ROUGHLY EDITED COPY

BIRNBAUN INTERPRETING SERVICES

FOOD & DRUG ADMINISTRATION

7th ANNUAL CLINICAL OUTCOME ASSESSMENT IN CANCER CLINICAL TRIALS PUBLIC WORKSHOP

JUNE 29, 2022

HARDEMAN REALTIME, INC.

* * * * *

This is being provided in a rough-draft format. Communication Access Realtime Translation (CART) is provided in order to facilitate communication accessibility and may not be a totally verbatim record of the proceedings.

* * * * *

WELCOME AND OVERVIEW

>> PAUL KLUETZ: Hi, everyone. It's 10 o'clock and I think we can advance the slide to the welcome slide. My name is Paul Kluetz and I'm a medical oncologist at FDA and it's my distinct pleasure to welcome you all to the 7th annual Virtual Public Workshop Clinical Outcome Assessment in Cancer Clinical Trials workshop. This is a series of annual events that we've hosted since 2016. We bring together experts in outcomes research to advance the measurement of symptoms and function for patients with cancer. Thankfully since 2010 when I started at the FDA, we're an in era of unprecedented number of effective cancer agents, but agents have toxicities and more so in oncology. And so the question always becomes if I'm one of the patients for whom this drug is working, how will I feel in function when I'm taking my therapy? And it's so important, especially in advance cancer setting to provide information to patients about that. It's a fundamental question and it's been a research and policy interest of mine. And certainly, the motivation for many dedicated scientists that we've interacted with in the research field some of whom we'll hear from today. So next slide, please. Next slide.

As I mentioned I'm the Deputy Director of on Oncology Center of Excellence. Whether they're drugs, biologics, or devices. The OCE mission is to achieve patient center regulatory decision-making and to that mission, one of the very first programs we set up was the oncology specific patient focused drug development program. This program is now led by Dr. Vishal Bhatnagar, who you'll hear from throughout the workshop and he's done an amazing job at really achieving many of the challenge, goals we set out to accomplish back in 2016. And addition to collaborating across these product centers, the patient focused drug program is explicitly to multi-disciplinary work. The work is most impactful not just by clinicians and statisticians and cycle nutrition but all three to find to achieve patient reported outcome analysis and be able to present important accurate data to patients if providers. So when I first presented this slide last year well, there was a lot of challenges. How do we use this data for regulatory purposes? And I identified some important challenges at the time. There were clearly no solid research objectives to the patient-reported outcome oftentimes. Oftentimes there's no standard set of concepts to measure, and, therefore, lack of agreements on instruments and how frequently to assess them. And this led to all sorts of different kinds of analyses and methods and data presentations, and we just had a hard time reviewing this. I think partly because of the lack of clarity, the data quality also suffered. PRO was put into clinical trials as an afterthought and we had in front assessment and high levels of missing data. So where are we now? What have we done to try to improve this situation? Next slide. Well, as I mentioned, this is now the 7th Clinical Outcome Assessment in Cancer Clinical Trials workshop. And so one of the first things I did was set up an annual engagement effort. That was very practical. It was very driven to address the issues that I previously presented. We tackled the outcomes. We looked at the existing quality of life instruments and identified those aspects of quality of life that were most relevant to the drug Understudy and least potentially affected by non-drug influences and came out with a core outcome set. We looked at the list of acceptable instruments and presented them to sponsors.

We looked at trial designs and assessment frequencies based on what we were trying to get out of patient-reported outcome, clarifying objectives. And this has led to endpoints analytics and visual points which is ongoing today. And so, today's workshop is a little different and we'll talk about it's a much broader topic than has been in the past. So in addition to engagement, we have a robust scientific program. We've taken advantage of the ORISE Fellowship Program that host early outcome research scientific that have been productive to help us investigate. Jessica Roydhouse came to us from Brown. Jeff Chen from Anderson came over and did incredible work looking at how subgroups with significant simple symptomatic toxicity affected quality of life. We also have current fellow Meena, as well as using PRO to assess frailty measures in the elderly. In addition to ORISE, we have CERSI. Of multiple modality of physical function assessment. We're looking at wearable devices, electronic captured reported outcomes, clinical assessments of physical function, and this is co-investigated with my colleague, Edith from the Mayo Clinic who is going to be on the panelist today. And finally we have FDA center for excellence broad agency announcement. Northwestern group is looking at overall side effect of particularly the FACITGP5 which is important for accountability and Jessica Roydhouse led that on both her panel today.

So through this engagement and the scientific portfolio, and through creation of consistent regulatory review and advice over the years, I am happy to say we've made what I think is pretty significant progress on those issues that I previously identified. We clearly now have a much more research objective-based PRO assessment strategy in cancer clinical trials. We proposed Universal is to provide safety and toll built. Unique patient centric concept. This is descriptive statistics and we do this with safety and certainly with patient reported symptomatic reported toxicity and side effect. Efficacy is also a subject and for functional improvement in select cases and compare tolerability is becoming an interest as project optimist comes online looking at randomized dose trial and see which dose is more tolerable. It's been said by the project optimist team and by Dr. Paster, it would be challenging tolerability to simplistic toxicity and how it's affecting them. And consistency among the concepts to measure and the instruments, the assessment frequency, we released the 2021 FDA guidance for industry core patient reported outcomes in cancer clinical trials. This again, wrote itself. Because we've been giving consistent advice over the last 2 to 3 years. And we helped push this guidance through and it really addresses what our core outcomes are that we like to see and how frequently they should be assessed and types of instruments that might be appropriate for assessments.

Perhaps because of all these efforts, we have seen continued improvement and thoughtfulness and patient-reported outcome captured in trials. So we have a lot of improved data and thoughtfulness and how the studies are performed. And now we go to the last issue, which is standard analytic method and data presentation. Now that we have the foundation, we can begin to really create consistency and analysis methods and data. We're participating an international collaboration like the and standard analyses for project voice for single item patient reported outcomes, and that, again, was really helped in large part from Ethan Basch and foundational work for PRO-CTCAE. I wanted to touch on couple of these important efforts. This is the blow up of the patient-reported outcome guidance. Again, this clarifies that our core set of

outcomes that we think we're going to help for regulatory review would be symptomatic adverse events, global side effect impact measure, physical and role function. And disease symptom where they're applicable. And it's a trunk loaded assessment frequency. So we like to see the assessment early on to capture the datas, potentially, and as many patient as possible for their acute toxicity and potentially acute symptom function improvement that we can see with highly effective cancer therapies. This is where the acute toxicity occurs. And it's where the least attrition holds. Next slide. One of the things when we were interested in moving to more of an item library approach to pick, the expected toxicity to assessment of patient-reported outcome, we were going from static instruments to more adaptive patient-reported outcome strategies. We were interested in PRO-CTCAE. And we heard from sponsors they were concerned this would be considered safety data. Perhaps CTC is in the name. So we did need to write a paper to clarify our thoughts on the topic. And it addresses some of these concerns I think very clearly. So this is a great resource for sponsors and one of the deployed CTCAE trial. Finally, I just want to do touch on project patient voice and this is a unique way to communicate patient-reported outcome. This is one of the core outcomes which is symptomatic side-effects. This was in the trial versus therapy. And were displaying these severities of symptoms experienced and side-effects for both arms. It's a publically available website as I mentioned.

With that, I will end my introductory comments. I hope I gave you sort of an overview where we are now. And where we're going, which is standard analytics and visualization and more communication. As well as, really looking heavy at tolerability. And today's workshop is going to be very interesting, because it's much broader topic. It talks about challenge of open label trials in-patient reported outcomes. And the reflex active thought that patient of reported outcome assessment is not going to be accurate. It's not going to be reliable. And we're going to really say that reality is that most of our trials with open label in cancer. We're always interested inside effects and clinicians are actually assessing and reporting. This is not new. This is just a different way of assessing tolerability and how to enable open trials related to patient-reported outcome and analyzing. And there's over is 1,000 registrants and we will not get to every single questions and comments. But we'll take them back and talk about them in the patient focused teams. And with that, I'll introduce Vishal Bhatnagar, and he's going to walk you through the agenda and begin session 1.

>> VISHAL BHATNAGAR: Thank you so much, Paul, for that introduction. My name is Vishal Bhatnagar. I'm associate director for patient outcomes in the oncology center of excellence and a medical oncologist, and we’re really glad to have an esteemed panel of discussants today to discuss this very important topic of open label bias in cancer clinical trials. We have the day structured into 3 sessions, and in the first session, we’ll discuss the reality of open label trials in oncology and use of patient reported outcomes in those trials. In the second session, we’ll talk about analytical methods to either detect or to mitigate the potential for open label bias. And in the last session, we’ll discuss ways to advance how patient reported outcomes can be incorporated into cancer trials, and then discuss novel ways that patient reported outcomes have been used in open label trials. So, with that, we’ll go ahead and get

started, and as Paul mentioned, we're excited for this workshop, and look forward to a robust discussion. Thank you.

SESSION 1

>> VISHAL BHATNAGAR: So, I'll just got ahead and get started with session 1. So, as Paul mentioned, we had a long way to go in terms of how patient-reported outcome have been incorporated in cancer drug development. So there are key characteristics of the current landscape for patient-reported outcome assessment on oncology. First, the patient-reported outcome in cancer trial has become a common place and collection. And majority of trials and oncology on the trial. And other reality, third reality we face is that there's a huge difference in how patient reported tolerability, descripted reports from patients and side-effects are compared to efficacy type PRO such as period of symptoms overtime. These concepts will be discussed at length today with goals of better detecting and potential for open label buy-in and consider ways to communicate parish commune patient-reported outcome outcomes trial. And I just want to make sure we're describing the oncology drug development experience. And that therapeutic areas outside of oncology have important consideration to be discussed with the relevant review decision. Next slide.

So to start the session off, we endeavor to review the existing evidence and literature that discusses the process of open-label process. And we found some dissipations and analysis that provide the bias to exist which helped inform this workshop. One of the earliest mention of this bias was Atkinson and colleague who conducted a literature review and identified 5 randomized trials in which patients experienced Hallmark of inadvertent unblinding. And in these trials, no difference was seen between the collected patient outcome and suggesting when patient were inadvertently aware of their treatment, knowledge of their treatment did not appear to influence their responses to PRO.

Dr. Jessica Roydhouse and FDA authors described the potential for and then in exploratory case example, a way to potentially detect open-label bias. And all three of these articles I mentioned were lucky enough to be joined by co-author of they see articles. They're all panelists in this workshop. Noted in colleague immune oncology studies, trial designs had no impact on PRO completion rate or baseline scores. Finally, a recent publication by Efficace, et al., trial design and trial result. Although there are potential demonstrations and analysis used, all of these trials share a commonality there is a potential for this bias to exist. And that they all undertook various methods to detect the bias, with that said, all did not find meaningful evidence of open-label bias. I'll provide you literature reference at the end of the presentation for further reading for those interested. So development occurs over a continuum. First in human trial all the way to post marketing and reaching clinical practice. We're advocating for measurement of patient symptom function to report to other pharmacology. And Paul referenced earlier, which optimal dose finding should include measurement of patient tolerability. But the reality is that many oncology drugs never have a blinded randomized control trial as part of their development. If they do, the

majority of development is done using open label trial to find the dose, so for efficacy and safety, and expanded cohort.

Therefore, we expect to receive increasing amount of patient reported symptoms and function data collected from open-label trials because of initiatives like project document.

Over the last 10 years, there have been 13 examples of patient-reported outcome included in oncology drug labeling. So some of these are patient reported improvement and symptoms such as pain measured by the pain inventory. One label includes pain tolerability and interference of visual disturbance faced by patients. And 3 labels include patient administration on dedicated trial results.

Next slide.

Of these 13 examples, 10 described patient reported outcomes collected from open-label trials and were disseminated to a product labeling. This slide is not meant to go through the rational and labeling for each of these products but simply to show open label trials are reality of oncology drug development and well collected PRO, clinically meaningful can and should be included in product labeling if from an open label trial.

Thinking of communication, as Paul mentioned, we launched patient voice in to 20 and we included an open-label trial that used PRO-CTCAE. Again, these results are from an open-label trial comparing chemotherapy. Because of the focus for project patient is to informed tolerability of CT cancer therapy. We included this patient outcome data from open-label trial.

So to summarize, we know that open-label trials are common in oncology drug development, but importantly, the collection of symptoms and function will increase in phases not previously done. We expect increasing amount of PRO data from all phases of drug development. If therefore, collected from open-label trial.

More research should be done to quantify and understand the magnitude to which open-label trial exists if at all. And most importantly, we all benefit from better collection, analysis, and communication of PRO from oncology trial. So this needs to be discussed in places besides the medical literature and our FDA review. Advocacy organizations, drug developers, clinicians, and most importantly patients need to be involved in the conversation, which is why we're having this workshop today.

Here are the references I mentioned earlier and these slides will be available in a few weeks if anybody is interested in referring back.

So I'm really excited to get the panel discussion going. We have a theme of guests. We're going to begin today's work starting with Terri Armstrong, who is an outcome researcher at the National Cancer Institute. Dr. Martha Donoghue, who is a pediatric oncologist. And Wenora Johnson, who is a 3 times cancer survivor. Dr. Bryce Reeve is a professor and population health sciences at Duke University. And last but not least, Dr. Gita Thanarajasingam is an Associate Professor of medicine and

lymphoma clinician at Mayo Clinic. So for purpose of this panel, we will go by first name for conversation purposes.

So for this session, we'll sort of dig into these questions that were included on the agenda. But I have here listed for reference. So the first question is going to be around why open-label trials are used in oncology drug development and whether they're prevalent in certain treatment settings. Next question is going to talk more about PRO and open-label trials and how they've been incorporated in open label trials. Thirdly, we'll talk about how the impact of open-label bias may differ on in tolerability or symptoms being set. And finally we're going to think about unique considerations for PRO core outcomes as we have designated in the guidance and as Paul mentioned in the outset. So I'm going to start with my colleague Martha Donoghue. So Martha, to kick things off, what are reasons you see open-label trial designs and oncology, and particularly, with your experience in the pediatric, and rare disease phase?

>> MARTHA DONOGHUE: Thank you so much, Vishal. And oncology, we touched upon this a little bit earlier and maybe in some of Paul's presentation as well. We often approve drugs based upon results of single arm trial that document clinically meaningful overall response rates that are durable. And single arm trials are open label by default and alternative treatment patients would be receiving. And we consider single arm design relying on overall response rate and duration of response endpoint appropriate to support approval for some cancer settings, particularly for rare cancers.

Although, a randomized trial would typically be our preference, there's circumstances in oncology where randomized trials may be infeasible or randomized trials are not ethical, circumstances when there's a lack of equipoise regarding whether it would be appropriate to randomize patient to alternative arm or when there isn't an appropriate or ethical comparator that the patient can receive. And in oncology, you know, one thing that differentiates us from other diseases, given that tumors generally do not shrink without therapeutic intervention, those end points of response rate and duration of response can be appropriate endpoints that are reasonably likely to benefit the patient which could support accelerated approval in many cancer setting or could be considered in some circumstances a direct measure of clinical benefit. So the prevalence of single arm trial, and therefore open-label trial is directly linked to the type of endpoint we use to support approval in oncology, particularly for rare cancers. And sometimes also in other context.

The other thing that I think differentiates oncology, when trials are randomized and double blind, that doesn't happen all that often in oncology, it isn't unusual for toxicity such as rashes which we see are epidural dermal. And so they may be open label irrespective of that design and it's intentioned originally for blinding.

And randomized trials can be between the test drug and therapy and dose register is particularly burdensome for patients. So we wouldn't necessarily want to go through the process of blinding administering placebo in that way. And there are other times we don't think blinding is necessary and that's why we don't see it very often.

>> VISHAL BHATNAGAR: Thank you so much, Martha. So Wenora, have you heard of in your experience the potential for in randomized trial, for patients to potentially know what they're receiving? Have you heard that from your colleagues or friends?

>> WENORA JOHNSON: Yeah, and I really want to say I appreciate Martha's deep explanation of that. Yeah, it's likely that patients can unintentionally identify what treatments they're given because of what they experience. So for example, I've gone through chemo treatment and I've sat in an infusion chair for 8 hours. And after being disconnected from that, I can recognize immediately the side-effects of chemo, because I'm nauseated, I feel if a fatigued and loss of appetite. If I'm in a blinded clinical trial but told I have to be connected to an IV or Ford infusion or so, I think I would unintentional identify that treatment as some form of chemotherapy due to the telling characteristics of the procedure itself. And then the side-effects of the experiences that I felt afterwards.

>> VISHAL BHATNAGAR: Thank you so much, Wenora. That's an interesting point. Although on project patient voice we have open-label trial, you can see quite an imbalance of some of the symptoms experienced and Hallmark symptoms. And I mentioned in the publication at the outset that looked at Atkinson's piece that looked at signature Hallmark symptoms that can inadvertent unblinding. So thank you for that perspective, Wenora. So Terri, I'm going to switch to you now, Terri. Look at your personal experience at the NCI and deploying single label trials, tell us about your work and also your experience in deploying PRO in open-label settings.

>> TERRI ARMSTRONG: Thank you so much, Vishal. It's such an honor to be part of this conversation. I appreciate the comments that my colleagues already made. I think the one point that I wanted to make is I'm currently an outcome researcher in the intramural program. But NCI has a long history of supporting this work with development of tools and methods. And even things like the PRO-CTCAE and others were supported by the NCI. So I've been privileged to lead the portion of studies in both large randomized trials but also single open-label studies, focused on primarily oncology. And in all these studies recognizing risk of bias and our own bias is important employing methods to reduce that is key. If you're not familiar with neural oncology, these are rare tumors and building on what Martha said, these diseases oftentimes, there's not a standard of care treatment. And understanding the clinical impact of these therapies as we develop them are incredibly important as they are with all cancers.

As an example, I've been able to lead some phase two studies where we have patient-reported outcome to assess the clinical experience of patients that are treated on these trials. And what we've been able to show in some of these studies is that patients may report symptoms and associate it with toxicity of therapy, but we've seen differential reporting in terms of symptoms associated with the disease. And that has become incredibly important in these diseases to be able to evaluate not only the clinical benefit, but the tolerability of that treatment in the trial, but also to provide that information to clinicians who may be prescribing it in the future and also patients undergoing this treatment. So in summary, I think we need to ask questions in a

scientific rigorous way. Hypothesis-based connections should be asked and adding these to traditional endpoint can provide meaningless full data.

>> VISHAL BHATNAGAR: Thank you, Terri. We're already hearing things like inadvertent blinding and just at the outset of these, many of these things will be discussed in the later session or within this session as well. So thank you for bringing up all those points, Terri, Wenora, and Martha. We're going to switch gears a little bit. And we're going to hear from Bryce Reeve about his experience implementing PRO in pediatric trials. Specifically we're going to look at ways there's accordance with reports from adults and pediatric patients. So I'll ask Rich to queue up the slides and we'll hear from Bryce.

>> BRYCE REEVE: Thank you so much. It's an honor for me to be part of the panel and the workshop to talk about these very important issues.

So obviously, what we're interested in is how a treatment impacts how a patient feels and functions. And I think there's been no debate in the field. There's been no debate in the field when it comes to ideally it should be the patients themselves that provide their own self-report about their symptoms and functioning. However, as we know for the question for today's workshop, there has been some concern about the potential bias that patients may have and reporting things if they know on the treatment arm or on the placebo arm. So that raises the question that is, okay, if there is concern about the potential for the self-reported data to be biased, should we be looking for this information about how patients feels and functions from other reporters from clinicians or caregivers? And the importantly, if we want to see this information from others, we need to ask if we are looking to them, what is the quality of their data? How well does it align with the actual patient self-report? So this issue has actually been explored in-depth among adult cancer patients. And as an example, I want to note some of the work that Ethan Basch has led, comparing the clinician report versus the patient self-report. And what you'll see in these cumulative instance curves in this article published by Ethan is the patients are arm red curves in terms of reporting accumulative adverse events, and we see there's a discrepancy as such symptoms among fatigue, appetite loss, nausea, and diarrhea. And in the adult cancer space, there seems to be agreement there is relative to adults clinicians tend to underreport both the number and severity of symptoms.

