

Standardize Your Research Data with the NIH Common Data Elements Repository - Transcript (10/04/22)

>>Welcome so much to Standardize Your Research Data with the NIH Common Data Elements Repository. The purpose of today's training is to familiarize you with common data elements, or CDEs, and the NIH Common Data Elements Repository, or the CDE-R. I'm Catherine Staley from the National Library of Medicine's Office of Engagement and Training, and here with me today are my colleagues Brittney Davis, Kate Majewski and Mike Davidson, also from the Office of Engagement and Training. And we are joined by Robin Taylor from the NLM's MEDLARS Management Section (MMS) who will be helping me to answer your questions. Robin has been with MMS for over three years as a Technical Information Specialist and is the product lead for the CDE-R.

During today's training, you'll have the opportunity to ask questions and practice searching in the repository. To keep us on track, we're providing a handout that you can download and follow along with, and Kate is putting a link to the handout in the chat now. If you're working on a small screen and it's not feasible for you to work on a handout during the presentation, feel free to just listen along and we'll share a link to the answer key at the end of class, which you can then use to review. We do have everyone muted to cut down on background noise, so a couple of notes about how we'll be communicating in Zoom today. If you have a question or a comment, please use the chat box and send it to everyone. I will not be able to answer messages that are sent directly to me. Brittney is going to be monitoring your chat questions throughout and I will pause frequently to see if there's anything that I or Robin need to address.

During class, I will also be occasionally asking for your nonverbal feedback using the Reactions feature. And the Reactions feature is accessible from the Reactions button at the bottom of your screen. So please find that button now and give me a thumbs up to indicate that you found it. I'm seeing lots of thumbs up. Thank you. Some devices may not allow that type of interaction, and that's totally OK. If you can't find the Reactions button on your device, please feel free to use chat to respond when I ask you to. Thank you for all those thumbs up. OK, I'm now going to turn off my camera and we will jump in.

Today, we'll explore a variety of ways that CDEs and the repository can support you and your work. But to get us started, I want to put them in the context of the NIH's mission. Both the NLM and NIH have identified Common Data Elements as a tool for facilitating the interoperability of data. Ensuring that biomedical research is interoperable is a step towards making that data FAIR. The FAIR principles recommend that data be findable, accessible, interoperable, and reusable. And the repository is one way that NIH supports researchers in aligning with the FAIR principles. Additionally, in the NLM Strategic Plan for 2017–2027, CDEs support the objective to “connect the resources of a digital research enterprise”. In the NIH Strategic Plan for Data Science, the use of CDEs is encouraged to “improve accuracy, consistency, and interoperability among datasets within various areas of health and disease research”. And finally, you're probably aware of the new NIH Data Management and Sharing Policy that goes into effect in January 2023. Common data elements can be a part of a compliant plan. So if you would like to learn more about that policy, Kate is sharing a link to the NIH Scientific Data Sharing Office

website in the chat. If you know how to locate CDEs in the CDE-R, you'll be better equipped to support all of these goals.

All that being said, today's training is going to specifically focus on how to navigate the repository and find CDEs that will assist you with your data collection process, and this is the agenda for today. First, we'll review what you need to know before using the repository, like what a CDE is, and special repository features that will aid your searching. Then I'll demonstrate the CDE-R live for you, and you'll have the opportunity to practice on your own, too. And finally, we'll review how to save CDEs for later and how to get help if you're stuck.

But before all of that, I would like to get a sense of who you are and your work. Some of you already said hello in the chat, so thank you for that. Please take a moment to tell us in the chat your title or your role and the name of the organization that you're with. Wonderful. I'm seeing a mix of organizations that folks are coming from and your different roles. Great. Great. Thank you for sharing where you're coming from and what your role is.

Next, I would like to know what your comfort level is with using CDEs. So please enter in the chat. The letter that corresponds to your response and your answer options are:

- A. Very comfortable! I use them all the time.
- B. Kind of comfortable. I've used them before but it was a challenge.
- C. Somewhat uncomfortable, I've attempted to use them but it wasn't successful.
- D. I know what CDEs are, but I've never used them
- E. I don't even know what CDEs are!

And that's OK, if you don't, we're going to go over that together. And I see our responses coming in. Looks like we have quite a mix. I'm seeing a lot of Ds. So many folks may be familiar with CDEs but haven't used them before. Excellent. Thank you for those responses.

So we can actually use the data that you just provided to explore some reasons that you might consider using CDEs in your projects. Let's say that I'm writing a report after this training to summarize how it went, and I want to include some data about you, the participants. I could start by entering each of your answers in a spreadsheet, and then I might need to standardize your organization's name and how it's represented. For example, do I write out the name as in the National Library of Medicine? Do I use the acronym of NLM? Or maybe both, with the acronym in parentheses beside the name? This is just a small set of data, so it wouldn't take me too long to standardize, but what if I wanted to compare the responses from all NLM trainings, not just this one? With more data, it could take a lot longer to standardize and I would have to ask other trainers to standardize their data too, so that we could combine our results.

Plus, what if another trainer asked about their audience's comfort level with CDEs, but they asked the question in a different way? So, for example, maybe they asked the audience to describe their level of comfort in their own words. It would be really hard to compare those free text responses to the responses that you just provided. We would have two very different sets of data like the ones on the slide. If my colleagues and I came together and created a survey that

asked all of our participants to answer questions using the same format, we would be able to combine and compare our data quickly and efficiently. We would have results sooner and we could adapt our trainings a lot faster based on who participates. And you probably want to compare your research results to those of your colleagues who are studying similar topics. CDEs can help you do that.

