OPEN PEDIATRICS™

Less is More: Oxygenation Targets in Critically Ill Children

In this episode Dr. Mark Peters discusses the Oxy-PICU trial, published in The Lancet in January 2024, which compared conservative (88-92% oxygen saturation) to liberal (>94% oxygen saturation) oxygenation targets in critically ill children. The study highlighted the importance of pragmatic trial design and the need for larger trials to confirm these findings.

SPEAKERS:

Mark Peters, MBChB, PhD, MRCP, FFICM, FRCPCH

Professor of Paediatric Intensive Care

NIHR Senior Investigator

UCL Great Ormond St Institute of Child Health

Hon. Consultant Paediatric Intensivist

Paediatric Intensive Care Unit and

Children's Acute Transport Service

Great Ormond St Hospital London WC1N 3JH

Jeffrey Burns, MD, MPH

Emeritus Chief

Division of Critical Care Medicine

Department of Anesthesiology, Critical Care and Pain Medicine

Boston Children's Hospital

Professor of Anesthesia

Harvard Medical School

Initial publication:

November 25, 2024

Sarah Marcley *00:04*
This podcast is a production of OPENPediatrics, an open access online community of healthcare professionals sharing best practices from around the world. Visit openpediatrics.org for more.

Jeffrey Burns *00:18*
Welcome to the World Shared Practice Forum on OPENPediatrics. I'm Dr. Jeff Burns, Emeritus Chief of Critical Care at Boston Children's Hospital and Harvard Medical School. We're very pleased to have with us today Dr. Mark Peters. Dr. Peters is professor of pediatric intensive care at University College London. He is also consultant in pediatric intensive care at the Great Ormond Street Hospital in London. Mark, welcome.

Mark Peters *00:42*
Hi, Jeff, thanks very much. Thanks for the opportunity to talk about our work.

Jeffrey Burns *00:45*
Mark, as you know, I've been interested in having you on this podcast for a year or so, at least, and it was really to catch up with you on your study published in Lancet in January of 2024 entitled The Oxy-PICU Conservative Versus Liberal Oxygenation Target Trial in critically ill children. And Mark this was clearly a noteworthy study in our field, and it's been widely commented on. And I wonder if we could begin by asking you, what was the background that drove the hypothesis to test this as a randomized, controlled clinical trial?

Mark Peters *01:25*
Thank you. Great question. It's worth noting that a trial like this has a long history. It has at least 10 years of build up until that paper is published. But the origin was some work I did with colleagues from Southampton University and adult ICU at UCL in a trip to the Himalayas in 2013. And the essence of that was to investigate adaption to hyperbaric hypoxia. And it's clearly true that humans can function at very low levels of ambient oxygen. They can climb massive mountains, after all. And this group had previously taken a femoral arterial stab at 8400 meters, and demonstrated that a number of the investigators had arterial oxygen tensions of less than 30 millimeters of mercury, and yet they were still climbing. You could argue their judgment might have been flawed if they were letting someone take a femoral arterial stab at that stage, but they were clearly functioning. And that formed part of a literature that was coming around at the time, that we may be underestimating the risk of high oxygen in some observational studies, but we may be also overestimating the risk of low oxygen. And so from that discussion, appears a U shaped curve. Our physiology is packed with U shaped curves and the characteristics of most of them in intensive care rules, we don't know where the lowest risk sits. And as soon as investigators start describing U shaped curves, then clinical trialist comes along and tries to identify two points on that curve to compare for clinical outcomes. So that was there as a sort of theory. About seven or eight years ago, a number of the adult investigators started to make comparisons. But I think the next thing that made it really clear this was feasible in pediatrics was some observation work we did look at current clinical practice. And we could demonstrate that certainly in UK, pediatric

intensive care units, the modal peripheral oxygen saturation values that were observed were 199%, so very much at non physiological oxygen saturation. So there was an extreme of practice that had become normalized, and helpfully, when we compare that with what was existing in a number of the guidelines, including ARDSnet recommendations, which included in severe disease, consider peripheral oxygen saturation of 88 to 92% or the European mechanical ventilation guidance published in Intensive Care Medicine gosh, about seven or eight years ago now, also suggested that in severe disease, at least, we should consider lower oxygen saturation targets. So that's a perfect setup for a trial. You have clinical practice at one end, guidelines at another, and highly variable clinical practice. So when you have that combination, it is relatively easy to persuade a funding body that this is an area of clinical uncertainty that's lacking data.