Next slide.

Our group was very interested to see if this level of discordance was similar in pediatric populations. So in our 9 pediatric cancer site across the country study, funded by the U.S. national cancer institute, we collected data from over 430 children and adolescents between ages of 7 and 18 years of age to report their symptom adverse events and we collected complementary data from their caregivers, as well as information from the pediatric oncology clinician that was most familiar with that child. The reporting system we used for the clinicians was the CTCAE and for the children, we used the pediatric version of the PRO-CTCAE. And for the caregivers, they completed a caregiver version of the pediatrician PRO-CTCAE. And we were able to align the scores in terms of CTCAE grade from 0 to 3.

The purposes of our comparison today, I will be presenting the data where the child and the caregiver and clinician completed the CTCAE version after they received either chemotherapy about 7 to 17 days later or radiation treatment, about 4 weeks later. This is at the point where the child would have experienced the worst symptom as result of a particular treatment.

And so, on this slide, what we're looking at is the level of agreement between the either the clinician and caregiver with the child with cancer. And along the y-axis is a level of agreement which could run from low to 0, or good agreement close to 1 at the top of the y-axis. Along the x-axis are 15 symptoms that we collected about the child from the child him or herself from the clinician and the caregiver. And on the screen, let's start with the first symptom on the screen, nausea. What we can see in the orange circle is the estimate of agreement between the caregiver -- I'm sorry, yes, the caregiver, the parent and the child, which is nor nausea is 4.0. And if we look at the agreement between the clinician and the child for nausea, which is the purple triangles, we see the agreements in the fair zone about .3. Now, to look at the trend across all the 15 symptoms, what's very clear is when you look at the clinician and child agreement in the purple triangles is that at best, the clinicians have fair agreement and there are times when the clinicians have very poor agreement with the child in terms of adverse events for such symptoms as fatigue, anxiety, and depression.

And as you can see, the caregivers have better agreement with the child than the clinicians. So now that we know there's more this score between the clinician and children there, now let's look at the magnitude of differences. So now what we're looking at on this screen are the same 15 symptoms along the x-axis. And now what we're looking at along the y-axis is average CTCAE grade. And in this particular sub-study, we only include children that were actually experiencing the symptom that is listed at the bottom on the x-axis. And so, again, starting with nausea as an example, the blue circles are the child. And so on average for nausea, average across the children that were experiencing adverse events, average was about 1.4. The caregivers are the green squares, and you can see for the caregivers for nausea is about 1.3. And then we see the red triangle for the clinicians for nausea, and that's somewhere around .90. Let's look at the trend across all the symptoms. And in particular, if you look at the red triangles, which represents the clinicians average grade for those kids who are reporting AE, consistently as you've seen in the adult space is consistently compared to the child the clinicians are downgrading both the number of symptoms, but also severity of symptoms compared to the child. So next slide, please.

So this brings me to my overall comments and findings take away points overall. And these take away points are not just reserved for the children, but also I would also think they're relevant for the adult space. We believe that patients, whether this is children or adults provide a distinct and important voice to reporting what symptoms they experience in a clinical trial. And so, therefore, it should be captured in open-label trials. Obviously, there is some value to also think about what we can collect from the clinicians and caregivers to provide a complementary perspective of what the patients are actually feeling. And importantly, this information is important especially when the patient is unable to report because they're too ill or too young.

If we already captured multiple perspectives, we just need to think about what would be the endpoint specification. The next point I want to make here is that all measures, all clinical outcome assessments whether it's self-report from the patient, clinician report, or from the caregiver, all these things are supportive. And all of them are prone to some bias and open-label trials. And importantly though, what we do know is especially in the past couple of decades, there's been an incredible focus and effort to patient-reported outcome measures that have undergone some extensive development and validation using quantitative and qualitative measures. If you look at such clinical measures like the CTCAE has never gone through the rigor like the self-report measures have. And so, again, overall, we need to collect this important distinct information, and more focus our comments today about what ways we can do to mitigate, adjust, and reduce the potential for bias in Piero open-label trials. And I'll turn that back to you for more discussion.

>> VISHAL BHATNAGAR: Thank you so much, Bryce. I think that presentation is just such a great platform to begin the discussion and we're going to continue thinking about these issues as we progress through session 2 and 3. So thank you, Bryce, for that presentation. And then there's also been some questions in the chat, which we'll get to end of this session in regards to your presentation. So hang close. So, Gita, I want to turn to you now and ask you have such a unique experience as a Pharma clinician, ODAC member and PRO researchers. But this question is more from your clinician hat. So what's your perspective on the potential advice on the clinician side of things when assessing PRO?

>> GITA THANARAJASINGAM: Yeah, you know, I really want to pick up on Bryce's second concluding point you're mentioning about subjectivity, bias of different clinical outcome assessments. As a clinician who is grading AEs both on open-label and blinded cancer trials, if there's a concern about patient reporting and bias in open-label trials, then we have to acknowledge there's bias inclination reporting. Open-label trials. We're trained to be objective evaluators, but we're human beings and we're subject to the same type of biases our patients are. But we tolerate this, because clinician reporting constitutes part of safety reporting that we need. But I really feel like vitally we need patient reporting as part of understanding tolerability. So if you're going to tolerate bias in one area, you know, tolerate it in another.

Another point I want to make is somewhat obvious, but patients who are receiving care in routine oncology clinics, not on clinical trials know what treatments they're on. And so when we collect patient reported outcomes in patients who know what treatments they're on, as a clinician, I feel like that data more closely mirrors what my real world patients are going through. It will likely reflect the experience of people who are treated with full knowledge of what they're on. So I think that's an important point to make.

And then thirdly, and sort of more generally speaking to kind of the topic of this meeting, I think we all understand that patient reported outcomes are vital component to understanding tolerability. Something I really want to emphasize that patients ask to much about tolerability information. They want to know how they will feel and function on treatment. And if they don't get this information from us in a rigorous scientific

manner, they are seeking it from other cancer patients, through avenue such as support groups, social media, and other resources. And I think those are important complements, but I don't want to feel inadequate as a clinician not being able to provide high-quality PRO information on tolerability to my patients. And we should be able to do this from rigorously high-quality PRO studies both in open-label and in randomized trials.

>> VISHAL BHATNAGAR: Thank you, Gita. So I'm going to ask Wenora, because Gita mentioned a lot of things. Some of what you talked about, Wenora, but we also want to hear from you, understanding you can't speak for all patients. But let's touch on things like where you received sources of information about symptoms and side-effects? And then also thinking about, you know, that sort of patient versus clinician reporting. So we want to hear from you, Wenora.

>> WENORA JOHNSON: Thank you. And I think Gita hit it right on the nail that patients will seek other sources of information. A lot of times that is in those Facebook groups or on websites. And the thing about it that you have to be really careful with, even as a patient, even for me when I'm looking up information is that it's correct and reliable information. My patient experience may be different how I reacted to a particular medication than another patient. So that's where you have to be really, really careful, too, as a patient. Patient-reported outcome are really important, because it can, you know, give us a better form of targeted therapy and intervention when it comes to, like say reporting pain outcomes. And I think for me, especially when I go back to my experience as a patient, and sitting in that chair, the only thing I could think about is how am I going to feel after? The fatigue and pain were the biggest concerns for me. And those are reported a lot on from various patients on sites. They want others to know. Hey, this is how you're going to feel.

>> VISHAL BHATNAGAR: I think Gita and Wenora, you bring up really good points. And advocate for why this data should be incorporated and collected. I want to ask Martha, you know, we receive a lot of data. Some of it regarding health outcomes like hospitalization. We receive a lot of clinical reported safety. But let's think about on receiving end of this type of PRO data, let's sort of demystify what happens in the regulatory process a little bit, specifically, when it comes to the importance of this patient-reported outcome likely to built data, some of the things like Wenora mentioned about like side-effects and symptoms from therapy. So Martha, if you don't mind sort of providing your perspective on that?

>> MARTHA DONOGHUE: Sure, thanks, Vishal. I'll do my best to demystify things. So we just heard a number of reasons why we would be remiss if we overlooked the opportunity to learn about patient experience directly from the patient who participate in clinical trials. In terms of tolerability, which is extraordinarily important as Wenora and Gita eloquently described. And at times, in terms of efficacy of the drug as well. And I think as Paul alluded to earlier, and as everybody can understand and hear from the ongoing conversation, our approach to FDA using PRO data and review process has evolved over the past several years.

And in generally, when we receive PRO data that are rigorously collected and using the appropriate instruments with applications, we do consider these data as part of our overall risk benefit assessment when we're reviewing application to help us in our evaluation and determination whether benefits of the drug outweigh the risk to the patient. And this risk benefit assessment is very complicated and context, contextual. And it takes into data element. We review investigative report information, including logic data and survival data, et cetera. But when we have well collected PRO data submitted, and we're seeing that more and more often recently, we use that data to help us gain a better understanding how patients may be benefiting from the drug, or how they may be affected by the drug to learn more about the side-effects they experience and how these side-effects impact their lives. So putting all these data elements together, including the PRO data can be very helpful when we're making these complex decisions about whether benefits of the drug outweigh their risk. There are also some rare cancer that have limited potential for tumor shrinkage, in these cases we're exploring ways in which PRO endpoint can play more pivotal role in establishing a drug can result in a direct clinical benefit to patients as well. So to summarize PRO data to add complementary information or play a stronger role in our risk assessment, depending on the situation, and importantly as you and others have stated very well, it provides potential to provide very useful information that patients can share with one another with regards to what to expect while in treatment. And I would imagine not being a patient myself, but I would imagine having an understanding as to what you can expect can be very helpful to help prepare you for those potential side-effects and I don't think the importance of that can be overstated.

Additionally, because the PRO data is captured longitudinally, it can provide a clearer picture to patients. Whereas oftentimes adverse data we see collected by sponsors is sort of did the adverse event occur? Yes or no? And what was the grading, but sometimes there's less of that flavor of experience overtime in that type of data. And I also want to emphasize the team during drug development interact closely with patient focus experts along the oncology for excellence and other clinical experts within the FDA, because of course that's where the expertise with respect to using PRO to support cancer drug development resides.

And so we get a lot of assistances as a review team from you, from others who have this expertise during the entire drug development process. And these discussions are really integral to how we use PRO data to support drug development in oncology. So I guess lastly I'll just say in general in my experience, we've considered analysis of PRO data to be supportive primarily and looked at the descriptive purposes, not always, when there's statistical attached to it and use the data differently. But irrespective whether it's descriptive or not, this information is helpful to help us enhance the understanding the effects of treatment or symptoms and related toxicity of treatment. And this is valuable to our overall valuable assessment and decision-making as regulators even when the trial is open-label or blinded or truly blinded.

>> WENORA JOHNSON: If I can add something really quick to that. I really appreciate what Martha just mentioned there. There's another component that's added

to this when it comes to the patient's experience, too, because when you kind of know ahead of time what is to be expected, your mental health side of it better-prepared. And you want to survive and you know what that side effect is going to be and you're prepared ahead mentally to fight it to get to the other side. So there's a mental component to that as well.

>> MARTHA DONOGHUE: The forearm, it's better to know what you're heading into. Sorry, Vishal.

>> VISHAL BHATNAGAR: I just want to ask, we're going to switch topics, but, really, stay with you here. Let's talk little bit more about this idea of what may be open bias and patients reporting their symptoms and functions and side-effects. And from your perspective whether or not this is even a thing or this bias exist. Or whether patients are being asked these questions and sort of what goes through a patient's mind when they're asked these PRO questions when it comes to knowledge of their treatment. Not to put you on the spot, but since we're going to talk about it.

>> WENORA JOHNSON: That's putting me on the spot, but that's okay. I mean, you know, the first thing that comes to mind as a patient is that you hope this information becomes important enough so that maybe the next person doesn't have to experience something so horrible that maybe you've went through or gone through. So you hope at the end of the day, it's going to be a good outcome. But that only comes with having that good relationship with that provider or that person who's taking care of during that point.

So making sure that that communication is there. It's just vital.

>> VISHAL BHATNAGAR: That's a really great point. And I mean, that's a huge component of why we created project patient voice, because we know patients provide their experience with their symptoms and side-effects, and a critical component is that they provide that with the assumption that that information is going to make it, yes to the drug company and yes to the FDA. But most importantly, it needs to be in the hands of other patients and other healthcare providers administering the therapy so they get a better understanding of what to expect. So I think that you're spot on and I fully agree with you. And that was a huge reason why we created project patient voice. Thank you for that, Wenora. Martha, going back to you, thinking about expansion of the way in which we assess patient reported outcomes, and thinking specifically about your experience with the PRO-CTCAE pediatric subcommittee, ODAC discussion that occurred recently. Thinking about expanding rigorous PRO collection to the pediatric population. Let's talk little bit about some of the topics that came up with that subcommittee and thinking about that same things you brought up and Wenora brought up in this point about sharing the experience of patients. So go ahead, Martha.

>> MARTHA DONOGHUE: Thanks, Vishal. So, you know, Bryce I'm sure would want to chime in too about this because he was pivotal to this discussion. So we have this pediatric subcommittee of oncology committee back in May and we devoted close to a full day of United States and capture of PRO data in pediatric for this specific focus on tolerability information being collected through PRO-CTCAE tool. And the

consensus among all members collection and use of this information using the PRO-CTCAE would be extremely helpful to help provide additional information to inform tolerability and safety of the drugs in pediatric patients. And this would be true irrespective whether the trial is randomized and blinded or open-label. And also throughout the entire stage of development, even in the very early phase 1, 2 stages. I think the committee recognized that getting the investigator perspective was only representing a piece of puzzle in that getting that patient or sometimes caregiver perspective as well through data collection using the PRO-CTCAE would really help augment a much better understanding of how patients are being affected by drug. And so they were very supportive. The entire stakeholder community that gave voice was very supportive of using that tool to make better decisions about drug development in general for pediatric patients and also for other purposes aside from drug development per se, such as helping families and patients understand better what type of side-effects has been experienced to date and what to expect. And how to support one another in a different way. And also I think there's potential down road to use that information as a clinician as well to better understand in real-time ideally when we have the right infrastructure to use that information, even in care decisions, medical decisions overtime for patients receiving treatment.

As you mentioned and this has been come up couple of times, it's important to distinguish the tolerability of the drug and how the drug improves related symptoms and either case, when we're using PRO data, we need to look at other pieces of information such as how supportive care that may be ongoing at the time can impact some of the data we're getting. For instance, use of pain medications, and how supportive care can impact the investigator assessment of symptoms or the patient assessment symptom.

>> VISHAL BHATNAGAR: Bryce, your name was mentioned. Thanks, Martha. What was your perspective on that subcommittee discussion and how it pertains here?

>> BRYCE REEVE: Thanks, Vishal, and thanks Martha for that wonderful summary. It was such a rewarding for me and Pam Heinz and our group there to be part of that subcommittee meeting there and to see the embrace for capturing the child's perspective, which I think will have a huge impact in the work that we do. It reminds me in terms of putting value on the child's voice there, doing our development of the pediatric PRO-CTCAE during qualitative interviews. This bright 9-year-old girl was being asked how difficult or easy was it to complete our pediatric PRO-CTCAE? And she responded very quickly, oh, it was very easy for me. And the interview probed further and asked why was it easy? Well, all the questions were about me and I'm the expert of my own life. So we felt like that's all we really needed to justify the use of the pediatric PRO-CTCAE. And we also got feedback from caregivers. We were doing interviews with caregivers as well, and they also came up to our CRAs and in particular I remember one parent said, thank you so much for giving me an opportunity to find different ways so I can understand what my child is going through, especially, on the mental things like depression, and anxiety. So, again, we're very excited about the embrace of the pediatric PRO-CTCAE. We know the child can provide a unique voice. But we also have to recognize there are challenges as well in pediatric trials. There are times when the child is too sick to self-report what he or she is feeling or just don't

want to complete a questionnaire. And we don't want to burden them with that. And addition, younger children are also maybe not able to read independently self-report. So there's still a lot more work we need to do to capture a child's voice for those particular population or adults, or pediatric where they can't self report, so I think it's important for us to get complement their information from the clinician or caregiver, and think creatively together how we can use the different reporters to inform our evaluation and tolerability, or benefit. If that will lead to the importance of making sure we're very clear about our endpoint specification in those types of settings.

>> VISHAL BHATNAGAR: Great. Thank you so much, Bryce. I think that point the patient made is very, very true. So thank you for that. Gita, I want to ask the question we're really trying to get to here is about, and we will expand on this as we go throughout the day. It's about the differences and potential bias when assessing, based on the trial objective. So when we look at tolerability information versus the have them E information. And you can bring your expertise as ODAC member and researcher. But we want to hear your thoughts on PRO objectives.

>> GITA THANARAJASINGAM: Yeah, Bryce and Martha made this distinction between PRO being used as looking at efficacy endpoint. But understanding how patients feel and function while on treatment, and their experience with adverse events is more use of PROs to evaluate the tolerability of therapy. Now, the former situation with the efficacy PROs is where the concern of open-label bias tends to be highlighted the most, but I still want to emphasize there's value to collecting PROs that might reflect improvement on disease related symptoms and open-label trials. The latter situation of using PROs as a key component to understanding cancer treatment tolerability is vital not only on randomized blinded control trials, but especially, in the more common open-label trials both late and early phase. And I want to hearken back to comments Wenora made which I feel like as an adult cancer doctor, we're equally as compelling to me as what the 9-year-old in the pediatric ODAC said. Wenora, you talked about pain and fatigue, and these are symptoms that across the line, it can be disease related or therapy related. But you said how important that was for you and that's really where it's at. We know tolerability information is informed by safety and treatment data, but centrally, we need these PROs in all types of trials. And I want to emphasize we're talking actually not only about open-label versus blinded. But I also want to introduce the concept of early phase versus late phase, and talk about the fact that understanding the patient perspective of the treatment experience even on Phase 1 trials help us understand what toxicity matter to patients, and which ones we need to follow closely as the drug proceeds through development. I'll mention and illustrate a point regarding where PROs were very helpful to me as an adult ODAC reviewer on a Phase 2 myeloma drug where we were looking at drug conjugate that caused ocular toxicity, significantly affecting the vision of patients.

And in that study, patients, this was the dream 2 study in myeloma. The investigators actually assess PROs using PRO-CTCAE and the EOR2C30 and excellent validated instruments. And the data suggested although the Vishal impairment could be severe, vision related patient reported outcomes worsen and then improved and among the patients with those symptoms, overall quality of life remain

stable. So the suggestion was that this was a drug which many patients on the study tolerated, you know, and did not have a substantial impact on their quality of life. To me, as an ODAC reviewer and as a clinician, that was really compelling and relevant PRO information in an open-label Phase 2 trial that spoke to tolerability and provided the patient voice on a key toxicity.

>> VISHAL BHATNAGAR: Gita, thank you so much. That example is one that people can take a look at in the ODAC material. And it was a really robust discussion when we had that Advisory Committee. So thank you for bringing that up and tying that in here because that's an example we haven't gone over so far today. So thank you. We're getting close to end of the session so I want to turn to the fourth question about unique considerations for PRO assessment when it comes to core outcome measurements specifically open-label trials. And just a reminder to put this in introductory slide, the poor outcomes, the disease symptoms, symbiotic adverse events related to treatment, physical function role function and other side-effects. And thinking about those 5 core domains, why don't we ask Terri, thinking about her work at the NCI, some of the domains you're collecting in terms of PROs may not fit perfectly into the core domain. Thinking about some of the work that you've done, what is your thinking around our core domains and also the patients that you see in neural oncology space?

