

Episode: Ethics Talk Podcast Transcript – Managing Health Care AI “Megarisks”

Guests: John Banja, MD

Host: Tim Hoff

Transcript by: Tom Wagner

Listen to the podcast [here](#)

Tim Hoff (Host): Welcome to *Ethics Talk*, the *AMA Journal of Ethics* podcast on ethics and health and healthcare. I'm your host, Tim Hoff. Risk managers who work in health systems don't like not knowing things. Their success depends on their keen understandings of clinical operations, patients' needs, third party payers' interests and more. Being able to predict how these stake holder's interests intersect to generate risk for an organization is key. Artificial intelligence has been applied to many problems in the health care sector, largely in the ways in that it uses machine learning to make predictions based on the massive amounts of data generated by health systems. But within risk management the optimism around AI is more measured. AI applications in risk management might change how risk is identified but the applications themselves also come with their own risks. Because AI tends to magnify and enlarge risks already present in health care organizations, these AI-induced risks to health, patient privacy, and more have the potential to be even more damaging than ever. These AI "mega-risks" are considered in an article from this month's issue of the *journal*. Dr John Banja wrote that article and he joins us this week to discuss the promises and perils of AI and risk management. Dr Banja is a professor and medical ethicist at Emory University and he is editor of the *American Journal of Bioethics Neuroscience*. His most recent book is *Patient Safety Ethics*.

Dr Banja: thank you very much for being here.

John Banja: My pleasure, Tim, thanks for having me.

HOFF: Many people see risk management as more closely related to health care administration rather than clinical patient care. Many folks, including health profession students, don't really understand what risk managers do. So to begin, can you tell us a little bit about where risk managers fit?

BANJA: Sure. Let me just refer to your question though where you said risk management is seen by many as health care administration. I suppose that that is true, I think it may have been especially true 30 40 years ago. But I will tell you that in the 20 years that I've been fairly close to risk management and patient safety, I have seen a trend toward the risk manager being thought of as not only the patient advocate but rather his or her functions as being folded in, or at a one, with quality improvement quality assurance. So the bottom line basically is that we're all in this to advance the welfare of our patients, to be advocates for our patients, and to create a reasonably safe environment for patients to receive care. That's what risk managers do.

HOFF: Great, thank you. It may be that the public perception of what risk managers do and even the knowledge of their existence in health care teams is a little different than the perception within the medical community itself.

BANJA: Yeah, so excuse me I really didn't answer the second part of your question [laughs] – what do risk managers do? So what risk managers do, not surprisingly, is they manage risk. When you talk about risk a lot of people just immediately leap to, "Oh it's about harm isn't it?" Well, it's really about how frequently or what is the probability of a harm happening or a harm materializing? So the risk, the "risk" of risk management, really entails at least two dimensions: one is how likely is it that a bad thing is going to happen and the second one is well if this bad thing happens, how bad is it? Risk managers have to be wary about that. Obviously they're going to be very, very concerned about high probability, high gravity, high harm events, those are going to be at the top of their list. So things like medication errors, things like diagnostic errors, errors in general which, presumably, we can prevent, preventable risks, those are the kinds of things risk managers are going to be especially sensitive to.

HOFF: Sure, given that there are risks present at essential every level of the health care interaction, at what point are risk managers actually brought in because obviously they're not integrated into primary care, go to your doctor's office and you talk to your physician they prescribe you something, you generally don't interact with a risk manager in that instance. At what level does the risk need to be for a risk manager to get involved, I guess is the question?

BANJA: Well interestingly enough, a harm doesn't need to happen for a risk manager to get involved. In fact, most errors that occur in health care do not result in harm to patients. But that doesn't mean that a risk manager isn't going to get involved. For example, a patient might get the wrong medication at 10:00 this morning, pharmacy has sent up the wrong medication, and the nurse didn't check it the patient receives the medication and nothing bad happens. I will tell you that risk managers going to be all over that one though. Because what we would call a harmless hit this morning could be a fatal medication error this afternoon. So the risk manager is going to ask herself questions – well how in the world did this happen? He or she is most likely going to do something like a "root cause analysis," although that's a bad name because it just indicates or connotes that there's only *one thing* that went wrong in all this. Well as a matter of fact what we know is that when disasters or catastrophes where nasty events that happen in a hospital that are preventable, almost invariably it requires multiple people making multiple mistakes for that medicine, that wrong medicine, to get to the patient. Essentially that's what that risk manager is doing. He or she is looking at all those variables those factors that enabled that error to happen and then is going to go back to the drawing board and say, "How can I make this more difficult for a future error to happen. How can I make future errors more difficult?"

