

Back in the day, PET scanners were basically research devices. And they imaged a single slice at reasonably high sensitivity and they were typically run by a research team and almost all of the applications back then, were in neuroscience. This is a picture of a PET III that was located at Brookhaven National at about 1980. But, in the last five or ten years, there's been a lot of growth in the field of PET. If you even look out on the floor here, there's a PET scanner now, which is new for the AAPM. And a lot of that has to do with three different factors: one is the, the, the increased use of PET for oncology and also the ability to have regional distribution centers for an F-18 FDG allowing one to do PET imaging without having your own cyclotron. The most important thing in the United States is that about five or six years ago, we started getting reimbursed for PET

scans, which certainly increased our business a lot. When you're not getting paid for them, it's hard to get too many people to order them. What we've seen as we've moved from PET scanning being done in the research lab, to the clinic, is we see that the scanners are easier to run, they tend to be more robust, they tend, as you'll see, to be multi-slice scanners now, that can image a 15 cm axial field of view and as I'll show you and as you already know moving towards multimodality, with the advance of the technology of PET-CT. Here is the outline of my talk today. I will talk a little bit on what you might consider design criteria, for a state of the art scanner, a little bit of basics of PET scanner design, I'll talk about PET data acquisition and reconstruction and then a little bit of a review of current PET instrumentation. Maybe the most important thing is is

image quality. How sharp are the images? How good is the resolution and how noisy are the images. In imaging, this trade off, particularly in nuclear medicine imaging is between resolution and sensitivity. And as you'll see, we're actually moving towards more sensitive scanners, rather than higher resolution scanners. There's vendors developing higher resolution scanners, but the scanners that you'll see on the floor is going to really try to emphasize having high sensitivity. Ease of use? We want this to be able to be run by technologist and not run by three physicists. Your ability to get to the patient or get the kinds of patients that you need to scan into the scanner is important, as you'll see. We want the scanners to be relatively fast, in the nuclear medicine sense of fast. We want to have high throughput and that's going to require having both good

sensitivity and good count rate capability. And we want them to be robust. In many cases, the PET scanner is going to be the only you're only gonna have one PET scanner in your institution. If it is not working, then you're sort of out of luck. It's not like you can just shift all the scans to a different scanner, like when one of the CT scanners goes down. And you want it to be not too expensive and that's certainly relative, but as you'll see, but it's a lot of hardware. And the expense in, in these machines are quite expensive. I mean, so just, you know, this is probably a review for, for most of you, but now let's say you have an FDG molecule over here and it has a fluorine-18 associated with it, it emits a positron or a beta plus and it meets up with an electron and they annihilate and part of their, and their mass is converted into energy in the forms of two photons going off back to back or almost exactly back to back, one over here and one over here. And

so, if we, if we detect it over here, and then over here, we can assume that the event occurred someplace in between these two detectors. And so, to first approximation, if someone tells you that the scanner has 6 mm detectors, you could guess that the resolution's going to be something pretty close to 6 mm or maybe, probably a little bit more. Now, you obviously don't get very much sensitivity if you only have two detectors, so what we typically do is surround the patient with many detectors and so now this detector over here, is not just in coincidence with one detector on the other side, but a whole bank of detectors, as you see here. In addition, you can improve on the sensitivity and also the axial sampling if you not only have the detectors be in coincidence with detectors within the same ring, as you see in this case, but see this

detector is actually in coincidence with detectors in the neighboring ring. We call it when it is within the same ring, a direct coincidence, and when it's in the neighboring ring, we call it cross coincidences and we'll talk more about that later. Now if you want to have more axial sampling, like you want to image, let's say to 15 cm axial field of view, then you stack many of these back to back, to back to back and your scanner starts to look like this. And so you can see that now a detector module, right here, is now basically a rectangular mosaic of many little detectors. And that is the case. Here is a detector from a GE Advanced NXI, and you can see that basically a detector block. This detector block is made of bismuth germinate, and has a bunch of many little detectors. It is a six by six array of detectors, where each detector is about 4 mm in this direction, about 8 mm in the

other direction, so in this direction, is the one going around the ring, which would define the transverse resolution, this would be into the ring. So, it's about a 4 mm detector. You can guess that the resolution of this scanner is about a little bit more than 4 mm and you would be correct. And you can see there are six of these things into a module and a number of these form a circle. So, at the end of it, you have a lot of hardware here. This scanner has twelve thousand of those little four by eight mm detectors, in a ring around the patient. Now there's one scanner that uses an alternate approach and you can see that this looks like a Gamma Camera, except for two things. One is that the crystal is very thick and it is curved. And basically what, what Phillips has done, is taken six of these Gamma Camera heads and put them in a ring and that provides a cost effective

approach to being able to do PET scanning. We'll talk about that later. This is the design of the Phillips C-PET Plus, which their moderate resolution scanner. Now, there are at least three different kinds of coincidences. We can have this one here where you have the annihilation here, this one head over here, gets detected; this one heads over here and gets detected. And so you can say that the event must have happened between these two detectors. And so I can characterize that by drawing, what I call the line of response, from this detector to that detector and since that passes right through, where the event occurred it is a pretty good estimate of what's going on. Keep in mind two things we're not estimating the location of the fluorine atom, , but where the annihilation occurred and they may not be exactly the same thing because of the positron range, that is since the

positron traveled a couple millimeters before it actually annihilated, that's going to lead

