

Meeting report

The PIRO Concept: O is for organ dysfunction

Jean-Louis Vincent¹, Julia Wendon², Johan Groeneveld³, John C Marshall⁴, Stephen Streat⁵ and Jean Carlet⁶

¹Debate moderator and Head, Department of Intensive Care, Erasme Hospital, Free University of Brussels, Belgium

²Consultant, Institute of Liver Studies, King's College Hospital, Bessemer Road, London

³Associate Professor, Department of Intensive Care, Institute of Cardiovascular Research, Vrije Universiteit Medical Center, Amsterdam, The Netherlands

⁴Professor of Surgery, University of Toronto, General and Critical Care Surgery, Toronto General Hospital, Toronto, Ontario, Canada

⁵Intensivist, Department of Critical Care Medicine, Auckland Hospital, Auckland, New Zealand

⁶Service de Réanimation Polyvalente, Hôpital Saint Joseph, Paris, France

Correspondence: Jean-Louis Vincent, jlvincen@ulb.ac.be

Published online: 7 May 2003

Critical Care 2003, **7**:260-264 (DOI 10.1186/cc2196)

This article is online at <http://ccforum.com/content/7/3/260>

© 2003 BioMed Central Ltd (Print ISSN 1364-8535; Online ISSN 1466-609X)

Keywords infection, organ dysfunction, PIRO, predisposition, response

This report is based on the transcript of a roundtable debate held at the 23rd International Symposium on Intensive Care and Emergency Medicine (ISICEM), Brussels, Belgium, 18–21 March 2003. The participants of the debate were Jean-Louis Vincent (Brussels, Belgium), Julia Wendon (London, UK), Johan Groeneveld (Amsterdam, The Netherlands), John C Marshall (Toronto, Canada), Stephen Streat (Auckland, New Zealand) and Jean Carlet (Paris, France).

[Jean-Louis Vincent] Thank you very much indeed, and let's move to the fourth letter – 'O'. The concepts have evolved over time, and I would like to start by asking the panel whether they agree that we need to stratify organ dysfunction, perhaps with several other levels of organ dysfunction, or if they prefer to stick to respiratory failure versus no respiratory failure, renal failure versus no renal failure, and that kind of dichotomous separation.

My next question of course will go to the various organs, and we may ask, for example, whether we can improve the ways to diagnose neurological status, with the Glasgow Coma Score, and liver dysfunction, with bilirubin. Perhaps we also need to ask whether we could better evaluate gut function or metabolic function – should we speak about tolerance to feeding, insulin requirements and that sort of thing?

But let's start with the first general question – again, do you agree that there is a need for stratification or would you be happy with the more dichotomous separation? I think this question is particularly relevant when it refers to renal dysfunction because I've spoken recently to a number of intensivists with a special interest in renal function and they say that, as you have acute lung injury and ARDS [acute respiratory distress syndrome] in respiratory dysfunction, maybe we need to have a common definition of renal failure. So do we agree with this proposal?

[John C Marshall] Intuitively, I've always liked the idea of describing organ dysfunction as an intervention – in other words, ventilation or what you have to do to treat renal failure – having a sense that things are basically OK or things have completely failed or you're somewhere in the middle. So to me, the acute lung injury model is a nice reflection of that. So I think there's probably merits to having several levels. At this stage I think our understanding is that the more refinement we have in describing something, the more potential you have to get out of it. Obviously, continuous variables give you more information than dichotomous variables when you use statistical analysis, and at this stage we're still trying to generate information – we don't really know where the cut point is.

[Julia Wendon] I'm a very simple person. I think that organ failure is, just by looking at a patient, variable – mild, moderate, severe, however you want to define it. To my way of thinking, a lot of the things that define those organ failures go across them, so that something that might be important in renal failure might also be important in the cardiovascular system. I agree absolutely that we need to make them broader than they are at the moment; they need to include metabolism – we need to look at gut function in more detail than we have in the past. We must also take account of interventions, and the scoring system must take into account whether we have intervened yet, how much we've intervened and to what level.

[Johan Groeneveld] If I might continue along those lines and come back to how I feel about treatment dependency or independency and how to judge organ failure, the best way to do it in my view is to be an observer, independent, to rule out a bit of the treatment dependency of your assessments. For example, one physician could spend half an hour longer waiting for intubation and mechanical ventilation in a patient with respiratory insufficiency than another. So, to improve objectivity it may be wise to remove a bit from treatment dependency, rather than incorporating it into the assessment itself, except if you accept that it could be a different type of assessment, independent from your organ assessment.

