

Transcript Details

This is a transcript of an educational program. Details about the program and additional media formats for the program are accessible by visiting: <https://reachmd.com/programs/the-convergence/raja-elie-abdulnour-nejm-clinician/39936/>

ReachMD

www.reachmd.com
info@reachmd.com
(866) 423-7849

"We have tools that can think alongside us and free us from a lot of burdens," says Dr. Raja-Elie Abdulnour from NEJM Clinician

Dr. McDonough:

Welcome to *The Convergence* on ReachMD, where innovators and physicians explore the technology transforming medicine. I'm Dr. Brian McDonough, and today, we're taking an honest look at the use of artificial intelligence in medicine—what it's doing well, where it's falling short, and what all of this means for clinicians right now.

And joining me to talk about this is Dr. Raja-Elie Abdulnour, Editor-in-Chief of the *New England Journal of Medicine* Clinician, and Assistant Professor of Medicine at Harvard Medical School.

Raja-Elie, welcome to the program.

Dr. Abdulnour:

Brian, it's great to be here.

Dr. McDonough:

You are in the ICU at Brigham while also building AI tools for the *New England Journal of Medicine*. You see this from both sides. Candidly, what misconceptions do clinicians believe about AI?

Dr. Abdulnour:

Because of my many hats—as an editor, a publisher, and a clinician, but also as an AI researcher—I have this vantage point where I get to see different perspectives. I get to see techno-optimists who think AI will be super good for humanity and really help us, and others who are highly skeptical. I think that the right stance is somewhere in the middle.

If we give utmost trust to these devices, this is where we risk automation bias. This is where we risk getting in trouble by believing a false recommendation. But then, on the other hand, if we don't trust them enough, we run the risk of missing the opportunity of taking advantage of their benefit. We don't want to be in either camp, right? So, this is where you make mistakes and, on the other hand, where you miss an opportunity to make a difference in the patient's care. I think that we need to be somewhere in the middle where we use them—when we think they're effective, we use them cautiously, always believing that there's always a chance of error.

Dr. McDonough:

Raja-Elie, could you share a little bit about your background and maybe go through a typical day? Clearly, you really have an interesting life.

Dr. Abdulnour:

Sure. So, at the *New England Journal of Medicine*, I've been making us all think critically about how artificial intelligence can help us as a journal and help us as clinicians, and how we can help artificial intelligence. So, I work with AI all the time, and in particular, large language models—all the chatbots, ChatGPT, Claude, OpenEvidence, and Gemini. I use them in many different tasks, both in trying them out, but also using them in my workday.

When I go to the ICU, I've incorporated them more and more. At the Brigham and Women's Hospital, where I practice, our intensive care unit has not deployed, for example, AI scribes, which have been deployed in many other hospital systems and clinical contexts, but not in the intensive care unit. So, what I do when I round with the residents is every time we talk about a patient, I pull up my phone, I bring up a large language model, and I engage with the team on how to use it and also show them where it can help us on some of our blind sides.

Dr. McDonough:

When you're starting a shift in the ICU, where does AI show up in your actual workflow today? Maybe you could walk us through it.

Dr. Abdunour:

I'm going to give you an example. It's a Monday morning in the ICU, and a patient arrives. No personal identifiers—it's a real patient. A lady in her 40s with advanced cancer—leukemia—is getting prepped for getting her treatment, and she needs to get an antibiotic desensitization. She goes to the allergy clinic, gets a pill, and she does okay. Half an hour later, she's in emergency room with shock and fever and an elevated white count.

Because of her background, everybody's thinking septic shock. And the allergist is saying, "Well, I don't think this is anaphylaxis. It doesn't look like anaphylaxis at all. There's no rashes or anything," The patient is admitted and put on a whole bunch of antibiotics. Everyone's puzzled. I'm puzzled. The ID docs are puzzled. The emergency doc is puzzled. The oncologist is puzzled. So, we treat her like we would treat anybody: a lot of pressors and a lot of antibiotics, even though she's allergic to many of them. And the patient's telling us, "Please don't give me antibiotics. These are not good for me." And I was telling her, "Well, unfortunately, we cannot rule out septic shock. It's too early."

