observational studies conclude that due to continual bleeding problems, for most patients, the risks associated with APC intervention seem to outweigh the benefits.\textsuperscript{238,242-244} There are currently four registered ongoing RCTs that hopefully will shed more light on this controversy (Table 1 in Appendix VI).

### 2. ANTI-INFLAMMATORY AND ANTI-COAGULANT STRATEGIES FOR ARDS OR ALI

#### 2.1 Inhaled nitric oxide for ARDS or ALI in children and adults

In this systematic review of 14 RCTs with 1303 participants with ARDS or ALI, 10 trials were of high risk of bias.\textsuperscript{245-248}INO showed no statistically significant effect on overall mortality [265/660 deaths (40.2%) in the INO group vs. 228/590 deaths (38.6%) in the control group (RR 1.06, 95% CI 0.93 to 1.22; \( I^2 = 0\%\)). The analysis on mortality showed absence of heterogeneity and robust results when performed on different subgroup and with various sensitivity analyses (For TSA on mortality, see figure 2 in Appendix IV). The 28-day mortality analysis showed 36% (208/578) deaths in the INO group and 32.7% deaths (165/504) in the control group (RR 1.12, 95% CI 0.95 to 1.31; \( I^2 = 0\%\)). We found a statistically significant but transient improvement in oxygenation in the first 24 hours, expressed as the ratio of partial pressure of oxygen to fraction of inspired oxygen (MD 15.91, 95% CI 8.25 to 23.56; \( I^2 = 25\%\)). This was confirmed by application of TSA (Figure 3 in Appendix IV). The oxygenation index was significantly lower in the INO group at 24, 72 and 96 hours but not at the 48 hours analysis. Additionally, the rate of severe respiratory failure decreased in the INO group (RR 0.21, 95% CI 0.05 to 0.79; \( I^2 = 0\%\)). Differences in mean pulmonary arterial pressure was initially significant at day one but no longer present on days two, three or four. Limited data demonstrated a statistically insignificant effect of INO on duration of ventilation, ventilator-free days, and length of stay in the intensive care unit and hospital. We did not find any statistically significant difference when examining the effects in subgroups according to duration of intervention; intervention among different populations (paediatrics, adults) and sensitivity analysis excluding trials only published as abstracts. The three paediatric trials with 162 patients were insufficient to demonstrate any benefits or harms of INO therapy in paediatric ARDS or ALI.\textsuperscript{247-249} Subgroup and sensitivity analyses assessing the impact of varied primary aetologies, reversal of ALI resolution of multi-organ failure, quality of life assessment and bias assessment did not result in statistically significant findings. However, INO appears to increase the risk of renal impairment among adults (Figure 9) but not the risk of bleeding or methaemoglobin or nitrogen dioxide formation.

Despite evidence of an initial but transient improved oxygenation in the INO group, these analyses were limited due to application of different indicators of oxygenation, different time points for oxygenation measurement, and demonstration of therapeutic effect in graphic form without adjacent numerical data in most publications, thus preventing adequate pooling of data. Even though a beneficial effect is true, oxygenation is only a non-validated potential surrogate outcome and it is uncertain whether it predicts any clinical benefits. Additionally, many trials were conducted before the general recommendation of the lung protective, low tidal volume ventilation strategy and application of high PEEP among ARDS patients.\textsuperscript{250,251} The latter combined with oxygen toxicity, surfactant inhibition and ongoing fibrosis as result of ARDS may have biased the results of these trials. The amount of used sedatives and muscle relaxants and the use of protocolized weaning could also potentially play a role.

However, since there was no difference in the mode of ventilation and overall treatment between the INO and control groups, this should not account for our findings of lack of benefit on survival and in potential harm. Improved oxygenation is not associated with increased survival since improved oxygenation does not necessarily indicate improved lung function, reduction of lung injury or resolution of the underlying cause of ARDS and the often co-existing multi-organ failure.\textsuperscript{252,253} Furthermore, NO is an important regulator of renal vascular tone and a modulator of glomerular function. Changes in NO production could potentially contribute to acute renal failure by altering the function of mitochondria, various enzymes, deoxyribonucleic acid and membranes.\textsuperscript{254} There is a need for large randomised trials with low-risk of bias with a sample size of more than 4000 participants to evaluate INO for adults and children before this intervention definitely can be
have a dual mechanism of action as they block the production of inflammatory cytokines and mediators such as TNF, IL-1, IL-6, IL-8, interferon-γ, and NO from a variety of cell types, such as leukocytes, endothelium, and epithelium, while at the same time enhancing the production of the anti-inflammatory cytokine IL-10 from T-cells. 253 The use of glucocorticoids in sepsis has remained controversial since its introduction. 254

The logic behind this therapy is based on the assumption that sepsis is complicated by relative adrenal insufficiency and peripheral resistance, which occurs among 50% leading to, impaired corticosteroid production and a reduced response to corticotrophin, thus making the patient susceptible to shock. 255 However, prolonged glucocorticoid therapy can have serious adverse effects such as skin atrophy, wound healing disorders, osteoporosis, glaucoma, diabetes mellitus (hyperglycaemia), adrenal insufficiency, hypertension and immunodeficiency caused by excessive lymphocyte apoptosis. 256

The latter may cause T- and B-cell deficiencies leading to hyperimmunosuppression and thus an inability to clear an infection and a subsequent increase in the occurrence of secondary infections. 260 Additionally, glucocorticoids may lead to excessive apoptotic loss of intestinal epithelium, exacerbating inflammation-induced intestinal epithelial barrier dysfunction and facilitating the dissemination of bacteria from the gut lumen into the systemic circulation. These side effects may contribute to morbidity and mortality of sepsis. 259 In 2004, Annane et al published a Cochrane systematic review indicating a reduction of 28-day overall mortality with low dose corticosteroid (five trials, n = 465, RR; 0.80, 95% CI 0.67 to 0.95), which has since been the rationale and the accepted evidence for corticosteroid therapy in sepsis. 262,263

In 2008 the results of the CORTICUS trial added more to this controversy. 264 In this multi-centre RCT, the patients were randomised to either intravenous hydrocortisone 50 mg or placebo every six hours for five days and then tapered during a six-day period. However, there was no difference in 28-day mortality between the two study groups independent of response to corticotropin stimulation. Overall 28-day mortality was 34.3% (95% CI 28.3 to 40.2) in the hydrocortisone group and 31.5% (95% CI 25.6 to 37.3; P = 0.51) in placebo group. Death during hospitalisation was more common in the intervention group (44.2 % vs. 40.8 %, P=0.47). Additionally, there were more episodes of superinfection including new episodes of sepsis and septic shock in the intervention group. The CORTICUS trial was stopped prematurely due to slow recruitment and expiry of the supply of study drug. Additionally, detractors stressed on several limitations of this trial such as high rate of inappropriate antimicrobial therapy, low mortality rate in the control group, use of etomidate and late use of hydrocortisone administration.

Subsequent to this publication, in 2009 Annane et al published an updated version of their Cochrane systematic review in JAMA. 258 The authors identified 17 RCTs (n=2138) and 3 quasi-randomized trials (n=246). The 28-day overall mortality in the RCTs was 35.3% in the intervention group vs. 38.5% in the control group (RR random effects model) 0.83). However, the subgroup analysis of prolonged low-dose corticosteroid treatment resulted in a statistically significant 28-day mortality reduction [12 trials, (236/629 [37.5%] vs. 264/599 [44%]); (RR 0.84; 95% CI, 0.72-0.97; P=0.01)] Additionally, corticosteroid treatment increased 28-day shock reversal (6 trials; 322/481 [66.9%] vs. 276/471 [58.6%]; RR, 1.12; 95% CI, 1.02-1.23; P=0.02; I²=4%) and reduced ICU length of stay by 4.49 days (8

### 2.2 Aerosolized prostacyclin for ARDS or ALI

In this systematic review, we were only able to include one trial with 14 critically ill children with ARDS or ALI that assessed the effect of aerosolized prostacyclin. 253,254 This is insufficient to demonstrate any benefits or harms of aerosolized prostacyclin therapy. We found two ongoing trials (NCT00314548, NCT00981591) but were unable to retrieve any data from the authors of these studies since they were not at a stage where they could disclose data. Based on the very limited data available, we were unable to show any benefits of aerosolized prostacyclin on survival or other clinical outcomes. The sparse data on mortality is not promising but is not evidence of the absence of a beneficial effect; nor do the data suggest the degree of a potentially beneficial or detrimental effect of inhaled prostacyclin. Despite signs of improved oxygenation, there is no statistically significant effect on mortality or other clinical outcomes based on this small paediatric RCT. 253

There is a need for large randomised trials with low-risk of bias and a required information size of up to several thousand participants (children as well as adults) to evaluate aerosolized prostacyclin before this intervention can be definitely rejected or accepted for use in critically ill patients with ARDS or ALI. However, more light will be shed on this matter when the data from the two ongoing RCTs are published enabling us to evaluate more evidence and providing more information regarding the need for future large RCTs with low-risk of bias.

