

**Health Information Technology Standards Committee
Summer Camp
Patient Matching Power Team
Draft Transcript
July 1, 2011**

Judy Sparrow – Office of the National Coordinator

Thank you, operator. Good morning everybody, and welcome to the Standards Committee's Patient Matching Power Team. This is a federal advisory call, so there will be opportunity at the end of the call for the public to make comment. And members, please remember to identify yourselves when speaking.

A quick roll call – Marc Overhage?

Marc Overhage – Regenstrief Institute

Present.

Judy Sparrow – Office of the National Coordinator

Judy Murphy? David McCallie?

David McCallie - Cerner Corporation

Here.

Judy Sparrow – Office of the National Coordinator

Nancy Orvis? Chris Roth? Walter Suarez?

Walter Suarez - Kaiser Permanente

Here.

Judy Sparrow – Office of the National Coordinator

Shaun Grannis?

Shaun Grannis - Regenstrief Institute

Here.

Judy Sparrow – Office of the National Coordinator

Lisa Gallagher?

Lisa Gallagher - HIMSS

Here.

Judy Sparrow – Office of the National Coordinator

Anybody from Doug Fridsma's office? Okay. With that, I'll turn it over to Marc Overhage.

Marc Overhage – Regenstrief Institute

Thanks Judy, and good morning everyone. What I thought we might be able to get through today is what we talked about last time we had an administrative call, focusing a bit on the Match Metadata – we kind of just got to that at the end – and then taking a walk all the way through the

recommendations – which I’ve tried to, in the slides, take the feedback that we received at the Standards Committee – although, I had to confess to Judy last night that the dog ate my homework. I lost one of my sticky notes that I had from that and may have missed something that I didn’t remember. So Walter or anyone else – Judy or anyone else who was on that call, if you spot something please remind me – as well as having looked back through the Power Team for Metadata and the Privacy and Security Tiger Team’s recommendation letter from February regarding patient identity and tried to round out some of the things that we had put together there. And then leave perhaps 15 or 20 minutes – cause we always seem to run out of time – to fill in whatever else we need. And then a few minutes as needed for public comment before we adjourn. So any – does that sound like an okay way for us to proceed today, or anybody have any suggestions for changes to that agenda?

Judy Murphy - Aurora Health Care

No. Mark, this is Judy Murphy. Just to let you know I joined.

Marc Overhage – Regenstrief Institute

Good morning, Judy.

Judy Murphy - Aurora Health Care

Good morning.

Marc Overhage – Regenstrief Institute

Alright. Well, let’s go ahead to the next slide then. Judy – and I apologize – I don’t know if you’ve got control or I do. And one more....

Judy Sparrow – Office of the National Coordinator

Actually, I guess Altarum will move it for you.

Marc Overhage – Regenstrief Institute

Great. So if we could go all the way to the last slide Judy, which is where – there perfect – you knew what I meant. We had just begun to talk a bit about – and certainly this was addressed in the Privacy and Security Tiger Team letter to a degree – what information besides “here are the matched patients” might be reasonable to ask whoever is doing the matching to return so that we could consider how that information might be structured and used. And the things on the slide here are just to be stimulating and thought-provoking. The one thing that the Privacy and Security Tiger Team letter seemed to come back to frequently was some information about – and I just missed one of the things that I forgot to change from the Standards Committee, which is: I used the term “Probability Level” and I think Shaun you suggested “Confidence Level” is probably a better term for us to adopt for how solid was this match. And may I use that as a starting point for discussion? Is that something that we would want and expect to come back with a match?

David McCallie - Cerner Corporation

So is it – this is David – I’m unclear if this is just an absolute assessment but you’re not given any other choices so you don’t know where this match stood with respect to other choices? It’s just a number? In other words, it would seem to me: One way to do something like this – not saying it’s the right way – but one way would just say, “Here are the top 10 closest matches and some scores that give you an idea of how close they were.” And I don’t think we’re taking that approach. But if you eliminate the other nine and just return the score without the notion of how close was the second best choice, do we gain much from that?

Marc Overhage – Regenstrief Institute

Well, let me just take a crack at that David. And the first thing you might be interested in knowing is how likely is it that the matched patients that were returned are false-positives. That would be a rate that – [interrupted]

David McCallie - Cerner Corporation

But you said [cut off] pleural. I thought we were kind of in this mode that we weren't returning a list of patients for fear of exposing unintended PHI.

Marc Overhage – Regenstrief Institute

Well there certainly may be more than one patient matched though, depending on the model that the particular organization had. An EHR, for example, might have a few patients registered who match with different ID numbers. I think across hospitals, for example, is what – 5 to 8% for duplicate registrations.

David McCallie - Cerner Corporation

Oh yeah. No, I'm sure there'd be a lot of multiple matches, but I was under the impression that – from maybe some other conversation – that we weren't – [interrupted-indiscernible]

Marc Overhage – Regenstrief Institute

Well let me make a distinction between – you may return a single – you may return patients that you believe are matches, and there may be more than one. You could – and this is where I think we had a previous discussion – you could return patients that were not – that you weren't sure were matches but might be. And I think that's what I heard you talking about, or maybe I'm – yeah.

David McCallie - Cerner Corporation

Yeah, I mean it's a subtle distinction between: "I think this might be a match at certain probability versus I think this is a match but here's the improbability of my certainty."

Marc Overhage – Regenstrief Institute

And that's where I was thinking more, but Shaun you might have some thoughts to share with that as well.

