0:0:0.0 --> 0:0:29.650  
Andrew Williams  
As with some of the other NLP focused studies, one of the main goals will be to demonstrate the value of using NLP to get additional information and to do that in an efficient way by building on an existing study and the outcomes of interest are going to be how much additional information do you get? How does that additional information allow you to do a better job of defining populations and exposures and outcomes that are used?

0:0:30.130 --> 0:0:38.80  
Andrew Williams  
In an analysis and then when you've done that, how does it affect the results of the original study that you're?

0:0:39.300 --> 0:0:45.520  
Andrew Williams  
Building off of and in this case, we're also gonna be taking advantage of the excellent work that's already been done.

0:0:46.260 --> 0:0:54.740  
Andrew Williams  
I they found and others VNA and others on the social determinants of health ontology that will allow us to.

0:0:55.910 --> 0:1:3.120  
Andrew Williams  
Map that data in a well formed ontology and to stratify in some.

0:1:3.840 --> 0:1:8.300  
Andrew Williams  
Uh. Convincing way that that is taking advantage of that.

0:1:10.200 --> 0:1:19.10  
Andrew Williams  
That better framework than people typically have, so even if you can get that data from NLP, there's the added problem of how do you represent it and use it in the analysis. And so the.

0:1:19.750 --> 0:1:22.170  
Andrew Williams  
Uh, I think the main goals are.

0:1:23.260 --> 0:1:23.940  
Andrew Williams  
To repeat.

0:1:24.930 --> 0:1:30.360  
Andrew Williams  
Show the additional information you can get show the impact of that information on superior.

0:1:31.180 --> 0:1:33.590  
Andrew Williams  
Definitions for populations, exposures and outcomes.

0:1:34.580 --> 0:1:42.660  
Andrew Williams  
Rerun the study and show how they are different in some dimension when you stratify on social terms of health dimension.

0:1:43.560 --> 0:1:48.830  
Andrew Williams  
How diabetes outcomes for a particular drug differ by.

0:1:49.980 --> 0:1:53.560  
Andrew Williams  
Some some measure, perhaps economic measure or other.

0:1:54.830 --> 0:1:56.290  
Andrew Williams  
Dimension of social terms of health.

0:1:57.120 --> 0:2:4.140  
Andrew Williams  
Umm. And in the process of doing all that, demonstrate the value of the social determinants of health ontology for.

0:2:5.110 --> 0:2:9.680  
Andrew Williams  
The the purposes of of representing that extracted data, that's.

0:2:10.440 --> 0:2:12.260  
Andrew Williams  
It's going to be in that.

0:2:13.330 --> 0:2:35.710  
Andrew Williams  
Overall description that people will react to, I don't think we've 100% nailed complete, you know, universal acceptance that that's the idea. But people, there's been mostly relying on head nodding and people saying, yeah, that sounds good. Let's do do that. But I think when we've got a document to respond to, we can people can.

0:2:36.600 --> 0:2:36.950  
Andrew Williams  
You know.

0:2:37.620 --> 0:2:44.100  
Andrew Williams  
Weigh in and say that this is missing something key or I don't have it yet. So does that make sense? You one does that.

0:2:47.460 --> 0:2:47.690  
Andrew Williams  
OK.

0:2:45.570 --> 0:2:50.710  
Lu, Yuan  
Yeah, that makes sense. Can you hear me OK? Uh, is it better now or is still unclear?

0:2:49.410 --> 0:2:54.980  
Andrew Williams  
It's still very it's. Yeah, your mic is not working well. Something about it is not happening.

0:2:54.710 --> 0:2:55.60  
Lu, Yuan  
Save.

0:2:58.670 --> 0:3:0.110  
Lu, Yuan  
How about now? Is it better?

0:3:0.750 --> 0:3:6.80  
Andrew Williams  
World's better I feel I feel lighter and and happier. Just hearing you speak.

0:3:5.180 --> 0:3:32.570  
Lu, Yuan  
OK. That's great. Sorry for that. I think that that's yeah, definitely makes sense. It seems we are going to do 2 fold, right? First demonstrate the feasibility of extracting the accurately extracting social, determine how information using NLP and then to show how does that information will affect the result. So in the discussion with Mark, I feel that like when selecting the exposure and outcome.

0:3:34.180 --> 0:3:36.840  
Lu, Yuan  
We may need to have discussion.

0:3:38.60 --> 0:4:5.690  
Lu, Yuan  
Related to our study goals, say if we want to demonstrate the value of and now P extracting SDO H then Mark was asking oh, maybe we should select the outcome. You could either select the outcome that is clinically most significant for diabetes or you can select some of the outcome that are poorly capture in the structure diagnosis code such as vomiting.

0:4:6.10 --> 0:4:35.690  
Lu, Yuan  
You know, like some of these bleedings some of outcome that are not, you know accurately captured by the diagnosis code that maybe there is higher potential for the NLP to be success to be successful. So I think in terms of the selection of exposure outcome, maybe it require additional discussion with clinician and with the teams as to like what what is our actual goal is and that may inform the selection of that.

0:4:37.750 --> 0:4:42.460  
Andrew Williams  
Right. And he and I think we all agreed, was a good idea to.

0:4:43.360 --> 0:4:44.670  
Andrew Williams  
We're going to invite.

0:4:45.290 --> 0:4:48.950  
Andrew Williams  
Uh, Rohan and Jody Ann.

0:4:50.610 --> 0:4:51.40  
Andrew Williams  
Uh.

0:4:50.510 --> 0:5:3.390  
Lu, Yuan  
I know we'll have you. Uh, yeah, we work in the same office. Like he also clinician at Yale in the Center for Outcome Research and Evaluation. I can. I can invite him to join. I don't know, Judy. So.

0:5:4.30 --> 0:5:5.10  
Lu, Yuan  
If anyone know?

0:5:4.590 --> 0:5:29.640  
Andrew Williams  
Jody Anne. Yeah, I can invite Jody. And she said, uh Colombia. She's kind of overall doing a lot to help organize things through the the office there at Columbia and helping to lead the SOS Challenge, etcetera. Does all kinds of great stuff. So and she's very interested in this work. And I will make sure she gets invited or Marty and I will both make sure that's happening.

