Supplemental Materials

Algorithms to identify outcomes

The following diagnostic criteria were used: 1) diabetes defined based on ICD-9-CM and ICD-10-CM codes presenting on two outpatient physician claims within 2 years, one outpatient physician claim plus a medication used to treat diabetes, or one inpatient claim (sensitivity= 92% and specificity= 97%)[1]; 2) hypertension defined based on two outpatient physician claims (any position) within three years or 1 outpatient physician claim plus a medication used to treat HTN (sensitivity= 73% and specificity= 95%)[2]; 3) myocardial infarction defined based on inpatient claim with a hospital stay ≥ 1 day (positive predictive value (PPV) > 90%)[3]; 4) stroke defined based on inpatient claims (primary position) with a hospital stay ≥ 1 day and a CT scan of the head the day before, day of, or day after the hospitalization date (PPV > 90%)[4]; 5) rheumatoid arthritis defined based on inpatient claim or three outpatient claims (diagnosed by a physician) and ≥1 by a musculoskeletal specialist within a 2 year period[5]; 6) psoriasis defined based on ≥1 physician-consultations with a dermatologist consistent with psoriasis (PPV 80-90%)[6, 7]; 7) inflammatory bowel disease defined based on ≥ 2 claims by gastroenterologist of Crohn's disease (CD), ulcerative colitis (UC) or mixed CD+UC codes within 1 year and had a colonoscopy, sigmoidoscopy, or bowel resection within 6 weeks prior to the first diagnosis (adapted from[8] PPV 81%-84%) ; and 8) irritable bowel syndrome defined based on any physician claim for irritable bowel syndrome in the absence of a diagnosis of inflammatory bowel disease, colorectal cancer, pancreatic cancer, chronic pancreatitis, cirrhosis, celiac disease, sprue or other malabsorption syndromes, ovarian cancer or prior treatment with mesalamine or medications commonly used to treat irritable bowel syndrome (PPV 83%-92%)[9]. List of ICD-9-CM and ICD-10-CM codes and additional lookback period requirements are listed in the **Supplemental Table S2.**

Supplemental Table S1. List of procedure codes for the identifications of fecal microbiota transplantation

CPT codes for the identifications of fecal microbiota transplantation

44705 = Preparation of fecal microbiota for instillation, including assessment of donor specimen
G0455 = Preparation with instillation of fecal microbiota by any method, including assessment of donor specimen
44799 = Fecal instillation by oro-nasogastric tube or enema

Supplemental Table S2. Summary of international classification of disease modification codes and requirement of lookback period for each of the study outcomes