And so then, I pulled up OpenEvidence. And again, another disclaimer is that the *New England Journal of Medicine* has an agreement with OpenEvidence on content. I personally have no stake in OpenEvidence. I don't get anything from them, but I use it, like more than 40 percent of US physicians do. So, I take OpenEvidence, and I say, "Hey, this is a patient with leukemia. She is going to undergo CAR T-cell. She has allergies to medications, and a few weeks ago, she got an immune checkpoint inhibitor. What could be happening? Could this be an immune checkpoint inhibitor-mediated augmentation of an allergy? Something that's not truly described?" And OpenEvidence says, "Yes, it could be," provides me a long explanation, and points me to a case report that's almost identical to this case.

I shared the information with the team, and 20 minutes later, we all agreed to completely change her care. And instead of treating septic shock, we're treating cytokine release syndrome. We're giving a bunch of immunosuppressants, and a few days later, she's out of the ICU.

That's an example where it changed the direction of this patient's care. All the top consultants in the field were baffled. There was no way I could find this case report in our more traditional resources. It just saved the day. So, that's an example where it makes the case that these tools are going to be extremely helpful.

Dr. McDonough:

So, obviously, you practice at a resource-rich academic center. What does the introduction of AI look like in community hospitals and rural practices? I already see it in my own practice, and I'm in a community hospital where I use OpenEvidence as well. So, I'm using that tool, and many of us—as you say, 40 percent—know about it, but are we at risk of creating a two-tiered system? Or is AI going to help things out?

Dr. Abdunour:

When we saw ChatGPT appear on our phones and our browsers for free, and now we have OpenEvidence being free, it would be foolish not to say that there's a good chance at democratizing access to knowledge and access to an intelligent copilot. It's a real opportunity. And for many health systems, whether in the states or outside the states, access to some of these expensive tools, expensive EHRs, and expensive online references was a real blocker, right? But now, these tools have made it very easy. Anybody could get these tools on their phone and start using them.

But in that democratization is a real risk—the risk being that a lot of these tools, many of which I've mentioned, are being used by clinicians and by patients outside of any oversight or supervision by other stakeholders, including the health systems themselves, who have really a true benefit in making sure that safe patient care is delivered. So, a lot of these tools are being used with very little data or research to show that they're fundamentally safe.

In fact, there's many recent studies, including a very recent one, where they've fabricated 300 cases, and they created fake tests—for example, a neurostatin as a blood test. And then, they ran these 300 cases by most large language models and asked, and the question was, would they pick up on the fabrication? In 80 percent of the cases, the AI just went with it. They assumed this was correct and answered the question as if this was a true test. These tools can be dangerous, so there needs to be careful oversight so that mistakes that can have real implications on patients don't happen.

Dr. McDonough:

So, Raja-Elie, if you take that concept, the framework that AI is wrong until proven otherwise, how should clinicians navigate that mindset while engaging with these tools?

Dr. Abdulnour:

Here's my take. Even before AI, I've always worked in an academic institution working with residents and fellows and medical students. And I've always used a sense of trust when working with a trainee, or even with faculty, but let's focus on trainees for now.

As you start working with a trainee, you can start getting a sense of how much you trust them. And, in fact, entrustability is a very well-used indicator of progress and development by educators, right? They use these EPAs—entrustable professional activities. And then, when I start working with a trainee, some trainees I find extremely trustworthy, and I have faith in them, and I can leave the ICU and have them run the unit. Others aren't trustworthy enough, and I need to be supervising everything that they do. I approach these tools the same way.

If I get to the point where an AI tool is consistently trustworthy in a particular task, I actually would delegate a lot of the work to it—for example, drafting emails to my friends, or giving me a recipe for a cocktail, right? Low stakes and really effective. On the other hand, if I don't have any idea if a particular tool is trustworthy, then I won't choose it. And I think in high-stakes situations, like in the clinical environment, the bar for trustworthiness is very high.

I'm assuming that whenever I'm working with an AI tool, it's actually a very good resident. I need to be in the unit. I need to provide some supervision depending on the situation. But then, at times, I need to jump in. I can't leave the ICU. And so, it behooves us as clinicians that when we are using these tools, the higher the stakes, the more supervision, and the lower the stakes, the less supervision.

