

Vocabulary Task Force
Draft Transcript
June 27, 2011

Operator

All lines are bridged.

Judy Sparrow – Office of the National Coordinator

Thank you, operator. Good morning everybody, and welcome to the task force on vocabulary. This is a federal advisory call, so there will be opportunity at the end of the call for the public to make comment. Just a reminder too; workgroup members please identify yourselves when speaking. I'll do a quick roll call – Jamie Ferguson?

Jamie Ferguson – Kaiser Permanente

Present.

Judy Sparrow – Office of the National Coordinator

Betsy Humphreys?

Betsy Humphreys – National Library of Medicine

Present.

Judy Sparrow – Office of the National Coordinator

Clem McDonald?

Clem McDonald – National Library of Medicine

Present.

Judy Sparrow – Office of the National Coordinator

Stuart Nelson?

Stuart Nelson – National Library of Medicine

Present.

Judy Sparrow – Office of the National Coordinator

Asif Syed?

Asif Syed – American Medical Association

Present.

Judy Sparrow – Office of the National Coordinator

Chris Chute?

Chris Chute – Mayo Clinic

Present.

Judy Sparrow – Office of the National Coordinator

Marc Overhage?

Marc Overhage – Regenstrief Institute

Present.

Judy Sparrow – Office of the National Coordinator

Dan Vreeman?

Dan Vreeman – Regenstrief Institute

Present.

Judy Sparrow – Office of the National Coordinator

Floyd Eisenberg? Donald Bechtel? Patty Greim? Jim Walker? Chris Brancato?

Chris Brancato – Office of the National Coordinator

Present.

Judy Sparrow – Office of the National Coordinator

Andy Weisenthal?

Andrew Wiesenthal – IHTSDO (SNOMED)

Present.

Judy Sparrow – Office of the National Coordinator

Bob Dolin? Ran Sriram?

Ran Sriram – National Institute of Standards and Technology

RAN SRIRAM: Present.

Judy Sparrow – Office of the National Coordinator

Ken Gebhart?

Ken Gebhart – National Institute of Standards and Technology

I'm here.

Judy Sparrow – Office of the National Coordinator

Lynn Gilbertson?

Lynn Gilbertson

Good morning.

Judy Sparrow – Office of the National Coordinator

Morning. Nancy Orvis? Anthony Oliver? Marjorie Greenberg? Did I leave anyone off?

Doug Fridsma – Office of the National Coordinator

This is Doug Fridsma.

Judy Sparrow – Office of the National Coordinator

Doug -- good morning. Alright, with that I will turn it over to Jamie.

Jamie Ferguson – Kaiser Permanente

Good morning everybody. Thank you for coming on the call with a relatively short notice. I think just about everybody on this call was on the call Friday, which we cut – well I mean we took the full amount of time, but it seemed that we were cut short, so we wanted to extend the conversation. We are talking with Ken Gebhart from NIST about the testing and certification of vocabulary subsets and talking about the functionality for testing those subsets; and we are starting with the LOINC subset discussion. So Ken do you wanna -- can I turn it over to you and you just sort of recap where we were.

Ken Gebhart – National Institute of Standards and Technology

Sure, I'd be glad to. So if we could get our slides up from the other day. While the slides are coming up, let me just mention that this conversation started in the early June Vocabulary Task Force meeting where there were some interesting thoughts about how some of the requirements being discussed could be tested. And so this slide deck basically represents some initial thoughts we had about the testability if you will of the requirements that were being discussed in that earlier call. If we could go to –let's see slide four, we talked for a little while about requirements on the last call that we heard discussed earlier. I think we generally had agreement during our call at the end of last week that these are the requirements, with Doug's friendly amendment that perhaps specific to lab tests, R2 might be an unknown code not just an unknown LOINC code. And if I recall correctly, Doug suggested that for example a lab vendor might send a proprietary code, and we'd want to make sure that the EHR was still capable of handling that reasonably without breaking. So I think we had -- at least my impression we had sort of general agreement that the R1, R2, and R3 seemed to make sense based on the discussions we have had earlier. Like anything else, the farther you dig into it the more you find use cases that either need a little bit of tweaking to requirement or represent new dimensions to it that hadn't been considered more generally. Also comment, if we go on to slide 5, that the thoughts that we have at this point are about what we generally call "black-box testing", in which we put inputs into an EHR, we examine the outputs either visually or electronically. Contrast this to what's generally called "white-box testing", where you might actually examine things like – pull open the master file table within the EHR and examine it to verify that all the expected codes are present, or you might review technical documentation of the technical architecture and data flow to verify that data coming into the EHR is stored appropriately and processed appropriately. Or you might even examine, in white-box testing, the data model that's inherit in the product to verify where white codes are stored and verify in fact that they are stored there. White-box testing is not part of today's conversation. Generally speaking, we stayed away from that in stage 1. And so the conversation at this point is mostly about the idea that it's black-box testing, where we provide inputs to the EHR , it does what it needs to do with it; and then we look at the outputs. So on slide 5, where we seem to be wrapping up the conversation the other day was question of which of these possibilities or are there other possible functions that would be tested in order to verify R1 that an EHR be capable of performing computable activities and receive LOINC codes. I don't know that we've reached a conclusion about this, and Jamie perhaps you'd like to state where you thought we were at the end of that conversation.

Jamie Ferguson – Kaiser Permanente

I'm sorry – what was the particular question?

Ken Gebhart – National Institute of Standards and Technology

I'm sorry Jamie -- I didn't mean to catch you off guard. So in R1 on slide 5, we were talking – I think you and Doug were talking about which of these capabilities or some other capability might be appropriate to....

Jamie Ferguson – Kaiser Permanente

Alright, well there was a question I think that was unanswered about the flowsheet capabilities of essentially all the certified products.

Ken Gebhart – National Institute of Standards and Technology

Right -- there you go.

