Subtitles and transcriptions

Subtitles and transcriptions are available for selected materials for purpose of helping users understand the contents of the educational sessions.

Uncertain words have been indicated with ?? before and after the part. Parts that could not be understood at all have been indicated as [Audio Not Clear].

Every effort has been made to faithfully reproduce the audio of the sessions as recorded. However, no responsibility is accepted for mistakes or omissions. ESO does not endorse any opinions expressed in the presentations.

Evidence-Based Medicine in Oncology Research

Prof Di Maio: Thank you very much. Good afternoon. And I wish to thank e-ESO because... Thank you, Luca. ...Thank e-ESO because this topic is very, very relevant, in my opinion. As clinicians, every day with every patient we visit, we have the problem of translating the evidence in our clinical practise. So, the very added value of clinical research is the value that we can transfer in clinical practise. So, I will try in these minutes to share with you some considerations about the value and the limitations of evidence-based medicine with a specific focus on oncology. Here, you can see my potential conflicts of interest. And let's start with the pyramid of evidence because all of us know that the large randomised controlled trials are considered the most robust, the most solid level of evidence for our decisions. Of course, many of our decisions in clinical practise are not based on randomised controlled trials, are based on a weaker evidence, case-control studies, observational studies, retrospective studies that of course contribute to the evidence. And in many settings, in many situations, it's difficult to perform a randomised control trial, for instance, because the population is very rare because, for instance, it's a molecularly selected population, or because in some cases it's difficult to propose the randomization to patients, especially, when we have to compare different treatments. For instance, surgery, with another approach, it's not easy to propose the randomization. However, we usually consider RCT as the most solid level. But it's not all gold that that glitters because even in the case we have the results of a randomised control trial, there may be some problems, some issue in the interpretation and in the applicability, in the value of those results. This table is taken from a very relevant, in my opinion, a very important review by Ian Tannock, published in the Lancet Oncology some years ago. Because here are summarised several problems, several issues that might limit the interpretation of randomised controlled trials. Each of these topics could deserve a long discussion. We have no time to discuss all these topics. I will rapidly discuss some of these. And for instance, we can start from a point that has been very, very much discussed also in the last decades, the evergreen problem of the dichotomy of the controversy between the statistical significance and the clinical relevance of the results. Trials conducted in oncology have been, in some cases, have been considered modest results that in some cases were statistically significant, but not very relevant in terms of the magnitude of the benefit, in terms of the clinical benefit. Maybe, in recent years, this is less true because we have observed very relevant results in oncology from a clinical point of view. But this point, in my opinion, is always very important to consider when we look at the literature and we look at the results of randomised controlled trials. I'd like to show you an extract of a publication by Irene Floriani, Valter Torri, and Silvio Garattini, published in Annals Oncology slightly more than 10 years ago. They, of course, emphasised that if a new treatment is to be introduced into the clinical practise, it should not be enough to show that it is better than the standard therapy regardless of the size of the effect. But it's necessary to demonstrate that the effect is clinically worthwhile. And of course, this is a clinical judgement that we have to apply every time we read a paper, every time we consider the application of those results in our clinical practise for our patients. Please remember that you can ask questions and send comments using the Q&A button, and we will discuss the questions later. Let's talk about the validity of the results. What is

