The Journal of Rheumatology 2023;xx:xxxx
doi:10.3899/jrheum.220990
First Release March 1 2023

The Journal of Rheumatology

A Phenome-Wide Association Study of Drugs and Comorbidities Associated With Gastrointestinal Dysfunction in Systemic Sclerosis

Rory H. Maclean[1], Fiza Ahmed[1], Voon H. Ong[1], Charles D. Murray[2], and Christopher P. Denton[1]

ABSTRACT. ***Objective.*** To explore the causes of and contributors to gastrointestinal (GI) dysfunction in systemic sclerosis (SSc) in a phenome-wide association study (PheWAS), using real-world clinical records data.

Methods. Twelve thousand five hundred thirty-five documented clinical assessments of 2058 consenting individuals with SSc at the Royal Free Hospital (UK) were available for detailed phenotyping. Diagnoses and drugs were mapped to structured dictionaries of terms (Disease Ontology project and DrugBank Open Data, respectively). A PheWAS model was used to explore links between 6 important SSc-GI domains (constipation, diarrhea, dysmotility, incontinence, gastroesophageal reflux, and small intestinal bacterial overgrowth [SIBO]) and exposure to various comorbidities and drugs. "Hits" from the PheWAS model were confirmed and explored in a subcohort reporting quantitative GI symptom scores from the University of California Los Angeles Scleroderma Clinical Trials Consortium Gastrointestinal Tract Instrument 2.0 (GIT 2.0) questionnaire.

Results. One thousand five hundred forty-six individuals were entered into the PheWAS analysis. Six hundred seventy-three distinct diagnoses and 634 distinct drugs were identified in the dataset, as well as SSc-specific phenotypes such as antinuclear antibodies (ANA). PheWAS analysis revealed associations between drugs, diagnoses, and ANAs with 6 important SSc-GI outcomes: constipation, diarrhea, dysmotility, incontinence, reflux, and SIBO. Subsequently, using GIT 2.0 symptom scores links with SSc-GI were confirmed for 22 drugs, 4 diagnoses, and 3 ANAs.

Conclusion. Using a hypothesis-free PheWAS approach, we replicated known, and revealed potential novel, risk factors for SSc-GI dysfunction, including drug classes such as opioid, antimuscarinic, and endothelin receptor antagonist, and ANA subgroup.

Key Indexing Terms: drug toxicity, gastrointestinal diseases, scleroderma systemic sclerosis

Systemic sclerosis–associated gastrointestinal dysfunction (SSc-GI) is a significant burden to many patients. It is among the most frequent manifestations of this multisystem disease.[1] It is believed that the central SSc disease mechanisms of vasculopathy, fibrosis, and inflammation are the drivers of GI dysfunction. However, at the patient level, there is significant heterogeneity in both the sites of the gut involved and the severity of GI symptoms.[2] Previous studies have identified that differences in SSc-related antinuclear antibodies (ANA) between individuals may explain the heterogeneity in the manifestation of SSc.[3] For GI involvement, anticentromere antibodies (ACA) and anti-RNA polymerase autoantibodies (ARA) have been linked to increased disease severity based on patient-reported outcome measures (PROMs).[4]

The treatment of SSc-GI is focused on symptom control.[1,5] Proton pump inhibitors (PPIs) are used for gastroesophageal reflux, laxatives for constipation, and prokinetics for dysmotility.[6,7] Small intestinal bacterial overgrowth (SIBO) can be managed with antibiotic therapy. However, as a multisystem disease, individuals are appropriately prescribed drugs for other organ indications, and off-target effects of such drugs may contribute to GI dysfunction. Further, links between the organ-specific manifestations of SSc and other apparently unrelated disease processes in the individual may reveal shared disease mechanisms. As such, the study of a breadth of drugs and comorbidities in a population with a single disease may further understanding of the disease process.

The phenome-wide association study (PheWAS) was conceived to look for associations in the opposite direction to a genome-wide association study, that is, from a target genetic variant to multiple phenotypic traits.[8] However, the approach has

This work was supported by The Royal Free Charity from a generous donation supporting the Dita Athanassoulia clinical fellowship in gastrointestinal scleroderma.

[1]R.H. Maclean, BMBCh, F. Ahmed, MBBS, V.H. Ong, PhD, C.P. Denton, PhD, Centre for Rheumatology, Royal Free Campus, Division of Medicine, University College London; [2]C.D. Murray, PhD, Department of Gastroenterology, Royal Free London, London, UK.