Shaun Grannis - Regenstrief Institute

Sure. There's – I think the question of if we're going to send a response with a single patient or multiple patients – the need for confidence might vary here. When you're given little information such as, "here's the single patient we think is the best match," providing some confidence measure of this single view into the other side's matching algorithm may make sense. If the model is to return multiple potential matches from which a person will choose – you know, guided by some sort of confidence interval – I frankly think there's less need for confidence if we're relying more on the human being to adjudicate upon the set. But confidence in my mind – and I think maybe David was getting at this, but I don't want to put words in his mouth – was to provide some ability for the human being to make a decision about whether or not to accept the response or not. If, on the other hand, we're talking about sending the top 10 best matches, a confidence measure of each of those matches would be useful but if the human being is gonna be sort of looking at the same identifiers that the algorithm used to make some sort of decision. So workflow is important here to understand how that confidence measure would be useful.

Marc Overhage – Regenstrief Institute

And I think where we were headed on that was that the idea – and this was sort of second – and maybe we should just go their now – item to talk about was the notion that if there were additional patients that might match, if you have some additional data, that part of what we might ask the algorithm to return would be – “and if you could tell me these two attributes, I can add some more potentially.”

David McCallie - Cerner Corporation

This is David. I was wondering that also. That maybe the score could be a proxy for matched criteria that was evaluated behind the scenes without exposing the PHI to the human who’s making the judgment. In which case, the score is telling you – “Trust me. I’ve looked at the stuff you can’t see, and this is a good match.” You know, so for instance, you could put in to query the last four digits of the social security number, take that into account in your algorithm, but not show any social security numbers of the patients who are putative matches but reflect that into the score. In which case, the score is communicating something really important without exposing PHI. But I’m not sure that’s what we’re talking about here.

Shaun Grannis - Regenstrief Institute

Yeah, I – again, depending on sort of the workflow, the design of the actual interface, and the algorithm one uses – all of those pieces will influence exactly what you want. Is there a way to take the task to say, “Some measure of confidence should be permitted to be conveyed but not necessarily –?” I’m just wondering what sort of strength to the recommendation we’re looking to make here.

Walter Suarez - Kaiser Permanente

This is Walter. I have a couple of comments. First, I’m concerned about the idea of sending out multiple patient – possible patient matches back to someone that is requesting or – [interrupted]

Marc Overhage – Regenstrief Institute

Walter, just to clarify – I think where we landed before is we would send back only patients that are matched. And there may be multiple of them.

Walter Suarez - Kaiser Permanente

Well, but the definition of matched is what becomes critical, because a match could be 95% or – [interrupted]

Marc Overhage – Regenstrief Institute

Well, and that’s right. So the two things are: One is there is – we have recommended that the Policy Committee establish a level of matching or a level of confidence that – [interrupted]

Walter Suarez - Kaiser Permanente

Well, I would say probably the Policy Committee should define what a match is based on levels of confidence, levels of thresholds, or whatever. But I think that’s – [interrupted]

Marc Overhage – Regenstrief Institute

So I’m assuming there is a threshold of false-positives here that we’re willing to tolerate that the Policy Committee will establish. I don’t know what that number – you know it’s 10%. So we’ve got that. So the algorithm presumably can cut off matches that would lead to that level of error rate and return only those matches that would limit the error rate to 10% or whatever the value is.

David McCallie - Cerner Corporation

But that could well include more than one patient. I don't think there is – [interrupted]

Marc Overhage – Regenstrief Institute

Of course.

Walter Suarez - Kaiser Permanente

Yeah, and again the concern of course is there is a level of tolerance from a Policy Committee perspective. There is a level of tolerance from a legal and regulatory and law perspective and risk perspective. So, I'm – but you know – [interrupted]

Marc Overhage – Regenstrief Institute

And that's not our job to set that level.

Walter Suarez - Kaiser Permanente

Yeah, no I understand. So, but that's my first – I guess – point. So, if the assumption is that we are going to depend on a definition of a match based on a Policy Committee-identified level of tolerance for false-positives, then that's the underlying assumption; and we will be seeing multiple patient responses or responses that include multiple patients that could be considered matches. Right?

Marc Overhage – Regenstrief Institute

Correct.

Walter Suarez - Kaiser Permanente

So, the next point I would make is: It seems to me that for each of those patients there would be different degrees and different – I guess levels of matches in terms of – For patient #1, we didn't have the middle initial and the phone number or some other characteristic. And so this is – you know it's sort of a metadata for each patient about what went into deciding the level of confidence. Right? It's not an over-arching formula that applies to, “Here is the 500 matches, and this is the level of confidence for all 500.” It's almost like a individual level of confidence for each single match.

Marc Overhage – Regenstrief Institute

Correct.

Walter Suarez - Kaiser Permanente

Okay.

Marc Overhage – Regenstrief Institute

I would agree with that. Others – [laughter]

David McCallie - Cerner Corporation

Yeah, this is David. I think that the spirit is to avoid the notion that you're just searching through a database of patients trying to find interesting things. On the other hand, you have to admit that you may very well return more than one actual identity that are above whatever threshold you set for the query; and you need something to disambiguate, or at least to be aware that you're running a risk of not disambiguating and alert people accordingly. So, I think the notion of discerning some kind of a score makes sense. I'm really nervous about calling it a probability level, because for it to really be that we'd have to do an awful lot of modeling of your populations at a level that I'm not sure is worth it for people to go through.

Marc Overhage – Regenstrief Institute

So let me – I think you’re exactly right. We need to get more precise about what this is. It seems to me that the notion of a false-positive – the “how often will we include an inappropriate patient?” – is measured at the population algorithm level and not an individual patient level.

David McCallie - Cerner Corporation

Right.

Marc Overhage – Regenstrief Institute

Does that make sense to everybody?

David McCallie - Cerner Corporation

Yeah. But it’s – [interrupted]

Marc Overhage – Regenstrief Institute

So, there is sort of a – what is our – and it does take work. Like you said David, that’s not something that you can magically compute. That’s something that the matching organization needs some assessment of. And the organization making the query would like to know [distortion].