### 3. ENDOCRINE-MEDIATED SYSTEMIC IMMUNE SUPPRESSION

#### 3.1 Glucocorticoids

Glucocorticoids are a specific class of corticosteroid hormones with potent immunosuppressive properties. 265 During infection or injury, elevated serum levels of pro-inflammatory cytokines activate the hypothalamic–pituitary–adrenal axis, leading to the release of adrenocorticotropic hormone (ACTH) from the pituitary gland into the circulation. 256 ACTH affects to the adrenal cortex, leading to the release of cortisol, which functions as a potent systemic immunosuppressant. 257 Glucocorticoids like cortisol

### Table: Randomized Clinical Trials

<table>
<thead>
<tr>
<th>Study in Subgroup</th>
<th>Events</th>
<th>Total Events</th>
<th>Total Weight</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucocorticoids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticosteroids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 9.** Forest plot of the effect of inhaled nitric oxide (INO) on renal function suggested by the randomised controlled trials, subgroup analysis based on the overall quality of the included trials.
the risk of gastroduodenal bleeding (13 trials; 65/800 [8.1%] vs. 363/703 [51.6%] vs. 308/670 [46%]; P<.001; I^2=0%). However, corticosteroids did increase the risk of hyperglycaemia (9 trials; 363/703 [51.6%] vs. 308/670 [46%]; P<.001; I^2=0%).

In 2009, Sligl et al published an additional systematic review on safety and efficacy of corticosteroids for the treatment of septic shock. The authors included 6 RCTs, one case-control study and one retrospective study. This systematic review does not adhere to Cochrane methodology and has several methodological shortcomings. The authors concluded that among patients with septic shock, corticosteroid therapy appears to be safe but does not reduce 28-day all-cause mortality rates despite significantly reducing the incidence of vasopressor-dependent shock, thus similar to Annane’s findings in his JAMA publication. Only one additional RCT (COHITS study) has been published since Annane’s latest systematic review. In this trial, 509 adults with septic shock were randomized to 4 groups examining the role of tight glycaemic control and glucocorticosteroids (appendix VI, table 3). There was no significant difference at hospital discharge in the overall survival between the various groups.

Current recommendations from Surviving Sepsis Campaign are to use low dose hydrocortisone therapy only in adult septic shock patients who are poorly responsive to resuscitation and vasopressor therapy. However, the controversy surrounding this topic is far from over since the scientific evidence to date is contradictory and underpowered for a balanced conclusion. This is clearly illustrated by the number of currently registered trials on the use of corticosteroids among septic population (Table 2 in appendix VI). Even more striking is the fact that the required sample size to resolve this controversy reliably may be more than 20,000 participants in order to detect or reject an intervention effect of 5% relative risk reduction (RRR) with a type I error risk of 5% and a type II error risk of 20%. Ongoing studies will hopefully confirm whether or not low dose hydrocortisone should be used in severe sepsis.

3.2 TIGHT GLYCAEMIC CONTROL (INTENSIVE INSULIN THERAPY)
Hyperglycaemia and insulin resistance is often present among critically ill patients and is almost universally seen in sepsis. This is believed to contribute to coagulopathy, induce apoptosis, impair neutrophil function, impair wound healing and increase the risk of infection and death. Intensive insulin therapy (IIT), is often defined as maintenance of blood glucose between 80 and 110mg/dl (4.4–6.1 mmol/l). By counteracting the pathological effects of glucotoxicity, insulin may have anti-inflammatory, anticoagulant, and antiapoptotic properties. In addition to the establishment and maintenance of normoglycaemia, insulin therapy may also have direct anti-inflammatory effects in critically ill patients. The latter is illustrated by lower serum C-reactive protein (CRP) and mannose-binding lectin (MBL) levels. Insulin appears to lower serum inflammatory cytokine levels, alters T-cell polarization and may have direct antiapoptotic properties independent of glucose uptake. Additionally, insulin’s intracellular signalling appears to play an important role in myocardial protection. Furthermore, IIT may prevent endothelial cell dysfunction and systemic hypercoagulation, thus preserving organ perfusion. The insulin receptor is expressed on resting neutrophils, monocytes, and B–cells, and stimulation of human mononuclear cells with insulin significantly suppresses the activation of NF-κB, which is likely to be of major importance since the activity of this transcription factor is essential to immune cell activation and the execution of inflammatory responses.

In 2004 the Surviving Sepsis Campaign incorporated into its guideline the need for tight glycaemic control amongst critically ill patients based on the publication by van Den Berge et al. This trial, conducted in a surgical ICU setting, reduced the risk of in-hospital mortality by one-third. It has since been endorsed by various professional societies and persists in the 2008 update of Surviving Sepsis Campaign guidelines. In 2008, Wiener et al published a systematic review in which the authors identified 29 RCTs (n=8432) who were randomly assigned to receive either tight or standard glucose control. They found no overall difference in hospital mortality between tight glycaemic control and usual care (27 trials, n=8315, 21.6% vs. 23.3%; RR 0.93; 95% CI 0.85 to 1.03). They conducted subgroup analyses based on glucose goal defined as very tight (≤ 6 mmol/l) or moderately tight (>8.3 mmol/l) and based on the ICU setting (surgical, medical or medical-surgical). They found no significant difference in mortality when stratified by glucose goal (very tight: 23% vs. 25.2%, RR 0.90, 95% CI, 0.77 to 1.04; moderately tight: 17.3% vs. 18.0%; RR 0.99; 95% CI 0.83 to 1.18). Furthermore, there was no difference in mortality based on the ICU setting, and no reduction in need for dialysis. The risk of hypoglycaemia increased almost 5-fold regardless of the ICU setting and was more commonly observed amongst patients receiving very tight glycaemic control (glucose ≤ 2.2 mmol/l; 15 trials, n=6613, 13.7% vs. 2.5%; RR 5.13; 95% CI, 4.09 to 6.43). In surgical ICU setting, tight glycaemic control was found to decrease the risk of septicemia compared with usual care (nine trials, n=3916, 10.9% vs. 13.4%; RR 0.76; 95% CI 0.59 to 0.97).

In 2009, the result of much anticipated NICE-SUGAR trial was published with 6104 included patients. Regarding 90 days all cause mortality, the authors found a statistically significant increase in the intervention group (27.5% (IIT) vs. 24.9%; OR 1.14; 95% CI 1.02 to 1.28; P=0.02). They found no significant difference in mortality between surgical or medical patients. Severe hyperglycaemia defined as blood glucose level ≥40 mg/dl (2.2 mmol/l) was reported in 6.8% in the intervention group vs. 0.5% in the conventional group (P<0.001). There was no significant difference between the two intervention groups in the median number of days in the ICU or hospital or the median number of days of mechanical ventilation or renal-replacement therapy. On the same day of this paper’s publication, Griesdale et al published an updated systematic review including data from NICE-SUGAR and two additional RCTs. However, despite the inclusion of the latter 3 RCTs, this paper only included 26 RCTs, did not adhere to Cochrane methodology and had serious shortcomings in terms of bias assessment and inclusion of trials. Nevertheless, this publication reinforced the general findings of Wiener et al’s. There was no significant difference in mortality with IIT compared with conventional therapy (RR 0.93; 95% CI 0.83 to 1.04). There was a 6-fold increased risk of severe hypoglycaemia by IIT (14 trials; m=12,347, 10.7% vs. 1.6%; RR 5.99, 95% CI 4.47 to 8.03, I^2=57%). Interestingly, the authors reported that the ICU setting was a contributing factor, with patients in surgical ICUs appearing to benefit from IIT with improved survival (RR 0.63; 95% CI 0.44 to 0.91). However, none of the new RCTs included in this systematic review were classified as surgical-ICU trials. The difference in findings between these two systematic reviews was thus based on their different inclu-
sion criteria. Griesdale et al excluded 3 unpublished trials (ab-

383-386 As Cochrane Collaboration recommends, trials published as abstracts should be included in the overall analyses in order to avoid publication bias. Thus I have chosen to update this meta-analysis with these excluded trials and included data from all the relevant latest RCTs. As clearly seen by Figure 10, there is no evidence to support the statements of Griesdale et al. Additionally, the NICE-SUGAR study included 2233 surgical patients with a significantly increased mortality in this subgroup, the majority of whom were admitted to the ICU following emergency surgery. This reinforces the findings of Figure 10, suggesting lack of benefit of IIT in patients treated in surgical ICUs. The latest systematic review on this topic was published in 2010. This paper included 7 RCTs examining the impact of tight glycaemic control (blood glucose 80-110 mg/dl). The authors concluded that there was no evidence to support the use of IIT as tight glycaemic control is associated with a high incidence of hypoglycaemia and an increased risk of death in patients not receiving parenteral nutrition. They found no benefit in terms of incidence of blood stream infections or the requirement for renal replacement therapy.

This topic remains highly controversial. Since 2009, at least 10 additional RCTs have been published examining the role of IIT in treatment of critically ill patients without altering the overall controversy (For description of these trials, see table 3 in appendix VI). Furthermore, there are currently six registered ongoing trials (Table 4 in appendix VII). In summary, the available evidence is still far from convincing. Despite the bias risk, the risk of random error is also substantial as approximately a required information size of 18 000 participants may probably be necessary to detect or refuse a RRR of 10% (alpha=0.05 and beta=0.20 and heterogeneity=18%) in a meta-analysis. However, TSA may reduce the ultimately necessary information size if the bounda-

ries for benefit or harm or futility is reached before the 18000 patients.

Figure 10. Risk ratios of mortality in clinical trials comparing intensive insulin therapy (IIT) to conventional glycaemic control stratified by type of ICU (surgical). Updated with latest RCTs.