CPD reports personal fees or research grants to his institution from GSK, Galapagos, BI, Roche, CSL Behring, Corbus, Horizon, and Arxx Therapeutics; all outside the submitted work. The remaining authors declare no conflicts of interest relevant to this article.

Address correspondence to Prof. C.P. Denton, Centre for Rheumatology, Division of Medicine, University College London, Royal Free Campus, Hampstead, London NW3 2PF, UK. Email: c.denton@ucl.ac.uk.

Accepted for publication January 12, 2023.

been broadened from genetic variant to target disease, symptom, and even laboratory result as opposed to a breadth of phenotypes. PheWAS methodology typically uses electronic health record data to generate phenotypes; for example, using coding schema such as the International Classification of Diseases, 10th revision. In this study, we had access to detailed clinical records from a single center and were able to create additional detailed SSc-relevant phenotypes. Additionally, we have included drug exposures in the scope of the phenotypes studied.

In this study we explored links between SSc-GI dysfunction and a large number of drugs and comorbidities using a hypothesis-free PheWAS approach. This was followed by confirmatory analysis using patient-reported GI symptom scores from the University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument 2.0 (GIT 2.0) questionnaire, a validated SSc-specific tool to measure the burden of GI symptoms.[9]

METHODS

Royal Free observational SSc cohort (Scleroderma Cohort). The first entry in the Scleroderma Cohort (SMART) was July 3, 2013, and the last was March 19, 2020. The archive of documented clinical assessments spanned 2002 to 2021. The SMART cohort[3] includes adults with SSc, morphea, or Raynaud phenomenon (RP), and healthy control subjects. In this study we included only individuals with a confirmed diagnosis of SSc, and > 90% of cases fulfill the 2013 European Alliance of Associations for Rheumatology (EULAR)/American College of Rheumatology (ACR) classification criteria.[10] Sixty-six (4%) were ANA negative and 41 (3%) had childhood onset disease.

The SMART study (UK SMART) database and tissue bank (version 7, February 17, 2020) was approved by the University College London Joint Research Office and the London-Fulham Research Ethics Committee (REC reference: 20/LO/0404, IRAS project ID: 279682). Participants gave written informed consent to join the study having read the patient information sheet. The SMART dataset includes demographic information and clinical information (diagnosis, disease onset, internal organ involvement, and test results).

Patient-reported GI symptoms subcohort. During the GIT 2.0 substudy period, from 2018 to 2020, consecutive patients with SSc fulfilling the 2013 EULAR/ACR criteria were recruited. Participants completed the GIT 2.0 GI symptoms questionnaire, a PROM focused on GI symptoms in SSc.[9] In this study, we used the GIT 2.0 symptom scores for reflux (8 items), distension/bloating (4 items), fecal soilage (1 item), diarrhea (2 items), and constipation (4 items). The total GI score also includes social functioning (6 items) and emotional well-being (9 items) domains.

Data extraction and processing. Documented clinical assessments were extracted from the electronic archive using patient identifiers from the SMART cohort. The clinical records follow a standard template for each episode that records a problem/diagnosis list, a current medication list, and then a narrative summary of current clinical issues and management together with a treatment plan. Semistructured text data were extracted from the documented clinical assessments, including dates, numbered lists of diagnoses, and numbered lists of drugs; negated terms were excluded.

A human disease ontology was accessed from the Human Disease Ontology website, under CC0 license,[11] and extracted with R package OntologyIndex. The dictionary was enriched for SSc-specific vocabulary, and regular expressions to match variations in syntax (eg, anti-ACA vs ACA).

The DrugBank drug vocabulary was downloaded from DrugBank under CC0 license[12] and synonyms were listed for each drug entity. This drug dictionary was enriched for SSc-relevant drugs including probiotics, trial drugs, nutritional supplements, total parenteral nutrition (TPN), and enteral feeds. Three-letter or shorter names were excluded. The generic terms "nutritional supplements" and "probiotics" were selected for this study, although it is likely that some relevant nonpharmacological interventions were not captured.

Mapping to drugs and diagnoses. The diagnostic and drug vocabularies were used to map the extracted text to standardized names, dealing with variations in naming and formatting. For each study subject, the exposure to drugs and labeling with diagnoses over time is summarized as "ever-labeled."

