



ICD-9 to ICD-10 Mapping for Research in Biologics and Biosimilars Using Administrative Healthcare Data

Mengdong He¹, Adrian J. Santiago Ortiz¹, James Marshall^{2, 3}, Aaron B. Mendelsohn^{2, 3}, Jeffrey R. Curtis⁴, Charles E. Barr⁵, Catherine M. Lockhart⁵, and Seoyoung C. Kim^{1, 2}

¹ Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Boston, MA; ² Harvard Medical School, Boston, MA; ³ Harvard Pilgrim Health Care Institute, Boston, MA; ⁴ University of Alabama at Birmingham, Birmingham, AL; ⁵ Biologics and Biosimilars Collective Intelligence Consortium, Alexandria, VA.



BBC  **C**

Biologics & Biosimilars
Collective Intelligence Consortium

PRIME

PROGRAM IN RHEUMATOLOGIC, IMMUNOLOGIC,
AND MUSCULOSKELETAL PHARMACOEPIDEMOLOGY



Disclosure

- ❑ This project was supported by the Biologics & Biosimilars Collective Intelligence Consortium (BBCIC).
- ❑ Dr. Kim received research grants to the Brigham and Women's Hospital from Pfizer, Bristol-Myers Squibb and Roche for unrelated studies.



Background



ICD-9 to ICD-10 Transition

- ❑ As of 10/1/2015, the Centers for Medicare and Medicaid Services (CMS) mandated the transition from ICD-9 to ICD-10 codes.
- ❑ The ICD-10 codes are very different from ICD-9 code sets.

Differences Between ICD-9-CM and ICD-10 Code Sets		
	ICD-9-CM	ICD-10 code sets
Procedure	3,824 codes	71,924 codes
Diagnosis	14,025 codes	69,823 codes

ICD-10 Code Structure Changes (selected details)		
	Old	New
Diagnosis Structure	ICD-9-CM	ICD-10-CM
	<ul style="list-style-type: none">• 3-5 characters• First character is numeric or alpha• Characters 2-5 are numeric	<ul style="list-style-type: none">• 3-7 characters• Character 1 is alpha• Character 2 is numeric• Characters 3 – 7 can be alpha or numeric
Procedure Structure	ICD-9-CM	ICD-10-PCS
	<ul style="list-style-type: none">• 3-4 characters• All characters are numeric• All codes have at least 3 characters	<ul style="list-style-type: none">• ICD-10-PCS has 7 characters• Each can be either alpha or numeric• Numbers 0-9; letters A-H, J-N, P-Z



BBCIC's ICD-10 Mapping Workgroup

- ❑ BBCIC uses a distributed research network (DRN) to generate post-marketing evidence for novel biologics and biosimilars.
- ❑ Active surveillance of pharmaceutical products' safety and effectiveness in DRNs requires a robust approach for converting ICD-9 to ICD-10 codes that are used to define study populations, covariates and outcomes.



Objective

- ❑ To convert from ICD-9 to ICD-10 diagnostic and procedure codes for various health conditions in clinical areas of BBCIC's interest and compare the prevalence of these health conditions before and after 10/1/2015



Methods



Clinical Areas of BBCIC's Interest

- 108 health conditions related to three disease areas:
 - Hematologic/oncologic conditions - related to granulocyte colony stimulating factors (GCFs)
 - Systemic inflammatory disease - related to anti-inflammatory drugs
 - Diabetes type 1 and 2 - related to insulin



Three Ways of Mapping

- Using the General Equivalence Mappings (GEMs) developed by CMS, we converted the ICD-9 to ICD-10 codes in three ways:
 - Forward Backward Mapping (FBM)
 - Secondary Mapping (SM)
 - Tertiary Mapping (TM)

CMS General Equivalence Mappings (GEMs)