Outcome	ICD-9 Codes	ICD-10 Codes	Algorithm	Minimum look back period without prior diagnosis or medications	Sensitivity, specificity, and PPV	Reference
Diabetes Mellitus	250.xx	E10.x–E14.x	2 outpatient physician claims within 2 years (nurse practitioner from category 1 or physicians from category 3 roll-up) or 1 hospitalization with the relevant diabetes ICD codes	TBD without diagnosis or treatment with metformin, sulfonylurea, meglitinides, DPP-4 inhibitor, GLP1 agonist, TZDs (PPAR gamma agonist) or insulin	Sensitivity = 92% Specificity = 97%	Chen et al. (2010)
Hypertension	401.x, 402.x, 403.x, 404.x, or 405.x	I10.x, I11.x, I12.x, I13.x, or I15.x	Two outpatient physician (nurse practitioner from category 1 or physicians from category 3 roll-up) billing claims for hypertension within a 3-year period	TBD without any diagnosis of prescription for HCTZ, ACEIs, ARBs, CCBs	Sensitivity = 73% Specificity = 95%	Tu et al. (2007)
Inflammatory bowel disease	555.x 556.x	K50.x K51.x except K51.4x	i) 2+ diagnostic code of 555 or K50 (CD), ii) 2+ diagnostic codes of 556 or K51 (UC), iii) mixed CD+UC codes. Diagnoses must be from a gastroenterologist (Category role-up 1= Gastroenterology). First diagnosis must be on or whtin 6 weeks of a lower endoscopy or bowel resection surgery.	1. Minimum 12 months without any prior IBD diagnosis 2. Has a colonoscopy, sigmoidoscopy, or bowel resection surgery with a code=1 for IBD diagnosis in the 6 weeks prior to or on the same day as the first IBD diagnosis by a gastroenterologist 3. Has minimum of two diagnoses of IBD from a gastroenterologist within 1 year 4. No prescriptions for mesalamine, olsalazine, or balsalazide prior to the first colonoscopy that is part of the diagnosis algorithm. 5. No prescriptions for azathioprine, 6MP, or vedolizumab prior to the first IBD diagnosis by a gastroenterologist 6. No prescription for methotrexate unless has a diagnosis of RA, Psoriasis, or psoriatic arthritis prior to IBD diagnosis. 7. No prescription for anti-TNF (adalimumab, infliximab, golimumab, certolizumab) unless there is a prior diagnosis of RA, Psoriasis, psoriatic arthritis, or ankylosing spondylitis prior to the diagnosis of IBD. Note this does not include etanercept	PPV of the previously published case-finding algorithm was 81% (95% confidence interval [CI], 78-87) and 84% (95% CI, 79-89) in two different organizations; Minimum of 2 diagnoses increased the PPV to 95%	Adapted from Herrinton LJ. (2007) Liu L et al. (2009)
Myocardial infarction	410 excluding 410.x2	I21.x	at least 1 inpatient claim and at least 1 night stay in hospital except if patient dies	1 year	PPV that exceeds 90%	Kiyota Y (2004)
Psoriasis	696.1	L40.0, L40.1, L40.2 L40.4, L40.8 and L40.9.	For PsO: ≥1 physician-consultations consistent with psoriasis (ICD-10 L codes). Although not part of original algorithm will require physician to	TBD without diagnosis or prescription for topical steroids. Also exclude if prior diagnosis with psoriatic arthritis ICD-9: 696.0 or ICD10 L40.5x, M07.0, M07.1, M07.2, M07.3 or M09.0	For ICD-9: Sensitivity= 80% PPV= 90% For CD-10: PPV= 81%	Löfvendahl et al. (2014) Asgari et al. (2013)

			be dermatologist (Category role-up 1= Dermatology) to be c/w Asgari algorithm.			
Rheumatoid arthritis	714.x	M05.x, M06.x	1 hospitalization arthritis code OR 3 physician RA diagnosis codes (claims) with ≥1 by a musculoskeletal specialist (Category role-up 1= Rheumatology or Orthopedics) in a 2 year period	Based on our pilot data, our algorithm is: Minimum 24 months without diagnosis or prescription for: (Sulfasalazine, hydroxychloroquine, Methotrexate) – unless there is a diagnosis of psoriasis or psoriatic arthritis prior to or on the date of the first dispensing. Anti-TNF drugs (etanercept, infliximab, golimumab, adalimumab, certolizumab – unless a diagnosis of IBD, PSO ankylosing spondylitis (720.0 or M45.x, Psoriatic arthritis (696.0 or L40.52) prior to or on the date of the first dispensing. Note that IBD does not apply to etanercept, just the other anti-TNF drugs Abatacept, tocilizumab, leflunomide tofacitinib , goldanakinra, rituximab (without a diagnosis of Lupus (L93.xx) any time prior or Lymphoma within the year prior (C81.xx to C88.xx, 200.xx – 202.xx except 202.5x-202.6x , 204.1x)	Sensitivity = 78% Specificity = 100%	Widdifield et al. (2014)
Stroke	430, 431, 433.x1, 434.x1, or 436 Plus a CT of the head either 1 day prior, the day of, or 1 day after the diagnosis	I60.x I61.x I63.x I69.0x I69.1x I69.3x Plus a CT of the head either 1 day prior, the day of, or 1 day after the diagnosis	in the primary diagnosis position of a hospitalization	1 year	PPV that exceeds 90% (note that the original algorithm did not include the requirement for the head CT)	Kumamaru H (2014)
Irritable bowel syndrome	564.1	K58.x	1 physician diagnosis with no IBS medications prior and no IBS exclusion diagnoses prior to the first IBD diagnosis		Sensitivity = 99% PPV = 90%	Sands et al. (2006)

