Herbert Y. Kressel, MD  
Hi. This is Herb Kressel and welcome to the June Radiology podcast. This month we have three really interesting discussions for you. First, my colleague Dave Kallmes who is the Deputy Editor for Neuroradiology will be speaking with Dr. Mayank Goyal who with his co-authors wrote a very important paper on the “Analysis of Workflow and Time to Treatment and its Impact on the Outcome of Endovascular Treatment of Acute Ischemic Stroke: Results from the SWIFT PRIME Randomized Controlled Trial.” The trial and this paper pre-print publication have received a lot of interest and I think our viewers and listeners will benefit greatly from this discussion. Next, my colleague Dr. Ali Guermazi the Deputy Editor for Musculoskeletal Imaging will be speaking with Dr. Johannes Roedl who with colleagues from Thomas Jefferson University Hospital reported on the “Potential Utility of a Combined Approach Deploying Ultrasound and MR Arthrography in Imaging Medial Elbow Pain in Baseball Players.” I think whether or not you’re a baseball player, this is an important topic and you will find the discussion of interest. And finally, I’ll be speaking with doctors Kathleen Brandt and Celine Vachon of the Mayo Clinic and Mayo school of medicine on their study “Comparison of Clinical and Automated Breast Densitometry Measurements: Implications for Risk Prediction and Supplemental Screening.” Doctors Brandt and Vachon reported extensively comparing different commercial software packages for this to conventional BI-RADS interpretation. I think this will be of interest particularly in the area that we are in of breast density notification in many states in the United States. As always I hope you enjoy this month’s podcast.

Mayank Goyal, MD, FRCPC • Ashutosh P. Jadhav, MD, PhD • Alain Bonafe, MD • Hans Diener, MD • Vitor Mendes Pereira, MD • Elad Levy, MD • Blaise Baxter, MD • Tudor Jovin, MD • Reza Jahan, MD • Bijoy K. Menon, MD • Jeffrey L. Saver, MD • For the SWIFT PRIME investigators

David F. Kallmes, MD  
Hello and welcome to this video podcast. My name is David Kallmes. I am the Deputy Editor for Neuroradiology. I’m joined today by Maynak Goyal who is an international neuroradiologist and professor of radiology at the University of Calgary and was the PI of the ESCAPE Endovascular Revascularization Trial. We are here today to discuss his paper entitled “Analysis of Workflow and Time to Treatment and its Impact on Outcome in Endovascular Treatment of Acute Ischemic Stroke Outcomes: Results from the SWIFT PRIME Randomized Controlled Trial.” Welcome Dr. Goyal.

Mayank Goyal, MD, FRCPC  
Thanks David. Thank you for the kind introduction. Good to chat with you.

David F. Kallmes, MD  
We should just start with your telling what you did briefly and what were your major findings?

Mayank Goyal, MD, FRCPC  
As probably your audience knows, the world of stroke has recently changed. There have been five recent NEJM published papers showing the superiority of endovascular treatment over what was previously a standard of care. As a consequence, the standard of care now has changed. I was significantly involved in two of the trials, ESCAPE and SWIFT PRIME, and subsequently while we were running SWIFT PRIME we did a very detailed workflow analysis and collected all kinds of times that related to how the workflow was happening and where the delays were. This paper is essentially analysis of workflow, where are the delays, where we can improve, and the impact of time on our outcome.

David F. Kallmes, MD  
And what did you find?

Mayank Goyal, MD, FRCPC  
And in terms of summarizing the main two results, one is something that is intuitively obvious and which we all know, that time is brain. The slower we are the less likely it is the patient having a good outcome. This is extremely, extremely important. It has a powerful effect. The second important finding that was there in this paper is the adverse impact of not going to the correct hospital the first time around. The way that it played out in this study, approximately two-thirds of the patients went directly to the comprehensive stroke center, and one-third of the patients went first to the primary stroke center, got their (inaudible) started and then were transferred over to a comprehensive stroke center. The median delay by not directly going to the comprehensive stroke center was around two hours which is massive for the population that we are talking about in terms of influencing outcome.

David F. Kallmes, MD  
Can you put into perspective for us say on a per 30 minute basis or so what does that do to expected outcome?
MG  I tell you the answer to this question in two different ways. One is as you know in any study we are limited by calculating this information based on the patients that were included in the study, and the study that we have done, for every 30 minute delay initially the effect is massive, but just to give sort of a ballpark figure to the audience is a ten percent reduction for every 30 minute delay. Now it’s not linear oriented, the delay is must more influenced. But the other part which is important to remember is what I call as the denominator fallacy and I’ve written about it extensively, which is the idea that all these patients were selected on the basis on imaging and as we get later and later, the likelihood of finding favorable imaging with get significantly reduced. That is based on experience and based on (inaudible) and based on the biology, so if you think about it if you’re one minute after the onset of stroke, hundred percent of the patients will have favorable imaging; and if you’re at 24 hours nearly all the infarcts will have evolved and sort of happened. So obviously there is a curve between one and 24 hours. We don’t know exactly the shape of the curve. But the other part which is important to realize is that the later you are the less is the likelihood of having favorable imaging which is a part which is not quantified in this paper, but is very important for everyone to understand.

