MEDINFO23

8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Understanding clinician EHR data quality to determine reuse in predictive modelling

Melinda Wassell

Central Queensland University

MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Melinda Wassell

M.Res Candidate (Comp. Sci)
Chiropractor
Certified Health Informatician
Adv.Dip OHS
ICSSD (International Sports Chiro Diploma)

Prof. Karin Verspoor
Prof. Kerryn Butler-Henderson

Dr James L Murray
Chaithra Kumar
For providing the research dataset

Ethics approval CQUHREC #0000023392.


MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA
@TheInstituteDH
#MEDINFO23

MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Objectives

To assess the consistency of chiropractors' data entry, to determine fitness for use in ML.

Secondly, to engage senior clinicians in the findings, to address sources of inconsistency and suggest improvements.

MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

The EHR datasource

Structured EHR

Highly structured and high amount of mandatory fields

Clinicians

1 organisation
61 Chiropractors, Physiotherapists and Osteopaths

7yrs

Data collection
2014-2021

57570

Patient care plans

Value based care

Focus on measuring outcomes

Client/Clinician Reporting

Reporting on health trends and outcomes

The data analysis

Variable Selection

Literature review
Used past ML model variables

Clinician Consulting

Consultation with senior clinicians to inform analysis and determine areas for improvement

Data analysis

Missingess
Distribution comparisons and outlier analysis
Alignment to local knowledge

Findings

15921 records

Missingness

All mandatory fields were missing <0.01

Non mandatory field completion depends on perceived value

Findings

15921 records

Missingness

All mandatory fields were missing <0.01

Non mandatory field completion depends on perceived value

Findings

15921 records

Missingness

All mandatory fields were missing <0.01

Non mandatory field completion depends on perceived value

Plausibility comparisons

Dependent on clarity of field definitions, clinician training and clinician skills

Alignment

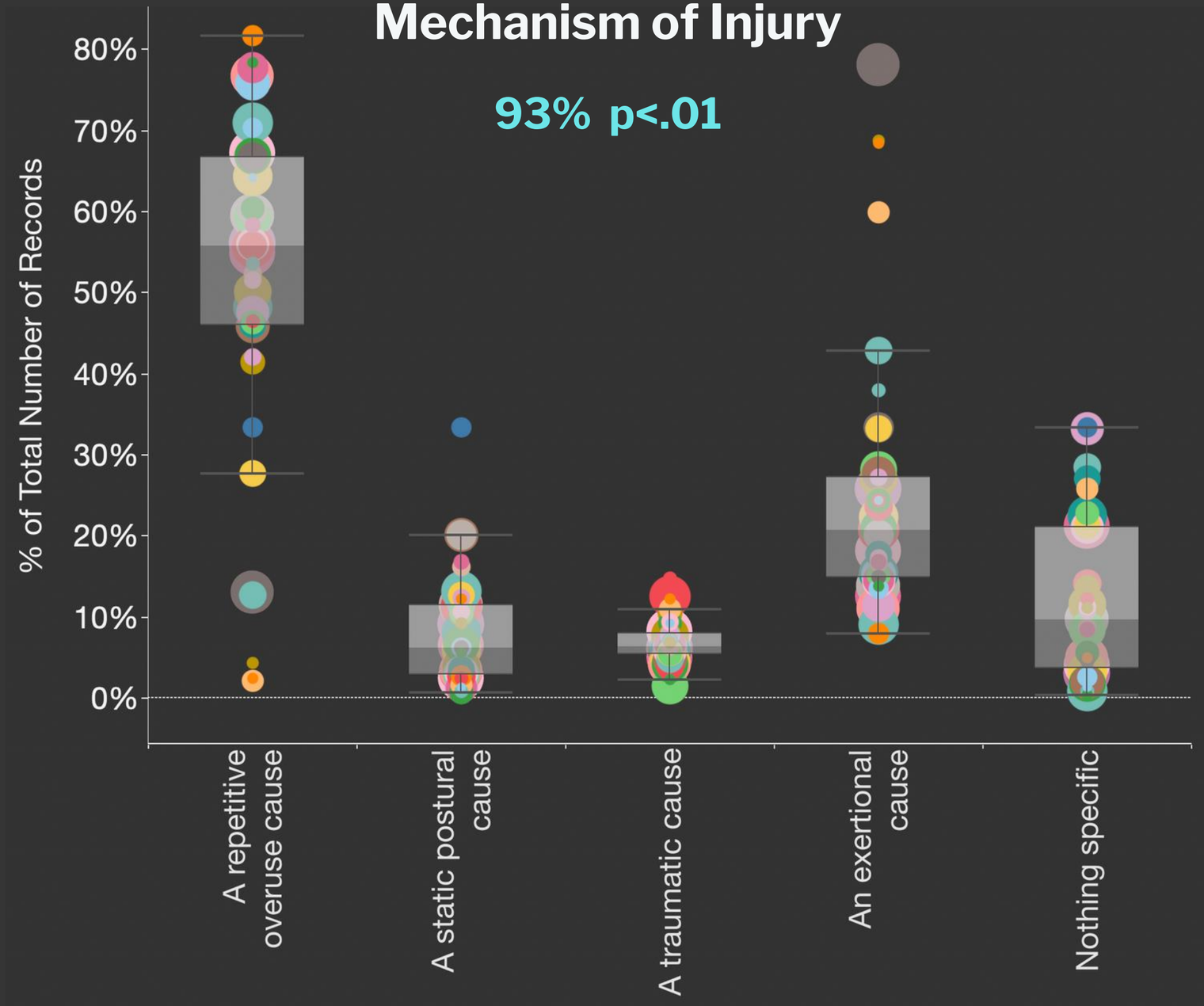
Less alignment on patient interview questions and those seen as unimportant for reporting/decision making