Forward Mapping

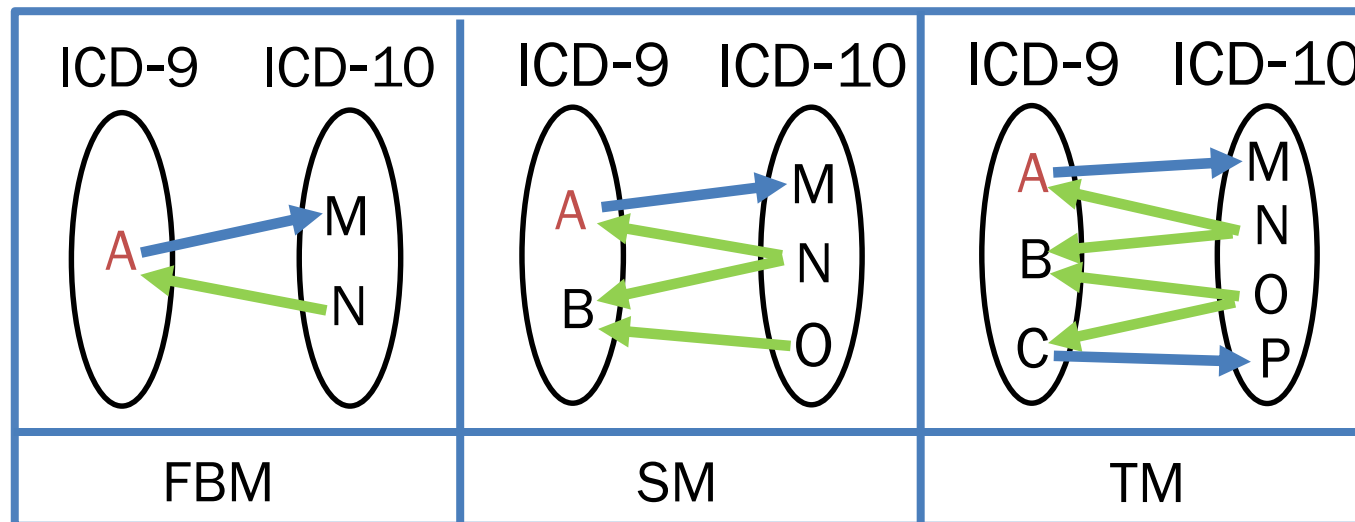


Backward Mapping



Three Ways of Mapping

- ❑ FBM: The simplest, using the direct links of forward and backward GEMs.
- ❑ SM: More complex, based on the ICD-10 codes identified by FBM.
- ❑ TM: The most complicated method based on an iteration of the SM.





Manual Review and Trend Analyses

- ❑ Physician expert (S.C. Kim) reviewed the relevance of ICD-10 codes from the three mapping methods.
- ❑ Prevalence of ICD-9 and ICD-10 codes from FBM were calculated in the pre- and post-ICD-10 implementation period (9/1/2012 – 3/31/2018)
- ❑ Harvard Pilgrim Health Care Institute team conducted the analyses in the DRN of 5 data partners and provided pooled results.
- ❑ We visually assessed prevalence trends of these health conditions and applied a threshold of 20% level change between the ICD-9 prevalence in 9/2015 versus ICD-10 prevalence in 10/2015.



Results

SM/TM vs FBM

- We observed a marked increase in the number of codes mapped by SM and TM for four conditions compared to FBM.
 - Most were not relevant or specific.
 - E.g. Type 1 diabetes: Additional ICD-10 codes were identified for other types of diabetes and atherosclerosis.

Condition	N of ICD-9 codes in the algorithm	N of ICD-10 codes from FBM	N of ICD-10 codes from SM*	N of ICD-10 codes from TM*
Type 1 Diabetes	20	84	327	281
Hypoglycemia	5	17	34	54
Inflammatory Bowel Disease	13	78	16	26
Tuberculosis	426	57	54	80
Myocardial Infarction	30	17	0	0
Rheumatoid Arthritis	11	451	0	0
Breast Cancer	11	54	0	0

* Unique addition of ICD-10 codes beyond codes identified by FBM

SM/TM vs FBM

- For conditions such as MI, RA and breast cancer, no additional ICD-10 codes were found by SM or TM.