DFK  So to paraphrase, for every 30 minutes delayed, there’s a 10 percent decrease in the proportion of patients who achieve a good outcome. Isn’t that what you’re saying? And the sample is biased to say that people who were delayed were excluded because of imaging exclusions. So really this is a best case scenario and maybe the real world is worse.

MG  Exactly, exactly, so the way I teach my residents and fellows is if your 30 minutes from onset and let’s say you have 100 patients with MR occlusion, likely 98 patients will be having good imaging. If you are at four hours probably 50 of them will be having good imaging. If you’re at seven hours probably 10 of them will be having good imaging. So in those ten at seven hours, maybe six of them will have a good outcome. But if you look at it at a population level, it’s six out of a hundred which will have a good outcome based on this. So it has a double impact; one is you get better as you go faster, but then the other part is you’re going to be treating more patients and getting more patients an opportunity to have better outcomes.

DFK  Sure. Speaking of imaging, I know there’s tremendous debate and the signs of influx about how much imaging do we need if time is brain and I’ve got a patient with a dense MCA and a good story, can I go straight to angio and when should I do even a CTA or a CPT?

MG  Okay so I’ll answer this question in a few different ways. First the thing is in this study we analyzed the data in terms of was the CTP used for decision making or not used for decision making and the way that the results played out, CTP did not have any impact. The patients that were treated without using CTP information, the effect size and overall impact was just as much as in the patients who did have CTP. The other part is on average waiting for the CTP and doing the analysis took an average of 18-19 minutes extra which is not a small amount of time. There was a difference in workflow. I have corroborative information in terms of that overall SWIFT PRIME workflow was on average lower than ESCAPE workflow by around 20 minutes or so and in terms of protocol one big difference between the two was that in ESCAPE CT perfusion was not required. That’s only one part of it. The second part of it is whether a CTA should be done or not. We’ve done this analysis in IMS3 where we looked at patients with CTA and without CTA and this is apples and oranges. It’s not a precise analysis. This was published in Circulation, so there’re two things that we found was that CTA can be consistently done in most centers within five minutes. The second part was because CTA allows you to evaluate the (inaudible) bifurcation this, that, and the other; there is a possibility that it naturally speed up the rest of the procedure that is what we found when we analyzed the data for our study. Now to answer your precisely, if you have a 45 year old where you are expecting reasonably (inaudible) straight vessel with a dense (inaudible) of 20, one hour from onset; sure go ahead and save on the CTA and go to the next stage sort of thing. But if you’re an 82 year old and things like that, I think that CTA must still be used for in terms of planning the rest of the procedure.

DFK  And CTP? When should I do a CTP?

MG  We don’t do it. We don’t recommend doing it. There’s some talk about whether in the later time windows CTP could potentially be of use. That is yet to be seen in ESCAPE. There were a total of 53 patients that were beyond six hours, the effects size on those 53 patients was the same as the rest of the trial although it did not reach a statistical significance. I do want to point out though that if you don’t do CTP one of the problems that is there which people face in the community is the ability to interpret the non-contrast CT scan and the ASPECTS scoring which as you know came out of Calgary. We do have sort of various solutions for it, one is to optimize the quality of the CT scan to see the grey-white differentiation and the second is to look at collaterals because collaterals goes hand-in-hand with the ASPECTS. So the way that we practice locally is you look at the non-contrast CT, you look at the collaterals and modified CTA, you go back and look at the non-contrast CT and use the collaterals to further enhance your interpretation of the non-contrast CT. And you will overall find you can do this in two to three minutes and works very well.

DFK  So are you using the multi-phase CTA as a perfusion scan or you’re simply using it as an adjunct to your ASPECTS scoring?

MG  We’re not using it as a perfusion scan for sure, we’re using it as a collateral scan, and the collaterals we’re using because good collaterals means good ASPECTS and bad collaterals means bad ASPECTS 99 percent of the time
and there are two exceptions to that. One exception is if the patient is 30 minutes from onset and has really, really bad collaterals, but the ASPECTS changes haven’t set in. The second exception is let’s say the patient was hypotensive on the scene, the collaterals collapse, the brain died by the time the patient comes to imaging, the blood pressure is back to normal and you can see the collaterals again. So those are the only two disconnects between the two but we have an experience of over 1000 patients now where basically the two go hand-in-hand. And the other part which I like to talk about is that ultimately there are only two decisions or one decision to make, go endovascular or not go endovascular. The way that we practice is you look at the CT head, you look at the collaterals and modified CTA, if the two go hand-in-hand just go. You make up your mind is the ASPECTS 8ish, 9ish, 10ish, good collaterals and the patient is sort of otherwise worth fighting for, go open the vessel.