>> TERRI ARMSTRONG: Thanks for that question. I think the core constructs have really allowed us to standardize inclusion of outcomes in trials. And the constructs that have been identified really occur commonly across cancers and represent things that are critical to our understanding of the impact, both the disease and treatment on how a person feels and how they function. And to amplify what Gita said, you know, this is what we do as clinicians. I'm also a nurse practitioner and I've seen patients on trial for 30 years in addition to being an outcome researcher. Both the disease and the treatment related symptoms occur in a person and can influence each other. Now, also about inclusion, I think we have similar data across studies to really compare and understand that impact in different populations.

And in our approach, we include these core outcomes in all of our studies, but we recognize in our population that there are other things that are important. And one thing is neural cognitive function. And inclusion of both objective measures of that and the patient's perceived cognitive effects really allows us to assess those foundational core constructs and the impact of those that are important in our population. And identifying what those questions may be early in the trial, following the guidelines put forth by the spear guidance and ISOQOL and other groups really help us what those questions are and inclusion of those can help us answer those questions of how they impact each other. I think another important thing is incorporation of these work in clinical care is really critical. And on my experience at the NCI, we found things that are helpful, including electronic data capture of the patient reported data. And finding a way to incorporate that to the clinicians so they're aware of that and use as a communication tool when they see the patient is really important. Clinicians are busy. And they have many things they're trying to do.

So more information inclusion, I think, is really key. And our studies, we have over 95% completion of these patient reported data, and that makes it even more

meaningful to understand what the impact is. And then lastly, I would say I would like to emphasize what Gita and Martha said about the importance of rare disease and early phase trials. And we formed the Office of Patient-centered outcome research, or OPCOR within the cancer research. And our goal is really to start incorporating these into the early phase trials where questions of tolerability often are the primary endpoint. And finding ways to evaluate some of these techniques and methods in our early phase trial portfolio. So thanks for the opportunity to comment.

>> VISHAL BHATNAGAR: Sure. Thank you, Terri. And it's been a pleasure working with you across government agencies in terms of your standing up for that very important initiatives. And I look forward to how things go over the next few years. Bryce, I want to ask you. Thinking about those core outcomes, again, as Paul sort of brought the slide up that showed that assessment frequency, you know, and also we're seeing in the Q&A Box as well. This idea of high-frequency assessment being potentially burdensome. Are there any sort of methods you can think of in terms of data collection that can help mitigate overall the burdensome idea?

>> BRYCE REEVE: Thank you, Vishal. Yeah, obviously for every single trial, the investigators need to take into consideration given what they're testing and the patient population, what would be best way to mitt gay the subjective bias, but also integrate this patient poor data in a way that yields important information at the right time to inform the research questions whether it's looking at treatment benefit or treatment tolerability. And importantly, just overall guidelines, I think it's equally important and needs to be said again there is to make that for anyone proposing a study, designing a study that they include multiple stakeholders. And importantly, I think having patient and patient advocates like Wenora to be involved in planning the study, they can provide unique feedback on how often and when to collect this data.

And importantly, it can also help to identify what outcomes to measure and can inform and help the decision about what are the appropriate measures of those outcomes. And then, again, thinking about open label trials, I would again, go to the patient advocates to ask them if they can provide a compelling information that you they can provide to the participants in the study, especially in the arm of what their role is and the value of participating in study. So again, including the patients as stakeholder would be critically important. And other considerations that need to be considered in longitudinal study, we always think of quantitative assessment collecting quantitative data. But I also think there's value to collecting qualitative data alongside when collecting the quantitative things. So you can subsample, run a subset of samples from each patient point and have more in-depth information about what's going in the patients' lives in terms of pretreatment and post-treatment there. So we can look at issues of expectations as well.

I agree that there is unique opportunity to do multiple assessments to get a reliable estimate. And investigators should consider multiple baseline assessments. It could be both pre and post randomization. And they can see for patients randomized the intervention arm, whether those scores change or differ from the control arm.

And then lastly, and I know this will be a topic for later panel is there are expectation measures which can also be included which gives a sense of the arms. Was there expectation for treatment benefits? So all these things together could help, maybe not remove, but mitigate the bias of patient outcome in open-labels.

>> VISHAL BHATNAGAR: Thank you to the entire panel for this discussion. We have done a really good job of answering the key questions we set out to. And we've also nicely tied in the second if third session. I'm going to briefly summarize what I heard and then we'll get to the Q&A. We have about 15 minutes or so to answer some audience questions. But overall, I've heard patients are experts in themselves, and provide their experience, and they do it for the trial, but mostly to help inform other patients and providers. So that's a really important point. We started to draw that distinction between the tolerability type of patient outcome and disease improvement efficacy outcome. And there's potential differences in the ways that bias can affect each of those.

So I think those are two really main points that I think are important. And then we've also clearly covered why open-label bias, sorry, why open-label trials exist and are feature of oncology drug development. So with that, I think we can go ahead and start answering the questions in chat. The first question I think that's an important question is, you know, whether collection of patient reported outcome from open-label trials, whether the PRO from open-label trials increases or decreases regulatory risk? So that's bit of a tricky question. I'll ask Martha who is one of the regulators on the panel and then I'll have Gita follow-up. So Martha, does open-label PRO increase or decrease regulatory risk from your perspective and your opinion?

>> MARTHA DONOGHUE: Thanks, Vishal, for the question. And you're right, it is a tricky question. I think it depends upon what you mean by regulatory risk. So, for me, regulatory risk means what is the risk that when we're making a decision about whether or not to approve the drugs that we're making the wrong decision based upon inaccurate information or incomplete information. And so if that is a definition of regulatory risk, then I would say that collection of patient reported outcome data from open-label trials decreases regulatory risk. Because it gives us that opportunity to have a more complete picture of the patient experience, what is this drug actually doing, and how is this affecting the patient? So if we have a high assurance and it's out to the U.S. patient population, then we have a good understanding as we possibly we can about the risks. So I would say it decreases our regulatory risk. I think how we use those PRO data will also impact sort of the degree of certainly that we may take the data in terms of whether there's bias that may adversely impact the interpretability of those data. So when we're thinking about tolerability, and use of PRO for tolerability, I would say there's a little less concern at least on my end of the impact of open-label bias or open-label bias will be there whereas if that PRO data is being used to support drug effectiveness on labeling claim, then I would say higher degree of scrutiny and would have a better level of understanding of what any bias may be applying into the assessment of that endpoint.

So I think in general, PROs add value to regulators, and that's just one piece of the reason why PO data collection is important.

>> VISHAL BHATNAGAR: Gita?

>> You had a comment. Go ahead.

>> WENORA JOHNSON: I'm sorry, Gita. I mean, it's just been burning inside of me, especially as a 3-time cancer survivor myself, identified with lymph syndrome and possible more cancers in my future. Every time I'm thinking about this information, I immediately just think of the analogy of, for instance, most of us have purchased things on Amazon. And a lot of times now, we don't purchase anything on Amazon without reading the reviews. And, sure, do those reviews help us make our decision? Yeah. At the end of the day, do we sometimes wind up purchasing it anyway? Yeah. And that's how I feel as a patient that, yeah, I know there's going to be some adverse effects out there, but at the end of the day, I want to feel comfortable enough that those adverse effects just don't affect my total quality of life, but provide enough information that's valuable to others out there who are looking for some kind of life-saving treatment.

>> GITA THANARAJASINGAM: Wenora, I think you're referring in very brilliant terms to high-quality PRO studies. Because when we all read the Amazon reviews, we look at the ones that were verified that have 600 thumbs up, right? And so I think when we think about doing this in a scientific rigorous manner in clinical trials, and we think about regulatory risk, which I'm not a regulator, so I don't know if I'm perfectly qualified to comment on that. But I think it has a lot to do with the quality of the PRO study. You know, Paul alluded in his introduction to, we get a lot of PROs that have been collected haphazardly in past. Someone mentioned we get PRO as an afterthought and they're not powered to differentiate clinical differences. And that's a comment I agree with. But when we have well-designed PRO studies, using well-developed validated measures, and seek, input from other regulatory colleagues or others with expertise in these assessments, many of which of us are on panel, these PROs are valuable. They're good quality and they're from a verified user. From my perspective, whether or not they're efficacy related PRO or tolerability related PRO or somewhere in between, because I think it's really difficult to bucket those two out. That data is extremely valuable. And, you know, speaking freely as a non-regulator, I feel like it would actually decrease regulatory risk as Martha mentioned, because we're scientifically factoring the patient perspective which is likely to give us important signals about a novel drug or regimen when it is released into the real world. As a non-regulator, I think of regulatory risk as FDA is the American population main defender of health and safety of new products coming out. And so if we can get that patient perspective in a high-quality well collected Amazon verified manner, that really has a potential to decrease regulatory risk.

>> VISHAL BHATNAGAR: Yeah, I totally agree with everything that Martha, Wenora, and Gita just said. I think that as a regulator, I think that PRO, particularly, when informing tolerability can provide regulatory reassurance essentially. As Dr. Pazzer has said recently, no drug sponsor will say their drug is not well-tolerated. And we really need to ask patients what their experience is when it comes to tolerability. So I agree with everything that was discussed and said, patient should provide their experience, particularly, when it comes to that tolerability component, because let's face

it, in oncology, we have very poorly tolerated medication and therapy. So I think that's a very important point when it comes to thinking about regulatory risk.

We have just a few minutes left. So I want to get through some of these questions. Wenora, I want to go back to you. There was a question about improving compliance rate when regarding PROs. And as somebody who has been administered like hundred questions of PRO and at a specific instance, are there ways to improve compliance rate? And after that, maybe, Terri, you can chime in with your experience deploying PRO in the clinic.

>> WENORA JOHNSON: I think the first thing is just making sure you set that clear goal. And then, again, it comes down to communicating. Communicating between the entire staff all way to the patient of what's expected.

>> VISHAL BHATNAGAR: Thanks, Wenora. Terri.

>> TERRI ARMSTRONG: I think amplifying what Wenora said, that's exactly right. And it's also communication not only between patient and the people doing the trial and the clinicians, but between the Pharma, the study team, and designing that study. So early introduction of these conversations, as the trial is being planned is really important. Making them part of the study and value to everyone involved in the study, there's a message there that we value this and you're not just filling it out and going into a box if never think about it. This really augments and standardized the questions we are going to ask you and need. And making it easy as possible for the patients to participate is really important with electronic data capture, and making sure they can use it, whatever type of device they use. And if they don't have access, providing that device is really important as well. Something definitely including that as patients come to the clinic. So the clinicians know there's value in that. If that it's part of the larger clinical trial question. I think all help.

>> VISHAL BHATNAGAR: Totally agree. I think including not just informed consent document, but having a trained staff and providing information in more than one modality to patients why they're answering these questions and advocating that to trial sponsors as well on our end. And thank you for providing your perspective and reassuring us. That's good advice. For a third question, we have a question about the P version of the PRO-CTCAE. There's a question regarding PRO-CTCAE P version not directly comparable, I'm sorry. This is looking at the PRO-CTCAE versus clinician reported CTCAE. So perhaps, Bryce, you can talk little bit about, since you were involved in the development of PRO-CTCAE and perhaps Terri and Martha after that, looking at the comparability of the PRO-CTCAE to the CTCAE clinician reported data and how those grades are described. Bryce.

>> BRYCE REEVE: Thank you, Vishal. Important point I want to make is that we're not trying to replace or remove the clinician's perspective in a reporting CTCAE. I believe, and I believe my colleagues who work in this area as well there believe that the patient outcome and caregiver version out there for us provides complementary information and provides a comprehensive evaluation of what's going on in that patient's lives. For the particular question over, for much more detailed and I can say on this

call there, I would refer to whoever asked that question to the actual article by David Freyer. And what we did in, and what I presented today there is, as you know, there's multiple attribute of PRO-CTCAE, like assessing like in frequency, severity, and interference. So we used a combination of the attribute to come up with a CTCAE-like grade. Now, the weighing of this was actually done, and you'll see in the citation in the Freyer article where we worked in the Delphi-like process. I don't remember, exactly like 200 experienced clinicians and we symptom-by-symptom, attribute-by-attribute, they informed how they ultimately aggregate the pieces of information. And clinicians tend to put a greater weight on interference, how that interfered with their daily lives. So what we see in the graph I showed you was the combination informed by clinicians themselves compared to other clinicians CTCAE. As a sensitivity analysis, we actually unpackaged each of those attributes and compared those to CTCAE and again, we see consistently the same trend of what I presented on the screen.

>> VISHAL BHATNAGAR: Thanks, Bryce. Gita, you wanted to follow-up?

>> GITA THANARAJASINGAM: Yeah, I just wanted to mention upcoming panelist Ethan Basch and Amy Dueck composed an analytic approach you can use to present PRO-CTCAE and CTCAE data and really get how that information complements clinician reporting. And do it in a concise way, because some of the PRO-CTCAE questions have multiple attribute. They ask about the severity and interference, and the frequency. Which is not something you get with clinician reporting. And so when we think about how can we present this data in publication, how can we boil this down? How can we create something that's a little bit comparable without reducing all the granularity of that PRO-CTCAE information, there are analytic approaches such as the composite grading method that is studied by Dr. Ethan Basch and Dr. Dueck.

>> VISHAL BHATNAGAR: As you mentioned, we're privileged to be joined by them in the later session. So teasing for the later session to hear from them as well. We have time for one last question. And I can take the first crack at it. So the question is what are the major risk sponsors for not including the PRO-CTCAE and other symptom assessment data, particularly, in early phase studies? And when is the most ideal time to start including the PRO-CTCAE in drug development program? That's a really great question and very timely because of many workshops and many public-facing events that Dr. Pazzer, Paul, myself and others have been involved with. And assessment of patient reported symptoms and function began as early as the first in human study.

If you are administering anticancer agent to patients, you should ask what their experience is. To answer the first part of the question, what is the risk? Well, the risk can be bear out that when you take a drug that is poorly tolerated and you take it into a late phase study, and although you may get a progressive response survival or you won't get an overall survival detriment and that may be due to poorly tolerated or safety issue with medication.

So, the risk is that you may take a toxic dose or intolerable dose and that's why project that's led by Dr. Shaw was born. And so, I think there's plenty of literature and

public-facing discussion about those potential risks. But to answer second part of the question, and then I'll kick it over to my colleague Martha, the marriage risk is that, sorry, the most ideal time would be early as possible. There are well evaluated tools and methods that can be used. Martha.

>> MARTHA DONOGHUE: Thanks, Vishal. I think you've encapsulated the risk quite well actually. But particularly with respect to dosing and making sure that you are getting the information that you into throughout drug development to help guide your further drug development decisions, the dosage regimen you have taken to your registrational trial is really important. And we started to see more and more examples in something we're increasingly focusing on is how is pill affecting patients? And if we're asking those questions late in the program, it's too late and too late to make corrections, almost too late in many circumstances. So that type of information really should be captured early on and throughout drug development. I think the other risk is that when we as regulators are looking at an application and we see a safety signal, if we don't have complementary patient information to interpret that safety signal, oftentimes we're going to be more conservative in our assessment of how likely that is going to impact patients. So getting the information, using another dimension, hearing about what the patient's take is regarding that particular adverse event would be helpful to us, and I think be more likely to result in a more favorable assessment if we feel more confident that we're getting the full picture regarding certain big safety signals that came up during the development.

So I do think it's a risk. And I do think there's always the added burden to the sponsor of collecting the data, but I do think the value of that additional information will outweigh the risk o sponsors as well.

>> VISHAL BHATNAGAR: Thank you so much, Martha. Thank you to all the panelists. We're just out of time. So Bryce, Wenora, Terri, and Gita, thank you so much for a really great panel discussion to kickoff the workshop day. Thank you to the audience and stay tune for session 2. We'll get into analytics and interpretation from data of open level trial. We'll resume session 2 in 15 minutes in 11:45 Eastern. So thank you, all.

[Break]

SESSION 2

>> VISHAL BHATNAGAR: Thank you for joining us today. I Vishal Bhatnagar and I'm the medical oncologist and medical director of Oncology Center of Excellence. Today theme is for Clinical Outcome Assessment in Cancer Clinical Trials on onocology. Next slide.

In the first session, we heard from panelists that open-label trials are a reality in oncology drug development, particularly, in area of rare cancer, pediatric trials, and early phase of cancer trucking. We've also heard that patient reported outcome can and should be collected in cancer trials regardless of trial design, and that they can be used informs of dose, and tolerability, and it be used as supportive evidence to

traditional clinical endpoints. Tolerability data from cancer trials, particularly, when theoretically collected to inform the impact of side-effects, and well-known validated measures are available and should be used in the pediatric or rare disease population. And, finally, patients are experiencing the treatment in their disease first-hand and inclusion of the patient experience and drug development can reduce bias and reported safety data.

And in the Ferris session, we also discussed the existing evidence from the literature in this space. We determined so what degree open-label bias exist from some of these literature and how this informs patient outcome. And some were written by FDA authors, and in 2019, my colleagues wrote a viewpoint article highlighting many of the issues we discussed in session 1 and we will be discussing throughout the workshop today. Specifically, this viewpoint article explorers ways in which potential bias can be detected, including detection of differential missingness between arms and pre and post examination PRO responses and sensitivity analyses to test the robustness of PRO findings.

And this article serves as a prelude to the presentation and discussion for this session where we're going to take an analytical approach of prospectus of open-label bias when interpreting PRO and oncology trials.

To that end, the goal of this session are to discuss the following key questions. First, whether it's possible to use available trial data to determine whether open-label bias has influenced PRO results. Next it's important from analytic interpretation standpoint to understand the differences between descriptive and comparative PRO analyses and ways in which open-label bias may impact either something we started in session 1 and we'll continue throughout the workshop. And third, we'll explore ways that open-label bias can be mitigated which includes trial design and analysis methodology. And then finally, as with the case in many areas, clinical trial experience has an impact on routine clinical care and vice versa, so we'll discuss how the analysis method and impact of that data can carry over to interpretation of PROs in routine oncology care and vice versa.

Next slide.

For this session, we're pleased to be joined by a knowledgeable and experienced group from diverse background to provide their perspective on these topics. Dr. Ethan Basch is a physician in chief at the North Carolina hospital and chief of oncology at the University of North Carolina. Dr. Selena Daniels is a clinical outcome assessment team leader at FDA. Dr. Mallorie Fiero is a lead mathematical statistician in the Office of Biostats ticks at the FDA. Dr. Jessica Roydhouse is a select foundation senior research fellow at the University of Tasmania and worked previously in FDA group as fellow. And Patty Spears is a scientific research manager and accomplished patient advocate in the University of North Carolina.

For this session, as we have been for session 1 and for session 3, we'll be going by first name. Before launching into the panel discussion, I want to take a step beyond the article I just mentioned from 2019. Jessica Roydhouse is going to take many of the

concepts of the early she authored and she actually published last year. And so after her presentation, we'll moderate the discussion panel. But Jessica is going to provide the presentation upfront in the session. So take it away, Jessica.

