

OHDSI内では、実名での活動になります。
Zoom参加時も「名前は実氏名で」お願いします。

OHDSI Japan evening conference #22

イブニング カンファレンス(第23回)

2021.10.28



OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

オデッセイ
ジャパン



OHDSI 2021 APAC Symposium

11月18日(木) 日本時間 9:00~15:30

参加登録⇒ <https://www.ohdsi.org/APAC/>



2021 APAC Symposium • Nov. 18

Nov. 18 (APAC Time Zone)	Time (Korea time)	Contents	Speaker(s)
Morning	9:00 - 9:25 am	OHDSI State of the Community	George Hlipcak/Patrick Ryan
	9:25 - 9:50 am	OHDSI APAC State of the Community	Mel Van Zandt
	9:50 - 10:00 am	Energy Break	
	10:00 - 10:25 am	OHDSI	Peter Rijnbeek
	10:25 - 10:50 am	FHIR and OHDSI Collaboration	Christian Reich
	10:50 - 11:00 am	Energy Break	
	11:00 - 12:30 pm	APAC Chapter Visions for 2022	Chapter Leads
Lunch Break	12:30 - 12:50 pm		
Afternoon (in GatherTown)	13:00 - 14:00 pm	Workgroup Sessions (Medical images, FHIR, CDM Tables)	
	14:00 - 15:00 pm	Collaboration Showcase	
	15:00 - 16:00 pm	APAC Study Sessions	



OHDSI関連論文

Pubmedで"OHDSI or OMOP"を検索



pubmed.ncbi.nlm.nih.govにて作成

全期間累計：9月215本→10月216本



➤ J Am Med Inform Assoc. 2021 Sep 30;ocab217. doi: 10.1093/jamia/ocab217. Online ahead of print.

Synergies between Centralized and Federated Approaches to Data Quality: A Report from the National COVID Cohort Collaborative

Emily R Pfaff ¹, Andrew T Girvin ², Davera L Gabriel ³, Kristin Kostka ⁴, Michele Morris ⁵, Matvey Palchuk ⁶, Harold P Lehmann ⁷, Benjamin Amor ², Mark Bissell ², Katie R Bradwell ², Sigfried Gold ³, Stephanie S Hong ³, Johanna Loomba ⁸, Amin Manna ², Julie McMurry ⁹, Emily Niehaus ², Nabeel Quresh ², Anita Walden ¹⁰, Xiaohan Tanner Zhang ¹¹, Richard L Zhu ¹¹, Richard A Moffitt ¹², Melissa A Haendel ¹³, Christopher G Chute ¹⁴, N3C Consortium

Abstract

Background: In response to COVID-19, the informatics community united to aggregate as much clinical data as possible to characterize this new disease and reduce its impact through collaborative analytics. The National COVID Cohort Collaborative (N3C) is now the largest publicly available HIPAA limited dataset in US history with over 6.4 million patients and is a testament to a partnership of over 100 organizations.

Methods: We developed a pipeline for ingesting, harmonizing, and centralizing data from 56 contributing data partners using four federated Common Data Models. N3C Data quality (DQ) review involves both automated and manual procedures. In the process, several DQ heuristics were discovered in our centralized context, both within the pipeline and during downstream project-based analysis. Feedback to the sites led to many local and centralized DQ improvements.

Results: Beyond well-recognized DQ findings, we discovered 15 heuristics relating to source CDM conformance, demographics, COVID tests, conditions, encounters, measurements, observations, coding completeness and fitness for use. Of 56 sites, 37 sites (66%) demonstrated issues through these heuristics. These 37 sites demonstrated improvement after receiving feedback.

Discussion: We encountered site-to-site differences in DQ which would have been challenging to discover using federated checks alone. We have demonstrated that centralized DQ benchmarking reveals unique opportunities for data quality improvement that will support improved research analytics locally and in aggregate.

Conclusion: By combining rapid, continual assessment of DQ with a large volume of multi-site data, it is possible to support more nuanced scientific questions with the scale and rigor that they require.

概要

背景 COVID-19を受けて、インフォマティクス・コミュニティは、この新しい病気の特徴を明らかにし、その影響を軽減するために、可能な限り多くの臨床データを集約するために、共同で分析を行うことで団結しました。National COVID Cohort Collaborative (N3C) は、現在、640万人以上の患者を対象とした米国史上最大のHIPAA制限付きデータセットであり、100以上の組織のパートナーシップの証となっています。

方法 N3Cでは、4つの共通データモデルを用いて、56のデータパートナーからデータを取り込み、調和させ、一元化するためのパイプラインを開発しました。N3Cのデータ品質 (DQ) レビューには、自動化された手順と手動の手順があります。その過程で、パイプライン内や下流のプロジェクトベースの分析において、いくつかのDQヒューリスティックが中央集権的に発見されました。現場へのフィードバックにより、多くのローカルおよび集中型のDQ改善が行われました。

結果 よく知られているDQの発見に加え、ソースのCDM適合性、人口統計、COVIDテスト、条件、出会い、測定、観察、コーディングの完全性、使用の適合性に関する15のヒューリスティックが発見されました。56サイトのうち、37サイト (66%) がこれらのヒューリスティックによって問題を発見しました。これらの37サイトは、フィードバックを受けた後、改善を示しました。

考察 連携チェックだけでは発見が困難なDQのサイト間差異が発生した。DQを集中的にベンチマークすることで、データの質を向上させるユニークな機会が明らかになり、ローカルおよび全体での研究分析の向上につながる事が実証されました。

結論 迅速かつ継続的なDQ評価と大量のマルチサイトデータを組み合わせることで、より微妙な科学的疑問を、必要とされる規模と厳密さでサポートすることが可能になります。



OHDSI環境構築

実際にやってみよう

7月: OMOP環境の実装手順(前半)

8月: SynPUFデータ準備

今回は後半、
いよいよATLASを動作させます。