0:5:31.170 --> 0:5:45.800  
Andrew Williams  
For future calls, but maybe there's a good point to discuss. I think we we'll wanna be deferring to clinicians and people have good background in the in the content area like like Rohan. But it's there's also just a strategic.

0:5:47.570 --> 0:5:54.420  
Andrew Williams  
Question about what to emphasize. I can feedback part of what you were saying in. In this way we could emphasize.

0:5:55.540 --> 0:6:3.150  
Andrew Williams  
The additional data that gets captured from NLP or emphasize the impact of that data on a clinical.

0:6:3.860 --> 0:6:7.130  
Andrew Williams  
Piece of evidence, a piece of clinical evidence. My.

0:6:8.810 --> 0:6:10.390  
Andrew Williams  
Preference I guess would be the latter.

0:6:11.140 --> 0:6:29.850  
Andrew Williams  
I would say even if you're not getting quite as much additional data, if getting it shows something important about groups in terms of the real question, that's more compelling. I think in general than just saying we got a lot more data because if that.

0:6:30.910 --> 0:6:34.780  
Andrew Williams  
Extra data you got doesn't end up mattering so much.

0:6:35.520 --> 0:6:35.930  
Andrew Williams  
It's.

0:6:36.710 --> 0:6:38.380  
Andrew Williams  
Of more academic interest than.

0:6:40.180 --> 0:6:47.570  
Andrew Williams  
Overriding importance. That's my two cents on on that. But what other people's thoughts on that question, the trade off, assuming there is one?

0:6:52.590 --> 0:6:56.780  
Dima Lituiev  
That makes sense to me. I do appreciate that. Maybe showing what?

0:6:58.250 --> 0:7:16.910  
Dima Lituiev  
How would we get impacts? The outcome is important. I wasn't clear why. Like we, I think you even brought up some other feels like vomiting legier which can be extracted from LP, but they are outside of the SDH scope and it wasn't quite clear whether that was the suggestion.

0:7:17.450 --> 0:7:24.420  
Andrew Williams  
I think it was just an example of something that's hard to get rather than something within that SDH domain. But Georgina, you're.

0:7:25.770 --> 0:7:28.380  
Andrew Williams  
Joining us in the wee hours of the what time is it there?

0:7:30.480 --> 0:7:31.870  
Georgina Kennedy  
6:15 good morning.

0:7:34.440 --> 0:7:42.290  
Andrew Williams  
Good morning. Thank you so much for joining us. Were you on one of these previous calls? So I know we had met in between calls to talk about this, but.

0:7:42.950 --> 0:7:45.670  
Georgina Kennedy  
No, this is the first one for this study that I have joined.

0:7:46.470 --> 0:8:9.360  
Andrew Williams  
Yeah, thought so. And you all are doing a lot of interesting work. And in in in another of these NLP calls, you brought up a topic I was going to raise, but you should be the one to kind of raise it and frame it here. So we talked previously maybe after you introduce yourself, I'll just ask you to talk about benchmarking. We talked briefly in this call as in some others about whether or not to have a comment.

0:8:10.0 --> 0:8:17.690  
Andrew Williams  
And LP acquisition strategy, common pipeline is used or let every site use it differently, whatever one they have.

0:8:19.60 --> 0:8:34.910  
Andrew Williams  
And in that context, I think on another what maybe the psychiatry and LP study we're doing or uncollege, maybe it is the oncology one you brought up the potential value of having a benchmark run at each site. So anyhow, maybe you just give a little introduction to yourself and then talk about benchmark.

0:8:35.490 --> 0:8:36.70  
Georgina Kennedy  
Yeah, sure.

0:8:49.100 --> 0:8:49.350  
Andrew Williams  
Yeah.

0:8:36.160 --> 0:8:52.980  
Georgina Kennedy  
And so I'm calling in from Sydney, Australia and I am a research fellow at the Ingham Institute and the University of NSW. It was definitely the oncology call because my research group is heavily focused.

0:8:54.10 --> 0:9:23.150  
Georgina Kennedy  
Exclusively focused in the oncology domain, so I'm not sure whether we will be able to fully participate in this city, except believe that the replication that you guys are doing is a diabetes study. So it's possible that we won't have the source data to work with, but we are very interested in social determinants of health space. We we work in the public health system here and we are focused on an area of Sydney that is very diverse.

0:9:24.40 --> 0:9:34.20  
Georgina Kennedy  
And we have a number of oncologists looking at populations of, you know, cultural and linguistically diverse populations, refugee populations.

0:9:35.220 --> 0:9:40.310  
Georgina Kennedy  
So a number of highly relevant domains and and one of the areas that we've been.

0:9:41.60 --> 0:10:11.60  
Georgina Kennedy  
Piloting is looking at interpreter use, but not only you know the the more formal book to interpreters, but, but I'm trying to figure out if it's feasible to get out of the natural language processing the use of its informal interpreters, so bringing along family members for those who who, who don't speak English, and and trying to figure out if it's feasible to have a baseline of of of that.

0:10:12.780 --> 0:10:18.890  
Georgina Kennedy  
But yeah, So what? What Andrew said that I guess was I was a bit concerned in the oncology group that we were looking for.

0:10:20.30 --> 0:10:25.580  
Georgina Kennedy  
Histopathology and endpoints and just to make sure that.

0:10:27.0 --> 0:10:28.710  
Georgina Kennedy  
Yeah, we don't have.

0:10:28.900 --> 0:10:43.930  
Georgina Kennedy  
And pathology reports and our source data where some other groups do. And the idea of having clinical notes at as the the data source compared to his histopathology reports.

0:10:45.550 --> 0:11:14.660  
Georgina Kennedy  
It seemed really important, I guess, to have some baseline, even if it's a very simple, naive kind of tool, that there are, you know, existing NLP pipelines or transform models for pathology reports that we could pull down and apply across the board and have a baseline of accuracy. But it's not necessarily also a baseline of accuracy, just of of coverage because you only know what?

