

Special Communication

Sex, obesity, diabetes, and exposure to particulate matter among patients with severe asthma: Scientific insights from a comparative analysis of open clinical data sources during a five-day hackathon



Karamarie Fecho^{a,1,*}, Stanley C. Ahalt^a, Saravanan Arunachalam^b, James Champion^c, Christopher G. Chute^d, Sarah Davis^a, Kenneth Gersing^e, Gustavo Glusman^f, Jennifer Hadlock^f, Jewel Lee^f, Emily Pfaff^c, Max Robinson^f, Eric Sid^e, Casey Ta^g, Hao Xu^a, Richard Zhu^d, Qian Zhu^e, David B. Peden^{c,h,i}, The Biomedical Data Translator Consortium

^a Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

^b Institute for the Environment, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

^c North Carolina Translational and Clinical Sciences Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

^d Johns Hopkins University, Baltimore, MD, USA

^e National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, MD, USA

^f Institute for Systems Biology, Seattle, WA, USA

^g Columbia University, New York, NY, USA

^h Center for Environmental Medicine, Asthma & Lung Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

ⁱ Division of Allergy, Immunology and Rheumatology, Department of Pediatrics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

ARTICLE INFO

Keywords:

Hackathon

Open data

Clinical data

Team science

Application programming interface

Multi-institutional collaboration

ABSTRACT

This special communication describes activities, products, and lessons learned from a recent hackathon that was funded by the National Center for Advancing Translational Sciences via the Biomedical Data Translator program ("Translator"). Specifically, Translator team members self-organized and worked together to conceptualize and execute, over a five-day period, a multi-institutional clinical research study that aimed to examine, using open clinical data sources, relationships between sex, obesity, diabetes, and exposure to airborne fine particulate matter among patients with severe asthma. The goal was to develop a proof of concept that this new model of collaboration and data sharing could effectively produce meaningful scientific results and generate new scientific hypotheses. Three Translator Clinical Knowledge Sources, each of which provides open access (via Application Programming Interfaces) to data derived from the electronic health record systems of major academic institutions, served as the source of study data. Jupyter Python notebooks, shared in GitHub repositories, were used to call the knowledge sources and analyze and integrate the results. The results replicated established or suspected relationships between sex, obesity, diabetes, exposure to airborne fine particulate matter, and severe asthma. In addition, the results demonstrated specific differences across the three Translator Clinical Knowledge Sources, suggesting cohort- and/or environment-specific factors related to the services themselves or the catchment area from which each service derives patient data. Collectively, this special communication demonstrates the power and utility of intense, team-oriented hackathons and offers general technical, organizational, and scientific lessons learned.

Abbreviations: API, Application Programming Interface; COHD, Columbia Open Health Data; CUIMC, Columbia University Irving Medical Center; EHR, electronic health record; FHIR, Health Level Seven International Fast Healthcare Interoperability Resources; ICEES, Integrated Clinical and Environmental Exposures Service; NCATS, National Center for Advancing Translational Sciences; PM_{2.5}, Particulate matter of size $\leq 2.5\text{-}\mu\text{m}$ in diameter

¹ Apart from the lead and senior authors, all other authors are listed alphabetically. Specific author contributions are listed under 'Author Contributions'.

* Corresponding author.

E-mail address: kfecho@copperlineprofessionalsolutions.com (K. Fecho).

<https://doi.org/10.1016/j.jbi.2019.103325>

Received 22 May 2019; Received in revised form 6 September 2019; Accepted 28 October 2019

Available online 30 October 2019

1532-0464/ © 2019 Elsevier Inc. All rights reserved.

1. Introduction

The Biomedical Data Translator Consortium currently comprises 11 teams, representing 28 institutions and ~200 team members. Consortium members have been working to overcome challenges in the sharing, reuse, and integration of data as part of the Biomedical Data Translator program (“Translator”), funded by the National Center for Advancing Translational Sciences (NCATS; [20,21]). The program aims to “bridge the current symptom-based diagnosis of disease with research-based molecular and cellular characterizations through an informatics platform that enables interrogation of relationships across the full spectrum of data types, from disease names, to clinical signs and symptoms, organ and cell pathology, genomics, and drug effects” [3]. The Translator program draws on the combined principles and practices of open science and collaboration, open source software development, agile science and software development, team science, and open community engagement [11,4,8,14]. Open access to clinical data is critical

as part of this effort. However, the many regulations that surround the use of clinical data, while essential to ensure and respect patient privacy, often hinder access to the data for research purposes. Moreover, the ability to share clinical data across institutions is central to open team science and the success of multi-institutional consortia [6]. Yet, numerous sociological and technical barriers challenge such efforts, including cultural norms that promote individual achievement over collective advancement, disciplinary nuances that present integration and communication challenges, and disparities and inconsistencies across resources that encumber efforts to share and reuse data [7].

To overcome these challenges and achieve the ambitious goals of the program, NCATS leadership has fostered a multi-institutional culture and community that promotes collaboration and team engagement, including regular in-person hackathons [25]. Herein, we describe several open approaches to share clinical data that have been developed as part of the Translator program—Clinical Profiles, the Integrated

Pre-Hackathon Planning

- Pre-hackathon planning meetings
- Space allocation
- AV and network planning
- Budget creation
- Hotel block rate
- Transportation arrangements
- Sign-up sheets for working groups
- Designation of working group leads
- GitHub repository
- Hackathon agenda
- Generic clinical workflow



Hackathon Implementation and Management

- On-the-fly coordination meetings
- Space coordination
- AV and network availability
- Open communication channels (e.g., Slack, Email)
- Generic clinical workflow
- Participant sign-up
- Coordination and facilitation of working groups
- Reorganization of clinical working group
- Specific clinical workflow instance



Post-Hackathon Activities

- Follow-up paperwork
- Budget resolution
- Follow-up communications and meetings
- Post-hackathon scientific analyses
- Manuscript preparation

Fig. 1. Hackathon flowchart, showing the three major operational components of the event and highlighting key example activities associated with each component. Pre-hackathon planning helped to guide group discussions during the first one and a half days of the hackathon. Hackathon implementation and management was aided by the pre-hackathon planning, although it did involve a nimble approach to respond to unexpected emergent issues. Post-hackathon activities focused largely on scientific outcomes. (The figure was created by Marcus Anderson.)

Clinical and Environmental Exposures Service (ICEES), and Columbia Open Health Data (COHD) (see [1] for overview). We focus on the successful application of these open clinical data sources in a research study that was conceptualized and executed over the course of a five-day Translator hackathon. While the hackathon writ large was focused primarily on software development and testing of the prototype Translator system, we describe the application of Translator Clinical 'Knowledge Sources' as one productive activity that took place during the broader hackathon event. We conclude with a discussion of technical, organizational, and scientific lessons learned.