Quality control. A subset of 90 cases were manually labeled with significant diagnoses (gastric antral vascular ectasia, renal crisis, cardiac scleroderma, rheumatoid arthritis, interstitial lung disease [ILD], pulmonary hypertension, cancer, inflammatory arthritis, and myositis) by experienced clinicians with access to the full clinical notes. Manual and programmatic labels were compared using sensitivity, specificity, and accuracy metrics.

Statistical analysis.

- *PheWAS.* Six SSc-GI outcomes were explored, based on the burden of SSc-GI disease assessed with GI symptom scores from PROMs. The selected outcomes were constipation, diarrhea, dysmotility, incontinence, reflux, and SIBO.

The PheWAS analysis involved multiple univariate logistic regressions. Each of the 6 key SSc-GI outcomes were taken in turn and regressed against every diagnosis and drug in multiple univariate models.

The P value threshold for PheWAS "hits," that is, variables that are significantly associated with a trait, was set using the Benjamini-Hochberg false discovery rate method. This was applied for each of the 6 key SSc-GI outcomes separately. An adjusted P value < 0.05 was the threshold for a PheWAS hit. This unbiased approach cannot identify causal relationships but does identify associations, in part because a temporal relationship is not evaluated. As such, our findings should be viewed as hypothesis generating and will require validation in future studies.

- *GIT 2.0 linear models of symptom scores.* Six outcomes were explored from the GIT 2.0 GI symptoms score: total GI symptoms score, reflux, bloating, diarrhea, constipation, and soilage (incontinence). For each GIT 2.0 outcome, the PheWAS hits were entered as predictors in univariate linear models. Using the GIT 2.0 symptom scores, PheWAS hits were confirmed if both the 95% CI did not include 0 and the effect direction was concurrent with the initial hit.

Analysis was conducted in R (version 4.1.314).[13]

RESULTS

Characteristics of the SSc observational cohort. From the digital patient records of SSc cases within SMART, 12,535 documented clinical assessments from the clinic visits of 2058 consenting participants at the Royal Free Hospital (United Kingdom) were extracted into a dataset. Of these participants, 1546 had at least 3 documented clinical assessments on record and were included in the analysis.

Subcohort data for GIT 2.0. In the GIT 2.0 subcohort, 370 participants completed the GIT 2.0 symptoms questionnaire, and these data were used to confirm and explore the PheWAS hits. These were unselected consecutive consenting patients attending the clinical service over a period of 6 months collected as part of a specific study on the burden of GI symptoms in SSc. The clinical and demographic characteristics of the study cohorts are summarized in Table 1.

Quality control of labeling. A subset of diagnostic labels was manually curated in 86 cases by clinicians with full access to the notes. This revealed a high accuracy of programmatic labeling

Table 1. Description of SMART observational cohort and GIT 2.0 confirmatory subcohort.

	n	Missing, n	Results[a]
SMART, n = 1546			
Female	1296	0	83.83
Limited SSc	1030	0	66.62
Diffuse SSc	516	0	33.38
Disease onset	1489	57	1959-2019[b]
Disease duration to end of study window	1489	57	16.5 (1.9-62)[c]
ACA	394	0	25.49
ATA	376	0	24.32
ARA	169	0	10.93
Anti-PM/Scl	71	0	4.59
U1RNP	89	0	5.76
U3RNP	61	0	3.95
ThT0	8	0	0.52
GIT 2.0, n = 370			
Female	308	0	83.2
Limited SSc	228	0	61.6
Diffuse SSc	140	0	37.8
Juvenile SSc	2	0	0.5
Disease onset	360	10	1962-2018[b]
Disease duration	360	10	13.7 (1-57)[c]
ANA	345	4	93.2
ENA	303	3	81.9
ACA	115	0	31.0
ATA	83	0	22.4
ARA	43	0	11.6
Anti-PM/Scl	18	0	4.9
U1RNP	22	0	5.9
U3RNP	14	0	3.8

[a] Units are in percent unless otherwise indicated. [b] Year, range. [c] Years, median (range). ACA: anticentromere antibody; ANA: antinuclear antibody; ARA: anti-RNA polymerase III antibody; ATA: autoantibody; ENA: extractable nuclear antigen; GIT 2.0: University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument 2.0 questionnaire; SMART: Scleroderma Cohort; SSc: systemic sclerosis.

(accuracy 0.91-0.98) and high specificity. Although sensitivity was lower, especially for rarer labels (eg, cardiac involvement), the rate of mislabeling was low.

Prevalence of drugs. We identified 634 distinct drugs in the SMART observational cohort. The most prevalent drugs were omeprazole (866/1546, 56.02%), losartan (758/1546, 49.03%), lansoprazole (649/1546, 41.98%), and mycophenolate mofetil (618/1546, 39.97%).