0:11:15.360 --> 0:11:43.660  
Georgina Kennedy  
You can get out of the text, you know, even if you're 90% accurate. If it's very poorly documented, you are going to be comparing apples and oranges from the source system. So yeah, I was just proposing looking for something that was publicly available and easy to deploy so that everybody could have a baseline and then use your own system that's been heavily tuned and localised and made much more relevant to your source data.

0:11:44.860 --> 0:11:51.110  
Georgina Kennedy  
For the main comparison but but to have to have this kind of baseline touch point to be able to know.

0:11:52.710 --> 0:11:55.990  
Georgina Kennedy  
The different data sources and kind of figure out what you're actually looking at.

0:11:57.430 --> 0:12:0.220  
Georgina Kennedy  
I'm not sure if that was sort of what you want me to describe.

0:12:2.30 --> 0:12:10.960  
Andrew Williams  
Yeah. Describe. And then, you know, have the group react. So I think it's an important question. And when you raised and I'm not sure we were able to.

0:12:12.840 --> 0:12:21.350  
Andrew Williams  
Come to a conclusion about it in the oncology group. But I think it's a, it's an excellent thing for us to discuss and start to put our.

0:12:23.300 --> 0:12:28.340  
Andrew Williams  
Heads together on so. I'll guess. What are other people's thoughts about the use of that kind of benchmarking strategy?

0:12:40.340 --> 0:12:51.190  
Georgina Kennedy  
I think it does. It does kind of depend what you looking for and I've come in and and I I don't really know your endpoints and and in the histopathology there do exist publicly available.

0:12:52.210 --> 0:12:58.980  
Georgina Kennedy  
Models and so it could be very, very different in this domain for the social determinants of health, because the you know.

0:12:59.700 --> 0:13:3.250  
Georgina Kennedy  
The applicability of additional public models.

0:13:3.350 --> 0:13:7.790  
Georgina Kennedy  
Yeah, could be infeasible and by comparison.

0:13:9.230 --> 0:13:18.40  
Andrew Williams  
I guess that number of things that we're going to be using in LP to get information on as part of the thing, the one of the values.

0:13:19.80 --> 0:13:24.940  
Andrew Williams  
One of the benefits of using the Legend Hypertension study as a way to do this is that.

0:13:25.820 --> 0:13:27.200  
Andrew Williams  
There's so many different.

0:13:30.820 --> 0:13:31.790  
Andrew Williams  
Things there.

0:13:32.610 --> 0:13:33.200  
Andrew Williams  
Uh.

0:13:51.320 --> 0:13:53.340  
Georgina Kennedy  
Yeah, you might not need 100% coverage.

0:13:54.120 --> 0:13:54.480  
Georgina Kennedy  
Yeah.

0:13:34.70 --> 0:13:54.520  
Andrew Williams  
That would be surprising to me if we couldn't find something where there's been prior work that could be used as a benchmark, and I don't know that that's true for sure, but there are just lots and lots of different exposures and lots of different endpoints involved in that study. And so it would at least open up the possibility that, sorry.

0:13:55.700 --> 0:14:3.50  
Georgina Kennedy  
I was just saying, yeah, you would. You wouldn't necessarily need 100% coverage if you had something representative and just picked a couple of important points.

0:14:4.260 --> 0:14:4.520  
Andrew Williams  
Well.

0:14:14.940 --> 0:14:15.290  
Georgina Kennedy  
Answer.

0:14:4.620 --> 0:14:28.370  
Andrew Williams  
Yeah. We're only gonna pick things. But, I mean, we could. We could have the availability of something that serves as a benchmark, be part of the heuristic that we use in picking something. If we decided that having a benchmark was a high priority, I guess that was the main thing I was trying to get to. But yeah, we're definitely not gonna try and cover all of the things that are in the legend, hypertension. We keep it simple as possible. And but we could.

0:14:30.340 --> 0:14:30.770  
Andrew Williams  
You know.

0:14:31.690 --> 0:14:33.970  
Andrew Williams  
Use the availability of eight benchmark.

0:14:34.650 --> 0:14:37.60  
Andrew Williams  
As one of the criteria for selecting something.

0:14:43.620 --> 0:15:5.100  
Lu, Yuan  
Yeah, I like the idea of of studying simple and straightforward at the beginning. So if we agree that pick a clinical important endpoint to start with, I would suggest using maze which is the primary outcome in legend diabetes. So there is either 3 point maze or four point maze.

0:15:6.310 --> 0:15:36.70  
Lu, Yuan  
Maybe using the three-point maze, because in the four point Mace the sudden cardiac death has lots of misclassification there. So maybe just focus on the three-point base as the primary outcome and the one that I think clinically will have the most significant impact on and for the exposure. I don't know. There's so many drug exposure out there and if we are thinking about stratifying the analysis by social determine how we may want to select.

0:15:36.680 --> 0:15:59.260  
Lu, Yuan  
Drug exposure that has a lot of drug users, like maybe some of the newer draft as their still taking time to increase their uptake in clinical practice. The actual number of users may not be bigger enough if we want to stratify analysis by many of the subgroups. So just something to consider in my.

0:16:0.490 --> 0:16:5.360  
Lu, Yuan  
Do we have enough sample size if we further stratify by different social determinant health support?

0:16:6.270 --> 0:16:6.670  
Andrew Williams  
Umm.

0:16:8.90 --> 0:16:9.980  
Andrew Williams  
I agree completely. Yeah, something that's.

0:16:12.370 --> 0:16:23.930  
Andrew Williams  
Yeah, that's that makes good sense. We don't wanna pick something where it's so small, probably for a number of reasons. But yeah, to to be able to stratify, you need a sizable subpopulation for each.

0:16:25.370 --> 0:16:25.860  
Andrew Williams  
Umm.

0:16:26.860 --> 0:16:31.610  
Andrew Williams  
Georgina is I really appreciate you making the time and sorry this is scheduled it it's a tough hour for you.

0:16:32.290 --> 0:16:36.730  
Andrew Williams  
The as you said, your group's ability to participate.

0:16:38.490 --> 0:16:41.580  
Andrew Williams  
Will depend on suitability for the population.

0:16:44.420 --> 0:16:44.750  
Georgina Kennedy  
Yep.

0:16:42.440 --> 0:16:47.930  
Andrew Williams  
You have, I'm not sure when you and I spoke earlier, if we.