Jeffrey Burns *04:25*
And Mark, could I just ask a brief follow up to that? I see your study as another seminal study in our field and in critical care in general. Less is best. You and I started training many, many decades ago, and at that time, you know, there was this concept of supra-normal oxygen delivery. Shove more oxygen at the mitochondria, and things will get better. And then we lived the era of normalization of values, and we certainly injured a lot of lungs by normalizing gas exchange. But now where are we on this field? Are we overshooting and and testing too much less is best, or are we actually just following the data step by step, and less is best? How would you put that in context?

Mark Peters *05:11*
So I think that's that was part of the thinking that made this an appealing option, not just for me, but for the UK group of pediatric intensivists to look at that the less is more version. That maybe were overdoing this and this and that came largely from the strength of the observational data that associated high oxygen with worse outcomes. But in the more general case, I think I'm right in saying that the only positive critical care trials in the last and 15 years have been of a less intensive treatment outperforming or being equivalent to the more intensive treatment. So that was certainly in the background in our thinking of designing this study. What else could we do less of. But obviously we need specific data relating to oxygen targets to make that credible.

Jeffrey Burns *06:05*
Could I [ask] you now? Could you take us into the trial design? Since this was a randomized control clinical trial at 15 participating centers who all contributed a significant number of patients to the final data analysis throughout the United Kingdom, and it was a pragmatic trial. Take us through that Mark. What is a pragmatic trial?

Mark Peters *06:27*
So a pragmatic trial attempts to model the intervention in the real world. So it's trying to understand what would happen if a guidelines changed and clinicians then went off with this idea and practiced with that as a starting point for their care. So what that means is you have to have inclusive inclusion criteria. It has to apply to a range of patients. And the strength of that is, it applies to a range of patients, and therefore the potential reach of the study, if you get a signal is greater. The downside is almost certainly some of those patients will benefit to a greater extent, and some to a lesser extent, and some may even be harmed by the intervention. So you're trading off the signal to noise or the eventual reach of the study. And what that means, in design terms, is, means you have to do quite a large study. Maybe it's beyond what the scope of this podcast, but there's a fabulous simplification of all of clinical

trials into one equation that you can you can express the confidence in a trial result on one side of the equation equals the ratio of the signal to noise times the square root of the sample size. The confidence equals signal divided by noise times the square root of the sample size. So we're accepting in this design that we're going to have a noisy trial. Because there are patients with all sorts of underlying conditions, with all sorts of acute conditions, but we're going to try and brute force the answer by the sample size. Just to restate that the advantage of this is, if you see a signal in this mixed bag of patients, it's likely to be real. Okay, you're less likely to see a signal because your signal to noise ratio is what it is. If we had done this purely in ARDS, for example, where perhaps the lessening of ventilation settings that you might achieve with a more conservative oxygen target, perhaps more directly, is on the causal pathway to outcome, then perhaps we would have seen a bigger signal. But the problem is, only very few of your patients have ARDS, so the impact of that study would be less.

Jeffrey Burns *08:43*
And so that leads me to ask, could you remind us what was the inclusion criteria? And importantly, what was the exclusion criteria of the children that you studied?

Mark Peters *08:52*
Okay, so the inclusion criteria were that these were emergency admissions, accepted to a participating pediatric intensive care unit. And you've already mentioned there were 15 of these, geographically spread around the UK. You couldn't be preterm, and you couldn't be after your 16th birthday, because that's the population that comes to the PICU in the UK. And then you needed to be recruited or randomized within six hours of meeting all of the following criteria. So accepted to the unit, receiving invasive mechanical ventilation via endotracheal tube or tracheostomy, on supplemental oxygen, and you're physically in the same room as pediatric intensive care or transport staff. And so those are the inclusions. There were quite a few exclusions, but the headline ones, if you like, are that we didn't include patients in whom brain pathology or injury was the primary reason for PICU admission. Now we didn't include patients with known or suspected uncorrected congenital cardiac disease or pulmonary hypertension. There was other, some sickle cell or previous recruitment, or in cases in which death was perceived as imminent, or that's an end of life care plan in cases. So there are, you know, those are more generic across all trials, but the brain pathology and cardiac disease or pulmonary hypertension has specific exclusions for this trial.

Jeffrey Burns *10:20*
And Mark, could you take us through the final salient data, how many achieved the target range and how many were in the liberal range? And could you remind us of what was the target conservative range?