Condition	N of ICD-9 codes in the algorithm	N of ICD-10 codes from FBM	N of ICD-10 codes from SM*	N of ICD-10 codes from TM*
Type 1 Diabetes	20	84	327	281
Hypoglycemia	5	17	34	54
Inflammatory Bowel Disease	13	78	16	26
Tuberculosis	426	57	54	80
Myocardial Infarction	30	17	0	0
Rheumatoid Arthritis	11	451	0	0
Breast Cancer	11	54	0	0

* Unique addition of ICD-10 codes beyond codes identified by FBM

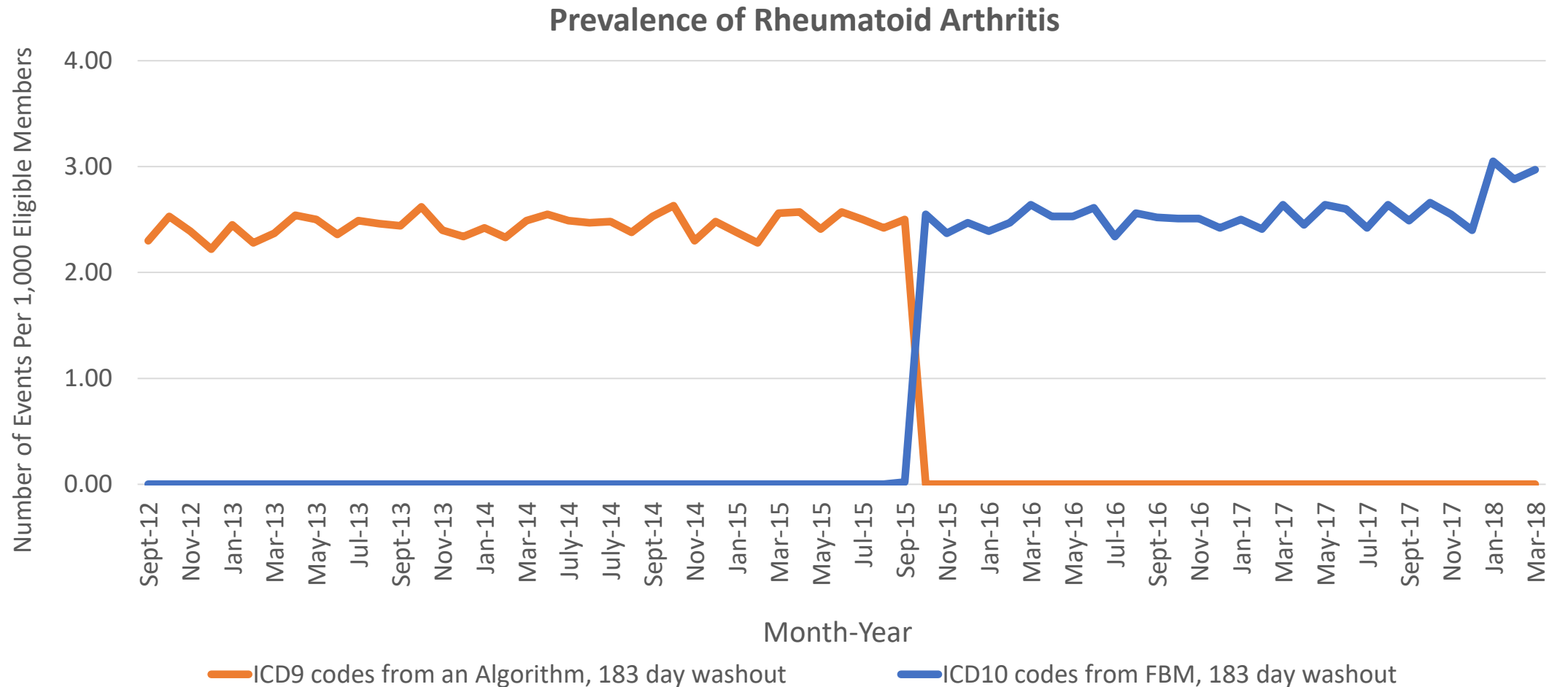
Manual Review Results

- 24% (N=26) conditions were considered mapped to problematic ICD-10 codes via FBM, for reasons including:
 - ICD-10 codes were too non-specific
 - E.g., ICD-9-CM 246.8 “Other specified disorders of thyroid” in the algorithm for thyroid disorders <-> ICD-10-CM E35 “Disorders of endocrine glands in diseases classified elsewhere”
 - ICD-10 codes were unrelated
 - E.g., ICD-9-CM 536.3 “Gastroparesis” <-> ICD-10-CM E0843 “Diabetes mellitus due to underlying condition with diabetic autonomic (poly)neuropathy”



Trend Analyses

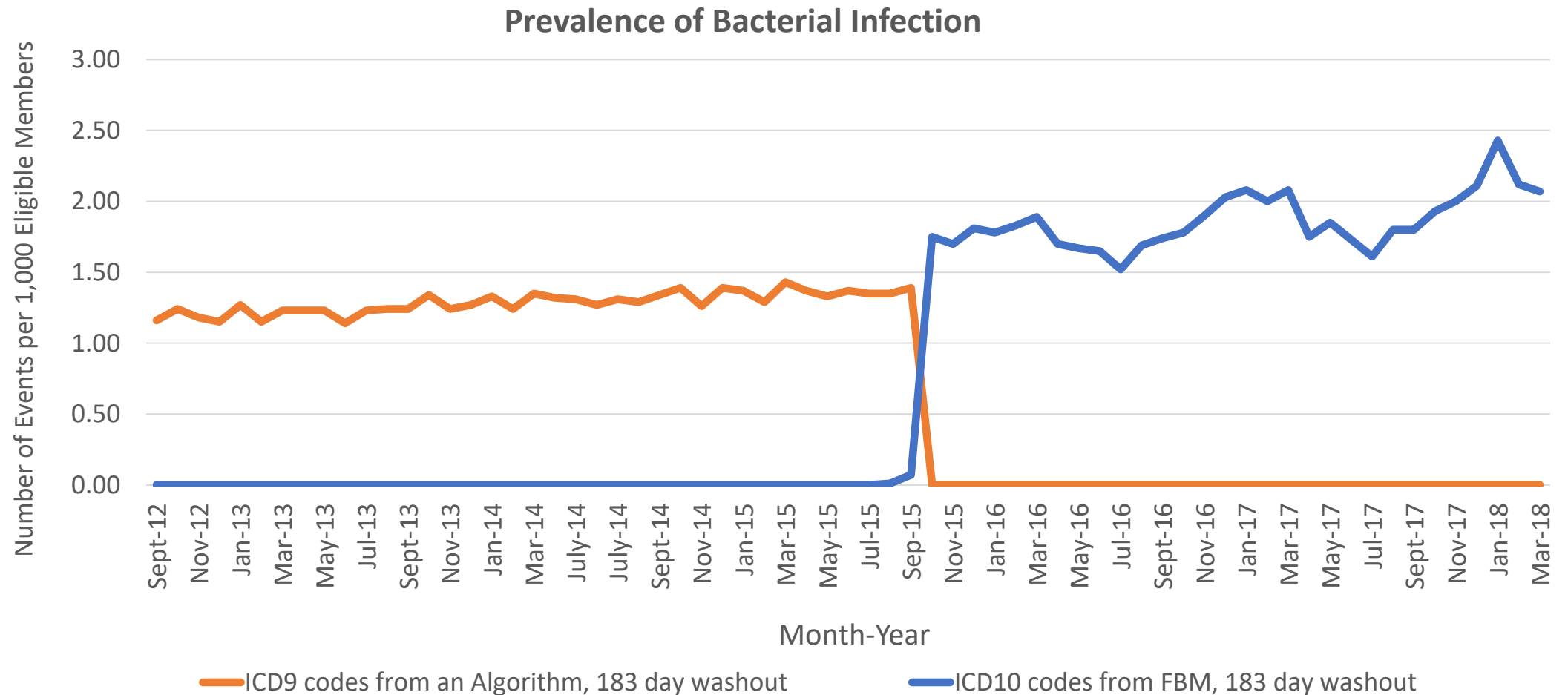
- 80% (N=86) conditions had visually comparable ICD-9 and ICD-10 trends after manual revisions of ICD-10 codes from FBM.