2. Materials and methods

2.1. Hackathon structure and methodology

As part of the Translator's grounding in open science, collaboration, and software development, the Translator program has held regular, in-person cross-team hackathons three times per year, for 2 ½ to 3 days each, from the program's inception in 2016 through 2018. In 2019, the program's leadership decided to change the structure of the hackathons and hold two five-day hackathons over the course of the year. This paper describes a subset of the activities and scientific applications that took place during the most recent hackathon, which was held on March 4–8, 2019 at the Renaissance Computing Institute, University of North Carolina, Chapel Hill, North Carolina, USA. (The hackathon agenda and other relevant event information can be found in the [supplementary materials](#).) The overall goal of the hackathon was to continue development and testing of the prototype Translator system in an effort to reveal new connections among existing data sources and facilitate discovery, new research hypotheses, disease classification, and intervention (see graphical abstract). However, the goal of the clinical working group that is the focus of this paper was to evaluate and compare the three open Translator Clinical Knowledge Sources that were developed as part of the Translator program (Clinical Profiles, ICEES, and COHD), and to determine whether those knowledge sources could be used to execute a meaningful research project over the course of the five-day event. The overall structure of the hackathon, as well as the key activities that we believe contributed to the overall success of the hackathon, are provided in [Fig. 1](#), and discussed in greater detail in [Section 3](#).

2.2. Scientific structure and methodology

The structure for the scientific research that is described in this manuscript and designed and implemented during the hackathon was minimal. However, a critical factor was that, prior to the hackathon, Translator team members had developed Clinical Profiles; ICEES; and COHD ([Table 1](#)). Each of these services is openly accessible via an Application Programming Interface (API) and provides access to data derived from the electronic health record (EHR) systems of different academic institutions. Briefly, Clinical Profiles represent statistical profiles of disease and associated phenotypic presentation, derived from observational data on patients from Johns Hopkins Medicine, adapting the Evidence-Based Medicine draft of the Health Level Seven International Fast Healthcare Interoperability Resources (FHIR) standard. Clinical Profiles are designed to be disease-agnostic and currently offer open access to profiles on patients with asthma, diabetes, or Ehlers-Danlos Syndrome. ICEES provides open access to observational data on patients from UNC Health Care System [9]. The clinical data have been integrated with a variety of publicly available data on environmental exposures (e.g., airborne pollutants, socioeconomic factors), using a complex, space- and time-dependent data-extraction and integration pipeline, termed Clinical Asset Mapping Program for FHIR (CAMP FHIR) and FHIR Patient data Integration Tool (FHIR PIT) [15]. Like Clinical Profiles, ICEES is designed as a disease-agnostic service and currently offers access to clinical data on patients with 'asthma-like' conditions. COHD provides open access to observational data on patients from Columbia University Irving Medical Center (CUIMC; [19]). Unlike ICEES or Clinical Profiles, COHD offers occurrence and co-occurrence rates of conditions, drugs, and procedures across all patients at CUIMC. Thus, all three open Translator Clinical Knowledge Sources bear similarities and differences, and each draws on patient data from a different catchment area, thereby allowing for comparisons across patient populations and inferences regarding environmental factors that may differentially affect patient subpopulations. (A more detailed description of each knowledge source is available via the hyperlinks provided in column 5 of [Table 1](#).)

While the development of these open Translator Clinical Knowledge Sources was by itself a valuable accomplishment, a detailed plan for how these knowledge sources would be applied during the hackathon in the context of a meaningful scientific question had not been developed

Table 1
Translator Clinical Knowledge Sources.

Translator Clinical Knowledge Source	EHR System	Brief Description	Available Cohorts ¹	Relevant Hyperlinks
Clinical Profiles	Johns Hopkins Medicine	Statistical profiles of disease and associated phenotypic presentation, derived from observational patient data	Asthma (~60,000 patients), Diabetes (~70,000 patients), Ehlers-Danlos syndrome (~1200 patients)	Clinical Profiles GitHub repository Clinical Profiles Data Model Clinical Profiles Example Output Files crepes LOINC2HPO tool
ICEES	UNC Health Care System	Patient- or visit-level counts of observational patient data integrated at the patient and visit level with a variety of environmental exposures derived from multiple public data sources	Asthma (~160,000 patients with asthma-like conditions; ~23,000 patients in year 2010 study period)	ICEES OpenAPI ICEES GitHub Documentation ICEES API Example Queries
COHD	CUIMC	Counts of observational clinical occurrences and co-occurrences (e.g., co-occurrences of specific diagnoses and prescribed medications), as well as their derived relative frequency and observed-expected frequency ratio	General patient population (5-year dataset of all patients at CUIMC)	COHD SmartAPI COHD Initiative

Abbreviations: API = Application Programming Interface; COHD = Columbia Open Health Data; crepes = clinical research engine for profile extraction & summarization; CUIMC = Columbia University Irving Medical Center; EHR = Electronic Health Record; ICEES = Integrated Clinical and Environmental Exposures Service.

¹ These are the currently available patient cohorts offered by the existing services. Each service is designed to be disease-agnostic and adaptable for any cohort of interest.