HOFF: What would be some sort of examples of the kind of steps that a risk manager would suggest, would it be changes to default settings in EHR's, things like that – what comes from that deliberative process of identifying risks?

BANJA: Right, so what's fascinating about it and what's fascinated me for 20 years, is that mal-occurrences in health care are very contextual. So, for instance what you have initiated . . . let's say we're having a lot of patient falls in our hospital. And we've discovered that some of them are attributable to our staff not using gait belts when they walk with patients. We had one incident last week where someone forgot to put up the bedrails on a patient's bed, and the patient tumbled out and fell. My point is what might work for a fall reduction program in a hospital may very well have no relevance whatsoever to ventilator-acquired pneumonias or diagnostic errors or the fact that our nurses are failing to have face-to-face communications with one another when one nurse is going off a shift and another nurse is coming on a shift so these communication kinds of errors. So what we're talking about here is very granular kind of particular type events each of which may require a different remedy.

HOFF: So as you just mentioned risk management has several key roles in identifying potential safety problems even before they arise, and your article in this month's issue considers how artificial intelligence applications might be used in health care to identify these patterns from the large data sets that health systems accrue over time that can help risk managers mitigate safety risks earlier. A critical point of your article is that these AI applications themselves pose key risks. Can you elaborate what some of these risks might be?

BANJA: Sure. You know in 2018, 2019, we were reading a research paper almost every week that was touting some new model that a bunch of computer scientists had worked years on in terms of better diagnosing, better identifying breast lesions, pneumonia lungs, brain cancers . . . you name it, if it can be imaged. One of the things that AI models are very very good at is image identification. 2018, 2019 we were let's just say getting this rush of articles, and every now and then there would be this prediction that this model is going to replace a radiologist, a pathologist, a dermatologist, something like that. Tim, I will tell you that 2020 is seeing an interesting reversal on that, that these models - which as a matter of fact in their test environments, where they were developed, function pretty well that is to say they function just about as well as a board-certified dermatologist or radiologist did - when we would take that model out of that environment and then use it in a different hospital system and provide the model with 10,000 mammograms and ask the model, "All right, which ones are cancerous which ones are not," the model did not do well.

What we are finding out right now, this is my first answer to your question, what we're finding right out now is the accuracy of a lot of these models is not anywhere near where we would hope it would be. Another way of saying this is the models don't generalize very well. They might work very, very well within the Emory health care system, but they don't work well in the Stamford health care system or the Massachusetts General health care system or the University of Chicago health care system. So, we're in the process, then, of working out those kinks right now, and by the way, it might take years to work this out. So, I think that accuracy is at the very top of the list here. But I'll also say that there is literature that suggests that hospitals and clinics some of them are purchasing these new models. They haven't been FDA approved we really don't know the quality of the model, but they're using them not only

for image recognition technology but for things like predicting complications or predicting readmission to the hospital. "Doctor, if you discharge John Banja tomorrow, there is an 83% chance that you are going to have to readmit him the following week" They're using these models, and frankly, we don't know how good they are, we don't know how accurate they are.