>> JESSICA ROYDHOUSE: Thank you, visual. It's quite early in the most of these for me so please be patient with me. Next slide, please. So, I want us to back a bit and talk about why we would care for trial open-label. And this is well covered by the panelists in the first session. But I wanted to go and think about the original FDA 2009 PRO guidance. And pullout couple of key points that are relevant. There's a focus on labeling claim as the quote above emphasizes. Second, there's implied difference in arm in particular if someone is in investigate nation active arm, there's the overestimation of benefit. Whereas, in the contrast of inactive arm, this underreporting improvement. And emphasis on how patients are responding to questionnaires. And looking at the evidence here, there's been several reviews on topic in quality. And one was large review by Page and colleagues. And what they have as subjective outcomes. In the substance encompass PRO and clinician reported outcome. And this was a broad review. It expand multiple clinical areas from pregnancy to critical care, to osteoarthritis and others and covered different interventions from pharmacological if non-pharmacological. And Burkeman and colleagues sound lack of blind may exaggerate effects for subjective outcomes. But they also found stronger effects for lack of double blinding in trials limited to specific area. And last, there has been recent work by Mustgaard et al, where they found no evidence indifference in treatment effects. Now, I think this point has been well made by the panelist in the first session, but I really like this table and I tend to put it up a lot when talking about this. The data are from Hirsh and colleague. And there's been talk about this oncology and why these matters so much, because we can see open-label trials are featured in the oncology landscape and it's much that way in oncology compared to non-oncology.

So if we think back to some of these larger reviews I mentioned in the early slide that cover both non-pharmacological interventions and very different indication such as pregnancy, there can be some challenges in extrapolating those finding with open label oncology. First, oncology drugs have a toxicity profile particularly when we compare it to other clinical context. And second, while there's talk, you might have seen literature about the placebo effects and role of placebo, a true Placido trial oncology and someone receives a placebo and receive nothing else is comparatively rare. We will usually have some kind of active agent on the control arm. And lastly, this point has been raised before, but I think it's worth mentioning when we're making a comparison of double bond open label there is this issue of possible unblind and side-effects in an oncology trial. And, therefore, we might be really talking about kind of the spectrum from double blind to partially unblinded, and then completely open-label never blinded. So now I'll be talking about the work that happened and was published last year.

So this is the question for this work was what is the impact of treatment for myeloma RRMM for short. We carried two trials from the FDA database. They had different investigational agents, but importantly they had the same controlled agents. They had similar compilation or RRMM and blinding strategies were different. And also, very interestingly, they were PRO collected from two time points before treatment.

First at screening or prior randomization and even patients in open-label trial who have not yet been aware of the assignment. And second baseline, after randomization, therefore, assignment could have been known.

Now, I want to briefly define what I meant by impact in study and the question is what is the effect of knowledge of treatment of assignment on patient's outcome? In this study, the focus was on efficacy and two arm randomized trials. And the FDA treat that as focused on label suggesting efficacy, but as panelist from the previous session indicated, question of bias is relevant beyond this and syllable arm trial are very important feature of the oncology landscape. So the focus on efficacy and true arm is meant to minimize those and this is just as important for the study. And consideration in the study were initial impact. So before treatment was received in the so-called treatment before knowledge of assignment. And what would happen if people were receiving treatment which is a little more challenging because there's also the question of the impact of the cancer treatment.

Now, for this study, the focus here was on mean score change which was supplement with responder analyses. And there were several domains of interest that we wanted to use to answer our research question. We looked at symptoms, in particular fatigue. And then we looked at multiple function domains, emotional, social, and physical. And finally we looked at global health status. And the time period were the initial impact prior to treatment. And then while on treatment which is 2 and 6 months.

Now I want to introduce some caveat. First this was an exploratory analysis with only two trials included. There were several statistical analyses that these aren't recommendations for people for trial design or anything else going forward. They were appropriate to answer the questions at study. I'm not saying everyone should do this. Second, domains evaluated are not likely recommendations. For example, social function not a common endpoint, but we consider it important to have address out of hypotheses. And choice is driven by the study goals.

Now the hypotheses were that whatever impact might occur largest early in the trial. Particularly if we saw something happening where there was nothing beyond someone knowing what treatment they got, a large change here could thus be possibly explained by impact of relative treatment assignment. We also hypothesized that whatever impact we would see would be greater in domain that were further away from the biologic impact of the therapy. This is particularly in emotional and social function as "the" were anticancer and not anti-sign language site therapies and large change could be a result of the impact of knowledge of treatment.

Now, looking at the baseline data in general we did not see much. So the change were quite small. And the social function domain had the largest effect, but what I like to mention is that effect was still quite small. In terms of directional, we saw slight worsening which was not significant for the open-label trial. And during treatment, again, we did not see much in general. And when we looked at social function at 2 months, there was actually inconsistent finding for the direction. And we had several methods used in study. And at 6 months, it was significantly worse for

open-label trial, but only with that time frame. And now, to me the important takeaway from this work first we did not find a mention full differential effect of knowledge of assignment on reporting. We did find limited support for the hypothesis about distal domains, but social function where the effects were strongest was not a score outcome. Now, I like to raise couple of other considerations and first is there are other mechanisms for impact. The focus in this study and in the guidance is about how people answer questions. But there are some other ways, and one is differential dropout prior to treatment. This has been seen in few trials as far as I'm aware it's not systemic assessment of health when thinks happens. But there it's not common either. And patients leave trial after knowing their assignment and before treatment. And couple of times it happened, it has been predominantly in the control arm. Second is about differential completion of PRO themselves for patients on the trial. This was another investigation undertaking and I put the link to that there on bottom of the slide. And first and positively, we found completion was generally good. And this is a strong message support for PROs in oncology trial which is great. We did find infrequent occurrence of differential completion and infrequent better completion for the investigational arm and open-label trials, but again, it's worth emphasizing this was not common.

And I think the response here would be mitigation through design, for example, site education, other approaches to emphasize the importance of PRO completion and minimize missing PRO data which is important for other special agents on onocology trial as well, not just this concern about open-label bias.

And so, therefore, I think the way forward-thinking about this is thinking about how we can mitigate it through both design and consideration through analytic approaches and that's what we'll talk about in the panel. So thanks very much for your time, everyone. That's the end of my presentation.

>> VISHAL BHATNAGAR: Thank you so much, Jessica. I think your experience with that exploratory analysis is really a great way to sort of set the foundation for this panel. And for those of you in the audience who are looking to read this paper more in-depth, this article was actually provided as background reading material in workshop and you can access it there in the PDF format. Now moving on to here from our panelist, I like to hear with my FDA colleague, Selena Daniels. So Selena, besides the method such as differential reporting as described by Jessica, are there other ways you can consider to assess whether there was potential open-label bias within the trial?

>> SELENA DANIELS: Thanks, visual. While talking to patient or obtaining patient input is always a good approach. One option we can explore is leveraging the industry or exit interviews or even surveys. Pretext entry interviews or surveys can be used to assess patient's expectation to the study treatment prior to the trial where exit interviews could be surveyed to issue patient knowledge to the effectiveness to the PRO measures. Using these interviews can be a mechanism. Another approach may be to include an instrument within the extent that measures the concept we may not expect to improve. Similar to the divergent validity and explore those responses.

>> VISHAL BHATNAGAR: Thanks, Selena. If going back to Jessica, since you published this analysis last year, have you considered any other analytical methods, perhaps even more advanced methods you can think of?

>> JESSICA ROYDHOUSE: So the first thing I think is one that was done in a publication by my former colleague Belinda and she looked at this single arm trial thinking of persistent of effect which is a good way to think about this. And the second I think is sensitivity testing. And thinking about this kind of quantitatively and also this issue of is this kind of spectrum of looking at this in trials where it was possibly double blind, but then there was unblinding which has also been done. So these are also helpful ways as well.

>> VISHAL BHATNAGAR: Thank you. And I think the other ways that we can think about is using sensitivity analyses as mentioned. And I just want to go back to Selena and think about whether besides PRO completion rates, thinking about things concomitant medication, have you seen this done Selena?

>> SELENA DANIELS: Actually I have, and I believe PRO results can be further supported by finding other endpoint by sensitivity or subgroup analyses comparing to the findings relative to other datas collected in the trial. Sometimes we see sponsors assessing pain intensity and when we see reduction in pain intensity, that's further supported by seeing reduced analgesic use. So that's one way. Or looking at symptom improvement and seeing it's more associated with significant tumor shrinkage. We can also compare PRO results with other COA that's less objective like performance outcome assessments or comparing it to digital health technology. Additionally, one can explore aspects of a concept. For instance, assess s intensity and frequency if these are important attribute for particular indication and we expect them to improve, and then comparing these results to see if there's any differences between those two endpoints. One could also think about optimizing the study design to evaluate whether or not knowledge of treatment assignment influences a patient's response. Feasible and ethical one approach would be to include a randomized draw phase where patients are considered responders if randomized treatment or randomized placebo or best care during that phase and look at the changes, if any, and how patients are responding to that symptom assessment. We can also look at the patterns of patient response, for example, examining temporal patterns and how it holds against a drugs mechanism of action, or each pharmacokinetic and Pharma profile.

And another approach could be use of cross over design where patients on placebo or best supportive care are allowed to receive active therapy after the primary has been met and look at change in the patients responses through the treatment arm. For single arm trial, one could consider learning data by incorporating historical control. However, from a measurement standpoint, every effort would be made to ensure patient between the assessment used and clinical trials with natural history control to a loud meaningful change overtime.

>> VISHAL BHATNAGAR: Thanks, Selena. I want to ask my colleagues from the FDA to weigh in here. So Mallorie, from your point of view, what are some things

you're looking at when examining PRO data from oncology trial and trying to assess open-label bias when you're looking at datasets?

>> MALLORIE FIERO: One of the main things we might first look at is patient attrition. Because in open-label study, one concern could be that once patients are unblinded to their treatment, that they could be with drawing consent at a faster rate for the control arm compared to the treatment arm. And something like this would impact how we would interpret patient required outcome and not just PRO but all he was endpoint would be impacted if we saw patient attrition.

The treatment effect itself if it's a large treatment effect of patient-reported outcome, might mitigate some concerns we might have about potential open-label bias. But remaining cognizant that there could be bias in overestimating the treatment effects. So that's something we would very carefully examine and I know that Selena provided a lot of great approaches kind of to evaluate open-label bias. One other approach is to have an open-label extension study where we have a double blind phase first and then there's an open-label phase afterwards, and then something we could do is we could evaluate whether we're seeing different PRO results in the double blind phase versus open-label phase.

>> VISHAL BHATNAGAR: Well, thank you, Jessica, Selena, and Mallorie for those points. Starting to think about the analytical method. We're going to switch gears here and hear from Patty and Ethan about some of the dimensions of how the potential for open-label bias differs between a descriptive tolerability PRO endpoint and a comparative efficacy type endpoint. As we mentioned this is going to be a thread throughout workshop. So Ethan, perhaps I'll start with you. How would you look at -- how would open-label bias look for a descriptive tolerability endpoint?

>> ETHAN BASCH: Thanks so much and thanks for the invitation to join in today for this really important workshop. You know, listening avidly to panel one, I have to say to the choir, my availably published data on the study of oncology trials, unblinding or knowledge of treatment mal indication or open label study does not meaningfully impact or influence patient reporting of PROs. So it is my belief based on, again, the available data, the available evidence that we shouldn't be worrying about this anymore. And that the patient voice should be considered valid and meaningful in any context, whether that's looking at comparative efficacy, comparative tolerability, or at descriptive tolerability.

I think I have some slides that we could load up. But you know, the other comment I'll make, and again, I think this is echoing comments made in panel 1 and implied by Jessica and also, Vishal made about is keynotes should be included in all cancer trial drug related development. And this particularly applies to tolerability. And while my slides are coming up, I'll make a quick note that measuring tolerability descriptively is substantively different from comparative tolerability, particularly, in the regulatory or drug development context.

So, descriptive tolerability is essentially asking patients or investigators to report on adverse events. And then reporting that information, again, descriptively and often

in reporting trials, in recent publications, there's no statistical testing done on those. Comparative tolerability is different. Comparative tolerability is similar to comparative efficacy. It's taking a tolerability endpoint, but looking at it using statistical testing and methodological rigor to try to delineate between study arms. In such cases, the same criteria that we would apply for looking at comparative efficacy would then be applied to comparative tolerability for the regulatory components of considering those endpoints.

I'm going to move ahead with these slides. So I'm specifically going to speak to this point of looking at PROs in open-label trials. First, for characterizing symptomatic adverse events such as PRO-CTCAE, again, restricting these comments to the descriptive collection of adverse event information to characterize the tolerability of a product, not to look at comparative tolerability.

So the first point I want to make, I've really already made in many respects, which is understanding the patient experience of adverse events is foundational in all human trials, regardless of trial design. Right? Patients are the ultimate users of these products. They're a component of the characteristics of these products that are only known by patients through their reporting of their subjective experience. There are many studies now showing that we, as investigators, are not able to rely or adequately capture and document that information in clinical trials. And therefore, the only way to capture this information is to collect it directly from patients themselves in the form of PROs.

PROs for adverse events should be included in all human trials for cancer drug development, certainly, all those that include some other mechanisms for adverse event reporting. If we think about cancer clinical trials in the drug development setting, the CTCAE is reported for trials. Where CTCAE is being employed, the PRO-CTCAE or another PRO tool to capture adverse events should also be applied, because without including a PRO tool for capturing adverse events, and only using CTCAE, or investigative reporting, we will have inadequate understanding of tolerability of that product.

Now, I'm going to show couple of slides in a moment about a potential model. Now, the standard approach to applying PRO-CTCAE or PROs in clinical trials is to collect that information, not to share it with site investigators or clinical staff and then to analyze it. But there is an emerging model to potentially share this information with investigators at the point of care to inform those investigators AE reporting, and also to potentially impact on clinical care. If this brings me to point number 2 on this slide. Which is that PROs have now been shown to positively impact outcomes in their own right. If you could go to the next slide, please.

We now have evidence from multiple prospective randomized clinical trial as well as from other large population data that collecting PRO data on symptoms and specifically symptomatic adverse events from patients during systemic cancer therapy and sharing that with the care team in real-time leads to improvement in clinical and utilization outcome and including survival. These are overall survival curves. We have seen in these studies, including some that were conducted by colleagues of mine or by my group that by sharing this information, there's an improvement in patient

quality life, symptom control, reduction in hospitalization, and emergency utilization, and improvement, again, in overall survival.

This is a strong rationale for systematically collecting PROs in clinical practice and sharing with the oncology team. So extending this model to clinical trials for adverse event monitoring, one can think about following model which my colleague Amy Dueck and colleagues employed in a national clinical trial through the NCI equal groups. This solicited a form for investigators to complete a number of different adverse events. Now, on this form, they can select the adverse event and grade at using CTCAE criteria. And an alternative approach is to collect AE information directly from patients and share it with the investigators as shown on this alternative rendering of the form. So here you can see that same form is the prior slide, but we have inserted into it the patient self-report of their own adverse events. This information comes from the PRO-CTCAE and can be used by the investigators to self-report. In this clinical trial, as reported at ISOQOL in the past, it was found that by showing or sharing this information with investigators, there was convergence of the clinician or investigator reports with patients. The clinician or investigator reporting of adverse events became more patient-centered. And we call this patient informed investigator reporting of adverse events. One does not have to use this approach, but one could use this approach, if this will then take advantage of the patient self-reporting to inform the investigators and also potentially to inform clinical care.

The last comment I want to make is based on a question through Vishal well do we introduce a problem with investigators? Are we going to change the finding of the study? And the answer to that generally is no. First of all, because the adverse event information has already been reported by the patient and recorded. And the second is because if we think about way adverse events are reported and the patient comes into the visit, they're in the room or on the phone, they tell the clinician or investigator about that problem, and then the investigator reports it. The investigator is aware of it already. So the standard approach to adverse events reporting for symptomatic is investigated. And this is merely an extension to make it more symptomatic and rigorous. I'll stop there.

>> VISHAL BHATNAGAR: Thank you. That was a great presentation and you answered one of those questions live. So that was a great question provided by the audience and a very timely presentation when it comes to answering that question. So thank you for that. Patty, I wanted to get your perspective on personally on what Ethan has presented so far. And also just thinking about what we've discussed in the workshop as a whole, but in particularly in this session, when it comes to using PRO as a critical element in understanding tolerability.

>> PATTY SPEARS: Thanks for that question. I have a lot to add. I think the first session was so good. And so much of it really overlaps, but it's complemented about what we're going to go over. But I think tolerability, you know, descriptive tolerability is fine for patients. I think that would be a step ahead of what we get already. And I think comparative would be good if you're comparing to treatments but it's not necessary for patients to the degree that it's necessary for the statisticians and the trialist and things. But patients are going to use it in a little different way. And I

think without the patient reported outcomes, you really can't even say anything about tolerability. And I think that, you know, there's an overuse of that term when PROs are not even included. So I think that we need to kind of change that language around there.

Because many patients are faced with choices when they go into the clinic on what they're going to you know what drugs they can take and what path they're going to go down. And the more they know about the treatments and how they're going to affect their daily living, the better choice they're going to make for their individual lives. So I think it's better for patients to know that ahead of time. And I think this was brought up on the previous channel, and I know in the chat, but it is really important for patients to know this ahead of time, because there are supportive care. I know when I went through chemotherapy, I was on a lot of supportive care. My oncologist was like let me know what you're feeling because there is a drug for that. And mentioned forewarned and forearmed, I definitely agree you can mitigate a lot of the side-effects you may be experiencing. And as a patient, you don't know what side-effects you're going to be experiencing because these are different drugs and anything you've really taken. So if something happens, you won't know if it's definitely a side effect and you won't know to tell your doctor and things like that.

So I think in reporting, so patient-reported outcome versus clinical reported outcomes, too, and the patient-reported outcomes, you're asked a specific question, so you answer them. And then off the bat, you know what's important to the doctor and what's important for the drug, and what's important to report. When the clinician reports their CTCAE, I'm pretty sure they don't does the patient all those questions and they just assume if they didn't mention it, they didn't have it which is a bad assumption. That might be where that bias comes in. In the model Ethan presented, that's really good focus so you know what to ask and what might not be a pro and you don't have to ask at the end. And mentioned in the previous panel, I think the PROs for the patient and what the patient is facing should be the gold standard. And that's a safety thing rather than how the patient is feeling and functioning.

And I think that avoiding bias and I think what we heard from the previous panelist about the different ways of doing that within the studies, I think it's about time to do that. I think we need to really concentrate on getting the PROs and best way we can back to the patient in a way they can use it. And I think it's really important to collect it in all clinical trials, regardless of our CTCAE blinded or single arms, and it's really good to see the NCI working to getting PROs into phase 1. I've been talking about that for many years and I've when harping on it many times. So it's nice to see the shift going. So we need to make sure that happens, because I think that if it's not caught early, then it will go down that path of developing, and then you can't really bring it back. It's too far gone, like we said previously. So I really think it needs to be in every type of study and in every type of trial and it will really benefit patients in the end.

>> VISHAL BHATNAGAR: Thank you, Patty. I think you summed up a lot of things we talk about in session 1. And I want to return to you in just a moment to hear your perspective a little bit more about the idea of open-label trials. But before we do, I wanted to ask Ethan, you know, your presentation provided a lot of great examples

comparing the patient reports to clinician report, but are there ways within the patient report realm to sort of provide comparison and support or verify symptom data? So I'm thinking use of things like summary measures, whether it's overall side-effects or use of comparison and things like symptom data to looking at function data from patient report as well. So to get a better understanding of the trajectory of a symptom.