Jamie Ferguson – Kaiser Permanente

And I think that's probably about where we left off. So I think there was general agreement that the ability to fire some decision support rule on a received white code was a good testable function, but I think the question that was unanswered was about the flowsheet capabilities. I mean I know certainly some of the products, but by no means do I even know a majority of them.

Clem McDonald – National Library of Medicine

There is a comment on the problem with the reminders -- [clears throat]excuse me, I'm still stuck with this pharyngitis -- is that if you are going to send random codes, which was part of the proposal, you may not hit any that have a reminder on it. So if you want to send random codes, I almost think you have to think flowsheet. I mean you might do both, but just keep in mind that the idea of sending random codes won't work very well as reminders cause there are not a million codes that are going to hit against the reminder.

Judy Sparrow – Office of the National Coordinator

Well I suppose you could do a randomization of the smaller set. You would trigger something.

Clem McDonald – National Library of Medicine

I think if the smaller sets -- well it's not a huge set. I don't think a test that would be used in the reminders -- in cholesterols and maybe blood gases. It would be worth maybe examining that, and you might send them all then.

Ken Gebhart – National Institute of Standards and Technology

I think you're right. We'll -- maybe this is a good point to jump to slide 6 because this gets us into sort of how to generate test data and both comments that just came up are certainly valid ones for this conversation. So looking at slide 6, I wanted to make a couple comments. First of all, what I'm showing of slide 6 is what we generally call an inspection testing approach, where you provide inputs to the EHR and then you actually examine an output that shows up on a screen or a printout. On the slide after this one, we'll talk about automated conformance testing, where you're actually asking the EHR to output some message or document and examining it for appropriateness. But this is ... [interrupted]

Unknown Speaker

[indiscernible] capture the order itself?

Ken Gebhart – National Institute of Standards and Technology

I'm sorry?

Unknown Speaker

Is [indiscernible] gonna capture the order itself. This is the result we are talking about.

Ken Gebhart – National Institute of Standards and Technology

I think that would be certainly a possibility. I think what's tricky in this one is -- I think -- let me offer this thought. I'm gonna hold onto the order thought. It will come up in the very next slide, and the question about orchestration of sort of the typical series of events is gonna be more broadly a conversation we need to have across a lot of testing. Let me just get through a few pieces of this slide, and then we can

sort of piece together all those thoughts on slide 7. So visual examination of the EHR – so in some fashion, we would use a test message generator that generates a lab message or a series of lab messages, as has already been mentioned here, depending on what we're actually testing that might be okay to be strictly random. It might also be that it's random within subsets or specific sets based on trying to hit a rule or reportable lab results for example, or something else.

Andrew Wiesenthal – IHTSDO (SNOMED)

Can I interject something here? This is Andy Wiesenthal. And the reason I'm -- just to disclose, I've been CCHIT commissioner for four years and now I'm a [inaudible - static], but that's the kind of question that does [indiscernible] whether you do, you know, subscribe to the way CCHIT does it or someone else. Now, I'm not sure what's happening here; but if we're reinventing testing, that's not necessary.

Ken Gebhart – National Institute of Standards and Technology

I'm sorry. I don't quite know what to do with your comment.

Andrew Wiesenthal – IHTSDO (SNOMED)

Well and I'm not sure – I'm sorry for that – I'm not sure what you're trying to address, and perhaps I'm just misunderstanding this conversation; but you're talking about how to generate test data and what kinds of codes to send and how to examine and test, and the testing organizations are doing that. And there are standards for that. There are approaches for that. They have test data. So either I'm confused – I'm sure I'm confused, but I'm not sure where, you know, why we are talking about -- it seems to me we are talking about creating something new when there is already a process that can be iterated in this time.

Doug Fridsma – Office of the National Coordinator

This is Doug Fridsma. You're right. There are, you know, CCHIT has a long history of doing testing and the like. However, CCHIT is not the only tester out there right now. There are at least five other organizations.... [interrupted]

Andrew Wiesenthal – IHTSDO (SNOMED)

No, no I understand that Doug. [indiscernible –speaking while Doug is speaking]

Doug Fridsma – Office of the National Coordinator

What we have to do is we have to make sure that there is consistency across the testing. Whether we leverage work that's already out there or not, I think we have to come to a common understanding of how that happens. And in the [indiscernible] legislation and the HITECH Act, NIST was sort of charged with making sure that the testing scripts were consistent across all of the certifiers. So that's the conversation that we're having.

Judy Sparrow – Office of the National Coordinator

So looking at this requirement potentially for stage 2 and stage 3, and whether we have the tests that match those.

Ken Gebhart – National Institute of Standards and Technology

This is Ken. I think we're -- you know, this will take a little while to get all on the same page. I think there are nuances to testing vocabularies that may be a little different than what's been done historically; but I certainly respect the fact that there is a history, particularly of generating lab messages and sending them to EHR and validating them. A couple things that I wanted to catch here; what standard or implementation guide are we gonna use for generating these messages. We know that the S&I framework is working on an implementation guide right now. So we'd need to settle -- is that the basis