the validity? It's a complex concept. We have to discuss the internal validity of a trial, and also the external validity. The internal validity is a very, very important condition. And it's the question, "is the research conducted correctly? Can we trust what we are observing as the results of that clinical trial?" And first important rule within a randomised trial is to check the adequacy of the control arm. That is probably the most important thing in a randomised control trial because a randomised control trial conducted with an acceptable, a suboptimal control arm maybe has a low-value because we cannot adequately interpret the results compared to the gold standard and compared to what we already have in clinical practise. A few years ago, these authors published a very interesting analysis of the quality of control arms in randomised clinical trials, leading to approval of cancer drugs by the FDA in the United States. And, as you can see, that analysis considered 96 anti-cancer drug authorizations during the years between 2013 and 2018. And as you can read at the end of the slide, 16 out of 96, that means 1 out of 6 of the approvals, according to the authors' opinion, were based on randomised clinical trials with suboptimal controls. That means that have a very limited internal validity because we cannot really interpret the value of the result for the setting we are discussing. And why the control arm was considered suboptimal? In some cases, because there were some treatments of proven efficacy for those patients that were not allowed by the protocol. And this is, of course, a mistake. It's, of course, an error. In 10 cases, the treatment was... The control arm, the treatment of the control arm was already proven inferior compared to newer options. And this is a situation that is quite common in recent years, because we have several new treatments that are compared against the old standard, but we have no direct comparison between them. So, we cannot interpret the relative value one against the other. So, please, consider always this important aspect of the trial when you look at the literature. Another important aspect of the validity of a trial, of course, is the endpoint that is chosen for evaluate the treatment. Traditionally in oncology, we have a scholastic distinction, a classic distinction between endpoints of activity against the disease and objected tumour response, but also, progression-free survival and time-to-progression have to be considered endpoints of activity because they are based on the instrumental evaluation of the disease. While on the right part of the slide, you can look to the endpoints which measure the clinical benefit, which measure the effect for the patient, of course, the overall survival and the quality of life. Because as some years ago was stated in a famous paper by the FDA, the objective of our treatments in oncology should be to live longer and/or live better. Of course, quality of life is an important concept from this point of view, because it's a balance taking into account, not only the efficacy of treatment, but also, its toxicity, also its adverse events. So, especially in patients with advanced tumours, it's very important to look at quality of life when evaluating the value of a treatment. And coming back to the panel, to the table by Ian Tannock, another problem in RCTs is that often they use surrogate endpoints that do not reflect necessarily a patient's benefit. That means overall survival or quality of life. Please, remember always to ask a question. Some years ago, I accepted, at the ESMO 2011, I accepted to defend this position, to defend that overall survival is not the only endpoint for drug approval. So, in my opinion, we must be flexible and we must consider with interest also surrogate endpoints and the points that are different from the overall survival. But it's important to discuss the rules to consider surrogate endpoints as a value. A primary condition, a primary rationale for using progression-free survival, for instance, as an endpoint in cancer trials is that independently of the surrogacy it could be considered a clinical benefit endpoint in itself, but with conditions. The treatment effect is sufficiently large, so, the magnitude of the benefit and that it's not only instrumental, the benefit, based on the comparison of CT scans, but it's also a clinical benefit documented by patient-reported outcomes, for instance. And so, another limitation, another issue is that often, very often, at least historically, randomised controlled trials in oncology fail to assess or inadequately assessed health-related quality of life and patient-reported outcomes. And this occurs also in settings like the advanced and the metastatic disease, where the goal of treatment often is not the healing, not the cure, but the palliation of symptoms and the control of the disease. So, it's important to have information about the impact of the treatment on quality of life and on patient-reported outcomes. And this is a very relevant topic. In recent years, the literature about the importance of patient-reported outcomes in clinical trials and in our evaluation of treatments for our decisions in clinical practise has grown and has grown also from the point of view of regulatory agencies. As