So as we just saw, some of us are very familiar with common data elements and the repository, and for others, it's a brand new topic. So let's review a definition of CDEs so that we're all on the same page, and this definition is the answer to the first question on your handout. Common Data Elements, or CDEs, are standardized, precisely defined questions paired with a set of specific allowable responses used systematically across different sites, studies or clinical trials to ensure consistent data collection. In other words, it's a standardized way of collecting a particular piece of information from a respondent. No matter who was asking the question, it's asked in the same way and respondents have the exact same answers to choose from. This applies to when respondents are filling out their own answers, or if the researcher is recording the answers on their behalf.

So let's look at an example of a CDE together. On the slide is a CDE from the repository. It's titled Usual Place of Healthcare Type. And it asks, "Is there a place that you usually go to when you are sick or need advice about your health? Select all that apply." That is the standardized, precisely defined question of the CDE. Respondents can answer with one of these specific allowable options:

- a doctor's office or community health center
- walk-in clinic, urgent care center, or retail clinic in a pharmacy or grocery store
- emergency room
- a VA Medical Center or VA outpatient clinic
- some other place
- there is no place
- don't know
- refused/not reported
- unknown

Here's another example of a CDE and this one is titled Date of Chest Examination. It asks: "What was the date of the imaging examination?" The specific allowable response for this one is the date, and you'll see other CDEs in the repository that allow responses like the time, a number, or a date. So after viewing those two examples, you might be thinking about how CDEs can fit into your workflow.

First, CDEs can save resources, including your time and your labor. Rather than writing your own instrument or data collection tool, you can find one in the repository. Using CDEs also advances science because it allows researchers to compare data and results across studies. Imagine if you could seamlessly compare your results with researchers across other institutions without having to coordinate your efforts because you're both already using CDEs. Or you might be required to use CDEs to comply with organizational policies or you're assisting researchers

at your institution with writing a data management and sharing plan for a grant. Many grants awarded by the NIH now require or strongly encourage the use of CDEs, so the repository can help you to meet that requirement.

Take the case of the Cure Sickle Cell Initiative from the National Heart, Lung, and Blood Institute, which uses CDEs for all three of these reasons. The initiative's goal is to accelerate the development of gene therapies for sickle cell disease. They're collecting information from those living with sickle cell disease about their attitudes towards clinical trials. To do this, they developed a set of CDEs for any researcher working on this project. With these CDEs already available, this gets the research off the ground faster. This map on the slide illustrates where that data collection has already taken place. And because they're using CDEs, researchers across the country can combine their data easily, regardless of where it was collected from. Additionally, their investigators are strongly encouraged to use CDEs if they've been awarded a grant from the NHLBI.

The focus of our training today, the NIH CDE Repository, or CDE-R, was developed to help you find CDEs. It was designed to provide researchers with CDEs that have been recommended or required by NIH institutes and centers or other organizations for use in research and for other purposes. The NIH CDE Governance Committee reviews and assesses CDEs submitted by ICs, and they work closely with the NIH CDE Repository team at the NLM. Currently, the repository has over 29,000 CDEs from across 18 different organizations and initiatives. In the next part of this training, I'm going to highlight two key features of the repository that can assist you with your search: NIH-endorsed CDEs and forms. And this will prepare you to actually search in the repository in a few minutes.

The repository includes NIH-endorsed and non-NIH-endorsed CDEs. NIH-endorsed CDEs have been reviewed and approved by an expert panel and they meet established criteria, and this is the answer to #2 on your handout, so I'll repeat it. NIH-endorsed CDEs have been reviewed and approved by an expert panel and meet established criteria. This is a new feature that is intended to help you identify CDEs that are ready to use. NIH-recognized bodies like institutes or research initiatives may submit CDEs to the NIH CDE Governance Committee for consideration for endorsement. When the Governance Committee deems the CDEs endorsed, the NLM will publish the CDEs in the repository and designate them as NIH-endorsed.

On this slide is the criteria for endorsed CDEs, and these are also #3 on your handout, so you can follow along there, too. These criteria were developed by the NIH Scientific Data Council. The criteria include:

- clear definition of variable measure with prompt and response
- documented evidence of reliability and validity
- human and machine-readable format preferred
- recommended or designated by a recognized NIH body
- licensing and IP status clear

The expectation is that the designation of CDEs as being endorsed by the NIH and making those CDEs discoverable and accessible through the repository will advance the use of CDEs and NIH supported and conducted research. So for you, this means that you can more quickly identify the CDEs that have met the set of criteria and are ready to use.

If the CDE is endorsed, you'll see this gold ribbon icon in the repository. You can search directly in the NIH-endorsed CDE collection, or you can limit your search results to endorsed CDEs in the repository, which we'll see how to do in a few minutes. Not all CDEs from the repository are endorsed, at least not yet. So when we start working in the repository today, I'll point out some other ways that you can sort through your options.

The second repository feature I want to highlight is Forms. Sometimes CDEs are combined to create Forms in the repository. Forms are a group of questions or variables with specified sets of allowable responses that are used as a set for particular research or clinical reasons. And this is the answer to #4 on your handout, so I'll say it again. Forms are a group of questions or variables with specified sets of allowable responses that are used as a set for particular research or clinical reasons. Forms can collect a lot of data about a participant all at once, like a family history, or they might assess something like pain or the severity of an adverse effect from a medication or procedure. Forms will often include a system for scoring a participant's responses. Once they've completed filling out or responding to a Form's questions, the administrator will tally up their score, and different scores will lead to different conclusions or interventions. (No. Here we go. Went too far.)