to some blurring of the data, maybe in the order of 1 to 2 mm depending on the design of your scanner. In addition, these, these photons may not go off exactly at 180 degrees. It may be like 180.5 degrees and, and depending again, on the configuration of your detectors, that's going to blur the data,.. So that in a whole body scanner, the highest resolution you could possibly get no matter how small you make the detectors, might be 2 or 3 mm. Now, as you'll see later on, when you bring the detectors much closer, you can get higher resolution and you'll see with small animal imaging, they're getting resolutions down into the 1.2 to 2 mm resolution range. And I'll show you some examples, if we have time, at the end of this talk. Now, another thing that could happen

is this. You have an annihilation, this goes over and gets detected just fine, but this one was heading over here, but then got scattered within the patient and got diverted because of Compton scatter. Now, when you draw the line of response, it doesn't pass through the event. And so, you can see that scatter, just as it does in single photon imaging, leads to sort of a blurring of the data, leads to a low frequency signal, a background on the image, and , and, reduces the contrast. Now, down at the bottom, you see at, at higher count rates, you can see that here you have an event that goes over here, gets detected just fine. This one comes over here and gets detected just fine. But, you know, they're not the same event. They're two totally random events. Two total different annihilations, but the scanner doesn't know that. If they happen at the same time (we'll talk a little bit

later about what "at the same time" means), it's going to draw it's line of response here, which isn't really characteristic of either of the events. And so, random coincidences and there are two things it does, it adds a to the background of the image that contains no information and also takes up some processing time, because you have to process these two events, even though it's not a real coincidence. And so randoms, can lead to a lowering of the, count rate capability of the device, by spending time processing events that really aren't very good. Now, you can estimate what the random rate is, and it's related to the count rate on one side of the scanner, times the count rate on the other side of the scanner, times what we meant by, at the same time. You can see $2\tau R_1 R_2$, where τ is the temporal resolution of the scanner. And let's say in a BGO scanner, that is

about 12 nanoseconds, if the two events occur within 12 nanoseconds, we say as we increase the activity within the scanner, two count-rates will go up, linearly with that, and true rate will increase linearly with that the random rate will increase as the square of the activity, in the scanner. And so, you can see that as we get higher count rates in the scanner, the random's rate increases as a square of the activity, whereas the true rate will only increase linearly with the activity. There are several different ways of getting rid of that. One is to acquire the singles rates and do this calculation. And many of the vendors are going in that way. Also, the other approach that's usually typically done is that you can use a delayed window approach. That if you acquire data in a prompt window, which may contr, contain both trues and random events and then you look sometime later after

the first event, so much later, let's say 40 nanoseconds later, so much later that the second

event that's in that delayed window cannot possibly be a true because it's, because the other photon must be like, you know, four rooms down, by that time. And then that's gonna be an estimate of how many randoms you have and if you subtract the events in those two windows, you will get an estimate of the trues. It's sorta like doing dead time correction. And that's an approach that's been used for probably the last 20 years. So, those are probably the two most common approaches. When you, when for those people using Gamma Cameras to do, you know, in coincidence to do, do fluorine-18 or positron imaging, then often in that case, they just use background subtraction. They just subtract it out, what they thought was an appropriate amount for the randoms. Now, for a long time, the, the scanners were almost always bismuth germinate or BGO. The one scanner

I talked about that has sort of the Gamma Camera heads, used sodium iodide. But, you've see, in recent years; scanners have been developed with these two new simulators, LSO, lutetium oxyorthosilicate and GSO, gadolinium oxyorthosilicate. And you can see that the big advantage, for example, OBGSO over sodium iodide, is obviously, that it has a higher Z number and a higher density, leading to a higher detection efficiency. Although it gave off less light, it had a higher detection efficiency. Now, these new simulators give you a reasonable detection efficiency. They have a higher effective Z, then sodium iodide and LSO actually, has a higher density than BGO. But their real advantage is here, that they give off their light, light much faster. And that helps you, because now, you can reduce the time that we said that it happens at the same time. So, instead of it

being 12 nanoseconds, we can maybe reduce it to 6 nanoseconds. And if you remember again, back to the formula for random coincidences that would drop the random coincidences just by a factor of two, just by doing that. So the real advantage of these, I think, is that you can run at higher count rates, and thereby, and reduce the randoms just by reducing with the timing coincidence window. That allows, for example, as we'll talk in a second, to be, these devices to be able to run in 3D more, maybe more effectively than bismuth germinate crystals, and we'll, we'll talk a little bit about that. Now, this slide is trying to talk a little bit about how we acquire data, in PET and, and one simple way of thinking of it, let's say this is supposed to be someone's head and that little triangle is supposed to be their nose. So that, so let's say they have a tumor sort of in the

anterior portion of their brain, up here. This other dot here is a reference point and that's just the center of the gantry. So, from that one tumor, let's say the first event happens, this goes over here, this goes over here. This line of response is then characterized by its angle of orientation, in this case, it's horizontal, we'll call it zero and how far it is from the center of the gantry, this distance here. And so we plot that on a plot over here, where we plot displacement from the center of the gantry on the X axis and angle of orientation on the Y axis, and A vend A gets plotted over here someplace. Now, B passes through that same point, but it's at a different angle and thereby passes slightly closer to the center of the gantry and it gets plotted over here. C is at even a steeper angle here. Let's say this is thirty degrees, let's say this is sixty degrees and it's passing even closer to the center of the gantry, and it's maybe here. Now, D, you can see passes through the same

point, but it's now, but it's now vertical and so it passes right through the center and it's plotted here. And if you plotted all the possible events, that had, went through this one point here, you would plot out a half of sine wave. And you can figure out where, just looking at the sine wave, where the event occurred because the amplitude of the sine wave tells you how far the event is from the center of the gantry. And the phase sort of tells you the, as _____ location. For example, in this case the anterior looks this way, if it happened to be lateral, it would've been a sine wave that sorta looked, maybe looked like that. And so, this has information about whether this happens. So you can see that at every point in the transverse plane, when you go to sinogram space plots into a half a sine