Trend Analyses

- 13% (N=14) conditions had a prevalence change greater than 20% in 10/2015, in addition to visual inconsistency.





Limitations

- ❑ We did not specify the diagnosis position (e.g., primary or any position) or the healthcare setting (e.g., inpatient or outpatient) for included ICD-9 and ICD-10 codes.
- ❑ Our estimated prevalence of health conditions were based on a group of ICD-9 or ICD-10 codes rather than the prevalence per each code.
- ❑ We did not validate the ICD-10 algorithms for the converted health conditions or assess their impact on study associations.
- ❑ Our conversion did not include non-billable ICD-9 or ICD-10 codes.



Conclusions

- ❑ FBM is generally the most efficient automated way to convert ICD-9 to ICD-10 codes.
- ❑ Manual review of the converted codes is recommended for all three methods.
- ❑ With manual revision, most ICD-10 algorithms from FBM achieved consistent prevalence trends compared to ICD-9 algorithms and had less than 20% level change in ICD-9 versus ICD-10 prevalence.
- ❑ Challenges present to empirically determine the quality of conversions due to a lack of guidance on comparing the performance of ICD-9 versus ICD-10 codes.



Thank you!

□ Email:

sykim@bwh.harvard.edu

mhe1@bwh.harvard.edu

Appendix

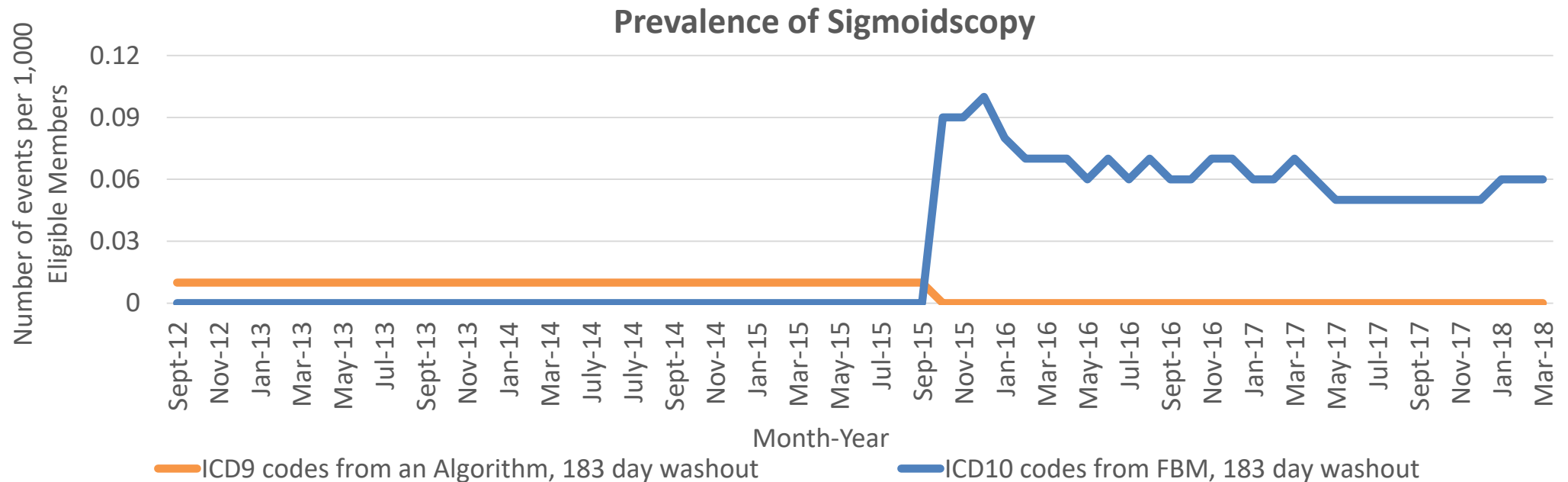
Mapping diagram: Type 1 Diabetes

	FBM	SM		TM	
	1 st ICD10 codes	2 nd ICD9 codes	2 nd ICD10 codes	3 rd ICD9 codes	3 rd ICD10 codes
N of codes input into GEM mappings	20 1 st ICD9 codes (Macro)	84 1 st ICD10 codes (Manually)	19 2 nd ICD9 codes (Macro)	327 2 nd ICD10 codes (Manually)	25 3 rd ICD9 codes (Macro)