I'll tell you one more that I think is a very serious one, and it's the fact that these models use voluminous stores of data. So there's a whole area of AI ethics that concentrates on what we call "big data" because that's the way you educate a model that's the way you educate an algorithm. You have to give it a 100,000 slides or images of breast cancers or whatever in order for the model to start identifying, diagnosing, distinguishing a cancerous lesion from a noncancerous lesion. That's one of the things the model can do. Another thing that maybe a model can do is they do natural language processing. There was a study in I believe it was Mount Sinai in New York not too long ago called Deep Patient where the model was fed millions of patient records as to see how good a diagnostic model we would get. The model knew everything about hundreds of thousands of patients: their medication history, their medical history, their family history, all of that, comorbidities, age, ethnicity, all of that kind of stuff. And the model got pretty good at saying well with if the patient presents with A, B, C, D, and E, very likely, and they always give a probability estimate, so there's an 83% chance the primary problem that this patient had was X and a 19% chance that he's also got Y and Z. Here's the problem, the problem is that these data streams, these huge amounts of data, these data banks, you can reuse those data banks for purposes other than patient care. That is to say, hospitals right now are being approached by data brokerage firms who are saying, "We would like to buy 100,000 of your brain scans, brain images, and we will pay you for these." For some clinics some hospitals that might be a very handsome revenue stream for them. But once you sell that data to someone else, and by the way, the data is going to be de-identified presumably that data is not going to be traced back to John Banja or Tim Hoff, but once you do that you really don't know - unless you have a contractual understanding with that data brokerage firm - you really don't know how they're going to use that data. They may just sell it to somebody else who may use it in ways that maybe you don't want it to be use. Maybe that somebody else wants to identify - I'll use some sexual intimate examples because they're hot button kinds of issues - maybe that data will be used to identify women who've had abortions or men who have erectile dysfunction or persons with a history with mental illness, in other words very intimate stuff that you would not want every Tom, Dick, or Harry to know.

And by the way one last thing, cyber hacks. There are voluminous data stores and consequently - the famous story about the bank robber Willy Sutton, Willy was asked, "Why do you rob banks," and Willy said it's because that's where the money is. Well if you're a cyber hacker and you're looking for data that is valuable, I mean, here's the target, here's the bank, this is where you would want to go. And so many very prominent hospitals have had these cyber hack incidences occur and of course the individual is looking for money. "I'll release the data back to you if you pay me...", however much money the hacker is worried about. Those are the kinds of things that, I think it's going to be very interesting to see if risk managers who traditionally have

looked at things like, I say, diagnostic errors, medication errors, those kinds of things, are they going to get involved with these risks from importing lots of AI technologies.

HOFF: I expected you to focus more on the sort of privacy angle of these massive data sets so I was interested to hear the concern about these private data firms buying up these data sets and using them in ways that either the individuals who are contained in the data sets or the hospitals themselves might not like. Do you have any examples of that happening already? Because I mean these data sets exist already and these massive accounts of patient information and things like that are already being collated.

BANJA: Actually there are some examples of this happening in the private sector where women who are pregnant have gotten advertisements of the, "Congratulations – we understand that you are pregnant! We're having a sale in our store of these items that your new baby may need or use." And these women are astonished because they didn't tell anyone about their pregnancy, and they wonder how in the world did this company find out about it. A lot of these applications, the really worrisome part of a lot of these applications are occurring outside of health care. So essentially businesses are using this data – guess what, to sell their products better. I'll give you a brief example then I'll come back to health care.

If I happen to know that in the pre-COVID days or the post-COVID days, when you got in your car and you drove off to work, that you went past my department store and you went past it twice a day, well, you're a prime example then of a person who I would like to send advertisements to announcing a sale or a special or something like that. I could use GPS data that we get from you to actually track your route, and by the way, this is even going to become more obvious when we have self-driving cars or autonomous vehicles when we know the route that these vehicles take.

In health care interestingly enough one of the big problems in companies finding out about patient data is the data itself, I said to you a few minutes ago that this data is de-identified. Often times there are glitches in that de-identification. There are some lawsuits out there right now where hospital and health care facilities have shared their data - they've shared it, not sold it now - but shared it with entities and the entities have looked through, again tons of data, and said, "You know what, a lot of this data is not scrubbed properly. There are patient identifiers in this data. We know the doctor of this particular individual. Here's a scan of a person's chest, we can see his or her jewelry on this image." It's a lot harder to de-identify data than you think it is.

So these are the kind of things, and from an ethics perspective too, we're in the infancy of figuring out how should we approach patients in terms of their consenting to the use of their data, in fact in ways that we can't even predict? Right now this data presumably is de-identified, and when it's de-identified, you don't need HIPAA protections anymore. All you need is the patient to consent to the use of his or her data, and then the hospital can de-identify and then do whatever they want with it. We're starting to wonder now, we're especially looking at things like the European Union's General Data Protection Regulation, the GDPR, the California Consumer Privacy Act, that take a much much more stringent patient protected kind of approach to this data. They're

looking to really ramp up patient consent to this data. Essentially do patient know what they're consenting to when they allow East Cupcake Hospital to use their data, even if it's de-identified.