>> ETHAN BASCH: Yeah, thanks for the question. So I think what you're asking is whether there's some benefit to adding summary measure in addition to the individual metrics for specific symptomatic adverse events. If this is useful in the overall context estimation of tolerability or side effect burden, and items like the fact that GP5 have become of interest. And so there are many cases in which individual PRO-CTCAE items are being implemented in a trial to look at specific AEs alongside the fact GP5 as a summary metric as an overall burden and validity check on the trajectory. I think that in trials, you know, sometimes there are, you know, if we're thinking about a comparative trial, there could be a trial in which there's worse diarrhea in one arm and worse neuropathy in other arm. We see this in GI trial with platinum drugs. And one has to trade those off, certainly, as a physician, I'm discuss those trade-offs with patients when selecting a treatment. But having an overall summary can help to understand the balance between arms in that setting, or in the case you presented just have an overall sense of the impact of tolerability. The potential value for summary metrics, if one includes items for individual AEs, but we miss important AEs, then having a summary measure might give us an overall sense of what's going on from the patient perspective. And I would add as an adjunct to that, I'm a strong advocate for including a free text or open item along with PRO-CTCAE to elicit from patients directly any additional symptoms if they may be experiencing during a trial that we may not have solicited from them in the PRO-CTCAE form.

>> VISHAL BHATNAGAR: Thank you, Ethan, for bringing that up. I'm going to ask you a question later on about that free text. So hang on to that thought. Before we move on to the next question, I like to sort of return back to you, Patty, and I think it was something important that Ethan said about the difference between a comparative tolerability versus just the descriptive tolerability question within a trial. And from the patient perspective, are you really interested in the comparative tolerability of a therapy? Or are you more interested in just the treatment that I'm receiving, what are the tolerability considerations? And the second part of the question is going to be whether that comes from an open-label trial or not, does that sort of impact how you view or absorb that data?

>> PATTY SPEARS: Yes, so I don't think patients are going to go back and look at what kind of trial it came from. I think both are important. And I think that, you know, descriptive is important. And I think comparative is important. But if there's so many barriers to comparative, that the descriptive doesn't get done, then I think that's a bad thing. I think they're both important and they both have a role. So I think that we need to keep an eye on, really, what we're doing and we're doing this for patients. If we can give them a descriptive first, and then a comparative, then that's fine as well. Like Ethan said, with different drugs, you can pick and choose different drugs based on your preference and the toxicity that it has. And your daily living and how it would

affect you specifically. So it's really important, because open-label studies are here to stay. I think the randomized trials and double blinded studies are just not very patients-friendly and patients like to know what drugs they had before or after. And so we need to get this right, because they're definitely here to stay. And I think that, you know, one of the things that also came up was about patients will report their toxicities. And I think they will report it as best they can regardless of whether it's a single arm study or a double blinded study. They don't kind of characterize the study based on things like that. So they will report back accurately regardless of what kind of study it is. And I think to ensure patients really report exactly what you're looking for is to inform the patient, when you start the study, what the purpose of the questionnaires are, and how important they actually are to the study. And if you really communicate with the patient about their importance and why you're doing it, I think they're going to really be more accurate in the reporting.

And also one of the other things was frequency as well. And reporting frequency, if you report more over time, if you tell the patient why you're doing that, I think they're more likely to kind of do it every time. But if you don't tell them every time, they might ask why you're asking them often. And if you give the patients the questions at the right time of appointment, I think that helps before-and-after the appointment. And we have electronic now, and rather than in the clinic when they're stressed and when they have a stressful appointment. So I think that will be well. And Ethan was talking about the summary type of measurement. I think summaries are good, but I think the individual type of side effect is very, very important to the patient. And you don't want that to be rooted out if it's one thing, it's really causing most of that problem. As a patient, I want to know it's mostly GI issues and not respiratory issues. Right? So I think they're both important, but we shouldn't lose the detail in the summary as well.

>> VISHAL BHATNAGAR: All those points are very well said and you have echoed something Wenora brought up when it comes to the measures, you lay the ground work which is very nice for presentations happening in session 3. So thank you for sort of providing that comment. It really is important for what we will discuss. Thanks, Patty. I will now move on to the other types of question we have for the session. Thinking about comparative efficacy type here and endpoint. Because that is one of the sort of major concerns when it comes to thinking about open label bias. So I'm going to ask my colleague Mallorie. Thinking about PRO that informs improving disease symptoms and those endpoints. Are there ways to examine this data for open-label or potential open-label bias and that type of data? Mallorie.

>> MALLORIE FIERO: Thank you, Vishal. One of the things I wanted to note which has already been in session 1, oncology trials, our primarily efficacy endpoint is generally things like overall survival. Progression free survival that we use for decision-making. And in general, patient required outcomes are secondary efficacy endpoints, and our supportive endpoints except for unique cases. And I think it's very important to understand patient experience regardless of whatever the trial design may be. Now, from a statistical perspective, there will always be some concern about the potential open-label bias, especially, for a comparative efficacy claim particularly if

there's a concern about false positive claims. And there are some ways to work with that. And I think one important note to make is that we can work with our division of clinical outcome assessment colleagues to better understand and measure we're looking at so we can feel that the instrument is reliable when we look at the PRO results. And one last thing I wanted to note from my perspective, open label trial randomized are different from single arm studies, and solid tumors for efficacy. That's because in solid tumor, control arm is generally necessary to be able to evaluate the treatment effect if that we know treatment effect is not just due to the natural history of the disease. That's why we use measures like objective response rate, because we feel like it's a direct measure of clinical benefit. And, certainly, for single arm studies, patient reported outcomes are very important for tolerability objectives as well as looking at some descriptive measures.

>> VISHAL BHATNAGAR: Thanks, Mallorie. Jessica, do you have any thoughts on this particularly analytical methods for open-label bias when examining the efficacy type endpoint?

>> JESSICA ROYDHOUSE: Thank you, I wouldn't say analytic method but as panelist emphasized, they're here and important. Rather than saying there could be bias, I think some of it which relates to other efforts in the field such as time of framework and missing data, this question of sensitivity and thinking about how robust the findings are to safe potential open-label. And that's a potential point and part of thinking kind of quantitatively is it a bias but question is how much bias is there and would it change our interpretation of the PRO data?

>> VISHAL BHATNAGAR: Thanks, Jessica. And lastly, I'll ask about any consideration how this differs from earlier when we were discussing tolerability concerns and your opportunity to chime in on the efficacy of this type.

>> SELENA DANIELS: I just want to emphasize the PRO objective. It's important to establish the objective of PRO will be used in the open-label studies, whether they're talking about PRO data to support efficacy labeling claims or capturing overall experience for the purposes. I think that that's key, as an agency may require more evidence for particular objective, for example, if we're talking about PRO data to support the objective of efficacy labeling claims, of course, we want the instrument to be fit for purpose of its particular context of use, but we also want confidence that the PRO response are interpretable and uninfluenced by the knowledge of treatment of assignment to ensure the claims are not false or misleading or if we're putting it in the product label. And maybe help us assess open label bias.

>>VISHAL BHATNAGAR: Thanks, Selena. That makes me think of sort of the way that Jessica ended her presentation and maybe Jessica, you can probably, you know, sum this up better than I can, but use of trial design. So I'm hearing from the last few comments that perhaps there is not necessarily ways to, you know, analytically detect bias, but perhaps trial design might be a better mitigation step. So Jessica, do you have any thoughts on that?

>> JESSICA ROYDHOUSE: So yes. I think some of the other panelists have

mentioned as well importance of communicating with patients, I think, also communicating with trial science. And this is also, I think, possibly beneficial for mitigating missing data, which is a very important thing anyway and very useful and in that regard, you know, sort of prevention has always been a cure for missing data. So I thinking through the design, how you will approach it, and having this part of the strategy going forward, this is a way that you can help many other ways with open label design.

>> MALLORIE FIERO: I want to add to that. When we are looking at a comparative claim, it is important that we respecify, you know, the [audio skipping] endpoint that would be included in the statistical analysis plan for a randomized controlled trial. And in general, I would recommend using all randomized patients when we are analyzing this just because we want to make sure we have a comparison between the treatment arms, and don't see any bias due to any differences between the treatment arms.

I definitely echo Jessica's comments about making sure we put in place strategies to mitigate missing data. This is a key issue that we will always look into whenever we are reviewing PR data. And the last note that I wanted to make is that, you know, the comments that we make here reiterate that we can't extrapolate these comments to other disease settings because we have heard in both sessions how unique the oncology space is, so we can't extrapolate these comments particularly when the COA endpoint or base endpoint is used as a primary.

>> VISHAL BHATNAGAR: Well, thank you, Mallory, Jessica, Selena for a robust conversation on that topic. So we will switch out to sort of the last question, I think, that was included in our key questions for this session, and thinking about the clinical trial experience, and how that informs your routine clinical care, but also vice versa, how in the routine clinical care space patients are aware and it is important to report outcomes, you know, outside of the clinical trial as well. So Ethan, with your research experience, what are ways that analytical methods can be used from the routine use in the clinical trial space and perhaps vice versa?

>> ETHAN BASCH: Thanks so much. Yeah. So you know, as I alluded to earlier there is now pretty robust evidence suggesting clinical benefits routinely collecting patients during cancer treatment and sharing that with the team to manage symptoms, thereby improving all kinds of outcomes. And there are a couple of, I think, related points to that to read into our discussion today. I think the first pertains to real- world data. Increasingly there is an interest in collecting information either retrospectively or prospectively, being classified as RWD. And historically, real world datasets have not included PROs, but as PROs are included more commonly as a part of routine clinical care, or as they are inserted into specific instances of clinical care intentionally with the future goal of harvesting real- world data, that information can become available to us.

But, you know, that information is subject to all the other limitations and other

real- world data sources, which it might not be as systematic or as rigorously applied as we would like in a prospective clinical trial. You know, that information can be used for all sorts of different purposes. It can be used to create synthetic or historical control arms for studies for single arm studies, it can be used to understand patterns of care as a basis for designing clinical trials, and could actually be submitted as evidence of the impact of drug as supplemental information in applications.

And you know, I think the biggest challenge in all in routine care for PROs is patient adherence. In clinical trials for PROs and prospective clinical trials we can see adherence rates from 90 to 95%, whereas in routine clinical practice it is more like 60, 70, maybe 80% if you really, you know, if you really use best practices.

But the principles are the same as in clinical trials, which is the -- has been mentioned earlier, there's great importance for teams to value the collection of the PRO data and communicate to patients this is an important part of what we are doing, and that generally engenders the highest compliance rates.

I don't know if there are other elements that you want me to highlight.

>> VISHAL BHATNAGAR: I think you highlighted real- world applications to clinical trials, and I think we have worked very closely with cross- governmental colleagues, thinking about ways to broaden incorporation of PROs into the routine clinical care space, and using the clinical trial experience to help inform that. I do hope that as we mentioned that compliance rates will improve in the routine clinical care space, but you are absolutely correct, that at least from our standpoint we have observed that compliance rates in clinical trials have gone up quite significantly, are north of 90% in almost all trial applications. So we hope that that improvement in compliance can certainly lead to improvements in routine clinical care as well.

I do want to have time for question and answers, so I think I will just have Patty provide the last word on the panel discussion, and you are welcome to give your perspective on what has been said so far by Ethan and his last point, and overall I do just want to allow you to provide your summary on this session and what you heard so far.

>> PATTY SPEARS: Yeah. So, you know, talking about, you know, what Ethan just said and presented in his presentation, you know, clinical care is the ultimate open label study, right? So it is definitely open label, and if they can do it in clinical care and do these great studies, I can't see why we can't do the open- label studies. But I can see a couple of important things that I wanted to bring forward. So I think it is really important that we gain information from patient- reported outcomes on tolerability in clinical trials. So that they can expect during their clinical care.

But, you know, in clinical care I think that, you know, Ethan showed that it really improves outcome of patients, and I really think that, you know, in clinical care is where you communicate the results back to the patient and you can mitigate the symptoms,

you know, at the time, so you might decrease hospital stays and different things. And then the patient really gets that comfort in knowing that they are reporting something back because they don't have to make that decision, should I call the doctor, should I not call the doctor, and you know, just report what's asked of me, and I know what to report. So it really is good for patients and it is good for patients across the board. And then I was just on a panel the other day and we were talking about the late effects of treatment. A lot of clinical trials do have limited followup, and you know, if you can get those late effect treatments the best place probably to get them if we can't do them with clinical trials because followup is too long and too expensive in a clinical trial, then clinical care in, you know, the normal practice of cancer, that's where we could possibly get some of that information because, you know, cancer patients are living longer and they are experiencing these late effects of cancer, and so if you do have a means of getting patient- reported outcomes through clinical care, I think we can inform that population as well.

I know when I walk into my room at the oncologist's office, it is a one- page laminated sheet that I go down and check. Very easy, very easy to do. I give it to the nurse that bring me back, she puts it in my chart, erases it, and it goes to the next patient. So you can make it so if if -- so if you make it easy for the parent, if you make it relevant for the patient, if you make it meaningful for the patient, I think we will get more patients actually filling them out as well.

And so I think that that's definitely a way that we can make it better and make it easier.

>> VISHAL BHATNAGAR: Thanks, Patty. And I think that's a great way to round up and finish this panel discussion. I would summarize what I have heard so far, but I want to go ahead and get the questions started and go back to you, Patty, and that's why I'm going to skip my summary until the end.

So one of the questions in the chatbox, that I would love to hear your perspective on, is preparing patients about what adverse events may occur is certainly important. Does this prior knowledge amplify the detection of certain events () at the expense of others? And I want to hear your perspective when you were administered a PRO.

>> PATTY SPEARS: Yeah. So every patient is different in how they interpret what they are told, right? So when you tell a patient, you know, you might experience this and this and this, they usually pick the top ones. They don't pick all of them, they are sort of rare type of thing. So if you have the information through, you know, PROs and collected on a collective group of patients, you would get those in a collective sense of, you know, what is really important to patients, kind of like that Amazon thing. What you don't want to do is have that small -- that rare site effect to be amplified by that complaining, that one person that has a loud voice, but nobody else is going to experience that. Kind of like the Amazon thing. You want 6,000 people, not 5, you know, things like that.

So, you know, I think if we do a better job of that systematic approach of PROs and then giving the information back to patients in a more comprehensive way and you can actually tell them, you know, the percentage of patients that experience different things and things like that, that it would really help a patient as well.

And then, like I said before, mitigating the side effects is really, really important for patients, and knowing that you can mitigate the side effects. Like if you experience this, let me know, and I can give you this drug to help you out.

>> VISHAL BHATNAGAR: Yep. Okay. Great. Thank you. The next question, I think, would be for Mallorie and perhaps Jessica, looking at is there a difference between compliance rates between open label and double- blinded studies. Have you observed that? And perhaps, Jessica, you can chime in.

>> MALLORIE FIERO: You know, I think primarily it will depend on the trial itself and certainly make sure that Jessica has time to speak because I know she has looked into this before. I think that it can open label trials can potentially impact differential completion between treatment arms, but as () mentioned, we haven't seen quite good completion rates or compliance rates regardless of the trial design.

So -- and again, this kind of goes back to the importance of looking at these completion rates for the specific trial to make sure we have good data quality before making any decisions based off of it. And I will hand it to Jessica.

>> JESSICA ROYDHOUSE: In the study we had a database, and we did really fine with completion, and that was a very positive finding. There was some evidence, but it was a really small subset of trials where there was asymmetric completion favoring the control and open label. But I want to emphasize this is really, really infrequent. I wouldn't ascribe it as a major concern, open label trials is bad. As I emphasize this issue of thinking about completion and validating the () PROs as part of the trial process is very important and mitigating through design and communicating the importance of filling in the PROs in a way that people think is meaningful I think would help address that in any kind of trial.

>> VISHAL BHATNAGAR: Great. Thanks, Mallorie and Jessica. The next question I will ask from the Q&A box, it is an FDA- directed question, the question is whether or not it is appropriate or possible to develop a new PRO instrument in a Phase 2 trial, and I think that the question really is highlighting the fact that heretofore we have talked about very well- known, you know, and validated PRO instruments. So the appropriateness.

>> SELENA DANIELS: So you could put it as a Phase 2 study. To be tested, you know, or piloted in the Phase 2 study. That's actually an opportune time to evaluate and accumulate evidence to support what's a meaningful support change in that instrument. It would probably be too preliminary to put in nonexisting items within a Phase 2 trial, if that's the question.

But also, if there isn't established or newly developed instrument that's in Phase 2, there are -- as I mentioned, you know, entrance or exit interviews where you could still obtain patient input. Or could still confirm, you know, the patient experience and the symptoms, to see what symptoms are important for them for that particular condition and make sure they are in the Phase 2 trial.

And that exit interview or survey stage, you could test to see what the newly developed tool is () measuring everything that it should be measuring for that particular patient, if it is important to them, as well that it, you know, they have experienced meaningful change within some of those items with that tool. So I will stop there. I will let you chime in on some of your thoughts.

>> VISHAL BHATNAGAR: I think it is really important to also think about, again, what the objective is in that Phase 2 trial. So if, you know, you are really trying to quantify tolerability, which makes sense, and we have just been talking about it at length advocating for early assessment of tolerability, I would say there is well- known libraries and other PRO instruments that can be used to assess tolerability. But my sense is talking about disease symptoms and efficacy.

And I would sort of say that it is probably too late at that point to develop a new instrument. That being said, it happens quite a bit. We have seen it, and if done well, then I think anything is possible.

But I would say if you are going to go down that road of developing a new PRO instrument, that you absolutely do need to interact with the FDA review division and folks like Selena Daniels very early on, even prior to that Phase 2 trial to understand what needs to be measured () and come to a consensus about the concepts that are being measured. And again, it is tolerability, a more efficient way of assessing symptoms and side effects related to therapy.

So that's how I would answer that question, is that yes, it is possible, but interactive with the FDA early and then come to some sort of consensus before deploying in a Phase 2 trial.

I'm going to take the next question. I will take first dibs on that and I will return back to you, Selena. How should we choose () for human study and is there an ideal that cannot be burdensome? And this is actually a question that we get a lot from drug developers when they may not know if the side effect profile is of their, you know, in their first () study.

What I will say is that, and we provide this advice freely, is that, you know, you may have some idea in a person human study based on clinical data of what types of safety, you know, concerns you may have. So, you know, usually those could be GI type or skin or other, but what I will say at a very bare minimum you should be looking at overall side effects further and you should try and select out the 8 to 12 most relevant

symptoms.

Overall side effects, even in the absence of known -- anything known about the safety of a specific drug can be quite informative of the and then the last thing I will say is use of the -- as Ethan mentioned, it has, you know, published on a free text item can be very powerful in a first human study because it can give you more information about unknowns. But when we developed our poor outcomes guidance we fully, you know, appreciated that we did not designate, you know, late () registrational. However, much of that guidance can be extrapolated to early phase study. Honing in on those core outcomes of treatment- related symptoms, expected ones, but perhaps also ones you may not expect. Physical function can capture a lot of tolerability information, and use of side effects as I mentioned.

So there are ways in which you can use the core outcomes in a first and human study. Selena, I will kick it to you, if you have anything to add.

>> SELENA DANIELS: No. I think you -- it is the challenges of () in the first human studies. I would agree with you, you probably want to select items that are consistent with some of the core side effect for that drug product or class of drugs, right? And in addition to maybe that free text item, if you could talk to patients after the trial, you could sort of get a sense of what side effects were most important to them and most bothered them and sort of use that to inform what sort to select for your next phase trial.

>> VISHAL BHATNAGAR: Thanks, Selena. And I will use this opportunity to pivot to Ethan. Ethan, there is a question about the free text item. So perhaps I can ask you that question, and if you have any other thoughts about those past questions. But the question is, Ethan, if free text is used within clinical trials what is the best way to summarize and report that data?

>> ETHAN BASCH: Thanks. Great question. So first, thank you. I will reflect on the prior question first, which is in that first in human or 1 trial, items to select. This is a quickly evolving area, but there have been multiple publications in this space, and I would echo what both Selena and Vishal said. I would advocate for including first any suspected AEs from preclinical or class effect, preclinical work or class effect that we might suspect ahead of of time.