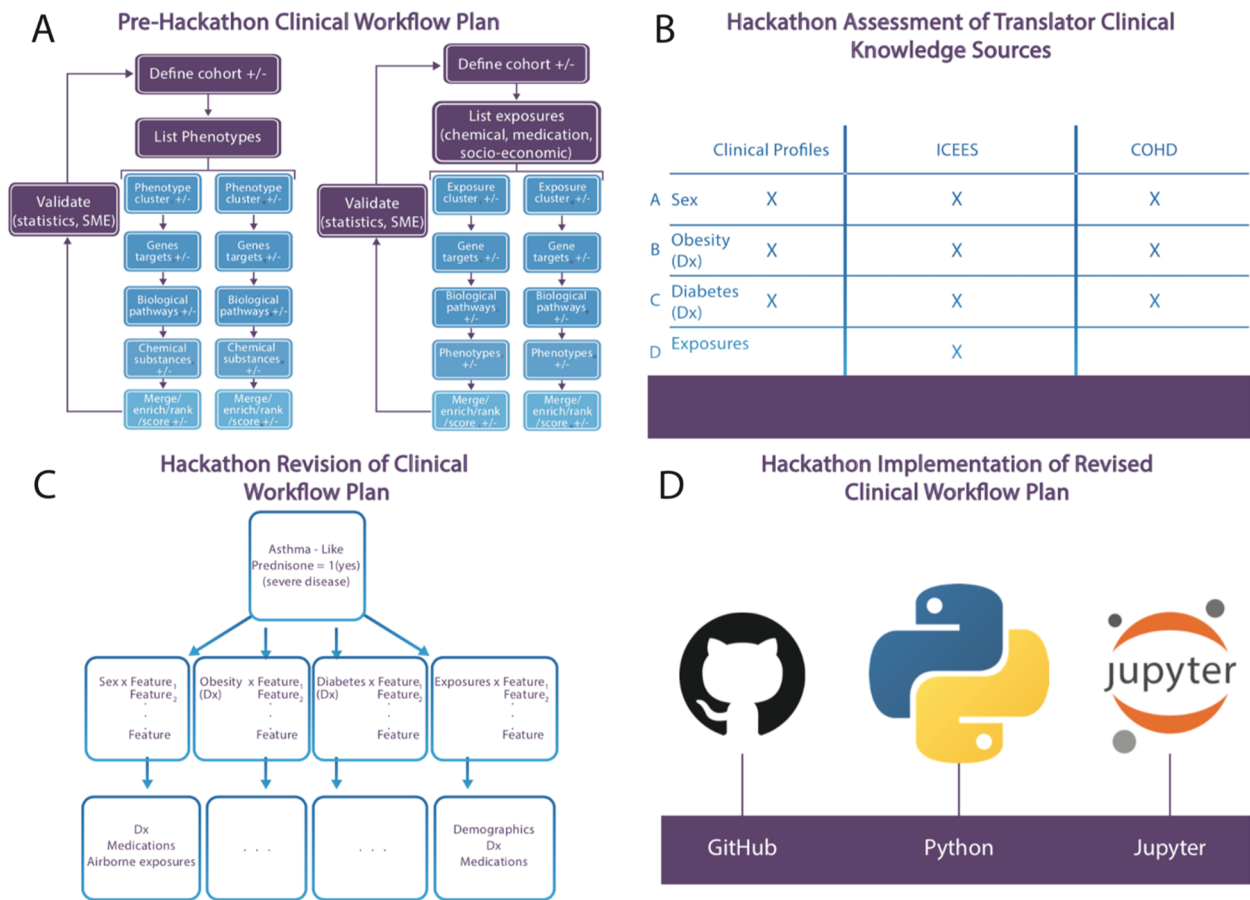


Fig. 2. Scientific flowchart, showing key hackathon aspects of clinical workflow development and implementation. (A) Pre-hackathon planning focused on the development of two generic clinical workflows; this planning helped to guide group discussions during the first one and a half days of the hackathon. (B) Initial hackathon activities focused on evaluating the three open Translator Clinical Knowledge Sources that were developed as part of the Translator program, in terms of the clinical workflows and the capabilities and data available from each knowledge source. (C) Subsequent brainstorming in the context of the generic workflows and the capabilities of the clinical knowledge sources led to the development of a specific instance of the generic clinical workflows. (D) The workflow was successfully implemented and executed over days three and four of the hackathon, and a first-pass analysis of the results was conducted by day five. (The figure was created by Marcus Anderson.)

prior to the hackathon. However, a generic clinical workflow plan [22] was defined prior to the hackathon (Fig. 2).

On day one of the hackathon, the generic workflow plan was reviewed and Translator Clinical Knowledge Sources were evaluated in terms of available data and limitations. This activity led to the development on day two of a specific instance of the generic workflows that aimed to explore the relationship between, obesity, diabetes, exposure to airborne fine particulate matter (particulate matter $\leq 2.5\mu\text{m}$ in diameter [PM_{2.5}]), and severe asthma. An analysis plan also was developed and agreed upon during day two of the hackathon.

Specifically, we employed patient use of prednisone as an indicator of severe asthma. A *post hoc* query of ICEES confirmed that prednisone is a valid surrogate indicator of severe asthma by showing that 16.67% of patients with an asthma-like condition who were prescribed or administered prednisone had ≥ 2 annual emergency department or inpatient visits for respiratory issues versus 5.58% of patients *not* prescribed or administered prednisone ($P < 0.0001$; data not shown, same cohort as presented herein).

Jupyter Python notebooks, executed in GitHub repositories, were used to query Clinical Profiles, ICEES, and COHD, using each service's open API. As a first-step exploratory analysis, the API queries were designed to stratify patients by sex (sex code = male or female), obesity (ICD diagnostic code = yes or no), diabetes (ICD diagnostic code = yes or no), and, for ICEES, exposure to PM_{2.5} (binned using pandas.qcut as [6.77, 47.06] versus (47.06, 114.94] $\mu\text{g}/\text{m}^3$ maximum daily exposure).

A Chi Square analysis, with the significance level set at $\alpha = 0.10$, was used to compare associations between the four stratification variables for Clinical Profiles, ICEES, Clinical Profiles + ICEES, and COHD. The Chi Square analyses were conducted in Python. Results for Clinical Profiles and ICEES were examined independently and also jointly because those cohorts were selected specifically to include only patients with severe asthma. Results for COHD were not combined with the results for Clinical Profiles or ICEES because the COHD cohort included all patients at CUIMC and was not restricted to patients with severe asthma due to limitations of the API (see Table 1).

3. Results

3.1. Hackathon findings

Close to 100 Translator team members attended the hackathon, with most participants joining the hackathon for all five days and self-organizing into six topic-based working groups (see [supplementary materials](#)). The largest working group, with over two dozen participants, was the clinical working group.

During the first day and a half of the hackathon, members of the clinical working group brainstormed to: (1) compare and contrast the capabilities and data available through each of the Translator Clinical Knowledge Sources; and (2) develop a specific instance of one or both workflows that could be implemented using all three services and

executed by the end of the hackathon. The brainstorming effort was guided by both an expert on asthma and the working group lead, who also was the lead developer of the generic workflows and thus was very familiar with them.

The group recognized that the design of COHD included all patients, without the ability to subset patients with severe asthma, and that only ICEES could provide data on environmental exposures. With those caveats in mind, the specific workflow instance that was developed aimed to examine the relationship between sex, obesity, diabetes, and exposure to fine particulate matter among patients with severe asthma. This decision was made at the end of the morning session on day two of the hackathon and was facilitated by pre-hackathon efforts on the generic workflows. Of note, the decision to focus on asthma largely reflected the fact that a subject matter expert had self-selected to join the group and lead the brainstorming effort. The generic workflows themselves were designed to be disease-agnostic.

Over the next two and a half days of the hackathon, the group implemented and executed the workflow, working largely in Jupyter Python notebooks and GitHub, and then conducted a first-pass analysis of the results. A presentation was then prepared and presented to all hackathon participants on the final day of the event. Of note, the activities and products described here represent just one of many tangible hackathon products of the clinical working group. Smaller sub-groups focused on other activities that were coordinated by the lead of the clinical working group.

3.2. Scientific findings

Clinical Profiles, ICEES, and COHD were queried for data on sex, obesity, diabetes, and exposure to PM_{2.5} among patients with severe asthma (Clinical Profiles, ICEES) or the general patient population (COHD).

