

## Problem



EMR Data is locked up in primary care

## Methods

- Co-design with 12 family physicians in Ontario
- CPCSSN 2015Q2 Dataset (1.2 M de-identified patient records, 1100 primary care providers, pan-Canadian)
- Natural Language Processing
- Coding to International Standards (SNOMED, ATC, LOINC)
- All cleansing is independently validated by 2 experts. Conflicts resolved by a third expert.
- Transformation to analyzable formats
- Provision in easy-to-use tools

## Results & Architecture

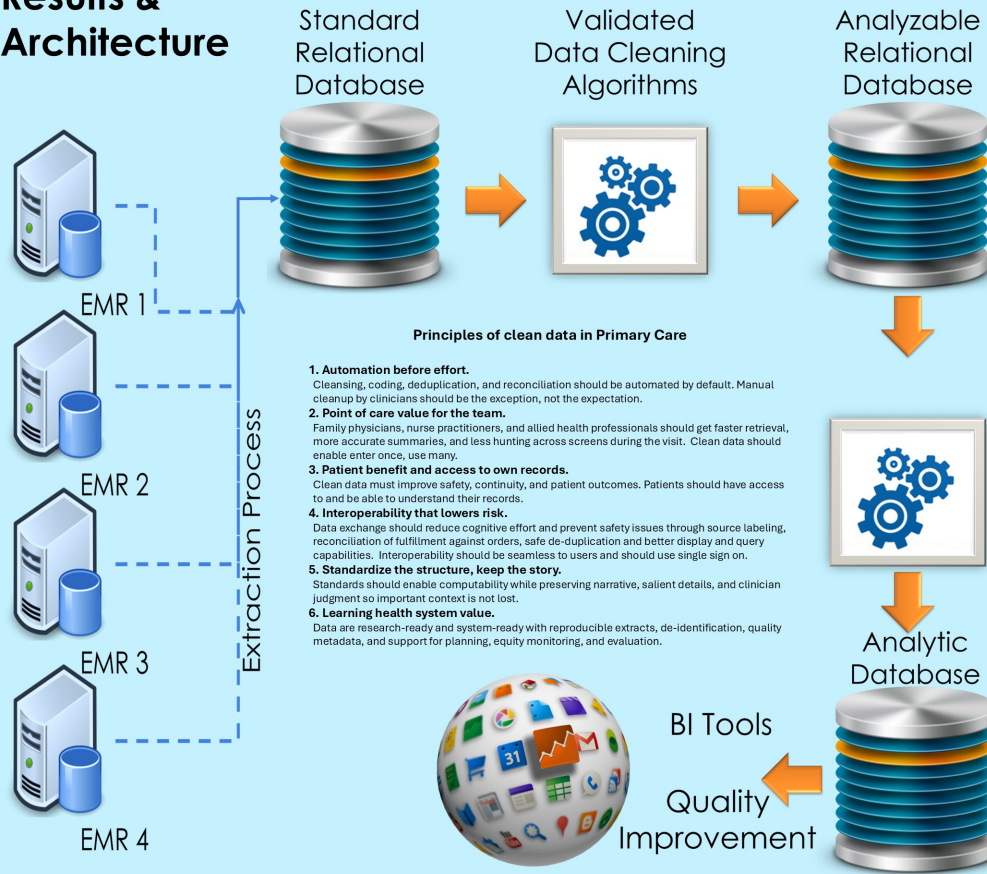


Table 1. Accuracy of data cleansing

Category	Sample Size	Irrelevant Records	Corrected Sample Size	# Matched	% Matched	# Coded	% Coded
Problem List	1,000,000	172,680	827,391	794,428	96%	783,801	95%
Past Medical History	1,000,000	192,621	807,379	728,593	90%		
Medications	500,000	6318	493,682	484,474	98%	484,474	100%
Allergies	100,000	16,091	83,909	74,374	89%	74,374	89%
Lab Results	5,000,500	338,162	4,662,338	4,521,517	97%	4,518,972	97%
Procedures	2,814,281	1,400,770	1,413,511	1,413,511	100%	1,413,511	100%
Vaccines	500,000	1236	498,764	498,240	99.9%	498,240	100%

## Benefits

- Physicians and allied healthcare providers benefit from higher quality notes, faster access to data and decreased cognitive load
- Patients can get access to their data with no additional effort on the part of physicians
- Researchers get access to large standardized data sets for their research projects
- Health systems benefit from less physician burnout and greater physician capacity

## Acknowledgements



Funding obtained through Canada Health Infoway Vendor Innovation Program

Platinum Medical Clinic for advice and guidance