0:16:50.410 --> 0:16:59.500  
Andrew Williams  
If you came to a conclusion, I think I mentioned it's possible to think about folks who are have comorbid cancer.

0:17:0.970 --> 0:17:6.670  
Andrew Williams  
And diabetes. But that's probably not optimal. It would have to sort of think about them as as a separate group.

0:17:7.330 --> 0:17:14.320  
Andrew Williams  
Uh, so we we talked about the possibility of having a maybe a parallel study. I I I think there's a fair amount of overlap.

0:17:15.420 --> 0:17:19.710  
Andrew Williams  
In whose participating in the oncology study and this one?

0:17:21.580 --> 0:17:24.650  
Andrew Williams  
So it's possible that we just think.

0:17:25.860 --> 0:17:37.70  
Andrew Williams  
More broadly about the two studies as having a different kind of relationship to each other, or if I think the strategy that we're moving towards there in the oncology studies to take.

0:17:38.450 --> 0:17:41.600  
Andrew Williams  
The cancer treatment pathways study that.

0:17:42.920 --> 0:17:45.790  
Andrew Williams  
Karthik has run and use that as the template.

0:17:46.990 --> 0:17:50.60  
Andrew Williams  
And do social determinants of health there as well?

0:17:51.50 --> 0:17:51.690  
Andrew Williams  
As in the.

0:17:53.30 --> 0:17:54.730  
Andrew Williams  
Type 2 diabetes study.

0:17:55.680 --> 0:17:56.410  
Andrew Williams  
Or.

0:17:59.110 --> 0:18:4.780  
Andrew Williams  
I guess potentially not use the Type 2 diabetes study and use the cancer treatment pathway study instead. I'm.

0:18:5.740 --> 0:18:15.630  
Andrew Williams  
Pretty interested in doing diabetes, but I guess we could also if especially if there's a large overlap and the number of sites participating.

0:18:16.450 --> 0:18:17.320  
Andrew Williams  
You know, run.

0:18:18.120 --> 0:18:19.570  
Andrew Williams  
On that same set of patients.

0:18:20.320 --> 0:18:20.780  
Andrew Williams  
Umm.

0:18:21.920 --> 0:18:29.660  
Andrew Williams  
The cancer treatment pathways study and stratified by social term of health and that group as well. But I don't know what are you.

0:18:30.540 --> 0:18:31.730  
Andrew Williams  
How do you see that?

0:18:33.950 --> 0:18:34.450  
Georgina Kennedy  
I.

0:18:33.800 --> 0:18:44.680  
Andrew Williams  
I guess can you illuminate you know how how much of an obstacle it would be to participate if it there's an A cancer outcome and what some of the strategies might be for doing something efficiently that does have a cancer outcome?

0:18:50.600 --> 0:18:50.810  
Andrew Williams  
Yeah.

0:18:45.550 --> 0:19:15.860  
Georgina Kennedy  
I I don't think I'd address suitable if if we're looking purely at the diabetes study, I'm keen to sort of come along for the ride regardless. I think because we are moving into the social determinants of health in an OP space already, and so we're going to have all these related considerations and so I'm happy to sort of just sticky beacon, but I'm an observer for in this particular study. And then if we do get down to the point in the in the oncology study where.

0:19:15.960 --> 0:19:25.690  
Georgina Kennedy  
There is sort of an applicable endpoint that we can draw sort of the lines together that might give us a comparison across the two studies, happy to kind of effect that in the other.

0:19:26.190 --> 0:19:47.400  
Georgina Kennedy  
And study as well and just kind of take a wait and see approach. I don't think it's wasted time for me to like just if you guys are happy for me to sort of participate passively in this one until such time as we can identify those synergies and pull something across if there's capacity for it.

0:19:52.750 --> 0:20:1.40  
Andrew Williams  
Yeah, a little bit more background to the cancer study we're thinking about. We haven't really confirmed this next meeting in oncology version of this would probably confirm it, but.

0:20:3.610 --> 0:20:9.940  
Andrew Williams  
In the United States, there is a division of the National Cancer Institute that's about population level.

0:20:10.690 --> 0:20:12.680  
Andrew Williams  
Strategies and effects and.

0:20:13.460 --> 0:20:20.80  
Andrew Williams  
The person there at NCI who sponsored the work that Kartik did.

0:20:20.910 --> 0:20:27.860  
Andrew Williams  
Is very likely to be interested in whether we could understand how these pathways differ by.

0:20:28.910 --> 0:20:41.290  
Andrew Williams  
Populations to find not only in clinical characteristics, but also on social determinants. And so it seems like a good fit, both from the sponsors point of view on our ability to show something and show the value of doing this.

0:20:42.440 --> 0:20:55.700  
Andrew Williams  
If there's a large enough sample of the sites involved in both, we could say you're already getting the social terms of data for the diabetes study. Apply it to this other study. So I don't know VPN.

0:20:57.50 --> 0:20:59.0  
Andrew Williams  
If you have a sense of how many of the sites.

0:20:59.840 --> 0:21:22.190  
Andrew Williams  
Involved, amartya. It's a pretty sizable overlap, right? When we were going through and saying everybody, raise your hand if you're interested in doing oncology and now social terms of health. And now psychiatry, right. It was kind of a similar group raising their hand over and over again. But I don't know exactly what that extent is, but do you have a sense or, Marty?

0:21:23.10 --> 0:21:30.720  
Vipina Keloth  
Yeah, I think for social determinants of health we have the maximum. We have around 5-6 institutes I guess.

0:21:31.860 --> 0:21:35.650  
Vipina Keloth  
But for the other, like many people who join.

0:21:36.500 --> 0:21:39.630  
Vipina Keloth  
Who are joining this meeting? We don't have them in the oncology and.

0:21:40.520 --> 0:21:41.490  
Vipina Keloth  
As I get everyone.

0:21:43.10 --> 0:21:43.340  
Andrew Williams  
OK.

0:21:48.510 --> 0:21:53.750  
Andrew Williams  
Anyhow, that still seems like a viable thing to explore, whether a subset of sites that are involved in both could run.

0:21:54.550 --> 0:21:55.10  
Andrew Williams  
Those.