David McCallie - Cerner Corporation

And isn’t that – I mean I’m not enough of a statistician to know – but isn’t that highly dependent on factors that will change from case to case, from patient to patient? In other words, if I supply absolutely everything that the algorithm would like to have, I’m gonna get a different false-positive rate for a given population than if I only can supply some of those fields or partial completion – “middle initial but not middle name, first four instead of all nine.” So I’m just – how would you calculate this in a meaningful way?

Shaun Grannis - Regenstrief Institute

There – depending on which model – this is Shaun. Depending on which model you use for the combination of fields that you supply, both present and absent, there are fairly robust ways to estimate the false-positive and true-positive rate for that particular kind of match.

David McCallie - Cerner Corporation

But that would be offline work you calculate against your particular population and against sample inputs?

Shaun Grannis - Regenstrief Institute

Right. So, we’re getting into some details – statistical details. One of the things I recommended in the ONC EHR-2 White Paper was that if we want to generate this sort of false-positive metric, we actually – so if Indiana wanted to match with patients in Kentucky, we would need to do an Indiana-Kentucky analysis of the combined dataset to establish a false-positive rate between those two data sources.

Marc Overhage – Regenstrief Institute

And you’d have to do that peer-wise with every other gateway in the entire country.

Shaun Grannis - Regenstrief Institute

Right. Now many – and we’ve observed this too – those rates tend to fall within a pretty reasonable band. I think we need more evidence to shore up that claim. But it may be the case

that with some fairly low investment, one can come up with fairly accurate estimates. But again, I personally think – and now I’m speaking as a researcher and so I’m biased – I think this is an area that we need to further investigate.

David McCallie - Cerner Corporation

Yeah, which is why – it’s again David here – even though I think this would be a fascinating number to have, I’m not sure the burden of trying to estimate an estimate and provide it to a human who’s trying to locate a patient is gonna be worth the – I don’t think worth the return for what that effort is. It might [cut off] thing to say that it’s a best practice that you establish a periodic assessment of your algorithm’s performance, but that’s different from saying it should be returned on a match request.

Marc Overhage – Regenstrief Institute

I hear you David, and I think that’s making sense to me. Unless – Shaun, is there any measure of the particular match that an algorithm just made that we know enough to make a suggestion for what to return?

Shaun Grannis - Regenstrief Institute

In terms of – you’re talking about the confidence level algorithm – [interrupted]

Marc Overhage – Regenstrief Institute

Well or anything else that is sort of a byproduct, if you will, of the matching process that would be useful to the person receiving the data to help them know how to interpret it. And that I think the discussion here is helping sharpen that it’s not – you know it would be nice to know what is the [distortion] this particular patient. But that’s probably not a reality.

Shaun Grannis - Regenstrief Institute

Right, yeah. In general, I think some characterization of the matching approach you’re taking could be useful. But I’m speaking in very generic terms here because characterization – different algorithms work differently, so some algorithms you can’t even put a continuous assessment of false-positive rate or likelihood of match because they are rule-based algorithms. And so, the method of measuring confidence is tightly tied to the algorithm or approach to matching. And so, if different organizations use different matching algorithms, then the measure of confidence will have different meanings across those organizations. Within Indiana, within their Health Information Exchange, we have access to and analyzed that data. So we can generate a confidence level that we can define and provide meaning for. But, using the Kentucky-Indiana match, if to match into Kentucky we use Kentucky’s algorithm, then it needs to be Kentucky’s confidence level. If we match using Indiana’s algorithm, then it needs to be Indiana’s level of confidence. If we so desire that that’s something we’d want to send.

Walter Suarez - Kaiser Permanente

This could even go further if one thinks of federated HIEs where the actual algorithm is – the HIE is basically a switch and the query goes from one provider to another provider and the algorithm is being run by the provider that is being asked to match – to do the match. So it’s not a kind of a statewide or regional level aggregation.

Unknown Speaker

Good point.

Walter Suarez - Kaiser Permanente

This is like multiplied by thousands of providers out there. If I'm in Minnesota where there is no central repository to run these algorithms, and I'm in [indiscernible] and I'm searching for a patient and I send it to five providers in the state who will represent 90% of the market, each of those five providers is gonna run its own algorithm and then come back with matches. Right?

David McCallie - Cerner Corporation

That's why this NHIN model that we currently have isn't gonna scale very well. But that's another conversation.

Marc Overhage – Regenstrief Institute

So in trying to bring us at least in a direction where we know what the next set of questions are – it sounds like Shaun, you're saying that one level of what you would like to know is: "How is the matching done, ideally?" Which is sort of a transparency – somewhere there ought to be a website that describes this and a way to get there – how this matching was done. And it's gonna be quite variable to Walter's point because it might say, "Go look at these 105 places/websites, depending on where it got matched." But that might be the limit of what we – would be reasonable today to actually ask would be a pointer to that information. The recommendation that we might make is that there is additional needs to sponsor or there needs to facilitate additional research – Judy, you'll help with wording there – that some kind of periodic assessment and reporting would be appropriate in the best practice, related to the false-positive rate and confidence levels, given that those are gonna be generic in the sense that they're not gonna deal with – "If I send you a name from San Francisco that's got a higher probability of having Asian heritage" – or something like that that might have different matching characteristics in my dataset than my local matching experience. Or something like that. Does that seem like a responsible path forward?

David McCallie - Cerner Corporation

This is David. I wonder if – I mean this is a little bit of maybe trying to sidestep the hard problem – but I think we all have a vague notion that a remote system could, without getting too complicated, reveal some kind of stoplight coding of "highly likely match", "not so sure match", and "really lousy match." It's almost like the best practice is that the remote system should return some indicator of the confidence in the match without specifying necessarily that it would be a literal estimated probability or that it be some kind of a score on a scale of 0 to 1 or whatever. But just to alert the remote querier: "This might not be a good match."

