On the Paclitaxel Meta-Analysis: A Conversation With Konstantinos Katsanos, MD

Dr. Katsanos discusses the methods and findings of his group's meta-analysis, the patient-level data presented in response, and potential next steps.

What led you to look into the possibility of increased mortality in patients randomized to treatment with paclitaxel delivery?

We have been using paclitaxel-coated balloons and stents since they first appeared on the market, and we have been very active in researching the space from the very beginning. With the accumulation of enough literature, we noticed that there was a signal in a number of trials and with different devices. We also noted that this signal was maintained in the long term, with the IN.PACT SFA (Medtronic) and the Zilver PTX (Cook Medical) randomized trials for example, which led us to look into this in a more systematic way.

Were there any such signs in your own practice?

No, but we did not conduct a controlled clinical trial to compare against plain control balloon angioplasty.

How did the authors decide on the data set inclusion criteria, and what were the challenges in making that decision?

That was actually an easy decision. To maintain the quality and rigor of the analysis and avoid confounding, we decided to include randomized controlled trials (RCTs) alone. We included all RCTs investigating any kind of paclitaxel device—a balloon or a stent—regardless of different formulations, doses, etc.

Can you tell us more about the decision to include data that were presented but not yet published, and if there was any concern as to "presentation bias?"

Yes, that's a very good point. To make sure that we included all available information and that the meta-analysis was contemporary and to increase the sample size and power, we decided to include presentations from congresses as well. The majority of those presentations were industry-sponsored studies. Even when we looked into publication bias, statistically, there was no evidence to suggest we over- or underrepresented certain parts of the evidence.

As a follow-up regarding publication versus presentation, the IN.PACT SFA data included were presented but unpublished 4-year data, and they were included in a "long-term" pool with published 5-year data. How did you decide to pool the 4- and the 5-year data together to create long-term data?

Those were the only data available at the time [of data collection]—the 4-year data from IN.PACT SFA and the 5-year data from THUNDER and Zilver PTX. So, we did pool those together to represent the very long-term time point.

Did you consider waiting for 5-year data from Medtronic?

The signal and the findings from the preliminary analysis and from the 2-year analysis were too strong to wait for a single trial update. I think it would have been unethical to have waited.

One of the primary criticisms of the meta-analysis was that it postulated a causal link, that of late paclitaxel toxicity, despite the design involving only study-level data. How do you respond to that concern?

Any meta-analysis, by definition, can be done at the study level or at the patient level. The best ones are those with individual patient-level data, but the starting point is always study-level data. We recognize that all of the studies were designed to investigate a treat-

ment effect in terms of patency, limb outcomes, freedom from target lesion revascularization (TLR), etc, and like most RCTs, they were not designed with a safety endpoint in mind.

I would argue that the value of the meta-analysis is the ability to pool the study-level data of individual trials that have been a priori underpowered to detect safety signals in order to increase sample size and be able to determine a more reliable signal in the pooled patient sample. Nobody in their sane mind would start a large randomized study to test this kind of hypothesis. In medicine, safety signals most often follow effectiveness signals. When you are working on a safety signal, you start by working with the study-level data.

However, despite this being a study-level meta-analysis, the baseline patient and lesion characteristics (also known as confounders or risk modifiers) were very well balanced between studies and also within the individual studies themselves. We presented a meta-analysis of a quite uniform and homogeneous patient population that mostly had intermittent claudication. So, even in the absence of patient-level data, I think that the quality of this study-level meta-analysis was good enough to show the validity of the signal.

The absence of a clear biological mechanism does not dismiss the validity of the information. We accept that paclitaxel reduces restenosis and TLR because we have controlled for all other factors (within randomized trials). Why would we accept that but not the higher death rate? We cannot violate the basic rules of modern epidemiology and evidence-based medicine.

Do you agree with the suggestion that meta-analyses, in general, are not designed to prove a hypothesis or causal relationship?

This is a question that has to be answered within an epidemiologic framework. If the proposed question is about efficacy, the meta-analysis usually increases the power, reduces the uncertainty around the treatment effect, and confirms the effectiveness of a certain intervention. However, as I said, these studies were not designed to test safety hypotheses, and if the question regards a safety endpoint, a meta-analysis is the first step to determine whether we need further studies. I do agree that this has to be interpreted as "potential" causality to inform further exploratory analyses like adjusted individual patient-level data comparisons next.