**DFK** Okay so you mentioned a two-hour delay for going to a primary stroke center as opposed to a comprehensive, what about delays within the radiology department for our audience? What should they take from this paper in terms of how to modify workflow to diminish any delays?

**MG** I like to use the this word which I call parallel processing in terms of from the time that the patient is about to hit the ER so one thing that we set up many, many years ago in our hospital is pre-notification where we are already notified before the patient arrives to the hospital. And if you think about it at the level of emerge once the patient hits the hospital, there are four or five things that need to be done. One is obviously the ABCs, the vital; the second is setting up a good IV line; the third is someone has to take history; fourth is someone has to organize imaging and those kind of things. Rather than one person doing it one after the other, the idea is to do it sequentially. The other part is for the radiology department to treat it as a super image to give it the same status as what you would in cardiology for an acute coronary syndrome or for that matter what an acute CTR trauma would get. And essentially the radiology department has to be sort of primed for it that is a super priority. The third thing that we put into action is we always have so to speak a fake ID ready for a stroke patient who comes in similar to a trauma so that we’re not spending time typing the patients ID in the system, there’s always so to speak an unknown ID ready to go and somewhere along the way the patients actually ID catches up with it. So we sort of just have an unknown ID ready for a stroke or trauma patient and just file the orders and it says acute stroke and there’s an acute stroke protocol which is built in which is an optimized CT head and a multiphase CT. So bang, bang, hang you get the patient to the CT, the CT is already set up, the contrast is loaded, takes a total of three minutes of imaging, non-contrast CT multiphase CTA and if is during daytime hours we all can go to the CT scanner, make the decision in the CT scanner and go straight from there to angio. Now obviously if it’s after hours, there’s a slight difference in workflow, but even at that point in time we try to sort of set up things in such a way that we minimize the time in radiology and try as much as possible to take the patient from CT to angio as opposed to from CT going back to emerge and then coming from emerge to angio again.

**DFK** Sure. So do you have any data from your center to try to quantify how much difference that approach makes over the standard approach?

**MG** Like in the sense that we don’t have data in terms of sort approach A versus approach B, but we do have data over the last ten years as to how we have speeded up things and essentially I wrote this paper in 2011 I think or 2012 in GNIS where at that point in time our CT scan to reperfusion, not groin puncture, CT scan to reperfusion was 84 minutes. So essentially that is for sure that across multiple centers it’s doable. Now that is not to say that it would be doable in every case and that is not to say that we should sort of sit on our laurels and not continue to try to improve.

**DFK** So what about the issue of administering IV tPA, should that be hanging as soon as the non-contrast head CT is performed? Should it get in the way of any other aspect of the patient’s care?

**MG** Basically the current guidelines are IV tPA should be given as early as possible based on IV tPA criteria. The way that we function during daytime hours is if are converging at the CT scan, the CT head gets done. If it’s (inaudible) for IV tPA the neurologist says okay let’s start mixing the tPA and whenever the tPA is ready to be given, they go ahead and give it whether the patient is on the CT scan table, it doesn’t matter. Practically speaking what happens is, by the time they’re mixing up the tPA the CT angio gets done, and right as the patient is coming off the table on CT angio they go and give the bolus. Now obviously after hours if the patient is coming in at 2:00 a.m. as you can imagine there’s slight differences and slight delays that get produced at the current moment if it’s after hours. But at the same time, I would say now in our workflow probably 50 percent of the time we get the vessel open before the tPA (inaudible.)

**DFK** Okay so while we have you on the video, can you tell our audience what’s coming down the pike in terms of stroke intervention research either in progress now or completed.

**MG** A couple of different things. One is as of now based on the current guidelines from the American Heart, the American Stroke Association, the treatment is approved from onset to groin puncture under six hours. In ESCAPE we go from zero to twelve hours but currently it’s ap-

Radiology 2016; 279:827–837
Johannes B. Roedl, MD, PhD • Felix M. Gonzalez, MD • Adam C. Zoga, MD • William B. Morrison, MD
Mika T. Nevalainen, MD • Michael G. Ciccotti, MD • Levon N. Nazarian, MD

Ali Guermazi, MD, PhD Good afternoon. Today a podcast about musculoskeletal radiology and the paper that we choose for this month is titled “Potential Utility of Combined Approach with US and MR Arthrography to Image Medial Elbow Pain in Baseball Players.” I have the honor and privilege to have with me on the other actual side of Philadelphia Johannes Roedl from Thomas Jefferson University, a hospital in Philadelphia. Thank you for being with us, Johannes.