The second thing that comes up following your original question – if you would favour a continuous level-wise versus a dichotomous assessment – I would say that is a matter of how you would prove objectively that the continuous level-wise assessment is more predictive – linearly additive – that one level is truly worse than the other level, in some objectively proven model. If you can prove that it's more predictive then I would go for the level-wise assessment. So it's highly dependent on what you want to predict as to how you would design it *a priori* for that purpose.

[Jean Carlet] I think we need some kind of global score giving an idea of the intensity of the severity of the disease. It's exactly what the people who developed SAPS [Simplified Acute Physiology Score] and APACHE [Acute Physiology and Chronic Health Evaluation] really wanted – to try to find a way to give a global indication of the severity. By the way, we can still use those – it's not necessarily organ system failures. But in addition I think we need information about the different kind of failures, because all are not comparable. I personally have a big problem to be convinced that putting shock at the same level as another one is appropriate – I am not sure, and this really worries me a lot. I was taught many years ago that shock is not only haemodynamic abnormalities but is far more – it includes the effects on the organs. So now if we define shock as just a fall in blood pressure, even if it's sustained, it's a problem. So I think we still have a lot of efforts to put in to describe those sorts of organ failures better, and I don't feel comfortable about the way we use it. It's important

because as you know the indications of activated protein C are now defined at a certain level of severity, and if you chose to select an APACHE score or the number of organs failing or the presence of shock, you will not necessarily pick up the same patients.

[Stephen Streat] I agree with Johan that theoretically one should seek a measurement system which is treatment independent, but for the reason that I've alluded to before about lead-time bias and the changing system it's practically impossible. And therefore I believe you'd need to have an equally objective measure of the treatment intensity. I agree with John's suggestion that, at the beginning, when one's trying to define whether a dichotomous or continuous process is likely to provide you with better information, you should start by collecting continuous information and seeking where the dichotomy should lie. My own personal bias is that we need something more than a dichotomy but much less than a continuous variable. Maybe we're talking about a 3-point or, at most, a 4-point scale; we're not talking about something which is enormous. I do share your view that we have focused on shock as if it were a cardiovascular problem whereas in fact it's a manifestation of a much more deep-rooted problem.

[J-LV] The therapeutic aspect is very important because, if I return to the reference to renal dysfunction, some people may use haemofiltration very early on, perhaps with high-flux filtration in the early stages of septic shock, whereas others may wait until there is a very positive fluid balance with a high urea or whatever before starting it. So just putting haemofiltration into the system may result in very different scores in one unit versus another. John, initially you were very reluctant to introduce any therapeutic variables into your score.

[JCM] My view has always been that these are two important related sides of a problem that give us different types of information in evaluating differing questions. So on the one level we have the question – is there a biological process going on here, is there a biological effect with an intervention, how severe is that biological effect? It's difficult for all sorts of reasons to find biology based on what the clinician does because we do different things with different biology. I agree completely that it's important to have a description of organ dysfunction that is, to the maximal extent possible, independent of therapeutic intervention. That was one of the original precepts of the MODS [Multiple Organ Dysfunction Score] score.

On the other hand I'm very much persuaded by the argument that it really doesn't matter what your $\text{PaO}_2/\text{FiO}_2$ [arterial partial oxygen tension/fractional inspired oxygen] ratio is, or what your creatinine is, or your urine output. If you're a patient what matters is, do you have a tube down your throat?, do you have a catheter in your groin? – it is organ dysfunction as

a clinical outcome. I think in any area of medicine we say that we've got a potential intervention – is there a biologic effect? And if so, is there a clinical benefit based on that? I think we need the next generation of organ dysfunction scores to integrate variables looking at biologic effect with variables looking at clinical consequences. That will not only tell us an awful lot about whether a therapy that works actually helps patients, but will give us wonderful insights into how we as a community of practitioners differ in how we treat deranged biology in patients, which actually gets back to the 'P' in the PIRO model.

This is a very interesting discussion because it's touching on some of the wonderfully complex issues. The other one related to this is, when is an organ dysfunction an outcome and when is it the disease that we're treating? In other words, in cardiovascular dysfunction, is shock the disease or is it the outcome of the disease that we're trying to measure? Is nosocomial infection actually a form of organ dysfunction or is sepsis the disease process that we're trying to treat? I don't know if it matters – but I think it's important that we address that question explicitly.

[Audience member] Don't you consider that organ dysfunction scores, APACHE scores, are about the same level – aren't we trying to define our populations a little better? Don't you think that APACHE scores do not tell you whether the kidneys are failing?

