

# Quantifying Time-Varying Medication Regimen Complexity Using a Machine-Learning-Based Pipeline in Patients with Type 2 Diabetes

Jun Gong<sup>1\*</sup>; Antoinette B. Coe<sup>1</sup>; Michael Dorsch<sup>1</sup>; Sarah Vordenberg<sup>1</sup>; VG Vinod Vydiswaran<sup>2,3</sup>; Corey Lester<sup>1</sup>

1 Department of Clinical Pharmacy, College of Pharmacy, University of Michigan  
 2 Department of Learning Health Sciences, Medical School, University of Michigan  
 3 School of Information, University of Michigan

PRESENTER:  
**Jun Gong**



## Background & Objective

- Patients with type 2 diabetes (T2DM) often manage **complex medication regimens** due to multimorbidity and polypharmacy.
- The Medication Regimen Complexity Index (MRCI) quantifies complexity based on dosage form, dosing frequency, and administration instructions.
- However, MRCI is **typically calculated manually at a single time point**.
- **Objective:** To develop a **scalable machine-learning-based pipeline** to calculate **time-varying MRCI** in patients with T2DM.

## Methods

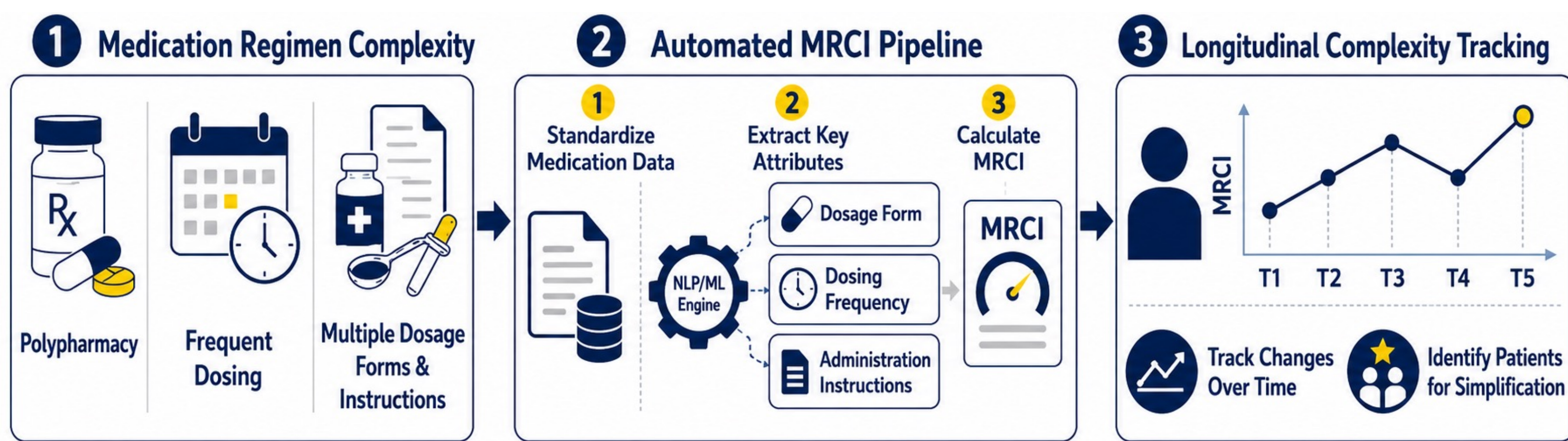


Figure 1. Conceptual Overview of Automated Time-Varying MRCI Measurement

Step	Description
<b>Design</b>	Retrospective cohort study; UM IRBMED approved (HUM00267637)
<b>Data Sources</b>	MGI EHR + Surescripts pharmacy fulfillment records, 2009-2019
<b>Population</b>	Adults ≥18 with T2DM and ≥1 post-index prescription
<b>Index Date</b>	Jan 1, 2009 for prevalent T2DM; first diagnosis for incident T2DM
<b>Medication Integration</b>	EHR orders + claims harmonized using RxNorm attributes, temporal alignment, and pharmacist-reviewed rules
<b>Timeline Construction</b>	Matched pharmacy claims retained; unmatched records preserved for longitudinal medication timelines

“**Medication regimen complexity in type 2 diabetes has always been a moving target.**”

Our pipeline tracks it automatically, over time, and at scale.

