Incorporating Real-World Data Research in Training First-Year Medical Students Using OHDSI OMOP and Atlas tools

Pavel Goriacko[2], Parsa Mirhaji[2], Jimmy John[1], Pavan Parimi[1], Erin M. Henninger[1], Selvin Soby
[1] Montefiore Medicine, [2] Albert Einstein College of Medicine at Montefiore

Background

Clinical informatics is a growing field that is impacting all aspects of clinical care. However, medical students have limited hands-on exposure to clinical informatics during medical school.[1,2] This gap in education can hinder their clinical and research skills during later stages of medical training, residency, and fellowship.[3] To address this issue, we implemented a clinical informatics summer internship program for first-year medical students, focusing on training them in the development of computable phenotypes using the Observational Medical Outcomes Partnership (OMOP) common data model and the OHDSI/Atlas platform. The internship program aims to introduce medical students with real-world data, computable phenotyping, informatics tools and methods for data harmonization, collaboration, and networked research, real-world data research methodologies, and provide mentorship necessary to walk through a directed study from beginning to until the results are validated and publishable.

Methods

First-year medical students with an interest in clinical informatics were recruited through the "IMPACT" student summer research internship program. The primary goal of the program was to provide mentorship and necessary guidance to the students in using OHDSI resources such as OMOP-CDM and OHDSI/ATLAS tools to build a computable phenotype that can be validated and shared with other clinicians to be used for real-world observational research projects. The students began by reviewing patient charts that met the phenotype definition, understanding clinical care and documentation in those instances. Subsequently, they learned about the common data types and domains in electronic health records (EHRs) and how they are mapped to the OMOP common data model and OMOP terminology system. With this knowledge, the students attempted to construct complex computable phenotypes in the Atlas platform.

To validate the computable phenotypes, we utilized our locally developed MRN export feature in our local Atlas implementation, enabling the students to export a small sample of patients identified by their cohort definition and manually validate them against data in Epic, our EHR system. This iterative process continued until the students developed and validated a computable phenotype (cohort definition) with high accuracy. The final computable phenotypes were published and made available as public concept-sets and publicly available and shared cohorts in our local Atlas implementation, making them available for reuse by other researchers.

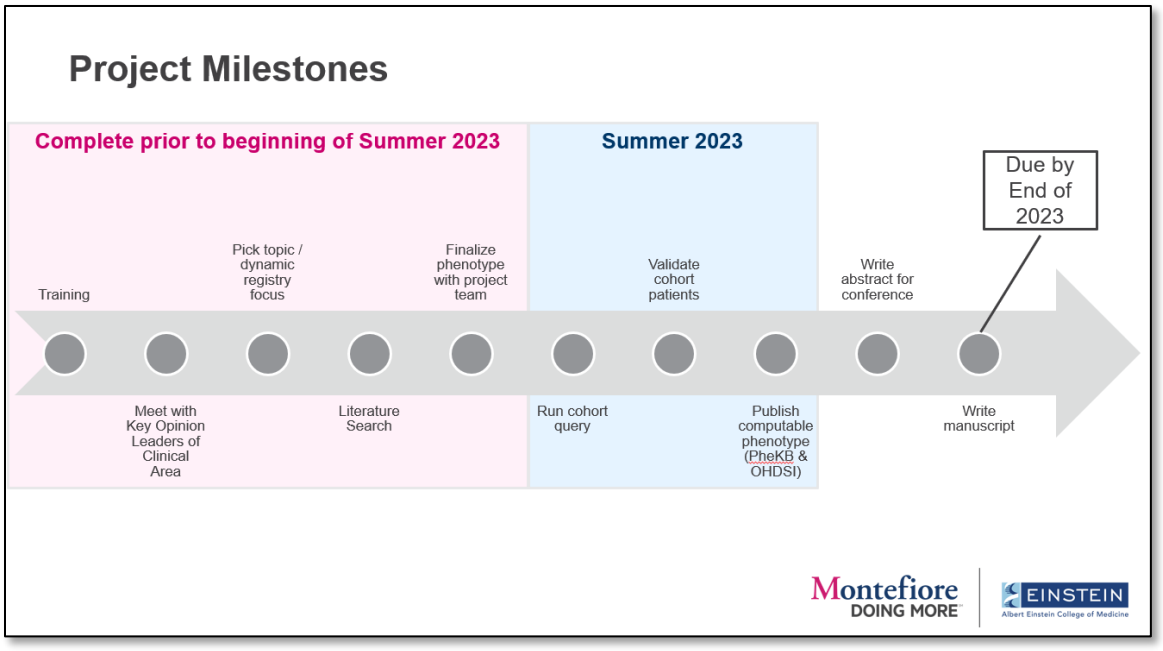

Figure 1: Project timeline, including key milestones

Results

In the pilot year of the program in 2023, five students successfully completed the internship. All of them expressed a strong interest in clinical informatics, with some considering computer science as an alternative career path. The students developed computable phenotypes across diverse clinical disciplines, including cardiology, oncology, interventional radiology, and general surgery. Examples of these phenotypes included new onset atrial fibrillation after non-cardiac surgery, chest port placement and subsequent infection, initiation of first-line therapy for acute myeloid leukemia, and diagnosis of acute gallstone pancreatitis followed by cholecystectomy.

Success factors for the program included effective communication between clinical and informatics preceptors, as well as connecting students with researchers who had tangible research needs for the computable phenotypes. However, the program also encountered challenges such as the students' limited previous exposure to EHR data and the need for IRB approval or exemption for these types of training initiatives. Additionally, clinical preceptors had difficulty understanding the OMOP model, as they were accustomed to local vocabularies.

Conclusion

The pilot program demonstrated the feasibility of providing first-year medical students with the concepts of clinical informatics. Based on our findings, we recommend that medical schools consider formally incorporating similar programs into their curricula. This will enhance students' clinical and research skills, better preparing them for future practice in an increasingly informatics-driven healthcare landscape.

References:

1. Jidkov L, Alexander M, Bark P, Williams JG, Kay J, Taylor P, Hemingway H, Banerjee A. Health informatics competencies in postgraduate medical education and training in the UK: a mixed methods study. BMJ Open. 2019 Mar 30;9(3):e025460. doi: 10.1136/bmjopen-2018-025460. PMID: 30928942; PMCID: PMC6475211.
2. Zainal H, Xiaohui X, Thumboo J, Kok Yong F. Digital competencies for Singapore's national medical school curriculum: a qualitative study. Med Educ Online. 2023 Dec;28(1):2211820. doi: 10.1080/10872981.2023.2211820. PMID: 37186901; PMCID: PMC10187105.
3. Zainal H, Tan JK, Xiaohui X, Thumboo J, Yong FK. Clinical informatics training in medical school education curricula: a scoping review. J Am Med Inform Assoc. 2023 Feb 16;30(3):604-616. doi: 10.1093/jamia/ocac245. PMID: 36545751; PMCID: PMC9933074.