

PCAST Public Call on Forensic Science

September 1, 2016 / 4:10 p.m. EDT

PCAST members in attendance

John Holdren, Eric Lander, Maxine Savitz, Wanda Austin, Christopher Chyba, Susan Graham, Chad Mirkin, Barbara Schaal, Daniel Schrag, Christine Cassel, Craig Mundie, Ed Penhoet, Mario Molina

SPEAKERS

Dr. John Holdren
Dr. Eric Lander
Dr. Maxine Savitz
Dr. Christine Cassel
Dr. Wanda Austin

PRESENTATION

Moderator Ladies and gentlemen, thank you for standing by. Welcome to PCAST Public Call. At this time all participants are in a listen-only mode. As a reminder, this conference is being recorded. Now I'd like to turn the conference over to Dr. Holdren. Please go ahead.

John Holdren Thank you very much, this is John Holdren, Assistant to the President for Science & Technology, and Co-Chair of the President's Council of Advisors on Science & Technology. This is a public conference call meeting of the full PCAST and it is, as noted, being recorded. It is on the record therefore, automatically and literally.

I want to open this discussion, which is focused on the PCAST report on *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature Comparison Methods* by noting that this study, which began last year, has been the focus of a tremendous amount of effort by the PCAST working group on the topic. It was chaired by Eric Lander. It included as well Jim Gates, Susan Graham, Michael McQuade, William Press, and Daniel Schrag. It was supported ably by a number of very hardworking

staff members, Tania Simoncelli and Kristen Zarrelli. I heard an interruption there, but Tania Simoncelli and Kristen Zarrelli at the Broad Institute for Genomics; and Diana Pankevich at OSTP.

In the course of the study the working group benefitted from inputs from a panel of senior advisors consisting of distinguished experts and practitioners in the law and in statistics. The group also received inputs from a wide variety of other experts and stakeholders, but I want to emphasize that the responsibility for the resulting report, its recommendations, its findings, and the rest of its content rests solely with the members of PCAST. Not just the working group, but the full membership of PCAST who have reviewed the report and will have an opportunity later in this call to vote on whether they approve the recommendations and the substance of the report.

I want to note, in fact I think we should, I should've done this at the beginning, but I think I should ask Ashley Predith to do the roll call so that the members of the public listening in will know which members of PCAST are participating on the call, and in the vote.

Ashley Predith I'm going to go down the list of PCAST members names, and if you could just say yes that you're here.

(Members present: Wanda Austin, Chris Chyba, Susan Graham, John Holdren, Eric Lander, Chad Mirkin, Maxine Savitz, Barbara Schaal, Dan Schrag, Chris Cassel, Craig Mundie, Ed Penhoet, Mario Molina)

Ashley Predith All right, we have a couple of people, two who will be calling in late, Mark Gorenberg and Michael McQuade; and then five people who were not able to join us today. We'll be hearing from them by email.

John Holdren But yes, we do have proxies by email from Rosina Bierbaum, Jim Gates, and Bill Press.

Okay, with that said, I want to turn the call over to Eric Lander to discuss further what went into the report, and what recommendations are coming out of it. Eric.

Eric Lander

John, thank you very much. Let me first take a moment to echo your thanks to the amazing support that we've received from our staff. Tania, Kristen, and Diana, thank you very much for an awful lot of work. And again, echo the comment about the fact that we received a tremendous amount of input, including in response to a public call for information where we got about 70 replies from the community, we're very grateful for that. And to underscore what John said, all of that input is gratefully received. It does not imply any endorsement of anything that PCAST itself concludes, we're really grateful for it.

We discussed this topic at several PCAST meetings already, so folks will be quite familiar with it. It's already broadly known what we've been working on, so I'm just going to give a very abbreviated background here. The PCAST report itself is a natural follow-on to the National Research Council's 2009 report on forensic science, which was a very influential report. It pointed out, among other things, that many forensic scientific methods lacked actual scientific evidence of validity, and actual measurements of accuracy, and expressed considerable concern about this. In effect, it turned the spotlight on the question of the need for empirical evidence to support forensic science methods. It was not that some of this was not known, but I think it was in many ways a wakeup call for many to see all of that organized in the work of the National Academy.