Angiotensin converting enzyme inhibitors, ACEIs; angiotensin receptor blockers, ARBs; calcium channel blockers, CCBs; dipeptidyl peptidase-4, DPP-4; Crohn's disease, CD; glucagon-like peptide-1, GLP1; hydrochlorothiazide, HCTZ; inflammatory bowel disease, IBD; International Classification of Disease, ICD; multiple sclerosis, MS; positive predictive value, PPV; psoriasis, PsO; rheumatoid arthritis, RA; thiazolidinedione, TZDs; ulcerative colitis, UC

Supplemental Table S3. Crude incidence rates in the mrCDI vs. non-mrCDI groups

Outcome	mrCDI				Non-mrCDI			
	N	Events	Person-years	IR (95% CI)	N	Events	Person-years	IR (95% CI)
Inflammatory bowel disease	3,183	5	7,310	0.68 (0.22-1.60)	108,454	127	263,610	0.48 (0.40-0.57)
Rheumatoid arthritis	3,303	11	7,678	1.43 (0.72-2.56)	109,977	445	267,238	1.67 (1.51-1.83)
Psoriasis	2,000	3	3,989	0.75 (0.16-2.20)	69,931	154	148,772	1.04 (0.88-1.21)
Diabetes mellitus	2,402	29	5,542	5.23 (3.50-7.52)	78,377	1,129	191,747	5.89 (5.55-6.24)
Hypertension	755	38	1,721	22.08 (15.63-30.31)	30,939	1,260	72,404	17.40 (16.45-18.39)
Myocardial infarction	3,281	59	7,840	7.53 (5.73-9.71)	110,747	2,161	273,836	7.89 (7.56-8.23)
Stroke	2,988	44	7,037	6.25 (4.54-8.39)	101,710	1,596	250,824	6.36 (6.05-6.68)
Inflammatory bowel syndrome	2,589	21	5,443	3.86 (2.39-5.90)	86,890	783	192,933	4.06 (3.78-4.35)

CI - Confidence interval, FMT - fecal microbiota transplantation, mrCDI - multiply recurrent C difficile, IR - incidence rate

Supplemental Table S4. Crude incidence rates in in the group with CDI treated with FMT vs. mrCDI without FMT

Outcome	CDI with FMT				mrCDI without FMT			
	N	Events	Person-years	IR (95% CI)	N	Events	Person-years	IR (95% CI)
Inflammatory bowel disease	881	2	1,534	1.30 (0.16-4.71)	3,234	5	7,277	0.69 (0.22-1.60)
Rheumatoid arthritis	986	5	1,770	2.82 (0.92-6.59)	3,355	11	7,654	1.44 (0.72-2.57)
Psoriasis	629	2	1,002	2.00 (0.24-7.21)	2,028	3	3,984	0.75 (0.16-2.20)
Diabetes mellitus	771	3	1,335	2.25 (0.46-6.57)	2,431	30	5,500	5.45 (3.68-7.79)
Hypertension	222	9	387	23.27 (10.64-44.17)	760	37	1,687	21.93 (15.44-30.23)
Myocardial infarction	1,031	24	1,811	13.25 (8.49-19.71)	3,328	59	7,796	7.57 (5.76-9.76)
Stroke	934	15	1,664	9.01 (5.04-14.87)	3,028	44	6,996	6.29 (4.57-8.44)
Inflammatory bowel syndrome	601	4	1,009	3.96 (1.08-10.15)	2,605	21	5,592	3.76 (2.32-5.74)

CI - Confidence interval, FMT - fecal microbiota transplantation, mrCDI - multiply recurrent C difficile, IR - incidence rate

Supplemental Table S5. Results from sensitivity analysis limiting the data to 1/1/2010 through 6/30/2019.