The second would be to include AEs from the core set of symptom items. That's different from the core domains, the core items for inclusion. And Rice Reed published some years ago a paper outlining core symptoms that are cross- cutting across context in oncology and hematology, there are around 10 of them, and those can be included as well. It was based on a systematic review of prior trials plus a large NCI dataset. So I would encourage including those, as well as free text, as well as a summary measure like GP- 5.

The other thing that one can do in early phase trials is have open- ended interviews and discussions with patients that can be challenging in early phase work, but can be considered to try to find signal to use in later phase studies.
In terms of analysis of the free text, which is the question that you posed, in general the way that free text items have been analyzed is to map them to existing standardized () cons, and Medra, mapping up to higher level terms, that has been done successfully in multiple prior trials, depending on the approach one wants to use, you know, I would recommend having two reviewers of that information who can each look at the data. And then in cases of disagreement, they can confer with each other for consensus. There probably isn't enough disagreement that you would need a third person to adjudicate that.

That's an approach that has been used multiple times, and then you will find that there are unmapped items, and those would have to just be reported individually, and you reduce them to a summary term or terms or you could report them verbatim. That's different from the question about how to use that information, you know, for choosing dose, but I will defer on that one.

>> VISHAL BHATNAGAR: Well, thank you so much, Ethan. I think we are just about out of time, so I appreciate the whole entire panel discussion. What we have heard here is in this session is really critical, I think. Thinking about, you know, analytical ways, I think Jessica started out the session, you know, with a really important accomplished exploratory analysis examining this prospect. It really moved the field forward. I think this discussion really honed in that good trial design and use of well- crafted endpoints and just good trial practices can help mitigate potential biases, and we started to broach the line of clinical practice and clinical trials.

So I think this has been a really informative and helpful panel discussion. We are going to take a 15- minute break. So we will rejoin at 1:15 Eastern, and in Session 3 we will talk a lot about where do we go from here. So again, I would like to thank the Session 2 panel for their input and discussion, and look forward to seeing you in Session 3. Thank you all.

[Break]

SESSION 3

>> VISHAL BHATNAGAR: Welcome, everyone, to the last session of today's workshop. Again, my name is Vishal Bhatnagar, Associate Director for Patient Outcomes in the Oncology Center of Excellence. This session will be moderated by myself and Paul Kluetz, who you heard from at the outset of the workshop, Deputy Director of the OCE. Next slide.

We are pleased to be joined by an excellent panel of guests. Yelak Biru, CEO of the National Multiple Myeloma Foundation. Melanie Calvert at the University of Birmingham. Angelo de Claro, division of hematology at the FDA.

Dr. Amylou Dueck, Associate Professor at the Mayo Clinic, and finally, Dr. Devin Peipert, Assistant Professor at Northwestern University. And we are going to go by first names for the panel, like in the first session. Next slide.

So we had a great workshop so far. Really robust discussion. I just wanted to take an opportunity to recap the workshop so far. So we heard in the first session about how open- label trials of various designs are a reality of oncology drug development and we heard there is a need to collect high quality from all types of trial designs. In Session 2 we thought about ways to measure, to detect the magnitude and also ways to mitigate -- sound trial methodology. Next slide.

Specifically in this session, which is meant to be a forward- looking session, we plan to discuss some examples of how PRO has been collected with open- label trials and has been communicated. We will then continue to discuss those differences between how described PRO data should be communicated versus comparative treatment effects from open- label trials. And how should PROs should be used to characterize tolerability moving forward. And we will discuss how PRO data to inform tolerability can be communicated outside of the drug label, such as methodologies like Project Patient Voice.

Before we begin the panel discussion just as we have gone in the previous sessions we want to hear a presentation from one of the panel members and then we will use that as a foundation for the panel discussion. So I would like to invite Dr. Angelo de Claro, who will look us through a useful example or set of examples of how patient- reported outcomes have been labeled from open- label trials to inform these improvements. Go ahead, Angelo.

>> ANGELO de CLARO: Thank you for the invitation to present. Next slide, please.

My objective for this presentation would be to review the use of patient- reported outcome endpoints for efficacy labeling in drug approvals for chronic graft versus host disease, my talk is as follows. Discuss -- we will review the disease, followed by description of the use of PROs in the clinical trial setting, as well as how this were used to support FDA approval, and a discussion of the framework. Next slide, please.

Disease overview, chronic graft versus host disease. This is a multisystem inflammatory disorder that occurs in about 40% of patients within the first year after hematopoietic transplant, stem cell transplant. This condition affects multiple organs including the skin, mouth, eyes, liver, GI tract, lungs, joints, and others.

Clinical manifestations include skin rash, thickening, oral or GI tract ulcers, dry, painful eyes, inflammation of the muscle/fascia, liver, and lung abnormalities. And the figures from the slide show some of the clinical manifestations of the disease. The top row is the dry eye and lung infiltrates. Middle row shows histologic findings, inflammation as well as fibrosis of the tissues, and the bottom row depicts the skin rash,

pretty extensive in this condition, as well as oral ulcers.

So based on these constellation of findings this is an asymptomatic disorder which makes it a great candidate for use of a PRO endpoint. Next slide, please.

The standard efficacy endpoint for this condition is a clinician- assessed response, and this is just a brief description of the response criteria. As you see, it involves assessment of multiple organ systems that are involved in the disease, and the clinicians will assess whether there is a complete, partial response, or progression in each category, and all of these are integrated to provide you an overall response rate. Next slide.

For chronic host disease there is an established instrument. A Lee Symptom Scale, LSS, which is a 30- item scale with 7 subscales which are described here, and these, again, correspond to the major organs that we described earlier to assess heterogenous symptom bother and impacts of the disease. This is a one- month recall period, and a 7- point decrease in the LSS score, as well as duration of that decrease is part of the evaluation. And a 7- point decrease is considered clinically meaningful.

There has been a modification to scale called the modified LSS, and this has -- there is 28 items, deletes 2 items based on supportive care needs, and the items deleted were use of supplemental oxygen and use of () nutrition. Modified LSS is 7- day recall period, and 6- point decrease considered to be clinically meaningful. So just to place in context, generally the [audio skipping] encountered for these in clinical trials, you are looking at baseline scores of 3 to 3.5, the score ranges from 0 to 100, so hence a 7- point decrease, if you are citing () a 20% decrease in the overall score. Next slide, please.

So this table depicts FDA approved treatments for previously treated and these were based on open- label trials. The approval was the first one in 2017, approved after failure of one or more systemic therapy. This was our first experience with -- this is the first FDA approval for this indication, as well as our first exposure to the use of the PRO for the LSS scale, so for -- to support the approval the basis. The efficacy was established based primarily on the overall response rate of 67%, which is clinician assessed.

We thought it would be useful to also include in the labeling the patient- reported outcome, the 7- point decrease which was reported in patients. We have applied this similar framework to support subsequent applications submitted and approved, including the Belumosudil, here again, the efficacy was established on clinician assessed response rates of 75%, and on the patient reported outcome endpoint response rate was 52%. Also approval in 2021, this was a randomized trial, open- label design, and this design allows you to describe overall response rate between the active arm in the best available therapy arm, and similarly also reported patient reported outcome as well for these trials.

So all of these examples were open- label, and certainly that was one of the concerns for the review team as we were reviewing this applications, and this can has been difficult to conduct a double- blind randomized trial due to the underlying disease, as well as toxicities associated with these various treatments. Next slide, please.

Discussion. So what were the factors that facilitated use of patient- reported outcome endpoints for efficacy to support these approvals. First was availability of a parallel, a clinician- assessed response endpoint. And the discussion in today's workshop, the patient reported outcome, it was complementary information that provided additional context to the efficacy findings.

We had established PRO instrument for this condition. Please note, this instrument is by no means perfect, but we thought this was a best place to start, and we have used this to support and provide advice to multiple sponsors to advise them on their development programs for this indication.

One of the advantages that we have prespecified time points for collection of the data as well as a prespecified analysis plan. This certainly facilitated the review of the data and also from the sponsor side for monitoring and minimizing the missing data issues.

And I think reflecting across multiple approvals as we have gotten more experience, I think we have become more comfortable with use of this framework, and I think it was () with the first approval, but I think we have shown with more experience that this has been a very useful measure for us in order to assess efficacy.

Now, challenges for use of PRO endpoints for safety labeling. I think one of the challenges would be needing to have a good idea of like what types of safety issues or tolerability, like what time point would this be most relevant to capture, and all of these probably would be better addressed if you have proper dose optimization and a good understanding of your drug's safety profile prior to the initiation of trials for registration. Next slide, please.

I can share with you one example where FDA oncology included a PRO in the safety section of a U.S. prescribing information, and this is for Crizotinib. So adverse reactions. Under disorders, you can see the text shows that we provide a description of the reported adverse reaction rate of () occurred in about 63% of patients. Majority of these were 95% were Grade 1. And subsequent text, we provided findings that the clinical trials also had visual symptom assessment questionnaires that were administered, and here we were able to provide more granular information. We were able to reproduce the majority of the patients had visual disturbances, and lasted up to one minute, and probably the patient reported this visual disorder have mild or no impact on (). So it is one example of like us using this in a safety side.
I look forward to the discussions. Back to you, Vishal.

>> VISHAL BHATNAGAR: Thank you so much, Angelo. I think that that was a

really interesting use case of, you know, how patient- reported outcomes can actually be used in a different way that we have talked about so far, from an open- label trial, and that's specifically looking at the efficacy of an agent, and two of those cases, those were actually single arm trials as well.

So just to go -- just to ask a specific question regarding, you know, those three examples that you provided, you know, was there any methodology that was used by the review team to consider, you know, whether open- label bias -- or examining this data overall?

>> ANGELO de CLARO: Certainly during the review's applications, for single arm trials we were dealing with like 40 to 60 patients, and having to deal with that open- label bias was something -- so it was of concern. I think what mitigated the risk was having that complementary like clinician assessed assessment. And they were moving in a similar direction, so that provided assurance that what was being reported by way of the patient was also being observed by clinicians. So kind of cross- check each other and provide you better context of like what the treatment effect was from an efficacy perspective.

It is also notable that PRO requires -- the way that the PRO was set up, you need a large magnitude of treatment () in order to meet that threshold. So it is interesting to see like the PRO endpoint response rates were actually lower than the clinician- assessed responses for the same disease. So I think this sort of mitigates the patient outcome measures were overestimating the treatment effect. So I think it is how you set up analyses, and initially we had multiple internal discussions, should we do this. And I think looking back it is good we did because I think this has further helped develop the field. We have high- quality data that we are able to review, and share experiences of that and see how this could be applied in other settings.

>> VISHAL BHATNAGAR: Thanks, Angelo. I think I very much appreciate those comments, and I know that much of those analytical methods and sort of that thinking that went behind incorporation of the PRO into at least the first case was well- described in the literature piece by our colleague, Belina. Very well- described and there are many that describe the thinking behind that labeling instance as well as measures that were taken to try and make sure that bias did not occur. So I appreciate that you brought that example forward.

I would like to welcome the rest of the Session 3 panel to get their perspective on what's been said at the workshop overall so far and then also Angelo's presentation. So first off, Yelak, what your perspective is how PRO is communicated, whether it is in the drug label, as Angelo just mentioned, or otherwise. Giving your perspective on communication of PRO and then dissemination of PRO to patients and healthcare providers.

>> YELAK BIRU: Thank you, Vishal, for the invite and thank the FDA for putting this important symposium together. Patients really want to look at two things when

they engage in shared decision- making with their doctors and choose the next best treatment. The one is does the drug work. Does the drug have efficacy. If it doesn't work, then I think almost no need to really talk about quality of life, because patients don't want that drug and we need to find a better drug.
But if the drug works patients want to know what quality of life they should expect, whether it is positive or negative, while being on this drug. So patient- reported outcomes of any sort, open- label or not in my opinion will add significant decision in choosing the best treatment in collaboration with our physicians.

In order for that to be effective, though, PRO from a scientific measurement to a real- world use. One is, which was talked about a little bit, a lay version of the PRO report prepared for audiences, and the second one is there needs to be probably a better, greater standardization of PRO assessment tools in communicating with patients. An example, I deal with several different myeloma measurement instruments, PHQ, FACT MM, and there are many of them. So I think there needs to be a lay abstract or a lay abstract type of communication that accompanies the PROs, and we should work towards better standardization of PRO measurement tools.

>> VISHAL BHATNAGAR: I think we are going to revisit your idea about patient- friendly language in abstracts that communicate that data. But before we do, I would like to turn to Melanie Calvert. Because I think the word standardization was used, and you have done a lot of work () to help improve standardization. So let's hear from you, Mel.

>> MELANIE CALVERT: Thank you. I think you have raised an important point now, and we have got a lot of tools now that are being developed, international methods that can help the community with this. So Vishal mentioned the protocol guidance, also the work of the () consortium, which is analysis guidance, but then there is also crucially relevant to these discussions in particular the consult guidance from reporting data and also the graphical presentation of trials and there has been work by Claire Schneider and Mike in that area. And all of that's really available on the consortium website. So you can access that and that's a really good source for the community.

And random reporting and spoke to the patient as well. I think in the drug trial snapshots for FDA where presenting results by age, by gender, by race, I think that's also potentially going to be more important moving forward, as we start to move forward with the ways we report PRO data, can we tailor that more to subgroups so people can really access that and understand what's going to be meaningful to them and their decisions around treatment. Thanks.

>> VISHAL BHATNAGAR: I think that's really important, and something you touched on, Mel, here about informative subgroups, and you know, I know you have done some work in the Diversity, Equity, and Inclusion space and, you know, thinking about that is there anything, hearing about open- label trials, early incorporation of PROs in general, are there any thoughts you have on diversity and informative

subgroups and reporting of informative subgroups from open- label trials?

>> MELANIE CALVERT: Yeah. So, I mean, I think it applies to all trials really in terms of thinking about diversity and inclusivity, and we have actually just recently published comment on that topic and talking about key actions. I think we are transitioning to use of electronic PROs, but we know people in society not everyone either have the skills, not access to devices, not comfortable using those devices, and perhaps you use interactive voice recognition or have iPads in clinic, or other modes of administration might be necessary. I mean, we have to remember that as we are moving forward, whether it be open- label trials or whether it be randomized trials or real- word, about how we make sure that we all are inclusive and equitable in the way we collect this data. Otherwise, we risk about having bias samples from subgroups of the population that might not allow us to evaluate care appropriately.

>> VISHAL BHATNAGAR: Thanks, Mel. And Yelak, so I will give you a chance to respond before we move on to the next few questions.

>> YELAK BIRU: Sure. Thank you. I think this is an important topic, and in order to be PRO in clinical trial or any normal clinical practice, it is important to make sure that PRO data () clinical trial and clinical practice include patients that are representative of the disease. So results are meaningful, intend to treat population. So whether it is race, ethnicity, gender, sexual orientation, economic makeup, all of that are important. And you mentioned, Vishal, doing a really nice partnering of partnering () patient organizations to provide best practices and guidance in this area, and one such example is recent guidance for multiple myeloma community. I think () was the last () that I attended before pandemic.

So for () clinical trial, though, whether it is pre or post approval of drugs, PRO, recommendation that the FDA has in my opinion needs to go () and say in order to get a drug approved, industry needs to provide -- prove that the drug has been tried on the intent to treat population or have a plan to followup to do that. So I really do believe that not just race and ethnicity, but all of the other factors that create a subpopulation in clinical trial or clinical practice and () patient reported outcome need to be considered.

>> VISHAL BHATNAGAR: Thank you both. I think that, you know, you have sort of taken a turn a little bit, but I think it is an important point to make, especially in this session. And I would be remiss to not mention much of the work that's been done in the States by () equity on that line. So I encourage the audience, if you are not familiar with that, to check that out.

I will now, you know, turn to key question number 2 for this question. Thinking about differences and how to communicate PRO data, but tolerability type data versus the () symptom data. So we are first going to think about this from a tolerability standpoint and communication of that.
First, I will ask Devin Peipert, you know, are there any differences that you can think of when it comes to communicating this type of information?

>> DEVIN PEIPERT: Yeah. Absolutely. So right away I think we are focused on one of the topics that we mentioned earlier, which has to do with the difference between descriptive analyses for tolerability versus comparative approaches. The descriptive is much better developed and can be really well- informed by great, great research that's already out there, as Mel mentioned, Claire Schneider, () as well as from Amylou about how to visualize maybe single items describing tolerability in ways that patients expect to and can comprehend very easily. Expect to see and comprehend very easily in settings that they like and that's really gotten quite detailed about the type of colors to use, etc.

And I that grows out of our descriptive approaches. So I think that would be a great place to start.

>> VISHAL BHATNAGAR: Thanks, Devin. Amylou, I'm going to ask you now, are there any considerations that you can think of, especially with all of the data visualization work that you have done, in terms of communicating the type of safety and tolerant data?

>> AMYLOU DUECK: Yeah. So we know that the typical analysis that's conducted on the clinical adverse events is lacking, so specifically when you report just the rate of Grade 3 or 4 adverse events, things like the onset, the duration of adverse events, this is not communicated, not putting aside all of the issues with kind of the overuse of the broad conclusions about generally well- tolerated as well as issues related to the physicians underreporting many symptomatic adverse events.

You know, but I think using PRO adverse event data and particularly descriptive approaches can really aid in our understanding of the patient experience, and you know, even beyond that doing the visualizations, as Devin mentioned, you can really dive into that data and get a rich understanding of what's going on beyond just summary measures.

>> VISHAL BHATNAGAR: Thanks, Amylou. So Angelo, this is something you mentioned at the end of your presentation, thinking about communication and tolerability information. Do you have any thoughts based on what Devin and Amylou just said?

>> ANGELO de CLARO: Sure. I agree with the previous comments, and just want to () really have a good understanding of like what is the correct dosing regimen and having a better understanding of the safety profiles, where you can adequately capture the data of interest. And you talk about are you looking for a safety issue or a tolerability issue, like an early concern during the treatment that goes away or is it something that has more like a delayed effect, or is it something just low level throughout? I mean, the collection methods and how you would describe these could be different. So I think having a good understanding of tolerability profiles would go a long way to designing an adequate framework in order to capture and analyze this

information.

>> VISHAL BHATNAGAR: Thanks, Angelo. You know, Paul, this is something that you and I have talked about at length, and this is something that's been, you know, discussed in previous sessions, I think, by Ethan. What are your thoughts about this type of communication or this type of tolerability data?

>> PAUL KLUETZ: Yeah. I mean, it is an interesting discussion when we talk about descriptive versus inferential statistics on tolerability, I think that's the first question is what's the objective. And when I mentioned earlier on that we are pretty solid with our objectives now, we think tolerability is a universal objective, I think we have done work on displaying single item over time with Project Patient Voice. And I think the issue of open- label trials with descriptive safety data has already been shown. We put safety tables in the FDA labels from open- label trials and those include symptomatic adverse events.

So I think descriptive tolerability data is well on its way to create, I think, a helpful visualization that's different from the safety tables. Comparative tolerability question, that's a more challenging question that I think is of great interest to a lot of folks and we are going to have to all put our heads together to how to show statistical superiority to one arm versus another.

>> VISHAL BHATNAGAR: Thanks, Paul. So we are going to keep things moving and think about ways about how to communicate this important tolerability information moving forward. So -- and that includes things like summary measures. So we are going to hear from Devin Peipert, who is going to provide a presentation to considerations about communication of this information moving forward, and then we will resume panel discussion after that. So take it away, Devin.

>> DEVIN PEIPERT: Thanks, Vishal. I should have some slides popping up here.

You can go ahead and advance. First just have some disclosures. So if you could just advance those.

We have heard some of this already today, but in case someone has joined late, I think the major distinction we want to make when we are thinking about tolerability is that, you know, in general PROs have been added into the trial protocol in a similar way as, say, efficacy endpoint that can be appropriate for an efficacy purpose, but of course there's tolerability purposes and, you know, I don't think we know exactly how to fit that into the protocol yet and how to treat, you know, as Paul mentioned just before. There is key issues around analysis differences and so we just want to think about those two things differently. Go ahead and advance.