So some repository forms recreate validated instruments that were created outside of the repository. So for example, this Audit-C Questionnaire is intended to identify alcohol use disorders, and it was developed by the World Health Organization in 1998, but it's now available in the repository and it includes 3 CDEs.

Other Forms were created in the repository. This includes a form like the BRICS Social Determinants of Health from the NINR, which the Audit-C Questionnaire that we just saw is actually incorporated into. This form is a collection of 63 CDEs related to social determinants of health. And in addition to the Audit-C Questionnaire, it also includes CDEs about access to care, food security, depression, housing and other topics. So the key thing to remember about Forms is that they are composed of individual CDEs. And those CDEs come together to measure or assess something. And because the repository has both individual CDEs and Forms, that means that there are more options for finding the data collection tool that you need. (OK. There we go. I'm having some-- sticky mouse here with my slides.) All right, so now you're ready to search the repository.

Before we dive in, I do want to review what we've covered so far. So Brittney is going to launch a poll with a few questions for you to answer that will check your understanding of CDEs so far. And you can also follow along with these on your handouts under the Check Your Understanding section. I'll give you a minute to answer those and then we'll review them together. OK, Brittney, let's go ahead and close the poll and view those responses. Thank you.

So yeah, you should be able to see everybody's responses on the screen now and we will go over these together.

So what can you find in the CDE-R? And the answer is CDEs and Forms. We're going to look at examples of both and how to find those today. Question #2: endorsed CDEs have been reviewed and approved by the NIH, so those are designated as endorsed. And finally, Question #3: to identify an endorsed CDE, you want to look for the gold ribbon. And don't worry, we're going to look at lots of endorsed CDEs today, so you will definitely be able to identify those by the end of today's training. All right. So thank you for participating in that poll.

OK. So this is a good time to see if anyone has any questions. So I'm going to pause here. Brittney, have we had any questions come in that Robin or I can address?

>>We had a couple of questions come in. The first one for Robin. Can you say Forms are part of CDE? Or, do Forms use CDEs?

>>Hi, I saw that, and it's the latter: Forms use CDEs. So you can think of CDEs as being the individual questions or variables, or those are like alternate names we sometimes call them. And a Form is a curated set of CDEs, so Forms use CDEs. Hope that clears it up for you.

>>And we have another question also for Robin from Cory. Where is the repository, i.e. where is the data stored?

>>So that's a great question. This is a different kind of repository. This is a repository of common data elements. So what we have here are the CDEs, the variables that you can use in research. We are not actually a data repository, so we do not store any research data in this repository. So what we have here are different kinds of variables that are hopefully going to be reused across multiple studies to make those data more interoperable downstream. And that's what we are storing and I hope that helps.

>>And we have another question from Catherine. Is there a relationship between CDEs and data standards?

>>Yes, I mean CDEs are, I mean in a way they can be a type of standard, especially if they, you know, a group for a research initiative mandates that they're researchers all use a certain set of CDEs and that becomes the standard for that project. In addition, many CDEs are linked to broader data standards. You'll see if you start exploring the repository. You'll see that some of them are linked to standard terminologies like LOINC and SNOMED.

>>OK, we're going to take another question from In Hye: It sounds like a CDE is a standardized question and answer, so both the question and answer must be present, not just the question. Is that right?

>>Very close. So the question and the answer-- and sometimes it's an answer list, like it's a standardized list of what we call permissible values, like answers that are allowed. But it might

also be a format. So it might be a question, and there's no answer list to choose from, but the answer might have to be a number, for instance, or it might be text or a date.

>>OK. And that's all the questions for now. We've had a couple more come in that we will take later on.

>>Great. Thank you for those questions, everyone, and thank you, Brittney and Robin. Yes, please keep your questions coming. We'll try to answer as many as we can today and I'll pause a few more times throughout to answer those.

OK, so we are now ready to search the repository. If you'd like, you can follow along with my demonstration on your own device. Kate is putting a link in the chat to the repository. I will try to make the screen bigger when I can, but you can also adjust the size of the demonstration on your own screen by going to the View Options at the top and adjusting the Zoom Ratio. OK, so I will open up the repository now. We will dive in.

OK, so anyone can search the CDE-R and see results for CDEs and Forms, but to access all of the repository's features you'll need a Unified Medical Language System Terminology Service account, or a UTS account. So you'll see the sign in button in the repository's upper right corner to either sign into an existing account or sign up for a new one. I'm already signed in, so my profile name appears there. If you don't have an account, I encourage you to wait until after this training to sign up. You can still follow along on your own device and see almost everything. And as we go I'll point out any differences between what the repository looks like when you are or aren't signed in. If you do have a UTS account, feel free to take a moment and sign in.

Next, I'm going to point out some features of the repository's home page. So the default search is for NIH-endorsed CDEs, and you can switch to searching for all CDEs or Forms by clicking on the tabs above the search box. You can search with a topic, a keyword, or an organization name. You can also select CDEs or Forms from the menu across the top of the page to search or browse either type, and we'll look at how to browse a little bit later. So let's start with searching within NIH-endorsed CDEs using keywords. Currently, there are 125 NIH-endorsed CDEs that come from the Project 5 Collection, which were developed for a wide spectrum of COVID-19 research, including translational, clinical and applied research.