for generating these messages or is some other standard or implementation guide gonna be allowed? You'll remember in stage 1, there was no particular implementation guide defined for this. There is this capability defined as a requirement in stage 1, but it was only for structured data. And then on the right side here, we are back to the same question we were talking about on slide 5 of what are we actually evaluating. So we won't belabor that point again, except to say that both in terms of how many messages get generated and how many get inspected visually, you know it depends on what target we are trying to hit here in terms of the functional capabilities on the right side. And remember that that list is just a teaser, if you will. Based on the conversation we had for slide 5, there may be other things that you conclude are the appropriate things to evaluate for computable activities. So this is sort of one way to look at it being done, where messages go to the EHR and then a human being actually looks at what the EHR did with it. We could go then to slide 7 – we'll talk about an approach that will probably look familiar to some of you, where we're actually stitching several requirements together in an automated testing fashion where we generate test messages somewhat the way we talked about a few minutes ago. The EHR receives these, and then the EHR is outputting a message that's also required by meaningful use. For example, a lab message to public health or a patient's summary record or a clinical quality measure report. And then a conformance testing tool; that's a piece of software that evaluates that output message to verify that it's conformant to the standard, for example a public health lab message standard, and includes the expected LOINC codes, patient's summary record including the LOINC codes we expected, and then the clinical quality reporting standards. This is mostly about the inclusion/exclusion criteria that are in the e-measures. So to make this work, the test message generator not only generates messages but it sends what it created over to the conformance tool so it actually knows what to look for. This possibility tends to follow more of a logical flow that you might expect in the real world, where lab messages are coming from a lab system (think of that as the test message generator) going to the EHR, the EHR is evaluating them based on rules, and then producing output messages, again based on the standards. This is the place where we would want to pick up that question about lab orders, because the way that EHRs probably work best, if you will, in terms of being able to manage this process, is if they have initiated an order so that when they get this message back from the lab or the test message generator, there is information in it that tells the EHR how to link the result with the order. So a refinement of this might be to stitch one more requirement together which is lab order-entry and start this testing scenario by having the EHR send a lab order to a test message generator and the generator send back a test message result. That sort of gets us into a stickier place about randomization and some of those conversations, but it's things that we could certainly explore further as this matures and our understanding of the requirements matures.

Jamie Ferguson – Kaiser Permanente

This is Jamie. I wanted to just comment, since you're talking about orders. There isn't a standardized order message that has been agreed upon, nor is there one that matches either the former or current ONC efforts at resulting or results delivery. So I think it might be problematic to do that. And one of the problems with going to the order is that you get into the area of the policy question really of whether it's desirable to have EHRs have the capability of both ordering and receiving results from multiple labs or just from one lab. And you know basically that's one of the differences between the different results specifications that are out there.

Ken Gebhart – National Institute of Standards and Technology

Good point. And this is sort of some of the creative tension we've seen all the way through meaningful use Stage 1, where people have repeatedly suggested that we needed to integrate things to represent a more complete clinical workflow, as you just pointed out, but some of the pieces may not be there to make it work yet.

Clem McDonald – National Library of Medicine

I think there is another reason to not count the orders too soon. In that, you'd like to be able to get results you didn't order. So I mean it's quite common for patients to go to the ER, and those results are sent by paper back to the office who is responsible for the patient. It would be really nice to get those electronically. Or patients in the hospital, you know to get the results from the hospital patient.

Jamie Ferguson – Kaiser Permanente

Right. And that's one of the, I think, key differentiators between the different result specs that are out there is that the previous ONC effort included the copy two list for the nonordering providers to get the electronic result.

Ken Gebhart – National Institute of Standards and Technology

Sorry. Did I cut somebody off?

Clem McDonald – National Library of Medicine

Well I was just thinking aloud that you've got some challenges and that maybe it would be useful, probably for you and those people who are doing testing, to have some dialogues offline just to see what the strategies are. To me it would seem you'd like to find a way to send all 2000 codes, if that's the subset that's focused on, and find a way to machine-based read, if they got them and interpret them right, by getting an output on the other end rather than having to have humans read over all of it. Is that imaginable in anybody's mind?

Ken Gebhart – National Institute of Standards and Technology

It's certainly one of the things that we've discussed internally here. It's highly desirable. I think we have to. In terms of understanding how to do that with meaningful use, we need to figure out what is an output of an EHR that's required for meaningful use.

Clem McDonald – National Library of Medicine

Well, I don't know that that's required to test the requirement to receive and store them. So I think when it's talking about the [inaudible] purpose, but is it required that you have to demonstrate through one of the meaningful use mechanisms. I mean, one way that might say send it all out as though it's sending it to a public health but just show you got them. Because we're gonna be getting a pinch at one end or the other end, and you're not -- it will make it a lot harder.

Nancy Orvis

So Clem I hear you saying that every receiver needs to have the base of 2000 lab tests to be able to take any of those.

Clem McDonald – National Library of Medicine

I'm not sure I wanna say that, but it depends on what the -- I thought, I mean it's possible that the minimum will be that they should be able to manage those 2000. That's not for me to decide. But whatever the subset is chosen, why don't you send them all, find a mechanism that's available to send something somewhere, and ask them to send them all out to that mechanism so you can do it with automatic message than have people reading screens.

Asif Syed – American Medical Association

Clem this is Asif. Most of the ordering transactions are based on CPT and the results are based on LOINC.

Clem McDonald – National Library of Medicine

We're not talking about ordering. I'm saying if this testing process is gonna say – say there are 700 or 2000 or 200 whatever it is, these tests are supposed to be results by lab systems and understand the LOINC codes. Why not just send them a set of all of them, find a mechanism in the EHR which is convenient for emitting them back out. This is speculation; but one might be delivery to public health, even though they don't need them all. But you have that machinery, and then you could just machine it and get it done. With anything else, this can be very, very, labor-intensive. But that's just a thought. Or put them into a CCR, I mean whatever the thing is their supposed to have available anyway, but not limit the set on the input and the set on the output to the same exact LOINC codes.

Ken Gebhart – National Institute of Standards and Technology

I think that's good input. I think we have a little bit of that on the right side of slide 7, but the adjustment that you'd be talking about is sending everything instead of just randomly send a subset.

Clem McDonald – National Library of Medicine

And then find some machinery that they can dump it back out and see that they've actually understood it and returned it.

Ken Gebhart – National Institute of Standards and Technology

Yeah. So one of the things that we had kicked around here at NIST, just as a way to think about this, was generate a hundred lab messages for ten patients and then ask the EHR to create the ten patient summary records and have the conformance tool verify not only that we got the ten summary records but that the LOINC codes are in the right summary records.