Mark Peters *10:35*
Okay, so just to be—it's really quite an important point that the intervention was the intention to hit a target range, except you can't always hit it. So the intervention was aiming for a peripheral oxygen saturation of 88 to 92% inclusive. And the standard pair group, if you like modeling or trying to mirror, what was being used widely in the UK, was peripheral oxygen saturation of greater than 94%. And we did not specify how clinical staff should change the ventilator or the oxygen in order to hit that target range. Ventilation practice is extremely varied, and trying to tell people what to do with the ventilator would have been a bigger intervention than telling them what target to try and achieve. It would be quite difficult to summarize the adherence data without a number of graphs, but essentially, we did achieve

separation between the at the time weighted averages of the peripheral oxygen saturations. But even in the low group, the observed oxygen saturations were frequently above the target range. The majority of those patients were breathing air, and it didn't seem reasonable to expect our colleagues to provide a hypoxic gas mix to get them down into range. So it was what was achievable on existing ventilator setups.

Jeffrey Burns *11:59*
And the outcomes were?

Mark Peters *12:01*
So the primary outcome we chose for the study was a composite of the number of days of organ support and mortality. And our thinking behind that is that mortality is such a rare event that it would be impossible to power the study on mortality alone. But actually the combination of how long you're on a machine on organ support and mortality is what families, and as patients told us they really cared about. And so what we did was we created a rank-based outcome of a composite of days of organ support and mortality, with mortality being the worst score, and then every number below that being the number of days of requiring organ support. And then we compare the probability of a random comparison between an intervention patient and a control patient, of having a better rank on that ranked outcome. It's a slightly unusual way of doing it, but actually it generated important additional power, and we think important additional relevance, because it describes both the severity of the course and the mortality risk.

Jeffrey Burns *13:07*
And as I recall, the salient outcome in terms of the number needed to treat adhering to the conservative oxygen target of 88 to 92 would lead to one less death out of 200 patients. Is that correct?

Mark Peters *13:26*
Yeah. I mean, that that's roughly right. For every 200 patients managed, we would expect to see one fewer deaths and 123 fewer days of organ support. And because in the UK, we well, in every healthcare system, we count the cost, but we're particularly attentive to the cost in the UK, it's about 400,000 Sterling saved for each 200 patients treated. So we scale out to the UK nationally, that would be 50 fewer deaths, 6,000 fewer days of organ support, and about 20 million Sterling saved. The way we presented the primary outcome measure was that something called a probabilistic index. So the probability of a better outcome in the intervention group compared to the control group, and there's a 6% difference, essentially. 53% chance of a better outcome in the intervention group versus 47 in the control group. And you might well say, well, that's a small effect. I think it is. It makes sense to me that the effect is at that scale. If there was an obvious harm in an individual patient of running at a higher or lower oxygen we would likely have learned this in our observational data sets before now, so it's likely to be a relatively modest scale. And the other point about is we said that the outcome measure was a composite of both organ duration, organ support duration and death, both the elements of that composite improved in the intervention group, which is an importance of internal validation. You're not looking at something that saves some people and harms others.

Jeffrey Burns *15:11*
And you're going to have to forgive me if my simple mind remembered basically the mortality benefit. And I certainly did not mean, because I agree with you, the significant burden of disease in terms of

organ support, on the child, on the family and on society, because it's our obligation to utilize these very expensive resources in the most safe, effective, efficient way possible. And so I agree, I think that that's an extremely salient finding from your study. Before we move on to where do we go with this, two questions on both the design, is there anything you would do differently on the design in retrospect? And secondly, what were the limitations? Or what have you heard from colleagues about what they see as the limitations of the trial as conducted?

Mark Peters *16:03*
Thanks. It's a great question. So no trial is perfect. You always wish you'd done things differently. Our choice of fields for the organ support element included some elements which are relatively low impact, including continuous infusion of a sedative agent, for example. We chose them because they were already collected and established in a national mandated registry called PIC Connect. It had a history. It had a validation and that supported them. What it leaves the trial open to criticism is that some small differences in sedation duration might have driven the differences you see in the study. Actually, if you dig into the data, most of it is driven by the ventilator time to extubation. So that probably isn't true, but that's something that we would I would do differently if I were doing this now. I would also have done a larger study, because we didn't know at the start that we would have been able to recruit ahead of time and target, and then we would have improved. I mean, the study was adequately powered. We achieved the primary outcome measure and we achieved statistical significance. But we only just achieved statistical significance. You could always argue whether that means your trial is perfectly powered and it hasn't wasted resources, or whether it's less compelling a result. There's always an element of estimation, again, to a power calculation. Had we done 3,000 patients rather than 2,000 we, I think, would have had a greater confidence in that relatively borderline result. But that's, you can't know that when you start really.