Dr. McDonough:

You mentioned trust. You brought me back to freshman year of medical school, when I asked a lab partner what a blood vessel was on the cadaver, and he told me and I trusted him. It came up on the test, and I identified it as such, and it was wrong. Later, I talked to him, and he said, "Oh, I thought that's what it was." From then on, the trust was gone.

Of course, the stakes get higher with patient care, and I think you bring up a really important issue for all of us, and that is trust. So, how do you personally decide when to trust AI output versus your own clinical judgment in the hospital setting? As you said, not just getting a cocktail recipe, but when it really matters for patient care? And how do you do it?

Dr. Abdulnour:

That's a great question. First of all, it all starts with how much I trust myself, right? So, going back to this patient I was describing, a case I've very rarely seen before, do I trust myself in knowing exactly what's going on? No. And that actually is the very first step that clinicians must take on, which is admitting that they don't know, right? Having this humility of saying, "I don't know," and therefore, asking for help.

And then, when you're asking for help, whether it's asking a consultant, a resident, or an AI, there are three determinants to trust. One is effectiveness. This other entity I'm going to be interacting with, how effective are they at the task that I'm going to be engaging in with them? So, for example, if I have a question about a rare infectious disease, I'm going to go for an infectious disease specialist consultant, not a pharmacist.

The second is transparency. If I'm asking someone for advice, they need to be able to back their advice with a reason, by some rationale. They need to explain to me where their answer comes from. So that's transparency.

And the third one, which is probably the most important one, is benevolence. For me to trust someone, I need to know that at some level, they have my best interest in mind.

And so, when you put these three together, the most trustworthy agents out there, whether they're human agents or AI agents, are effective and transparent, and have my best interest in mind. Think about the friend you mentioned in school—same thing with the AI tool. So, when I'm interacting with an AI tool, I need to make sure that it's effective, transparent and benevolent. And then based on the sum of all three, I make a judgment on how to trust it.

When we're talking about AI tools in particular, effectiveness can be measured by data. Is there evidence out there? Has this tool been benchmarked? Has it been evaluated? Has it been compared to other tools? And we've all seen the literature studies that looked at how good ChatGPT is at answering USMLE questions, taking on clinical problems, et cetera. So, these are markers of effectiveness, and these tools are indeed effective.

Dr. McDonough:

When you look at the difference between AI as a crutch or a tool to help you, it's such a fine line.

Dr. Abdulnour:

It is a fine line. In the literature, it's been described as a mode—what's going to be my mode as a user, like my brain? There are two

sorts of user behaviors. One is a cyborg, and the other one is a centaur.

I am in a cyborg mode when I'm using AI for a low-stakes task, where I know I'm good enough at it that I can quickly detect an issue. And so, for example, when using ChatGPT to draft an email, I'm a cyborg. I just put in some prompts, and it gives me the output. I take the output, make some tweaks, and send it away. I'm a cyborg.

In high-stakes situations where I don't know for sure that the tool is 100 percent trustworthy, I'm a centaur. I will take on most of the tasks, especially decision making, myself, but I will use it as the resident. "Hey, can you confirm this thing for me? Can you look up this evidence for me?" But ultimately, I'm the decision maker.

And so, shifting between cyborg and centaur based on the task is really what makes for a good AI user—this ability to go back and forth between using AI, embedding it in your mind, and replacing you in some ways versus using it to help you out in situations, but you keep the decision making to yourself.

Dr. McDonough:

I know you've heard people say AI will replace doctors. What is your response as someone who's both using and building these tools and educating the rest of us?

Dr. Abdulnour:

Whether it's AI or not, technology has replaced many human tasks, and to good effect, by liberating our bandwidth to do more things, right? So, think about any tool, whether it's a calculator, a word processor, a spell-checker, et cetera.

And I don't think things are different here. I think if we look at all the tasks that clinicians do—taking a history, writing a note, obtaining prior authorization—all of these are tasks that, if they can be delegated and made easier, we should welcome that support with open arms.

In my mind, the key thing that will prevent human clinicians from being replaced is the issue of benevolence. It's going to be very hard for humans to put their lives in the hands of a machine if they don't know that a human is overseeing it—a human who has their best interest in mind.