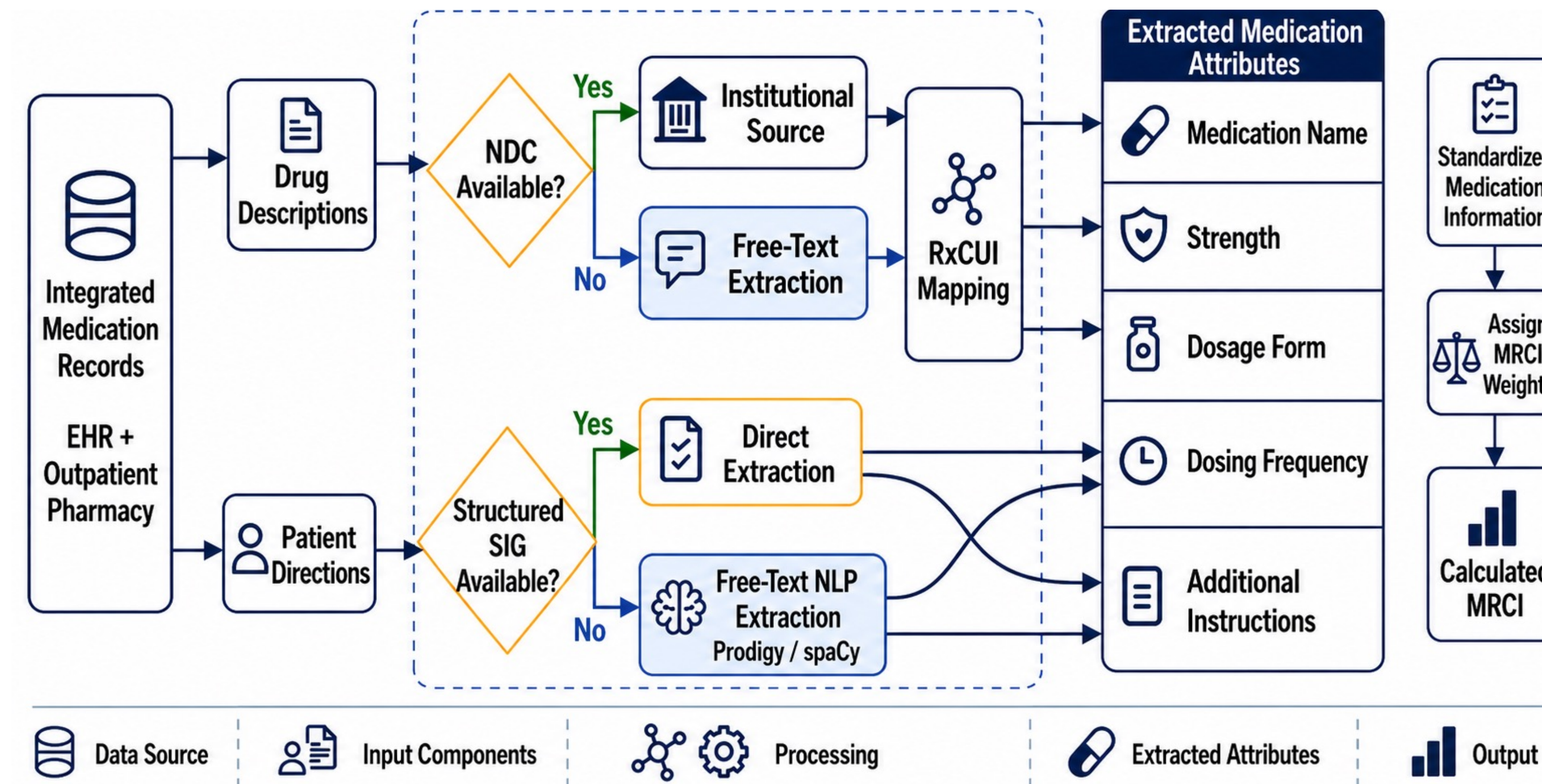


Figure 2. Automated Medication Information Extraction and MRCI Scoring Workflow

MRCI FORMULA	EXAMPLE MEDICATIONS	MRCI SCORING SUMMARY																																								
$DF = \sum_{i=1}^n \text{Dosage Form};$ $DF_q = \sum_{i=1}^n \text{Dosing Frequency};$ $AI = \sum_{i=1}^n \text{Additional Instructions};$ $MRCI = DF + DF_q + AI$	<table border="1"> <thead> <tr> <th>Medication Name</th> <th>Dosage Form</th> <th>Frequency</th> <th>Additional Instruction</th> </tr> </thead> <tbody> <tr> <td>Metformin</td> <td>Tablet</td> <td>Twice daily</td> <td>Take with meals</td> </tr> <tr> <td>Insulin glargine</td> <td>Injection</td> <td>Once daily</td> <td>Take at the same time every day</td> </tr> <tr> <td>Atorvastatin</td> <td>Tablet</td> <td>Once daily</td> <td>Take in the evening</td> </tr> </tbody> </table>	Medication Name	Dosage Form	Frequency	Additional Instruction	Metformin	Tablet	Twice daily	Take with meals	Insulin glargine	Injection	Once daily	Take at the same time every day	Atorvastatin	Tablet	Once daily	Take in the evening	<table border="1"> <thead> <tr> <th>Medication Name</th> <th>Dosage form</th> <th>Frequency</th> <th>Additional Instruction</th> </tr> </thead> <tbody> <tr> <td>Metformin</td> <td>1</td> <td>2</td> <td>1</td> </tr> <tr> <td>Insulin glargine</td> <td>3</td> <td>1</td> <td>1</td> </tr> <tr> <td>Atorvastatin</td> <td>1</td> <td>1</td> <td>1</td> </tr> <tr> <td><b>Total sub-scores of MRCI</b></td> <td><b>5</b></td> <td><b>4</b></td> <td><b>3</b></td> </tr> <tr> <td><b>Total MRCI</b></td> <td colspan="3"><b>5 + 4 + 3 = 12</b></td> </tr> </tbody> </table>	Medication Name	Dosage form	Frequency	Additional Instruction	Metformin	1	2	1	Insulin glargine	3	1	1	Atorvastatin	1	1	1	<b>Total sub-scores of MRCI</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>Total MRCI</b>	<b>5 + 4 + 3 = 12</b>		