Walter Suarez - Kaiser Permanente

So here is again the point I was trying to make before. So again, I'm sending a query for information about this patient. And there is a hundred organizations out there that are going to be responding, or whatever number. And each one is going to run its own. And there is gonna be – in my mind I guess – there's gonna be perfect matches; matches that represent "a perfect match", which again can be defined. And these would be defined, whether it's 99.99% of confidence or 100% of confidence, whatever. So there's that perfect match. And then there is the immediate next level which is everything else, which begin to say: "This is not a perfect match, but these are – the following are matches up to 99% confidence level." And each of them individually – each of those matches have to have the level of confidence and the type of data that went into it. So my thought was – I mean in this day and age of metadata – "What if there was a way to create a standard metadata-matching package that goes with every turn?" That for each entry – I mean I'm sending back – let's say my organization was asked, and I found 15 people. Five of them are perfect matches, and they seem to be all of them the same – it's just it varies a little bit with the

name or something or middle initial, whatever. Or two of them are perfect matches, and then the other 13 are matches that look like they're almost perfect – you know, the next level up to 99% confidence. And for each of those 15, there is a metadata package that tells all of the characteristics of the match; whether – what data was used, what confidence level was met, what matching algorithm was used. Those kind of things.

Marc Overhage – Regenstrief Institute

And Walter, I think that's the discussion – David and Shaun correct me if I'm wrong, or others – that I heard from David and Shaun was that that's probably – we probably can't provide that number for each match, because that's not a number that can be computed on a [indiscernible – interruption] query if you tell me the number.

Walter Suarez - Kaiser Permanente

I wasn't thinking of a number but more a series of data elements in a metadata package that describes – not a single number – it just describes level of confidence and describes matching algorithm.

Marc Overhage – Regenstrief Institute

Well, then how do you describe level of confidence? So red, blue, green; but that's got to translate to some kind of: “98% is green and 92% is red.” Right? I mean behind the scenes.

Walter Suarez - Kaiser Permanente

Yeah. I suppose, yeah.

David McCallie - Cerner Corporation

Yeah, somebody still has to put a number in somewhere. You actually come up with a number. And that's gonna be hard given the populations. You know I'm thinking of : You might have somebody who travels a lot and has had some encounters in remote cities where they – you know in an emergency room or something – that are important and need to be able to be located. But where they only provide a subset of their totally completed [interrupted]. And it can be fairly hard to know that those secondary matches may not score as high because they didn't provide as much information in that remote place, but it's still the same person. So even the scores aren't gonna tell you –

Marc Overhage – Regenstrief Institute

Right.

David McCallie - Cerner Corporation

It's a really tough problem. And I guess I'm anxious to say let's not over-engineer or over – you know, establish requirements that are just too burdensome to be met, unless there is just a really clear benefit of doing it. And I'm not sure what that benefit is yet.

Walter Suarez - Kaiser Permanente

And actually, that takes to me to the other point I wanted to make which is: There might be – depending on the purpose for the request – there might be differences in what I respond. So for example, if I'm requesting these for treatment purposes, I really need a perfect match or I need a – But if I'm requesting these for secondary use as research or some other purpose – operation, you know payment operation part of the TPO – then I might actually be fine with receiving a whole host of data that I can then do my own match.

Marc Overhage – Regenstrief Institute

Right. And we went through that I think early on; and we all agreed that we're gonna focus this particular set of recommendations around the direct patient care scenario, recognizing that it could be expanded out.

Walter Suarez - Kaiser Permanente

So it seems to me this is a – [interrupted]

Marc Overhage – Regenstrief Institute

Pardon?

Nancy Orvis – DOD Military Health System

Marc and Walter, this is Nancy Orvis. I'm thinking just exactly what you were saying. The credit retrieval matches. If I say I have four characteristics, and I put that out there. And then I get a list of people; and I can tell that this is where I got 100% match, and these only matched on three of the four, and these were two of the four. And you saw what the differences were on where they did not match – not the same birthday, middle initial wrong – wouldn't that be – that's kind of the algorithm I'm thinking that Walter is talking about. I've seen this. Wouldn't that be where you get your "confidence?" Cause you'd absolutely know that if you've got the right birth date and the first and last name, you've probably got – you know, or something like that. The person judging could say, "Yes. I will take this match." Because – [interrupted]

Marc Overhage – Regenstrief Institute

Shaun, I'd like you to comment. But going back: Number one, I don't think that we want to rely on human impressions here. We want to rely on some level of statistical confidence.

Nancy Orvis – DOD Military Health System

Well, I agree. The statistics will be there, but the statistics would need to tell you which things they had the 100% confidence on. That's all. You know, that says I got a 100% match on first/last name and date of birth. But I got – [interrupted]

Marc Overhage – Regenstrief Institute

What I was gonna say further: I don't think that it's a – some algorithms are going to have a level of confidence you could assign to the match. Cause it's not just saying the name matches, right? It's the name matches somewhat, and it's the combined probability that informs you. But I can see your point about – so what I hear is: For each data or patient attribute that I provided as a requestor, you're asking whether it would be reasonable feedback to get at least this element match – and again, you're gonna have to pick some threshold – above some threshold.

Nancy Orvis – DOD Military Health System

Right. And I'm thinking of all the patient matches I've done through [distortion] with CHCS in DOD. I mean, I can get 25 John Smiths real easily.

Marc Overhage – Regenstrief Institute

True. Of course.

Nancy Orvis – DOD Military Health System

And that's exactly – but I go – if I'm looking for the right medical record, I have to make some judgment calls somewhere in there. But, I'm – [interrupted]

Marc Overhage – Regenstrief Institute

Yeah. And again, we're not asking the person to make the judgment call here. We're saying the algorithms made the judgment call and said, "Here's the five patients that match." There's not one.

Nancy Orvis – DOD Military Health System

Yes. Okay, but I think – can't the algorithm come back and tell you which of the things matched exact – you know, the ones that aren't so confident? You want that algorithm to come back and tell you what did or did not quite match. And – or else, are you saying that anything that doesn't quite match is illegal to pull?

