

The title of this talk is Experimental Verification and Clinical Application of Monte Carlo. So what are some of the questions, we'd respect your experimental verification, what experimental verification is needed to show that Monte Carlo's working correctly and with respect to clinical application, why Monte Carlo, what are some of the physics issues and which anatomical sites is Monte Carlo most beneficial. I'll talk over a bit about statistical uncertainties and how we can interpret them for treatment planning and finally, I'll give some sort of summary on clinical benefit of Monte Carlo. I guarantee not to answer any of these questions, just to provide some perspective on some of the issues. So let me begin by making this maybe bold statement that a Monte Carlo treatment planning system is just another treatment planning system, it's not a time

machine, it's just another treatment planning system and as such it should be subjected to the same testing and verification requirements as any other planning system such as provided in documents like AAP and TG-53 for example. However, that said; let's remember also that with Monte Carlo, we are able to transport particles both in the treatment head and within the patient and inherently implying a higher degree of accuracy in radiation characterization in the treatment head and the patient. What this means then, is that perhaps test should be designed to verify the stated level of accuracy, so if we're saying that it's going to be more accurate in these more complex, more sophisticated cases, then we should actually verify that it is accurate in these situations. So one strategy then, is to design measurements to test the algorithm performance with

emphasis on simulation of the LINAC treatment head components, and typically this is done in homogeneous phantoms where we test issues such as the beam energy and then modeling of components such as flattening filter, beam modifiers components, such as the MLC, etc. And the second part is radiation transport accuracy in the patient, so how well is the algorithm actually working and this is typically done in heterogeneous phantoms. With Monte Carlo we want to use small field sizes and we want to use non-equilibrium conditions. Some might argue that the main benefit or utility of Monte Carlo is in these types of situations so we want to be able to verify that. I should also point out that there's been a fair amount of work done in the area of verification, and in the few next minutes I'm gonna tell you what's been done and also at the same time talk about

some of the issues. With respect to verification in homogeneous phantoms, I'm showing some depth dose curves here, from the Peregrine group and what I want to point out here is the range of field sizes for 6 and 18 MV photons. They went from a 2 by 2 to a 38 by 38 and the advantage of this is that it really tests the various components, at the small field sizes, you're checking primary radiation. As you get bigger in field size, scatter becomes more important and we really need to include this range of field sizes for testing. With large field profiles, 40 by 40, you see the entire beam profile, and this is sensitive to how well you model structures such as a flattening filter. In this D-max region you really the effects of the flattening filter in this particular case. What about surface dosages, which are important for head and neck planning for example. Monte

Carlo gives us a platform to really get this right and if you're doing source modeling, then you better be sure you're getting it right. This is an example where we use the source model and it shows the calculation with no electrons here in this dash-line and with electrons and you can see fairly big differences if you don't account for electrons correctly. What about verification of beam modifying devices like the multi-leaf collimator, there are several different ways to do this, the Peregrine group has proposed doing this with explicit MLC transport so we're transporting all particles through the multi-leaf collimator. This may be not as efficient, but it is shown to be quite accurate and this example, you can see is a very complex shape going perpendicular to the beam

and they're getting all the transmission effects correct. It shows pretty good accuracy. You can also use approximations and Jeff Siebers does a first Compton scatter model approximation, so he's doing some transport, pseudo-explicit transport we can call it, and it's more efficient and in this case it's also shown to be more accurate. Actually Jeff has shown that his model works well for complex MLC field shapes as well. At Michigan we use sort of an implicit transport model called a source model, an intensity map approximation in which we break up fluence into grid elements and then we use some empirical methods to correct for, to account for field shapes, arbitrary field shapes. What about inhomogeneities and this is a study from Arnfield from the Virginia group, showing a slab type geometry and we see water here along equivalent slab and water 4 by

4, 18 MV and you can see some significant differences here even comparing Monte Carlo with collapsed cone convolution. In this region, the difference is on the order of 15 to 20 percent so showing that the issue of lateral electron transport is very important. High density inhomogeneities from Charlie Ma's group, another slab phantom experiment, water, bone, water, and he's compared Fast Fourier transform cone convolution with superposition convolution and also Monte Carlo; but again, you see that there are some differences on the order of five to ten percent. Other types of geometries, perhaps as we move Monte Carlo closer to the clinic, we need more clinically realistic verification geometries and as some would argue that with Monte Carlo we should be testing the code to prove that it works on the non-equivalent conditions. Here are two example

geometries shown here, thorax phantom on the left and the lung tumor-type phantom on the right. Using small field sizes really poses a strict situation and it's a difficult situation for most algorithms to get right and I would argue it's a good test to make sure that the transport is working correctly. Finally, we can't underestimate this point, accurate measurements are very difficult and I can show you two examples of that; These are some measurements in the buildup region showing differences between a parallel plate ion chamber, two different cylindrical ion chambers and a stereotactic diode and you can see that even amongst the various measurements there is a fair amount of variability. So how are we going to accurately verify algorithms if we don't have a consistent sets of measurements? Here's another example of issues with measurements from a paper by