4. THE IMPACT OF POINT OF CARE ASSAYS ON TRANSFUSION REQUIREMENTS, AND THEIR POTENTIAL INFLUENCE ON COAGULATION AND INFLAMMATION

4.1 Thrombelastography (TEG) or thromboelastometry (ROTEM) to monitor haemotherapy versus usual care in patients with massive transfusion

This recently published systematic review consists of nine RCTs with 776 patients at high risk of bleeding due to cardiac surgery and liver transplantation. Only one of the included trials had a low risk of bias. Compared with standard treatment, TEG or ROTEM showed no statistically significant effect on overall mortality (3.78% versus 5.11%, RR 0.77, 95% CI 0.35 to 1.72; I² = 0%). However, only five trials provided data on mortality and none were powered to detect any difference in mortality event rate and the result of these analyses are additionally influenced by the short period of follow-up in majority of the published papers and the very low proportion of post operative mortality among cardiac surgery patients. 

We found two ongoing trials among adult cardiac surgery population with high risk of bleeding and an estimated population of 300 to be included but failed to retrieve any data from the investigators at current stage. We were unable to conduct many of our predefined subgroup and sensitivity analyses due to limited data. Subgroup and sensitivity analyses assessing the impact of duration of intervention; length of stay in ICU and hospital, duration of mechanical ventilation, rate of surgical re-intervention, and amount of blood products transfused did not result in a statistically significant finding. Outcomes such as complications to treatment or transfusion requirements and other clinically relevant outcomes were inconsistently reported. Authors were contacted for missing data. Few responded, but did not provide much additional information other than that originally published.

However, our analyses demonstrated a statistically significant effect of TEG or ROTEM on the amount of bleeding (MD -85.05 ml, 95% CI -140.68 to -29.42; I² =26%). Despite evidence of reduced bleeding, our analyses were disabled as various trials described the effects of the intervention differently, and thus preventing adequate pooling of data. Even though a beneficial effect is true, it may only be a surrogate outcome and it is uncertain whether it predicts any noticeable clinical benefits. Thus, it is questionable if a bleeding reduction of 85 ml is clinically relevant and whether the included patients are truly at high risk of bleeding. Additionally, TSA indicates that despite this indication of reduced bleeding by application of TEG/ROTEM, there is still need for additional RCTs before firm evidence may be reached (Figure 4 in appendix IV).

Our chosen primary outcome measure (overall mortality) in this systematic review may be contested by many. However, the large multi-centre RCT on the effects of tranexamic acid on death, vascular occlusive events, and blood transfusion with 20,211 randomized adult trauma patients should serve as an inspiration. Some may argue that technologies such as TEG/ROTEM have not been designed as life saving instruments but rather as qualitative point of care tools, assisting clinicians in the interpretation of whether the transfusion/substitution strategy is adequate to ensure optimal fibrin formation in patient’s blood. However, the choice of overall mortality as the primary outcome measure summarizes ultimate harms and benefit simultaneously. We acknowledge, of course, that other outcomes may have great clinical importance as well and accordingly we analysed the effects of the intervention on several of these in our systematic review. However, it should be equally recognised that as long as none of the surrogate outcomes have been evaluated thoroughly and proven to be relevant surrogate outcomes adequately reflecting patient important outcomes we should be careful not to incorporate evidence of benefit solely on these as arguments for using the intervention.

4.2 There is an urgent need for several large RCTs with low risk of bias to evaluate the use of TEG or ROTEM among populations and clinical settings such as paediatric/neonatal, septic, trauma, critically ill and other surgical populations with massive transfusion such as aneurism repair, liver surgery and post partum haemorr-

DHANIL INE MEDICAL BULLETIN 18
tionally further trials need to focus on other relevant outcomes such as long-term survival, duration of stay in intensive care units and hospital, rate of infection, adverse events and quality of life assessment.

**DISCUSSION**

Limitations and strengths of our systematic reviews

The four systematic reviews include 44 trials with a total of 5537 patients. The methodological quality was low in most trials in our systematic reviews. Among the 44 trials, 30 trials (68.2%) had high risk of bias. On average such trials overestimate intervention’s beneficial effects. Two of our systematic reviews reported larger beneficial treatment effects among trials with high-risk of bias compared to trials with low-risk of bias, however, the other two did not detect a difference of the intervention effect estimates in trials with high- and low-risk of bias.

The systematic review on the use of AT III in critically ill patients included 12 trials with high-risk of bias and eight with low-risk of bias with a total of 3458 patients included. The duration of intervention varied from less than 24 hours to four weeks. Follow up ranged from seven days to 90 days. We classified two trials as obstetric, three trials as paediatric, two as trauma and the remaining trials consisted of mixed populations of critically ill participants, mainly septic. We did not find a statistically significant effect of AT III on mortality among all the trials and in any subgroup but AT III increased the risk of bleeding events. Multiple subgroup and sensitivity analyses including one on the use of adjuvant heparin did not show differences in the estimates of intervention effects. There is insufficient evidence to support the use of AT III in any category of critically ill patients and the effect on all-cause mortality will not exceed a 10% RRR and probably not 5% RRR either.

The systematic review on the use of INO for ARDS or ALI included 10 trials with high-risk of bias and four with low-risk of bias with a total of 1303 participants. We classified three trials as paediatric trials while one trial included a few children and the remaining trials consisted of mixed populations of critically ill adults with ARDS or ALI. The sample size varied from 14 to 385 participants and the duration of intervention varied from less than 24 hours to four weeks. INO showed no statistically significant effect on overall mortality and in several subgroup and sensitivity analyses, indicating robust results. Conversely, INO increased the risk of renal failure among adult population. It transiently improved oxygenation, only for the first 24 hours.

Our systematic review on the use of aerosolized prostacyclin for ALI/ARDS was only able to include one trial with low risk of bias including only 14 critically ill children with ALI and ARDS. The duration of intervention was less than 24 hours and the length of follow up was 28 days. Thus, we have insufficient data to demonstrate benefits or harms of inhaled prostacyclin therapy. We found two ongoing RCTs but were not able to retrieve any data from the authors of these trials since they were not at a stage where they could disclose data. Based on the very limited data available, we were unable to show any benefits of aerosolized prostacyclin on survival but found significant improvement in the oxygenation index in the prostacyclin group.

The systematic review on the benefit of TEG or ROTEM in order to monitor haemotherapy included eight trials with high-risk of bias and one with low-risk of bias with 776 adult patients. The sample size varied from 28 participants to 224 and 8 trials were conducted in cardiac surgery settings and one in liver transplant setting. Only one trial applied TEG beyond the first 24 hours to guide transfusion. We found insufficient data on survival and none of the five trials providing mortality data was powered to show benefit on survival. Only three trials had long-term follow-up for mortality (length of stay in hospital). There is weak evidence to support the use of TEG or ROTEM in cardiac and liver-transplant surgery population as only our secondary outcome parameters showed beneficial effects with a reduction in number of patients actually transfused.

In summary, the currently available reliable evidence has not shown beneficial effects on primary clinical outcome (e.g., mortality) in any of our systematic reviews. However, the trials and reviews on these interventions are under-powered to draw firm conclusions and the confidence intervals include both possible beneficial and possible detrimental effects. Our results are based on few trials with sparse data. Trial selection bias and outcome reporting bias should therefore be considered. Application of trial sequential analyses in these systematic reviews does not support firm evidence for any of the interventions on primary outcome measures and indicates need for additional large RCTs in order to provide evidence of harm or benefit.

But if based on surrogate outcomes (i.e., improved P/F ratio, oxygenation index, proportion of patients receiving blood products, etc) there is evidence favouring some of the interventions. It is important to remember, that such analyses (secondary outcome measures), identifying a subgroup of patients which particularly benefit from the interventions in investigation, are only exploratory and hypothesis generating, especially when the results of the primary outcome measures are statistically insignificant and with broad confidence intervals not excluding substantial harm.

The research on the effects of treatments in our selected group of critically ill patients is limited by several factors such as the lack of shared definitions (i.e., ALI or ARDS), inadequate bias control in the trials as for example lack of properly blinded trials, high-risk of random errors as sample size were generally low, and due to inappropriate design of the trials. This dissertation comprises four systematic reviews, which were based on pre-specified, peer-reviewed, and published protocols. In all reviews, we performed comprehensive searches of major databases and contacted authors and pharmaceutical companies and manufacturers. We appraised the quality of all included trials and emphasised the results of trials with low-risk of bias in our conclusions. We conducted meta-analyses using the Cochrane Collaboration methodology and trial sequential analyses (TSA), and adhered to the GRADE and PRISMA-guidelines when conducting these systematic reviews. Nevertheless, our systematic reviews may still be prone to both publication and reporting bias. Therefore, the results may well tend to overestimate the possible benefits of the interventions evaluated in the present systematic reviews. No systematic review is stronger than the trials it comprises but the reviewing process may reveal exactly that setting and the stage for future research.