Clem McDonald – National Library of Medicine

Yeah. That would be one good way, I think.

Ken Gebhart – National Institute of Standards and Technology

And you could suggest that you do a 100% of them. And we're [interrupted] ...I'm sorry, go ahead.

Nancy Orvis – DOD Military Health System

I think I would like us to explore that a little more too, because having – this is Nancy Orvis. I've done that kind of testing before with [inaudible - distortion] on the drug side. Like I wanted to take my top 2000 formularies and put it through a test. And that was quite effective. I think that number is because you're not quite sure of which [inaudible - distortion] .

Ken Gebhart – National Institute of Standards and Technology

Nancy, I got the first part of that which was that you've done the same thing with the top 2000 meds I believe; and then your cell phone started to break up. But it sounded like you were being supportive of this idea. Is that right?

Unknown Speaker

Maybe her phone broke up altogether.

Ken Gebhart – National Institute of Standards and Technology

I think we lost her. Is that what other people heard?

Clem McDonald – National Library of Medicine

That's what I thought she would have said if we heard it all.

Ken Gebhart – National Institute of Standards and Technology

Yeah, okay.

Clem McDonald – National Library of Medicine

I mean the thing is the only reason you are doing random is cause you got labor in between, and you didn't want to have to do a lot.

Ken Gebhart – National Institute of Standards and Technology

I agree. I think, you know, what becomes the challenge is when we extend this conversation down another slide or so to other vocabularies, then send-all may not be really practical.

Clem McDonald – National Library of Medicine

Well I think generally they're looking at subsets, and maybe make a tighter subset based on frequency of those that are so large it wouldn't be practical to run the test.

Ken Gebhart – National Institute of Standards and Technology

Alright.

Clem McDonald – National Library of Medicine

But random makes it harder in some ways.

Ken Gebhart – National Institute of Standards and Technology

It does.

Andrew Wiesenthal – IHTSDO (SNOMED)

This is Andy. I'm with Clem. I think you can identify or have high-users identify those vocabulary codes that are used all the time and make sure that those can be sent and then dumped back out in some output form. You know, there may be 50,000 diagnosis codes; but you really only need 600 and maybe a smattering of the others at some random basis. Any time you can get people's eyes out of the middle of the testing process, you are much better off. I'm just stating this, given that we all understand.

Ken Gebhart – National Institute of Standards and Technology

So I will echo what I believe I've heard on a few of the vocabulary calls recently, which is that this group is trying to identify subsets of vocabularies exactly for this purpose -- for meaningful use stage 2. Is that a fair statement?

Clem McDonald – National Library of Medicine

All I know is I've heard the subsets described in terms of the SNOMED subsets and the LOINC subsets. And I don't know that it was specifically for testing purposes, but if there are going to be proposed to be some minimum, you would like to test them.

Ken Gebhart – National Institute of Standards and Technology

That was the assumption I was working from as well.

Betsy Humphreys – National Library of Medicine

This is Betsy Humphreys. I think that really the subsets that we have been working on, which many of which are frequency-based or public health requirement-based or whatever -- I think our first purpose in developing these was actually to provide tools that would facilitate adoption of the standards; get people to focus on the things that were likely to matter the most to them. But also, I think that certainly Doug and others have said that this would also serve as a useful basis for testing.

Ken Gebhart – National Institute of Standards and Technology

Okay. So I'm just gonna make one other comment on slide 7 -- I appreciate everybody's thoughts here -- and that is that in terms of actually generating lab result messages, there is more to it than just plugging a LOINC code in, so NIST and whoever is developing this would need help from the industry from stakeholders to develop the test data. We've had experience, speaking of NIST, for meaningful use stage 1 of developing a small, very small set of test data for labs including LOINC codes, and found that there is -- we got a lot of feedback. I'll just say that it demonstrated that we're not the domain experts, and we should defer to the domain experts to do that; not attempt to become them ourselves. So getting the test data right is a key part of all of this testing, and I would want to make sure that the right people are involved in that process. I wanted to ask one other thing. It's down in the lower left corner here. I mentioned it a minute ago. I just want to make sure we understood the comment. We need a standard and an implementation guide to generate these messages. I am making a general assumption, but I don't know whether it's appropriate; and that is that the work that's underway in the S&I framework right now for the lab results interface implementation guide is the target for meaningful use stage 2. I don't know if I need an answer right this second, but I'm just trying to make sure we understand that there has to be a decision about that to drive this.

Unknown Speaker

Doug can you comment on that.

Doug Fridsma – Office of the National Coordinator

Sure. I think one of the things that we want to get clarity around is sort of improving our implementation specification. From a standards perspective, there really is no change from meaningful use stage 1 and the work that's going on within the S&I framework around electronic lab interfaces. It's still the HL-7 251 standard there. What is different has to do with the implementation specification. So to the degree that the implementation specification will help us I think with testing to be a little bit more -- a little clearer in what's expected and probably get us a little closer to what we'd like to see with interoperability. There is that relationship there. One would hope that the work that happens in the S&I framework around electronic lab reporting can feed in and make our testing scripts a little bit better. We certainly should be able to leverage the works that's going on within that team to kind of get their input and their ideas about how best to take that implementation guide that will be coming out of that and use that to help support testing.

Ken Gebhart – National Institute of Standards and Technology

Yeah. I think Doug the pragmatic question -- again, not trying to put you on the spot -- is whether that's gonna be a standard specified for meaningful use 2.