Prevalence of diagnoses. We identified 673 distinct diagnoses in the SMART observational cohort. The most prevalent diagnoses were limited SSc (1005/1546, 65.01%), RP (773/1546, 50%), ILD (557/1546, 36.03%), and overlap syndrome (417/1546, 26.97%). It should be noted that when RP is listed, it means that RP was a particularly significant problem for that patient. Thus, absence of association with RP as a separate entity does not imply absence of RP; indeed, previous analysis confirmed that almost all SSc cases in SMART have RP.[3]

PheWAS analysis. PheWAS analysis demonstrated 88 hits across the 6 SSc-GI domains (56 drugs, 26 diagnoses, 6 ANAs), involving 37 distinct drugs, 18 distinct diagnoses, and 3 distinct ANAs (Figure 1).

Constipation. For constipation, 12 drugs, 4 diagnoses, and 1 ANA were significantly associated. The largest effect sizes were prucalopride (odds ratio [OR] 50.50, $P < 0.001$), bisacodyl (OR 36.70, $P < 0.001$), docusate (OR 23.10, $P < 0.001$), migraine (OR 13.80, $P < 0.001$), and aspiration (OR 14.80, $P = 0.005$). Of interest, amitriptyline (OR 3.31, $P = 0.01$), nicorandil (OR 12.80, $P = 0.05$), fentanyl (OR 11.50, $P = 0.002$), ANA negativity (OR 4.84, $P = 0.046$), and fexofenadine (OR 5.09, $P = 0.01$) were also linked to constipation.

Diarrhea. For diarrhea, 4 drugs, 4 diagnoses, and 1 ANA were significantly associated. The largest effect sizes were cimetidine (OR 11.60, $P = 0.02$), rifaximin (OR 10.10, $P = 0.005$), ambrisentan (OR 9.51, $P < 0.001$), loperamide (OR 8.45, $P < 0.001$), and constipation (OR 7.74, $P < 0.001$). Of interest, pulmonary hypertension (OR 3.81, $P = 0.001$), and ACA (OR 3.04, $P = 0.01$) were also linked to diarrhea.

Dysmotility. For dysmotility, 14 drugs, 4 diagnoses, and 1 ANA were significantly associated. The largest effect sizes were