wave turned on its side. And so, the sine, back in SPECT, we used to acquire data that sorta looked like this, we acquired a series of projection views at each angle around the patient. And so we acquired an image at each angle, and each image that we had, had all the slice information for a particular angle. And you had a ber, bunch of these for each angle. Sinogram data is the same data, just, or, just sort've organized in a different fashion. Now, you have an image for each slice, where this go, which has image information about every angle at that slice. So in this case, you have an image for each slice and it has information across all angles. Whereas, here you have an image for each angle, with information across all slices. So, you can actually go between these two. You can take all the data you have, the projection data you have and either orient it in

sinogram views, or projection views. Now, when you look at the sinogram, the other interesting thing to look at, is, if you look at the events that occur from one block, let's say here, at block, I've sort of... this has eight blocks and here's block seven. If you look at all the data from block seven, they all are oriented in this sort of diagonal band across the sinogram, going in this way. Now, block seven, is in coincidence with say, with blocks two, three and four in the opposite side, the events for blocks two, three and four, are in diagonals going in the opposite direction, on the other side, sort of across the other side of the sinogram. So, for example, this line of response going across here is basically, all the events from this block, go along this diagonal and all the events from this block, go on this diagonal. And where those two lines cross, here, this little dot here, is

representative of this particular line of response. So, you can see that every pixel in the sinogram, represents a line of response through the...within the scanner. So, there's a one to one match between every detector pair that could have a possible coincidence and a pixel in the sinogram. So, another thing that's useful about that, is, if you do a sinogram, let's say you, you image sort of a uniform source, which we do each morning, and you see a black streak going across your sinogram here, in this case, you would say "Well, Geez, it looks like one of the detectors associated with block seven, is not working." So you can, you can not only figure out that you're, you don't have uniform response, but you can actually figure out which block, which detector, is in fact not working and tell you service engineer that information, so they can come in and fix it.

So, just to, to, to, just to summarize here, a point and transverse slice maps to a sine

wave. The sinogram is a plot of the displacement on the X axis and angle on the Y axis. Each row in the projection is a projection through the object, at the particular corresponding angle. Each detector is mapped along a diagonal and each pixel in the sinogram corresponds to a particular line of response that is to a particular detector pair that's in coincidence. Now, I'm switching gears a little bit here and talking about the data with respect to being in the axial direction. So, this is supposed to be the patient's head here and his neck and his body. So now, this direction is the axial direction into the scanner and this is supposed to, each of these sort of pairs across here represents a ring of the detector. And so as we said before, we could have this detector being in coincidence with detectors within the same ring, this would be a, what we would

call a direct coincidence or we could have this in coincidence with detectors in the neighboring ring. And this would be a cross coincidence. And the placement of this cross coincidences would be sorta of halfway between, let say in this case, let's say that's one, two, three, four, between two and three. And so you can see, that, for that cross plane we have twice as many counts because we have ring two being in coincidence with ring three and ring three being in coincidence with ring two. And so, it would look like this with a combination of directs and crosses. Now, we can look at that, a way of looking at the sampling is to plot out what we call the Michelogram, named after Christian Michelle, who works at CTI, and you can see that on the X axis here, we have one through eighteen, which represents the rings on one side of the scanner, and let's say

here, and here you have one through eighteen, which represents the rings on the other side of the scanner. And so for each detector pair, let's say here that is in coincidence, you can see that if it's a direct, like here, two to two would be represented here, two to two, whereas the crosses, two to three, three to two, are represented here, two to three, three to two and then connected by a line that say that those are actually combined into one, into one sinogram. So, you can see that in this case, the even sinograms have twice as much data, because they connect two ways, then the odd ones and we would have to correct that, which we do. Now, you'll see some scan, and this is referred to as a span of three. If you add the number of lines in response that are used for the odd number slices, in this case one, it's a direct, to the even number of slices, which in this case is two, one

plus two is three and we call that the span of the scanner for the axial sampling. Now, you could do more axial sampling by not just having direct, odd scans be just the directs, but also sort of the second order crosses on either side. So now, you not only have two to two, but you have one to three and three to one. Now you have three times as many counts being combined into that particular slice. And then the even, you're not only have two to three, three to two, but you have one to four, four to three. You see, know we've combined more data into the, into the sinogram and you can see that now you have three and four in the middle of the scanner for a span of seven. However, if you're at the very, very top slice, there's no detector on the side to, to combine. So, you can see that the first slice is gonna have, what's a relative sensitivity of one, second slice two, three, four and

then it goes back and forth, three, four, three, four, three, four, until you get to the other

win, three, four, three four, three, two, one. So, that's the sensitivity if a scanner has a span of seven in the GE mode, this is what's referred to as high sensitivity acquisition. And the span of three would be high resolution, in the axial direction. This is a series of sinograms from a GE scanner, that's acquired with a span of seven and I didn't correct these for normalization. So, so that you can see the difference in the number of counts in here, relatively speaking the first sinogram is a relative sensitivity at one, two, three, four and if you look closely, you can see that these alternate, three, four, three four, three, four. You can see they sort of alternate between being one brighter and then the next one, not as bright, until you get to the other end, four, three, two, one. Now, when you