When the data were stratified by sex (Table 2), obesity was found to be significantly more common among female patients than male patients across Translator Clinical Knowledge Sources (Clinical Profiles: 30.52% vs 17.50%, $P < 0.001$; ICEES: 18.01% vs 10.65%, $P < 0.0001$; Clinical Profiles + ICEES: 27.07% vs 15.29%, $P < 0.0001$; COHD: 4.80% vs 3.75%, $P < 0.001$). Moreover, Clinical Profiles and ICEES had higher overall rates of obesity than COHD. Results for diabetes were less consistent across Translator Clinical Knowledge Sources. For instance, diabetes was more common among females than males with Clinical Profiles (18.62% vs 13.13%, $P < 0.001$), less common among females than males with COHD (4.82% vs 6.35%, $P < 0.001$), and equally common among females and males with ICEES (22.14% vs 22.52%, N.S.).

When the data were stratified by obesity (Table 3), the results showed that the proportion of females was significantly higher among obese patients than non-obese patients (Clinical Profiles: 79.18% vs 64.75%, $P < 0.001$; ICEES: 74.71% vs 61.58%, $P < 0.0001$; Clinical Profiles + ICEES: 78.32% vs 63.73%, $P < 0.0001$; COHD: 62.67% vs 56.46%, $P < 0.001$). Likewise, diabetes was significantly more prevalent among obese patients than non-obese patients (Clinical Profiles: 29.63% vs 12.32%, $P < 0.001$; ICEES: 41.57% vs 18.79%, $P < 0.0001$; Clinical Profiles + ICEES: 31.93% vs 14.40%, $P < 0.0001$; COHD: 25.06% vs 4.59%, $P < 0.001$).

When the data were stratified by diabetes (Table 4), obesity was found to be more prevalent among patients with diabetes than among patients without diabetes (Clinical Profiles: 46.36% vs 22.38%, $P < 0.001$; ICEES: 28.60% vs 11.53%, $P < 0.0001$; Clinical Profiles + ICEES: 40.11% vs 19.36%, $P < 0.0001$; COHD: 19.85% vs 3.44%, $P < 0.001$). Clinical Profiles and ICEES had higher overall rates of obesity among patients with diabetes than COHD. As with the results shown in Table 2, the relationship between diabetes and sex was not consistent across Translator Clinical Knowledge Sources.

ICEES was used to examine the relationship between exposure to PM_{2.5} and sex, obesity, and diabetes among patients with severe asthma

Table 2
Associations between obesity, diabetes, and sex as revealed by Translator Clinical Knowledge Sources: stratification by sex.¹

	Clinical Profiles				ICEES				Clinical Profiles + ICEES				COHD ²			
	Male		Female		Male		Female		Male		Female		Male		Female	
	Total	X ² P value ³	Total	X ² P value	Total	X ² P value	Total	X ² P value	Total	X ² P value	Total	X ² P value	Total	X ² P value	Total	X ² P value
Obesity Dx	1414		2597		4011		1170		1900		3767		5911		935,389	
	82.50%		69.48%		73.57%		81.99%		84.67%		72.93%		76.81%		95.66%	
	300		1141		1441		257		344		1398		1785		75,229	
Yes	17.50%		30.52%		26.43%		18.01%		15.33%		27.07%		23.19%		4.80%	
															4.34%	
Diabetes Dx	1489		3042		4531		1111		1744		4153		6275		1,636,891	
	86.87%		81.38%		83.11%		77.86%		77.72%		80.41%		81.54%		94.52%	
	225		696		921		316		500		1012		1421		94,967	
Yes	13.13%		18.62%		16.89%		22.14%		22.28%		19.59%		18.46%		5.48%	
Total	1714		3738		5452		1427		2244		5165		7696		1,731,858	
	31.44%		68.56%		100%		63.59%		100.00%		67.11%		100.00%		56.73%	

Abbreviations: Dx = diagnosis; ICEES = Integrated Clinical and Environmental Exposures Service; COHD = Columbia Open Health Data.

¹ Note that column percentages are provided in the table, except for the total row, which lists row percentages.

² Unlike Clinical Profiles and ICEES, COHD results are not restricted to patients with severe asthma, but rather reflect all patients.

Table 3
Associations between obesity, diabetes, and sex as revealed by Translator Clinical Knowledge Sources: stratification by obesity.¹

Clinical Profiles										ICEES					Clinical Profiles + ICEES					COHD ²				
Clinical Profiles										ICEES					Clinical Profiles + ICEES					COHD ²				
	Obesity Dx = No	Obesity Dx = Yes	Total	X ² P value	Obesity Dx = No	Obesity Dx = Yes	Total	X ² P value	Obesity Dx = No	Obesity Dx = Yes	Total	X ² P value	Obesity Dx = No	Obesity Dx = Yes	Total	X ² P value	Obesity Dx = No	Obesity Dx = Yes	Total	X ² P value				
Sex	1414	300	1714	102.50, P < 0.001	730	87	817	21.69, P < 0.0001	2144	387	2531	132.24, P < 0.0001	721,240	28,081	749,321	1130.30, P < 0.001	721,240	28,081	749,321	1130.30, P < 0.001				
	35.25%	20.82%	31.44%	P < 0.001	38.42%	25.29%	36.41%	P < 0.0001	36.27%	21.68%	32.89%	P < 0.0001	43.54%	37.33%	43.27%	P < 0.001	43.54%	37.33%	43.27%	P < 0.001				
Female	2597	1141	3738	64.75%	1170	257	1427	61.58%	3767	1398	4024	63.73%	935,389	47,148	982,537	56.73%	935,389	47,148	982,537	56.73%				
	64.75%	79.18%	68.56%		61.58%	74.71%	63.59%		63.73%	78.32%	52.29%		56.46%	62.67%	56.73%		56.46%	62.67%	56.73%					
Diabetes Dx	1014	4531	226.40, P < 0.001	1543	201	1744	87.28, P < 0.0001	5060	1215	6275	280.03, P < 0.0001	1,580,517	56,374	1,636,891	58173.80, P < 0.001	1,580,517	56,374	1,636,891	58173.80, P < 0.001					
	87.68%	70.37%	83.11%	P < 0.001	81.21%	58.43%	77.72%	P < 0.0001	85.60%	68.07%	81.54%	P < 0.0001	95.41%	74.94%	94.52%	P < 0.001	95.41%	74.94%	94.52%	P < 0.001				
Yes	494	427	921	16.89%	357	143	500	18.79%	851	570	1421	14.40%	4,59%	18,855	94,967	5.48%	4,59%	18,855	94,967	5.48%				
	12.32%	29.63%	16.89%		18.79%	41.57%	22.28%		14.40%	31.93%	18.46%		4.59%	25.06%	5.48%		4.59%	25.06%	5.48%					
Total	4011	5452	100%	1900	344	2244	100.00%	5911	1785	7696	100.00%	1,656,629	75,229	1,731,858	100%	1,656,629	75,229	1,731,858	100%					
	73.57%	26.43%	100%	84.67%	15.33%	100.00%		76.81%	23.19%	100.00%		95.66%	4.34%	100%		95.66%	4.34%	100%						

Abbreviations: Dx = diagnosis; ICEES = Integrated Clinical and Environmental Exposures Service; COHD = Columbia Open Health Data.