Improvement Suggestions

Update to gold standards, standardisation of questions, definitions and clinician training

MED INFO 23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Clinician Data Variation

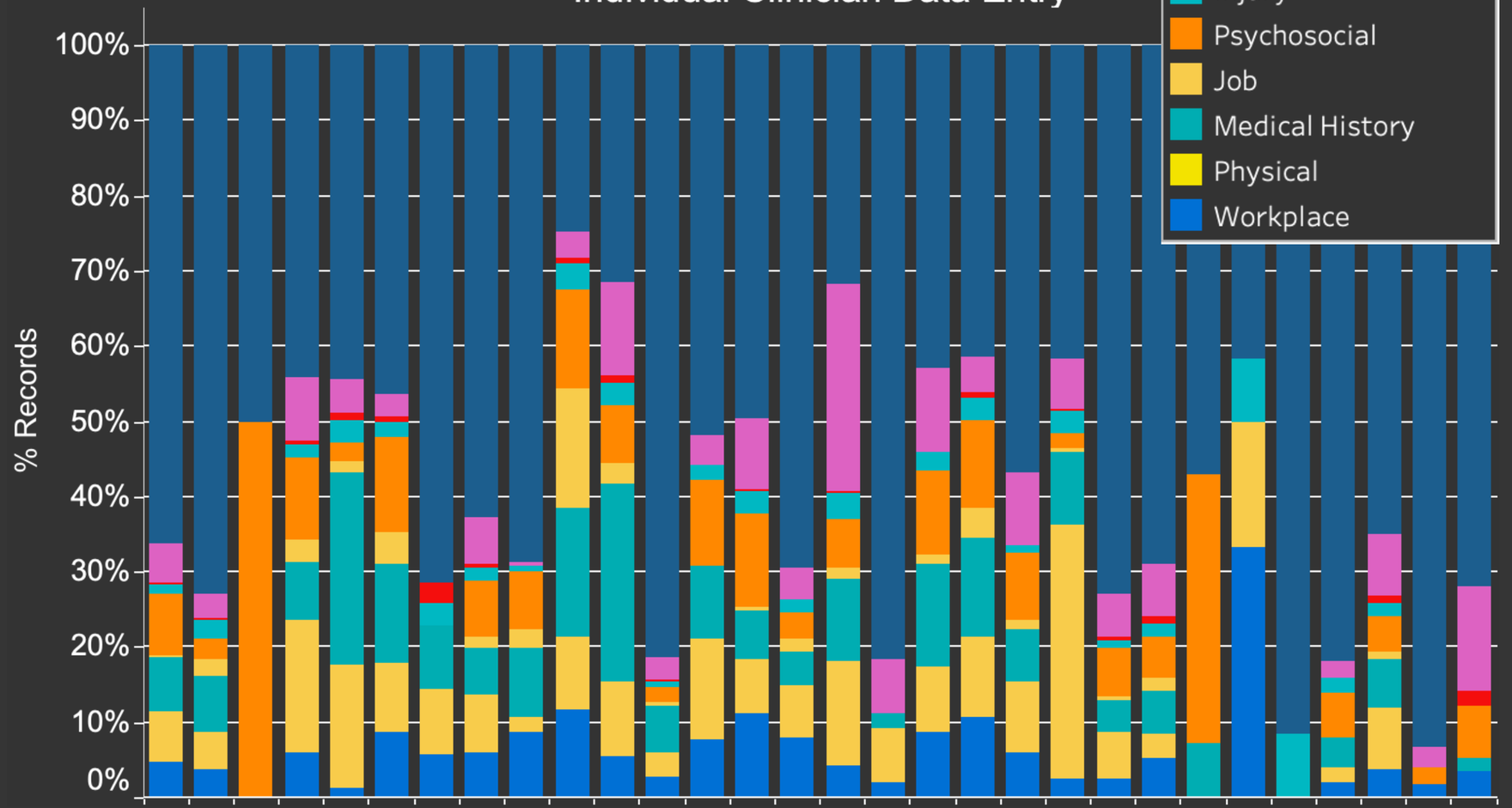


MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Clinician Data Variation

Obstacles to recovery

Individual Clinician Data Entry



Columns = individual clinicians, color = category of obstacle
70% practitioner p values <.01

MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Conclusions

Statistically relevant for reuse vs clinically trusted for reuse.