references, in this slide, you can see a couple of documents by USFDA and by European Medicine Agency that emphasise the role of patient-reported outcome inclusion in the clinical trials, in oncology. Patient-reported outcomes are a direct report of a patient's condition. So, without the filter, without the interpretation, without the modification from the health operators. And nowadays, are considered the goal standard for the assessment of subjective symptoms and subjective conditions, both in clinical trials and in clinical practise. But today, we talk about, of course, we talk about clinical trials. And some days ago... Some years ago, sorry, we published this analysis, this systematic review of the inclusion of quality of life and patient-reported outcomes in oncology randomised phase III trials. Considering the trials published in a five-year frame between 2012 and 2016, and looking at deficiencies in this assessment and reporting, in terms of both inclusion of quality of life among the endpoints of the trials, and in terms of publication of quality-of-life results in the study papers, in the study publications. And one of the most important results of this analysis was that in almost half of the trials, we considered the 446 trials, as you can see, but in almost half of the trials, 47.1%, quality of life was not included among the trial endpoint. So, this was very disappointing from our point of view, considering that we were looking to this endpoint as primary, but also secondary or exploratory endpoint within the trials. And another important aspect was that considering the trials where quality of life had been included as an endpoint, so, it had been analysed, had been collected, questionnaires were collected from the patient, out of 231 publications of these trials in 38.1%, quality of life results were not included in the publication. So, they were under-reported. And in many cases, they were subsequently published in a secondary publication after months, and in some cases also after years from the primary publication. And of course, this was another disappointing aspect because it is a proof of the under-rating that in many cases, we, as a scientific community and as oncologists, have in patient-reported outcomes and in quality of life. This is the graph of the time to secondary publication. And you can see that even after two years, three years, four years from the primary publication of the trial, most of the results of quality of life were still unpublished. So, this is a big limitation to our possibility of evaluating the global value of the treatment. Because in many cases, we have all the information of a surrogate endpoint, like progression-free survival, without the information of quality of life. That would be very important to use for decisions and for communicating the value of the treatment to our patients in our hospital. And even in the case we have some treatments with a clear benefit in overall survival, it's also important to understand what is the price the patient has to pay in terms of toxicity, in terms of quality of life. So, the evaluation should not be based on a single endpoint, but on the total evaluation of all the endpoints of the study, including, possibly also, patient-reported outcomes. And this was the main message of this editorial that I had the pressure to write on the Lancet Oncology some years ago, because in recent years, most important scientific societies like ASCO in the States, or like ESMO in Europe have included quality of life among the parameters that are needed to calculate the ASCO framework and the ESMO magnitude or clinical benefit scale, which are different, but similar from some point of view, attempts of evaluating the value of anti-cancer treatments. So, if we do not have the information about quality of life, we cannot adequately estimate the value of the treatment. Remember the questions, remember the Q&A button. So, I have rapidly discussed some aspects of the internal validity, which means the study design, which means the study endpoint, the control arm, et cetera, but another important aspect, another very important aspect for translating the literature into clinical decisions is the external validity of a trial, which is the question, "are these results really applicable, generalizable to the real world?" Because we all know that clinical trials, in some cases more, in some cases less do practise a selection of patients who do not necessarily represent those we see in everyday practise. This was the last, but, of course, not the least point in this panel by Ian Tannock. And also, regulatory agencies have emphasised this aspect. Some years ago, FDA conducted with ASCO these important analysis of the investigation, new drug applications in 2015. And you can see in how many cases there was a restriction in terms of performance status. Most trials include only fit patients, do not include patients with performance status-2, for instance, according to ECOG, but this represents a very important proportion in our clinical practise. Most trials exclude patients with brain metastases, most trials exclude patients with HIV infection, but now this is not a competitive cause of risk for patients with cancer because the viral infection can be