Dr. McDonough:

Both of us are involved in residents' education and training young doctors, and I'm sure many of our listeners are as well. So, let's say you're supervising some residents, and they've never practiced without AI tools. That's going to become more and more common. What are they doing right with AI that more experienced clinicians could learn from them?

Dr. Abdulnour:

Good question. I think the one thing that more experienced, seasoned clinicians can learn from trainees is using AI. The trainees are just more open to using any tool. They're in learning mode all the time. Everything is new to them, including these tools. And so, they are more readily open to using them, experimenting with them, and learning from them. And I think this is critical. This avoids the risk of a missed opportunity—in other words, the risk of not choosing a tool that can actually benefit you and the patient. I say it all the time: we have to be using these tools. We need to learn how to use these tools. We need to figure out what works and what doesn't, and when to use it and when not to.

Now, on the flip side, the biggest risk for trainees in using the tool is that they may not have the required skills to detect when a tool is giving them a false judgment or not. And they run the biggest risk of automation bias and taking an AI's judgment as truth even though it's incorrect. And this is where experience and knowledge matters. We've seen this before; this is where the student becomes the teacher, and the teacher becomes the student, and it becomes this partnership between the more seasoned educator and the budding trainee where together, they need to learn from each other on how to use these tools effectively.

There are different risks depending on the level of expertise, but for trainees, the biggest risk is never upskilling—in other words, not learning an important skill. One example is writing a note. I can tell you that when I reasoned through a patient, I almost have the template of the H&P that I had in residency where it had the HPI, medication, surgical history, assessment, and plan. And that's almost how I think. Now, with AI scribes, I'm actually fairly concerned that if a medical student never learns to write a note and is using AI scribes all the time, they will not learn how to organize their knowledge. I may be wrong; we just don't know. So, this is where there's a significant risk of never-skilling.

Dr. McDonough:

That is what a great point. You're right, we were all trained the same way. We pretty much wrote our notes. We did a SOAP note, and everybody did it the same way. And you went back and did it the same way every time, almost like reading an x-ray; you just went one

step to the next, skip things. I do see residents coming in, and they're using their phone because the EMR is on their phone, and they're scrolling, and I'll say, "Any allergies?" And they're flipping up and down looking, and I go, "Well, isn't that kind of in your head?" We used to know that walking into the room, and they now don't have to do that.

So, for attending physicians listening, can you provide a practical framework for how they should supervise trainees using AI? What suggestions do you have?

Dr. Abdunour:

I co-authored a review article on educational strategies in supervising the use of AI. And the first thing is to recognize a moment where a human-AI interaction is happening and where there's a potential risk for never-skilling, a risk for de-skilling, or a risk for mis-skilling, which is learning the wrong skills.

So, the first thing is to recognize it. It's a bit like sepsis; if you think it's sepsis, call it sepsis once you recognize it. There are several frameworks that educators should use, all of which are centered on the same thing, which is to promote critical thinking. If there is one skill that we should teach more, not less of today, it's critical thinking.

Now, how do you promote critical thinking? There are a few frameworks out there, and some have been validated. One is DEFT, which stands for diagnosis, evidence, feedback and teaching. And it's a framework where for each one of these, you ask the question to the trainee. These are frameworks that were developed for clinical reasoning and diagnostic reasoning. These same frameworks could be applied for the use of AI, which is why, in the paper I wrote, we called it DEFT-AI.

So, for diagnosis, what I would do is, for example, if a student says, "This patient has a status asthmaticus," and I know that they've used ChatGPT to give them the answer, I say, "How did you get to this diagnosis? Did you use an AI? What AI did you use?" I want them to commit to surfacing, to make it explicit, what their assessment is, which AI they use, and how they use it. Surfacing this is really important for everyone to realize that we've used an AI tool.

A quick tangent here. Again, I was in the ICU, and a student was presenting a case to me. I asked the student, "Should we give clindamycin to this patient who may have a parapharyngeal abscess?" And the student is thinking out loud, and then I say, "Hey, let's ask OpenEvidence." I asked OpenEvidence, and then I showed the answer to the student, and the student read the answer and assumed it's correct. This is a critical mistake because, again, never assume it's correct. And so this is where I paused. I realize there's a risk here, and then I engage with DEFT-AI with the student. So this is diagnosis.