Doug Fridsma – Office of the National Coordinator

Well, that's one of those things that I can't really say necessarily, but clearly the activities that we have within the S&I framework are aimed at trying to create better implementation guides and things like that. Ultimately, it will be the HIT Standards Committee that will review that material and decide whether or not to recommend going forward with that or getting sort of broader input. But, you know, the issue that we have there is that the implementation specification that HITSP developed as part of the HL-7 251

standard was relatively underspecified in the sense that it had a tremendous amount of optionality; there were a lot of variability that was possible there. That makes it hard to test. It also makes it hard for that test to assure that when somebody was to send a message that they would have interoperability. So, you know, if we can create additional constraints and more clarity around what's expected, both in terms of what we expect people to send and what we expect people to be able to receive, that's gonna make, I think, the testing a little bit easier. And so once we have some of that information out of the S&I framework activities, and we anticipate that just in the next couple of weeks, we should probably have something. It's one of our summer camp activities. Once we have that sort of feedback, I think it will make it a lot easier for us to then sort of roll that in, have it evaluated by the HIT Standards Committee, and if necessary then roll that into our testing strategy.

Ken Gebhart – National Institute of Standards and Technology

Right.

Clem McDonald – National Library of Medicine

One other thought. The testing of being able to take – I don't wanna say this wrong – you could construct a test to see that people receive LOINC values and could report them back out more easily than you could construct a test for all of the HL-7 functionality. So I think you wanna be careful that you don't have to do everything to do something. And the second thing is that there is a – I don't know if you're familiar with the [MIRF] interface tool which can be used to generate stuff, and you could generate messages out of it fairly easily. We have a copy. We know how to use it if you're interested in that space.

Ken Gebhart – National Institute of Standards and Technology

I think both are good points. I think one of our challenges we be whether we want to – and this will come to the way that ONC writes the criteria. I don't want to put anybody on the spot today about that. But, if there is a specific requirement that describes specifically validating LOINC codes and it's separate from generating or receiving lab messages, then doing what you suggested makes sense. If it's embedded in the requirement for receiving and processing lab messages, then we pretty much need to do it all together.

Clem McDonald – National Library of Medicine

Well I wasn't suggesting you could test them without an HL-7 message. It's just you wouldn't have to test so many different things.

Ken Gebhart – National Institute of Standards and Technology

Yeah, I'm with you.

Clem McDonald – National Library of Medicine

We haven't talked about [interrupted] ...

Jamie Ferguson – Kaiser Permanente

This is Jamie. I just wanted to say from the perspective of a lot of the current discussions in the HIT Standards Committee, the committee is very hesitant to recommend for adoption and a final rule any standards that have not been basically proven in a number of real-world implementations. And so a new implementation guide specification would fit into that general direction from the Standards Committee. However, a subset of LOINC, which is sort of widely used, I think would be easy to pass a recommendation out of the standards committee. So in other words, from a Standards Committee perspective, we might end up having the ability to recommend a vocabulary subset for immediate

adoption in stage 2; but I think it would be unlikely for the committee to recommend a particular implementation guide for anything other than piloting if it hasn't already been used.

Clem McDonald – National Library of Medicine

That would be discouraging.

Doug Fridsma – Office of the National Coordinator

Jamie, I think the challenge we will have, of course, in the HIT Standards Committee is that both of the standards that are under consideration in ELR have been widely disseminated and used. The question really is -- one is less constrained and therefore less likely to get us to our goal of interoperability; and the one is highly constrained and very, very specialized but has use in California for example.

Jamie Ferguson – Kaiser Permanente

I understand. And that one is constrained to a single EHR in a single lab, because it uses the proprietary local codes. But the issue that I'm talking about is from a Standards Committee perspective. I'm just reflecting Doug on the recent discussion around the metadata recommendations, where what seemed like otherwise reasonable recommendations received a lot of pushback because they hadn't been vetted through a specific real-world exercise.

Doug Fridsma – Office of the National Coordinator

Right. And I think the good news is that much of the work around laboratories and vocabularies, the standard is a well-established standard. The implementation guide actually has been used fairly broadly. The difference between what's coming out of the S&I framework and those things that are already implemented, you know, it isn't probably the same as what's come out of the metadata report. So I think those are the kinds of conversations that will be important in the HIT Standards Committee. But I think clearly when it comes to the vocabularies, the subsets, those sorts of things, there has been fairly broad use in adoption by LOINC and HL-7 251, both in its HITSP implementation guide as well as in the [\[E-links\]](#).

Ken Gebhart – National Institute of Standards and Technology

Okay, this is Ken. I'm gonna try one more thing to explore here on slide 7, just for a second; and that's the notion of clinical quality measure reports. This occurred to us as we were looking at this overall and saying, you know, there are a number of measures that use LOINC codes as inclusion or exclusion criteria. And so one of the things that would be worth exploring here at some point would be does validating the EHRs ability to produce clinical quality reports actually give you the ability to satisfy this requirement. There is obviously an existing toolset, the popHealth tool MITRE is working on, that essentially does everything that you see on this little flow diagram up here. And so it might be interesting conversation to have at some point about whether or not if an EHR can adequately produce clinical quality measure reports that discriminate based on LOINC codes – that it's met this requirement. Any thoughts about that?

Clem McDonald – National Library of Medicine

Well, it would be worth doing an inventory of what percentage of codes that are there and whether it would be required. Because that would then impede the utility for presenting data to clinicians in an efficient and fast way, if there is just a very small subset identified to satisfy the requirement. My intuition is that this is not a huge set of lab tests, but I haven't done the count.

Ken Gebhart – National Institute of Standards and Technology

I gather Floyd's not on the call. He would have probably been able to give us some hints. But, you know that's the kind of question that we could explore to kind of refine this idea further. Okay, anything else on slide 7. Alright, let's then move on to slide 8. We're moving now to the second requirement, which is don't break if the EHR receives an unknown LOINC code, or using Doug's friendly amendment, an unknown code period. And here, the requirement that we discussed was that the EHR needs to be capable of displaying the text description of the lab test if the received LOINC code is not recognized by the EHR. And that's the text description as provided in the received message. We are back to the same model we used before. What functionality would you test? Certainly one of them would be receiving and displaying the lab results, including the text of the lab result. That pretty much fits with the CLIA requirements. CLIA doesn't actually, as I recall, doesn't talk about LOINC, it just talks about pieces of data that have to be presented. So that would be a visual inspection. I don't know if you're thinking flowsheet display is part of the requirement if you don't know the LOINC code. I struck through some of the others that were on the earlier list, assuming that they didn't make sense based on this requirement, but I'd be interested in your thoughts of what would you expect an EHR from a functional point-of-view to be able to do if it does not get a LOINC code it recognizes.