actually acquire the data, it corrects for that. So, it boosts up the number of counts, in this, by let's say a factor of, you know, four thirds, compared to this, to make it all look uniform, but you'll notice, that at the edges of the field of view, you're not gonna have near as many counts and so when you do more than one bed position, for example, if we scan a cancer patient from here to here, what we do is we overlap, the last two slices of each bed position in order to, to make up for the fact, that this last one doesn't have quite as many counts. So, if you overlap by a couple slices, then this one that only has a relative sensitivity, of let's say, two, gets added to a slice on the next bed position that's a relative sensitivity of one, to sorta make it look uniform, even across the transition between bed positions. Now, another thing is, in, in many scanners, you have the ability

to have septa in them here, these yellow things are supposed to be septa, interplane septa, between all the detectors. Again, this is the axial direction, as it was in the previous slides. And what that does is cut down on two things. It cuts down on interplane scatter. So, if a scatter event, if one event came over here and got detected, fine and the other one went down here, it would get absorbed in the septa, which are either LAD or tungsten. And not allowed to be detected. So, it eliminates a lot of the interplane scatter and that, in so, if you use septa, you reduce the sensi... you reduced this scatter fraction from about 40% down to about 10 or 15%. In addition, let's say you're imaging in the lungs, and the brain's up here, you may know that FDG goes to the brain. If you don't have septa in it, then activity photons from the brain could come into the scanner, and what

could, that could lead to is high level, although you're not gonna get any coincidences because it's out of the field of view, it could lead to a high number of random coincidences because it just increases the count rate on both sides of the detector, as we've said before. So, the septa reduced not only scatter, but it reduces the number of randoms, as well. However, if you take the septa, and also you can see that the septa limit, how, how you can acquire the data to pretty much, you know, maybe directs, crosses and maybe those second order crosses, but you aren't gonna be able to get a coincidence from this detector, down to that detector. If you take the septa out, like here, then you allow this detector to be in coincidence with all the detectors, on the other side, as you can see here. And so, now you've greatly increased the sensitivity, by like, a factor of four or

five, by requiring, in what we call 3D mode, and as you'll see some scanners can acquire

in only 3D mode, but some have the option of using septa and acquiring, what we call 2D mode. Now, here is a 3D sinogram, remember we only sort of went across the diagonal before, now you see we've filled in this whole thing. The slide disorientation goes this way, if you remember, so now you see that, if you're, if you're in the middle of the field of view, you get a lot more of the, of the boxes covered, then you did before. And so you see, you get very high sensitivity, in the middle. However, if you go to the very edge, like the first plane, it's just really direct, right? One to one. So, you get, don't get really more counts at the edges of the axial field of view, but you get a whole bunch more in the center. So, if you are gonna do a multiplanes, a multibed position study, you're gonna probably overlap by a lot, a number of slices, not just one or two, but maybe five or six,

or seven, in order to make up for the fact that the sensitivity in the first few and last few slices, is much, much less than it is in the middle. In addition to that, you may also, sort of segment the, the, the, the 3D sine... michelogram into looking at different data and this gives you'll, again, a little bit of an insight on how this data is acquired. If you look at this middle segment here, what I'll call segment one, it looks pretty much exactly like the two day data with a span of seven, you see the one, two, three, four, three, four, three, that looks like what I showed you, sort of earlier. So that sample is the whole brain in pretty much uniformly, except for the few slices, at the very end. If you look at the next segment, which I'm calling segment two, which is this part of the michelogram and this part, what that provides you is sort of angular views through the brain, cutting off the top

and the bottom. Okay? So that's the advantages. Now this detector can be in coincidence way down here, but you're cut, you're not getting any more information on the top and the bottom, then you did before. Here and here. And then you can go to the very ends here, then, what I called segment two, and you'll get very angled views to the brain, that are very truncated, both at the top and the bottom. So you can see in 3D, what happens is you're sam, you're sampling the middle of the brain, axially, much more then you are sampling the edge of the brain. And that's where the increase in sensitivity is. Is in the middle of the axial field of view. Now, here is the, the data from a GE NXI Advance, it acqui, it sort of oriented in a projection view and here is the michelogram associated with that. So this part, right here, is this. And you can see that in the middle,

in this case, in GE, it does what we call, the span of three before, one, two, one, two, one, two, one, two. It looks like a 2D view, and you can see the striations. You can see, like, it looks striped across it, because each other's pla, each other's slice in this case, has twice as much counts as the other slice. You can see that. As you go to the other views, this is the second image, third, fourth, fifth, sixth, seventh, you can see we start to get sort of truncated views, of the brain, cutting off the top and the bottom and by the time we get to the very edge here, you can see we're only looking at the middle and not seeing the top and bottom. And so, again, you can see from this, that what 3D gives you is much higher sensitivity, in the middle of the field of view, in the middle of the brain here, and not as much, you know, and really no more than 2D on the very edges of the axial field of view.

Just to summarize, with respect to CD, 3D, you see that the sensitivity drops off towards

the edges of the axial field of view. Not towards the edges of the brain, but the edges of the axial field of view, but it does lead to about a four to five times increase in sensitivity over all, much higher maybe a factor of ten, in the middle of the field of view, but four or five overall. However, it also leads to an increase in scatter from about 15% scatter fraction to 40, to maybe 40 or 50%, depending on how big the patient is, if you take out, if you acquire in 3D mode. Also leads to an increase in randoms of outer field activity. We'll talk a little bit later about the fact that you need rebinning algorithms in order to, to, to reconstruct, reconstruct this and we'll talk about that, in a little bit later. Some devices, as you'll see later, when I look at the instrumentation, can acquire only in 3D.