Laub and Wong for a 1 by 1 field size and you see that the output factor drops dramatically as the active diameter of the chamber increases and this is a significant effect. In going from a diode to a six millimeter detector size for example and you see on

the right here the effects of volume averaging on the profile that can be significant also. So measurements of small field sizes in homogeneous media are even more complicated. We need to keep this in mind as we are developing new verification geometry. So some other remarks.. With Monte Carlo, there are several new parameters that can influence stimulation results within the treatment head and the patient and I would argue that vendors and users should be cognizant of these issues and I'll show you two examples. This is a paper by Darius Shiekh-Bagheri and Dave Rogers that talks about sensitivity

treatment head simulation to the beam energy in this case and you can see that these off-axis ratios in air as a function of energy will change dramatically. In this example in going from 6.6 MEV to 5.5 MEV is roughly a 15 to 20 percent change. So what happens if you have the wrong vendor blueprints, and this figure really speaks for itself, it's a 10 by 10, 15 MV profile and this line is a profile with copper design that was specified in by the vendors; it turned out that the real material was tungsten. So, maybe this is more of an issue for the vendors, when they're giving you these systems, but, what do you do if you end up in a situation like this; do you go to the vendor and say we need more details on the treatment head geometry, at least we need some sort understanding or some consensus on how to do deal with some of these issues and hopefully the task group will

be able to provide some recommendations on these issues. Enough on validation, we'll move on to clinical application of Monte Carlo and I will show some examples of treatment plans in different anatomical sites, talk about statistical uncertainties and treatment planning and discuss a little bit about the clinical utility of Monte Carlo. What is the main dosimetric issue, and you can see that this is a 6MV oblique fields in the lung. The main issue is underdosage of the PTV due to lateral transport of electrons and you see that the solid line Monte Carlo is underestimating the PTV coverage here, the 100 percent line relative to the, EPL, the equivalent path-linked algorithm. You also see that it's nicely in the sagittal cut and then on the right, the influence on the DVH's. We also have this issue because of lateral transport of penumbral broadening and on the left you

see an equivalent path length calculation on the right Monte Carlo and you can see the issue with the, in the penumbral region here, it's much wider for the Monte Carlo calculation. And for that particular case here are the DVH's for the CTV on the left and for the left lung on the right and for both these cases, both the CTV and the left lung we find some significant differences. This becomes important particularly if we want to correlate how these different algorithms will correlate with clinical outcome. What about the head and neck. The variety of tissue densities in the head and neck may mean that you might see some differences between these algorithms. This is a 19 field plan, that's what it looks like and here are the DVH's and we've included convolutions, super position here in white, Monte Carlo in green and equivalent path length in red. We see

clear differences between Monte Carlo and EPL. Convolution is closer to Monte Carlo but there are still some differences. What about prostate, this is a four field plan and prostate is one area that we think it's fairly homogeneous except for maybe the rectal region and the femoral heads, and you can see relatively good qualitative agreement here between the equal in path length and Monte Carlo. If you look at the DVH's, you see at

the PTV DVH's are falling on top of each other, almost coincident but the rectal DVH's are different and this is due to air in the rectum which is also an issue for accurate dose to the rectum. The liver is another that's fairly homogeneous and again this sort of shows this result. The DVH's for the PTV in the normal liver and you can see that the differences are much smaller in this case because it's a fairly homogeneous type

geometry. Radiosurgery's another complicated issue because of the small field sizes used in radiosurgery. And this is a nice paper by Solberg that shows differences between conventional methods for the 1 cm collimation and Monte Carlo on the right and you can see some differences here in the air spaces for this cavernous sinus tumor. And Solberg also points out that lateral electronic disequilibrium will always exist whenever the field size is smaller than the range of the secondary electrons. What that means for radiosurgery is this, that even if you are in a homogeneous situation, in a roughly homogeneous area like here for example, you're still going to have significant lateral electron issues which will be harder to model with conventional algorithms and when you add inhomogeneities it becomes even more complicated. So actually I want go back for a