Johannes B. Roedl, MD, PhD Thank you very much for the invitation.

AG Actually my first question and I do think it would be the question of anybody, what prompted you to do this study?

JBR We at Thomas Jefferson in Philadelphia we have a pretty high population of baseball players and baseball pitchers especially. We do the imaging for the professional sports team, the Phillies, and many years ago my colleague and a pioneer in musculoskeletal ultrasound, Dr. Nazarian, went to spring training to Clearwater in Florida with the Phillies and he started MSK ultrasound there and he really on an initially experimental basis tried the

proved from zero to six hours and there are a couple of studies that are running from the beyond six hour window and it’s likely that they will get finished in the next couple of years and we’ll see further data in that regard from a late window perspective. But I do want to stress the fact that even if the late window trials are positive, you cannot get past the denominator fallacy that we may be able to show that endovascular treatment is great for certain selected late window patients but that does not mean that we should not continue to make efforts to get the correct patient to the correct hospital as fast as possible. So that’s a study that is ongoing which we should see the results of. The second part that people are starting to talk about is that in a typical sort of MI occlusion do we really need to give tPA or not an instead of case for doing a trial in which we randomize tPA versus no tPA and there’s are few trials that are being talking about; one in the Netherlands, one in U.S. and we’ll see how that all plays out in terms of sort of how better we can save the money and the complications and the infrastructure required for giving tPA. Those trials should probably start soon. To my mind what is the most exciting thing in my personal opinion is that for the first time in human history we have a temporary MCA occlusion model in which a patient comes with an MCA occlusion and we are able to consistently open it. And to my mind this is a time to go back to world of neuroprotection and is there something that we can do that holds the core together, doesn’t allow it to expand, the neurons stay alive while you go and open the vessel. That is what we in Calgary are planning to devote to the next four or five years of our life that will be an ESCAPE 2 trial or the ESCAPE NA1 trial we will be working with this compound, NA1 we’ve been working with it for many years. We did a previous study, Michael Hill was the PI which was called ENACT. It was in aneurysm coiling patients and that was sort of proof of principle study and a safety study which insured that the drug is traumatically safe and now we are embarking on this study which will be sort of a phase 3 study, to test neuro-protection along with the revascularization. So that is another exciting opportunity that is open to us collectively in terms of being able to help a greater number of patients.

DFK Alright well Mayank anything that we didn’t talk about that you want to touch on before we finish?

MG Another interesting thing that we’re doing is that there were five trials that were published recently and in fact there are two other trials, one is called TRACE which is from France which hopefully should be published soon. Another one which is from U.K. which is called BEAST. The way that it all played out was that we were able to set up a collaboration across all the five trials which is called the HERMES Collaboration in which we have a patient level database that we are putting together of all the five trials and hopefully all the seven trials. I’m sort of chairing that committee together, the HERMES Committee. The first paper from HERMES got published a couple of weeks ago in Lancet. The second paper from HERMES would be a time is brain analysis, that’s what we’re doing right now, but from the point of radiology, the very interesting thing that we’re doing is we put together an imaging database of these 1,286 patients right now and maybe we will add the TRACE and BEAST patients which make it to a 1,700, 1,800 patient database, imaging database, and we will sit and reinterpreting all that imaging and we’ll have a much better feel for issues like what is it that is a predictor of bad outcome; what is a predictor of tPA is not going to work? That will be the best database on this subject ever. Watch out for some exciting publications that start to come out of the HERMES database.

DFK Alright well Mayank thanks for joining us and congratulations on your paper and please send us more great papers in the future.

MG Thanks Dave. Good to chat with you.
ultrasound on a few players. That's how it all started. That was about 15 years ago and slowly he published some important papers initially and slowly we incorporated his findings in clinical practice. This study includes patients from 2003 to 2013. That was 2003 was where we seriously in general had baseball players imaged in that way meaning MR arthrography and stress ultrasound. But the initial ideas of stress ultrasound and the idea for the study really came from Dr. Narzarian and with baseball players from the Phillies in Philadelphia.

**AG** I do think the span is multiple years but it's impressive because the number of players that you included is huge. Let's actually be brief and I want to ask you the same question, what are the main findings of your study in brief?