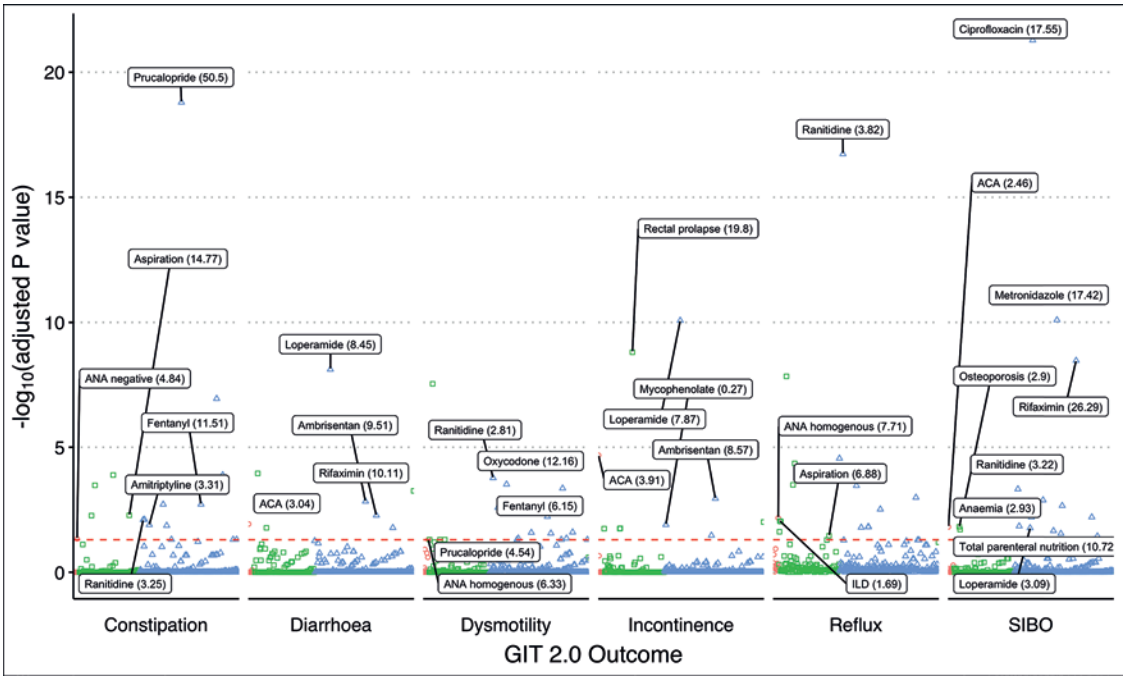

Figure 1. PheWAS plot showing key clinical and drug associations with GI complications of SSc. PheWAS analysis generated "hits" with drugs, diagnoses, and ANAs across 6 SSc-GI outcomes. ANAs are represented by red circles, diseases are in green, drugs are in blue, and the red line is adjusted $P < 0.05$. Certain associations are labeled, with ORs in brackets. Constipation was linked to fentanyl, amitriptyline, migraine, ANA negative, aspiration, and incontinence. Diarrhea was linked to ambrisentan and ACA. Dysmotility was linked to ANA homogenous, tamsulosin, fentanyl, and oxycodone. Incontinence was linked to ACA, ambrisentan, and inversely linked to mycophenolate. Reflux was linked to ILD and ANA homogenous. SIBO was linked to ACA, total parenteral nutrition, osteoporosis, and anemia. ACA: anticentromere autoantibody; ANA: antinuclear antibody; GI: gastrointestinal; GIT 2.0: University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument 2.0 questionnaire; ILD: interstitial lung disease; OR: odds ratio; PheWAS: phenome-wide association study; SIBO: small intestinal bacterial overgrowth; SSc: systemic sclerosis.

oxycodone (OR 12.20, $P < 0.001$), barium (OR 9.43, $P = 0.05$), docusate (OR 9.07, $P = 0.03$), bisacodyl (OR 7.04, $P = 0.04$), and tamsulosin (OR 6.92, $P = 0.006$). Of interest, iloprost (OR 2.55, $P = 0.003$), RP (OR 1.97, $P = 0.05$), ANA homogenous (OR 6.33, $P = 0.05$), SIBO (OR 2.99, $P = 0.05$), zoledronic acid (OR 3.27, $P = 0.05$), and thiamine (OR 6.33, $P = 0.05$) were also linked to dysmotility.

Incontinence. For incontinence, 4 drugs and 5 diagnoses were significantly associated. The largest effect sizes were rectal prolapse (OR 19.80, $P < 0.001$), ambrisentan (OR 8.57, $P = 0.001$), loperamide (OR 7.87, $P < 0.001$), gamolenic acid (evening primrose oil; OR 5.47, $P = 0.03$), and constipation (OR 5.33, $P = 0.001$). Of interest, ACA (OR 3.91, $P < 0.001$), limited SSc (OR 3.85, $P = 0.02$), and mycophenolate (OR 0.27, $P = 0.01$) were also linked to incontinence—noting that mycophenolate had decreased odds of incontinence compared to the others.

Reflux. For reflux, 9 drugs, 7 diagnoses, and 1 ANA were significantly associated. The largest effect sizes were famotidine (OR 13.10, $P = 0.001$), ANA homogenous (OR 7.71, $P = 0.007$), aspiration (OR 6.88, $P = 0.03$), diarrhea (OR 4.62, $P < 0.001$), and dysmotility (OR 4.27, $P < 0.001$). Of interest, ILD (OR 1.69, $P = 0.009$), RP (OR 1.61, $P = 0.02$), digital ulcers (OR 1.84, $P = 0.01$), doxycycline (OR 2.86, $P = 0.02$), domperidone (OR 2.39, $P < 0.001$), and metoclopramide (OR 2.69, $P < 0.001$) were also linked to reflux.

SIBO. For SIBO, 13 drugs, 2 diagnoses, and 1 ANA were significantly associated. The largest effect sizes were rifaximin (OR 26.30, $P < 0.001$), ciprofloxacin (OR 17.50, $P < 0.001$), metronidazole (OR 17.40, $P < 0.001$), TPN (OR 10.70, $P = 0.006$), and docusate (OR 8.91, $P = 0.04$). Of interest, ACA (OR 2.46, $P = 0.02$), anemia (OR 2.93, $P = 0.02$), and osteoporosis (OR 2.90, $P = 0.02$) were also linked to SIBO.