Of course, it really matters, because the law does indeed take seriously the question of what sort of expert evidence might be used in courts. I know that forensic laboratories take seriously the standards that they are going to apply, and so it's very important to make sure that there is a scientific foundation for the things that are being used in forensic science.

In particular, the law has a set of criteria laid out, including in the Federal Rules of Evidence, that expert testimony should be based on reliable principles and methods, and that they should have been reliably applied to the facts of the case. All of that, including Supreme Court decisions and lots of discussion, point to the question of needing to understand scientific validity. That's the backdrop there, and the president and the administration have responded to the questions raised by that National Academy Report NRC 2009 by taking a number of steps including the creation of the National Commission on Forensic Science, which meets

regularly. One of our PCAST members, Jim Gates, is a member that commission. Under NIST there was created the Organization for Scientific Area Committees that deals with questions of best practices in forensic science, and trying to ensure the development of standards documents, so these are all very important steps.

A big issue, though, raised in the academy report that isn't addressed by those two steps is the scientific question of scientific validity for particular forensic science methods. How does one think about it? What does it mean to have methods that are scientifically valid, and to know their reliability? Reliability means you actually know what its accuracy is.

I think the most interesting thing that has transpired since the 2009 report by the National Academy has been a growing change in the forensic science community itself. There was a traditional view—and I'll caricature just a little bit here in the shortness of time—but there was a traditional view held by some in the forensic community that empirical evidence of scientific validity and empirical measurements of the levels of accuracy were not really necessary. That in fact, examiners could assert, based on experience, or training, or other things, that their conclusions were justified.

Indeed, I think it's fair to say there's been a lot of shift here, but there was a time even the United States government held the position that things like latent fingerprint analysis was "infallible." We had documents that said that. I think we've been seeing a shift in the forensic science community toward what one might call a new empirical movement that says, yes, we actually really do need to possess scientific evidence of validity through properly done studies, and empirical measurements of accuracy for methods that we're using routinely in forensic science. I think it's really remarkable to watch a number of groups that have jumped in to—since the 2009 report—supply very thoughtful studies.

Here, I think PCAST takes its hat off to the FBI, which has been a model in its work on latent fingerprint analysis. There's been a series of truly elegant papers from the FBI that have done what are called 'black box' studies, where they measure how do examiners acutely perform in circumstances; and accompanying them, 'white box' studies to ask, how

do examiners make the decisions that they're making? They have revealed lots of interesting things, including areas that need improvement; like examiners looking back and forth, and changing their view of a fingerprint based on how well it matches. I just have to say again how wonderful it is to see the FBI take on questions like that in a very rigorous, scientific way, publish them in the *Proceeding of the National Academy of Sciences*, and really be willing to hold a mirror up to practices, and where they're not optimal, talk about how to improve them.

This black box study reports actual error rates. That's very helpful to know. Fingerprints aren't infallible, but they're quite good, they're measured at an error rate of about 1 in 600 as a best estimate. If you include a statistical confidence limit, as one should in fact, about 1 in 300 error rate. All that's a beautiful, beautiful development, and PCAST clearly supports this empirical movement in forensic science, and thinks it is indeed necessary.

PCAST, in this report, aims to take the next step from the National Academy in saying that from a scientific point of view, you can't call a method valid and reliable unless you've empirically tested it and measured its accuracy under conditions relevant to its intended use. The PCAST report though focuses on, well, what are the scientific criteria for testing validity and reliability? We speak about two kinds: foundational validity, which is the scientific parallel to this legal concept of reliable principles and methods; and validity as applied, which is a scientific concept parallel to the idea that a method has actually been applied in practice.

We've tried in the report to lay out a clear, scientific definition of the scientific criteria for validity and reliability in the case of one particular class of methods, forensic feature comparison methods. Examples we'll come to in a second there. We did that first, because we think there hasn't always clarity about it, and we see the field moving in this direction, and recognize it's absolutely essential that it does, if it's going to have scientific validity.