**JBR** The main findings are really that the usual so-called gold standard MR arthrography that is used for imaging of medial elbow pain in baseball players is a highly accurate study but it can further significantly be improved by adding an ultrasound to the imaging of medial elbow pain, and specifically stress ultrasound and conventional ultrasound. Adding those three tests meaning stress ultrasound, conventional ultrasound, and MR arthrography together had substantially higher accuracy in diagnosing patients with medial elbow pain that each would now be by itself.

**AG** Okay in essence why do you think this is an important finding? In another actually word if you want, how your findings would impact or would change care in these baseball players?

**JBR** It is extremely important since the surgery for an ulnar collateral ligament injury for example, UCL tear, Tommy John surgery, is first of all technically complicated and the recovery time is long. It takes at least a year, 12 months, and the return to previous level of play is not guaranteed so many players they come back to play but they never perform as they did before. It is important to avoid unnecessary surgery and one has to be certain that there is a UCL tear before a radiologist or an orthopedic surgeon sends the patient to surgery. And it is important any study that would increase the accuracy to diagnose UCL tears is really important especially for baseball players.

**AG** I have a question there, Johannes, in your opinion and I know this is not actually part really of actually what you are going to publish for this paper for May in Radiology, in your opinion what is the relationship between the player age, the number of pitches thrown or activity level if you want, and the relative utility of the ultrasound and MRI or MR arthrography?

**JBR** That is a great question. Our study really shows that MR and ultrasound together are very strong for diagnosing UCL tears. However when it comes to findings that are more age related including posterior medial impingement at the elbow which is cartilage loss at the posterior medial aspect of the elbow, and common flexor tendon and muscle injuries; ultrasound doesn't add much. You can imagine that cartilage defects can't be seen on ultrasound. That's really MR arthrography strong. In older patients experienced pitchers that these findings you know muscle tears, strains, and cartilage issues become more important. So the older the pitcher gets, the less value the ultrasound had. In general ultrasound can be strong in younger, middle-age patient, where UCL tears are the most common injury. In older patients of course it's so important if the UCL is a primary culprit, but if cartilage issues come into play or if ulnar nerve problems, then MR is really strong and ultra adds a little less.

**AG** Thank you. Actually looking at the images in your manuscript and paper, I found really the images of the MRA and also the ultrasound impressive. I would like you particularly to comment on Figure No. 4 especially why stress would change the opinions on ultrasound of the UCL tear?

**JBR** Yeah so on that Figure 4, you can see on Figure 4b the upper image that's the rest ultrasound image; and that white line outlines the ulnotrochlear joint that's the medial aspect of the elbow joint. That's under rest and then image 4c really is under stress so you can see how the ulnotrochlear joint gap widens to in this case to about 7.7 mm and previously 2 mm so there's substantial widening of the ulnotrochlear joint with stress and the reason is that the ligament expands the joint, the ulnar collateral ligament is torn and therefore is lax and the joint shows increased gapping so that's the whole idea of stress ultrasound. You really try to show joint gapping in there by indirectly diagnosing a ligament tear.

**AG** That is quite impressive. When I look actually at the paper as a whole as I said before, there is the kind of actually MR arthrography which is compared to the ultrasound, and someone will maybe ask the question, the follow-up question, and especially in Europe or actually in Asia, they will say do you thing MR alone without arthrography is sufficient for such a diagnosis or you really need actually the MR arthrography? In your experience, what you find actually that MR arthrography will give you more than simple MRI of the elbow?

**JBR** Right that's a good question as well. We are big believers in MR arthrography not only for elbow but also other joints including the hip or the shoulder if indicated. The reason for elbow MR arthrography is that the cartilage is better imaged in general in terms cartilage defects for posterior medial impingement or vessels (inaudible) as well as intrarticular bodies of (inaudible) after contrast is put in the joint. And for the UCL tear you can see it on Figure 4a as well how a contrast really undermines the tear and trickles into that proximal UCL tear and that will be difficult to see without contrast. So we think really contrast adds substantial values especially in professional or collegiate players.

**AG** I looked also at actually at the diagnostic value of ultrasound and I saw that the sensitivity is around 81%, 91% specificity and accuracy of 88% when you use ul-
trasound alone, so for me the question would be do you think that portable, on-field ultrasound would be helpful in making quick diagnosis, enhance medical decision in injured baseball players? We know in Europe we know that they do in soccer, soccer players, so is this something that you can foresee for baseball players with an ultrasound on field?

**JBR** Yes I can absolutely see that in the future. Our study really shows the accuracy is high and I think it’s an ideal tool that is portable. As I mentioned, Dr. Narzarian goes to the spring training and examines the players there and would be very suitable for on bench on the field diagnosis, whether it’s in baseball or any other sport. I think ultrasound will go into the stadiums.

**AG** My next question if you are actually really happy with the ultrasound and you think it’s really accurate, sensitive specific, etc. to the level that is really acceptable here and more than acceptable; now my question would be simple. Why do we need MRI?