0:21:57.330 --> 0:21:58.120  
Andrew Williams  
And.

0:22:0.740 --> 0:22:13.310  
Andrew Williams  
As opposed the protocol for both those studies gets firmed up and the number of sites that are investigating the support in their own data starts to get clearer, we'll have a more concrete answer to that. But I guess it seems like a.

0:22:14.360 --> 0:22:17.30  
Andrew Williams  
Something to keep in mind, Georgina. We might.

0:22:17.810 --> 0:22:18.980  
Andrew Williams  
Be able to run a version.

0:22:19.690 --> 0:22:20.880  
Georgina Kennedy  
Yeah, that sounds good.

0:22:20.770 --> 0:22:21.390  
Andrew Williams  
Across both.

0:22:30.40 --> 0:22:39.680  
Andrew Williams  
Because I don't think we had a an agenda, I knew I wanted to follow up on those conversations that we've had with Mark and with Georgina and others.

0:22:42.40 --> 0:22:43.110  
Andrew Williams  
In the mean time.

0:22:45.20 --> 0:22:45.690  
Andrew Williams  
UM.

0:22:47.530 --> 0:22:52.640  
Andrew Williams  
Party. Or maybe, you know. Did you have other to DOS that came out of that last meeting that we ought to bring up here?

0:23:3.160 --> 0:23:3.610  
Andrew Williams  
Umm.

0:22:52.210 --> 0:23:6.740  
Vipina Keloth  
Yeah. We had a couple of action items. One is this that you discussing it with Mark and then coming up with a concrete paragraph on the study scope and which would be used by Marty to build an IRP template.

0:23:8.610 --> 0:23:19.490  
Vipina Keloth  
Uh, then uh. We were supposed to collect all available SDH pipelines, so there is one document that initially I have added some existing studies.

0:23:20.670 --> 0:23:31.800  
Vipina Keloth  
Uh, and some tools that are available for, you know, SDH extraction. So that part is going on. I think I have a bit more that needs to be filled up.

0:23:33.280 --> 0:23:34.730  
Vipina Keloth  
Yeah then.

0:23:36.50 --> 0:23:43.80  
Vipina Keloth  
You and me and Doctor Sue were supposed to develop the initial protocol, so all these were the action items from the last meeting.

0:23:45.300 --> 0:23:49.900  
Alvarez, Marta  
And I sort of think the protocol might be something we would need before we work on the IRB template.

0:23:52.850 --> 0:23:53.80  
Vipina Keloth  
OK.

0:23:54.780 --> 0:23:55.130  
Andrew Williams  
Yep.

0:24:0.20 --> 0:24:8.500  
Andrew Williams  
Are there any people who want to present something that they've done? I mean, I'm going to start putting some things in the chat that probably go into this lit review that are.

0:24:9.280 --> 0:24:28.590  
Andrew Williams  
Sort of. You know, ohm up focused stuff. Here's one, but dimma, you've done work in this space and you gave us the link for the hugging face site where stuff. Was there any more detail you wanna go into on that or anybody else?

0:24:37.550 --> 0:24:38.510  
Andrew Williams  
How he did.

0:24:34.860 --> 0:24:39.150  
Vipina Keloth  
I think Dema left, he said. He had some other conflicting meeting here.

0:24:40.80 --> 0:24:42.230  
Andrew Williams  
Yeah, I'm busy grabbing links.

0:24:45.80 --> 0:24:47.920  
Andrew Williams  
Here's the link to his staff.

0:25:4.380 --> 0:25:9.30  
Vipina Keloth  
Yeah, I have added the meeting notes to the files in the channel.

0:25:9.780 --> 0:25:12.370  
Vipina Keloth  
So this information is already there.

0:25:15.0 --> 0:25:15.490  
Andrew Williams  
OK.

0:25:18.400 --> 0:25:18.850  
Vipina Keloth  
This.

0:25:29.450 --> 0:25:29.930  
Andrew Williams  
Yeah.

0:25:33.730 --> 0:25:38.200  
Andrew Williams  
Georgina, there things you've done in this edema looks like you is still here.

0:25:41.790 --> 0:25:42.40  
Andrew Williams  
No.

0:25:43.680 --> 0:25:44.260  
Andrew Williams  
Maybe it's.

0:25:43.430 --> 0:25:50.300  
Alvarez, Marta  
That's weird. It shows up funny on my screen too. Like there's a box where he used to be. That's not filled in, but.

0:25:51.160 --> 0:25:52.130  
Alvarez, Marta  
Yeah. So.

0:25:57.700 --> 0:26:2.440  
Andrew Williams  
We wanna go through the SDH ontology a little bit more. Like what? That.

0:26:4.400 --> 0:26:12.200  
Andrew Williams  
Covers. I know it covers both sort of behavioral and social determinants, right? And we talked about wanting to do both of those.

0:26:13.40 --> 0:26:16.830  
Andrew Williams  
I have slight misgivings about that breadth, but it's OK.

0:26:17.940 --> 0:26:19.560  
Andrew Williams  
I can live with slight misgivings.

0:26:33.170 --> 0:26:34.40  
Georgina Kennedy  
Yes, definitely.

0:26:29.390 --> 0:26:35.30  
Andrew Williams  
George, it sounds like a lot of the work that you been doing is on language and accessibility from people's.

0:26:35.340 --> 0:26:38.190  
Georgina Kennedy  
Yeah, culture, cultural, population bases.

0:26:36.910 --> 0:26:40.940  
Andrew Williams  
What language they speak at home and and language. They're fluent and that kind of thing.

0:26:41.940 --> 0:27:5.120  
Georgina Kennedy  
Yeah, country of birth and trying to figure out the immigration aspects as well. We've got we we have a small subset of patients linked to just raw taxation and visa based data as well which helps with the special interests of refugee populations, which can be very, very small but are highly concentrated in now geographical area.

0:27:21.820 --> 0:27:22.340  
Andrew Williams  
Umm.

0:27:6.580 --> 0:27:34.650  
Georgina Kennedy  
Yeah, that's that's the primary area of interest. But we are also, I was sorry I was in and out the screen looking at the links you had sent and I saw that there's a bunch of, you know, kind of burden measures as well in that baseline model on having faced with anxiety and financial burden and that kind of stuff. I think the psychologists that we work with would be very interested in piloting that and seeing how effective it is in our set as well.