Clinicians are EHR user subject matter experts - involve them to achieve better analysis

Take aways

The more governance, the more focus

Clinical outcomes improvement is more exciting that 'data quality analysis project'

How can we make this valuable for the clinician right now?

References

[1] Waring J, Lindvall C, Umeton R. Automated machine learning: review of the state-of-the-art and opportunities for healthcare. Artif Intell Med. 2020 Apr;104:101822, doi: 10.1016/j.artmed.2020.101822.
[2] Yoon A, Lee YY. Factors of trust in data reuse. Online Inf Rev. 2019 Sep;43(7):1245-62, doi: 10.1108/OIR-01-2019-0014.
[3] Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res. 2018 May;20(5):e185, doi: 10.2196/jmir.9134.
[4] Zwaanswijk M, Verheij RA, Wiesman FJ, Friele RD. Benefits and problems of electronic information exchange as perceived by health care professionals: an interview study. BMC Health Serv Res. 2011 Dec;11:256(2011), doi: 10.1186/1472-6963-11-256.
[5] Kahn MG, Callahan TJ, Barnard J, Bauck AE, Brown J, Davidson BN, Estiri H, Goerg C, Holve E, Johnson SG, Liaw ST. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. eGEMs. 2016 Sep;4(1):1244, doi: 10.13063/2327-9214.1244.
[6] Liaw ST, Guo JG, Ansari S, Jonnagaddala J, Godinho MA, Borelli Jr AJ, de Lusignan S, Capurro D, Liyanage H, Bhattal N, Bennett V. Quality assessment of real-world data repositories across the data life cycle: a literature review. J Am Med Inform Assoc. 2021 Jul;28(7):1591-9, doi: 10.1093/jamia/ocaa340.
[7] Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. eGEMs. 2017 Sep;5(1):14, doi: 10.5334%2Fegems.218.
[8] Henley-Smith S, Boyle D, Gray K. Improving a secondary use health data warehouse: proposing a multi-level data quality framework. eGEMs. 2019 Aug;7(1):38, doi: 10.5334/egems.298.

References

[9] Bates DW, Auerbach A, Schulam P, Wright A, Saria S. Reporting and implementing interventions involving machine learning and artificial intelligence. Ann Intern Med. 2020 Jun;172(11_Supplement):S137-44, doi: 10.7326/M19-0872.
[10] Kuchinke W, Ohmann C, Verheij RA, van Veen EB, Arvanitis TN, Taweel A, Delaney BC. A standardised graphic method for describing data privacy frameworks in primary care research using a flexible zone model. International J Med Inform. 2014 Dec;83(12):941-57, doi: 10.1016/j.ijmedinf.2014.08.009.
[11] Kahn MG, Raebel MA, Glanz JM, Riedlinger K, Steiner JF. A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. Med Care. 2012 Jul;50(0), doi: 10.1097/MLR.0b013e318257dd67.
[12] Van der Bij S, Khan N, Ten Veen P, de Bakker DH, Verheij RA. Improving the quality of EHR recording in primary care: a data quality feedback tool. J Am Med Inform Assoc. 2017 Jan;24(1):81-7, doi: 10.1093/jamia/ocw054.
[13] Owens Jr EF, Esposito J, Hosek RS, Sullivan SG. Assessing the feasibility of using an electronic records database system in use in a group of private chiropractic clinics for practice-based research. J Can Chiropr Assoc. 2021 Apr;65(1):94-104.
[14] Thuraisingam S, Chondros P, Dowsey MM, Spelman T, Garies S, Choong PF, Gunn J, Manski-Nankervis JA. Assessing the suitability of general practice electronic health records for clinical prediction model development: a data quality assessment. BMC Medical Inform Decision Making. 2021 Dec;21(1):1, doi: 10.1186/s12911-021-01669-6.
[15] Taggart J, Liaw ST, Yu H. Structured data quality reports to improve EHR data quality. Int J Med Inform. 2015 Dec;84(12):1094-8, doi: 10.1016/j.ijmedinf.2015.09.008.
[16] Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc. 2013 Jan;20(1):144-51, doi: 10.1136/amiajnl-2011-000681.

MEDINFO23
8 – 12 JULY 2023 | SYDNEY, AUSTRALIA

Thank you

Dr Melinda Wassell
e: drmelinda.wassell@gmail.com
l: linkedin.com/in/melinda-wassell-17b9ab32/