So let's imagine that you're designing a research study and you want to collect data about the medications that a patient being treated for COVID-19 received. We can search within the endorsed group with the keyword medications to see if any CDEs are available on that topic. So I will type medications into the search box and then click the magnifying glass or hit Enter on the keyboard to search. I will make my screen a little bigger. So here are the results.

And I'm going to point out some features of the results page now. Notice the yellow ribbon that indicates a CDE is NIH-endorsed. Also on the results page you can see the CDE title, its steward, who its used by, and a preview of the permissible responses, if they're available. Every CDE has a steward, which is the organization or project that owns the CDE. So if this CDE ever needs an update or it changes in any way, the steward is responsible for working with the

repository to make those updates, and that's the answer to #5 on your handout. If the CDE ever needs an update or changes in any way, the steward is responsible for working with the repository to make those updates. Typically the steward is also who the CDE is used by, although anyone can use it. Below the title of the CDE is its definition, which is a short description of its purpose. And the definition is helpful for determining if a CDE fits your topic. Some CDE definitions also include information about their development. So I'm going to click on the second one in our list titled COVID-19 specific medication type so we can see its full record.

All right, so this is the record for the CDE. At the top, we see a lot of the same information that we saw on the results page, like its steward and who it's used by, and its definition. If you will recall our definition of CDE earlier, it should have a precisely defined question and a set of specific allowable responses, and we need to scroll down to see those. So I'm going to scroll. And the first thing I come to are the permissible values. We're going to come back to those in a moment, because if you keep scrolling, you'll see this naming section. And this section includes the preferred question text for the CDE. So in this case it is "What medications did the patient take to treat COVID-19?" That is the precisely defined question. The permissible values are the specific allowable responses to that question, so if I scroll back up we can take a look at those. They're arranged in a table, so let's review each part of the table. The first column in the table is labeled "Value" and it includes the allowable answers to the question. So in other words, this is the specific data that is being collected and these are the responses that someone answering the question can choose from. The next 4 columns are related to the code including "Code Name," "Code," "Code System," and "Code Description." By default the code system is the NCI Thesaurus, but you can change to different code systems by selecting the boxes beside them and the system options include LOINC, NCI Thesaurus, SNOMEDCT US, and UMLS. You should be aware that if you're not logged into a UTS account, then the permissible value section might be missing some information. If you keep scrolling down the record of the CDE, you're going to see a lot more details about it. We don't have time to explore all of this in depth today, so I'll refer you to the repository help guide for more information, and Kate is putting a link to the help guide in the chat.

OK, so let's try out searching. So I have an exercise for you. On the screen I'd like you to search within the NIH-endorsed CDE collection. For a CDE that collects employment status of the respondent. And once you've found it, give me a thumbs up, with the Reactions button. I'm already seeing some thumbs up, but I'll give everybody a second to do that search. Wonderful. All right. Lots of thumbs. So now let me know, in the chat, one of the permissible values for that CDE. So what is one of the permissible values listed with the CDE that you just found? The answers are rolling in. There's quite a few permissible values to choose from with this one. OK. Thank you for completing that exercise, I'm going to review so we can all take a look at how we found that together.

So I'll open the repository back up. And I'm going to go back to the top of the screen and click on the repository logo so I can start fresh with a new search. OK, so I'll start by searching within the NIH-Endorsed CDEs search box for employment status. I'll search. And this is giving me four results. Based on the definition of the first one, which is "A textual description of a person's employment status." I think that sounds like a pretty good fit, so I'm going to click on the title to

see its full record. And I'll scroll down to permissible values and view the different allowable responses that are listed in that value column. So I think between everybody we listed them all. But as you can see, there are quite a few different options here, like working without pay, employed full-time, employed part time, etcetera. Great. Thank you.

OK, so next we are going to explore searching for all CDEs in the repository. So you'll recall that the repository has over 29,000 CDEs and 125 of them are NIH-endorsed for now and that number will grow. So you might want to search outside of the endorsed group for certain topics.

So I'm going to return to the repository home page and on this home page, I will click on the Search All CDEs tab from above the search box so that we can search the entire CDE collection. OK, this time I want to find CDEs that measure a person's physical activity, so I will type "physical activity" into the search box. And search. And I get a lot of results, almost 400. And because it's so many results, I can use the filters that are on the left hand side of the screen to help find CDEs that actually match what I need. The first thing I'm going to do is try limiting to NIH-endorsed CDEs by clicking on the box beside "NIH-Endorsed." Now when I do that, this gives me four results. But I don't think that any of them are measuring a person's physical activity just based on their titles and their definition.

So in that case I will take the NIH-endorsed filter off by clicking on the X beside that filter name. And now I'm back to my original result group. So if you can't find an endorsed CDE that matches what you need, there's a few other criteria that you can use. First, you can look at the Collections options. These are the organizations or initiatives with CDEs that are related to your search. So if your organization is listed there then you might see if it has any related CDEs. So if I click on the NHLBI to limit to CDEs from them, notice now that I have two more available filters to specific projects and topics from within the NHLBI like the Sickle Cell Disease project. If I want to remove the NHLBI filter from my search, I can just go back up to the top of my search and click on the "X" beside its name, and now I'm back to my original results.