**JBR** Yeah, very good question as well. When it comes to medial elbow pain there are many other diagnoses that have to be ruled out or considered including posterior medial impingement which again you can barely see on ultrasound, you can’t see cartilage defects and intraarticular bodies. The muscle injuries, tendon injuries, these are only really seen in MRI and well as stress bone marrow edema which is a common injury in pitchers where at the distal humerus or the olecranon, we see a lot of stress bone marrow edema in pitchers and that can cause medial elbow pain. So you need really the big picture information from the MRI. In addition, I think that even for the UCL I feel much more comfortable and we do a lot of ultrasound at Jefferson, but we always have the MRI at the same time. It’s really more like a complimentary exam. The ultrasound is extremely useful, but we use it as an add-on for MRI. The other reason why ultrasound I think is still in a way controversial is that it is operator dependent so you need some experience. It’s not easy to learn ultrasound from books. You have to actually practice and perform ultrasounds before you can make a comfortable diagnosis, I think when the bottom line is that it’s a combination that is key in the future because stress ultrasound combines the anatomy from ultrasound with the anatomy exam from a physical exam so in a way stress ultrasound is really a fusion modality between physical examination and ultrasound. I think that’s the beauty of stress ultrasound.

**AG** Yeah there is always actually that learning curve as you know for ultrasound and you delineated that very well and I do think that’s absolutely the case here for the baseball players. So my last question here since Radiology is an international journal as you know and distributed worldwide. Baseball is only popular in the United States in Cuba and Japan so we’re left with a lot of actually part of the world that would not be maybe interested I would say perceived by actually baseball, but there are other sports, throwing sports, as you know for example handball that can cover actually the rest of the world excluding the U.S. but now actually Europeans are very interested there. What do you think actually this paper would be or these findings would be actually also translate to actually someone who is throwing but (inaudible) player elbow?

**JBR** Yes it’s the same pathology. Any sport that is overhead throwing and that puts valgus stress on the elbow really is affected from UCL tear so in Europe it’s handball, it’s tennis where we see UCL tears so it is really any sport that has overhead throwing or racket sports that have valgus stress involved that can really benefit from the study and from stress ultrasound.

**AG** Thanks, Johannes very much. This was actually a pleasurable and I hope that you get more information for us next time so we hope for the next paper about the baseball players maybe and thank you very much for your time. Thanks for all the team for all the hard work and thanks for being actually a leader and an author for Radiology. Thank you.

**JBR** My pleasure. Thank you very much Dr. Guermazi.

**AG** Thank you.

---

**Comparison of Clinical and Automated Breast Density Measurements: Implications for Risk Prediction and Supplemental Screening**

Radiology 2016; 279:710–719

Kathleen R. Brandt, MD • Christopher G. Scott, MS • Lin Ma, MS • Amir P. Mahmoudzadeh, PhD
Matthew R. Jensen, BS • Dana H. Whaley, MD • Fang Fang Wu, BS • Serghei Malkov, PhD • Carrie B. Hruska, PhD
Aaron D. Norman, MPH • John Heine, PhD • John Shepherd, PhD • V. Shane Pankratz, PhD • Karla Kerlikowske, MD
Celine M. Vachon, PhD

Herbert Y. Kressel, MD Hi. This is Herb Kressel and welcome to the June Radiology podcast. Today I’m joined by Dr. Kathleen Brandt who is Associate Professor of Radiology at the Mayo Clinic School of Medicine, and Dr. Celine Vachon Professor of Epidemiology at the Mayo School of Medicine as well. These two were co-authors of a very provocative study entitled “Comparison of Clinical and Automated Breast Density Measurements: Implications for Risk Prediction and Supplemental Screening.” Welcome Dr. Brandt.

**Kathleen R. Brandt, MD** Thank you.
HYK Welcome Dr. Vachon. Thanks for joining us.

Celine M. Vachon, PhD Thank you. Thanks for having us.

HYK Let’s get started. Historically, the determination of breast density on mammography to help guide screening has been made with BI-RADS, was BI-RADS effective in this role? What were some of the problems with BI-RADS? Dr. Brandt?

KRB Well you’re right. Traditionally mammographic breast density has been assessed visually by radiologists at the time of the mammogram interpretation and it’s assessed into four categories of increasing density based on the BI-RADS lexicon. The top two categories are what we consider dense breasts. Mammographic breast density has become a very important issue. We know that it’s an independent respecter for cancer with the women with the highest breast density having three to five-fold increased risk compared to those with the lowest density and we know that it’s inversely related to the sensitivity of mammography due to masking effect. For these reasons currently, as of this week, 26 states have now passed breast density laws that specify that women must be notified, in writing, of their breast density and ideally they would have supplemental screening tests considered per discussion with their primary healthcare providers. Today breast density is a very important issue for women undergoing mammography.