One potential weakness of our systematic reviews may be the lack of inclusion of uncontrolled observational clinical studies, since the observational studies play an important role in the evaluation of rare adverse events. Observational studies provide reliable evidence if interventions have dramatic effects (e.g., insulin in diabetic ketoacidosis). But in clinical settings with moderate or small intervention effects, some might argue that the human processing of data, unsystematic data collection, and the human capacity to overcome illnesses spontaneously limit the value of uncontrolled observations. Despite acknowledging the importance of experimental studies for estimation of toxicity and their necessity for increasing our pathophysiological understanding
of the underlying biologic mechanisms, their main shortcoming is the required extrapolations, which may lead to the wrong conclusions. Overall, the observational studies tend to overestimate the beneficial effect of the interventions and their benefit could be attributed to periodical changes, recall bias, and differential measurement errors.\textsuperscript{119}

Although there was minimal heterogeneity among trial results on mortality in our systematic reviews, we are aware that we pooled heterogeneous trials in terms of age, patients, settings, and treatment regimens. Thus, the validity of our meta-analysis may be criticized. However, all trials included patients with clinical conditions of similar inflammatory pathways. Therefore, we think that there is a good biologic reason to perform broad meta-analyses, which also considerably increases the generalisability and usefulness of the review. Further, a broad meta-analysis increases power and precision, reduces the risk of erroneous conclusions, and facilitates exploratory analyses, which can generate hypotheses for future research.\textsuperscript{209} This will be discussed in more detail in the following sections.

**Limitations and strengths of systematic reviews in general**

Systematic reviews provide a more objective appraisal of the evidence than traditional narrative reviews with a more precise estimate of a treatment effect and may explain heterogeneity between the results of individual trials. Additionally, they may highlight weaknesses within the research field and generate important research questions to be addressed in future trials.\textsuperscript{7} However, systematic reviews have their limitations since the nature of systematic reviews is that of a retrospective observational study with all the corresponding bias risks. Systematic reviews may have considerable limitations due to meta-analysis of results of heterogeneous RCTs, including systematic error (bias), random error, inadequate update and rarely incorporating evidence from non-RCTs.\textsuperscript{6} Additionally their validity may suffer from publication bias, confounding and selective reporting of outcomes in the included trials, which can lead to false positive conclusions, as previously described in detail.\textsuperscript{310} In order to minimise the risk of bias and to enhance transparency, a good systematic review should be based on a pre-specified, peer-reviewed published protocol. This contains clearly formulated questions and descriptions of explicit methods in the identification, selection, and evaluation of included trials. Often, a systematic review will include a meta-analysis, which offers a quantitative summary of the results from individual trials. When conducting a meta-analysis, the sample sizes, populations and interventions may differ since the trials may have been conducted over a long period. Additionally, external factors such as advances in overall health care support, advances in surgical and medical interventions, changes in antibiotic treatment and resistance may limit the validity of the evidence.

Thus, it may be argued that the settings are not identical among trials in a meta-analysis and the results of a new trial are not independent of the results from previous trials. In other words, since systematic reviews bring together trials that are diverse both clinically and methodologically, heterogeneity in their results is to be expected. Heterogeneity is likely to arise through diversity in doses, lengths of follow up, study quality, and inclusion criteria for participants. However, in large multicentre trials analogous problems may arise and they may be analysed in a fixed-effect model despite risk of even substantial clinical and statistical heterogeneity among recruiting centres. This has led to the suggestion of multicenter RCTs taking heterogeneity into consideration when estimating the sample size.\textsuperscript{311} Among the strengths of Cochrane reviews are not only the publication of a pre-specified and peer-reviewed protocol but also electronic availability, regular update and strict methodology based on The Cochrane Handbook.\textsuperscript{7}

The Cochrane Collaboration recommends that Cochrane reviews should provide the evidence rather than direct recommendations, as it may be very difficult for authors to put the evidence into context of any setting (external validity).\textsuperscript{7} Subsequently the internal validity of the review should be assessed by the clinicians, which must then critically decide how applicable the evidence is to their setting. The latter decision depends not only on the type of patients, but also on factors such as costs, resources, and the local assessment of the weights of benefits and harms.\textsuperscript{7} Additionally, health-care characteristics should also be considered since trials are often conducted at specialised centres in high-income settings with capabilities and expertise exceeding those available at routine patient care or in low-income settings. Finally, additional trials on the same intervention for the same condition may be considered. Despite agreement or disagreement between systematic reviews and guidelines they may both be wrong. In many cases, there may not even be any large trials but several small trials with low statistical power focusing on non-validated surrogate outcomes.\textsuperscript{109,312} Meta-analyses, especially on ‘hot’ topics, are updated several times. Additionally, there is an increasing acceptance that meta-analyses ought to be conducted both before and after each new trial.\textsuperscript{118} Therefore, statistically significant results due to repetitive analyses represent a real problem. This is the rational for applying trial sequential analysis (TSA, Appendix II). TSA provides information on which meta-analysis needs further participants (new trials) to obtain firm evidence and which may have reached sufficient information. Obviously, the decision to conduct more RCTs must be taken on a ‘meta-analysis by meta-analysis’ basis, incorporating all knowledge on the specific therapeutic area. However, as previously discussed, application of TSA provides more conservative conclusions compared to ‘traditional’ meta-analyses.\textsuperscript{196} This may delay the use of potentially beneficial interventions by as much as 2-3 years but it should be weighted against the risk of introducing useless or even harmful intervention based on inconclusive evidence.\textsuperscript{196} However, application of TSA in a prospectively manner in planning of future trials may reduce this delay.\textsuperscript{192}

**Randomized clinical trials: challenges and future perspectives**

Archie Cochrane stated: “The RCT is a very beautiful technique of wide applicability, but as with everything else there are snags”.\textsuperscript{113} Clinicians making decisions on the basis of RCTs need to be cautious of small trials (even those properly randomised) and systematic reviews of small RCTs, both because of risk of systematic and random error and especially risk of biased reporting.\textsuperscript{310} We acknowledge the challenges and obstacles of performing large RCTs on interventions among critically ill patients. Some might state that it is overoptimistic to aim for large multicentre RCTs in critical care since most RCTs are organized and funded in the high-income nations where chronic diseases are responsible for most morbidity and mortality. Thus, we can anticipate only small to moderate treatment effects from our interventions.\textsuperscript{314} Providing definitive answers in the face of small-to-moderate treatment effects often necessitates large sample sizes. Organizing large trials involves big challenges, as does monitoring the quality of enrolment and data collection once the trials begin. Funding for such large trials is an obvious challenge if we desire not to be restricted to industry sources. Nevertheless, it is possi-
ble to conduct large multicentre trials with appropriate statistical power, if international groups of investigators collaborate. But does size of the trials really matter? After all, power calculations are often inaccurate and can have serious implications. In my opinion the answer is YES. If our aim is to detect an intervention effect that is both likely to influence clinical practice and is feasible, then the sample size ought to be large (powerful) enough to detect that effect with reasonable confidence, allowing for acceptable errors (typically 5% (or 1%) for a type 1 error, and 20% (or 10%) for a type 2 error). The effect size decision is more challenging than an acceptable error estimate decision since trialists often have to choose between a very large intervention effect with the potential to change clinical practice and a more realistic estimate based on emerging empirical evidence (i.e., systematic review or a pilot study) about the likelihood of a particular effect size.

However, if trials are too small, they may miss realistic but moderate and clinically important treatment effects, and a potentially useful intervention may therefore be dismissed. The consequence of the latter at its extreme would be lack of initiation of further trials if considered clinically or commercially not feasible. On the other hand, if an underpowered trial does find a statistically significant effect, there is a great risk that it will overestimate the size of the effect if there is any effect at all. This occurrence may also stop initiation of further trials since clinicians might feel that it is unethical to randomise more patients if a treatment appears to be beneficial. Thus, some might argue that no trial regardless of its size and quality should stand alone.

We should remember that systematic reviews cannot replace RCTs but should rather be populated by preferably high quality large RCTs. However, large RCTs may still have methodological shortcomings despite our very best attempts. As I will discuss later, large RCTs using sound research design and methodology are strongly needed in the field of critical care. Trials should focus on patient-relevant outcomes like clinical improvement, recovery, mortality, length of stay in hospital with strictly defined discharge criteria, adverse events and quality of life. Future trials should report their data according to the recommendations of the CONSORT Statement. Additionally, better reporting on surrogates in RCTs might reduce unwarranted conclusions and uncritical acceptance of new treatments.

One special concern is the relative rate of increase in the number of RCTs for adults compared with those for children despite financial and legislative incentives to promote conducting paediatric trials. The number of RCTs with the participation of adults published annually in high-impact general medical journals has doubled during the past 20 years. Conversely, there has been virtually no change in paediatric trials and the gap seems to be widening in almost every major clinical specialty including critical care, anaesthesia and emergency medicine (Fig. 11). This issue is of major concern despite acknowledging the logistical, financial, and ethical barriers to conducting trials in children.

Given the difficulties in identifying unreported outcomes and in contacting investigators for further information, all trials should at their initial stage be registered in a public database. Complete data for all pre-specified trial outcomes, independent of their results should subsequently be made publicly available. Discrepancies between outcomes in the methods and results sections of publications should also be addressed during peer review. Trial protocols should be made available in the public domain before trial completion and editors and reviewers should compare submitted manuscripts with the published protocols when considering these for publication. This will lead to a transparent choice of outcome and analysis plan minimising the risk of selective reporting. One might argue that the latter strategy is not feasible for journals due to space restrictions. However, this is hard to see as increasing number of journals adopt internet sites, which can alleviate these concerns.