adequately controlled for many decades and many years. So, these patients should be included in cancer trials. And many, many trials exclude patients with cardiovascular disease also in the history, not necessarily current. And this is another problem of applicability of results to our clinical practise, because the efficacy should be not necessarily the same when we apply the treatment in our patients. And also, the toxicity could be different if we administer the treatment in patients who are more unfit and with more comorbidities and who assume more concomitant drugs, for instance, with the risk of interactions. And from this point of view, FDA, but also the EMA has emphasised the importance of integrating, not substituting, but integrating the evidence coming from randomised controlled trials with the evidence coming from real-world evidence. Real-world data and real-world evidence are very commonly discussed in these recent years and it's a very important opportunity to produce data that could increase the evidence in settings that are not completely explored and not completely covered by the results of randomised controlled trials. And if you want, you can read this paper, this review which I had the pleasure of writing with my colleagues and friends, Franco Perrone and Pierfranco Conte, a couple of years ago, because, of course, real-world evidence in oncology, like in other fields of medicine is a very timely opportunity, but has also, of course, limitations because in many cases, we cannot substitute the conduction of a well-done randomised controlled trials. But we can integrate those two ways of producing evidence. Please, remember the questions. The last point that I would like to discuss today is a very common point in publication and a very common point that we have to discuss when we try to select the best treatment for each patient in clinical practise. Especially in the era of personalised medicine and precision oncology, subgroup analysis is very, very, very often discussed because it seems a valuable tool with the aim of optimising treatment choices. And in some cases, subgroup analyses also impact regulatory decisions. Let's think about the recent approval of durvalumab, of the immunotherapeutic durvalumab, as consolidation after chemo-radiotherapy in patients with locally advanced non-small cell lung cancer. And this was very relevant because the EMA, European Medicine Agency, decided to restrict the approval of durvalumab, based on the expression of PD-L1, based on an analysis that was not pre-specified and pre-planned but was post-hoc, was based on a different cut-off compared to the cut-off that was pre-planned. So, it's an example of the importance of subgroup analyses in some cases also for regulatory decisions, not only for our single clinical decisions in clinical practise. But why to conduct, why to perform subgroup analyses that we can read in the vast majority of application? Within a study with an overall positive conclusion subgroup analysis might help to better identify patients who benefit more, patients who benefit less, or patients who don't benefit at all with the aim of describing a treatment heterogeneity, according to patients' characteristics that could, in theory, optimise our decisions. Within a study with overall negative conclusions, on the other hand, subgroup analysis might help to avoid "Throwing the baby out with the bath water," and identifying certain groups of patients in whom the experimental treatments, which did not show benefit in the overall population could work and could be of value. But we must always remember that despite the aim of the subgroup analysis is valuable. They have many risks because of the risk of false positive and false negative results due to multiplicity of tests. We decide of not performing a single test within the trial, but to perform two, four, five, ten, twenty, in some cases, many, many, many statistical tests in different groups and this, of course, increases the risk of false results. So, just a slide to remember some key-words, that in my opinion, should be always remembered when looking at subgroup analysis. The first is caution. Subgroup analysis should be looked at as hypothesis-generating. Please remember the problem of multiplicity with the risk of false positive and false negative results. Always look at consistency among studies, because if we are lucky and we have several studies conducted with similar drugs or the same drug in the same setting, it's important to look at consistency of results of subgroup analysis. This could increase our trust in those results. Always look at plausibility because subgroup analysis should be plausible. It's not useful to look at zodiac sign of the patients and find that some patients benefit and some others not because of course this would not be plausible and this would be a clear example of a false positive, or a false negative result. And we should always look at the interaction test that in many cases is not present in the field, is not present in the table or the publication, but it's a way of testing, in a formally correct way, the presence or absence of heterogeneity between the treatment effect and the characteristics

that we are exploring. So, just a minute for some take-home messages. The first is that randomised controlled trials are the gold standard for the production of evidence, but remember that not all that glitters is gold. And of course, also that, in many cases, we do not have randomised trials on which to base our decisions. Please, look always at the adequacy of study design, the adequacy of control arm and of study endpoints. Please consider that the external validity, which means the applicability of results to the real-world patients is not obvious and should be carefully discussed. And consider and remember that subgroup analysis is legitimate to optimise treatment personalization, but has risks. So, please consider its risks when looking at subgroup analysis. So, this was the last slide. These are my contacts. So please, you can also, after this webinar, you can also contact too, for comments or exchange of opinions. And thank you for your attention.

Dr Bertolaccini: Thank you. Thank you, Massimo. Thank you to Professor Di Maio for this beautiful lecture about some critical points. Okay. Katarina, from the floor asked, I don't know if it was mentioned, but why do we have a different interpretation if subgroup analyses are or not pre-planned? And can you explain the differences between planned or not?

Prof Di Maio: Thank you, Katarina. This is a very important question. And the problem is that, in many cases, when we read a publication, we do not know if those analyses were pre-planned or not. In some cases, we have the study protocol and we can check, but in other cases, no. So, it's difficult. The problem is that why a subgroup analysis should be pre-planned? Because it means that the authors had strong hypotheses to test those subgroups, because if I have a strong hypothesis to test, I can decide when I perform the study design and the sample size, I can also estimate which will be the power to test the comparison among treatments in that specific subgroup. And so, I know if I can verify that hypothesis. And the risks when we look at a very high number of not pre-planned analyses, but post-hoc analysis is that the risk of false positive results is very, very high because, by chance, if we accept a P value of 0.05, 1 out of 20 will be positive by chance and not because there is a real difference between treatments. So, the highest the number of subgroups, the highest the risk of these false results.