Marc Overhage – Regenstrief Institute

Correct – the latter.

Nancy Orvis – DOD Military Health System

Okay fine. Then that makes it really – [interrupted]

Marc Overhage – Regenstrief Institute

Well, "it doesn't quite" is a relative term, right?

Nancy Orvis – DOD Military Health System

Right.

Marc Overhage – Regenstrief Institute

That's the problem here. So, Shaun and David – thoughts about the idea of sort of each – that Walter and Nancy are talking about – where each data item would get a red, green, blue – or I'm sorry, red, yellow, green.

Shaun Grannis - Regenstrief Institute

This is Shaun. When you say every data item – I just want to clarify. Are we talking about assessing the relative agreement among each field, so last name, first name, gender all get a red, green, blue?

Marc Overhage – Regenstrief Institute

That's what I think I hear people suggesting. Yes.

David McCallie - Cerner Corporation

I was – this is David – I was suggesting an overall score, not field by field. Because I think that, Marc as you identified, you could have three fields with partial match; and if you overwhelm a user with six different colored lights, one per field, they'll just – they'll never even look.

Marc Overhage – Regenstrief Institute

Although, this may not be a user thing either, right? This may be just data that gets kept in the background and analyzed periodically.

Multiple Unknown Speakers

Yeah. Right.

Walter Suarez - Kaiser Permanente

We're talking about overall score per person, right?

David McCallie - Cerner Corporation

Per match. You don't know if it's a person or not.

Walter Suarez - Kaiser Permanente

Well whatever, yeah per record. I mean if I'm returning 15, I'm gonna get 15 "overall scores"; one for each of those 15 matches.

Marc Overhage – Regenstrief Institute

Well, I think Nancy was saying – backing up and saying, "Tell me whether the first name matched. Tell me whether date of birth matched. Tell me whether the other identifier matched."

Walter Suarez - Kaiser Permanente

Yeah. That's what I was also suggesting as part of this metadata per-record-match.

David McCallie - Cerner Corporation

Right. Nancy sounds like she was asking essentially for the ability to explain why you thought this was a match. "I thought this was a match because first name was a match, last was a match, date of birth was match except for flipped month and day."

Unknown Speaker

Right.

David McCallie - Cerner Corporation

And I think that would be interesting, but it presupposes the algorithms are – you know, I mean they may not work that way. And if you're doing a machine-learning approach, the machine is just gonna say, "This is the best score. I don't know why. I can't put it into English."

Shaun Grannis - Regenstrief Institute

Right. Yeah, you just have to present the neural network cause it's [indiscernible - interrupted]

David McCallie - Cerner Corporation

And that's why I was trying to kind of say, "Well maybe we can't come up with a number that a human could look at and say I know what that means." If the number comes back at 0.87 in the context of that remote system, what does that mean? You don't know. You don't know whether the score ranges from 0 to 1 or whatever. So if the system could at least say, "In my system's algorithmic experience, this is a pretty good match"; and give the relatively naïve user a clue that they were in safe territory or if they should be – you know, if possible, collect additional data and try again.

Walter Suarez - Kaiser Permanente

That's a point of discussion we had at the last call. And Marc, you were arguing that. So, you know, if I've submitted the request and I receive back 15 matches or possible matches, what would the receiver do with that data because there is no – I mean, supposedly the receiver of that data who submitted the request would not be able to do anything else with the data to try to do more matching.

Shaun Grannis - Regenstrief Institute

This is Shaun. From the confidence measure perspective, I agree with keeping it simple for the user if there is some way to do that to provide a single number. But, what I'm hearing though is that there are different factors that may go into that. And, you know we are talking field level.

Absolutely, field level factors will go into that. What I see as a value of some sort of confidence measure is – I think about: What do I know about this match when I’m going along? There are certain things that the sending query system will know. They know information about their data; at least they should know. And one of the things that we actually have for some of the matching – probabilistic matching we do – we generate a ton of metadata. We generate how closely does each field match from 0 to 1. We have a percent agreement based on different comparators. We actually do a root mean square comparison of different comparators. So, there’s lots of ways to generate metadata about how these fields behave. But the question is how it rolls up into that confidence level. And fundamentally, I think the value of providing that is that confidence level is telling you something about the data source into which I’m querying that I don’t know about. I know Indiana’s data, and we make assumptions about our accuracy within our system. Now as Kentucky can tell us something about their data. So, “Smith” may be very common in Indiana. But if I query for a “Smith” in Kentucky: If there is some way for them to convey -- “Whoa, Smith is incredibly rare here” – that influences my confidence in the algorithm. Now, I may not need to present the fact that Smith is incredibly rare; but it may factor into how I generate a confidence level. And so one of the things we’ve contemplated and done a lot of thinking about is just these sorts of: “Well, we know how discriminating Smith is our database. How discriminating is it in yours?” And of course, then it would be incumbent upon the service – the matching service providers to generate those sorts of metadata statistics. But I think those sorts of things could be useful if presented in a digestible fashion to the end-user.

Marc Overhage – Regenstrief Institute

And what I hear you saying is those type of data describing the matching are gonna be, if you will, generic to the data source and not specific to the query.

Shaun Grannis - Regenstrief Institute

Yeah.

David McCallie - Cerner Corporation

Yeah. They are functions of the data source. But my concern is: Do we want to burden every system with the requirement that they calculate a confidence score, number one? Maybe we should. And number two: How do you interpret the meaning of those confidences if you don’t know the algorithm that the remote system is using? How do you communicate that this is a good confidence versus a bad confidence – high confidence versus low confidence? Because not everyone’s gonna use the same system.

Shaun Grannis - Regenstrief Institute

Right. Yeah, it absolutely depends on your knowledge of what algorithm is being used.