Clem McDonald – National Library of Medicine

The bigger question – going back to Doug – if it doesn't get a code it recognizes, and I don't know how it can flowchart if it doesn't recognize it. I mean it has to know it's the same of something else so it has to recognize it in some way. Maybe there is a way.

Betsy Humphreys – National Library of Medicine

This is Betsy. I would have thought that on the assumption that they have to tell you what the lab test is, even it's not standard and there is no LOINC code. Why would we not want them to pass it out on a patient summary?

Clem McDonald – National Library of Medicine

Good question, yeah.

Betsy Humphreys – National Library of Medicine

It would seem to me that if the patient's summary is supposed to include lab test results, then it should go there.

Ken Gebhart – National Institute of Standards and Technology

That's probably a good – that's a good catch.

Clem McDonald – National Library of Medicine

Typically, what systems will have is they will have a report that shows it soon as it comes across. Like, you know like the paper report or something close. And then I don't think we've heard what percentage of EMRs will have flowsheets. Are they required to? Does anyone know? Because then they do have to understand the codes at some level.

Ken Gebhart – National Institute of Standards and Technology

As we mentioned right at the end of the last call, I'm not sure – I don't recall any requirement in meaningful use stage 1 about flowsheet capability being required for meaningful use.

Clem McDonald – National Library of Medicine

Does anyone know if HL-7 requires it in its functional specs? I mean I think the big hospital ones all show them, I just don't know about the office practice ones.

Ken Gebhart – National Institute of Standards and Technology

Right. I would expect to find it in the EHR functional model, if that's what you're asking.

Dan Vreeman – Regenstrief Institute

This is Dan Vreeman. One other thing to consider in this space is maybe the case where what's sent is both the local code and the LOINC code at the same time, but yet the LOINC code is not recognized. Maybe it's a valid code, but it's not in their top 2000. I'd think you'd wanna ensure that the system preserves that alternate code and could use it if it was generating a patient summary. So not that it necessarily recognizes it and can do something smart with it, but at least it shouldn't throw it away.

Ken Gebhart – National Institute of Standards and Technology

Good point, good point. Let's move on the slide 9 just quickly. I think we've touched on most of this, but again how would you test. This could be an inspection test if you're trying to look at what the EHR does with the lab message in the EHR. If we take the path of producing patient summary records, as was mentioned a minute ago, then there could be some automation testing to back this up. Then let's move onto slide 10 for just a second. Trying to get to sort of the big list of questions here. So if you're trying – the third requirement was to transmit LOINC codes and the appropriate messages and documents. We won't spend a lot of time on this. There is some of this that is already being done in stage 2 with conformance testing tools. And pretty much, I think the process for validating those is pretty straightforward. The one nuance that would be different than what's been done in stage 1 would be the test message generator that has the ability to send the EHR the information that it's supposed to be evaluating. As you might imagine, we did get some significant pushback from organizations that were asked to input laboratory result data into their EHRs in order to generate output saying, "We do that through an interface from a lab system. Why won't you allow that." We did leave that possibility open in the test procedures for meaningful use stage 1; however, there was plenty of confusion about whether that was an option or not. So we would just make it the default method for this part that feeds the EHR the input lab messages, and then it could generate the outputs.

Clem McDonald – National Library of Medicine

I don't follow. You're saying that the lab would generate the test messages?

Ken Gebhart – National Institute of Standards and Technology

The test message generator would generate the messages.

Clem McDonald – National Library of Medicine

I didn't understand what the pushback was.

Ken Gebhart – National Institute of Standards and Technology

Oh, the pushback was in the test procedures for stage 1. Because we didn't have a test message generator available, people were upset that the test procedures required that they enter lab test results into the EHR. And we tried to provide the option for individuals and organizations to key the data in or absorb it through an interface. But that point about absorbing it through an inbound interface got lost in many people's minds, and so they were objecting to the notion that they'd have to key in laboratory test results when of course they get them in a message.

Clem McDonald – National Library of Medicine

Thank you. I get it.

Ken Gebhart – National Institute of Standards and Technology

Yeah, so it's lessons learned from how do you write these as well as what are people expecting. They're expecting a real-world workflow is what it amounts to. Okay, so I'm gonna move on to slide 11 then. I think we've touched on some of these things, but as we start to think about what does this mean for other meaningful use 2 vocabularies, here's a starter list of questions we could start to kick around. First one is, do we have the same R1, R2, R3 for all vocabularies, or are there some differences in EHR capabilities that need to be considered between and among vocabularies. Obviously, this tends to be somewhat whether you're thinking of just testing that the vocabulary is present or whether you're thinking about it in the context of what functionality the EHR is doing with it. Anybody want to share any thoughts about this?

Clem McDonald – National Library of Medicine

Well, I think on the first batch you'd like to be able to take whatever that subset that's defined or some -- maybe a smaller -- well some subset and see if you can send it to the thing. It stars it, and it can put it back out in another message. And the summary report might be the best cause you could use it for all of them.

Ken Gebhart – National Institute of Standards and Technology

Yep.

Clem McDonald – National Library of Medicine

And I think you know, at least in HL-7, there is always a text-plus-code. I don't know -- so that it's an expected thing.