The second one is evidence, where I say, "Okay, you've used ChatGPT. Is there evidence that ChatGPT is good for diagnosis? What is that evidence? Is there evidence of risk? Is there evidence of bias?" And depending on what I hear from them, I know whether they used it blindly or they were familiar with the tool.

Then, there's "F" for feedback, which is asking them, "How would you assess your use of AI? Do you think you used the right prompts? Do you think you were critical? Do you think you should have been critical or less critical?" So, again, open-ended questions engage the student in self-reflection.

And at the end, "T" is teaching, where based on everything I've learned, I give a few minutes of teaching or more. Often, I've realized that even a resident, a super savvy and super digital native, has no idea how these tools work. So, a lot of my teaching is spent on how these models work and why they shouldn't be trusted.

And then the AI at the end is a recommendation, meaning, "If I recommend the use of AI in the future, this is how you should be using AI, you should be using this AI, not that AI," and so on and so forth. So, DEFT-AI, which is in the paper we've authored, is a simple question-based framework. It takes a few minutes to do at the bedside. It works like a charm. It's been validated in clinical reasoning, not in AI, but it's definitely useful.

Dr. McDonough:

For those just tuning in, you're listening to *The Convergence* on ReachMD. I'm Dr. Brian McDonough, and I'm speaking with Dr. Raja-Elie Abdunour.

We have a very large clinician audience, and there are some tough questions I'm certain they would like to ask. First of all, liability—if you use AI and miss something, are you more or less legally exposed? And what do clinicians need to know?

Dr. Abdunour:

The question of liability, when it comes to the use of AI, is an open question. The greatest minds in the nation are thinking about it, and it's unsettled. But fundamentally, in a recent conversation I've had with a lawyer on the topic, when it comes to the liability of the physician, it's whether or not the physician breached the standard of care.

So, for example, it's a quirk in the tort law where if a physician accepted a false recommendation by the AI and that led to harm, if that action, even though it was not appropriate for this patient, was within the standard of care, then they're not liable. At the end of the day, it just comes down to whether or not that action breached the standard of care, whether they use AI or not.

So, what does it mean for us as clinicians? I think about this all the time on rounds—what is the standard of care for that particular patient? And then I can decide whether or not I need to individualize my care for that particular patient. Sometimes, individualizing care may not follow the standard of care. You may want to use a drug off label, use a different dose, or do an intervention or a test that is outside the guidelines but where you think, as a clinician, it's what's best for this patient, right?

So, we often take these considerations and make decisions that may be outside the standard of care. That's the bottom line—whether or not what we're doing is a breach or not of the standard of care. The use of AI is actually irrelevant here.

Dr. McDonough:

So, we know documentation tools are everywhere. We're more or less drowning in them now. From what you've seen, which ones actually work, and which ones are really just adding noise?

Dr. Abdulnour:

That's a great question. There's been a number of studies now, including recent randomized controlled trials, where they've randomized clinicians to using a scribe versus not. And they looked at different measures of wellness and measures of efficiency in an EHR, and then all of these studies use different scribes.

To my knowledge, no one scribe has been shown to be better than the other. But what I've seen consistently are two things. One: at best, there is a slight improvement in efficiency, meaning clinicians are spending maybe one or two minutes less in EHR per patient. Now, if you have 50 patients a day, that can amount to a significant saving of time. The gains are maybe statistically significant in these studies, but whether or not they're clinically significant in the real world is unsure.

But the most striking findings, and these are showing up repeatedly in these trials, are the impact on wellness. So, when you ask clinicians through surveys, and if you ask them in many different ways, "Has your quality of life improved by using these scribes?" The answer is a resounding "yes." So, even though it's not necessarily helping their time, it's impacting their wellness and their sense of burnout in a very positive way.

Dr. McDonough:

Now, I helped lead efforts to get EMRs into our hospitals, and for a lot of physicians, one of their biggest complaints was they just couldn't stand the alerts and all the different things that slowed them up. Does AI reduce this cognitive load or add to it from your perspective? Especially things like the inbox?