¹ Note that column percentages are provided in the table, except for the total row, which lists row percentages.

² Unlike Clinical Profiles and ICEES, COHD results are not restricted to patients with severe asthma, but rather reflect all patients.

Table 4
Associations between obesity, diabetes, and sex as revealed by Translator Clinical Knowledge Sources: stratification by diabetes.¹

Clinical Profiles										ICEES				Clinical Profiles + ICEES				COHD ²					
Diabetes Dx = No		Diabetes Dx = Yes		Total	X ² P value	Diabetes Dx = No		Diabetes Dx = Yes		Total	X ² P value	Diabetes Dx = No		Diabetes Dx = Yes		Total	X ² P value	Diabetes Dx = No		Diabetes Dx = Yes		Total	X ² P value
Sex																							
Male																							
1489	225	1714	25.30, P < 0.001	633	184	817	0.04, P = 0.9977	2122	409	2531	13.30, P < 0.0003	701.725	47,596	749,321	1921.60, P < 0.001								
32.86%	24.42%	31.44%	P < 0.001	36.30%	36.80%	36.41%	P = 0.9977	33.82%	28.78%	32.89%	P < 0.0003	42.87%	50.12%	43.27%	P < 0.001								
3042	696	3738	67.14%	1111	316	1427	63.59%	4153	1012	5165	57.13%	935,166	47,371	982,537	56.73%								
67.14%	75.57%	68.56%		63.70%	63.20%	63.59%		66.18%	71.22%	67.11%		57.13%	49.88%	56.73%									
Obesity Dx																							
3517	494	4011	226.40, P < 0.001	1543	357	1900	87.28, P < 0.0001	5060	851	5911	280.03, P < 0.0001	1,580,517	76,112	1,656,629	58173.8, P < 0.001								
77.62%	53.64%	73.57%	P < 0.001	88.47%	71.40%	84.67%	P < 0.0001	80.64%	59.89%	76.81%	P < 0.0001	96.56%	80.15%	95.66%	P < 0.001								
1014	427	1441	22.38%	201	143	344	18.79%	1215	570	1785	14.40%, P = 0.0003	56,374	18,855	75,229	5.48%								
22.38%	46.36%	26.43%		11.53%	28.60%	15.33%		19.36%	40.11%	23.19%		3.44%	19.85%	4.34%									
4531	921	5452	100%	1744	500	2244	100.00%	6275	1421	7696	100.00%	1,636,891	94,967	1,731,858	100%								
83.11%	16.89%	100%		77.72%	22.28%	100.00%		81.54%	18.46%	100.00%		94.52%	5.48%	100%									

Table 5
Associations between sex, obesity, diabetes, and airborne fine particulate matter among patients with severe asthma, as revealed by ICEES: stratification by exposure to relatively low versus high levels of fine particulate matter¹.

	ICEES Average Maximum Daily PM _{2.5} Exposure Bin: [6.77, 47.06 µg/ m ³]	Average Maximum Daily PM _{2.5} Exposure Bin: (47.06, 114.94 µg/m ³]	Total	X ² P value
Sex				
Male	461 35.27%	352 37.81%	817 36.41%	1.52, P = 0.6781
Female	846 64.73%	579 62.19%	1427 63.59%	
Obesity				
No	1129 86.38%	765 82.17%	1900 84.67%	7.43, P = 0.0593
Yes	178 13.62%	166 17.83%	344 15.33%	
Diabetes				
No	1051 80.41%	687 73.79%	1744 77.72%	13.77, P < 0.01
Yes	256 19.59%	244 26.21%	500 22.28%	
Total	1307 58.24%	931 41.49%	2244 100.00%	

¹ Note that column percentages are provided in the table, except for the “total” row, which lists row percentages.

(Table 5). The results showed that obesity and diabetes were more common among patients exposed to relatively high levels of PM_{2.5} than among those exposed to relative low levels of PM_{2.5} (obesity: 17.83% vs 13.62%, *P* = 0.0593; diabetes: 26.21% vs 19.59%, *P* < 0.01). There was no relationship between exposure to PM_{2.5} and sex.

4. Discussion

4.1. Hackathon lessons learned

We assert that intensive hackathons provide an opportunity for large, multi-institutional consortia to engage in team science and collaboratively address relevant scientific questions, producing results that are informative and scientifically meaningful (Fig. 1). The clinical working group that convened as part of the Translator hackathon was highly productive; we believe that several key aspects of the hackathon contributed to this success.

First, participants embraced the tenets of open, agile, and team-based science and software development [11,4,8,14], as promoted by the Translator program. In that spirit, the hackathon was largely self-organized and self-led, and an open GitHub repository served as the focal point for software development and related activities. The choice to participate in the hackathon was partially driven by geographical location and the availability of travel funds, but it was largely driven by enthusiasm and desire to participate, as well as positive experiences during prior hackathons.

Second, while NCATS developed a loose agenda for the hackathon (see [supplementary materials](#)), the event was mostly unstructured and comprised of pure ‘hacking’ sessions, although key logistics were carefully prepared prior to the event, including space allocation, hotel reservations, transportation, sign-up sheets for working groups, etc. NCATS set the goals for the hackathon and assigned a lead to each working group in advance of the event, but other team members were free to assign themselves to a working group of their choice.

Third, in addition to the working group lead, the specific scientific question that is the focus of this manuscript was guided by an expert on the use case, i.e., asthma. This helped to keep the interdisciplinary team focused on clinically relevant questions. Note, however, that the choice

of subject matter simply reflected the fact that an expert on asthma self-selected to join the group and lead the brainstorming effort. The generic workflows and the specific workflow instance that was implemented during the hackathon could very well have focused on another subject.

Fourth, the self-organized clinical working group was quite large at the start of the hackathon, and it quickly became clear that the entire group would not be able to efficiently coordinate their efforts on any single activity and thereby produce a viable hackathon product. Thus, with some direction from the group leader, and with the willingness of the group, as well as support from NCATS to reorganize on an *ad hoc* basis and function as more of an ‘unconference’ [5], the larger group divided into smaller sub-groups focused on different scientific questions and hackathon activities. This nimble ‘divide-and-conquer’ approach proved to be quite successful, and other outcomes of the clinical working group will be reported elsewhere. Moreover, the discussions that took place during the hackathon, and the experience itself, led to the conceptualization and initiation of several new projects. For example, several members of the clinical working group are now working to develop a shared Translator Clinical Knowledge Source, one that adheres to the Translator program’s open API standards and allows users to select functionalities and/or services and directly compare results across services, without requiring separate API calls.