PheWAS hits were confirmed and explored using GI symptom scores (GIT 2.0). The 29 distinct PheWAS hits (drugs, diagnoses, and ANAs) were confirmed and explored formally using the patient-reported GI symptom scores (GIT 2.0) across GI domains: reflux, bloating, diarrhea, constipation, soilage (incontinence), and the total score. This allowed a more detailed analysis of associations, as summarized in Figure 2, including both the direction and the magnitude of the effect on patient-reported SSc-GI symptoms (Table 2 and Table 3).

Confirmed associations with drugs. Twenty-two drugs were significantly associated with GI symptom scores. Many drugs were associated with multiple symptom domains, including the total score: ranitidine (6 domains), domperidone (5 domains), omeprazole (5 domains), amitriptyline (4 domains), and loperamide (4 domains), for example (Figure 2). The largest increase

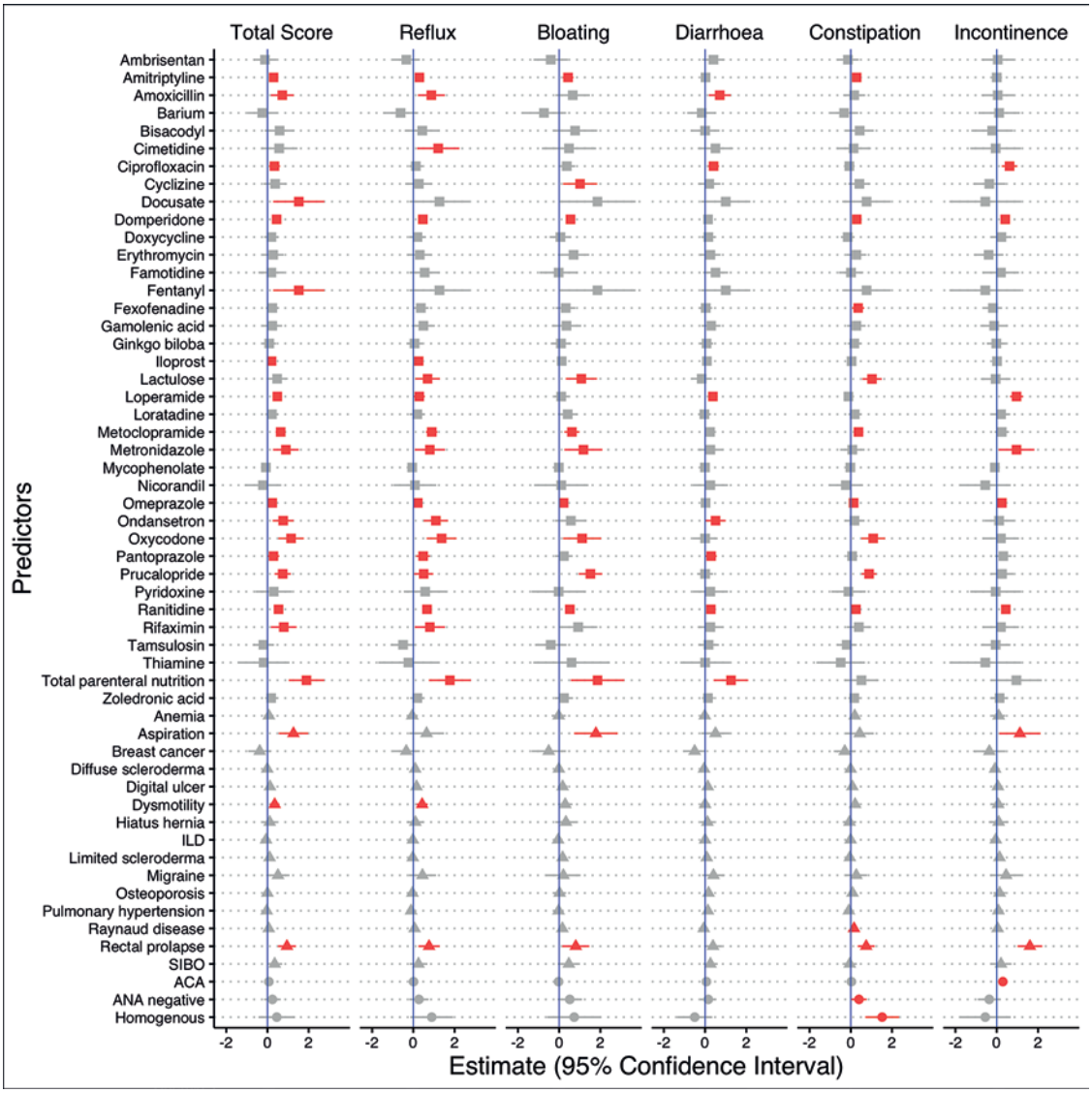

Figure 2. PheWAS hits confirmed and explored with GIT 2.0 symptom scores, including total score, reflux, bloating, diarrhea, constipation, and incontinence. Estimates with 95% CIs are presented; red indicates $P < 0.05$. Drugs are represented by squares, diagnoses are represented by triangles, and ANAs are represented by circles. ACA: anticentromere autoantibody; ANA: antinuclear antibody; GIT 2.0: University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument 2.0 questionnaire; ILD: interstitial lung disease; PheWAS: phenome-wide association study; SIBO: small intestinal bacterial overgrowth.

in symptom scores were seen with TPN (total, bloating, diarrhea, and reflux), docusate (total), prucalopride (bloating), and oxycodone (reflux, total, bloating, and constipation). Drugs not indicated for GI symptoms included the following: amitriptyline (bloating, reflux, total score, and constipation), fentanyl (total score), fexofenadine (constipation), iloprost (reflux and total score), and oxycodone (reflux, total score, bloating, and constipation). For PheWAS domains with a matching GIT 2.0 domain, associations were confirmed with greater certainty: fexofenadine, prucalopride, lactulose, amitriptyline, domperidone, and ranitidine (constipation); loperamide (diarrhea); loperamide (incontinence); and ondansetron, ranitidine, metoclopramide, pantoprazole, domperidone, and omeprazole (reflux).