Some give you the option of being able to do 2D and 3D. Now, if you think about it, which, what applications may 3D be most suited for, it would be those that maybe you think have less scatter and maybe not as many randoms out of field activity and the place that where it really is applied very well, is in the brain. Where the brain is a little bit in a, little bit smaller part of the brain, then if you try and measure through the thorax, particularly on a patient that looks like me, and so in that case, if you're tryin to measure in the abdomen, you may not wanta do 2D. 3D works very well in the brain. 2D maybe better in the whole body, particularly for large patients, where you have a lot of scatter, but as you'll see, some scanners will only allow you to do through 3D and you apply 3D in the whole body, as well. Attenuation correction, just like in a SPECT, we have to

acquire, we have to correct for attenuation because the probability of, of measuring something from the middle of the, of the patient, is less than if you measure it from the edge of the patient. And basically the four approaches that you can use, is one just to do a calculated approach, something maybe like a, in a way, sort of like a chain, we'll talk about that in a second. Do some measurements that allow you to, to, to estimate what the attenuation, is. You can do, you can take that measured, and it tends to be real noisy, as you'll see and do, in segmented in order to reduce the noise. You can use a singles approach, as I'll show. And then, more recently, you can use CT. In a calculated approach, let me just point something out, before, before we move onto, from here, though. In PET, you have to measure; you have to measure both photons, in order to

have an event. So, if the events in the middle and this one gets sort of attenuated by half of this distance, across the patient, this one gets attenuated by the other half, across the patient. And so the amount of attenuation along this line of response is dependent on how the thick the patient is along that line of response. Now, if you go over to this event, let's say here, you can see that this photon didn't get attenuated by very much, but the other one had to travel through the whole patient. So it turns out that it doesn't matter where along this line of response, the event occurred. You get the same amount of attenuation along that line of response, irrespective of whether the source was there, there, or even here. And so, you can do a correction by just knowing, you have to know what the shape of the patient is, and what the material is, but if you know that, you can

than do a calculation, that I know how deep, thick this is, let's say that's D here, and that,

and then they, in this case, they also estimating that bone's slightly different in red. You can do a calculation to correct for attenuation, it maybe in the brain or someplace where you have uniform attenuation, that might work, pretty well. However, another probably, more accurate approach and works better, let's say, in the thorax where you have different kinds of materials, lungs and things like that, and it, your better off using a measure attenuation approach. And this has been a very common approach that we use, in, in PET. And in, in this case, what we do is before the patient even gets into the scanner, we acquire a blank scan. Now, you can see along this line, and see, you can see that there's an external source here. There's external sources, in the scanner that will rotate around the patient, in order to allow you to make these transmission measurements. And so, here you do a blank, without the patient there, and let's say along this line of

response, particular line of response, you get a hundred counts per second. When you put the patient in there, and you do the same measurement, you can still see the sources out here. No, let's assume, no activity in the patient, you don't get as many counts, because they've been attenuated through the patient. And by taking the ratio of these two, 100 by 15, you get an estimate of what the, what the attenuation correction factor needs to be, at, for this particular line of response. And then, when you do the emission scan, what you do is take the counts that you get, in the emission scan N here and just boost them up by this fraction. And you do that for every line of response and if you remember to the sinogram, we said that every line of response was a particular pixel, in the sinogram. I don't, let me just see if I have... yeah! Okay, in the, every pixel in the sinogram and then

you just do the corrections. So here is a blank sinogram, you can see it looks somewhat uniform, this cross acting is because I didn't correct for normalization, but let's look in the middle of the brain. You can see that, let's say, it looks pretty uniform here. And then when you, come over and do the transmission scan, you can see it has a lot of less counts where the brain is. And so, let's say, you take the ratio of those counts, maybe this is a 100 this is 20, that would be a correction factor of five, when you do the emission scan, you take that same pixel, one same pixel, and boost it up by a factor of five, to correct for attenuation. Outside the body, let's say over here, that's 100, that's still 100 correction factor of one. So, you don't get a correction, if for, or basically more than one, if your outside the body. And this will work even if you're going through the

lungs and different material. So measured the attenuation correction in PET works quite well, even for nonuniform attenuation, and which is, as you may know, a difficult thing for us to do in SPECT. However, the problem with this, it's, it's a probably an accurate correction, but it can be noisy because you're taking a noisy sinogram, dividing it by another noisy sinogram and using that to correct a, an even third noisy sinogram. And so, we used to, when we used to do just simply, as I've described it, you spend as much time acquiring transmission scans as we did emission scans. We do five minute transmission scan and a five minute emission scan. We were scanning the patient half the time for just for the correction for attenuation. So, one approach to getting around that was to do segmentation. Here's the transmission scan. You can see, I used to say this was a poor

man's CT scan. It's actually sort of an expensive CT scan, but it's just a bad CT scan.

However, if you do a segmentation on that, let's say you say "Well, all the stuff that looks like this, will just say said a soft tissue. Then you can make it look nice and smooth. Even when you started out with noisy data and then you can use this nice and smooth sinogram... image here, to generate new sinograms that you used for the measured attenuation. And so, by doing that, you're able to use very noisy data, to do a reasonable correction of, of attenuation. So, now instead of acquiring the transmission scan for five or seven minutes, we can acquire for two or three minutes and get a reasonable transmission, reasonable attenuation correction. And so that's what a lot of people have done. But you're still acquiring for two or three minutes per bed position.