0:27:45.200 --> 0:27:50.960  
Andrew Williams  
So assuming the broad strategy.

0:27:52.840 --> 0:27:57.390  
Andrew Williams  
That I described at the beginning is something we roughly agree to. We're going to have to kind of figure out.

0:27:58.570 --> 0:27:59.90  
Andrew Williams  
Which?

0:28:0.300 --> 0:28:8.430  
Andrew Williams  
Social determinants are mostly of interest, So what are people's thoughts on how to on the process for doing that?

0:28:9.720 --> 0:28:10.970  
Andrew Williams  
After figuring those out.

0:28:13.860 --> 0:28:22.230  
Andrew Williams  
Is it based on a lit review? We kind of say this is what literature suggests are most important, or you do it on the things that we think we're best able to.

0:28:23.380 --> 0:28:31.330  
Andrew Williams  
Collect data on on things that are what? What are the ways in which we should be helping to?

0:28:32.650 --> 0:28:35.420  
Andrew Williams  
Might move the conversation forward about what's the highest priority there.

0:28:36.630 --> 0:28:39.610  
Andrew Williams  
I assume maybe I should never assumed I was assuming.

0:28:40.340 --> 0:28:42.90  
Andrew Williams  
That we don't wanna look at everything.

0:28:44.320 --> 0:28:46.390  
Andrew Williams  
That looking at everything would be.

0:28:49.910 --> 0:29:0.210  
Andrew Williams  
Confusing, but maybe that's bad assumption on my part. What? What are people's thoughts? Should we try and look at everything? Should we select and something? And if so, how to prioritize it?

0:29:2.100 --> 0:29:3.870  
Chinthala, Lokesh Kumar  
You have maybe if we can.

0:29:5.260 --> 0:29:11.40  
Chinthala, Lokesh Kumar  
I mean, look into some subset of SDH, for example social support or.

0:29:11.750 --> 0:29:12.420  
Chinthala, Lokesh Kumar  
Ohm.

0:29:13.780 --> 0:29:17.840  
Chinthala, Lokesh Kumar  
But I mean they cannot incoming stability so.

0:29:18.800 --> 0:29:23.870  
Chinthala, Lokesh Kumar  
Some of other these two things would contribute to uh, type 2 diabetes or.

0:29:24.840 --> 0:29:25.480  
Chinthala, Lokesh Kumar  
Maybe.

0:29:26.460 --> 0:29:29.250  
Chinthala, Lokesh Kumar  
The major risk factors, so if you can.

0:29:31.160 --> 0:29:33.570  
Chinthala, Lokesh Kumar  
Restrict or at least start.

0:29:34.350 --> 0:29:42.760  
Chinthala, Lokesh Kumar  
Focusing on these two small topics and then eventually I think we can add to the SHR.

0:29:45.310 --> 0:29:51.260  
Chinthala, Lokesh Kumar  
I mean pull of contents that we can, I'll look in, look into like from the OR extract from the clinical notes.

0:29:57.480 --> 0:30:4.520  
Liu, Feifan  
Yeah, I agree with the. Starting from a smaller scope, sorry. Can, can you hear me OK?

0:30:5.220 --> 0:30:5.620  
Andrew Williams  
Yes.

0:30:6.980 --> 0:30:13.710  
Liu, Feifan  
Yeah. So I think, you know, last time we touched base some you know first we need to kind of.

0:30:20.670 --> 0:30:21.90  
Andrew Williams  
Umm.

0:30:14.370 --> 0:30:34.700  
Liu, Feifan  
Clag to the list of tools and that can be ready to run. I think that is the because with the smaller scoop and we can find the things that you know can be because it this study involves the different institutions and won't make the tools and to be.

0:30:35.210 --> 0:30:40.730  
Liu, Feifan  
Umm. As easy as possible to be to be used by each institution.

0:30:41.350 --> 0:30:59.280  
Liu, Feifan  
Umm, I think some some tools, you know it may need relatively complex configuration and some tools is easier to use and we want to know what kind of SQL which is can be extracted using the existing tools. So then we can.

0:31:0.620 --> 0:31:2.880  
Liu, Feifan  
Determine you know which.

0:31:3.830 --> 0:31:7.440  
Liu, Feifan  
Uh. Among them can be focused on.

0:31:8.270 --> 0:31:9.150  
Liu, Feifan  
For this study.

0:31:16.480 --> 0:31:20.570  
Vipina Keloth  
You can share the screen with where we are regarding the tools.

0:31:22.0 --> 0:31:26.20  
Vipina Keloth  
So this is a few basic ones location at added this moonstone.

0:31:26.660 --> 0:31:29.130  
Vipina Keloth  
Ohh from University of Utah.

0:31:30.480 --> 0:31:35.380  
Vipina Keloth  
And Doctor Peng last time sent me the N2C2S DoH challenge paper.

0:31:37.190 --> 0:31:44.420  
Vipina Keloth  
The third one is uh. This is from University of Pennsylvania. They did a classification of.

0:31:45.370 --> 0:31:49.630  
Vipina Keloth  
Text classification. Basically sentence classification for these factors.

0:31:50.390 --> 0:31:58.140  
Vipina Keloth  
Then this one. Yeah. This is from the University of Florida. They have something called.

0:31:58.980 --> 0:32:8.820  
Vipina Keloth  
Your uh soda or natural language processing package to extract the social determinants of health. I think they're 1619 of them.

0:32:10.260 --> 0:32:15.620  
Vipina Keloth  
Then this is dima's paper. Uh, these are the factors that they are concentrating on.

0:32:18.90 --> 0:32:20.40  
Vipina Keloth  
Then I found a couple more.

0:32:21.10 --> 0:32:32.500  
Vipina Keloth  
Yeah, this is a game based on the IN2C2 challenge where they are doing any R relation classification and subtype classification. I think one of the institutes that participated in that study, this is again another one.

0:32:33.720 --> 0:32:36.950  
Vipina Keloth  
So yeah, this was basically.