Jeffrey Burns *17:46*
I know you did a wonderful point counterpoint talk, but also in PCCM, Pediatric Critical Care Medicine. What have been the challenges to the study?

Mark Peters *17:57*
Yeah, so a number of commentators have said you didn't hit the—I think they perhaps misunderstood, or we've inadequately explained that the intervention was targeting this saturation, not necessarily hitting it. And so some commentators have said, well, clinicians know which patients it's safe to do this in and which aren't. And the ones in which you're non-compliant, meaning the saturations sitting in their middle 90s, say, intervention group and the ones in which clinicians have rescued them from the dangerous intervention. I think that's guesswork. I think, remember, most of this practice is controlled by the bedside nursing staff adjusting the inspired oxygen fraction, and these are busy people. They have many tasks, and they like goals to work with, and I think it's relatively unusual for the clinician to override that without a separate justification, so I don't accept that as a criticism. And if people do believe that they should do their own study to, to refute what we've found is my sort of simple response.

Jeffrey Burns *19:05*
Mark, where do we go next? What's your follow up to the study? To repeat it in other environments? To see if it's generalizable? To do a larger study? Or are you—and I know you're moving on to other things as well—but where do we go next with this hypothesis and these data?

Mark Peters *19:23*
So there are every decent bit of science generates more questions than it answers. I'd love to know mechanism. I'd love to know the optimal because we only compared two points on the curve. It doesn't tell us what's the lowest risk on the curve. But actually, if you're designing research trials, you have to go for where you think the greatest impact is. And to repeat this trial with two different targets would be very challenging in the likely scale required to inform on the difference between, let's say, 88 and 92 and 93 to 97. There's less clear blue water between the groups, therefore the effect size is likely to be smaller. Therefore the power would be a problem, and the sample size would be much larger. So our next, we've actually moved on. We're doing a blood pressure targets trial called PRESSURE. We're doing studies of deimplementation of gastric residual volume measurement. Both of those are more than halfway through. And we are in the process of trying to examine a series of other routine interventions in pediatric critical care, with the aim of giving everybody a starting point for refining practice, for example, for sedation focused entries and thresholds for fluid administration and deresuscitation. We accept that these are noisy answers that are average treatments affected across large populations, but that's probably more valuable. Will have greater impact than trying to refine further an oxygen target study. And the other part of that is that there are two very large adult trials of oxygen targets that will be completing in the next 12 months. UK-ROX, which is 16,000 critical adults, that's run by Dan Martin, I think it's nearly completed recruitment. And the MegaROX, which is led from ANZICS group Paul Young from New Zealand, which I think is 40,000 critically ill adults. And so there will be a signal of some sort from those trials, which we will have to consider before we go back into oxygen again. I think it would be unwise to try and design a study without sight of the results of those very large trials.

Jeffrey Burns *21:41*
Well, I know there was a study of adults with head injury published in JAMA just last week, a non-inferiority, which did not find any harm in the lower oxygen cohort as I recall it. Am I right about that?

Mark Peters *21:56*
I think so. I thought you were going to point out to the HOT-COVID study, which is the adult trial, has similar protocol to ours, which showed a similar advantage to ours in critically ill adult patients, manages conservative oxygenation with COVID ARDS. And how you interpret those data depends perhaps on how you you feel about our trial. We observe that's the only adult trial that has used a similar outcome to us, both the combination of organ failure, duration and mortality, and they saw the same effect. The majority of the other adult trials have used mortality alone and haven't seen an effect. So it could be just a methodological similarity that's allowed it to work. Yeah, there are several oxygenation trials in head injury. I've not seen one that shows a clear signal. Interestingly, Paul Young's pilot data, so that's the ICU-ROX that had 1,000 patients from Australia, New Zealand. They had a hint of a signal in favor of low oxygen in post cardiac arrest care. And so this field is shifting. And actually reference back to the question you asked earlier about what might we do differently? I think the equipoise has changed to the point we could have included neurological injury in our inclusion criteria now. When we were starting this study, everybody was a little bit more nervous of including CNS injury, but I think that equipoise has shifted so it could be done on a broader population.