We would be amiss to ignore the organizational and technical challenges presented by the hackathon. First, the logistics of organizing a hackathon, particularly one involving a relatively large number of participants, need to be sorted out well in advance of the event itself.

Second, the physical location of the event needs to be conducive to participation by as many team members as possible. This is especially important for multi-institutional bicoastal consortia such as The Biomedical Data Translator Consortium. In this regard, NCATS has rotated the physical location of hackathons between the East and West Coast of the continental United States. While NCATS has opted for in-person hackathons, other options such as online hackathons or community DREAM challenges are possible. These alternatives have the benefit of avoiding travel costs, but they have the potential disadvantages of remote, asynchronous participation. Careful consideration of cost versus efficiency and anticipated outcomes should be considered prior to organizing and hosting any in-person hackathon.

Third, the length of the hackathon needs to be carefully considered. For the Translator project, moving from a 2 ½- or 3-day hackathon to a 5-day hackathon presented a risk, in terms of lack of participation and hackathon burnout [17,18]. Most participants attended all five days of the event, although the majority of participants were from the East Coast, i.e., close to the physical location and therefore requiring less travel time to attend the event. In terms of burnout, informal discussions suggest that participants experienced a certain amount of mid-week burnout, but the energy of the event led to renewed engagement toward the end of the week.

Fourth, prior to the hackathon, a high-level clinical workflow had been developed by the leader of the clinical working group and other team members. In addition, the group leader spent time considering specific instantiations of the workflow and other possible hackathon activities and preparing for them. A challenge was that the composition of the group was not finalized until day one of the hackathon. Nonetheless, the pre-hackathon preparation allowed the group leader to suggest possible workflow instantiations and other hackathon activities at the beginning of the event, something that proved critical when the larger group divided into smaller subgroups.

In terms of evaluating the success of the hackathon, we received informal feedback, but NCATS did not conduct a formal post-hackathon survey. This decision was made largely because prior post-hackathon surveys received a low response rate and were not very informative. We did, however, receive quite a bit of positive verbal and email feedback. For example, participants provided email comments such as: “this was the best Hackathon yet”, “really liked the meeting spaces and the food”, “the meeting felt ‘chill’ despite the intensity and the high bar set by

NCATS regarding expected hackathon deliverables”, “the whole event was awesome”, “this was the best...most engaging and collaborative”. While this feedback was qualitative and perhaps biased, the indication is that participants viewed the hackathon favorably and derived value from it.

4.2. Scientific lessons learned

With Clinical Profiles and ICEES, we were able to target queries on sex, obesity, and diabetes to patients with severe asthma and compare results to those of the general patient population, using COHD. We used these services to replicate established or suspected interactions between sex, obesity, diabetes, exposure to particulate matter, and asthma. For instance, Assad et al. [2] found that body mass index predicts incident asthma (unadjusted hazard ratio of 1.17 per five index units; $P < 0.001$). When those investigators stratified patients by sex, a significant effect was identified for women (unadjusted hazard ratio of 1.19 per five index units; $P < 0.001$), but not men (unadjusted hazard ratio of 0.98; $P = 0.60$). Greenblatt et al. [10] used EHR data to examine the incidence of asthma exacerbations (defined by ICD code for asthma and an order for oral corticosteroid) and identify contributing factors. This group found that females were overrepresented among patients with asthma exacerbations (74.7% versus 59.5% in general population). After controlling for sex and several other demographic factors, significant predictors of asthma exacerbations included chronic bronchitis, sinusitis, emphysema, fluid and electrolyte disorders, class 3 obesity, and diabetes (odds ratios of 2.70, 1.50, 1.39, 1.35, 1.32, and 1.28, respectively). A study by Requia et al. [16] examined energy generation and fuel sales in 117 regions in Canada between 2007 and 2014 and used an over-dispersed spatiotemporal Poisson regression model to estimate risk of diabetes, asthma, and high blood pressure. These authors found a significant association between a two-year increase of $10 \mu\text{g}/\text{m}^3$ $\text{PM}_{2.5}$ and increased risk in the incidence of diabetes, asthma, and high blood pressure (increased risk [95% confidence intervals] of 5.34% [2.28, 12.53], 2.24% [0.93, 5.38], and 8.29% [3.44, 19.98], respectively).

We identified similar results that are shared among the three Translator Clinical Knowledge Sources. We also identified specific differences in findings across the services, suggesting cohort- or environment-specific factors related to the catchment area from which each service derives patient data. For instance, we identified higher rates of obesity and diabetes for Clinical Profiles and ICEES when compared to COHD, suggesting that obesity and diabetes are more common among patients with severe asthma than among the general population, similar to the findings of Greenblatt et al. [10]. COHD has low rates of recall for most diagnoses, however, so any conclusions should be tempered. Nonetheless, a *post hoc* analysis of the raw COHD data behind the API confirmed higher rates of obesity and diabetes among patients with asthma or severe asthma (data not shown). In terms of sex differences, we found that obesity was more common among females than among males for all three Translator Clinical Knowledge Sources, although the association was more pronounced for patients with severe asthma than among the general patient population, similar to the findings of Assad et al. [2]. We also found that diabetes was more common among males than females when using COHD, a finding that is established in the literature (e.g., [12]); however, when we focused on patients with severe asthma, we found the reverse relationship with Clinical Profiles (i.e., higher rates of diabetes among females than males) and no relationship between diabetes and sex with ICEES, suggesting that the presence of asthma influences the relationship between sex and diabetes. Finally, with ICEES, we were able to confirm the results of Mirabelli et al. [13] and Requia et al. [16] on exposure to $\text{PM}_{2.5}$ and risk of diabetes and asthma, and we extended the results to demonstrate an association between $\text{PM}_{2.5}$ exposure and obesity.

4.3. Conclusion

We demonstrated as proof of concept that intense, five-day, in-person hackathons can be used to conduct meaningful science, in addition to their traditional application for software development. While hackathons are not always viewed favorably [17,18], if executed properly, and with the appropriate level of leadership and team commitment, these events can be quite successful. Specifically, the use case described herein demonstrates the ability to leverage the nimble structure of the hackathon and the engagement of participants to conceptualize and execute a research study that used three open Translator Clinical Knowledge Sources derived from the EHR systems of three major academic institutions to replicate known or suspected associations between sex, obesity, diabetes, and exposure to airborne fine particulate matter among patients with severe asthma.