second, I want to use this as sort of a segue. You see these jittery isodose lines and these jittery isodose lines are due to the stochastic nature of Monte Carlo. Here's an example of that, let's say we had a voxel where we deposited a mean dose of one gray and in this case let's say 5 percent uncertainty in that, so we're going to have a .05 gray uncertainty associated with this dose. The uncertainty's proportional to one over the square root of N, where N is the total number of particles and I should also point out that there are two sources of uncertainty. Uncertainty comes from both the treatment head simulation of the phase based simulation, called the latent variance and which is sort of a systematic uncertainty that's propagated through the entire calculation. Then we also have uncertainties that are more random in nature from the patient simulation. I'm going to

show an example now for a three field lung case, it's a clinical photon beam example, and I'm going to show you what happens as we increase the number of particles. And on the left here you see the range of uncertainties five to seven percent of blue line is the PTV, so within the PTV we're on average five and a half percent but if we look in the contralateral lung, we're at ten to 20 percent at 10 million particles. As we increase 150 million particles, now on the left inside the target we're roughly one and a half percent. In the contralateral lung we're at two to six percent and if we go to 1.5 billion particles now in the entire region we're between zero and one percent and about half percent of the target. Okay? So what is this how does this look in terms of what the isodose lines? This is the 95 percent isodose line, at 10 million, 50 million, 150 million and 1.5 billion.

So this sort of gives you a sense of how statistics can sort of smooth things out. So there are few remarks from this example, and one is that the point doses are subject to large statistical fluctuation, isodose lines also have fluctuation but have less fluctuation and so we have to think about another way to prescribe doses. Prescribing doses to a line may obviate the statistical issues with point doses. But there's a real difference in the paradigm when we prescribe doses with Monte Carlo. The second issue is where do we report statistical uncertainties in Monte Carlo treatment planning and this is something

that the task group is now fighting about, or I should say debating about, that was a Freudian slip. So should we do this at D-max? Should we do this at doses larger than 50 percent? Or should we look at uncertainties where we want to see them, structures of

interest? And this is not an easy question to answer. You could argue for each one of these and this is something that we will hopefully have some recommendations on in the task group. So moving on to the clinical benefit of Monte Carlo and how we can study this, I would argue that retrospective studies investigating correlations of dose effect for both tumor and normal tissue, may help to give an early indication of the clinical benefit of Monte Carlo. Retrospective studies may also help physicians determine how to use Monte Carlo doses more effectively in planning and finally data from retrospective analyses will help guide us in using the data in prospective studies. I'm going to show you an example from some of the work we're doing at Michigan; this was based on the University of Michigan lung dose escalation protocol, a quick review of this. There were

120 patients, treated, with non-small cell lung cancer, treated with conformal radiation; the dose escalation schema was based on V-effective. We used five V-effective bins and the doses ranged from 63 to 100, roughly 103 gray. The treatment planning was initially performed with inhomogeneities correction using equivalent path-linked algorithm and we're now re-calculating these using Monte Carlo so I want to point out that this is a preliminary study so you have to take some of the results with a grain of salt. This is target dose and regions of recurrence and on the left you see the 95 percent line with Monte Carlo, the PTV's in red and the EPL algorithm in white and you can see the underdosage here and the region of recurrence, shown here. And on the right here, the red line is the Monte Carlo 100 percent significant reduction in the 100 percent line,

relative to the EPL and then the recurrent tumor showing up here. These results are suggestive, clearly not conclusive at this stage, but suggestive that that dose maybe correlated with regions of outcome but again We need to investigate this further. What about the normal tissue doses, and I show here a picture of the percent difference in mean lung dose and this is Monte Carlo minus EPL and as a function of the mean lung dose calculated with the EPL and I've done this for both inhomogeneities on, so the actual tissue densities versus off, the in the water only case, and you can see that with inhomogeneities on, the 95 percent confidence ranges from minus six to minus ten percent, but you also notice that when you turn the inhomogeneities off, that these differences are still on the order of minus four percent. So what this is saying is that

some of these differences are due to beam model differences and not due to algorithmic differences. That's a very important point, we need to be comparing apples to apples when we're trying to correlate outcome, dose with outcome and if we don't do that correctly, then it's going to be a biased comparison. In conclusion, although we want Monte Carlo in the clinic ASAP, I would argue that Monte Carlo is not a plug and play type of algorithm. I consider myself a Monte Carloist but I'm also working in the clinic and I understand some of the issues. So, we must do this sensibly and understand the issues. Finally, routine clinical implementation of Monte Carlo will not occur without

strong clinician support and studies of clinical utility of Monte Carlo are encouraged. I want to acknowledge the following individuals including Dr. Spring Kong who is going to be talking here today. Thank you.