Jeffrey Burns *23:31*
Could I ask you, do you think there are potentially responders and non-responders to low oxygen therapy?

Mark Peters *23:39*
Yes, definitely. And that's kind of a mechanistic question, isn't it? So I think we haven't touched on why this might work, and there are two main candidates in mind that this might work. Essentially, it might work on the patient, or it might work on the medical staff. And if it works on the medical staff, it could just be as simple as if you're allowed to be gentler with your patient, you're happier to stop sooner. If they're already in your target range, you may be comfortable to pull the tube out because the oxygen inspired oxygen fraction is lower. But there is quite a credible biological mechanism by which lower oxygen concentrations in the face of systemic inflammation, causes less oxidative injury. And we did a very small methodological study that was previously published in PCCM, Gareth Jones was the first author, showed there is a difference in the redox status even at these relatively small differences in oxygen targets and some of the defense mechanisms you see in hypoxia. So hypoxia inducible factor gene expression increased in the conservative group. So both are in play, I suspect so the question is in answer to your who's a responder question, you have to perhaps qualify by which mechanism. I think that the medics effect is available to the whole population. The oxidative injury effect perhaps matters more to those who are more seriously ill and have more—are struggling to resolve their organ failure. And with the eye of the enthusiast, if you look at our results, we present a spectrum, if you like, of where the organ failure distribution has shifted a little. You could interpret that, but what seems to have happened is the very short stayers stay even less time, and the very long stayers and the deaths are also slightly less common in the intervention group. But the middle of the population hasn't shifted much, and so I'd like to suggest that that means that the doctors are more comfortable to discharge patients sooner at one end, they're less than well patients. And there is also a population who are super sick, who are there for a long time, for almost a month, in multiple organ failure, who also have a small benefit. So we're going to try and do some work to characterize the subgroups that benefit more and benefit less. But just just eyeballing the data, there's a hint of that that might be present.

Jeffrey Burns *26:15*
It's fascinating for you. My personal belief is that our field pediatric critical care, but critical care in general has a long way to go on implementation science, you know, we do this great research, and, you know, there are signals. No signal is perfect, as you said, but there are signals. And then we, we kind of move on to the next. We don't pause and say, let's make sure we're implementing, you know, these, this hard won knowledge. In your own practice, first and then secondly, throughout the UK, has there been any change in practice that the standard of care should be the adoption of 88 to 92 unless special populations are at play?

Mark Peters *27:00*
So yes, in my own practice, I've definitely changed. My staff are very aware that I'm attentive to 88 to 92 and the majority of my colleagues follow that also. Around the UK, numerous colleagues have also changed practice and tell me so. But actually, the step to make this stick will be its inclusion in guidelines and the surviving sepsis for children guidelines are in the in sort of latter stages of preparation, and that's one of the first opportunities that will have to be through a wider review and appear. So, you're quite right, implementation science is crucial. It's not something we have extensive experience of, part because we have so few clear signals from our clinical trials. And we're continuing to explore how we do make it stick.

Jeffrey Burns *27:53*

Can you give us 30 seconds out what the PRESSURE trial is before I let you go?

Mark Peters *27:56*
I can, but if I but 30 seconds, this is not just my work, it's a huge group of clinical trials specialists and the pediatric critical care society study group UK. So, you know, just so happened I was lucky enough to be the name on this grant, but it's a huge group contributed. So the PRESSURE trial is essentially a very similar design. It's of age specific blood pressure targets percentile versus usual care in patients who are receiving vasoactive medications as part of an emergency admission to the acute. We're aiming to recruit almost 1,800 patients. We're at about 1,200 now. So we would hope to have an answer with your David Inbold as the chief investigator, would hope to have an answer for you within about 18 months.

Jeffrey Burns *28:44*
Terrific. Professor Mark Peters, University College London, Great Ormond Street Hospital London. Mark, thanks for being here today. You know I speak for colleagues around the world that you are clearly one of the thought leaders in our field, and that the clinical investigations that you've done to date, and the work that you continue to do is really just so impressive. You are helping to advance our field, and so we thank you for everything you're doing, and thanks for coming today to our podcast.

Mark Peters *29:11*
Thanks, Jeff. It's very kind. It's not just me. It's a huge team.

Sarah Marcley *29:15*
This has been a production of OPENPediatrics. You can find the resources and journal articles referenced in this podcast in the description. We have more podcasts like this one available everywhere you get your podcasts. Visit openpediatrics.org for more information.