Recommendations of interventions based on statistically significant outcomes (P < 0.05) due to use of unvalidated surrogates may lead to the introduction of interventions of considerable costs with no true value or a potential harmful effect. A recent publication indicates that only about one third of authors of RCTs that used a surrogate as a primary outcome reported adequately on the surrogate. This may explain the often observed disagreements between meta-analyses and large trials and between large trials on the same topic. RCTs focused on surrogate outcome measures are less directly applicable to patients than trials focused on patient-important outcomes. Thus readers should be aware of conclusions in RCTs that are based on unvalidated surrogates.

Furthermore, we need to abandon our convenient, yet insufficient and inappropriate strategy of claiming conclusive research findings solely on the basis of statistical significance, typically a p-value < 0.05. This explains the high rate of non-replication (lack of confirmation) of research discoveries in majority of published research claims. Since it is near impossible to know with certainty what the “truth” is in any research question, the pure “gold” standard may be unattainable. However, approaches to improve the post-study probability such as better powered evidence, e.g., large RCTs or low-bias meta-analyses, may close the gap to the unknown “gold” standard.

Figure 11. Trend in publication of RCTs in selected specialty journals during a period of 20 years.

Systematic reviews: challenges and future perspectives

Despite their potential shortcomings, when conducted adequately, systematic reviews can assist clinicians keep up-to-date and provide evidence for policy makers to judge risks, benefits, and harms of health care interventions. Additionally, systematic reviews locate and summarize related research findings for patients and their carers while assisting developers of clinical guidelines. Finally, systematic reviews provide summaries of previous research for funders wishing to support new research and help editors judge the merits of submitted reports of new trials.

Systematic reviews addressing topics pertinent to critical care medicine appear to have an overall poor quality of reporting. This may reduce the confidence of clinicians in applying evidence-based practices and reduces the validity and applicability of systematic reviews in general. The main methodological areas that were reported to be deficient in critical care reviews were the conduct of a comprehensive
search, the avoidance of bias in the inclusion of studies and the appropriate assessment of the validity of the included studies. The quality of the systematic reviews in the critical care literature appears to be comparable to the quality of reviews published in the emergency medicine and anaesthesia literature. The number of systematic reviews is increasing in the field of critical care and physicians cannot simply rely on browsing critical care journals for systematic reviews, as the majority of these reports are not published in critical care journals. Systematic reviews have the potential for guiding future clinical research and ideally, no clinical trial should be planned without a thorough knowledge of similar, existing research.

However, empirical evidence suggests that this is far from the case and authors of primary trials do not consider a systematic review when designing their trials. Given the potential role of systematic reviews in decision making, authors should be transparent about their funding, the role of funders in the reporting of the systematic review, and should declare any real or perceived conflict of interest. Due to the poor reporting quality of non-Cochrane systematic reviews, and some Cochrane reviews, there is an urgent need to raise awareness of the principles of evidence-based medicine among literature consumers and authors, and promote familiarity with The Cochrane Library and Cochrane Collaboration methodological standards. To improve transparency, authors of systematic reviews should be required to publish a pre-specified and preferably peer-reviewed protocol. Access to the protocol should easily be available. Currently, non-Cochrane protocols can be registered free of charge at the UK national research register through the Centre for Reviews and Dissemination in York, UK. Protocols for Cochrane reviews are published in the Cochrane Library.

Health care journals, editorial groups and reviewers play a major role in enhancing the quality of systematic reviews by promoting higher methodological standards by scrutinizing the methodology of each individual systematic review before accepting its results. They should formally endorse guidelines for transparent reporting such as PRISMA and CONSORT statements. The methods used for searching for eligible trials, the inclusion of trials, the validity assessment of included trials, data extraction and analysis of included trials should be improved to ensure a proper basis for evidence-based clinical decision making. However, while introduction of guidelines such as CONSORT may improve reporting for more recent trials as increasing number of journals and authors adopt these guidelines, systematic reviewers will continue to face issues arising from poor reporting when they include studies from the era before the guidelines.

The inflammation and coagulation axis: challenges and future perspectives

As discussed earlier, the interaction between coagulation factors and the inflammatory response is quite complex and our knowledge is still evolving through basic research. The bidirectional interplay between coagulation and inflammation provides potential mechanism(s) for modification and thus development of new interventions to minimize the detrimental inflammatory response and uncontrolled coagulation (at its worst DIC). However, despite encouraging results from basic science and preliminary trials, the disappointing results of subsequent larger RCTs and systematic reviews firmly illustrate the gaps in our current knowledge about the complexity of the potential and multiple actions of the proteins, such as protein C and AT III. Administration of intrinsic/synthetic analogues of coagulation cascade components (e.g., protein C, AT III, factor Vlla or Xa), coagulation inhibitors (e.g., heparin or analogues), or agonists or antagonists of the protease activated receptors in order to modulate the coagulation/inflammation balance, can result in responses ranging from beneficial to detrimental.

Due to the heterogeneity of patients with critically ill syndromes such as ARDS or sepsis, it may be overly optimistic to presume that any new intervention resulting from evolving knowledge of the bidirectional contribution of the clotting cascade and inflammation can be universally applied. Thus, many patients will surely have contraindications to the new intervention or to certain dosages of new drugs due to pharmacokinetic or pharmacodynamic variability. There is an apparent need for knowledge about these contraindications and variables. Some may be obvious (i.e., coagulation disorders) but others may not be anticipated before being put to practice since alteration of complex physiological interplay may be influenced by many, as yet unknown, factors in the individual patient. An example is the concomitant administration of heparin and ATIII in septic patients, which antagonizes the anti-inflammatory effect of ATIII and greatly accelerates its anti-coagulant abilities and increases bleeding events. It is safe to say, that a more complete understanding of the impact of concurrent conditions, confounding therapies and the patient’s genetic structure on outcome will be required if we in the future are to obtain the maximum benefit of any new intervention aiming to alter the balance of such complex interplay as coagulation and inflammation.

Critical care: challenges and future perspectives

Critical care or intensive care medicine is a nascent specialty, dating back to the 1952 Copenhagen poliomyelitis epidemic and the subsequent realisation that prevention, care, and when possible cure of the acutely unwell patients is often required. Intensive care medicine is challenged by the consumption of considerable and increasing proportion of health-care resources in order to meet the needs of an ageing population and cope with the consequences of conflicts, natural disasters, inadequate primary care, and high-risk treatments for very sick patients, at a time of economic constraint. Intensivists are faced with challenge of utilising a powerful ability to manipulate physiology, biochemistry, and immunology to improve outcomes for their patients. Critical illness syndromes (i.e., sepsis, ALI, ARDS, coagulopathy) cannot be diagnosed with simple tests and their definitions are based on clinical, laboratory, radiological, and physiological criteria, derived by consensus panels, and under continual debate and revision. These syndromes have a brief prodrome and high short-term mortality, thus reducing the number of available patients for a trial at any given time. Mortality after critical illness is related to a complex interplay between disposable resources, ICU capacity, cultural and religious perceptions, clinical decision on the level of intensive care, the consequences of the disease and finally the availability and intensity of health services beyond ICU. Regardless of regional differences in ICU set-up (surgical, medical or mixed ICU) and irrespective of the initiating insult (i.e. infection, trauma, massive bleeding), the final common pathway for many patients in ICU (i.e., multiple organ failure and DIC) is similar. However, the mortality rate differs greatly and is much higher among medical patients in ICU since surgery often is a curative procedure. Despite the proliferation of local, national, and international academic research networks organising multicentre trials, and despite the emergence of various international guideline committees, intensive care medicine is still trailing significantly compared...
to other specialties in terms of evidence based practices. Even when there is sufficient evidence to support a treatment, universal implementation has been difficult to achieve due to uncertainty about the generalisability and applicability of trial results to individual patients and perceptions of potential harm. The results of many multicentre trials in intensive care have been discouraging with either negative or even harmful effects in previously theoretically considered beneficial treatments. In many cases, initial optimism based on results of RCTs have been replaced by disappointment due to negative findings of subsequent trials (i.e., tight glycaemic control and activated protein C). This raises questions about issues related to the complexity of the disease processes; patient heterogeneity; definition and diagnostic challenges; trial designs; insufficient knowledge of underlying pathophysiological processes; and subsequent variations in interpretation, acceptance, and implementation of study findings. Designing RCTs among heterogeneous populations (i.e., sepsis, ALI, ARDS, patients at high risk of bleeding) is challenging since they include several clinical disorders in which the timing and extent of any immunomodulation might be crucial to outcome. Furthermore, the results may be confounded by wide variability in the management regimens such as sedation, nutrition, transfusion strategies, fluid balance and requirements of certain physiological endpoints (i.e., vasopressors for certain blood pressure). As a consequence, trials require enrolment of large number of patients in order to detect even moderate differences in survival. Some of these barriers have even lead some opinion leaders astray questioning the value at all of RCT’s in the ICU setting. These statements may be considered “loosing the edge of underlying pathophysiological processes; and subsequent variations in interpretation, acceptance, and implementation of study findings. Designing RCTs among heterogeneous populations (i.e., sepsis, ALI, ARDS, patients at high risk of bleeding) is challenging since they include several clinical disorders in which the timing and extent of any immunomodulation might be crucial to outcome. Additionally, there is little evidence to support that our conventional approach of organ support in ICUs for organ failure (i.e., mechanical ventilation, dialysis, vasoactive and inotropes) in itself is curative. The only true exception is perhaps antibiotic therapy and source removal in patients with sepsis. As a consequence, the application of these techniques will at its best probably only support the patient during their stay in the ICU in the hope that extra time will enable the patient to selfheal. Despite the allocation of a great amount of time, energy, and resources, the search for a cure-all magic bullet in treatment of critically ill has failed so far. There is a need for a greater understanding of the basic processes involved in the care of critically ill patients.