Medication Name	Dosage Form	Frequency	Additional Instruction																																							
Metformin	Tablet	Twice daily	Take with meals																																							
Insulin glargine	Injection	Once daily	Take at the same time every day																																							
Atorvastatin	Tablet	Once daily	Take in the evening																																							
Medication Name	Dosage form	Frequency	Additional Instruction																																							
Metformin	1	2	1																																							
Insulin glargine	3	1	1																																							
Atorvastatin	1	1	1																																							
<b>Total sub-scores of MRCI</b>	<b>5</b>	<b>4</b>	<b>3</b>																																							
<b>Total MRCI</b>	<b>5 + 4 + 3 = 12</b>																																									

Figure 3. Example MRCI calculation

### Time-Varying MRCI Construction

- 1 Define Active Intervals**  
Use prescription start/stop dates
- 2 Link Fills & Aggregate Daily Scores**  
Adjust overlap and short gaps and carry scores forward
- 3 Generate Patient-Month MRCI**  
Average daily scores within patient-month bins

### Evaluation & Longitudinal Modeling

- **Named Entity Recognition (NER) Model Evaluation**
  - Fine-tuned Prodigy™/spaCy NER model
  - Training/refinement: 1,000 + 3 x 2000 annotated records
  - Test: independent 1,000-record gold standard
  - Primary metric: F1 score
- **Longitudinal Trend Analysis**
  - Outcome: patient-month total MRCI
  - Primary model: quadratic linear mixed-effects model with patient-specific trajectories
  - Output: fitted population trajectories with 95% CIs

## Results

Characteristic	Overall cohort
Patients, N	7,671
Medication records, N	1,209,933
Age at index	56 (46, 64)
Prior T2DM duration	0.00 (0.00, 3.86)
Follow-up time	8.86 (5.10, 11.00)
Female	3,646 (48%)
Caucasian	6,242 (82%)