Dr Bertolaccini: Thank you, Massimo. An anonymous attendee has asked why in your pyramid of evidence there is no a meta-analysis? I found meta-analysis at the top.

Prof Di Maio: You're right. You're right. Of course, maybe it was a too simple pyramid because in the more complex pyramid, you can always read large randomised controlled trials and/or meta-analysis. Of course, I completely agree. I have performed a lot of meta-analyses, also individual-patient data meta-analyses in my career, so, I completely agree about the importance of meta-analyses, which are important especially when we have several trials which are small, which have small sample size but taken together can produce a definitive answer to the comparison, or when we have several trials with apparently different results. For instance, one is positive, one is negative. and so, the best way to estimate all the evidence is to put all together and to produce a meta-analysis. The comment is completely right. It should be added meta-analysis to that level of evidence. It's difficult to perform meta-analysis nowadays, when we talk about pivotal trials of new drugs, because most of these drugs are of the companies. So, it's difficult to put together data from different trials, but it would be very important from a scientific point of view.

Dr Bertolaccini: Massimo, a personal question. You are a Professor of University of Turin, the university where I studied medicine and surgery, so. How to teach the methodology of the research to medical students, in which manner?

Prof Di Maio: You are right. This is a very, very, very important topic, probably, and I talk also about my personal experience. We did not receive an adequate education during the course of medicine in this issue. Also, other issues, for instance, we did not receive adequate education for the communication with the patients. And now, there have been some improvements in this, and maybe also in the methodology of research. I know that many universities have a dedicated course for the methodology of clinical research. In my experience, many fellows, many residents, which enter in the specialty school in oncology do not have an

adequate education from their university studies. So, I try to integrate and to dedicate several lessons to this issue, but from the perspective of how to correctly read the literature and how to correctly interpret the results, but also on the perspective of how to correctly plan a research and how to correctly conduct a clinical trial. Because in my opinion, we have these two distinct, but integrating skills. We have to look at the literature to decide for our clinical practise, but we also can be scientists, we can be researchers and produce the evidence.

Dr Bertolaccini: I totally agree with you. And last comment. It's about when we think about a trial, we have to define the minimally clinical relevant difference when we design the trial, because as you stated, the P value, the significant of a P value could be nothing for a patient.

Prof Di Maio: Yeah. I agree.

Dr Bertolaccini: So, could you leave a comment about the relevant minimal difference?

Prof Di Maio: Yes. Thanks Luca. We have a famous example in oncology. Some years ago, there was a phase III trial conducted in pancreatic cancer, in advanced pancreatic cancer with erlotinib and formally, it was a positive trial because the addition of erlotinib to gemcitabine have been producing a statistically significant improvement in overall survival. The problem was that that improvement was in terms of a couple of weeks of median overall survival. So, considering that we were talking about a drug with side effects, a drug that is also a cost in economic terms, of course, this has not been considered a standard for pancreatic cancer. But in other cases, we have several examples of drugs or treatments that have been introduced in clinical practise based on modest survival advantage, maybe not two weeks, but one or two months in median OS. And this is also a critic that we often received in the past from other specialists that could say, "You oncologists are consuming a lot of money for treatments that have a modest impact on the life expectancy or on the outcome of the patient." And considering that economic resources are limited and need to be shared between different specialties and different diseases this is an important critic to consider. So, maybe, in recent years, the concept of the clinically relevant difference has been more often considered in our discussions and also, in the evaluation of clinical trials. And now, in recent years, we have introduced many treatments in clinical practise, which have demonstrated a very relevant benefit in terms of proportion of patients who are alive after many years, in terms of median survival, which is not improved by weeks or months, but by 6 months, 12 months compared to the control arm. I'm thinking about the trials in prostate cancer or in hormonal treatment for breast cancer, or... I could list several other examples. So, maybe, now is better than some years ago, but the bar should be set always very high because treatments are toxic, treatments are expensive. So, only those associated with a very important value for the patient should be introduced in clinical practise.

Dr Bertolaccini: Thank you, Massimo. We have not any other question or comments from the attendees. I would like to thank you for this beautiful afternoon. It was a real topic and a real beautiful lecture. Thank you, Massimo. Thank you. Thank you. And thank you for the attention to all our attendees.