Can you further explain how a collection of trials that are not designed or powered to show long-term safety can be grouped to then power an analysis that does?

This is the beauty of a meta-analysis and the essence of the science of statistics—you may combine individual data that come with some uncertainty, and the pooled data have significantly less uncertainty, such that you can rely on these data more than that of the individual study. Every study's treatment effect comes with its own sampling limitations and statistical uncertainty. By combining the individual treatment effects from numerous studies, we come out with a single effect that has reduced uncertainty around the pooled mean effect size. In other words, a meta-analysis will pool the individual, underpowered, diverse, and uncertain effects into a single more solid, more certain effect. And, we use random-effects models to account for variations (in study design or population heterogeneity) that may not have been previously identified.

But, it is also true that every meta-analysis is only as good as the studies that it combines. If you start with bad apples, you are going to get a rotten apple pie.

Was there any hesitation to group the drug-coated balloon and the drug-eluting stent cohorts?

Yes. I did discuss this internally with my team time and again. At the end of the day, the overarching question was whether paclitaxel is safe or not. We decided to pool the studies because paclitaxel was the key differentiator between the active and control arms. Having said that, within the published manuscript, we have done what we call sensitivity subgroup analyses and have calculated the pooled treatment effects and risks of death for the stents separately from the balloons and the same for subgroups of the individual drug doses.

What is your overall impression of the data that were presented at LINC in support of the safety of paclitaxel?

The companies need to be congratulated on how quickly they collected an impressive volume of data, because the community needed this information. Of course, individual companies reported on their own individual data. I have reservations and questions about the statistical approaches and the presentation mode of certain studies. These questions remain to be answered once the studies are published in the major journals, when we can scrutinize the finalized data in detail. For example, there was a study that violated the intention-to-treat principle. There was a study that combined patients with different lengths

of follow-up into a single curve. At the end of the day, we didn't get answers on the specific time points for all the studies. Some studies included these answers, but not all devices showed the exact comparison of deaths at the specific time points that we have published. And that remains, in my eyes, significantly different.

What are your thoughts on the potential confounding aspect of control patients who are later treated with paclitaxel? In other words, after randomization, they later required revascularization and were treated with a paclitaxel device.

By definition and per COCHRANE recommendations, the meta-analysis maintains the allocation of the subjects to the treatment that was initially allocated to them at the time of randomization. The meta-analysis cannot look into what happened at the TLR if the patient has another intervention 9 months or 2 years later. The intention-to-treat principle is the foundation of RCTs, wherein if the patient is initially allocated to treatment A, they have to belong to treatment A for the duration of any kind of statistical comparison. However, I accept that follow-up treatments introduce further confounding (especially if they involve application of paclitaxel) when analyzing long-term follow-up at 5 years.

Patient-level data evaluations presented at LINC **found no statistically significant differences between the doses of paclitaxel and increased mortality, in some instances even showing better survival in patients who had higher dose ranges. How do you interpret those data in light of your study-level data regarding dose effect?**

First, the subgroup regression analysis within our meta-analysis is only exploratory. In the subgroup tests, we try to make sense of the associations and relationships in a quantitative way.

We have shown that increased initial dose relates to improved freedom from TLR, but also a higher death rate—that there is a highly significant linear relationship between freedom from TLR and death rate. In other words, the more effective paclitaxel is, the more deadly it may be. We will be presenting new information on dose-response models in the near future.

I also think that some of the presented analyses have produced mostly misleading conclusions. Taking the Medtronic analysis (recently published in *Journal of the American College of Cardiology*) as an example, an independent analysis was performed with the conclusion that no link could be found between paclitaxel and mortality. However, the dose-response tests were performed within the active arm alone, excluding the referent control arm that did not receive paclitaxel, hence reducing a randomized study to a single-arm analysis. A formal dose-response test would be expected to test the odds or risk ratios of death for different doses of paclitaxel.

Do you believe it is more so the dose of the drug and not the drug itself that leads to increased mortality risk? Do you have firm thoughts on the threshold at which dosage is safe?