The second thing that you can refer to is the Registration Status. So I will scroll down a little bit so you can see that here - Registration Status. This indicates where a CDE or Form is in the review and vetting process. So if you can't find an endorsed CDE, check to see if any standard CDEs meet your criteria. A standard registration means that the CDE might have been developed by a standards development organization like LOINC, or vetted by some other recognized authority. So I'm going to try limiting to Standard now by clicking on the box beside it. And so that gave me 24 results. And if I just glanced at their titles, they seem to be more related to physical activity, my topic.

There's one more way to filter your results that I want to show you. You can check to see if a CDE is used with any Forms that you're familiar with or are from a widely recognized survey or instrument. So remember that Forms are sets of CDEs. I'm going to show you an easy way to see if a CDE is a part of any Forms. So first you'll click on the CDE that you're interested in. I'm going to choose the third one in our list titled "To what extent are you able to carry out your everyday physical activities such as walking, etcetera." I'll click on that and if I scroll down a little bit, I see a blue button that says "Linked Forms." And when I click on the Linked Forms button, if

it appears on any Forms, I'll see this pop-up window. And it lists the Forms that use the CDE. So here we have two. So if these Forms are ones that you're familiar with, then this might be a good CDE option. And I saw somebody give a nice surprise Reaction emoji to that. Yeah, so great feature to use. All right, we're going to search specifically for Forms in a few minutes too, so you'll see how to do that as well.

OK. This is a good time to pause for questions. So, Brittney, have we had some other questions come in?

>>We have had quite a few questions come in. The first one for you, Catherine, is from Manuel, who, for the searching, wants to know why is physical activity bringing those results?

>>Oh, good question. So. And Robin can correct me if I'm wrong, but when I type in "physical activity" it is searching all of the different fields in the CDE record for those terms. So if "physical" or "activity" or "physical activity" is appearing anywhere in that CDE, then it's going to pull that into my search results, which is why those filters are really helpful. And Robin, if you have more to say about the search mechanics, feel free to jump in.

>>No, I think you've got that right so far and I know we have a lot of other questions too,so.

>>OK, the next question, from Lena: Is there a difference between the CDE Forms and validated instruments from PsycTests, etcetera.? That one's for Robin.

>>Hi, yes. So the short answer to that is that all validated instruments in the repository are Forms, but not all Forms are validated instruments, if that makes sense. That's where you would find the repository has these two sections - CDEs and the Forms - and you would find validated instruments within the Forms. That's not to say that we have, you know, definitely don't have all the validated instruments that you might be looking for. You'd have to search and see if we do, but that's where you would find them. But some Forms, as Catherine pointed out, are, you know, created by groups. They're not particularly validated instruments, but they're sets of CDEs that, you know, sort of belong together and might be used for research purposes.

>>OK. Another question also for Robin, from Adam: Do CDEs respect FAIR principles? If yes, how do you define their persistent identifier?

>>So yeah, great. CDEs do make data more FAIR. They make it more interoperable, which of course is the "I" in FAIR and interoperable data is more reusable. So that's how they help support FAIR data in a very—you know, putting it very simply. And every CDE and Form in the NIH CDE repository has a persistent identifier. It's called the NLM ID. And when you're looking at the details page later you'll see on the identifiers that that's present on every page and doesn't change.

>>OK, we have another question for Catherine. Seonyoung wants to know: Can anyone open an account to use CDE-R?

>>Yes, you'll just need to sign up for a UTS account, and if you go in the repository to the sign in button in the upper right hand corner, you'll see the different login options, and so you can create an account, and then you'll be able to explore all of those features.

>>And she had a follow up question: How do you find out who is the steward for the CDE?

>>Sure, yes. So the steward should be listed with the CDE right below the title. Let me just show you here quickly on the screen. So if this is the record for the CDE, then I have steward listed right below the title so you should see that for each of them.

>>OK, another question for Robin: Are the CDEs mapped to SDTM?

>>They're in the repository and we, NLM, on the repository team, we don't do any mapping to SDTM. However, I had to do a quick search for "SDTM" and found a few results showing me that some CDE working groups that submit their data to us are doing that. It wasn't a lot, but there is some there.

>>And we're going to take one or two more questions before we move on. The next one is from Katie. What is VSAC value set? And that is for Robin.

>>Sure, VSAC, everyone may not be familiar with it. It's the Value Set Authority Center. It's another NLM product which is a tool for creating value sets - basically those answer lists I was talking about earlier, their list of permissible values, and so if you were using a CDE in the repository, or if somebody wanted to put a CDE in the repository and for the answer list reference one of those value sets, they could do that. We do have a link there.

>>And last question for Robin. Eileen wants to know why are there multiple codes and how do you use them?

>>So I think Eileen was talking about multiple codes in the permissible values table. And those codes that's shown there are concept codes and they help describe the meaning of that permissible value. So there are codes that come usually from like the UMLS or from the NCI Thesaurus, and sometimes it will take more than one code there to help them adequately describe the meaning of the value.

>>OK. That's all the questions for now, just in the interest of time.

>>Sounds good. Thank you, Brittney. Thank you, Robin. Yes, please keep your questions coming. We will try to answer as many as we can today and I'll also make sure that you know how to get help to answer any questions we don't address at the end of this training.

