All Payer Claims Database Data Release Committee
Meeting Minutes

Meeting Date	Meeting Time	Location
January 14, 2020	1:00 pm – 3:00 pm	OPM Conference Rm. **2A**, 450 Capitol Avenue, Hartford CT 06106 **Webinar option**: Join Zoom Meeting https://zoom.us/j/176265523 Meeting ID: 176 265 523 Dial by your location +1 646 876 9923 US (New York) +1 669 900 6833 US (San Jose)

Committee Members					
Dr. Patricia Checko, Chair	X	Kristen McClain	X	Lisa Freeman	X
Michael Girlamo	X	Henry Jacobs		Michael Fields	X
Justin Peng	X	Anthony Dias	X		
Sheryl Turney		Kun Chen	X		
Supporting Leadership					
Allan Hackney		Adrian Texidor	X	Tina Kumar	X

Agenda			
	Topic	**Responsible Party**	**Time**
1.	**Welcome and Call to Order**	**Dr. Patricia Checko**	**1:05 pm**
	Pat welcomed the committee and guests and called the meeting to order. A quorum was recognized.		
2.	**Public Comment**	**Attendees**	**1:10 pm**
	No public comment.		
3.	**Review and Approval of October 8, 2019 Minutes**	**Committee Members**	**1:10 pm**
	Pat asked for a motion to accept the October 8, 2019 Minutes. Anthony Dias created the motion. Michael Fields seconded. The motion passed.		
4.	**Application Review and Discussion:** **UConn Health Center, Division of Behavioral Sciences and Community Health**	**Dr. Patricia Checko, Adrian Texidor, Committee Members**	**1:15 pm**
	The applicants from UConn Health Center, Dr. Robert Aseltine and Kun Chen were present. Dr. Aseltine stated this application is an expansion of the previously approved protocol from the August 2016/2017 application. Adrian clarified according to the APCD Policies & Procedures, we are treating this as a new application, rather than an expansion. Every application is treated as a new proposal. The policies and procedures state that when the applicant is finished using the data for that approved purpose, it is to be destroyed with proof of destruction. This application is a new UConn proposal to the Data Release Committee, and not associated with the Office of Health Strategy in regard to what data is available for use. The current application proposal requests the use of the data to be used along with CHIME data (hospital data) from the Department of Public Health (DPH), and opioid related data from the medical examiner's		

	office. CHIME data availability has opened up a lot of avenues for researchers. It has identifiers which have been shared by DPH as part of their agreement. Pat explained the first issue with this data request, is the intention to merge data which we (DRC) are unable to do. The second issue is, the data request is looking for unmasked service dates. These issues were discussed with Dr. Aseltine, and a modified proposal was requested. Dr. Aseltine explained the original application that was submitted, mirrored the methodology and approach of what was approved for risk modeling for suicide that was approved in August 2017. When the original request was submitted, Dr. Aseltine's team had hoped the limited data set would be available for the obvious utility of the data provided. The limited data set the applicants (UCHC) currently have, can only be used for specific state purposes (OHS) initiative approved projects. Such OHS related projects are approved separately, directly by OHS and not part of the DRC process for external requests. Any data integration approach would require use of a limited data set for any kind of matching. This is off the table. There are obvious limitations under the current proposal that should be addressed. For academic purposes, Dr. Aseltine shared that his team is intrigued by methods that allow them to borrow information from data sets that don't allow actual integration, because that is really hard to integrate and procure. They are interested in methodologies where they can statistically borrow information by computing similarity scores; essentially computing correlations based on similarities of patients across the data. They intend to learn about outcomes that they don't have in our data sets by looking at that similarity. Dr. Aseltine said that this request also has a public health interest, as opioids are a hot topic. There is important information in the APCD with the prescribing data and characteristics for providers and patients, and their health outcomes. This can be informative in understanding where opioids are being used appropriately with being prescribed, and what the health consequences and costs of it are. Pat asked what the reference was to fusion. Dr. Aseltine explained this is a data fusion approach, which is under a broader framework of transferred learning. In their initial application related to suicide risk, that was the motivating methodology; and that this is an innovative approach that allows you to have incomplete comprehensive data sets and take advantage of similarities and improve modeling based on the information. Dr. Aseltine clarified there will be no integration of patient level data across the data base, they are requesting use of APCD data that may be informed by what they receive. Dr. Kun Chen further explained the application and the intention of doing transfer learning to transfer one data set to another without integration. Dr. Robert Aseltine shared an example of where they have done this. They are working with CT Children's Medical Center and they have their entire EHR database to look at, 43,000 in service patients from 2018 with limited data sets. In this environment, all you have to use from your own data base is limited knowledge of patients who are typically seeking comprehensive specialty care. Dr. Aseltine added that there is a ton of information available from businesses, like the APCD with similar patients who have data (ambulatory visits, pharmacy info). This information can be used to calculate similarities across the data bases and doing this have generated a better prediction model than from one limited database alone. The motivation, a group like CCMC would not have access to this comprehensive info across care settings and across provider type. If we