HYK Why do we actually need automated breast density measurements if we’ve been using BI-RADS? Dr. Brandt?

KRB BI-RADS density categories have been shown to be associated with breast cancer risk, but they’re also very subjective. There’s been many studies that have looked at the inter observer variability in the radiologists assessment of the BI-RADS density and it’s only moderate. So the idea behind the automated method is that now because we have colloquial digital mammography in most institutions we can now apply these automated programs that will automatically assess breast density and reduce the variability.

HYK Dr. Vachon anything you want to add about this?

CMV I just want to mention kind of from the research perspective, we’re also interested in assessing changes in mammographic density due to treatment effects such as tamoxifen. In order to do that the BI-RADS may not be the most appropriate measure and a poor category assessment. Having an automated measure that’s systematically objective and reproducible across centers would be ideal for assessing these types of changes that can inform women’s future efficacy of treatment.

HYK I gather apparently there are now two commercially available products that will do these automated breast density measurements. At least two that you assessed in your study, but I was surprised to read that there were some differences in the way they actually go about determining the metric of breast density. Dr. Brandt what are some of the differences between the two approaches?

KRB The two systems that we evaluated, Quantra and Volpara, do have different ways of calculated breast density. Some of this information is proprietary and we aren’t aware of how they do it. For example how they calculate breast thickness is a proprietary situation that they don’t let us know how they do it. But we do know that Quantra will sum the maximum density measurement for each breast and then give an average; while Volpara will do the breast all together. So they do have different ways of assessing.

HYK And then presumably since they’re commercial products they may alter some of their proprietary software as they get different releases. Is that right Dr. Vachon?

CMV Yes that’s true. In fact the version we used of Quantra for instance at the time of this paper has now been updated. So we’re actually looking at the updated version now. There’s also some subtle differences like including the skin line, not including the skin line, but more importantly I think with this paper and these results is how they assess the BI-RADS-like categories. We call them BI-RADS-like because they are in the clinical radiology perspective, but they have different cut points for how they do that also.

HYK I see. Going back to your study now Dr. Brandt, what was the specific rationale for the study that you reported on?

KRB We compared these two automated methods for assessing density – and we used their BI-RADS-like categories in addition to the dense volume and we compared it to the BI-RADS categories determined by the radiologist and we looked at factors such as risk association with breast cancer and also the number of women that are put into the dense category versus the non-dense category with each method.

HYK Okay and Dr. Vachon could you go into some depth of sort of what did you actually do in your study? What was the method and how did you tackle these questions?

CMV Sure we were fortunate to be collaborating with another individual, Karla Kerlikowske, at UCSF who has done a lot of studies in breast density and together we all have a grant to evaluate these automated measures. What we did was we recruited or through a retrospective study, we obtained about 2000 cases that were new cases. That is newly diagnosed and then matched them to a set of about 4200 controls.

HYK So they were diagnosed with invasive breast cancer. Is that correct?
CMV About 70% were invasive cancer. Thanks for asking. So about 4200 were controlled so unaffected. Importantly in this study what we really wanted to do was get a pre-diagnostic mammogram so we could really look at that prediction or association of breast cancer before the person was actually diagnosed. What we did then was assess – we used the clinically available BI-RADS categorizations from practice. So we did not have one or two radiologists assess all these images, but we did run them all through the Volpara and Quantra software systematically as Dr. Brandt mentioned to assess volumetric percent density, dense volume, and then the categories of BI-RADS by these measures.

HYK What did you find?

CMV We found actually regarding the risk association that is how these measures were associated with risks were very similar. And in fact if you look at Figure 3 from our results, what you can see are all these measures and their association with future breast cancer. Importantly, our take home was that they are as far as their association with risks it is very similar across the BI-RADS category, the clinical BI-RADS category, which is in the top set of four rows in which you see essentially a 2.3 fold increased risk in the highest or densest category relative to the second or somewhat average risk category. And those differences are similar in Volpara BI-RADS categories, the second set of rows, in which you see the highest, extremely dense, about a 1.8 fold increase breast relative to kind of the average for BI-RADS 2 category, and similarly for Quantra you see about a 1.9. They were similar but interestingly even with all the noise inherent across all the radiologists, we still saw a little bit stronger discrimination of breast cancer risks among the clinical BI-RADS density categories. That is this measure could better partition who was going to be a case and who was going to be a control. That was one of our main findings. The other finding was actually how you classify dense breasts. Dr. Brandt alluded to the fact that BI-RADS 3 and 4 is basically a woman who has a dense breast, and so what we found was that although there was some similarity in the majority of women that were dense hovered around 50%, there was about 14% difference actually between the BI-RADS clinical density category and the Volpara and Quantra categories. That is if you were going to say that a woman has dense breasts that measure could change if a woman went to a radiologist that was using Volpara versus a clinical BI-RADS measure versus Quantra. That variability actually is very relevant, but at the same time you might expect this from multiple measures and even within measures.