We believe that the mortality risk is associated not only with the actual drug dose delivered but also with the exact properties of the medium of delivery. The latter relate to whether a balloon or stent platform is used, to the chemical properties of the excipient, and above all, to the exact paclitaxel formulation applied (the more crystalline it is, the longer the tissue half-life because of poorer solubility).

Some have criticized the means of calculating dosages administered, raising the possibility that the dose/time relationship may have overestimated the drug exposure, and that the equations should have differed based on the nature of the devices (ie, the varied surface areas of stents vs balloons). How would you address these concerns?

This is a valid concern. However, meta-regression analyses are exploratory by definition, and we pursued to include both time and dose in a single function. Again, we will be presenting more new data on proper dose response models shortly that demonstrate the very same findings.

What do you think of the alternative hypothesis that follow-up compliance, which was shown to be higher for percutaneous transluminal angioplasty in the IN.PACT study patient-level analyses, may be associated with lower mortality?

This is a finding that has been raised by one study out of the 28 we included. In our meta-analysis, we did sensitivity tests—for example, we removed the IN.PACT SFA study—and even without the IN.PACT SFA data, the rest of the studies still produced a highly significant risk of death at 2 years. Theoretically, I accept that there may be differences in compliance,

medications, etc, but I do not think that those subtle differences may fully explain the absolute magnitude of the difference we observed—a nearly 3% excess risk of death at 2 years. We accept that paclitaxel produces better TLR rates because we have controlled for all other factors. Why do we accept that but not the higher death?

How do you think paclitaxel should be evaluated in currently enrolling trials and those yet to be designed?

I think currently enrolling trials need to continue with some caution, of course, and with the appropriate ethical responsibilities and information being communicated to the patient. For new studies, this is a question that is going to be very difficult to answer. If I were to design a randomized study today, it would be difficult to choose the comparator. I don't really know.

Have you performed or reviewed any research on applications of limus drugs in the periphery? What are your thoughts on whether devices delivering another antiproliferative drug should require the same level of scrutiny and consent of ongoing trials involving paclitaxel?

With the exception of the SIROCCO RCT that tested a sirolimus-eluting stent in the superficial femoral artery, there is not much else regarding limus-coated or -eluting devices in the periphery. I understand that several companies are on the run now to get limus-coated balloons and stents for the superficial femoral artery. It will definitely be an interesting development; however, the expected standards now have clearly changed, and limus devices will have to be put to even more stringent tests (eg, longer-term RCTs) before deciding on their safety and effectiveness.

Do you think VIVA's Vascular Leadership Forum initiative will yield meaningful data?

The VIVA Physicians initiative is very welcome. They have promised to maintain independence and consult with an independent group to address any concerns of conflicts of interest. My only suggestion is that this safety interrogation must be combined with the effectiveness of the devices. This is really critical, and nobody has discussed it, but there is no point in combining data in terms of safety if those devices are not being compared on the basis of their effectiveness as well. In medicine, every drug or device has an efficacy/safety profile, and there is always a trade-off to some extent.

What should be the next focus of paclitaxel-related study?

First, I do not think that we will go back to the Stone Age of uncoated balloons and stents. Some kind of antirestenotic drug is necessary for superior results in peripheral artery disease (PAD) treatments, and this is here to stay in some form or another. Nonetheless, in terms of the safety concerns, I think we need to also check paclitaxel outside of PAD, and we need to see relevant evidence in other vascular fields, such as in the coronary vessels.

Based on the data from your meta-analysis and those presented in response, what do you currently think are the appropriate scenarios for using paclitaxel delivery devices in PAD, and how will you decide in your own practice?

Since our findings, I have restricted my use of paclitaxel-coated balloons and stents to the absolute minimum. I'm not treating long lesions because that would amount to a higher cumulative dose. I am not using the higher-dose devices because that amounts to a higher cumulative dose. I am trying to restrict this to the absolute minimum, no more than one device per patient and preferentially lower-dose devices. ■

Konstantinos Katsanos, MSc, MD, PhD, EBIR
Assistant Professor, Interventional Radiology
Patras University Hospital
School of Medicine, Rion
Patras, Greece
katsanos@med.upatras.gr
Disclosures: Over the last 12 months, personal fees from Medtronic and Boston Scientific Corporation; grants from Medtronic; all outside the submitted work.