Ken Gebhart – National Institute of Standards and Technology

Yep. Why don't we talk about the text for just a second. There is -- this is one of the things I've been scratching my head about, because we all know that virtually all the standards that have specified so far have some place to put a text description. It would be good to do the due-diligence about whether the implementation guides tell us what is supposed to be populated in that text field. I bring this up mostly because if you think about LOINC, there could be a presumption that's it the LOINC short name or LOINC long name or the local text description of the sending system. And we want to be clear about what is expected. If you get back in your future conversations about the medication allergies letter that you were discussing in the last call, there is a similar point about text; but it has a subtlety to it that the letter is, as I understand from George Robinson -- talking to him directly, the letter is suggesting that the text description that the clinician actually saw on the screen when they selected the allergy is what should be conveyed in the message. So, you know there is variations like that that we'll have to dig through and get to the bottom of as part of the requirements.

Clem McDonald – National Library of Medicine

Can I direct some points? I don't know what might be constrained by decisions in other domains, but there are two fields in most cases -- I mean two triplets, a local and a standard; and one might argue presenting both.

Ken Gebhart – National Institute of Standards and Technology

Right. And the big thing, from a conformance testing point-of-view is having a clear conformance statement that says what is expected. So, alright. A couple of other things which I know you've grappled with early in the Vocabulary Task Force conversations. I don't personally recall the conclusions; but how will official versions of each vocabulary be posted, provided, distributed, available for testing is sort of my

question. And also, how will versioning differences between these versions of vocabularies be handled during testing?

Betsy Humphreys – National Library of Medicine

This is Betsy Humphreys. I would think that in the near term – Doug can say what he thinks, but I would assume that NLM could be the source of the official versions if that would be helpful. We'll all have to decide which version it is you're supposed to be using. Right? At any given time.

Ken Gebhart – National Institute of Standards and Technology

Right. I do remember some of those early conversations in this workgroup, and I didn't know if they reached closure or it sounds like maybe there's still a little bit of closure needed.

Betsy Humphreys – National Library of Medicine

I wouldn't doubt that.

Ken Gebhart – National Institute of Standards and Technology

But to get to testing, we have to sort of resolve this to get down to the point of what versions are we testing, where do they come from, and if there are – you know the classic problem of release a new version and how long do people have to get up to speed; and all that kind of stuff needs to be thought through.

Lynn Gilbertson

And Ken, this is Lynn Gilbertson. You might have extended situations with the versioning. Like if you think about the exchanges of national drug codes and the fact that the compendia are holding that information inside their products and the versioning is not really exchanged as part of the message. So you may have – you may not have that ability to really know which is the official version unless you dig into where you got the source from. And if it's one off, it may not be even germane.

Ken Gebhart – National Institute of Standards and Technology

I agree. I think there is a – there are people who have kicked this around in a variety of settings that I've listened to; and all I'm trying to suggest here is in order to get to a testing strategy, these are things we need to nail down.

Lynn Gilbertson

Right. Or at least the floor perhaps or something.

Ken Gebhart – National Institute of Standards and Technology

Yep. Again, it's a bigger challenge when you start extending this thought to multiple vocabularies not just LOINC and – I'm not sure how big this list is of how vocabularies are gonna be specified for meaningful use stage 2. So the more there are, the more we have to get sorted out. One last comment here that is on this list is in order to do the automated conformance testing, it would be important to understand what data transport standards are gonna be specified for meaningful use. If you'll recall in meaningful use stage 1, there were early conversations about data transports standards; and then those were removed from the final rule. If there are transport standards identified for stage 2, then we'd need to apply them for this automated testing. If transport standards are not identified, then we need to, in terms of developing a test method, provide some alternatives in the test suite for EHR vendors to send data or receive data using various transport standards. So that's another one of those perquisites that you'd like to get settled early; not late in the game. Are there other sorts of questions that this has prompted people to think we should add to this list to try to get sorted out?

Unknown Speaker

Did you have any – I think we ran into that with the e-prescribing where the test scenario -- there was some conundrum with whether or not the EHR system actually supported, like for example the type of drug. Is there any of that concern with LOINC or SNOMED, where it's a test scenario that is not a part of that business and so is difficult for the implementer to try to test?

Ken Gebhart – National Institute of Standards and Technology

So probably an analogy here would be the ophthalmology specialty systems that don't need all these codes. Wouldn't use them.

Unknown Speaker

And wouldn't know how to build a test – you know do they have to have their own kind of test scenarios or what?

Clem McDonald – National Library of Medicine

Well, if you dealt with just it can pass through, understand, and ship them out – I get it, they shouldn't have to, but it shouldn't be – I mean it makes it real easy to test and it shouldn't be that hard to implement. But I don't know. No, it's not saying you have to put them on your menus for the users to use. It's not saying you have to show them in your own local flowsheet. It's just saying you should be able to process them.

Unknown Speaker

Right. And we just ran into some questions with if this isn't part of my business, all I'm doing is monkeying up some data and sending it. Is that enough of a test? So that was just one of the things I was thinking of as we went through phase 1.

Doug Fridsma – Office of the National Coordinator

I think one of the things that we may want to consider is we have a workgroup that is specifically charged with taking a look at both the implementation of the EHRs with regard to meaningful use stage 1 and some of the certification criteria. So that particular workgroup – I think conversation that we've had with this group should feed into that workgroup and provide them some additional feedback and information about things that have worked and that haven't worked and how the vocabulary testing strategy might fit into the feedback they're gonna get from those folks that have already gone through the certification process. And so, Judy Murphy is one of the co-chairs there, and Liz. And I think those two probably would do well if we could summarize some of this conversation and perhaps give them that information, because they'll be able to fold that into some of the other feedback they've got about some of the testing strategies and people dummifying up data and things like that. And they'll be able to devote some time, not only around the vocabulary issues but around the broader business processes and things like that.