So thinking about tolerability, just looking quickly at a standard definition, the ICH defined tolerability as the degree to which overt adverse effects can be tolerated by the

subject, disease, sign, or symptom, that's the paradigm we are coming from. If you advance, I will show a more patient- oriented definition from an effort, including many folks on this call, led by the friends of cancer group. They offered a really exciting new, as I say, patient- oriented definition of tolerability.
We can read it here. I think the underlying portion, a complete understanding of tolerability should include direct measurement from the patient on how they are feeling and functioning while on treatment. That echoes something we have actually heard a patient advocate. If you aren't talking to the patient what do you learn about tolerant. So I think that helped us move in that more patient- oriented paradigm, of course, you know, the use and the incorporation of PROs historically in cancer trials was pushing us in this direction as well. Go ahead and advance.

The same paper included this really nice and instructive display around how to divide up the types of data we would collect to get at different aspects of tolerability, highlighting standard assessments as well as patient- reported tolerability assessment, which integrates the patient experience. These include on the patient reported side symptomatic adverse events, overall side effect burden, patient reported physical function and other types of functional assessments. These map on pretty nicely in the FDA's most recent core patient reported outcomes cancer clinical trials. I would like to note also in that guidance one key distinction for tolerability as we are thinking how that might differ from efficacy is recommendations around assessment frequency, whereas, you know, if we are trying to get a read on tolerability, we may want to include much more frequent assessment of a short set of tolerability items than would be administered for just the efficacy endpoint purpose. Go ahead and advance.

Let's dig a little bit deeper what we mean by patient reported adverse events. There are a lot of great options out there, but I think they are what they sound like intuitively, which is the types of adverse events we are used to, but that can be reported on by patients because they are about the patient symptoms. And so we have PRO- CTCAE. All get to different did I things mentions, and get at a rich understanding of symptom frequency, severity, and interference in daily life. Go ahead and advance.

I love this slide, which came from a paper led by Paul that gets at another category, which is overall side effect burden, and we have heard some pros and cons to this potential way to tap tolerability in the previous session, but it can capture the impact and potentially cumulative impact of multiple different types of toxicities, not just within one treatment, but across treatments where the toxicities may differ, and () in the previous session as well can be informed by () as well.

So this is where my research focus is most, and I think I'm very interested in how it might evolve this to the point where it is included in sort of a standard approach to understanding tolerability along with other options we have. Go ahead and advance.

Just looking at some specific ways you might get at overall tolerability with the overall side effect burden, probably the most common approach is to use the FACT Item GP5, I am bothered by side effects of treatment. So this is already routinely

assessed in many cancer trials. There is a similar item from the EORTC system, IL46, slightly different wording, both are options. And this is sort of what we mean by overall side effect, getting at something more global which could capture multiple different toxicity impacts. Go ahead and advance.

Thinking about tolerability in the context of open- label trials, which we know is the dominant type of trial, looking at this paper by authors, as well as Dave and many others, you know, I think that paper sheds light on sort of the specific, you know, potential impact of open- label for tolerability because they looked specifically at what they call proximal PROs, which include things that we know to be like patient- reported toxicity, fatigue, nausea, vomiting, pain, dyspnea, etc. And I thought it was really a clever analysis, looking at whether or not blinding had an impact on whether or not the active treatment favored PROs, and .08, not statistically significant. Just like for all PROs, these sort of toxicity or tolerability type PROs didn't really show any, you know, apparent bias. These types of analyses could be extended to include what I just showed in terms of the overall side effect burden items like GP5 and that could probably extend our understanding of potential for open- label biases. Go ahead and advance.

Early phase trials had come up a little today, too. There is great work emerging, including a fantastic paper that came out yesterday in Oncologist, I would suggest looking at that, but at the time preparing slides it wasn't available, but other research here in the basic rundown, I think we have all come to sort of identify and talk about today the strong potential for patient reports of trying to define tolerable dose, but it is not done. So a systematic review from a few years ago identified only 15 trials with PROs of over 1300. Phase 1 trials. So I think this is an area that requires more research in terms of the appropriate methodology to include PROs in early phase, especially dose finding trials, but I think we can get there. Please advance.

And finally, just to wrap up by highlighting that I did with a patient advocate, Marylou Smith, where we looked at, you know, differentiating some important concepts within tolerability. So, you know, tolerability, we have been talking today about how to capture the patient experience side of that, which I think is most important in terms of selecting a tolerability measure for inclusion in trial, but as we learn about tolerability there is also going to be analyses to try to understand what influences tolerability, what really defines tolerability from the patient's perspective, and leads to whether or not they stay on treatment. And there is a whole element of patient disposition, which is to say what the patient walks in with and things like their overall willingness to stay on treatment, preferences, and attitudes that likely influence their experience and ultimately their tolerability of any given cancer treatment.

But I think we need to crack that open as well. And with that I will close.

>> VISHAL BHATNAGAR: Thank you so much, Devin. I think that really provides a really unique perspective on things like a summary measure. We have talked a little bit about the session panels, but I think you have provided a really nice summary of much of the work that's been done. I would like to hear from Amylou.

You have done a lot of work with the consortium, and there is a lot of Tolerability Consortium members on the call. So I would like to hear from you a little bit about the work that the Tolerability Consortium is doing at our sister agency.

>> AMYLOU DUECK: The Tolerability Consortium, 4 teams of investigators funded through the NCI Moonshot Initiative to investigate tolerability endpoints and analyses in cancer trials. We also incorporate NCI, FDA, patient advocates, and other stakeholders to help our work. We are also designing a website right now to try to disseminate all the work that the 4 teams are working on. So be on the lockout for that. In general we are -- each team has a little bit of a different flavor so, you know, my group is particularly interested in developing standardized tables, graphics, and summary measures. Devin's group, the Evolve team, is looking at overall symptom burden, as well as variety of other analyses. Supria, University of Rochester, is looking at measure in older adults, and then Patty and Andrea are looking at how individual patient co- variants impact measures of tolerability in the relationship there.

So focusing on my own team, what we are working on is proposing standardized tables, graphics, and summary measures, and initially we focused on CTCA type of analyses, and -- but ultimately the analyses that we are working on now go beyond this. So these are -- they appear similar to these maximum grade type approaches but account for the patient baseline symptoms. They also can be applied at the individual attribute level or can be applied at the symptom level when applied to PRO- CTCAE type data.

We also support doing a deeper dive into analyses, so summary measures are useful for summarizing your data, but also look at the longitudinal visualizations. We dive into other visualizations as well, because it is such rich data and available to a wide range of stakeholders.

There is a lot of work still needed to communicating about missing data. Also integrating measures of variability, integrating patient characteristic, and what type of sample sizes do you need.

>> VISHAL BHATNAGAR: Thanks, Amylou. Yelak, we were talking about this this session. You had a really interesting comment that I would love for you to share about these types of overall summary measures, specifically overall side effect. So could you provide your perspective and share with your group what you think about some of these summary measures?

>> YELAK BIRU: Sure, Vishal. So pre- pandemic if I was going to go to a family reunion and my uncle asks me how are you doing, the answer I give is usually good or okay, and I move on. And I don't really tell him the full story. But if he asks me did you have a flat tire when you went to see your Aunt Mary, my answer would be more specific, yeah, I had a flat tire, I went to the gas station, I paid this much, and I will give -- I give this () the specificity of the question matters. So many questions like how are you doing can be answered vaguely, and lead to uninformed or inactionable

resulted.

But on the flip side, exhaustive 120 questions, 20, 25- minute answers may also water down what the patient's response may be, and I may be responding more mechanically to finish the questions than to be able to answer that question more specifically.

So I think there needs to be a good balance between simple and repetitive. But if I can challenge us also to aspire to deburden the patient, as was spoken earlier, with the use of mobile apps, with technologies, and digital technologies, I think it would help make the answers more meaningful, deburden the patient, but also allow us to be able to potentially reply to the patient more realtime or near realtime to their PRO needs. I think there is data that shows that data that get near realtime answer to their PRO distress are more likely to better outcomes than those who don't. So I think they should make this available to patients, to look after or look back, in the decision- making doctors as their doctors, I think we should also aspire to continue to challenge ourselves to use PROs in a more realtime and create a balance between the one answer question and the half hour question.

>> VISHAL BHATNAGAR: Thanks, Yelak. I think that this really segue ways, this is something that Paul and I and many others have discussed, you know, internally and I think it is worth describing here. You touched on a few of those things. What's the use of a single item versus multiple items? And Paul, I would love to hear more thoughts about that, but the other is use of digital and wearable technology to help inform realtime PRO.

So Paul, I would love to hear your thoughts on that.

>> PAUL KLUETZ: That's great. Such a great comment, Yelak. And one of the things that we have seen a lot of and has kind of played the field was using an even more broad question to try to get at side effect bother, which was how was your quality of life. And that is a single question that's actually the most commonly used health- related quality of life tool that's in most of commercial trials, and we don't infrequently see differential toxicity in two arms with the patient () outcome conclusion there was no meaningful difference in quality of life between the two arms are for instance. So has always been an inherently tolerability question. So while it is still broad, I'm bothered by the side effects of my treatment, it is still much closer to what that actual objective is, which is are you tolerating your therapy, you know, you are benefiting.

On the question of a single question versus a scale, which is really the question you are asking here, the single question clearly has the advantage of it being () I love to use, you know. It is very easy to administer, but it has the drawback potentially of not being very sensitive to catch smaller, maybe incremental situations of tolerability.

One of the interesting challenges that we have with PROs, even if you have a

very sensitive scale, you then end up asking yourself whether the small change you saw was meaningful. And therefore, we put an anchor to determine whether it is meaningful, and that anchor is a single question. So there is a really challenging PRO dilemma that's almost like a circular argument. So across the board I think a single question is fine when you have an objective that's really related to comparative tolerability, maybe you should have both, you should have a scale that's more sensitive, and you can also have a single item to help you understand the magnitude of the different things you see.

>> VISHAL BHATNAGAR: Okay, Paul. So, you know, thinking about the need to elaborate the use of PROs in general in early phase trials, which is something that came up in the first session quite a bit, but also in the second session, I think it is worthwhile to hear from this session panel to hear more about the use of PROs in early phase setting. And again, I think this was something that was very popular in the comments and Q&A box.

So Melanie Calvert, do you have any thoughts in initiative you have in terms of increasing PROs in early phase studies?

>> MELANIE CALVERT: Yeah. Thanks, Vishal. This is something that we have been working on and writing about recently. And my experience with this is in the nononcology setting recently, so just to put that out. But the single question or multiple questions that Yelak said, this work was published, () has successfully been used. In an early phase setting we have found that it was very challenging to identify which [audio skipping] symptoms, () focus within that, where we got very limited data at that point, you know. We looked at the literature, we considered the mechanisms of action, but it was quite hard to know whether there is unanticipated side effects that might be cropping up. So we actually opted to go for the whole use the free text option that Ethan mentioned and we are exploring that.

So I can let you know maybe in a few years how that's gone, but I think the other thing is I do see value, real value in collecting early phase setting. Devin mentioned the work done with Christina Yap (sp) on toxicity trials, think that might be an area for future work that's really important. And I think beyond that it can give you preliminary data on efficacy and safety, and also the feasibility of assessment as well.

And I think Selena mentioned in the second session talking about, you know, the qualitative work that you can do alongside that understanding whether they can complete the measures appropriately, whether there are any challenges.

So the preliminary data that you can do, monitor your processes, but barriers to participation, you can then modify for your next phase of work. So I think there are multiple benefits there. Thanks.

>> VISHAL BHATNAGAR: Early phase measurement of patient- borne outcomes, and we have been honing in on this, and it works hand- in- hand in the work

that we do. So there is data coming and I appreciate that you have done so much work in the field to improve the methodology there and advocating on your end for early collection of PRO in early phase studies. So thank you for that.

This also relates, I think, to the next sort of key question, thinking about communication of that information if it is collected in early phase trial. Ways to do so. And this has cropped up in the Q&A box as well, about well- collected PRO data from early studies and about whether it is in the drug label or outside of the drug label. I mentioned this before, but Project Patient Voice is a publicly available website. That is an example of a way to communicate well- collected PRO data.

So before I sort of answer that question, and I want to save that for the end, perhaps Amylou, Project Patient Voice, what are your thoughts on something like Project Patient Voice that's outside of the ()?

>> AMYLOU DUECK: Yeah. So I'm a big fan of Project Patient Voice, and kind of other places to report this data, mostly because like traditional places like manuscripts, you know, are fairly limited in the real estate you have to actually put a lot of this data. And particularly for PRO- based adverse events, the data is so rich, and I think that's why there hasn't been kind of this deeper dive into that data because there hasn't been a place to actually report that data. So specific to Project Patient Voice there are many opportunities to even expand what could be reported so, you know, it could be expanded to include, say, you know, reporting of overall side effect burden, measures of physical function, or other kind of common elements that studies could incorporate.

You know, using standardized tables and graphics can aid folks in developing a level of comfort in interpreting those kind of data visualizations, and I think it is important still to include a variety of graphics so then kind of user preference can come into play. We know based on the work of Claire Schneider (sp) Michael Brandridge (sp), patients tend to like the pie charts () line charts a little bit more. The online platform also enables, you know, further flexibility to do things like interactive graphics. Again, in like a static, you know, paper manuscript you can't do these sorts of things, or drug label you can't do these sorts of things where you can actually go in and select specific patient characteristics and generate a tolerability profile that would actually match a specific patient.

This isn't without challenges, so there is patient confidentiality and obviously like small subgroup estimation that we have to worry about, but I think there is many opportunities to doing kind of these online type data presentations.

>> VISHAL BHATNAGAR: Thanks, Amylou. Devin, you know, what's your perspective of communicating this, whether it is inside the label or outside the label?

>> DEVIN PEIPERT: This could be a recycling or a continue to sing the praise, but I think we have learned so much from the research that's out there, and you know,

Amylou has been a strong proponent of the bar chart. Of course, she has mentioned nice things about the, you know, the work citing patient preferences versus clinician preferences and so I think bringing that all in and, you know, implementing it in a way that patients can receive it well, whether that's coupling with educational programs, providing additional context, providing someone who can help, you know, explain the results further, you know, might be ways to further extend.

>> VISHAL BHATNAGAR: Thanks, Devin. You know, there has been a lot of work that we did with Project Patient Voice. I would like to give Paul an opportunity to, you know, summarize a little bit about what's been said in terms of Project Patient Voice and ways in which we envision using Project Patient Voice moving forward. So Paul, go ahead.

>> PAUL KLUETZ: Yeah. This was actually generated historically speaking from a side conversation I had with one of my colleagues, geez, how can you report this information, the label is black and white, it is too small, there is all sorts of reasons why that's not great. And as an aside, why don't you put it up on the website? This was 5 or 6 years ago, and we came to the conclusion that was actually possible. And it all came down to data quality and visualizations and analytics that would make sense and also scoping it with one of the core outcomes, which happened to be symptomatic adverse events. There is definitely room for future growth. We welcome industry colleagues to bring their data in so we can put it up on the website. We already have the visualizations and the analytics ready. But you have to come to us with those PRO data collected in a way that, you know, that we can use it.

The other thing we want to do is put up physical function and overall side effect bother, but again, those are going to be analytics that we need to create. And then one last thing about Amylou's point about kind of -- actually, it might have been Mel's, about patients like me, subgrouping out specific sets of patients. I mean, we sort of dipped our toes into the water with that, Project Patient Voice, patients that had no symptoms at baseline. To me, that's super valuable and interesting analysis. Because what you are doing is heightening your attribution of what you are seeing to the drug. I didn't have diarrhea when I started this trial, I now have diarrhea after initiating the drug. And as we all know, attribution of side effects versus the disease versus something else is really a challenging thing.

So all you can do, get into small subgroups, Project Patient Voice, there is a lot of future potential for Project Patient Voice.

>> VISHAL BHATNAGAR: Thanks, Paul. And then as we have had with this last few sessions, I would like to give Yelak, the patient advocate on the panel, the sort of final word for the panel discussion. And Yelak, why don't you give your perspective on what you thought when you viewed Project Patient Voice?

>> YELAK BIRU: Thank you, Vishal. If I can digress like for 4 seconds and say I stand by the movement statements like the treatment was () well- tolerated, and

presentations, and then I also stand by the movement that describes patients subjects we should try to not necessarily disallow, but limit that work as it relates to patients as subjects.

So specifically your question, to answer it, we have come from the day of () the treatment combination, which I was treated in 1995 and 1996. With approval of 2 () therapies over the last years. And I understand in the 1960s, when () was being used to treat Hodgkins lymphoma, anybody who talked about cure was kind of labeled a renegade in the field of cancer treatment.

So today in my opinion we talk about patient- reported outcome, quality of life, tolerability because we have options that are important in extending lives, and we could choose either for approval purposes or for what the next best treatment should be, so it is because the scientific community has advanced treatment and treatment options for various cancers and various ailments forward. And in my opinion, Project Patient Voice is really important. Patients today got their patient- reported outcomes and tolerability data ad hoc, meaning when a patient is considering a new treatment they ask their friends, people that they know, they go on social media and ask other patients what other side effects of a particular treatment they are about to have is.

So I see this in various forms, right, including in- person support groups and social media. So the standardization of this through patient -- Project Patient Voice is really important, and in my opinion helps move the field even further forward.

>> VISHAL BHATNAGAR: Well, thanks so much, Yelak. And thanks again for the entire panel for your perspective on communicating patient () open- label trials. Thank you to Devin and Angelo for your presentations.

We have about 25 minutes, and so we will go through and do the question and answer portion for this panel, and we have quite a few questions that have piled up from earlier sessions. So with the kindness of this panel to answer some of those questions that have built up, I will start first with Paul, you know. One of the questions that was in the question and answer box from the last session was provided that there are domains that are not core concept to oncology, such as emotional function and financial function, questions only add burden and how would you suggest eliminating that burden for patients to only collect data that's going to be analyzed?

>> PAUL KLUETZ: Yeah. This reminds me of like the first workshop or maybe the second workshop where we really looked at all the tools and one of the really nice things about QRQC 302 has individual -- so it is actually set up in a modular way.

And while I don't think it is feasible to strip it down completely just for FDA's core outcomes because let's remember, single trial that's going to need to give their data to all global regulatory and pair groups. So I understand that there are others that want those domains.

What I do know is that created a QRQC 17, removing the symptom scales. So that's, you know, a savings of multiple questions. So that they can deploy alongside of that in item library, or even disease- related symptoms.

So I think there are things you can do, but we have to remember that FDA is just one stakeholder who is going to be using this data.

>> VISHAL BHATNAGAR: Yeah. I think that's a really good point, Paul, and I totally agree that although there are questions that were designated () guidance, many other bodies besides the Food and Drug Administration care about some of those questions, and are useful for them, such as HPA body. So I think that's an important distinction to make there.

So Mel, do you have any comments or followup? I see that you may have a comment that you would like to make here.

>> MELANIE CALVERT: Yeah. I think, again, it goes back to the point that Selena made earlier in the session about importance of patient involvement in the co- design of trials, and actually we can't realistically assume actually that people won't experience burden. I think we have to ask, explain while we are collecting the data what the rationale is, how it is going to be used, and whether they would like to see it there. Because there is evidence in the literature when people use shorter forms, when they discussed it with patients they felt that there was stuff missing that mattered to them. So I think we should do research, we need to work collaboratively with patients to capture what they need to know as well. Thanks.

>> VISHAL BHATNAGAR: Great. Thanks, Mel.