CONCLUSION

There is an urgent need for applying evidence-based practice to the treatment of critically ill patients. It is a daunting task trying to convert anaesthesiologists and intensivists away from their conventional practices, but the future appears bright since ever more journals, international scientific organisations and societies and editorial boards appear to be embracing systematic reviews as the highest standard of evidence when dealing with interventions and treatment of patients. It is important to remember that guidelines and treatment recommendations based on small trials or meta-analyses of small numbers of trials can be misleading due to random error, study design and bias. We need to focus on more than just a p value < 0.05 or a confidence interval which both may have to be corrected by sequential methods in cumulative meta-analysis when conducting systematic reviews or when focusing on guidelines and recommendations based on few RCTs. Often, when larger and methodologically sound trials appear, the evidence for beneficial effects becomes less convincing and subsequently disappears or even turns around to become evidence for a harmful effect. This certainly emphasises the need for careful estimation of the amount of low-bias information we need before we decide to implement interventions. We should view any gathered information as a mere step in the right direction urging us to evaluate its significance in the light of good bias control, relevant design, and results from sequential methods to reduce random error. Unfortunately, very few of our treatment regimens in critical care are founded on the basis of solid evidence. Hence, there is need for a critical open mind rather than blind adherence to guidelines, which optimally should take on the form of updated evidence rather than expertise-based recommendations. When confronted with various decisions in our daily practice, we should constantly ask ourselves whether the patients would benefit from our practices when they are not evidence-based. Fortunately, even intensivists appear to become more focused on applying evidence-based practices.

Let us help our patients by addressing the fundamental question of quality standards and start basing our practices on solid and robust evidence. In summary, let us conduct and base our recommendations for interventions on large high quality, large RCTs as well as high quality systematic reviews with meta-analyses of such RCTs rather than reinforcing implementation of interventions with dubious evidence. We should dedicate ourselves to help close the information gap.

APPENDIX I: GENERAL METHODOLOGICAL CONSIDERATIONS

To quantify the estimated effect of various interventions, we conducted meta-analyses using the Cochrane Collaboration-methodology, trial sequential analyses (TSA), the GRADE and PRISMA-guidelines when conducting our systematic review. All reviews were performed according to published protocols following the recommendations of the Cochrane Handbook for systematic reviews of interventions.

Trial selection: We searched in the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE, Science Citation Index-Expanded, The Chinese Biomedical Database, and LILACS. We included all randomized clinical trials. We hand-searched reference lists, reviews, and contacted authors and experts for
additional trials. We searched ClinicalTrials.gov, Centre Watch Clinical Trials Listing Service and ControlledTrials.com for missed, unreported, or ongoing trials. We screened bibliographies of relevant articles and conference proceedings and wrote to trialists and pharmaceutical companies producing the drugs in question.

Data selection and extraction: AA screened the titles and abstracts for relevant studies. At least two authors independently evaluated whether identified trials fulfilled the inclusion criteria. AA and a co-author independently extracted data from the retrieved trials. Disagreements were resolved by discussion among all the authors involved. Trial authors were contacted for additional information.

We only included randomized clinical trials comparing the interventions as previously mentioned. Inclusion was regardless of publication status, language, outcomes reported, or blinding status. We evaluated the validity and design characteristics of each trial and bias risk components (random sequence generation, allocation concealment, blinding, incomplete data outcomes, selective outcome reporting, sample size and power calculation, and the ability to perform intention to treat analysis). Trials were defined as having a low-risk of bias if they fulfilled the above criteria according to Cochrane Handbook definitions. Our primary outcome measure was mortality in all publications. Secondary outcomes were examined as defined in various papers.

Statistical analysis: Data were summarized as relative risks (RR) with 95% confidence intervals (CI) for dichotomous variables and the mean difference with 95% confidence intervals (CI) for continuous outcomes. We used fixed- and random-effects models for all meta-analyses. Heterogeneity was explored by visual inspection of the forest plots and by using a standard Cochran’s Q-test and $I^2$. $I^2$ values of 50% and more indicate a substantial level of heterogeneity. In the case of heterogeneity ($I^2 > 10\%$), we reported results from the random-effects model. We analyzed data by intention to treat and included all patients. All forest plots and meta-analytic estimates were calculated with RevMan 5.

We carried out multiple number of subgroup and sensitivity analyses to assess specific benefits or harms of each intervention in various settings. If analyses of various subgroups with binary data were significant, we performed a test of interaction. We considered P < 0.05 as indicating significant interaction between the intervention effect on mortality and the subgroup category. To assess publication bias and other types of bias we created funnel plots for mortality.

**Trial sequential analysis (TSA):** Meta-analyses may result in type-I errors due to an increased risk of random error when few data are collected and due to repeated significance testing when repeatedly updated with new trials. To assess the risk of type-I errors, we used TSA. For more detailed description of TSA, see Appendix 2.

**GRADE-criteria:** We summarized the evidence applying GRADE-levels (high, moderate, low, and very low) by evaluating design, quality, consistency, precision, directness and possible publication bias of the included trials using GRADEpro-version 3.2.2-software.

**APPENDIX II: TRIAL SEQUENTIAL ANALYSIS (TSA)**

Meta-analyses of randomized trials increase the power and precision of the estimated intervention effects. However, meta-analyses may result in type-I errors (false positive results) or overestimate treatment effects due to increased risk of systematic errors (bias) and random errors (play of chance). Increased risk of random error may arise due to repeated significance testing when updated with results of new trials due to repetitive testing of accumulating data, which inevitably, sooner or later, lead to type-I errors.

The required information size (the required cumulated number of participants for a meta-analysis should be at least as large as an adequately powered single trial since results from adequately powered and adequately bias-protected trials often fail to confirm findings from statistically significant small trials. Trial sequential analysis (TSA) combines information size estimation for meta-analysis (cumulated sample size of included trials) with an adjusted threshold for statistical significance in the cumulative meta-analysis. The latter, called trial sequential monitoring boundaries, analogous to interim monitoring boundaries in a single trial reduce type-I errors. In TSA, the addition of each trial in a cumulative meta-analysis is regarded as an interim meta-analysis before the required information size is reached and helps to clarify whether additional trials may be needed or not. Trial sequential monitoring boundaries adjust the level of statistical significance that is required to minimise the overall type-I error risk and to obtain firm evidence considering the number of participants and events in a meta-analysis P-value. The fewer participants and events, the more restrictive the monitoring boundaries are and the lower P-value is required to obtain statistical significance. The idea in TSA is that if the cumulative Z-curve crosses the trial sequential monitoring boundary, a sufficient level of evidence has been reached and no further trials are needed (Fig.1). If the Z-curve does not cross the boundary and the required information size has not been reached, there may be insufficient evidence to reach a conclusion.

Application of TSA on meta-analyses has some apparent and potential benefits such as the potential to minimise false positive results due to random errors if the sufficient information size is not reached (Fig.2). There is increasing evidence that meta-analyses should be assessed with TSA as rigorously as sample size estimation and sequential monitoring boundaries are applied to a single trial. Additionally, TSA provides information about when reliable evidence is obtained which can stop implementation of redundant trials. Finally, in non-significant meta-analyses TSA with both boundaries for beneficial and harmful effects and futility may provide information on whether more trials are needed (absence of evidence) or not (evidence of absence of effect).
Figure 1: Example of upper half of two-sided trial sequential analysis. TSA is a symmetric two-sided analysis as the cumulative Z-curve can obtain negative and positive values. The cumulative Z-curve was constructed with each cumulated Z-value calculated after including a new trial according to publication date. Thus, when comparing two different interventions, either the upper or the lower monitoring boundary needs to be crossed to obtain firm evidence adjusted for random error risk for a significant difference. Brok et al. J Clin Epidemiol. 2008;61(8):763-9. Reprint with permission from publisher.

If more trials are needed, TSA can help re-estimate the additional required cumulated sample size of patients in future trials to obtain firm evidence in the meta-analyses given the estimates of heterogeneity and intervention effect. TSA is applicable to any type of binary or continuous outcome meta-analyses. It provides more conservative conclusions, which may delay clinicians’ use of potentially beneficial interventions. However, such delay should be weighted against the risk of introducing useless or even harmful interventions based on insufficient evidence.

Figure 2: Examples of upper half of two-sided trial sequential analyses. The cumulative Z-curves (A-D) from four different meta-analyses were constructed. Crossing of Z=1.96 provides a “traditionally” significant result (A). Crossing of the monitoring boundary before reaching the information size is needed to obtain reliable evidence adjusted for random error risk (B). Z-curves not crossing Z=1.96 indicate absence of evidence if the information size is not reached (C) or lack of the predefined intervention effect if the information size is reached (D). Brok et al. J Clin Epidemiol. 2008;61(8):763-9. Reprint with permission from publisher.