Dr. Abdulnour:

Great question. I think it depends on the tool and the AI. If you go back to the scribe, one of the benefits is a drastic reduction in cognitive load because you're no longer both typing and looking at the computer, listening to the patient, and scribbling notes, worrying about missing something. By having a scribe that is transcribing the whole thing, it really opens up your mental bandwidth on just focusing on the patient in front of you, and then you leave the editing of the note for later. So, that's an example where cognitive load is decreased, which is, I think, why wellness is increased in these studies.

With EHR alerts, the way that they're currently designed is a huge problem. And if we start developing AI tools that will keep making more alerts, there's definitely going to be alert fatigue. The opportunity there is to use intelligence systems that can adapt their threshold to alert a clinician based on the individual case, meaning not everyone whose heart rate is above 100 and who has a white count above 12 needs to run an alert for sepsis, for example. An intelligence system may spend a little bit more time and gather a bit more data, and then increase specificity when it comes to alerts without knocking down its sensitivity.

So, it all depends on the implementation. And Brian, it's a great point. Sure, these tools are fantastic, but if they're not implemented well, they can actually make things worse, not better.

Dr. McDonough:

In my experience, when electronic records were starting out, the vendors would come in, and they promised the world—"You can do this, you can do that." But it would be a lot of money following many of those things. If you're a chair or practice leader, what questions do you ask AI vendors when they come to your facility and they're offering tools?

Dr. Abdulnour:

Let's go back to my three determinants of trust. I need the vendor to tell me that the tool is effective at the task it was designed to do,

and it's designed to support—so, some measure of effectiveness, for example, for a tool that screens for sepsis. Show me that it's sensitive and specific.

I need some transparency. I need to know how these tools work. And now, a lot of these vendors will have proprietary systems, but some measure of what training data was used. Did they do some tests for safety and bias? And so on and so forth.

The last thing is benevolence, meaning, can the vendor tell me that this tool actually made a difference in the patients' lives or in the clinicians' lives? And this is all about implementation, which is the hardest thing to prove because implementation design is not necessarily the top priority of a lot of these vendors.

Dr. McDonough:

Let's swing into practical guidance. What's one thing every clinician should start doing with AI immediately, just after hearing this? What would you suggest?

Dr. Abdulnour:

What I would suggest is consider AI a copilot. Rule number one: a copilot, by definition, does not replace you. A copilot, by definition, is not someone you have blind faith in. It's not someone who you don't take their recommendation seriously. The copilots are in this gray zone: I trust them enough that I can engage with them. And that should encourage clinicians to use AI. This is the number one, two, and three thing I say to folks that haven't used it yet: use it. The only way you're going to get better at it and build the muscle and the skills is to try them out. Use different prompts, read about how to use it, learn about it, and listen to a podcast like this one.

The second thing is when you're using an AI, again, just like working with a resident, assume that they may be better than you. Assume that they may know more than you—new things that they've learned. And I think it is important for you to be open to the fact that AI is going to make you better.

And rule number three is assume AI is wrong until proven otherwise. And that's the hardest—how do you work with something that can be better than you? But you need to assume it's wrong unless proven otherwise. So, that's where I land on this concept of adaptive practice. In the same way that I see patients with sepsis all the time, right, do I pause and think critically in an hour on every one of these patients? No. When I feel that it's a familiar presentation, I just go with the flow and have a low level of critical thinking. I'm a cyborg. But when it doesn't feel right, that's when I pause and I engage my critical thinking. Same thing with AI. If you're using a tool in a high-stakes situation where you're really unsure that this AI is going to be good for you, be a centaur. Be cautious, and engage your critical thinking. For example, when using OpenEvidence, look at the references.

The key thing to bring all this together is, how do you work with a tool that can be better than you at the same time you can trust? It's really to know when to pause and engage with your critical thinking and when to switch from a cyborg mode, where you're using the AI as if it's part of your brain, to a centaur mode where you are using the AI, but you leave the final word to you. You are the one who's making the decision. So, we need to learn how to switch from a centaur mode to a cyborg mode. Like I said earlier, I'm a cyborg when I use ChatGPT to send an email to my friend. I'm a centaur when I'm using OpenEvidence to help me make clinical decisions in the ICU.