Confirmed associations with diagnoses. Four diagnoses were significantly associated with GI symptom scores (Figure 2). Aspiration was associated with increased bloating, total score, and incontinence. Rectal prolapse was associated with increased incontinence, total score, bloating, reflux, and constipation. Dysmotility was associated with increased reflux and total score. RP was associated with increased constipation. For PheWAS domains with a matching GIT 2.0 domain, the following associations were confirmed with greater certainty: rectal prolapse (incontinence) and dysmotility (reflux).

Confirmed associations with ANAs. Three ANAs were significantly associated with GI symptom scores (Figure 2). ANA homogenous pattern and ANA negativity were associated with increased constipation. ACA was associated with

Table 2. Confirmed associations with matching PheWAS outcome and GIT 2.0 domain.

			PheWAS Model		GIT 2.0 Linear Model		
Outcome	Type	Predictor	OR	Adjusted *P*	Estimate	95% CI	*P*
Constipation	ANA	ANA negative	4.84	0.046	0.39	0.02-0.77	0.04
Constipation	Drug	Amitriptyline	3.31	0.01	0.28	0.11-0.44	0.001
Constipation	Drug	Domperidone	3.28	0.01	0.29	0.14-0.43	< 0.001
Constipation	Drug	Fexofenadine	5.09	0.01	0.35	0.06-0.65	0.02
Constipation	Drug	Lactulose	5.64	0.002	1.03	0.56-1.50	< 0.001
Constipation	Drug	Prucalopride	50.50	< 0.001	0.89	0.52-1.25	< 0.001
Constipation	Drug	Ranitidine	3.25	0.008	0.25	0.12-0.38	< 0.001
Diarrhea	Drug	Loperamide	8.45	< 0.001	0.38	0.16-0.61	0.001
Reflux	Diagnosis	Dysmotility	4.27	< 0.001	0.43	0.16-0.69	0.002
Reflux	Drug	Domperidone	2.39	< 0.001	0.46	0.28-0.64	< 0.001
Reflux	Drug	Metoclopramide	2.69	< 0.001	0.89	0.62-1.17	< 0.001
Reflux	Drug	Omeprazole	2.22	< 0.001	0.22	0.07-0.37	0.005
Reflux	Drug	Ondansetron	4.11	0.003	1.09	0.50-1.68	< 0.001
Reflux	Drug	Pantoprazole	2.47	0.02	0.48	0.15-0.81	0.005
Reflux	Drug	Ranitidine	3.82	< 0.001	0.67	0.52-0.81	< 0.001
Incontinence	ANA	ACA	3.91	< 0.001	0.29	0.09-0.50	0.005
Incontinence	Diagnosis	Rectal prolapse	19.80	< 0.001	1.61	1.01-2.21	< 0.001
Incontinence	Drug	Loperamide	7.87	< 0.001	0.96	0.64-1.28	< 0.001

ACA: anticentromere autoantibody; ANA: antinuclear autoantibody; GIT 2.0: University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument 2.0 questionnaire; OR: odds ratio; PheWAS: phenome-wide association study.