We then, having laid out definitions, illustrate these definitions by applying them to seven topics. The DNA analysis of single samples and simple mixtures is the first. The second is the DNA analysis of complex

mixtures. The third is bite marks. The fourth is latent fingerprint analysis. The fifth is firearms analysis. The sixth is footwear analysis with respect to individualizing the source of objects, as opposed to what are called class characteristics. The seventh, we didn't undertake a full study, but we did an evaluation of some guidelines that had been issued with respect to hair analysis.

That's the backdrop of the report. Let me turn now to the recommendations that emerged from having done all of this work, and gotten all of this input.

The first recommendation concerns ongoing assessment of foundational validity. I'm going to read a portion of the recommendation, and then it's supported by additional text. In the interest of time, I'm not going to go through all of the supporting text. But it's important that ongoing scientific evaluations of the foundational validity of forensic feature matching methods, or feature comparison methods, be conducted to assess on an ongoing basis. To ensure that scientific judgments are unbiased and independent, such evaluations must be carried out by a science agency that has no stake in the outcome. The National Institutes of Standards and Technology, NIST, we say, should perform such evaluations, and should issue an annual public report evaluating the foundational validity of key forensic comparison methods.

That's an important statement that this does need to go on, because there'll always be a continued evolution, and NIST is the obvious home for this because pattern feature measurement and comparison is metrological method, and NIST is the world's best metrological organization, and certainly the United States Government's metrological organization. In support of this, under Recommendation 1, we also recommend that the president request and the Congress provide increased appropriations to NIST to support these evaluation activities, and increased research activities at NIST. That's Recommendation 1.

Recommendation 2 concerns the development of objective methods for certain forensic feature comparison methods that are today subjective. One of the things discussed in the report is that, for example, latent fingerprint analysis and firearms analysis are subjected methods today,

rather than being based on objective criteria. PCAST recommends that NIST, the National Institutes of Standards & Technology, take a leadership role in transforming such subjective methods into objective methods. We can talk more about that later in the call, but we think now is a very good time to be able to do that, and it will be valuable for all.

In Recommendation 3, we talk about ways of improving the organization for scientific area committees that I had referred to. We recommend that NIST should improve the organization of scientific area committees which was established to develop and promulgate standards and guidelines to improve best practices in the forensic science community. We specifically recommend the development of a metrology resource committee and other changes within that work there. We also recommend that NIST ensure that the contents of the standards that are developed through this process be freely available to any party that may need them. All of that is discussed further, and people could ask questions if they want.

The fourth recommendation concerns R&D strategy for forensic science. Given the importance of forensic science and this evolution that we are seeing toward a true appreciation of the critical nature of empirical science, we recommend that the Offices of Science & Technology Policy should coordinate the creation of a National Forensic Science Research & Development Strategy, and we lay out some elements that should be considered within that. We say that in preparing the strategy, OSTP should seek input from appropriate federal agencies including the Department of Justice, the Department of Defense, National Science Foundation, the National Institutes of Standards & Technology, federal and state forensic science practitioners, forensic and non-forensic science researchers, and other stakeholders. So that's a set of recommendations pertaining really to science-based agencies.

The next recommendations pertain specifically to the FBI laboratory. In those recommendations, again, I want to underscore that we are extremely impressed that at least some units within the FBI have been undertaking rigorous studies and it's a real model, so our recommendations here pertain—Recommendation 5—to expanded forensic science agenda at the FBI. We recommend that the FBI should undertake a vigorous research program to improve forensic science, building on its recent important

work on latent fingerprint analysis, and we discuss that at some length there.

We discuss the need for black box studies, and urge the FBI to assist in the design and execution of additional black box studies for subjective methods, including additional studies for latent fingerprint analysis and firearms analysis, because it's important to be able to have such measurements, and the FBI has shown itself able to do that kind of work at a rigorous level. We also say—although we've recommended that NIST take the lead on the development of objective methods – we believe that the FBI should play an important role in transforming these subjective methods I was referring to before into objective methods. We also believe the FBI should play an important role in increased rigor in forensic testing by moving toward routine blind proficiency testing within the flow of casework in its own laboratory, and assisting other laboratories in doing so as well, and encouraging routine access to and evaluation of the tests that are used in commercial proficiency testing.