0:32:38.20 --> 0:32:46.180  
Vipina Keloth  
What I found and see tax is used by many people as a sort of a baseline to compare to other machine learning and other models.

0:32:47.60 --> 0:32:52.780  
Vipina Keloth  
So probably that is one thing that we could also look into if we want a baseline.

0:32:55.80 --> 0:33:2.650  
Vipina Keloth  
So I I have added this while also to a a teams channel. Everything is available there now.

0:33:5.640 --> 0:33:12.970  
Chinthala, Lokesh Kumar  
Sorry, I wasn't sure if I. I mean if we should add quickly on my list like we did a small poster where we compared.

0:33:18.720 --> 0:33:19.170  
Vipina Keloth  
Mm-hmm.

0:33:13.310 --> 0:33:32.270  
Chinthala, Lokesh Kumar  
Umm. Let's see. Takes quickly on my list and meta meta map on how it extracts. I mean how they extract as your edge information, but at least the document. I wasn't sure if we should include all this applications.

0:33:34.950 --> 0:33:38.500  
Vipina Keloth  
Yeah, I think we can add those and later device on it.

0:33:39.730 --> 0:33:42.10  
Chinthala, Lokesh Kumar  
Yeah, maybe there's another one. Clamp.

0:33:42.780 --> 0:33:43.610  
Vipina Keloth  
Umm yeah.

0:33:42.620 --> 0:33:46.170  
Chinthala, Lokesh Kumar  
So that is really good for smoking status so.

0:33:47.220 --> 0:33:47.530  
Vipina Keloth  
Yeah.

0:34:0.280 --> 0:34:3.320  
Vipina Keloth  
Doctor Wong, though, was. Was there anything from your group?

0:34:25.940 --> 0:34:26.270  
Vipina Keloth  
Can.

0:34:6.820 --> 0:34:35.790  
Wang, Yanshan  
So the student in my group who was leading the SDH work recently withdraw from the university, so that's why I need to identify another student to work on the SH part. We focus on people with disabilities. So yeah, I think there's for maybe we need to have a smaller scope of study so that you know.

0:34:35.870 --> 0:34:37.660  
Wang, Yanshan  
Then we can make it more feasible.

0:35:10.900 --> 0:35:14.800  
Wang, Yanshan  
I think everyone is busy with Ms. submission today.

0:35:16.480 --> 0:35:25.430  
Wang, Yanshan  
I have a A2 minus request too many street from my student. Did you for my comment by the end of today.

0:35:33.10 --> 0:35:36.680  
Andrew Williams  
There was a scoping review also, right?

0:35:37.60 --> 0:35:38.550  
Vipina Keloth  
Yeah, that was uh.

0:35:39.610 --> 0:35:44.600  
Vipina Keloth  
They're gonna write breja. Yeah. From Doctor Pathak's group.

0:35:48.850 --> 0:35:53.670  
Andrew Williams  
Maybe that's a different scoping review that I was thinking about. This wasn't this scoping review by Hong Fangs Group.

0:35:55.480 --> 0:35:55.820  
Andrew Williams  
That.

0:35:57.10 --> 0:36:10.420  
Andrew Williams  
Wow was also on. I'm not sure if you were on my, my thinking about the wrong topic. I'm getting a little confused. There was a scoping review on NLP for. I thought it was DMH.

0:36:11.960 --> 0:36:14.390  
Vipina Keloth  
Yeah. One systematic review was there from.

0:36:15.700 --> 0:36:18.570  
Vipina Keloth  
Their channel, so that one I remember.

0:36:21.460 --> 0:36:24.580  
Wang, Yanshan  
There is one from University of Minnesota.

0:36:25.930 --> 0:36:26.600  
Vipina Keloth  
OK.

0:36:26.930 --> 0:36:27.420  
Wang, Yanshan  
Yeah.

0:36:55.340 --> 0:36:56.750  
Andrew Williams  
Yeah, I think this is the.

0:36:59.620 --> 0:37:0.370  
Andrew Williams  
The one.

0:37:7.300 --> 0:37:7.990  
Andrew Williams  
Yeah, maybe not.

0:37:16.540 --> 0:37:17.230  
Andrew Williams  
The event.

0:37:21.360 --> 0:37:28.660  
Andrew Williams  
People now or others. What? What do we think is the strategy for coming to grips with the set of things once we identified them?

0:37:29.980 --> 0:37:32.370  
Andrew Williams  
What kind of choices do we need to make and how do we make them?

0:37:35.930 --> 0:37:41.180  
Vipina Keloth  
I basically feel depending upon the you know the up population, what are the main.

0:37:43.50 --> 0:37:56.220  
Vipina Keloth  
You know, sort of things. And so, for example, if I say psychiatry like all this, uh, social support, all these things are the main factors I think that should be one thing that should determine which social determinants that we.

0:37:57.370 --> 0:38:7.700  
Vipina Keloth  
Are focusing on and the second should be the availability of tools that can actually extract them. These two factors would basically.

0:38:8.560 --> 0:38:9.930  
Vipina Keloth  
Inform my choices.

0:38:13.110 --> 0:38:17.540  
Andrew Williams  
OK, let's write up a little something that kind of gets that.

0:38:32.520 --> 0:38:32.970  
Vipina Keloth  
Umm.

0:38:24.140 --> 0:38:48.180  
Andrew Williams  
I guess some version maybe not something that's definitive, but something we can respond to similarly to like my paragraph that I I'm gonna produce something that says. Here's here's what the decision making process is. That kind of is going to narrow down what we're looking at based on the tools and how we should decide amongst them because it's a potentially complex conversation. I think some guidance on how we go into it is going to be useful.

0:38:48.870 --> 0:38:50.230  
Andrew Williams  
That seem right.

0:38:50.950 --> 0:38:51.240  
Vipina Keloth  
Yeah.

0:38:53.80 --> 0:38:58.970  
Andrew Williams  
And if we can have that by our next call, we can start take a look at what we've got and and start to go through that process.

0:39:29.70 --> 0:39:31.730  
Andrew Williams  
Well, that might be enough for today's call. What do you think?

0:39:33.400 --> 0:39:33.870  
Vipina Keloth  
Umm.