An approach that Phillips uses, is say, well, let's just assume that all the counts get measured over on this side, so why do you need to even worry about them? There is nothing attenuating it. There all gonna get counted. So use a single, source, single photon source, like cesium, and only measure the ones on the other side. Since your, since you, now shu, they're actually, they put a little shield here so they can cut down on the dead time on the detectors, on this side. They also, since they apply this in the, let say in the CPET, they maybe able to have good enough resolution that you can actually separate the two peaks and do them at the same time. You can get a high count of singles, much higher count in singles, then you do in coincidences. So, you able to inc... reduce the noise, just because you're able to get more events. And so, Phillips has used

this approach, which actually works quite well. And of course, in the last few years, they used a CT. So, you get a nice CT here. It obviously tells you something about the transmission through the, through the patient, but not necessarily exactly what the transmission is for 5/11. So, what do you need to do about that? Well, it does a pretty good job here. You noticed one thing that happens is here's where the data was acquired. Here's where we're interested in doing the attenuation correction. You can see that everything that's sort of soft tissue ends up looking all pretty much the same, way out here at 5/11. But there is a difference in bone and air and so if we can correct for the differences in bone/air by using the CT scan, by doing some kinda transformation, for the Hounsfield units in this energy range, up to what it would be for mU values, up in this

energy range, then maybe we could use the CT. And so here is sort of the, the transformation that sort of applied. You can see that it's sort of bilinear. And you can see that they use this correction for the Hounsfield units to mU values in this range, and then a slightly different one about, above that range. And actually, they're, they're many other companies are now experimenting with even a third one up here, to deal with contrast medium, where you end up having Hounsfield units in the 3000 range. So you can see that you get, if you do that map, then you can use the CT scan. Now, I said it took it, we used, even with segmentation, we used maybe three minutes per bed position and if you're imaging, let's say from the eyes to the thighs for a, a, oncologic PET scan, that's maybe six or seven bed positions. So, even at three minutes per bed position, let's

say it's seven, that's 20 minutes you're using to acquire the transmission scan. You, you

all know that acquiring a CT scan, a Helical CT scan from the eyes to the thighs, takes a minute. So, you're cutting down the time for the whole PET scan by, you know, at least, you know, 20 minutes by using PET CT attenuation correction, rather than using, the, the, the rotating rod sources, positron emitting rod sources in the scanner. However, it's maybe not exactly, not exactly as good with respect to accuracy, because you're not using 5/11, you're making some transformation. But it works reasonably well, it does have other issues that have to deal with registration between this PET and the CT, but it works quite well and really improves your, your, your throughput, by reducing your PET scan by a factor of like 15 or 20 minutes. I'm, I'm gonna change gears here a little bit and talk a little bit about reconstruction. In the last few years, you know, when I, when I first got into PET about 15 years ago, you determine that there's an, that there is some difference,

that you then back project and use that to make a better estimate of what you started with. And you do this a number of times, in order to estimate what the activity distribution is. And so, you may use the maximum likelihood, maximum likelihood algorithm, which maximizes the likelihood that the estimated, your estimate of the activity distribution within the body, which is what you're trying to image, at, it, that is the reconstructed transaxial slices, would lead to the measured projections, which is the raw data, that you have. And the algorithm that you use to try to make that, to, to estimate this, is this expectation maximization algorithm. And so for every, let's say, this is the transverse slice through the body, and here's where we get our projection data, around the body. For every pixel within the transverse slice, where you could, you're gonna be estimating,

you want to figure out what the, you have to estimate what the probability is that you could measure data that was admitted from here, in this bin, in the projection bin. And that probability that you estimate is called this A_{ij} , so before you start, you have to have some estimate of what that's going to be. Now, that estimate of A_{ij} can combine a whole bunch of things. It can combine, if you have attenuation, you can combine that into it. You could combine scatter, like this may more likely go here, but because of scatter, it could maybe go over there, and you can combine a whole bunch of physics into these weighting factors. So the A_{ij} 's can contain physical information affects the spatial resolution scatter attenuation, but then you apply those, you, you use those in the EM and L algorithm, and you can maybe do a more accurate reconstruction. This is sort of what

the, what the formula looks like. You can see here's the, here's the real projection data is D . The estimated, calculated or estimated projection data from your last estimate of what the activity distribution is D' . And so you can see that, if you take the ratio of those, that is a comparison of how close, D' is to D . So, you're gonna eventually make your new estimate based on that. However, you're gonna weight that by this weighting factor, which is the probability that, that data actually is real, you know, that if it's way out and it's not very likely that, that, that pixel is going, actually got any data, then you're not gonna give this too much of a, a weight, but if it's in a place where it may, it could really contribute, then you weight it a little more. And then you normalize for all the weights just down here, and now you use that to update your, your next estimate. And

you do this on an iterative approach and with maximum likelihood, you used to take

about 50 iterations, which took about 50 times longer than filter back projection, in order to get a good estimate. So, we went to a new algorithm, the OSEM algorithm and now what we do is, we, we, reconstruct or we look at just subset of the projection data, at any one particular time. So, at each step, we now project and back project only a sampling of the angles. That's maybe say, a quarter of the, of the, of the angles. Let's say the first, let's say the first angle, the fifth angle, the ninth angle, the fourteenth angle all the way around the patient. You perform the first step of the iter, iterative reconstruction. You update your estimate then, and then you go on to the next subset, the second, the sixth, the tenth, on and on. And so, what happens here is by the time you get through the whole

set of projection data, you've already star, you've already started to, to converge. And so this algorithm converges much quickly, more quickly than just going through the whole set of projections before you, you update your estimate of what the data is. And so, you can see that the data starts to converge, even before you finish the first iteration. And so, this tends to, start to converge in very few, maybe I say three to ten here, maybe even as few as two, as you see down here, iterations, instead of 50. And so this very much speeds up the, the algorithm. So now, we've gone from maybe half an hour, for a reconstruction, down to a few minutes. Maybe it's two or three times as long as a filtered back projection, but not 50 times longer. And for our GE scanner, we use 28 subsets, 12 projections per subset and two iterations for our reconstructions. And you can see here is