Again, with respect to the FBI, the FBI has thought very carefully about how to improve latent fingerprint analysis, including rules that are called linear ACE process. Basically, where examiners document their description of a fingerprint before turning to the potential match that there might be. We urge the FBI to vigorously promote the adoption of this best practice linear ACE process in fingerprint analysis. That it also promote regular public reporting of quality issues in case work. There is very useful work here in the Netherlands that I think led the way, and we think the FBI would ideally follow those practices. And then budget, we recommend a significant increase in the budget of the FBI's laboratory for forensic science research activities, so that it can be able to undertake the very good work described there.

The next two recommendations are to the Attorney General. In the Recommendation Number 6, we recommend that the Attorney General should direct attorneys appearing on the behalf of the Department of Justice to ensure that expert testimony in courts about forensic feature comparison methods meet the scientific standards for scientific validity. Of course, the Attorney General ensures that testimony meets the legal standards that are always required, but meeting the scientific standards for

scientific validity is very important because it's in the interest of the United States broadly to ensure that evidence is scientifically valid.

We suggest that the Department of Justice should undertake an initial review, with assistance from NIST, of subjective feature comparison methods used by the Department of Justice to identify which methods, beyond those reviewed in the report, may lack appropriate black box studies that are necessary to assess foundational validity. Where relevant methods haven't yet been developed, the Department of Justice should encourage and provide support for appropriate black box studies to assess foundational validity, and make actual empirical measurements of reliability.

In Recommendation 7, we turn to some specific guidelines that were recently issued by the Department of Justice with respect to uniform language for testimony and reports by forensic examiners. In those guidelines that were issued over the last several months, the Department of Justice took an extremely important step. In the past, there was the ability of many examiners to make sweeping statements that methods were infallible, 100% accurate, could never fail, etc., etc., when in fact there was no scientific support for that at all. To its enormous credit, the Department of Justice states clearly in its guidelines that examiners should not be able to make such scientifically invalid claims.

PCAST, however, has a concern here that in doing that, the guidelines also say that examiners should not be able to make valid scientific claims about measurements of accuracy. Of course, with regard to the empirical necessity here, it's crucial that empirical measurements of accuracy be stated. So we recommend that the Attorney General should revise and reissue for public comment the DOJ's proposed uniform language for testimony and reports, and the supporting documents, to bring them into alignment with the scientific standards for scientific validity that include the necessity that a statement that a method is foundationally valid include actual empirical measurements of its accuracy.

Finally, in the last recommendation, we make a recommendation to judges who have a role in this as well, because judges have to make a decision about whether expert testimony is admissible, that's entirely a legal

decision that has to be made on legal criteria, but those legal criteria have, as their overarching inquiry, when it's a question of scientific evidence, they have the question of scientific validity. So without telling judges anything about the legal criteria, we do say that federal judges should take into account the appropriate scientific criteria for assessing scientific validity. They will take that into account and then apply that as they see fit with regard to the law, but we think it's important that in reasoning about this, judges are well aware of the scientific criteria for assessing scientific validity.

We also recommend that federal judges, when permitting an expert to testify about foundationally valid feature comparison methods, should ensure that the testimony that is given about the actual accuracy of the method and the probative value of composed identifications is scientifically valid, in that it is limited to what the empirical evidence supports. In this regard, we support what the Department of Justice has already said that claims like zero error, vanishingly small errors, essentially zero error, negligible error really should not be allowed because they're scientifically invalid statements. Again, federal judges will have to decide what to do, but these are our recommendations here.

To assist judges, we suggest that the Judicial Conference of the United States, through its standing Committee on the Federal Rules of Evidence, should prepare with advice from the scientific community, a best practices manual, and an advisory committee note providing guidance to federal judges concerning the admissibility of expert testimony based on forensic feature comparison methods. The Federal Judicial Center, which runs among other things educational programs, should develop programs concerning the scientific criteria for scientific validity of forensic feature comparison methods.