David McCallie - Cerner Corporation

It’s like if I told you the speed limit was, you know 80, but I didn’t tell you if that was kilometers or miles. You know, you could get a ticket.

Shaun Grannis - Regenstrief Institute

Right.

Marc Overhage – Regenstrief Institute

Right.

Walter Suarez - Kaiser Permanente

I guess in more practical ways, an entity that receives a query is gonna probably – I mean more from a legal and again risk perspective – they’re gonna do one of two things: They’re gonna return a perfect match – whatever the definition of perfect match would be for them – saying this is the only record I have that is a perfect match because everything matches perfectly. Or they’re gonna say we have several possible matches, but I’m not gonna give you any of them. I need additional information before I can release any of them to you. I don’t think, again – now in the context of where we are with respect to regulatory restrictions and now with the accounting of disclosures that we have to make to make sure – you know, adding burden to the release of information about patients, I think people are going to be very worried about just releasing a whole host of things that say, “Well this one matches, next one matches almost perfectly, you know green/red Jello levels matches.” I think, again, what we – what would be best is perhaps in our discussions to say, “Okay, return perfect match. And perfect we define.” And then secondly, for non-perfect matches, submit a query for additional information that helps the entity that is inquiring about these people to send back additional information to help define the mismatches. I think the reality is probably gonna be there, more than anything. Now again, if we’re talking about research and academic databases and other things that would – public health analysis, sure there would be a different type of exchange in which I will send you a bulk of data that then you can decide what is a match.

Marc Overhage – Regenstrief Institute

So what I – and I wanna close this off so we can take a quick skim through the slides and have a chance for public comment. Given this discussion, what I wonder if might be a reasonable recommendation for us and what I – is number one: It would be helpful to – it may be a reasonable thing to ask anybody who responds to queries to have available some – and that has to be debated – but some descriptive information about their matching. And it’s gonna be a text blob, day one. And it may just say – you know, the first version of it is: “We used, you know, so-and-so’s product.” [laughter] Might be all it really gets to. The second thing that seems like we could recommend is – this really is a – you would like to see more there characterizing it. And it might eventually end up in a standardized set of metadata about the matching environment that says “how common is Smith” and “how common is Jones” and “how common is Suarez” that somebody could use to help infer things. But that’s probably down the road, and we don’t know what those things are yet. And so that’s work that ought to be done. So sort of a two-part direction; transparency by making at least a level of information and a contact person for it. So, who do I talk to if there’s an issue or problem or question? But to be expanded in an effort to further expand that data. Is that maybe – is that all we can say at this point perhaps? [Pause] And then we can ask Doug and – you know, whether that’s something that he wants the – more stuff to be elaborated now or that’s something that he’s okay with sort of putting on the list to do at some future date.

Walter Suarez - Kaiser Permanente

So this we’re saying to whom? To the Standards Committee?

Marc Overhage – Regenstrief Institute

This is our recommendation to the Standards Committee. Yes.

Walter Suarez - Kaiser Permanente

Oh, okay.

Marc Overhage – Regenstrief Institute

If that’s okay with folks, Judy if you could navigate us back to the Principles slide. I just wanted to quickly walk through cause they have hopefully refined – changed anyway – a couple of slides

in the deck here. And just – maybe we can briefly walk through these and see if folks are comfortable with this; the first being that we need a policy decision about the rate of false-positives, David’s important point that we need to make it easy to include additional patient attributes which may prove useful in the future, and that we need to align efforts to improve data at the time of capture – is what I was intending there – with the importance of the data for matching. So sort of, if you don’t really know date of birth year, you may capture it but at least flag it as a “not exact.” Or if you don’t the social security number or the patient doesn’t have one, don’t require a number in the field but allow a flag that says, “I asked. I tried, but there is no number available.” They kind of promoted that to a Principle, if you will, just cause it seemed like it didn’t fit where we had it before.

David McCallie - Cerner Corporation

Mark, this is David. Do we need a Princ – and maybe this is your third Principle I’m trying to [indiscernible] – Do we need a Principle about the importance of getting the data right at the initial capture? Is that what that third point is?

Marc Overhage – Regenstrief Institute

That is the intent. And we clearly need some wordsmithing on all these, but I’ll capture that idea of getting the data right as the phrase.

David McCallie - Cerner Corporation

Yeah. I wonder also – and maybe this is just out of scope – but we keep dancing a little bit around the notion of the tradeoffs between disclosure – privacy disclosure versus confidence of the match. And there’s kind of – it seems like there’s an unwritten principle here that we’re being extremely sensitive to inappropriate disclosure, but we may do so at a cost to certainty of the match. Maybe that’s just – [indiscernible -interruption]

Marc Overhage – Regenstrief Institute

Right.

David McCallie - Cerner Corporation

But somewhere in there you decide what’s more important to you.

Marc Overhage – Regenstrief Institute

Right. And we kind of said that in the first Principle, but yeah we may wanna – I hear what you’re – elaborate that more and it may be separate.

David McCallie - Cerner Corporation

I know that was a huge debate in the PCAST. You know, post PCAST recommendations about statistical matching. You know, I mean I think at some point you just have to be practical and say, “Look there’s health on the line here; and there may be some minimally disclosed – there may be some disclosure here but it’s to appropriate personnel operating in a healthcare organization, and that’s what any sane person would want to happen if their life is on the line.” I’m leery of bending over too backwards to prevent disclosure if it means we sacrifice quality – sacrifice certainty that we’re treating the right patient.

Marc Overhage – Regenstrief Institute

Right.

David McCallie - Cerner Corporation

And I don't know how to capture that. We haven't really had a lot of debate on it. We're probably – none of us – the right people to even have that debate. But it's got to fit in somewhere in terms of how you prioritize these decisions.

Marc Overhage – Regenstrief Institute

Yeah.

Walter Suarez - Kaiser Permanente

Well, what you're doing David – I think – is you're lowering the threshold of or increasing the ability to include false-positives, I suppose.