the same exact data reconstructed with filtered back projection and with OSEM. And the big advantage, you can see, is that you don't see the streaks that you see in filtered back projection, in particular, let's say, here is a transverse slice through the bladder, you can see that it really has reduced the artifact of streaks that you get from the bladder. So, it handles streaks and noise much nicer, it looks much nicer. There is still some question with respect to noise in the liver, for example, that can be a little bit difficult if you're, when you're first starting to use the OSEM. But once the, the physicians get used to using it, they tend to like these images that look much nicer than, like those images. 3D reconstruction, we talked about 3D before, how do we reconstruct it? As you probably know, reconstructing 2D slices as we've just talked about is much easier than dealing

with 3D data, like in a cone beam, cone beam kind of configuration. So how do we deal with that? Well, the way we deal with it, we could do a fully 3D reconstruction, but that is not the current way of dealing with it. As we somehow rebend the 3D data to look like 2D data and then reconstruct it, using the 2D algorithms. So, we take the 3D object and we acquire 3D data. And remember 3D data had all these angled, not just straight across, sinograms, data from straight across sinograms, but also a bunch of angled sinograms, as well. And we end up with a whole bunch, maybe, let's say M^2 , that would be the opti, maximum, you could get sinograms, a bunch. You then rebin those into a bunch of 2D and you see, you've reduced the number to $2M$ minus 1 now. And then reconstruct that with a 2D algorithm and this is much faster than reconstructing the data with a 3D

algorithm. However, in order to go from a 2D to estimate where the 3D data is, can be a

little bit inaccurate. The simplest way to do that is to use, what's called Single-slice rebinning, which is, if you have a line of response that goes from here to here, just assume that you take the half the distance between this one and this one and you say the data really came from here. Well, you can see that, that works not too bad in the middle of the patient, but not so good out here. You can see that, saying that this event came from here, really isn't doing very well. Now, it, it's not, I sort of exaggerated it, because the ring is so small, these, these detectors are so close to the head. In fact, really, these detectors are a meter away from each other and so really, it's not quite as bad, this angled one maybe is more like that, but still at the edges of the field of view, it's not gonna do such a great job. This is fast, and it works for the center of the field of view, but not so

good on the peripheral field of view. So the more common approach we're using now, is Fourier rebinning because it turns out that your ability to do that rebinning and Fourier space, works much better, than doing it in real spaces, you do it with single slice rebinning. And so the approach we use is, we take, we, we initialize a stack of Fourier transforms and 2D space and then for each oblique sinogram, remember we have sinograms that are straight across, but a lot better angled because of the 3D acquisition, we take each sinogram, we then do a Fourier transform of it and we determine where that should be placed, according to this formula here. And you can see that, that it iteration is really only a 1D iteration here now. It's related to the angle of orientation of that particular slice, and the, you know the 48 transform of the X direction and the angle

direction of the... in the Fourier transform of the sinogram and that tells you, gives you a new estimate of where you should put that data. You add that all up as you go on for each pixel in the sets of sinograms. You normalize for the fact that you're gonna be over sampling some areas, and under sampling some others, you correct for that. You then take the inverse Fourier transform of that and now you have a set of 2D sinograms that you then just reconstruct. Sounds pretty complicated, but it's a lot faster than doing a 3D algorithm. Now, I, I've gotta about, I think about eight minutes left, so I'm going to move on to talk about PET instrumentation, just a very quick review of PET instrumentation. And I've sort of broken these up into two groups, the high resolution scanners that are out there and medium resolution scanners. And you can see here that

the high resolution scanners, I've listed at the top. You can see that the detectors are basically 4 mm, in the trans, around this ring, in the transverse direction and maybe slightly longer in the axial direction. And because their about 4 mm around the ring, as I've said, you can estimate that the resolution is gonna be a little bit more than 4 mm and your, you be it right, about 4.6, 4.8. You can see that there's a lot of detectors, 18 thousand for the Siemens HR Plus, 12 thousand for the GE, 17 thousand for the Phillips Allegro. They all have pretty good resolution. You can see that the detector material is BGO, BGO, GSO, as we've talked about before. In the sensitivity numbers, the first number is, is 2D sensitivity and the second one is 3D. You can see the difference here; let's say here is about a factor of five difference between 3D sensitivity and 2D, here

about four and a half. You can see that I didn't give a 2D number because this machine only runs in 3D, okay. Now when you go down to sort of the medium resolution, you can see now the detectors are about 6 mm, resolution's about 6 mm. You can see here, that this has, you know, 18 thousand, 12 thousand, 17 thousand, 9 thousand, 9 thousand, six. This is the one that had the six Gamma camera looking head, so they're very big detectors and there's only six of them, and it's sodium iodide. Here's and LSO scanner, resolution about the same, but you can see the sensitivity for that one that sort of looks like Gamma camera heads is much less than these. It's about twice the sensitivity in 2D, but, only about half the sensitivity in 3D and you cannot acquire 2D data, it has no septa

in it. Here's the inside of our GE Advance that we have at Children's Hospital. This is what the guts look like, a bunch of, the blocks are down here. A lot of photo multiplier tubes, a lot of hardware, which is the reason that these are so expensive. Of course when you put the covers back on, it sort of looks like a CT scan. PET CT scanners, a bunch of different models, and I'll talk a little bit about those. And, and the approach pretty much has been to take a high end PET scanner, combine it with a high end CT scanner, into a combined unit. This is the GE Discovery LS. You can see that the, that you can get the CT part of it, in 4 to 16 slice flavors and you can get a high end PET scanner. This device, has retractable septa, so you can acquire both in 2D or 3D, and you get, this is the image I showed earlier, but you get a nice CT scan. Here's the PET scan, you see