Author contributions

Karamarie Fecho led the clinical working group, contributed to study design and data analysis, and prepared the first draft of the manuscript; Jennifer Hadlock and Qian Zhu contributed to study design and data analysis; David B. Peden led study design, served as expert on asthma, and contributed to data analysis; Casey Ta, Hao Xu, and Richard Zhu implemented and executed the clinical workflow instantiation, including *ad hoc* data harmonization and Python programming code to enable, coordinate, and combine calls to COHD, ICEES, and Clinical Profiles; Qian Zhu, Saravanan Arunachalam, James Champion, Christopher G. Chute, Kenneth Gersing, Gustavo Glusman, Jewel Lee, Emily Pfaff, Max Robinson, Eric Sid provided general intellectual input and resources to support the design and implementation of the clinical research study described herein. All authors reviewed and approved the manuscript for journal submission.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors wish to thank the staff at the Renaissance Computing Institute for hosting the hackathon. The authors also acknowledge and appreciate the leadership and support provided by the National Center for Advancing Translational Sciences, including Christine Colvis, Noel Southall, Tyler Beck, Grayson Donley, Tyler Peryea, Sarah Stemann, and Mark Williams. The authors note that Christine Colvis, Tyler Beck, and Sarah Stemann, in particular, were instrumental in the planning, implementation, management, and overall success of the hackathon. Finally, the authors also wish to acknowledge the intellectual input and hackathon camaraderie provided by Debbi Adelakun, Vinicius Alves, Stephen Appold, Alejandro Valencia, Joyce Borba, Maureen Hoatlin, Eugene Muratov, Charles Schmitt, Lisa Stillwell, Nicholas Tatonetti, and Alexander Tropsha. While these persons did not contribute to the work described herein, they were members of the clinical working group and engaged in other productive activities during the hackathon.

Funding statement

This work was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, grant numbers OT3TR002026, OT3TR002020, OT3TR002025, OT3TR002019, OT3TR002027, OT2TR002517, OT2TR002514, OT2TR002515, OT2TR002584, OT2TR002520, and UL1TR002489. Funding support also was received from the Environmental Protection Agency (EPA CR 83578501).

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jbi.2019.103325>.

References

- [1] S.C. Ahalt, C.G. Chute, K. Fecho, G. Glusman, J. Hadlock, H. Solbrig, C. Overby-Taylor, E. Pfaff, C. Ta, N. Tatonetti, C. Weng, The NCATS Biomedical Data Translator Consortium, Clinical data: sources and types, regulatory constraints, applications, Clin. Transl. Sci. (2019), <https://doi.org/10.1111/cts.12638> *Authors are listed alphabetically. <https://ascp.onlinelibrary.wiley.com/doi/full/10.1111/cts.12638> [E-pub ahead of print].
- [2] N. Assad, C. Qualls, L.J. Smith, A. Arynchyn, B. Thyagarajan, M. Schuyler, D.R. Jacobs Jr, A. Sood, Body mass index is a stronger predictor than the metabolic syndrome for future asthma in women. The longitudinal CARDIA study, Am. J. Respir. Crit. Care Med. 188 (3) (2013) 319–326.
- [3] C.P. Austin, C.M. Colvis, N.T. Southall, Deconstructing the translational tower of babel, Clin. Transl. Sci. 12 (2) (2019) 85, <https://doi.org/10.1111/cts.12595>.
- [4] L.M. Bennett, H. Gadlin, Collaboration and team science: from theory to practice, J. Invest. Med. 60 (5) (2012) 768–775 <https://www.ncbi.nlm.nih.gov/pubmed/22525233>.
- [5] A. Budd, H. Dinkel, M. Corpas, J.C. Fuller, L. Rubinat, D.P. Devos, P.H. Khoury, K.U. Förstner, F. Georgatos, F. Rowland, M. Sharan, J.X. Binder, T. Grace, K. Traphagen, A. Gristwood, N.T. Wood, Ten simple rules for organizing an un-conference, PLoS Comput. Biol. 11 (2015) e1003905 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4310607/>.
- [6] S.M. Burwell, S. VanRoekel, T. Park, D.J. Mancini, Office of Management and Budget, Executive Office of the President. Memorandum M-13-13, Open Data Policy-Managing Information as an Asset, May 9, 2013. <https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2013/m-13-13.pdf>.
- [7] J. Cutcher-Gershenfeld, K.S. Baker, N. Berente, C. Flint, G. Gershenfeld, B. Grant, M. Haberman, J.L. King, C. Kirkpatrick, B. Lawrence, S. Lewis, W.C. Lenhardt, M. Mayernik, C. McElroy, B. Mittleman, N. Shin, S. Stall, S. Winter, Zaslavsky, Five ways consortia can catalyze open science, Nature 543 (7647) (2017) 615–617.
- [8] B. Fecher, S. Friesike, 2014. Open science: one term, five schools of thought, in: Opening Science, pp. 17–47. https://doi.org/10.1007/978-3-319-00026-8_2. http://book.openingscience.org.s3-website-eu-west-1.amazonaws.com/basics_background/open_science_one_term_five_schools_of_thought.html.
- [9] K. Fecho, E. Pfaff, H. Xu, J. Champion, S. Cox, L. Stillwell, C. Bizon, D. Peden, A. Krishnamurthy, A. Tropsha, S.C. Ahalt, A novel approach for exposing and sharing clinical data: the translator integrated clinical and environmental exposures service, J. Am. Med. Inform. Assoc. (2019), <https://doi.org/10.1093/jamia/ocs042> <https://academic.oup.com/jamia/advance-article-abstract/doi/10.1093/jamia/ocs042/5480568?redirectedFrom=fulltext> [E-pub ahead of print].
- [10] R.E. Greenblatt, E.J. Zhao, S.E. Henrickson, A.J. Apter, R.A. Hubbard, B.E. Himes, Factors associated with exacerbations among adults with asthma according to electronic health record data, Asthma Res. Pract. 5 (2019) 1, <https://doi.org/10.1186/s40733-019-0048-7> eCollection 2019. <https://www.ncbi.nlm.nih.gov/pubmed/30680222>.
- [11] J.S. Hammond, Best Practices: Improve Development Effectiveness Through Strategic Adoption of Open Source, Forrester Research Inc., Cambridge, MA, 2009.
- [12] A. Kautzky-Willer, J. Harreiter, G. Pacini, Sex and gender differences in risk, pathophysiology and complications of type 2 diabetes mellitus, Endocr. Rev. 37 (3) (2016) 278–316, <https://doi.org/10.1210/er.2015-1137>.
- [13] M.C. Mirabelli, A. Vaidyanathan, W.D. Flanders, X. Qin, P. Garbe, Outdoor PM_{2.5}, ambient air temperature, and asthma symptoms in the past 14 days among adults with active asthma, Environ. Health Perspect. 124 (12) (2016) 1882–1890, <https://doi.org/10.1289/EHP92>.
- [14] National Research Council, Committee on the Science of Team Science, Board on Behavioral, Cognitive, and Sensory Sciences, Division of Behavioral and Social Sciences and Education, N.J. Cooke, M.L. Hilton (Eds.), Enhancing the Effectiveness of Team Science, The National Academies Press, Washington, DC, 2015.
- [15] E.R. Pfaff, J. Champion, S. Cox, H. Xu, K. Fecho, A. Krishnamurthy, C.G. Chute, C. Overby Taylor, S. Ahalt, All roads lead to FHIR: an extensible clinical data conversion pipeline. Accepted as a conference paper and podium presentation for the AMIA 2019 Informatics Summit, March 25–28, 2019, San Francisco, CA, USA. Pfaff et al. AMIA Summit 2019 Abstract (last accessed March 19, 2019).
- [16] W.J. Requia, M.D. Adams, P. Koutrakis, Association of PM_{2.5} with diabetes, asthma, and high blood pressure incidence in Canada: a spatiotemporal analysis of the impacts of the energy generation and fuel sales, 584–585, Sci. Total Environ. (2017) 1077–1083.
- [17] V. Siva, Are hackathons good, bad, or overrated? hackerearth blog, June 6, 2018. <https://www.hackerearth.com/blog/innovation-management/hackathons/good-bad-overrated/>.
- [18] N. Swanner, Is it time to rethink the hackathon? Dice, February 21, 2018. <https://insights.dice.com/2018/02/21/time-rethink-hackathon/>.
- [19] C. Ta, M. Dumontier, G. Hripsak, N. Tatonetti, C. Weng, Columbia Open Health Data, clinical concept prevalence and co-occurrence from electronic health records, Sci. Data 5 (2018) 180273, <https://doi.org/10.1038/sdata.2018.273>.
- [20] The Biomedical Data Translator Consortium, The biomedical data translator program: conception, culture, and community, Clin. Transl. Sci. 12 (2) (2019) 91–94, <https://doi.org/10.1111/cts.12592>.
- [21] The Biomedical Data Translator Consortium, Toward a universal biomedical data translator, Clin. Transl. Sci. 12 (2) (2019) 86–90, <https://doi.org/10.1111/cts.12591>.
- [22] K.M. Unertl, L.L. Novak, K.B. Johnson, N.M. Lorenzi, Traversing the many paths of workflow research: developing a conceptual framework of workflow terminology through a systematic literature review, JAMIA 17 (3) (2010) 265–273, <https://doi.org/10.1135/jamia.2010.004333>.