N after applying GEM mappings	135 ICD10 codes: ✓ 1T1 and B_combo ✓ F_combo	232 ICD9 codes: ✓ 1T1 and F_combo ✓ B_combo	602 ICD10 codes: ✓ 1T1 and B_combo × F_combo	933 ICD9 codes: ✓ 1T1 and F_combo ✓ B_combo	784 ICD10 codes: ✓ 1T1 and B_combo ✓ F_combo
N after removing duplicate codes from GEM mappings	84 ICD10 codes: ✓ 1T1 and B_combo × F_combo	39 ICD9 codes: ✓ 1T1 and F_combo ✓ B_combo	389 ICD10 codes: ✓ 1T1 and B_combo × F_combo	44 ICD9 codes: ✓ 1T1 and F_combo ✓ B_combo	559 ICD10 codes: ✓ 1T1 and B_combo × F_combo
N after removing duplicate codes as used previously	N/A	19 ICD9 codes (No duplicates as in 1 st ICD9 codes): ✓ 1T1 ✓ B_combo	327 ICD10 codes (No duplicates as in 1 st ICD10 codes): ✓ 1T1 and B_combo × F_combo	25 ICD9 codes (No duplicates as in 1 st and 2 nd ICD9 codes): ✓ 1T1 and F_combo ✓ B_combo	281 ICD10 codes (No duplicates as in 1 st and 2 nd ICD10 codes): ✓ 1T1 and B_combo × F_combo

Acronym: 1st ICD9/10 code=Primary ICD9/10 codes; 2nd ICD9/10 codes=Secondary ICD9/10 codes; 3rd ICD9/10 codes=Tertiary ICD9/10 codes; 1T1=One to One Mapping; F_combo=Part of Combination Codes From Forward Mapping (ICD9 to ICD10); B_combo=Part of Combination Codes From Backward Mapping (ICD10 to ICD9).

Appendix

- 13% (N=14) conditions had a prevalence change greater than 20% in 10/2015, in addition to visual inconsistency.
 - Some were caused by inherent differences between the two coding systems.
 - E.g. Sigmoidoscopy: ICD-9-PCS 45.24 “Flexible sigmoidoscopy” <-> ICD-10-PCS ODJD8ZZ “Inspection of Lower Intestinal Tract, Via Natural or Artificial Opening Endoscopic”



Appendix

- Three conditions (meningitis, chronic liver disease, and ankylosing spondylitis) had their ICD-10 algorithms further refined.
- The new algorithm of meningitis achieved consistency.