something there, you see something in the PET scan, you can do a fuse, fusion of, of the two informations over each other and see that the PET data lines right up where the tumor is. GE, I guess about a year and a half ago, came out with a new version of their PET CT called EST, it's a, it has a bigger hole, you might be able to tell here, but it's an interesting other wise, interesting, com, choices that they made. Here with the covers off, you can see the CT part in the front, there's the x-ray tube down here. Here's the bank of detectors, over here, then you move around to the back and you can see the PET part, which looks very similar to the GE scanner, I showed you before, slightly different, but you can see it looks very similar. So, basically it's just a CT scanner, sort of nailed to a PET scanner. On the difference between the ST and the LS, you can see is that this is

the newer model, the ST, has larger detectors, six versus four. So, not, might not be surprising that the resolution is a little bit worse, actually with the newer scanner than it was with the older scanner. It's more like 6 mm rather than 4.8, shorter septa. It's still does do 2D, but it doesn't have as long septa. Those septa, again, I should've said this when we talked about septa, don't really think of them as a collimator in the nuclear medicine sense, but think of them as an antiscatter grid in the, in the, in the radiology, in the radiography kind of sense. And so you can see it, by shortening the septa, you are allowing more scatter in, but increasing the sensitivity. So that leads to a higher sensitivity by about a factor of two, in, in 2D and about a 30% and you'll even see, because the detectors are also closer in, about a 30% increase in sensitivity in 3D mode.

A larger patient bore, which makes it real, much nicer maybe, for, for radiation therapy applications and things, has no rod sources. CT is your only option for doing attenuation correction, which may be just fine. But it does allow you to use both 2D and 3D PET

acquisition. You can see here, that this is the Siemens Biograph HR Plus, was a BGO scanner, a large port, you can see that I got this from Siemens, it's got an optimum bed design. But they have recently gone to a, to, a higher resolution LSO detector, more like the HR Plus, but with LSO. You can see the detectors are smaller here, leading to higher resolution. And you can see the difference in quality between the high res scanner here, and the HR Plus here and their more conventional PET CT down here, much better, maybe for imaging the brain, maybe not that different for imaging the whole body. And

then Phillips there's actually one of these out on the floor, if you'd want to look at it, has, here's the CT part, here's the PET part, they can actually separate these two, for, for maintenance and things. And they would say, tell you that the patient feels a little bit better because they're in this... the tubes starting to get a little bit longer, but they do see some white through it. And here's a combination, I mean a comparison. You can see on the seam inside, and it had a BGO scanner that was high resolution. It had an LSO scanner that wasn't and in a recent bout, I guess about a year and a half ago, introduced a high res LSO version, which has many, 23 thousand detectors now. Only acquires in 3D, all these no septa, only 3D and only CT for attenuation correction. GE we talked about, they have two flavors now. We can see ones higher resolution than the other, just like

these two, but allow you both 2D and 3D, so larger patients, you may want to do 2D, smaller patients you can probably do 3D. And then the Phillips over here, only in 3D, good resolutions here. So you can see sort of a comparison between all these. And I think I have maybe a minute to talk a little bit about special PET devices that are out there. Here is a CTI who make their scan, make the scanners that Siemens sells. They developed a head only, very high resolution scanner here. Now you can see this has 120 thousand little detectors and they're only 2 mm. So this has a reconstructed resolution for brain imaging of about two and a half mm, very high resolution for imaging the brain and very images if you get a chance to see them. Some people have developed, we worked a little bit of this when I was at Wake Forest, TEM technologies and a number of other

investigators have developed what they called positron emission mammography, where now they take two little detector blocks, put them on either side of the breast, and if you could see now, we're sort of, we've sort of incorporated this into a stereotactic mammography unit here? So, why, right after you acquire the mammogram, you can put these two little blocks under your side, and acquire sort of a little PET scan, to see whether the FDG matches up with what you see on the, saw on the mammogram. And so here is an image I got from Wake Forest, when I was there. Here's what they saw on the mammogram. Here's what they see on the FDG scan, and maybe it helps them better place the needle for biopsy, in where the tumor actually is. And lastly, very interesting developments, with respect to small animal imaging, this is the micro PET developed by

Concord. You can see the very excellent images of rats. This is a fluorine-18 bone scan, fluorine-18; fluoride is a bone scanning agent, of, of a rat. And you can see very good detail. Of course, this is a good size rat. This is a good, you know, good size Boston rat, but you can get really good images and not only Concord, which is now part of CTI? So, from Knoxville? But also GE has a small animal scanner that they, has a, about 1.6 mm

resolution? And Phillips has a scanner, which is about 2 mm resolution. So you can see that this, it's an interesting thing because it's driving in many ways, the technology. I mean obviously, you're not, can't really image it. It's, very, you may be, you can fit someone's wrist in here, but you're not going to be able to fit a human in there, but we're learning a lot more about what the limits are cause, you can see we're getting down into

one 2 mm, that's down where we think that the pot, you know, the fundamental resolution of PET is. And so we're learning a lot of things by looking at these small animal imaging devices. And then they go for about maybe three quarters to a million doll, million dollars for those devices. So just in summary, scanners are being optimized for clinical imaging, in particular, oncologic imaging, cause that's really driving the field. So, they're trying to have higher sensitivity, higher count rate capability. New detector materials you've seen, if you looked at that, you could see that the newer detectors are usually maybe LSO or GSO, Seaman, GE is still staying with BGO at least for the time being. PET CT is really coming into the field, I think I'm told by the vendors that now half of the PET scanners that they sell, maybe more than that, are PET CT scanners. And

then small animal imaging is very interesting and that's really, I think, pushing the technology, as well. So, thank you very much.