Trials with high-risk of bias (i.e., inadequate generation of the allocation sequence, inadequate allocation concealment, or inadequate double blinding) tend to overestimate the intervention effects compared to trials with low-risk of bias. Additionally, meta-analysing data from multiple trials as if participants originated from one trial may be biased because of trial heterogeneity. Since many of the published systematic reviews do not apply the more conservative random-effects model to their meta-analyses in case of considerable heterogeneity, there is a potential for larger number of analyses obtaining traditional significance (P<0.05) despite lack of a “true” intervention effect. In case of bias in a substantial proportion of trials, TSA would also provide misleading results. Therefore, it seems more reliable to base the TSA calculation of an effect size magnitude and the information size on the intervention effect estimated by the trials with low-risk of bias adjusted for heterogeneity. Therefore it is apparent that increase in heterogeneity will result in a larger required information size: the low-bias, heterogeneity/diversity-adjusted information size (LBH(D)IS). TSA-LBH(D)IS seems more reliable, as it is based on data from previous low-bias risk trials. However, in case of publication of only few small low-bias risk trials, even the TSA-LBDIS may be inaccurately estimated.

If the monitoring boundaries for TSA-LBDIS are crossed before the information size is reached, there is indication of larger intervention effect in high-bias risk trials compared to low-bias risk trials. However, combination of TSA with bias-risk assessment is recommended in, for example, subgroup analyses, funnel plots, and meta-regression analyses.

We applied TSA in our systematic reviews to reduce the risk of increased type-I errors, to assess the level of evidence and to estimate the number of patients needed in further trials.

APPENDIX III: CATEGORIZATION OF SYSTEMATIC ERROR (BIAS) OF CLINICAL INTERVENTION STUDIES INTO LEVELS OF EVIDENCE

<table>
<thead>
<tr>
<th>Category</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1a</td>
<td>Meta-analysis of randomized trials with low risk of bias</td>
</tr>
<tr>
<td>Level 1b</td>
<td>Randomized trial with low risk of bias</td>
</tr>
<tr>
<td>Level 1c</td>
<td>Meta-analysis of all randomized trials</td>
</tr>
<tr>
<td>Level 1d</td>
<td>Randomized trial with high risk of bias</td>
</tr>
<tr>
<td>Level 2a</td>
<td>Meta-analysis of cohort studies</td>
</tr>
<tr>
<td>Level 2b</td>
<td>Cohort study</td>
</tr>
<tr>
<td>Level 3a</td>
<td>Meta-analysis of case-control studies</td>
</tr>
<tr>
<td>Level 3b</td>
<td>Case-control study</td>
</tr>
<tr>
<td>Level 4</td>
<td>Case-series</td>
</tr>
<tr>
<td>Level 5</td>
<td>Expert opinion</td>
</tr>
</tbody>
</table>

APPENDIX IV: TSA FIGURES OF INCLUDED INTERVENTIONS
Figure 1. Trial sequential analysis (TSA) of all trials of the effect of AT III on mortality. Cumulative z-curve in blue does not cross the trial monitoring boundaries constructed for an information size of 3317 patients in the meta-analysis (full red line with open diamonds) with a RRR of 10% (alfa = 0.05) and a power of 80% (beta=0.20). With the total number of accrued participants in randomized trials being 3458, we are able to reject a RRR of 10% with a power of 80%. Reprint with permission from publisher.

Figure 2. TSA of all trials of the effect of INO on mortality (longest follow up, minus zero event trials). The a priori heterogeneity-adjusted required information size (4679 patients) is determined by an assumed 10% RRR. The cumulative z-curve (blue line with filled squares) at the current accrued information size of 1233 patients does not cross the trial monitoring boundaries (red lines with open diamonds) constructed for an information size of 4679 patients (indicated by the vertical red line) in the meta-analysis. A boundary crossing is necessary to detect or reject a relative risk reduction or increase in overall mortality of 10% with a type I error of 5% (alpha = 0.05) and a power of 80% (beta = 0.20) taking multiple up-dating and early “random high” in meta-analyses into consideration. The horizontal lines at cumulative z = 1.96 and at cumulative z = -1.96 indicates conventional levels of statistical significance corresponding to a P = 0.05 (double-sided) at the required information size (n=4679 patients). Reprint with permission from publisher.

Figure 3. TSA of all trials of the effect of INO on P/F ratio (longest follow up). The a priori low-bias heterogeneity-adjusted required information size (2236 patients) is determined by a mean difference (MD) of 14.94. The cumulative z-curve (blue line with filled squares) at the current accrued information size of 614 patients crosses the trial monitoring boundaries (red lines with open diamonds) constructed for a low-bias heterogeneity-adjusted information size of 2236 patients in the meta-analysis in order to detect or reject a mean difference in P/F ratio of 15 mm Hg suggested by the trials with low-risk of bias. This analysis indicates statistical significance in favour of improved oxygenation even with adjustment for repetitive testing on accumulating data in the cumulative meta-analysis since the z-curve crosses the trial sequential monitoring boundary. The horizontal lines at cumulative z = 1.96 and at cumulative z = -1.96 indicates conventional levels of statistical significance corresponding to a P = 0.05 (double-sided) at the required information size (n=2236). Reprint with permission from publisher.

Figure 4. TSA of the effect of TEG or ROTEM on bleeding (longest follow up). The a priori required information size of 2717 patients is determined for a mean difference (MD) of -85 ml. The cumulative z-curve (blue line with filled squares) at the current accrued information size of 776 patients does not cross the trial monitoring boundaries (red lines with open diamonds) constructed for a required diversity-adjusted information size of 2717 patients (D^2=0) in the meta-analysis in order to detect or reject a mean difference in bleeding of 85 ml. A boundary crossing is necessary to detect or reject a relative risk reduction in bleeding of 85 ml difference with a type I error of 5% (alpha = 0.05) and a power of 80% (beta = 0.20) taking multiple up-dating and early “random high effect” or “early random low P-value” in meta-analyses into consideration. The horizontal lines at cumulative z = 1.96 and at cumulative z = -1.96 indicates conventional levels of statistical significance corresponding to a ‘cumulative’ P = 0.05 (double-sided). This analysis indicates that despite indications of reduced bleeding of 85 ml by application of TEG/ROTEM, there is still need for additional RCTs before firm evidence may be reached. (RIS= required information size). Reprint with permission from publisher.

APPENDIX V: ACTIVATED PROTEIN C

Figure 5. APC versus placebo, 28-day all cause mortality. Subgroup analysis based on population.
Table 1: Ongoing trials on the use of activated protein C (APC) among critically ill

<table>
<thead>
<tr>
<th>Study Year</th>
<th>N IIT</th>
<th>Death</th>
<th>Population/ Trial Description</th>
<th>Intervention</th>
<th>Main Findings</th>
<th>Over-all risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Savioli, PM 2009</td>
<td>90</td>
<td>45</td>
<td>Septic shock, sepsis, adults, mixed ICU</td>
<td>IT (BG 80-110 mg/dl vs BG 180-200 mg/dl</td>
<td>ICU mortality: 20% vs 18% (p&lt;.05), 90 days mortality: 31% vs 29% (p=.83)</td>
<td>High</td>
</tr>
<tr>
<td>Zuran, PM 2009</td>
<td>29</td>
<td>14</td>
<td>Septic shock ARF, adults, mixed ICU</td>
<td>IT (BG 4.4-6.1 mmol/l vs BG &lt;11.9 mmol/l)</td>
<td>6 months mortality: 26% vs 28% (p&lt;.02), rate of hypoglycaemia (p&lt;.001), rate of infection &amp; ICU stay ↓</td>
<td>High</td>
</tr>
<tr>
<td>Blotta, MM 2009</td>
<td>483</td>
<td>241</td>
<td>Critical neurosurgical, adults</td>
<td>Mixed ICU, adults</td>
<td>No difference in mortality or hypoglycemia</td>
<td>Uncertain</td>
</tr>
<tr>
<td>Wuytacka, JMM 2009</td>
<td>40</td>
<td>20</td>
<td>Septic shock</td>
<td>Mixed ICU, adults</td>
<td>IT (BG 80-110 mg/dl vs BG 180-200 mg/dl</td>
<td>High</td>
</tr>
<tr>
<td>Vasseelaars, JMM 2009</td>
<td>700</td>
<td>349</td>
<td>Septic shock infants &amp; children</td>
<td>Mixed ICU, adults</td>
<td>IT (BG 2.8-4.4 mmol/l vs BG ≥11.9 mmol/l)</td>
<td>Low</td>
</tr>
</tbody>
</table>