0:39:35.550 --> 0:39:36.20  
Andrew Williams  
All right.

0:39:38.720 --> 0:39:44.810  
Vipina Keloth  
So we are still to decide on what exposures and finalize those things too, right?

0:39:47.930 --> 0:39:48.280  
Vipina Keloth  
OK.

0:39:46.750 --> 0:39:50.650  
Andrew Williams  
Absolutely. I think we also, yeah, maybe.

0:39:51.450 --> 0:39:54.550  
Andrew Williams  
Postponed for a little while. A decision as to whether or not it's going to go across.

0:39:55.230 --> 0:39:59.660  
Andrew Williams  
Two things, both cancer and diabetes, but within diabetes definitely want to.

0:40:1.610 --> 0:40:2.450  
Andrew Williams  
I think probably.

0:40:3.320 --> 0:40:4.140  
Andrew Williams  
Following what?

0:40:4.970 --> 0:40:16.690  
Andrew Williams  
One said earlier what definitely pick something that's used a lot so that any kind of stratification within the populations is going to have enough patience to be meaningful at each of the different sites that are doing it.

0:40:18.90 --> 0:40:27.130  
Andrew Williams  
And get Rao and other folks kind of weigh in on what some of the most important clinical outcomes are that we might wanna use to select.

0:40:27.880 --> 0:40:28.430  
Andrew Williams  
That.

0:40:29.700 --> 0:40:30.590  
Andrew Williams  
And then you'll.

0:40:31.340 --> 0:40:39.610  
Andrew Williams  
Have some guidance for us and like how we're reviewing the tools that are available and using that as part of that selection process as well.

0:40:45.30 --> 0:40:51.470  
Andrew Williams  
I cannot remember if Hong Fang's group was going to be participating in this or not. So.

0:40:53.80 --> 0:41:1.590  
Andrew Williams  
Again, I'm just confused. We've got so many NLP studies in Odyssey going to parallel the same time. I'm not. I'm not able to keep all of them straight. I know she said that for.

0:41:2.460 --> 0:41:5.370  
Andrew Williams  
Uh psychiatry and for.

0:41:6.710 --> 0:41:14.480  
Andrew Williams  
Something else, but I can't remember if that's something else was either was cancer or social terms of health. I thought it was this one and I thought.

0:41:15.590 --> 0:41:19.940  
Andrew Williams  
She was one of the authors of this scoping review that I'm looking for and not finding but.

0:41:23.150 --> 0:41:24.360  
Andrew Williams  
Probably misremembering.

0:41:31.900 --> 0:41:36.930  
Vipina Keloth  
They didn't invite her for this, that they remember. So we. Yeah.

0:41:52.430 --> 0:41:52.820  
Andrew Williams  
OK.

0:41:59.430 --> 0:41:59.820  
Andrew Williams  
Mm-hmm.

0:42:9.900 --> 0:42:10.160  
Andrew Williams  
Mm-hmm.

0:41:40.760 --> 0:42:10.650  
Alvarez, Marta  
I can look at the emails that we've sent or that we've, you know, participated in for the three different NLP related studies and just sort of come up with a quick list of who's on what. But I think this meeting in particular originated with Hua, or at least that's what the meeting invites coming from. So I just since it's not us, I just wanna make sure we highlight that it sounds like you, who's going to reach out to invite Rohan and then Andrew, you said you might invite Jody Anne, is that sound?

0:42:10.770 --> 0:42:11.240  
Alvarez, Marta  
Correct.

0:42:11.750 --> 0:42:12.30  
Andrew Williams  
Yep.

0:42:12.620 --> 0:42:15.510  
Alvarez, Marta  
OK. Just so that if it's originating with.

0:42:21.120 --> 0:42:21.400  
Andrew Williams  
Yeah.

0:42:21.460 --> 0:42:23.150  
Alvarez, Marta  
The people are also being added OK.

0:42:23.260 --> 0:42:24.870  
Vipina Keloth  
Yeah, I have noted down that.

0:42:25.380 --> 0:42:34.520  
Alvarez, Marta  
OK, I'll I'll send an e-mail to you. The Pina and Andrew with the three of us just to see where the at least what I can find on the emails.

0:42:37.660 --> 0:42:38.70  
Andrew Williams  
OK.

0:42:39.240 --> 0:42:39.670  
Andrew Williams  
Great.

0:42:41.360 --> 0:42:44.590  
Andrew Williams  
I'm gonna put Jody Anne's e-mail in the.

0:42:45.710 --> 0:42:56.360  
Andrew Williams  
Chat UM. I will send an e-mail to her later, but you'll have it right now for collusion in any invitations going forward.

0:43:1.160 --> 0:43:6.340  
Andrew Williams  
Her name is Jody Anne Mclagan. It's Jody, is spelled with an I.

0:43:16.410 --> 0:43:20.390  
Andrew Williams  
Ohh no, it's not. I don't know where I got that information from. Sorry. I'll put her name in.

0:43:21.540 --> 0:43:24.330  
Andrew Williams  
Rather than describe it, since I'm doing that wrong.

0:43:28.660 --> 0:43:28.980  
Andrew Williams  
OK.

0:43:31.840 --> 0:43:33.450  
Alvarez, Marta  
You say she's that Colombia, Andrew.

0:43:34.280 --> 0:43:35.320  
Andrew Williams  
She is.

0:43:35.730 --> 0:43:36.130  
Alvarez, Marta  
OK.

0:43:45.390 --> 0:43:45.720  
Andrew Williams  
OK.

0:43:47.450 --> 0:43:50.910  
Andrew Williams  
Alright, let's get everybody 10 minutes back. We have our to DOS.

0:43:52.390 --> 0:43:53.30  
Andrew Williams  
Thanks everybody.

0:43:52.930 --> 0:43:53.410  
Vipina Keloth  
OK.

0:43:54.340 --> 0:43:54.530  
Andrew Williams  
Like.

0:43:54.90 --> 0:43:54.730  
Vipina Keloth  
Bye.

0:43:57.0 --> 0:43:57.630  
Georgina Kennedy  
Thanks so much.

0:43:57.530 --> 0:43:57.840  
Liu, Feifan  
Aye.