All right, so I'm going to open up the repository and go back to the repository home page, because the next thing we're going to do is review strategies for when you are searching for CDEs and you're struggling to find one related to your topic. So let's say you want to use CDEs that measure pain or have to do with pain in some way. So I'm going to switch to searching for all CDEs by clicking on the Search All CDEs tab above my search box, and I'm just going to

type pain into my search box and search. So we get over 1,000 results with the search. On the one hand, this is great because it means that we have a lot of options for finding a relevant CDE. On the other hand, that is way too many results to sort through.

So take a moment to glance at these results and consider how you might narrow down these results. What are some filters or other strategies that you could use to narrow your results down? Let me know in the chat one way that you might narrow these down and think about some of those strategies we just reviewed together. Oh, so good answers here. Yes, someone said additional keywords. We're going to look at that. Joanne says NIH gold, which I think referring to NIH-endorsed, yes, we can filter to NIH-endorsed CDEs. Someone says that we can select the steward, and one thing that we could do is select the Collection. So if we want to limit to the organization or initiative that we're associated with or familiar with, we can do that. Excellent. And the other thing is limiting to Standard under registration status. So those are the three strategies that we reviewed earlier together, and those are also the answer to #6 on your handout, so I'll go over them again. So you can limit to NIH-endorsed CDEs, limit to a specific Collection, or limit to Standard under registration type.

OK. So another strategy is to get more specific with your search by adding additional terms, which somebody suggested in the chat. And that's exactly right. So for example, is there a type of pain or pain in a specific part of the body that you're concerned with, like a limb? Is there a condition or disease that the pain is associated with, such as a fracture? Is there an activity or movement that you're trying to gauge pain during, like walking? You can add these terms to your search to see if a more specific CDE is available, so I'm going to add walking to my search, in the search box at the top of the page. And then I'll click search. And so this retrieves 26 results. So by getting more specific with my search terms, it really helped me to narrow my focus and my options. And keep in mind that you can combine these strategies to further narrow down your results.

OK, let's do another search. Maybe I want to find CDEs that record body mass index. So I'm going to go back up to the top of my results page. I'm going to delete my previous search, and this time I'm going to type in body mass index so I can search all CDEs for body mass index and I'll click search. So with this I get about 20 results and in those 20 results you can probably find what you need. However, unlike our previous example with pain, when we got over 1,000 results and we wanted to narrow our focus, we might want to broaden the search a little bit to make sure that we're seeing as many potential CDEs as possible for our topic. So one way to broaden our search is to consider any alternative terms, abbreviations or acronyms that could be used for your term. So some CDEs might use a different word to describe the same topic that you're searching for. So what's another term or abbreviation or acronym that someone could use instead of body mass index? Tell me what you think in the chat. What could be an alternative?

Great. Yeah, BMI. I'm already seeing that from a couple of people. Yeah. So the acronym BMI is a pretty common acronym for body mass index, so we can try searching with that as well. Maybe there are CDEs that use that term.

So one way to also search for BMI in addition to body mass index is to use the word "OR" to connect the words in the search box, and this will search the repository for CDEs that contain either term. And "OR" is the answer to #7 on your handout. That to explore-- or excuse me, to expand your search, add similar terms or acronyms to your search, and connect each term with "OR." So I'm going to go back up to my search box and I already have body mass index. I'm going to type in "OR" and "BMI." And then I'll click search. And I went from 22 results to 48. So it expanded my search when I searched for both of those terms.

Another feature that I want to point out that is useful for finding relevant CDEs is the More Like This button. So I'm going to click on the first CDE in my list, titled "Body Mass Index (BMI)." We can see its full record. And I'm going to scroll down until I see that familiar Linked Forms button that we looked at earlier. Beside it is the More Like This button, and when I click on that, a window opens showing the CDEs that are similar to the one that I'm viewing. So this is a great feature if you find a CDE that's on your topic, but it's not quite what you need and you want to see similar options, and this is #8 on your handout. Click on the More Like This button to see similar CDEs or Forms to the ones that you're viewing.

OK, so now we are going to explore searching for Forms, and that works very similarly to searching for CDEs, so I'll go back out to the repository home page. And this time I'm going to click on the Search Forms button above the search box. And let's say that you want to screen study participants for depression, and you specifically want to use NIDA's self-administered version of the PRIME-MD tool. You can use the information that you have about that tool to try and find it in the repository. So I'm going to type into my search box NIDA PRIME-MD and I'll search. This gives me one result, so I'm going to click on the title to view its details and confirm that this is the Form that I'm looking for. So right away I see our Form at the top of the record. Below that, I want to draw your attention to the description section. The first sentence of the description reads, "The Patient Health Questionnaire (PHQ) is a self-administered version of the PRIME-MD tool for common mental health disorders administered by health professionals." And if I keep reading, I also see that this particular Form is the PHQ-2 which comprises the first two items of the PHQ-9. So maybe after reading this description you decide that you want to see the version with all nine items. What would you search in the repository if you wanted to see the tool that includes all nine items? What's the keyword that you may search in the repository? Let me know what you think in the chat.

Yeah, exactly. Lots of folks are getting it: the PHQ-9. So we see in the description, but that's what the nine item version is called, so we can try searching for that. And the point is, the descriptions can be really useful for not only making sure that you found the CDE or the Form that you're looking for, but also if there are similar ones or different versions of them that you can use for your data collection.

OK, so let's try another exercise. This time we'll try looking for a Form, so see if you can locate in the repository the Clinical Decision Support for Substance Abuse form that comes from the NIDA. And once you've got it, tell me in the chat how many questions it has. So again, looking for a Form. And once you found it, let me know in the chat how many questions it has. I'll give everybody a minute.