HOFF: Earlier you were talking about the influx of AI technologies specifically related to imaging, and how there was this great promise about how it would change the field forever and put radiologists out of a job and everything like that. Obviously that has not happened in the way that it was expected to. Going forward, what sort of criteria do we use to judge which AI applications could actually have this kind of amazing promise for health care and which are over-hyped – how do you determine where it would be most applicable?

BAMJA: You know I'm kind of chuckling right now as I'm starting to answer your question because I think most of the AI in health care today right now is over-hyped.

HOFF: [laughs] That's probably true.

BANJA: I think an answer to your question, we are getting reality check here in 2020. As I talk to health care professionals and they say you know I got a new algorithm to try out to test out last week, and frankly, I'm quoting one, she said "It absolutely sucked." I think an answer to your question, we are in the trial and error stage of this new technology. Just like a drug might be in phase two or phase three trials. We're figuring out on whom does it work. What are the glitches in the system that we may have to attend to? On whom does it not work? Just as phase two, phase three trials are continuing to tryout that drug, and even if the drug is FDA approved, we like to say in research ethics, it then goes into phase four trials, right? The drug may have been approved, may have been test on ten or twenty or thirty thousand patient subjects, patient participants. Now it's going to be tested with ten or twenty or thirty million people, and now we're really going to see how well this works. Well I think that's where we're at right now with AI.

I think the best answer to your question is though, if these technologies continue to show a lot of promise the primary criterion that we're going to use compare the quality of these technologies is Dr Jones, nurse Smith. What is their error rate, how accurate are they? And if this technology shows, again after a considerable testing period, that it is in fact as good as Dr Jones or nurse Smith. Well then I think hospital are going to be very very inclined to want to purchase these technologies. And that Tim is going to usher in a whole new era of health care. Because remember these technologies, you don't pay them vacation leave, you don't pay them paternity leave, they don't go on vacation, they work 24-7. It will be interesting to see how the complexion, how the landscape, how the work flow, how the staffing ratios change when these new technologies emerge on the scene. I don't think that's going to happen though for a good, at least, for a good five to ten years and maybe more like 20 or 30.

HOFF: Some hospitals and clinics are eager to offer patients who are willing to be research subjects all of these latest cutting-edge technologies, human device implantation, for example. And these things, like you were mentioning, might not be FDA approved or might be approved differently than the use for which they're being

prescribed, things like that. For devices that don't have an existing risk profile, how can AI help risk managers and clinicians try to estimate risk?

BANJA: I think that the greatest promise of these technologies is going to be, as I said earlier, making error harder to happen. These technologies for example don't fatigue, and you don't want to be the last patient of the day that Dr. Smith is seeing. I think where the great gains in these technologies are going to be is, for example, reminding health care professionals that, "You know what you ordered a mammogram on Mrs. Jones it has come into the office you need to look at it now." And probably a thousand other kinds of reminders; a help to radiologists for example in terms of, "Doctor, these particular scans look very suspicious and you need to look at them. These scans however look absolutely clean and perhaps you can just skim over them." So it's time savers, those kinds of things. Of course, the AI, when my grandchildren thirty, forty, fifty years from now going to a primary care provider I predict that what's going to happen is they're going to put their health card into a computer, just as you and I put our credit card into a gas pump, and that credit card's going to have all of our medical data, all of our medication history, all of our health history, all of our DNA on that card, and the machine will read it and perhaps come up with a better diagnosis and treatment plan than a board-certified primary care provider can today. I think that's decades in the future, but I also think it's inevitable. It's going to happen, these technologies are just going to get better and better and better and better, but not any time soon.

HOFF: Dr John Banja, thank you very much for joining me this week and sharing your expertise.

BANJA: My pleasure, Tim, thank you very much.

HOFF: That's our episode for the month. Thanks to Dr John Banja for joining us. Music was by the Blue Dot Sessions. For more on risk management ethics visit JournalofEthics.org to read this month's issue of the journal. Follow us on [Twitter @journalofethics](https://twitter.com/journalofethics) for all of our latest news and updates. And we'll be back with you next month for an episode on brain death. Talk to you then.