David McCallie - Cerner Corporation

Yeah, I mean what I was really kind of thinking is – we haven't talked a lot about it – But, is there some difference in behavior depending upon the context of who the user is making the query? So, for example, you know someone in the insurance company is trying to track down someone – you may be willing to disclose more or less than someone in an emergency room who's dealing with a patient, and all you know is first name and last name because the neighbor dragged them in unconscious from the front yard. And that's all they know; there's nothing else. And you may be willing to expose more in an effort to find out what's going on. I don't – I'm not sure how to characterize that, but it's a Principle-level thing that basically says privacy is traded off against the certainty of the match based on our willingness to expose or not expose PHI.

Walter Suarez - Kaiser Permanente

Based on – well that's the critical part of the purpose.

David McCallie - Cerner Corporation

Right.

Walter Suarez - Kaiser Permanente

So based on purpose, in some instances – based on purpose of the request, in some instances it's okay to send more information because the life of a person may be at risk if it's for treatment and it's an emergency situation. But if it's for payment or if it's for some operation, some research, or some other things, there might be other –

David McCallie - Cerner Corporation

Right. Yeah, no I'm just not sure all that is captured in the notion of a sensitivity-specificity tradeoff. It's sort of: What's the setting in terms of what data elements are you willing to expose and use as part of your match? That may vary as well.

Walter Suarez - Kaiser Permanente

Yeah, I suppose another Principle would be “purpose” is an important factor in determining specificity and sensitivity levels.

Marc Overhage – Regenstrief Institute

Yes, but again we're focusing here on the one-use case, which is the healthcare provider. We tried to scope ourselves a little bit. [laughter]

David McCallie - Cerner Corporation

Maybe that's one of – maybe that should be on the – [indiscernible - interrupted]

Marc Overhage – Regenstrief Institute

So, part of what I think we need – I'm jotting down some of this discussion, but I'll need your help as we turn this into a letter. I think this is part of it that will obviously get elaborated a lot more, and we'll have a chance to add some language that clarifies.

David McCallie - Cerner Corporation

Okay.

Marc Overhage – Regenstrief Institute

If I may, in the interest of time – but I think this is, as you said, an area we're gonna come back and spend some more time on in terms of turning this into a letter – see if we've got the core of these next things right. If we can move to the diagram slide. I didn't change that at all. The next slide, the “Matching Fields” slide, I did some considerable surgery on. And this will generate some discussion I'm sure. It was suggested that in the core we might want to include “zip code” because it's relatively available and – although, certainly not always. And I struggled with this issue of – you know, we spent a lot of time talking about the temporal information and so on. And I – it seems we've got to allow or provide – and I'm looking back at the Metadata Group's letter. They sort of said the same thing, which is you need some other identifier. And they were crafty in their language, and we may want to borrow their crafty language. But, when you're querying for data where the patient's last contact with the provider may have been more than a year or so ago, it seems like you're starting to lose with not including the zip code or the social security number so much sensitivity, with 15% of people changing zip codes. It's 12 1/2 to 15% every year, according to the Census Bureau. You start to lose so many people. So it seems like you've got to include something like that to keep the sensitivity up.

David McCallie - Cerner Corporation

What about – we had a little bit of a discussion – this is David again – about cell phone number or phone number? I mean I've been hearing people talk about cell phone numbers being the best match; the most likely number that stays constant just because people have such a strong incentive.

Marc Overhage – Regenstrief Institute

But the data that we can find says a 30% per year turnover.

David McCallie - Cerner Corporation

In cell phone numbers?

Marc Overhage – Regenstrief Institute

Yeah.

David McCallie - Cerner Corporation

Yeah, that's not – [interrupted]

Marc Overhage – Regenstrief Institute

No, it's terrible data. But, it's from telemarketers.

David McCallie - Cerner Corporation

They don't get cell phone numbers.

Marc Overhage – Regenstrief Institute

Well, some of them – [laughter]

David McCallie - Cerner Corporation

Yeah, but very few. That would be a really bad sample.

Walter Suarez - Kaiser Permanente

Well, that's fine. But, I – [interrupted]

Marc Overhage – Regenstrief Institute

I agree. It's crummy data about turnover. But – and the other question that some folks had some feedback about: We have to be thoughtful about some of the disadvantaged populations that perhaps move more, use prepaid cell phones – you know, where we've got to be thoughtful and not disadvantage some of those folks too.

David McCallie - Cerner Corporation

Yeah, I know. These are all “if available.” I mean there might not be a zip code either if they don't have a home. You know, but I just – every system I know captures a lot more than this.

Marc Overhage – Regenstrief Institute

Well, and you can provide other attributes. But the other question David is: It may capture more, but the matching systems don't necessarily capture or use it. But they're [indiscernible], right? I mean that's your point of we need to allow, as we've outlined here, that we shouldn't inhibit other attributes that are defined in HL7. In other words, anything that HL7 thought of to describe a patient we should allow to be sent. And then we're not excluding attributes here, and you can send the cell phone number.

Walter Suarez - Kaiser Permanente

Marc, why is the street address, city, and state optional instead of core. I mean I would think that that should be a core, except it's only if they don't have it they don't send it. But – [pause]

Marc Overhage – Regenstrief Institute

We could certainly promote it. It's not very useful in matching – if you get the zip and what does it add beyond the zip.

Walter Suarez - Kaiser Permanente

Sure, okay.

Marc Overhage – Regenstrief Institute

Shaun, do you have thoughts about that one?

Shaun Grannis - Regenstrief Institute

Which – I didn't hear which field was mentioned.

Walter Suarez - Kaiser Permanente

Street address, city, and state.