[JCM] We want both. Initially you said it's a way to describe the intensity of care and severity of the patient, but it's not something that is meant to predict mortality. But in fact it does. So we are always balancing between a system to describe the picture, the intensity including the intensity of what we are doing, and another score. Which is really the most appropriate to describe severity, so that we can predict mortality, at least in big groups? And we are always in-between – we never decide exactly what we want to do. In this case, we probably should decide. If you want to describe severity I am very convinced by Jean-Roger Le Gall, Stan Lemeshow, and Bill Knaus, that we should not put therapy in it. I am really convinced.

If you heard John's remark about physiological variables and therapies used, does it mean we should have a 2-level scale, with respiratory failure, $\text{PaO}_2/\text{FiO}_2$ ratio, and at level B, the type of respiratory support, for renal the degree of increase in creatinine ... Johan, you don't agree?

[JG] I like your idea, as always, but my feeling is that we should be more specific, we should design a system and see objectively whether that works in terms of predictive value; design your organ failure system separately from your treatment system. I'm not sure if there are data available on that score, but it should be looked into, in some large databases perhaps.

[J-LV] In the SOAP study, for instance, mechanical ventilation was a very strong predictor of outcome.

[JG] Independently of the respiratory SOFA [Sequential Organ Failure Assessment]?

[J-LV] Yes, independently of the respiratory SOFA. Now there may be some ethical questions associated with it, because perhaps there are not many patients dying in the ICU [intensive care unit] without a tube in their throat.

[JW] Doesn't that also raise the point – does the treatment actually just make the figures that we measure better or does it actually impact the organ dysfunction – and that's perhaps where even treatment needs to be separated somewhat. And even if we can't do it now, the aim of collating such data would be to enable us to answer that question over time.

[SS] This is what concerns me the most – the possibility of 'gaming' the scores by the therapies that you do. Now does that relate to outcome? Are you better predicted by the gamed or the ungamed score, or does it not matter? It's an empirical question but we don't have that information.

[J-LV] Maybe changes over time would be valuable. There was a paper from our group (Lopes *et al.*) in *JAMA* last year showing that the delta SOFA score is highly correlated with outcome – and actually we're using it now in our clinical trials.

[Audience member] But this is not treatment, it changes upon treatment – but what you're referring to is some additional treatment variables.

[J-LV] It's still difficult to separate a change in function in time from a change in therapy.

[JC] Since you mentioned this issue of trends, I think that the trends in the first 12 hours are really the key. So we should probably use those scores – SOFA and some other ones, it doesn't matter to me. There are some data showing that the trends in the first few hours or days are far more important than the absolute level at the beginning, and we should use it more.

[Audience member] There may be a drawback here because when you are approaching death it is far more easy to predict it than in the beginning, so this is a two-edged sword. The closer you get to death the more easy it is to predict. The change in SOFA score over a limited time period may be useful, of course it should be – more useful than the initial one. This is not unexpected though.

[J-LV] We are studying it on the data from the PROWESS study. It's a little too early to show you the data, we're still working on it, but it's amazing to see how we can predict

outcome in the first 24 hours. It seems that beyond the 24 hours the prognosis is already largely established.

[Audience member] When we spoke about markers before, some, if not all, of them are probably mediators, so I wonder whether in the search for a better definition of organ failure we should incorporate markers? For example, in the vascular compartment the elevated markers probably have the role of regulating the systemic part of the response of the host response, but when we do the same measurements in, say, a bronchoalveolar lavage we see different levels, and markers in different compartments do not necessarily correlate. So I wonder whether, with our definitions of ALI [acute lung injury]/ARDS, we could do a better job in defining respiratory failure by including parameters that we could measure?

[J-LV] There is a lot of biological variability there and I'm sure that our panelists will say that we are very interested in clinical parameters as well.

Julia – we are not quantifying liver dysfunction very well – we just look at bilirubin, etc. – but what can enzymes tell us? In many disease states enzymes can be elevated, so it's fairly nonspecific; are there any other simple tests we can do?

[JW] Simple tests are difficult, as you say. You can look at conjugated bilirubin – it's got to be conjugated, people don't even understand that. The transaminases go up in all sorts of diseases – it's hard to say it's liver specific. The hepatologists would say that we should look at albumin – but we know that that's pretty useless in critical care. If you're using a starch versus an albumin as your colloid or crystalloid you're going to get very different results, so I've got concerns there. No one in reality at the bedside is going to put hepatic vein catheters in to look at hepatic uptake or splanchnic blood flow. I think perhaps what we've got to do is to return to simpler things, maybe a lactate flux test. In the old days we used to use lactate buffers in sick patients – infusing lactate was a great way to assess hepatic lactate metabolism – but we don't do that any more. In addition, maybe – and I think we must do more work on it – we could think about using indocyanine green [ICG] by finger probe. It's a composite measure; it has its problems, but lactate and base deficit are composite measures and it doesn't stop us using those. So I think over the next few years we will see an expansion of liver assessment.