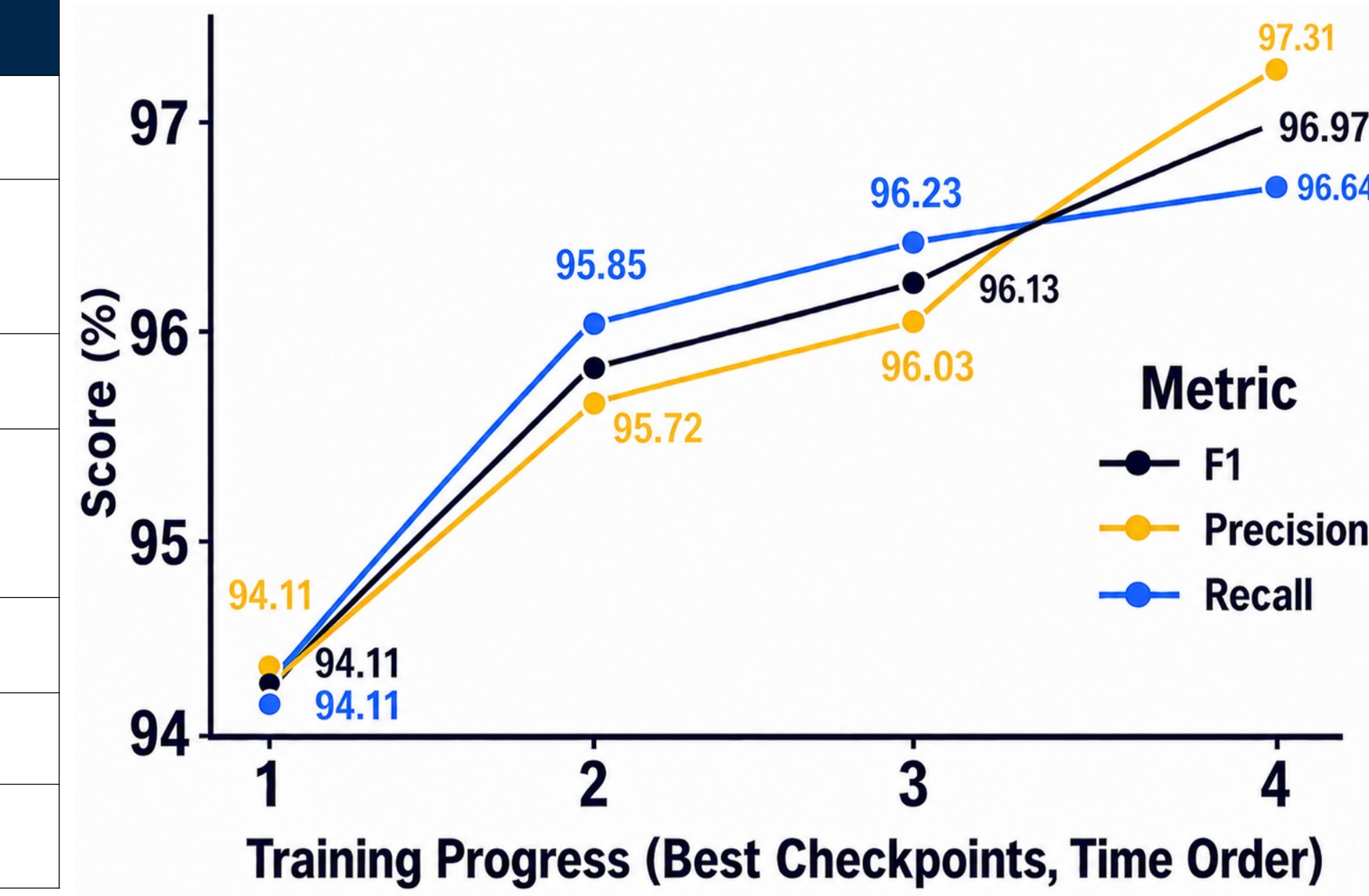


Table 1. Baseline Characteristics  
 \*Values are median (Q1, Q3) (years) or N (%).