That was a mouthful, but that is a description, John, of the background and of the recommendations, and I will turn it back to you to ask if members of PCAST have questions that we could answer on behalf of the working group.

John Holdren

Great, well thank you, Eric, for that summary, and thank you even more for the extraordinary work you did as chairman of the working group, not

to mention your many other contributions as my co-chair of PCAST. At this point, we do have time for some questions from members of PCAST. The PCAST members do have, and have had a draft of the report, and so they have had the opportunity to read the whole thing, but they may have some questions that they'd like additional insight about. I open the floor to PCAST members to propose those questions.

Maxine Savitz Hello, John, this is Maxine Savitz, I have a question. Eric, I want to thank you, and also the whole working group and staff also, for the excellent work and commitment to this. In one of the recommendations, you mentioned being able to transform some of the important feature comparison methods, such as fingerprint analysis and firearm, from subjective methods into objective methods. This would coordinate this working with the FBI labs and others. Could you go into a little more detail about what's really involved to make this transformation so that these can be objective, and is this a very lengthy process?

Eric Lander That's a great question. I apologize, of course, because of limited time, not to have gone into that, but let me try to elaborate a little bit.

John Holdren And there's still limited time, so don't elaborate too much.

Eric Lander Just a bit, John, just a bit. Basically, looking at a fingerprint or looking at markings on a bullet is a visual comparison problem. This is something that was very hard to think about objectively before, but image analysis has undergone a just dramatic transformation in the last four or five years, such that as I think many of us know for example facial recognition technology has just increased by leaps and bounds. The ability to parse images into their segments suggests strongly to PCAST that—it was the PCAST working group—that similar methods applied to fingerprints and to firearms, for example, would allow pretty fully objective methods; or in any case a very vigorous program should develop toward it, because it would largely eliminate any of the questions around validity if one could take a large supply of samples, run them through highly objective programs, and demonstrate their accuracy. This was not a reasonable thing to expect years ago; I think it is a reasonable thing now, and we strongly urge it, because we think it will lead to increased quality.

John Holdren

Good, other questions from PCAST members?

Chris Cassel

Eric, this is Chris Cassel, I have a question. First of all, let me add my congratulations to you, and the working group and the staff put together this really thoughtful and very deeply analyzed report. It's just got me thinking, this is probably a topic having to do with another day and another PCAST, but whether you've been very careful, even though the issues of evidentiary science are very clear and standardized and universal, you've been very careful to limit your observations here to feature comparisons. It's interesting how much of judicial evidence comes under that category. There are other kinds of things that people use for evidence, even lie detectors and similar kinds of things. I'm wondering if these same principles might apply to other kinds of evidence that come forward, or either that are already in use or that you might foreseeably see coming down the pike where the same principles would apply.

Eric Lander

Again, a great question, I think it's important to say that when we talk about scientific validity the analysis is not a one-size-fits-all analysis. It's flexible enough that it has to apply to the specific field being analyzed. If it was arson analysis, fire patterns in a building, or shaken baby syndrome or something, one would have to make sure that the analysis was appropriate to the area.

We limited ourselves deliberately to forensic feature comparison methods, measuring and comparing features because they all fall into the field of metrology, essentially measurement science. So I want to express appropriate scientific caution here that we've tried to analyze that one problem—which frankly, is a big enough problem, as you say, to bite off and chew – and has also been the problem where most of the issues have arisen, for example, that have led to exonerations based on incorrect forensic statements. I'm sure that the thinking can be broadly useful, but the analysis is specifically directed to feature comparison methods.

John Holdren

Thank you, I think we have time for maybe one more question from a PCAST member.

Wanda Austin

John, this is Wanda Austin. If I could, I'd like to ask a question.