[J-LV] Any proposal for the evaluation of gut dysfunction? Of course the ability to tolerate feeding may depend on the type of patient, and perhaps your own protocols, but is there a way? Not really.

My next question is, should we include insulin requirements?

[JCM] I think the notion of incorporating some measure of endocrine function is intuitively a very appealing one. I'm not

an endocrinologist and I'm not sure what an endocrinologist would say about the notion that the entire endocrine system was one and we could equate insulin requirements to thyroid function to sex hormones to releasing hormones to adrenal status. Essentially, what one wants to do in describing organ dysfunction is to come up with a construct that looks like the dimensions of the patients that we treat in the ICU. So we think about a lung failing, a heart failing, about neurologic dysfunction, and that's one important way in which it differs from a generic severity measure like APACHE. Increasingly, the endocrine system is one of those systems, so I don't really know what the outcome measure would be.

One of the curious things you find when you do that is any one you look at is no longer significant when you look at it in the context of all the others. We've seen that with ARDS, for example, patients don't die of ARDS, they die of multiple organ dysfunction. But you can do that with any other organ dysfunction and it simply reflects the fact that, although we have to think about them in isolation, we are treating patients who are entire species.

[SS] Can I take a slightly different viewpoint? Going back to what you told us at the beginning about the independent deleterious effects of positive fluid balance in the SOAP study, I'm intrigued about the possibility of measuring something that is related to the disposal of intravenous fluid – the Trans Escape Rate of albumin, maybe, something to do with interstitial fluid compliance, something to do with the dynamic state of the solid-gel phase in the interstitial site – whatever it is that drives this behavior which is independent of vasodilatation and cardiac depression but is part of the syndrome of fluid disposal.

[J-LV] This is interesting but it gets into quite sophisticated analysis that Johan knows about well ...

[JG] It increases in sepsis but I'm not sure whether this is a bedside measurement with any predictive value. It has some pathophysiological significance but I'm not sure whether it could be incorporated into this system as an independent, easily obtainable predictive measure. So I'm hesitating.

[JW] Staying along those lines, it's not a direct measure of capillary leak. But if you've got someone who's leaking who's got a lot of fluid on board – perhaps also someone who also isn't tolerating their enteral feeding as well, even if you've got a post-pyloric tube down there as well so you don't know what's happening to that food in the small bowel – should we start using intra-abdominal pressure more as a prognostic indicator, and incorporate that into the definition of gut failure?

[JCM] That's a really interesting point; what about the delta between the intra-abdominal pressure and central venous pressure?

[J-LV] What about neurological function? We're still left with the Glasgow Coma Score – we haven't found anything better – and of course it's difficult in the patient treated with sedative agents because we have to use an 'assumed' score. Is there anything better that we can consider, any opinions about how to evaluate neurological function in ICU patients?

[JG] Even if the Glasgow Coma Score [GCS] has some drawbacks, I think it's a fairly well-validated system in terms of prognosis for a wide variety of neurological conditions. So before getting rid of a system and designing a new one or an additional one, we must be very secure about losing a validated system that has been used for so many years already.

[JCM] For all its weaknesses, we looked at the GCS in the MONARCH anti-TNF study, and the strongest signal of organ dysfunction was a change in the GCS. And that was in a blinded study where the GCS was being measured at 157 different sites, with people coming up with their best estimate. So it actually does work in practice, in a variety of conditions.

[JC] I think we can use many different scores – my problem is not to use a score rather than another one. I think we should look at trends and incorporate this into our models as soon as possible. The other thing I wanted to say is I'm not comfortable with using hypoxaemia if the initial disease is pneumonia, and I think this is a big issue. It will be a big issue for the ALI [acute lung injury] studies. So I think we have to say something on this. I am not comfortable with the level that we selected – 200 or whatever.

[JCM] Can I just make a point? One of the things that was in the original ACCP/SCCM [American College of Chest Physicians/Society of Critical Care Medicine] consensus paper from 1991–1992 was the notion that we need to think about organ dysfunction as primary or secondary. I forget whose idea it was, but it really got thrown by the wayside – the idea that primary organ dysfunction, for example pulmonary contusions or pneumonia, is different from a secondary pulmonary dysfunction in ARDS. Maybe that's an idea that should be revisited.