Wonderful getting lots of responses. 28 is the consensus, and that's right, yes, this Form has 28 questions included with it. So let's review how to find that together. I'm going to open the repository back up. There we go, and I'm going to go back out to the home page. And I will click on the Search Forms tab. And I'll type in "clinical decision support for substance abuse." And click search and then I see this top one which the name matches. I also see that the steward is the NIDA. And it tells me that there are 28 questions. And I can click on the title to actually view those 28 questions in the record. Great.

OK, there's one more way to find CDEs and Forms in the repository, which is browsing. So I'm going to go back out to the home page, start fresh. You can use the CDEs and Forms buttons from the top of the home page screen to browse. So let's look at Forms. I'm going to click on Forms from that menu. And from here I'll scroll below the search box and here are the different Collections with available Forms that we can browse. So let's click on the NICHD to see what forms they have. So I get 15 forms. And I want to point out that there are filters for limiting to different projects within this Collection on the left hand side. So if you're looking for a specific group of Forms then you can look here. And I can click on one of those and then I see the Forms within that specific project.

If you want to browse CDEs, this process is the same except I will click on the CDEs button from across the top of the home page. And then if I scroll down I can see the different Collections and I can click into those to browse from there. So let's try browsing together now.

I've got a third exercise for you. What is the name of one of the NEI CDE collections that you can browse in the repository? So we're looking at CDEs and we want to browse and what are the different collections within the NEI that we can browse? Let me know in the chat.

Great. Getting a lot of answers here. We'll give folks another couple seconds to take a look. OK, great. Yeah. So let's take a look at how to find those together. And I think lots of folks have gotten the answer in the chat and we'll just take a look at how to do that as a group. So I'm going to open the repository back up. And so I am already on the page to browse CDEs. And so what I want to do to find what collections are within the NEI is scroll down and I see NEI and I'll click on that. And so now I'm seeing all of these CDEs from the NEI and they have two Collections within the NEI that I can limit it to. So the LASIK Quality of Life Collaboration project has 126 CDEs and the eyeGENE has 109. So those are the two collections that you can then limit to within NEI. OK, great.

All right. So the final feature that I want to demonstrate is how to save CDEs and Forms that you find so that you can come back to them later. The CDE-R has a feature called My Boards where you can create a board which is kind of like a folder to store CDEs and Forms that you need later. And "My Boards" is the answer to #9 on your handout. In order to save and access boards, you'll need to be logged in to your UTS account. And like I mentioned at the beginning of the training, you can sign up for a UTS account if you don't already have one, but you will need it in order to save CDEs and Forms and return to them later.

OK, so I'm going to go back out to the home page to walk through using My Boards. So once you're signed in to create and access your boards, click on "My Boards" from the menu across the top of the page. And it actually signed me out, so I'm just going to pause my share and sign back in really quick. OK, here we go. Just timed out on me there.

All right. So again, I'm going to click on "My Boards" from across the top of the page. And when I'm signed in, I will then see the boards that I've created. So I have three here. Notice that one of them in the middle is for Forms and the other two are for CDEs. You can't save them to the same board, so you'll need to make separate boards for each type. I'm going to select my Forms board so you can see what that looks like. And so here are the different Forms that I've saved to it, and it gives me a brief description of them if those are available and I can just click on the title of the Form to go view it in the repository. I'll go back to my board here. One other thing to show you on the screen. If you want to delete a Form from your board, you just click the little pin symbol beside its title and that will unpin it from your board. But what you probably want to see is how to add them to your boards. So let's go back to our search and find some CDEs and Forms that we can add to our boards. So I'm just going to do a really quick search of all CDEs for diabetes, because I know that there will be a lot of CDEs on that topic. And when I find a CDE in the results list that I want to add to a board, I can just click on the little pin beside its name. And when I click on that little pin, it'll ask me, which board do you want to add it to, and then I can click on the board I want to add it to and it will add it. We'll do that again. Click on the board I want to add it to, and then it'll tell me that it's been pinned to that board. You can also click into a CDE to view its entire record and then select add to board and you'll get the same pop-up window and from there you can click the board that you want to add it to. And Forms can be saved in the exact same way.

OK. I bet we've got some more questions, so I think it's a good time before we wrap up. Brittney, have any more questions come in?

>>We have quite a few. The first question is from Seonyoung: Is there any CDE that is more relevant to basic science researchers? That one's for Catherine.

>>>So I think actually I may need Robin's help answering that question. If there are, I think the question is if there are Collections related to basic science.

>> That is a great question. I don't have a quick answer for you, I'm afraid. Yeah, that's all I can say right now. I can't think of off the top of my head any particular collections that are aimed at basic science, so you might have to explore a bit.

>>Browsing might be a good way to kind of take a look at what is available and seeing the Collections and the projects and initiatives within Collections might be a good starting point for that.

>>OK, so we have another question, also from Seonyoung: if stewards need to update CDEs, do they simply log in and have the edit option then?

>>I can take that one. That used to be the case, but since the new governance policies and procedures were created just like, I mean this last year, I guess is when we really started them, the answer now is no, because everything that's going to get published from here on out in the repository has to go through the endorsement process. And once it's been endorsed, the stewards are not allowed to log in and make any changes; those changes have to go back through the Governance Committee and be approved.

>>And the next question is for Robin. Adam wants to know: what if you choose ONC? It shows birth date as an option. Any idea why is that?