Shaun Grannis - Regenstrief Institute

Yes, well I don't see street address used consistently because it's so highly variable and it has to be preprocessed and parsed, and most systems just don't have that today. What one of the questions this raises for me – and I think this relates to the cell phone number as well – is: I think as technology improves, I think cell phone may become a better identifier in the future. Is this group going to make any future, you know leaning recommendations regarding these fields – [interrupted] – may become better in the future?

Marc Overhage – Regenstrief Institute

Good point. That's a good point. So we may include cell phone and a – as a leading indicator, if you will, in the example of optional attributes.

Shaun Grannis - Regenstrief Institute

Yeah.

David McCallie - Cerner Corporation

That would be good. I mean in the notion of future – I mean, you know NIST is trying to get us all to go get trusted cyber identities.

Marc Overhage – Regenstrief Institute

Right.

David McCallie - Cerner Corporation

So presumably, for those people that are willing to go do that and aren't afraid of it, for whatever reason, that would be a powerful matching token.

Marc Overhage – Regenstrief Institute

For those who have – I mean you got to have it on both sides though is the interesting thing. So it'll be a long decision.

David McCallie - Cerner Corporation

I agree. We're talking about future. I mean – [pause]. I mean, in other words, how could you make this better without establishing a National Patient Identifier? Since that is not on the table, one of the things you can do is leverage some of these emerging identifying tokens that are – have less turnover than “address.” But that's where – you know, after you get past date of birth, the turnover starts to go up pretty high.

Marc Overhage – Regenstrief Institute

Yes it does. Let me – and I venture Judy you may have some guidance here. I'm watching the clock. I've one or two other things that I wanted to at least highlight. What I think we do next with this is begin to craft a letter where we can have – as Walter loves to do – some good, careful, thoughtful wordsmithing to make sure we've kind of captured some of these things and where we've got some verbiage to support or explain these things. Maybe what we do here is – let me just highlight these two other things so that folks will be sensitized to them. Then we open up for public comment. And then I think our next step here is to draft a letter version out of this that we can then circulate to the committee for sort of bilateral feedback. You know, sort of gather everybody's feedback. And then decide from there whether we need a follow-up meeting to finalize it. Does that sound okay Judy?

Judy Sparrow – Office of the National Coordinator

Yeah. That's very reasonable. Yeah. I think so.

Marc Overhage – Regenstrief Institute

So on the next slide – the Data Quality slide – the one thing that – the only thing that I added or changed here I think materially is: I’m looking at the Privacy and Security Tiger Team letter. They had this notion – and it struck me as reasonable when on Data Quality – that we would encourage (and we’ll have to figure out the word here) methods to allow the patient to check the entry, such as sharing the entry screen with them, a printed summary, or online access. Since, presumably, they might be more able and likely to identify errors or say, “Gee, no I actually don’t use that phone number anymore” type thing. Seems like a reasonable Data Quality process to at least encourage to be thought about. And then the – one of the – on the Data Formats/Content slide, following along with the good work that was done by the Metadata Group, CDA R2 header formats are representing patient attributes is what they have been recommending and seems to cover representation of even complicated name structures for example, as well as accommodating some of the uncertainties in dates and some of the other things we’ve talked about. And so it seemed a both coherent and reasonable recommendation. I think we still have some thinking about – there was some public feedback and some other questions that folks raised about – as we think about sort of what the exchange or protocols might be like; whether the IHE PDQ Implementation Guide versus the NwHIN Patient Query Implementation Guide as a starting point. And again, I don’t think – unfortunately, we don’t have time on the call to work through that, but that’s something that we need to come to some grips with I think. So with that, Judy let me ask you to open up for any public comment or questions. And then, we’ll try to sort of wrap this up.

Judy Sparrow – Office of the National Coordinator

Okay, thank you Marc. And Operator, can you check and see if anybody from the public wishes to make a comment?

Operator

Yes. If you are on the phone and would like to make a public comment, please press *1 at this time. If you are listening via your computer speakers, you may dial 1-877-705-2976 and press *1 to be placed in the comment queue. [Pause] We do not have any comments at this time.

Judy Sparrow – Office of the National Coordinator

Okay, thank you. Marc, thank you.

Marc Overhage – Regenstrief Institute

Great. So, I will take a crack at drafting this into a letter. I will, I am sure, in that letter highlight some areas for a number of you to expand or contribute more, you know in a focused way. And like I said, I think if we could – we’ll see how that goes. And if necessary, I suspect we could schedule an administrative call in mid-to-late July to discuss items there if we feel like we need to. But hopefully, we can – we’ve had enough [indiscernible] from tremendous discussion – we’ve had enough good discussion that we can get most of it done. I think – we’re gonna say I’m hoping. Is that agreeable to everybody?

Multiple Unknown Speakers:

Yeah, that would be great. Yeah.

Marc Overhage – Regenstrief Institute

I’m sure you all want another call. [laughter]

Judy Sparrow – Office of the National Coordinator

Right.

Walter Suarez - Kaiser Permanente

Yeah, that would be great Marc. Could I send you a suggestion of some – I think the letter needs to cover briefly the description of the Use-Case, briefly also description of the Assumptions, and then the description of the Principles. And the Recommendations which I see are like in four categories: Data Elements, the Query Format and Content, the Data Matching, and the Data Quality. And then some needs from like the Policy Committee and others to define; for example, provide definition of match, the confidence level, the false-positive levels. Things like that. So – I'll send you an email with some thoughts about that.

Marc Overhage – Regenstrief Institute

Great.

Judy Sparrow – Office of the National Coordinator

Okay, thank you.

Marc Overhage – Regenstrief Institute

Thanks everyone.

Judy Sparrow – Office of the National Coordinator

Happy 4th everyone. Happy 4th.

Nancy Orvis – DOD Military Health System

Thank you. Take care.

Walter Suarez - Kaiser Permanente

Bye-bye.

David McCallie - Cerner Corporation

Bye.

Shaun Grannis - Regenstrief Institute

Goodbye.