So the next question we have, I'm going to field for Amylou, and this is a little bit open- ended, but feel free to respond how you like.

So this question is what do you recommend for approaches for analyzing PRO- CTCAE, and I think you are probably one of the world's experts on analysis of PRO- CTCAE. So it is a good question for you. So go ahead, Amylou.

>> AMYLOU DUECK: Okay. So, you know, we typically start with, like I said, kind of a CTCAE type of approach. So we tend to focus on a safety type population. So these are patients who initiated treatment from a PRO perspective. We do typically look at patients who have completed the baseline assessment and then also completed at least one post baseline assessment. We usually start with a high level summary, so we do focus on a summary measure that kind of mimics the maximum score post baseline. But we do take into account the patient baseline score, the preexisting symptoms, and to apply that which we call the baseline adjusted approach, we basically only incorporate scores post baseline which are worse than the patient's baseline score.

We do apply what we call our mapping algorithm, which consolidates data kind of

across attributes, but we do look at the individual attributes, so there are cases we have found where you might get a different signal when you look at different attributes within an individual symptom. So we think it is important to dive down into the individual attributes.

And then once we do kind of all of that then we do look at the longitudinal plots. We also might generate additional summary metrics, you know, if there is a particular, again, depending on the profile that we are observing we may describe a particular summary measure like time to, you know, time to a peak or something like that.

So that kind of sums up like a general approach. Again, depending on the particular study we might take a slightly different approach depending on what we see.

>> VISHAL BHATNAGAR: Great. Thanks, Amylou. I see there is other questions about PRO- CTCAE, so don't go far. But before we do, I will ask you, Devin, about PRO- CTCAE. So the question from the audience is the PRO- CTCAE mainly focuses on treatment- induced toxicity. Would it also be a useful tool to assess treatment improved symptoms?

>> DEVIN PEIPERT: Yeah. Paul brought up attribution earlier. You need to identify for the PRO- CTCAE and answer more generally about all PROs that we are using for this purpose. There is overlap between disease symptom and side effects of treatment, right, and getting at that attribution where they occur is very difficult. I won't sit here and say you are going to come to a solution that's going to give you X percent of the fatigue is due to this and X is due to that.

One thing is you look to your trial design, look to the specifics of your context and try to learn and make some hypotheses about, you know, whether or not a particular symptom is really a side effector is it from the disease. And I think you can also use overall side effect burden to try to understand that within the context of your own trials. So take something like GP5, if you have got fatigue and that's really correlated with the GP5, that's probably a treatment side effect. But if it is not correlated with GP5, maybe that's still due to the disease. I hate to give a researcher answer to say it depends, but I follow up with another research answer which is to say try to learn and try to do research and bring some more evidence to the table for your particular situation to build an evidence- based argument.

>> VISHAL BHATNAGAR: Great. Thanks, Devin.
So the next question is for you, Yelak. If patients know what their adverse event or PRO -- that their adverse event or PRO information will be viewed, presumably by clinic staff or trial staff, how will it affect how and if they report? And perhaps you can also provide your opinion on whether this differs by () in the era of multinational trials. So go ahead, Yelak.

>> YELAK BIRU: Yeah. I think the short answer is that patients are generous with their data in sharing, especially if they know it will be used for good and help

someone who comes after them. I think a more detailed answer is I think one of the fear patients have is that their patient- reported outcome report will be ignored or disregarded and not that they will be overshared. And I think the associated fear with that is that patients have is not about the PRO reporting, but the fact that they will be labeled, they have disease X or cancer Y, with treatment side effect Z, and marginalized or options and doors starting to close on them both in the professional and personal areas.

So I think when it relates to clinical trial or practice, patients are more generous than we think we are.

>> VISHAL BHATNAGAR: Great. Thank you very much, Yelak. So Amylou, I told you I would return back to you about PRO- CTCAE. There is a question in the Q&A box that says, how do we visualize with so many time points that the PRO- CTCAE is collected? So I guess it is a question about data visualization and high frequency of assessments.

>> AMYLOU DUECK: Yep. And I can tell a true story that in the early days of doing PRO- CTCAE in trials I think we asked too many questions and assessed too many time points because we didn't know too much about the tool. So right now I'm analyzing a study where we did weekly assessments and we had some patients that stayed on treatment for, you know, out to 3 years. So I have, you know, a study where I have a tremendous number of time points and I'm definitely dealing with this sort of issue right now. So I will say in my preliminary analysis, I print out bar charts with multiple rows and I do look at every single time point because I think the patient provides me data, I'm going to look at it.

My final graphics do not actually print out that far, you know. I'm working on kind of consolidated graphics. I will use -- if the pattern continues on I will use like some little hashmarks to kind of show that they continue on. I'm also doing some visualizations with lines, so kind of like think of them as like smoothing like, you know, kind of a smoothing lines are to try to show patterns. But to kind of consolidate so you don't have to have full bars, but showing some numbers at certain time points to kind of give a flavor of what the sample sizes are further out. So, again, some ideas around how to show many time points, but not cluttering out the graphics, but still trying to be true to the data.

>> VISHAL BHATNAGAR: Paul, do you have a followup to that? The frequency of assessment?

>> PAUL KLUETZ: Yeah. I just -- I think it is -- the reason we wrote the guidance is where we have high frequency to 6 months and then every 3 out to a year is to say after a year we have a lot of information about safety and tolerability and even efficacy for the acute- phase. And I think you have to be really thoughtful about how you look at long- term patient reported outcomes because I think things like time () where you are going to serially assess for perhaps years is something we have to think

carefully about whether we are getting value out of that versus -- for the research perspective.

>> VISHAL BHATNAGAR: Definite live. So now, Mel, question for you. If there is no existing PRO that fits the disease, is it suggested to add one or several items to the existing PRO or add the disease specific to each item in the existing PRO as a modification? So I will start with you, Mel, and maybe Devin, you can chime in there as well.

>> MELANIE CALVERT: Yeah. So, I mean, again, we have experienced this in the normal oncology () didn't exist. And really it was about going back to the development team, so I contacted () good chat with her about the approach that the developers wanted us to use in terms of developing new items. So working collaboratively. So I think my stance in this would be very much that if there are things really important elements that the patients or the clinical team have identified as being important may come up, and you may require, that you have already got some really excellent banks that are available from existing measures and I think it is going back to the developers and working with them to discuss how you might modify items. Of course, they wouldn't be validated, and need further validation work going forward of the thanks.

>> VISHAL BHATNAGAR: Thanks. How about you, Devin?

>> DEVIN PEIPERT: I agree with Mel wholeheartedly. We are living in an era now where fit for purpose is really what we are aiming for, regulatory concerns are on the table. Especially if you have qualitative data that was collected prior to the trial that says there is some issues that patients really care about that are not in the instruments you have available to you, which is a frequent occurrence, it is much better to try to get at that in some way than to ignore it.

So there are many strategies that you can use to build that in, and assess the best you can that instrument or whatever you have added, psychometric properties, but I think the exit interview is a nice thing you can do to help look at newly added items, perhaps, and see how they are received and see what patients think is meaningful about them to help inform the analyses of that trial data, using the new items.

>> VISHAL BHATNAGAR: Okay. The next question, I think, just came in, is with () be appropriate for () so I will ask Angelo and then maybe Amylou to opine on that type of visualization since I think these are two people who really are very good at visualization.

>> ANGELO de CLARO: () diagrams are a way to graphically depict like what's happening for the population over a period of time. I think this is an excellent way, an alternative to additional -- when we present our information and prescribing information, this is snapshots, they are providing adverse event rates at that particular data (). And it is probably not -- not useful to patients who are like, well, what am I

expected to experience during this trial? At what time point will I get these adverse events, etc. So I think having a way to visualize, the Sankey and other ways to visualize this is more informative. But these also can get pretty complicated rather quickly, depending on categories. So I think we really need to engage with the patient advocate or like the patient communities for like what visualizations and presentations would be most useful for them.

>> VISHAL BHATNAGAR: Totally agree, Angelo. Amylou, your perspective on Sankey?

>> AMYLOU DUECK: So I'm a big fan of the Sankey plots, but I'm also a visualization nerd. I think they are cool because they show dynamics of score changes. So different groups change over time and how scores change over time which you can't really see from, say, the bar charts. You can't see how patients move from different categories over time.

But, you know, I think what is challenging with the Sankey plot, if you haven't seen many of them or you can't really follow the little lines, you know, it can be very challenging to interpret them. So it is definitely kind of a give and take with the Sankey diagrams. And there is a big learning curve. And some of them may not be that useful, so I think there is probably a context for them, but it may not be, you know, for perhaps like a more lay population. But I do think that they are pretty nifty plots.

>> VISHAL BHATNAGAR: Totally agree. I have a couple more questions. I think there is two sorts of groups of questions. So the first group questions that's in the Q&A box is actually thinking about fatigue, and I think some folks are really thinking about fatigue as a disease- related symptom as well as a tolerability issue.

So I will start first with Devin and then maybe Amylou, again, so thinking about any methods to disentangle disease- related fatigue versus fatigue related to treatment, and if that's even possible. Go ahead, Devin, first.

>> DEVIN PEIPERT: Yeah. Yeah. I will harken back to a response I gave a few questions ago, and that's to say I don't know that there is any one absolute way to get a for sure answer, and especially quantify how much, say, fatigue is due to disease versus treatment.

But I think going back to the idea of correlations with items representing, you know, known side effects, it could be an important way. I mean, using the trials data to learn about, you know, that specific context or a specific observation is probably -- that's what I do first and try to get, you know, in this particular situation what are we seeing there, and of course the clinical area, what's known, all of that stuff to make your best guess.

>> VISHAL BHATNAGAR: Amylou?

>> AMYLOU DUECK: Yep. I would agree. I wouldn't say that there is any one particular, you know, approach to completely disentangle it, but you do have to be careful about, you know, which approach you use. So specifically like our baseline adjustment approach is specifically looking only for worsening. So it is not going to be able to look for any kind of improvement, so it would give kind of a false readout if you were specifically looking for any type of improvement. Because those are all going to get zeroed out. So if you are specifically wanting to see, you know, perhaps some patients who are worsening, some patients who are improving, you probably want to be looking for, you know, explicit change scores or, you know, potentially looking, you know, looking for some other type of analysis that allows for movement in both directions.

But I do think kind of pinning analysis on change can potentially be used to disentangle, you know, whether it's, you know, disease related or kind of more of a toxicity, although not necessarily always true, but that's how I tend to think about it.

>> VISHAL BHATNAGAR: Thank you both. So I fully recognize that I have been handing out the questions and I haven't answered any. So I will take the first crack at the next one.

The question is can high- quality descriptive tolerability data be included in a label if the drug is approved only in a single arm Phase 2 trial. I think the answer to that question really depends on how that Phase 2 trial is used in the drug application and submission, and also there is very -- you know, there is very specific rules about what can and can't be included in a drug labeling, and my colleague, Paul, is the guru.

If there is well- collected, clinically meaningful, descriptive information about tolerability from an early phase trial that there are ways to disseminate that. One could be the label, and Angelo mentioned, you know, there has been inclusion of tolerability information in Section 6 of the label, which reflects the safety. There is an opportunity to also use Project Patient Voice, again, understanding that this needs to be a trial that would use as probably the totality of information to evaluate () risk. And then there is also well- documented examples in the literature of how early phase studies can support or complement later phase studies and there is that experience that's been documented in our FDA reviews as well as in the medical literature that will point you to the FDA approval summary of (), which includes early phase tolerability information that may not have been included in the label, but was included in FDA review summary.

So it is possible, and there are ways to do so. I just mentioned 3.
Paul, do you ever any additional thoughts on inclusion of tolerability from early phase studies and inclusion of that in the product labeling?

>> PAUL KLUETZ: Well, in that we provided efficacy information on a single arm () I don't see any reason why we wouldn't be able to incorporate well- collected, rigorous tolerability information. I think makes the key point that the data has to be well- collected, assessed at the right frequency, and make sense we will review it and

add it. So the big picture about communicating () data, as we get standard analytics, have a different type of review summary in the published literature where we will focus on the patient () data, but that's, you know, another aspiration that we can look at, and we will.

>> VISHAL BHATNAGAR: Thanks, Paul. And just to clarify something I said in terms of inclusion of tolerability in product labeling, you know, it hasn't been done necessarily for things like the () PRO- CTCAE symptoms, and we have Project Patient Voice for that reason. It can provide us that high fidelity assessment for specific treatment- related symptoms. But we haven't included that type of data in product labeling because of the limitations in terms of color and phase, etc.

With that being said, if one were to think about inclusion of patient- reported symptoms related to treatment, to inform tolerability, then use of Section 6 as opposed to more traditional clinical trial results in Section 14. This hasn't been done necessarily and there has to be a lot of caveats that go along with that, but for that reason we really think Project Patient Voice really seems like a federal -- a constellation of symptoms need to be included from a trial.

So just wanted to clarify that one point.

The next question, I think, is to perhaps Mel and then Yelak. The question is in terms of PRO, how to evaluate the impact due to culture, income, and social situations, especially running global trials? Any advice considered to design it? So Mel, go ahead.

>> MELANIE CALVERT: I think it is about thinking about whether you have got any hypothesis at the start that () alongside different groups and if you have, I think, collecting some really robust data. There is elements of baseline so you know you have really well- characterized your population so you can do some exploratory analysis to see whether those different treatment effects in those different subgroups of the population might be useful.

And I guess that also, you know, applies to things like health, economic evaluations. I know that's not the scope today, but thinking whether you do that for different countries and so on and capture this data, which also might again lead to that. So those are a few considerations from my perspective, and you might want to add to those.

>> YELAK BIRU: Yeah. So my experience in () setting is there is quite a large difference in terms of drug access and real- world based clinical practice settings drug access due to economic situation of a particular country or region and affordability of drugs. So clinical trials, () is high, but not necessarily availability of the drugs. So I think it is important that PROs are looked at with that light and patients need to be encouraged to report the true and real PRO impact because they may want to do whatever it takes to stay on the clinical trial, and they may not want to be "kicked out" of

a clinical trial. So I would be interested to see others on the biological differences globally, if there are differences in PRO reporting with U.S. and () U.S. data.

>> VISHAL BHATNAGAR: Thank you, Mel and Yelak. So I think we have time for just one more question. I will point this one to Paul. I think the question is what is the agency view on realtime reporting, but I think they might be referring to realtime monitoring of PRO- CTCAE data during the conduct of a clinical trial?

>> PAUL KLUETZ: We get into this a little bit in our paper that I mentioned from 2018 that it is not the expectation that there is clinical monitoring during the trial of PRO- CTCAE or any () if it is not in the protocol, that's what they are intending to do. That's not to say it couldn't be done and that, you know, the investigator decided they wanted to say, monitor diarrhea through patient report, and they created a plan to do that. I think it adds a complication to the clinical trial conduct, but I think it could be done. We haven't seen it.

And another, I think, situation that's the other side, this idea of realtime kind of assessment, where you have an app on your phone and you are assessing it whenever you want, whenever you feel like reporting you report. I think that has challenges, could add some heterogeneity to how the data comes out, and we haven't seen that. The last thing I will mention on realtime clinical monitoring is someone had put a question, would you use PRO for dose modifications, because that's kind of the natural next step, you are monitoring. If there is a high rate of diarrhea, you dose modify, you know, I think to me when I look at making clinical decisions, we almost never go straight from a lab value, for instance, to a clinical action. So if I see hyperkalemia in the clinic I go and talk to the patient, I think about repeating the lab, I really have to think about things before I initiate a supportive care. The same for this, you would want to have the patient come in based on their signal and talk more about it and see if it warrants a dose modification.

>> VISHAL BHATNAGAR: Thank you, Paul. So with that I think we are at the end of our time for this session. And you know, this was a really critical session to help orient us for the next few years as far as a few key items, which I will go through in just a minute when I talk about the conclusion of the workshop wrap- up. But I would like to thank the Session 3 panelists for their participation of the workshop.

CONCLUDE AND ADJOURN

>> VISHAL BHATNAGAR: And the conclusion slide. Thank you so much. So I thought this was an excellent workshop in the sense that we went through many of the sort of realities that are a little bit challenging, but many of the things that we needed to discuss in the field. And as Paul mentioned, this was a bit outside of the orthodox of our workshop subtypes, and so appreciate the flexibility of the panelists and the willingness of the audience to participate in this workshop.

I might have a few takeaways, and I will ask for Paul's key takeaways. But my

initial takeaway with this overall workshop, I think that trial design methodology and good, sound trial fundamentals are key to reducing bias of all types, whether that's open- label bias or other types of bias. I think that we have heard loud and clear that patients provide their experience truthfully and they provide their experience with the medication that they are being provided, and they do so for altruistic reasons because they want others to learn from the experience of the clinical trial that they are enrolled in.

I think that my biggest key takeaway is that this will redouble our efforts of patient outcomes. We began much of this work with the friends of cancer research white paper, but as time has gone on we are honing in on early phase studies. We have heard a lot about that this year, and based on the conversation of the workshop, I think it should be clear to the audience that early phase PROs are here to stay and will be examined, analyzed, and potentially communicated.

And lastly, I think that one thing that I will take away from this workshop is that inclusion of the patient experience actually reduces bias in a trial. So I think that with those 4 points in mind, I would like to ask my friend and colleague, Paul Kluetz, to provide his workshop wrap- up.

>> PAUL KLUETZ: And the biggest takeaway for me is it is possible for your audio to be muted, not just your voice to be muted. So I apologize for the technical challenge in the beginning of the workshop. I want to mention broadly this is about open- label trials and the fact that PRO data can be evaluated. And I also want to mention, we don't want to suggest they are great no matter what. We would want to have a blinded trial if the key outcome was a patient () outcome where feasible. We just know oftentimes this is not feasible and we want to make sure people realize you don't need to preclude that.

I heard loud and clear, it is obvious that they are common in oncology. I heard that people are really agreeing with the idea that tolerability is a key objective of patient- reported outcomes. You know, we certainly heard from Patty, one of our advocates, across the board she would like to see descriptive information on tolerable, as we do in safety data. And I think that certainly, as Vishal said, we can use this data in early drug development, where we have very little information about safety very early on, and actually help us to design both the dose of the drug going through the registration trial, as well as the descriptive information that we receive.

I also heard from Angelo that efficacy is not off the table for patient- reported outcomes in open- label trials. Now, this is a very unique context. I want to say that oncology, as Yelak said, we want to know the drug works first and foremost. And the uniqueness with oncology, a lot of our objective measures, we are very comfortable that the drug has antitumor efficacy, that the drug is working. Once we know that then we can use patient- reported outcomes data that's complementary to that efficacy, and we can do that even in open- label trials.

Finally I will end, hopefully I showed you we really have the tools and guidance to do this well. I know we are going to see this data in high quality over the next few years, and it is our responsibility in the field to generate those standard analytics for physical function, overall side effect impact, so that we can get it up or on Project Patient Voice and in some cases the FDA label.

So with that I want to thank Vishal for putting pretty much this workshop all together. He did a phenomenal job. I want to point out, Erica, our colleague, was unable to make the workshop, but she was also very, very involved from the beginning and did an incredible job and we will be seeing her in next year's workshop.

>> VISHAL BHATNAGAR: Thanks, Paul. And next slide, please. We would like to thank all the staff that was involved, including Erica, but Caitlin, Rich, and Joan were very instrumental in the execution and planning of this workshop. And we would like to thank them. And many other FDA staff, OCE leadership. Thank you to Paul for starting this workshop 7 years ago, and we are going to continue this workshop. Again, really like to thank all of those that were involved in the planning and execution of this workshop, and as well our panelists and speakers. So thank you for the joining us this year, and we will see you next year. Bye- bye.

[Workshop concluded at 2:51 p.m. eeastern standard time]
[Human captioning provided by HRI]