- John Holdren Please.
- Wanda Austin Eric, again, thank you for your leadership and for the efforts of your team in leveraging science and technology to the benefit of our society. My question was, however, did you consider or examine the use of forensics in the investigation of civil cases?
- Eric Lander We, again, tried to be very careful here. We focused on the use in criminal cases, particularly, not civil cases, simply for needing to bound the scope of our study. I should also say that our statements pertain to the phase where you're actually bringing evidence to court, where there is a particular standard that has to be met. With regard to say the investigation phase of a crime where you may be looking broadly, I think we all agree that a much broader view of what types of methods, including very early developmental methods might suggest a hypothesis, for example, of who might be connected to a crime. I want to be clear that our statements here are not meant in any way to cast doubt on the broadest possible use of methodologies to help develop information that could lead to the solution of a crime.
- John Holdren Good, well thank you very much for all of that. We now have, as we always do in our public meetings, a brief time for public comment. My understanding is that we have two people on the list, and I'm going to turn it over to over to our executive director, Ashly Predith, to run the public comment part of this program.
- Ashley Predith Sure, so our first public commenter will be Bridget Lewis, from the International Association for Identification. Bridget Lewis' line can be unmuted, and if she could have two minutes for her comments.
- Bridget Lewis Thank you very much, Ashley, thank you very much for this time. As the oldest professional organization on forensic science, we do totally support the efforts of PCAST to put more research. We totally believe that research is going to build all of the comparison sciences.
- One thing we are concerned about is there is also a need for training and the money that is in research also needs to be shared. The FBI has done an outstanding job in their research efforts, and they used to be—years

ago, I'm dating myself—the premiere place for comparison scientists to get their training. We would like to suggest that you also look at that aspect. Thank you very much for your time.

John Holdren Thank you.

Ashley Predith Our second public commenter is Emily Rice from the New Hampshire State Police. So Emily Rice, you'll have two minutes.

Moderator Ms. Rice, please go ahead. Emily Rice, your line is open.

Emily Rice I'm sorry, I think there was a miscommunication. I didn't have a comment.

Ashley Predith Okay, and then just to check, there was a Brian Lee, who originally registered also from the Forensics Exploitation Directorate. I want to check and make sure that he's—is he by chance back on the line?

Moderator Brian Lee is not on the line.

Ashley Predith Okay, thank you.

John Holdren Okay, well it then comes to a time to vote. I emphasize that this vote is on the recommendations and the substance of the underlying report subject to final edits, which is always the way that PCAST votes. As we pour over this extensive report, we will doubtless find a few typos, infelicities, and matters of clarity that need adjustment, but know following the vote, if the is for approval, there will be no edits that will change the substance of the report or the recommendations.

We do have, I think some additional PCAST members who are on the call now. Will you identify them, Ashley?

Ashley Predith Yes. Is Mark Gorenberg, could you unmute and tell us that you're on line. Okay, Michael McQuade, if you're on the line, could you unmute (Michael McQuade). Great, thank you, Michael.

John Holdren Okay, so we did the roll call before, and now we have an addition, Michael McQuade. After the folks who are on the phone vote, I will reveal what proxies we have, but let me again now ask for the ‘aye’ votes of those who are prepared to approve these recommendations and the substance of the underlying report, please say, ‘aye.’

Members Aye.

John Holdren And those who are not prepared to approve the recommendations, and the substance of the underlying report, please say, ‘nay.’ Thank you. I’m also pleased to report that the proxies of Rosina Bierbaum, Jim Gates, and Bill Press are all ‘aye’ voting in favor, so with that, we have approved the recommendations and the substance of the PCAST report on *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature Comparison Methods*.

I should add as we move to close this call that originally we had thought we might be prepared to discuss the recommendations and vote on the ongoing PCAST study on bio threats, but as it happens, work is continuing on those formulations. We expect we will have the vote at our public meeting in Washington D.C. at the end of the month on the bio threat report.

Let me just ask, do any other PCAST members have comments before we close this call? Hearing none, big thanks again to the working group, to the staff, to the full PCAST who have read and approved the recommendations and the substance of this report. Thanks to all the members of the wider community who took the time to join this call. We very much appreciate your interest. This call and this public meeting by conference call of the full PCAST is hereby adjourned.

Moderator Thank you, John, thank you, Eric.