Figure 4. NER Model Performance Across Training Checkpoints

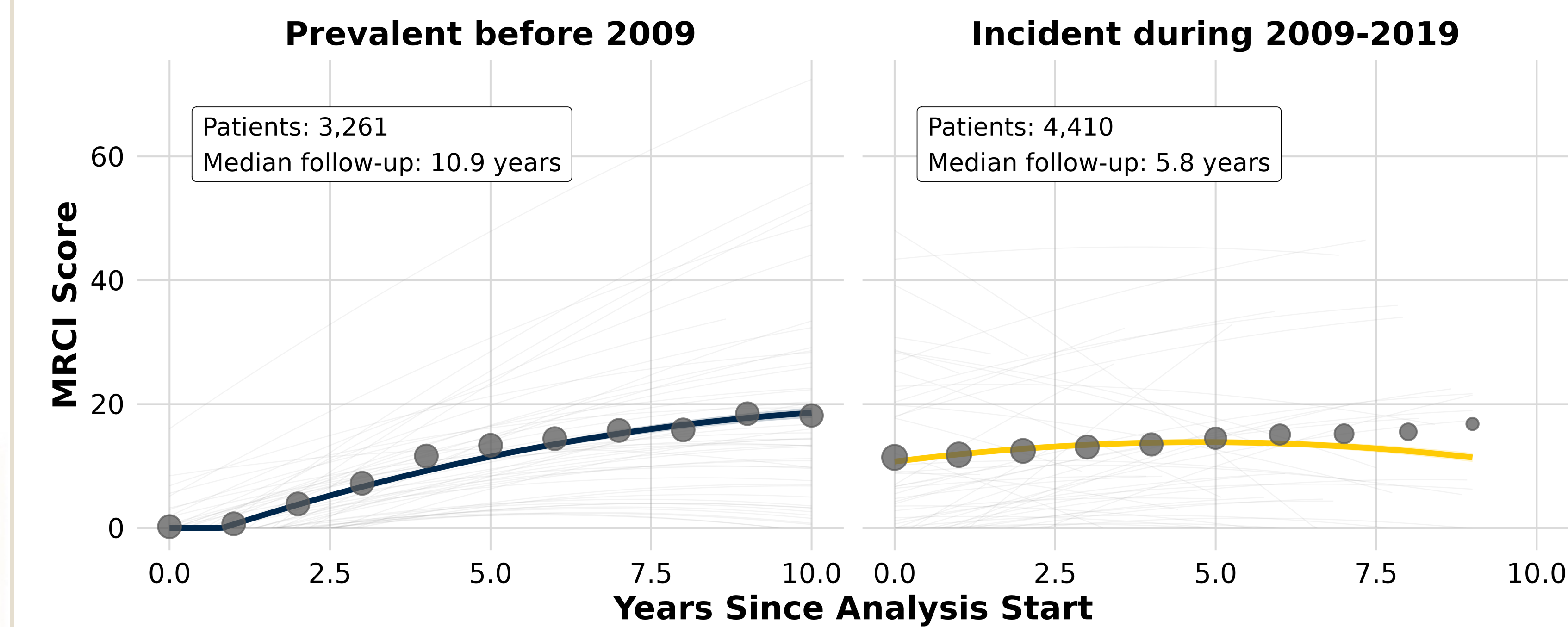


Figure 5. Medication Regimen Complexity Trajectories by Cohort  
 \*Prevalent early values may be underestimated because pharmacy capture begins in 2009.

Model-based estimate	Prevalent before 2009	Incident during 2009-2019
<b>Predicted MRCI at year 0</b>	0.0	10.7
<b>Predicted MRCI at last supported year</b>	18.6 at year 10	11.4 at year 9
<b>Total change, year 0 to last supported year</b>	+18.6	+0.7
<b>Average yearly change, year 0 to last supported year</b>	+1.9/year	+0.1/year

## Conclusion & Future Work

- Time-varying MRCI increased over follow-up in patients with T2DM.
- The ML pipeline enabled scalable longitudinal MRCI calculation from real-world medication data.
- Score discrepancies suggest the need for MRCI refinement and human oversight for complex directions.
- **Future Work:** 1) Validate the pipeline against PharmD-calculated reference scores; 2) Compare time-varying vs static MRCI for outcome prediction.