A) Unadjusted and adjusted estimates of each outcome in the mrCDI groups vs. non-mrCDI

Outcome	Unadjusted HR (95% CI)	Adjusted HR (95% CI)	Model adjusted for the listed variables
Inflammatory bowel disease	1.06 (0.26-4.34)	1.23 (0.30-5.02)	age
Rheumatoid arthritis	0.54 (0.20-1.45)	∏	no covariates changed HR by 10% or more
Psoriasis	1.19 (0.38-3.74)	1.32 (0.42-4.19)	US Census division
Diabetes mellitus	0.77 (0.40-1.49)	∏	no covariates changed HR by 10% or more
Hypertension	1.51 (0.99-2.31)	1.23 (0.80-1.89)	age
Myocardial infarction	0.77 (0.55-1.09)	∏	no covariates changed HR by 10% or more
Stroke	0.94 (0.65-1.37)	∏	no covariates changed HR by 10% or more
Irritable bowel syndrome	1.00 (0.60-1.67)	∏	no covariates changed HR by 10% or more

Confidence interval, CI, hazard ratio, HR
∏ no further adjusted was made as none of the covariates changed the hazard ratio by ≥ 10%

B) Unadjusted and adjusted estimates of each outcome in CDI treated with FMT vs. mrCDI without FMT

Outcome	Unadjusted HR (95% CI)	Adjusted HR (95% CI)	Model adjusted for the listed variables
Inflammatory bowel disease	1.49 (0.14-16.41)	∥	
Rheumatoid arthritis	3.08 (0.76-12.56)	∥	
Psoriasis	1.72 (0.28-10.32)	∥	
Diabetes mellitus	0.56 (0.12-2.58)	1.53 (0.24-9.69)	inpatient place of dx, calendar year, # ambulatory care visits, # inpatient admissions
Hypertension	0.88 (0.35-2.20)	1.42 (0.52-3.86)	antibiotic use 90 days before, antibiotic use 90 days after
Myocardial infarction	1.89 (1.10-3.26)	1.70 (0.95-3.05)	inpatient place of dx, number of inpatient admissions
Stroke	1.41 (0.73-2.69)	1.40 (0.68-2.91)	inpatient place of dx, antibiotic use 90 days before, number of inpatient admissions
Irritable bowel syndrome	0.98 (0.32-2.98)	1.11 (0.33-3.76)	antibiotic use 90 days before , antibiotic use 90 days after, Charlson score, number of ambulatory care visits, number of inpatient admissions

Confidence interval, CI, hazard ratio, HR
∥ Not calculated due to the small number of events

References

1. Chen G, Khan N, Walker R, et al. Validating ICD coding algorithms for diabetes mellitus from administrative data. Diabetes Res Clin Pract 2010;89:189-95.
2. Quan H, Khan N, Hemmelgarn BR, et al. Validation of a case definition to define hypertension using administrative data. Hypertension 2009;54:1423-8.
3. Kiyota Y, Schneeweiss S, Glynn RJ, et al. Accuracy of Medicare claims-based diagnosis of acute myocardial infarction: estimating positive predictive value on the basis of review of hospital records. American Heart Journal 2004;148:99-104.
4. Kumamaru H, Judd SE, Curtis JR, et al. Validity of claims-based stroke algorithms in contemporary Medicare data: reasons for geographic and racial differences in stroke (REGARDS) study linked with medicare claims. Circ Cardiovasc Qual Outcomes 2014;7:611-9.
5. Widdifield J, Labrecque J, Lix L, et al. Systematic review and critical appraisal of validation studies to identify rheumatic diseases in health administrative databases. Arthritis Care Res (Hoboken) 2013;65:1490-503.
6. Asgari MM, Wu JJ, Gelfand JM, et al. Validity of diagnostic codes and prevalence of psoriasis and psoriatic arthritis in a managed care population, 1996-2009. Pharmacoepidemiol Drug Saf 2013;22:842-9.
7. Lofvendahl S, Theander E, Svensson A, et al. Validity of diagnostic codes and prevalence of physician-diagnosed psoriasis and psoriatic arthritis in southern Sweden--a population-based register study. PLoS One 2014;9:e98024.
8. Herrinton LJ, Liu L, Lafata JE, et al. Estimation of the period prevalence of inflammatory bowel disease among nine health plans using computerized diagnoses and outpatient pharmacy dispensings. Inflamm Bowel Dis 2007;13:451-61.
9. Sands BE, Duh M-S, Cali C, et al. Algorithms to identify colonic ischemia, complications of constipation and irritable bowel syndrome in medical claims data: development and validation. 2006;15:47-56.