Table 2: Ongoing trials on the use of glucocorticoids among critically ill

<table>
<thead>
<tr>
<th>Study Year</th>
<th>N IIT</th>
<th>Death</th>
<th>Population/ Trial Description</th>
<th>Intervention</th>
<th>Main Findings</th>
<th>Over-all risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Savioli, PM 2009</td>
<td>90</td>
<td>45</td>
<td>Septic shock, sepsis, adults, mixed ICU</td>
<td>IT (BG 80-110 mg/dl vs BG 180-200 mg/dl</td>
<td>ICU mortality: 20% vs 18% (p&lt;.05), 90 days mortality: 31% vs 29% (p=.83)</td>
<td>High</td>
</tr>
<tr>
<td>Zuran, PM 2009</td>
<td>29</td>
<td>14</td>
<td>Septic shock ARF, adults, mixed ICU</td>
<td>IT (BG 4.4-6.1 mmol/l vs BG &lt;11.9 mmol/l)</td>
<td>6 months mortality: 26% vs 28% (p&lt;.02), rate of hypoglycaemia (p&lt;.001), rate of infection &amp; ICU stay ↓</td>
<td>High</td>
</tr>
<tr>
<td>Blotta, MM 2009</td>
<td>483</td>
<td>241</td>
<td>Critical neurosurgical, adults</td>
<td>Mixed ICU, adults</td>
<td>No difference in mortality or hypoglycemia</td>
<td>Uncertain</td>
</tr>
<tr>
<td>Wuytacka, JMM 2009</td>
<td>40</td>
<td>20</td>
<td>Septic shock</td>
<td>Mixed ICU, adults</td>
<td>IT (BG 80-110 mg/dl vs BG 180-200 mg/dl</td>
<td>High</td>
</tr>
<tr>
<td>Vasseelaars, JMM 2009</td>
<td>700</td>
<td>349</td>
<td>Septic shock infants &amp; children</td>
<td>Mixed ICU, adults</td>
<td>IT (BG 2.8-4.4 mmol/l vs BG ≥11.9 mmol/l)</td>
<td>Low</td>
</tr>
</tbody>
</table>
Table 3. Characteristics of recently published RCTs on the use of tight glycaemic control. ARF: acute respiratory failure; BG= blood glucose; GOS: Glasgow outcome score; ICU: Intensive Care Unit; IIT: Intensive Insulin Therapy; MD: mean difference; n: number of participants, TBSA: total body surface area. §: For risk of bias assessment, see appendix IV. §§: Jeschke and Jeschke represent the same cohort of randomized patients. In Jeschke the authors include some non-randomized patients as well. The truly randomized population are described in Jeschke. This information obtained after contacting the lead author of the trials. EE: Patients were randomly assigned to 1 of 4 groups: continuous intravenous insulin infusion with hydrocortisone alone, continuous intravenous insulin infusion with fludrocortisone plus hydrocortisone alone, or conventional insulin therapy with hydrocortisone plus fludrocortisone, conventional insulin therapy with hydrocortisone alone, continuous intravenous insulin infusion with hydrocortisone. $$: To convert to millimoles per liter, multiply by 0.0555. Thus, 80-200 mg/dL = 4.4-6.1 mmol/L and 180-200 mg/dL = 10-11.1 mmol/L

<table>
<thead>
<tr>
<th>Study</th>
<th>Protocol</th>
<th>Outcome Measure</th>
<th>Condition</th>
<th>Study Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young et al. 2009</td>
<td>Severe traumatic brain injury, adults, ICU</td>
<td>ITT (BG 4.4-6.1 mmol/l) vs. BG &lt; 11.1 mmol/l</td>
<td>6 months mortality: 52.1% vs. 53.4% (p=0.03)</td>
<td>High</td>
</tr>
<tr>
<td>Coester et al. 2010</td>
<td>Severe traumatic brain injury, adults, ICU</td>
<td>ITT (BG 80-110 mg/dL) vs. BG &lt; 180 mg/dL</td>
<td>Hospital mortality: 28.2% vs. 27.5% (p=0.3), Hypoglycaemia: 82.1% (IIT) vs. 17.5%</td>
<td>High</td>
</tr>
<tr>
<td>Jeschke et al. 2010</td>
<td>Severely burned ICU children (TBSA burn &gt; 30%)</td>
<td>ITT vs. control (not specified)</td>
<td>ICU mortality: 4% vs. 11% (p=0.14), ↓ rate of infection &amp; sepsis (p=0.05)</td>
<td>High</td>
</tr>
<tr>
<td>COSTS Study, 2010</td>
<td>Septic shock, adults, mixed ICU</td>
<td>ITT (BG 80-110 mg/dL) vs. BG 180-200 mg/dL, EE</td>
<td>Hospital mortality: 45.9% vs. 42.9% (p=0.50), IIT: ↑ risk of hypoglycaemia (MD 0.15, 95%CI 0.02-0.28, p=0.003)</td>
<td>Uncertain</td>
</tr>
</tbody>
</table>

**SUMMARY:**
Cochrane systematic reviews with meta-analyses of randomised trials provide guidance for clinical practice and health-care decision-making. In case of disagreements between research evidence and clinical practice, high quality systematic reviews can facilitate implementation or deimplementation of medical interventions into clinical practice. This applies especially to treatment of critically ill patients where interventions are most often costly and the clinical conditions are associated with high mortality.

**Objectives:** To assess the potential benefits or harms of 1) anti-thrombin III (AT III) for critically ill patients; 2) inhaled Nitric Oxide (INO) for acute respiratory distress syndrome (ARDS) and acute lung injury (ALI); 3) aerosolized prostacyclin for ARDS and ALI; 4) thrombelastography (TEG) or thromboelastometry (ROTEM) to monitor haemotherapy versus usual care in patients with massive transfusion.
Methods: We performed four systematic reviews of relevant randomised clinical trials. To quantify the estimated effect of various interventions, we conducted meta-analyses, where appropriate, to determine intervention effects using the Cochrane Collaboration methodology, trial sequential analyses (TSA), the GRADE, and the PRISMA-guidelines when conducting our systematic reviews. All reviews were performed according to published protocols following the recommendations of the Cochrane Handbook for systematic reviews of interventions. We performed multiple subgroup and sensitivity analyses with regard to methodological quality and various clinical outcomes. Trials were identified through Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE Science Citation Index-Expanded, The Chinese Biomedical Database and LILACS. We included all randomized clinical trials. We hand-searched reference lists, reviews, and contacted authors and experts for additional trials. We searched ClinicalTrials.gov, Centre Watch Clinical Trials Listing Service and ControlledTrials.com for missed, unreported, or ongoing trials. We screened bibliographies of relevant articles and conference proceedings and wrote to trialists and pharmaceutical companies producing the drugs in question.

Results: Four systematic reviews included a total of 44 trials with 5551 patients. Only 15 of the trials were classified as trials with low risk of bias (high methodological quality) regarding generation of the allocation sequence, allocation concealment, blinding, follow-up and other types of bias.

1) Compared with placebo or no intervention, AT III did not significantly affect overall mortality (relative risk (RR) 0.96, 95% confidence interval (CI) 0.89 to 1.03). No subgroup analyses on risk of bias, populations of patients, or with and without adjuvant heparin yielded significant results. AT III significantly increased the risk of bleeding events (RR 1.52, 95% CI 1.30 to 1.78).

2) INO showed no statistically significant effect on overall mortality (RR 1.06, 95% CI 0.93 to 1.22) and in several subgroup and sensitivity analyses, indicating robust results. Limited data demonstrated no effect of INO on duration of ventilation, ventilator-free days, and length of stay in the intensive care unit and hospital. We found a statistically significant, but transient improvement in oxygenation in the first 24 hours, expressed as the ratio of $P_{\text{O}}_2$ to fraction of inspired oxygen (mean difference (MD) 15.91, 95% CI 8.25 to 23.56). However, INO appears to significantly increase the risk of re- nal impairment among adults (RR 1.59, 95% CI 1.17 to 2.16) but did not significantly affect the risk of bleeding or methaemoglobin or nitrogen dioxide formation.

3) We found only one small low risk of bias paediatric trial examining the role of aerosolized prostacyclin in ALI or ARDS. Based on this limited amount of data, we were unable to support or refute the routine use of this intervention in ALI or ARDS.

4) Compared with standard treatment, TEG or ROTEM showed no statistically significant effect on overall mortality (RR 0.77, 95% CI 0.35 to 1.72) but only five trials provided data on mortality. Our analyses demonstrated a statistically significant effect of TEG or ROTEM on the amount of bleeding (MD -85.05 ml, 95% CI -140.68 to -29.42) but failed to show any statistically significant effect on other predefined outcomes. However, whether this reduction has implication for the patients clinical condition is still uncertain.

Conclusions: We did not find reliable evidence to support the clinical use of the assessed immuno-coagulatory interventions for general use in critical care based on the available evidence. A large proportion of the trials had serious methodological shortcomings, small number of patients, and short trial duration. The sparse data provided in the included trials may be or may not be promising but is not necessarily evidence of absence of a beneficial or harmful effect, because many of the outcome measures have not been adequately addressed so far. There is an urgent need for several randomised clinical trials with low risk of bias and low risk of random error to evaluate the use of the assessed interventions.

REFERENCES:


Anderson MR. Update on pediatric acute respiratory distress syndrome. Respiratory Care 2004; 28:237–244.


Anderson MR. Update on pediatric acute respiratory distress syndrome. Respiratory Care 2004; 28:237–244.


114) Beckman JS, Beckman TW, Chen J, Marshall PA, Freeman BA. Apparent hydroxyl radical production by peroxynitrite: implications for endothelial injury from nitric oxide and superoxide. Proceedings of the National Academy of Sciences of the USA 1990;87(4):1620-4.


244) Afshari A, Brok J, Møller AM, Wetteslejer L. Inhaled nitric oxide for acute respiratory distress syndrome (ARDS) and acute lung injury in


286) Chan et al. 2007. For more detail, see reference 31


309) Gotzsche PC 10977820. Why we need a broad perspective on meta-analysis. It may be crucially important for patients. BMJ 2000;321(7621):585-6.