Translator March 2019 In-Person Hackathon

Renaissance Computing Institute at UNC Chapel Hill
Europa Center, 100 Europa Drive, Suite 540, Chapel Hill, NC 27517
March 4th-8th, 2019

Monday, March 4th

8:30-9:00 am – Breakfast/Registration
9:00-9:10 am – Introductions and Logistics – RENCI Staff
9:10-9:40 am – State of the Translator – Noel
9:40-10:00 am – What is and is not a TIDBIT? – Tyler B
10:00-10:15 am – Group structure/Room assignments
10:15-10:30 am – Break
10:30-12:00 noon – Hack/Work on Deliverables
12:00-1:00 pm – Lunch
1:00-3:30 pm – Hack/Work on Deliverables
3:30-3:45 pm - Break
3:45-4:40 – Hack/Work on Deliverables
4:40-5:00 pm – 18F presentation - Jessie Posilkin
5:30-7:30 pm – Welcome Reception/Networking Event – 1st Floor Atrium, Europa Center

Tuesday, March 5th

8:30-9:00 am – Breakfast
8:30-8:45 am – Morning Announcements/Logistics Updates
8:45-9:15 am – Presentation from group leads on expected deliverables
9:15-9:35 am – 18F Presentation - Aditi Rao
9:35-10:15 am – Hack/Work on Deliverables
10:15-10:30 am – Break
10:30-12:00 noon – Hack/Work on Deliverables
12:00-1:00 pm – Lunch
1:00-3:30 pm – Hack/Work on Deliverables
3:30-4:00 pm – Break
4:00-5:00 pm – Time for Optional Reports/Requests for Feedback (Check #hackathon-general Slack channel after 3:30pm for agenda)

Wednesday, March 6th

8:30-9:00 am – Breakfast
8:30-8:40 am – Morning Announcements/Logistics Updates
8:40-10:00 am – Hack/Work on Deliverables
10:00-10:30 am – Break
10:30-12:00 noon – Hack/Work on Deliverables
12:00-1:00 p – Lunch
1:00-3:00 pm – NCATS communications strategies – NCATS Office of Communications and publications committee (4th floor IE Conference Room - open to all)
1:00-3:30 pm – Hack/Work on Deliverables

3:30-4:00 pm – Break

4:00-5:00 pm – Time for Optional Reports/Requests for Feedback (Check #hackathon-general Slack channel after 3:30pm for agenda)

5:30-7:30 pm – Happy Hour at [Imbibe Gastropub, 108 Henderson St. Chapel Hill, NC 27514](#) (One-way transportation provided)

Thursday, March 7th

8:30-9:00 am – Breakfast

8:30-8:40 am – Morning Announcements/Logistics Updates

8:40-10:00 am – Hack/Work on Deliverables

10:00-10:30 am – Break

10:30-12:00 noon – Hack/Work on Deliverables

12:00-1:00 pm – Lunch

1:00-3:30 pm – Hack/Work on Deliverables

3:30-4:00 pm – Break

4:00-5:00 pm – Time for Optional Reports/Requests for Feedback (Check #hackathon-general Slack channel after 3:30pm for agenda)

Friday, March 8th

8:30-9:00 am – Breakfast

8:30-8:40 am – Morning Announcements/Logistics Updates

8:40-10:00 am – Hack/Work on Deliverables

10:00-10:30 am – Break

10:30-12:00 noon – Presentation of TIDBITs & next steps from groups 1-4

12:00-1:00 pm – Lunch

1:00-1:30 pm – Presentation of deliverables and next steps from groups 5 & 6

1:30-2:00 pm – Final thoughts

2:00 pm - Adjournment

Working Groups for March 2019 Hackathon in Chapel Hill, NC
Please add your name to the working group that you expect to START OFF WITH on Monday during the Hackathon. (Please also highlight your name with your team color! See Tyler's entry for an example.) You may find the need to jump between groups as the week goes on. If you feel there is another topic that could make for a full working group, please email Tyler Beck at tyler.beck@nih.gov with the idea for consideration.

Group name	1: Biology/ Genetics-focused Modules	2: Chemistry/ Drug-focused Modules	3: Clinically-focused Modules	4: Workflows 1 & 2; Normalize Ranking/Scoring, Create TIDBITS	5: Knowledge Graph Standardization	6: Parallel Module Results Comparisons
------------	--------------------------------------	------------------------------------	-------------------------------	---	------------------------------------	--

[illegible]