>>I think that was during the search exercise that Catherine did, and I could no longer remember what was happening on screen during that. Catherine, do you know?

>>Ohh, let me take a look. I might be able to come back to that question, Brittany. Let me take a look at-- yeah.

>>Next question from Nancy: having a wide range of CDEs for the same precise concept sort of defeats the purpose of common, doesn't it? Does NIH-endorsement take into account the overlap with other endorsed CDEs?

>>And I will take that. And so yes, absolutely NIH-endorsement definitely takes into account overlap. The aim, of course, is in the name, it's in "Common." We want things to be reused as much as possible, so that they are common to many, many studies, and makes that data that much more interoperable. So sometimes you may see multiple CDEs about the same topic or domain. If you see multiple endorsed CDEs about the same topic or domain, there are probably some subtle but key differences in there, in what's being captured or measured. So I think somebody in a previous training gave the example of employment status. There are several CDEs about employment status, but they're all kind of, they're coming at it from different angles or they have different possible answers and so that's what differentiates them. And we are working towards a way of grouping these similar ones so you can, you know, pull up ones that are similar but not identical, and compare them and see which is best for your purposes.

And while I'm talking about endorsement, I saw there were a couple other quick questions. Somebody asked why the PROMIS measures are not NIH-endorsed. And the answer to that is that the endorsement process is relatively new and it does require that CDE working groups submit their data in a particular format. So we do hope that PROMIS and other NIH groups will all choose to submit their CDEs for endorsement and that will eventually be published in the repository. But it just, you know, it just hasn't happened yet.

And somebody also asked if we endorsed-- if the NIH CDE Governance Committee endorsed the CDEs that are not validated. The answer is that they would prefer that CDEs be validated. But it's--we're also--it's understood that with things especially like with COVID-19 for example, a lot of the, you know, the CDEs that were arising for use in studies about COVID-19 were so new that they couldn't possibly be validated yet, so you know they do take that into account as well.

>>OK. And the next question is for Catherine. Does the order of search terms matter?

>>No, the order of your search terms should not matter. I do want to point something else out. We had talked earlier about when you search, for example, physical activity, what is being searched, and it's each of the different fields of that CDE record, you can put a phrase in quotations, in quotation marks, and that will search for the exact phrase. So that's another strategy that you could try as well.

>>I'm going to jump in, sorry, and just add one comment to that. The order of your search terms should not change the number of results that you're getting if it's the same terms, but it may affect the relevancy rankings of those results. So for instance, I put a search for "birth date" and then "date birth" and I got the same number of results both times, but there were different results on top because the order does matter a little bit to the algorithm.

>>Great. Thank you, Robin.

>>And the next question for Catherine, can you download the Form questions?

>>Yes, you can download the Form questions - either from the record or from your board. And I think on the help guide, which I believe Kate shared a link to that earlier, maybe we can put that back in the chat, there are directions for the different ways to export Forms and CDEs out of the repository.

>>And we have another question from Seonyoung: what is the difference between Linked Forms and More Like This? Clarification: CDE seems mostly for clinical researchers.

>>I can answer the question about the different buttons. Good questions. So Linked Forms is going to show different Forms that the CDE is a part of. So if you're looking at a CDE and you click Linked Forms, it's going to show you, is it a part of any Forms. Same if a Form is within another Form. The More Like This button is going to show similar CDEs. So are they about a certain kind of topic or a similar topic? So that's the difference between those, Robin, maybe better equipped to address the other part of that question.

>>No, I think, I mean, I think she's answered it or-- I'm sorry, they, I don't know. They have answered it themselves. Yeah, the CDEs in the repository are mostly for clinical researchers. That's true. And that's why I was having trouble thinking of examples of basic science ones earlier, because I'm not aware of any Collections off the top of my head for basic science.

>>"If stewards need to update CDEs, do they simply log in and have the edit option then?" That's a follow up.

>>Oh, I think we-- oh, is that a follow up to the answer I gave earlier? I see that. So they're asking if they could log in and then have that change be reviewed by the Governance Committee. That's not the workflow right now. I'll just say that we're working on our workflow, so it doesn't work quite that way yet.

>>OK. And I believe that's all the questions as of right now. If anybody has any questions, please put them in the chat.

>> Great. Thanks, Brittney. OK, so to wrap up, I want to-- Well, goodness gracious, let me just get my slides back in order. OK, there we go.

So to wrap up, I do want to point out a couple ways that you can get help with the repository. So at the top of the repository is the Help menu, and if you hover over that, you'll see a Guides page which provides in-depth instructions for using the repository, and we've shared the link to that a couple times. Keep an eye on the New Features page for any updates, and you can also use the Contact Us button to write to the NLM help desk with your questions.

I will ask Kate to put the handout answer key in the chat so that you have that answer key to take with you in case you missed any of those and to use for your review. Thank you for attending today's session. We are also putting a link in the chat to an evaluation survey and would appreciate you letting us know how we can improve our training in the future. At the end of the survey, you will see a screen that will allow you to get an enrollment code that you need and instructions on how to obtain your MLA CE if you're doing that today. So we do appreciate you filling out that evaluation, and the link is in the chat now.

And thank you so much for attending today's session, and we hope to see you at future NLM webinars. Thanks everyone. I will leave this open for just another minute to make sure everybody can get that evaluation link. Great. All right. Thank you. And we'll close the webinar now. Thanks so much. Bye, bye.