HYK If I remember correctly Dr. Vachon there was more consistency within machines rather than across machines. Is that correct?

CMV We actually didn’t assess within or across machines, but we can be assured and we did assess concordance of a set of mammograms which we ran on software at multiple sites, and no matter where you run the software if it’s the same version then you’re going to get the same result. And as you know if we take a clinical mammogram and give it to five radiologists that’s not going to be the case.

HYK Okay now Dr. Brandt you didn’t specifically look at the variability in BI-RADS, the critical BI-RADS determinations in this group of patients, you used the existing clinical BI-RADS category, but you did state earlier that it’s basically there’s moderate agreement that’s what you said in the beginning. So taken overall how would you compare this to the level of agreement in the automated methods you evaluated?

KRB I think one of the main points of our study is that if you use the same automated method you’re going to get the same density measurement. It’s reproducible. But if you use a different automated method you will not necessarily get the same density measurement. Between automated methods it’s similar to the variability you see with a radiologist’s visual assessment of the BI-RADS density, but within the same system the variability should be minimal to zero.

HYK I see. Now thinking about sort of the future state if we go more and more to tomosynthesis where you can actually get 3D pixel voxel measurements for attenuation and relative density, do you think this problem will go away in terms of the variability between machines?

KRB Actually I think we’re going to end up with some of the same issues that is depending on vendor and depending on version of software, we will potentially see changes. However, as stressed by my colleague in discussions, density has become so much more relevant to clinical practice now such that there may be an effort by the manufacturers and the vendors to come up with a more systematic assessment of density. So a patient can have a tomo at one place and receive a precise measure of density that’s translatable to another location.

HYK I can’t imagine that they would have approved a CT scan. The FDA if everybody’s water measurements didn’t scale on the same scale we would be really in a lot of trouble with CT.

KRB I think one of the other questions of tomosynthesis is will it do a better job of associating density with risk and because it may be that is more accurate because you have a 3D image, a relatively 3D image, rather than a 2D image. I think that’s a big question with tomosynthesis that remains to be answered.

HYK In thinking about your results of the study, you alluded to this, we’re a very mobile society and patients likely will move from one location to another or another neighborhood and perhaps will have their breasts imaging done at different centers using different methods. In terms of communication with patients in terms of their...
results or communication with referring doctors, what
do you think radiologists should do to make everyone
aware of this issue?

KRB  I think ideally, and hopefully someday, the ven-
dors won’t have as much variability that no matter where
you have your mammogram and no matter what system
is used, you’ll have the same density assessment because
it’s going to be very confusing for patients if they get
a density letter one year and then not the next year.
They’re not going to know what to do. I don’t think we’re
there yet, but hopefully someday and maybe tomosyn-
thesis will be the answer. There will be less variability
between vendors. But I do think that’s probably where
we’re headed where density will be assessed with these
automated methods.

HYK  We ran into something similar with bone density
measurements when we were using a number of differ-
ent approaches and I was at Beth Israel Deaconess at
the time and people would actually note specifically the
machine that it was done on with a little note that this
may not be reproducible on some other system. Do you
think we should be doing that for these measurements
or not necessarily?

KRB  I think maybe just increasing awareness so that
especially if you’re following longitudinal changes in the
same individual over time like say assessing response to
tamoxifen, just for the individual interpreting density to
know that there’s variability in the systems.

HYK  Very good.

CMV  And just to follow up on that, something that we
can do nowadays is bring images on CD or DVD to an-
other location so really if we get to the point where we’re
assessing longitudinal changes at a point that matters for
treatment we can actually have a reassessment of that
image on a different software that’s being used at that
institution.

KRB  That bring up another point though that Volpara
and Quantra that we assess is done with a raw image
information, well if you bring your CD to another institu-
tion, you’re not going to have the raw data. Ideally you’ll
have a system that performs just as well that can work
off the processed image.

HYK  I think this is very thought provoking. I can easily
see a lot of anxious patients kind of dealing with, but
last year I was at this density level and this year I’m that,
nothing happened to me. Anyway I think we’ll be sorting
this out for a few years to come. I want to thank you both
for participating in the podcast. I certainly learned a lot
and I’m sure our listeners and viewers will as well. Thank
you very much.

CMV  Thank you.

KRB  Thank you for inviting us.

HYK  Bye-bye.