	can do this in a fusion approach, they never have to integrate the data. Even if it was legally allowable, it would be incredibly time consuming and challenging. Dr. Aseltine explained that the most germane to their request, is that there is a lot of shared characteristics: refer to document. (External and target data box). There is an outcome only in one of them, they want to be able to predictor that outcome using target data, where they don't have the outcome. The calculation of similarity scores across the databases enables us to use the information without ever integrating-- the data. Pat assumed that the external data base, would be APCD. The target data base will be the CHIME data to look for the similarity. Are the scores then used as a link to the target data set? Dr. Aseltine answered it's a combination of the similarity and risk scores, and it's really the similarity score that is the joint data. Pat asked after I get the risk score, how do I get the similarity score? Kun answered for to determine the similarity score, (look at sheet for example) look at the correlation and it will tell you similarity. With two different sets of patients look at frequency for each patient and compute for two patients. (*need clarity) Pat asked for an outcome example. Dr. Aseltine provided an example in looking at hospitalization for opioid overdose. Here there are a lot of diagnostics codes, and more limited data on that in a hospitalization only database than there is in one that includes ambulatory care. There is pharmacy date from only one of the data bases. These are things we'd want to pull out and make sure we evaluate it understood to all individuals for all persons hospitalized for an opioid overdose. Scores will be created on an individual patient level. Adrian suggested the applicants provide a full methodology of the protocol so that the committee is clear to be as well informed as possible. Dr. Aseltine agreed to provide as much information as necessary. Pat said this method was used in the suicide study and asked if a limited data set was used. Dr. Aseltine answered no, they only had it for 6 weeks. It was done with the DRC approved deidentified data set with 18 patient identifiers removed and masked dates of service. Justin Peng shared that he is torn and is also looking for the research method proposal. He questions what this may lead to, and in coming up with the risks of similarity scores linked to age and gender, how is there no intention of relinking when it is in fact linking. Anthony Dias agreed that the applicants provide an analytic plan to explain this to provide clarity. How is the APCD data set going to be utilized and is this permissible. Lisa freeman commented that in her understanding of this, they are requesting to take data from the APCD database and creating a new database with similarity and then comparing to other sets of data. Her perception is that once we have created the similarity scores, the original data becomes irrelevant.		
5.	**Committee Vote: Data Release Application**	**Committee Members**	**2:00 pm**
	Pat asked for a motion to table the vote of the application at this time and schedule a special meeting once the requested information is available. Anthony Dias created a motion to request additional detail to provide technical clarity related to the methods for arriving at risk scores, and /or similarity scores So the committee can determine whether this rises to a threshold level of co- mingling of the data. Justin Peng seconded the motion. Michael Girlamo and Kun Chen abstained. All in favor, motion passed.		

All Payer Claims Database Data Release Committee
Meeting Minutes

6.	**Update on Data Release Survey**	**Dr. Patricia Checko**	**2:15 pm**
	Pat shared an update on the Data Release survey. Pat thanked Sheryl Turney and Michael Fields for spending time to work on this. The survey will be sent to all of the applicants who ever submitted a proposal whether or not the application was submitted for review and approval. Reasons for developing the survey: • Committee's interest in how the data is being used, and its usefulness. • Committee is concerned with the actual ability of researchers to use the deidentified data set with all of the parameters around it. Pat went through the survey draft and asked the committee to review it and share their feedback with survey team and will finalize it and aim to distribute in February. Chris Wyvill suggested to include a question relating cost of the data, and its value.		
7.	**APCD Application Review**	**Dr. Patricia Checko**	**2:15 pm**
	Pat commented that the current APCD data release application itself is woefully inadequate and insufficient to understand the request, until we talk to them face to face. There will be efforts to organize a work group around this, and the discussion will continue. Adrian Texidor shared that we have received four data request application in the last 24 hours.		
8.	**Wrap up and Meeting Adjournment**	**Dr. Patricia Checko**	**2:30 pm**
	Pat asked for a motion to adjourn. Lisa Freeman created the motion, Michael Fields seconded. The meeting adjourned at 2:30 pm.		