Dr. McDonough:

Right now, we're in a situation where some community hospitals don't necessarily have all the AI tools. They have some free things, but they don't have others. What do you do in that situation to start to take advantage of the opportunities that are out there?

Dr. Abdulnour:

It's a tough question, Brian, because on the one hand, these tools have been so successful at solving such painful problems for clinicians that saying, "Hey, don't use them until your health system has approved it and has vetted it" is difficult. I can say it for myself. With some of these tools that I'm using and I'm very savvy with, so I know when to use it and when not to, it's hard to say, "Don't use it."

What I would say is, if you're a clinician working in a hospital where your system doesn't have the resources or hasn't vetted a tool yet, practice it. Practice using the tool very cautiously, and never put in patient health information. You can use de-identified cases or prompts. Try it. And if you think it's really making your life easier, keep knocking at the door of your hospital's administrators and tell them, "Hey, we need guidance on this. We need an approval process. We need some oversight. This tool is really making my life easier, and I want to make sure I use it safely." That's my advice.

Dr. McDonough:

This is a question I'm really excited to ask you. Let's go into the future. Five years from now, what will define an excellent clinician? In other words, what's the blend of using AI and growing with this, and what will be different from today?

Dr. Abdulnour:

As a patient, an excellent clinician is going to be the clinician who, when they enter the room, is spending the entire time looking at me,

asking questions, and doing a physical exam—not worried about a note or anything of that nature. At the end, there is an AI-generated report that I can go through with the clinician, and the clinician can astutely tell me what is right and what is wrong and answer my questions, signaling to me that they know what they're doing, and they're not relying entirely on the AI.

Dr. McDonough:

What gives you hope and gets you excited about AI in medicine?

Dr. Abdulnour:

What gets me excited—AI has been there for a while, and what makes large language models in particular, and the tools that use them, like OpenEvidence, the AI scribes, and ChatGPT so different, is that they have, in the words of Yuval Noah Harari, these tools have hacked humanity's operating system. They've hacked humanity's code, which is language, right? So, these tools aren't processing images or sounds; they're processing words and are able to co-think with us. So, for the first time in history, we have tools that can think alongside us and free us from a lot of other burdens. And all of a sudden, our efficiency, our wellness, and our productivity have multiplied. I feel it every day, and I feel myself being a better editor, a better clinician, and a better teacher. And a lot of it has to do with how I've incorporated the use of these tools in my day-to-day.

So, that's the hope and the excitement, and that's why I'm really excited by it. I want to work more on it. And I just want to make sure that we do it safely. We can only do it safely if we do the right research and the right testing, if we engage with the patients and with the stakeholders, and if we have a transparent conversation on how we govern these, how we use these, and how we implement these.

Dr. McDonough:

You mentioned safety, and that may be your answer, but what keeps you up at night?

Dr. Abdulnour:

What keeps me up at night is the premature implementation of these tools driven by business needs and financial needs, and a market competitiveness without any oversight and without any say from patients and from clinicians. And we'd end up being in the same place we were when EHRs were rolled out, where business incentives were met, but the burden of care has increased, the quality of care has decreased, and we actually may be in a worse place than we were before.

Dr. McDonough:

Wow. That is a really, I think, prophetic statement, and something that worries a lot of us as well.

Raja-Elie, people have heard what you've had to say, and I know they want more.

Dr. Abdulnour:

They can follow our work on NEJM clinician, which is a very clinician-oriented sister publication to the *New England Journal of Medicine*. I'm on social media. I'm on Instagram, Twitter and Facebook, speaking a lot about AI. They can engage, and they can send me direct messages. I'm always happy to have conversations. I'm always happy to talk about AI. I think it's a very important topic to discuss.

Dr. McDonough:

You've been listening to *The Convergence* on ReachMD, where innovators and physicians explore the technology transforming medicine. To hear about other technological advances shaping the future of medicine, visit [The Convergence](#) on ReachMD.com, where you can Be Part of the Knowledge. Thanks for listening.

We'll see you next time.

And Raja-Elie, thank you so much for taking the time to join us.

Dr. Abdulnour:

My pleasure, Brian.