Ken Gebhart – National Institute of Standards and Technology

Okay. This is the conclusion of everything I was thinking to share with you today. It's sort of initial reaction to the requirements and thoughts about the testability of these requirements based on what we know right now. The underlying message I believe from this would be there is more specificity needed in the requirements in order to actually design the test. The extent to which test tooling is helpful kind of depends on how we shake that out. It's been great to hear your comments today; and I'll be glad to update this deck to sort of reflect those comments or at least summarize them so we have them in print, if you will. I don't know whether the workgroup would like NIST to do anything more at this point with this, or [interrupted] ...

Jamie Ferguson – Kaiser Permanente

This is Jamie. Do we still have Doug on the call?

Doug Fridsma – Office of the National Coordinator

I'm still here.

Jamie Ferguson – Kaiser Permanente

Alright. So I have a question for you that relates to the maintenance of the subsets. And so if the adopted standard in a rule is the entire vocabulary, but we're using subsets both for convenience purposes and for certification testing, is it possible that those subsets that are used in certification can be updated on a maintenance schedule unrelated to the publication of the rule? So basically, the adopted standard is the whole thing; but for certification we can maintain a subset differently.

Doug Fridsma – Office of the National Coordinator

I think that is always an issue when we start talking about what we can do in regulation and what we can do in subregulatory mechanisms. I don't know if I have a specific answer for you. I think we're still sort of exploring what's the right thing to do. But you know, if there is something that is published as part of the regulatory process in the rule, then updates to that will likely need to occur through rule-making.

Clem McDonald – National Library of Medicine

I think the question – can you reference a source in a standard. I thought I'd seen that; you know where you say it will be whatever as [interrupted]

Jamie Ferguson – Kaiser Permanente

Yeah. My question, I think Clem, is slightly different. So I'm proposing that the adopted standard in the rule would be a version of the entire vocabulary, but that for certification there might be a subset that would be managed on a different maintenance schedule.

Clem McDonald – National Library of Medicine

Oh, okay.

Doug Fridsma – Office of the National Coordinator

I think one of the issues is that suppose you adopted the entire set of codes represented in LOINC as your vocabulary. It would impractical, except if we did white-box testing, to confirm that a particular EHR has every single one of those codes represented. And I think the kinds of things that NIST is doing in this discussion to try to get that functional conversation that we send you a LOINC test, you're able to repackage that and send that as a clinical care summary, putting together scenarios that stream together functionality and achieve kind of the policy objectives about information. Those sorts of things aren't gonna show up in the rule, but those will show up in the way in which the testing criteria occur. So, we know that even for the testing, even if we didn't put any identified subsets, we would probably have to figure out some way of demonstrating functionality that meets our policy objectives that uses the standards but doesn't necessarily require you to show you all 500,000 SNOMED codes or all of the thousands of LOINC codes that might be out there as well. So I think there is a way that we can help provide – make it easier for the vendors to know where it is that we're going, what are the things that they're going to be held accountable for. And I think certainly the strategy about saying, listen you have to be able to receive a LOINC code that you know about or that we've identified but you also can't break if there is either a new LOINC code or something that you don't recognize. In some sense, you're doing through certification and testing future-proofing the functionality of the EHR, if you will, because if there is

an update to the LOINC codes, then those systems should continue to operate even if they aren't able to do decision support or other things on it.

Jamie Ferguson – Kaiser Permanente

Okay. Is there anything else for this call then?

Clem McDonald – National Library of Medicine

I'd just like if Ken could give us a sense of what he heard from us.

Betsy Humphreys – National Library of Medicine

Well, I think he was gonna update the deck and send it around. Right?

Ken Gebhart – National Institute of Standards and Technology

That's what I'd like to do if you don't mind. That I would – it will help me sort of gather my thoughts and maybe listen to the conversation one time and play it back to you to see if it still makes sense.

Betsy Humphreys – National Library of Medicine

This is Betsy Humphreys. I think that would be very helpful, and Ken if you are able to maybe add something at the end that lists the things that you still don't know that you would need to know in order to do this right.

Doug Fridsma – Office of the National Coordinator

And let me just also echo my gratitude to NIST and to Ken for the work that they've done with regard to this. One of our objectives a year ago, as we were going through the meaningful use stage 1 and as we were sort of putting together the operations around the standards and interoperability framework, was to really engage NIST in the testing strategies early in the process, before we sort of had identified everything and then sort of handed it over to NIST. So I'm just delighted to have NIST involved early in this process, cause I think at the end of the day we will have a much better and much more coherent testing strategy that will, I hope, achieve our policy objectives that we've got. So, I'm just really grateful for NIST's participation early in this process.

Ken Gebhart – National Institute of Standards and Technology

Glad to do it.

Jamie Ferguson – Kaiser Permanente

Yes, thank you Ken.

Judy Sparrow – Office of the National Coordinator

Public comments.

Jamie Ferguson – Kaiser Permanente

Yeah.

Judy Sparrow – Office of the National Coordinator

Okay, operator can you check and see if anybody wishes to make comment from the public please?

Operator

Yes. If you're on the phone and would like to make a public comment, please press *1 at this time. If you are listening via your computer, you may dial 1-877-705-2976 and press *1 to be placed in the comment queue. [Pause] We have a public comment.

Judy Sparrow – Office of the National Coordinator

Great. Can you please identify yourself?

Carol Bickford – American Nurses Association

This is Carol Bickford at the American Nurses Association. I appreciate the thinking that's going forward, particularly in light of establishing a methodology that will work for the LOINC testing but also has applicability to later terminology work. It is great to think that the workgroup is addressing forward-thinking action.

Judy Sparrow – Office of the National Coordinator

Thank you, Carol. Any other comment?

Operator

We have no more comments at this time.

Judy Sparrow – Office of the National Coordinator

Okay, thank you. Back to Jamie.

Jamie Ferguson – Kaiser Permanente

Okay, thank you everybody for a great discussion today. Actually, I think that's it for this call then.

Judy Sparrow – Office of the National Coordinator

Yes, thank you.

Unknown Speaker

Thank you very much.

Judy Sparrow – Office of the National Coordinator

Bye-bye.