Table 3. Confirmed associations with GIT 2.0 total score.

			PheWAS Model			GIT 2.0 Linear Model		
Type	Predictor	Outcome	OR	Adjusted *P*	Outcome	Estimate	95% CI	*P*
Disease	Aspiration	Constipation	14.80	0.005	Total score	1.26	0.55-1.97	< 0.001
Drug	Amitriptyline	Constipation	3.31	0.01	Total score	0.29	0.12-0.47	0.001
Drug	Docusate	Constipation	23.10	< 0.001	Total score	1.52	0.28-2.76	0.02
Drug	Domperidone	Constipation	3.28	0.008	Total score	0.44	0.29-0.59	< 0.001
Drug	Fentanyl	Constipation	11.50	0.002	Total score	1.52	0.28-2.76	0.02
Drug	Prucalopride	Constipation	50.50	< 0.001	Total score	0.74	0.34-1.13	< 0.001
Drug	Ranitidine	Constipation	3.25	0.008	Total score	0.54	0.41-0.66	< 0.001
Drug	Loperamide	Diarrhea	8.45	< 0.001	Total score	0.48	0.25-0.71	< 0.001
Drug	Rifaximin	Diarrhea	10.10	0.005	Total score	0.79	0.17-1.41	0.01
Disease	Aspiration	Reflux	6.88	0.03	Total score	1.26	0.55-1.97	< 0.001
Disease	Dysmotility	Reflux	4.27	< 0.001	Total score	0.35	0.12-0.58	0.004
Drug	Domperidone	Reflux	2.39	< 0.001	Total score	0.44	0.29-0.59	< 0.001
Drug	Metoclopramide	Reflux	2.69	< 0.001	Total score	0.64	0.40-0.88	< 0.001
Drug	Omeprazole	Reflux	2.22	< 0.001	Total score	0.23	0.10-0.36	< 0.001
Drug	Ondansetron	Reflux	4.11	0.003	Total score	0.76	0.25-1.27	0.004
Drug	Pantoprazole	Reflux	2.47	0.02	Total score	0.29	0.00-0.57	0.0498
Drug	Ranitidine	Reflux	3.82	< 0.001	Total score	0.54	0.41-0.66	< 0.001
Disease	Rectal prolapse	Soilage	19.80	< 0.001	Total score	0.94	0.51-1.38	< 0.001
Drug	Loperamide	Soilage	7.87	< 0.001	Total score	0.48	0.25-0.71	< 0.001

GIT 2.0: University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument 2.0 questionnaire; OR: odds ratio; PheWAS: phenome-wide association study.

incontinence. For PheWAS domains with a matching GIT 2.0 domain, the following associations were confirmed with greater certainty: ANA negativity (constipation) and ACA (incontinence).

DISCUSSION

In this study, we explored GI involvement in SSc, mapping drug exposures, comorbidities, and ANA. Using real-world data, we revealed both the expected and appropriate prescribing of drugs to treat GI symptoms, but also potential causes and contributors to GI dysfunction in SSc. We confirmed and explored initial associations (hits) with detailed patient-reported GI symptom scores, confirming multiple hits and exploring the effects on GI symptoms.

The PheWAS analysis included 1546 individuals; the majority were limited SSc subtype, and ANA subgroups reflected a typical SSc population. The median follow-up was long (16.5 years), supporting the ascertainment of SSc-GI outcomes that can develop late in the disease. The confirmatory GIT 2.0 cohort (n = 370) was slightly enriched for diffuse SSc.

Almost 700 distinct diagnoses were identified in the clinical records. This permitted a hypothesis-generating approach, with the diagnostic space not limited to prior associations with SSc. However, as expected, the most frequent diagnoses and comorbidities were those directly related to SSc, including limited SSc, RP, and ILD. However, diagnoses not specific to SSc, such as migraine, osteoporosis, anemia, and breast cancer, were associated with SSc-GI in PheWAS analysis, underlining the scope of this approach.

Over 600 distinct drugs were identified in the clinical records. As expected, the top drugs identified were those indicated for the management of SSc manifestations, including PPIs, angiotensin receptor blockers, and mycophenolate. However, most medicines were prescribed for a non-SSc indication, and as such we were able to assess the potential contribution of such drugs to SSc-GI dysfunction.

Eighty-eight hits were generated by the PheWAS model by domain (no. of hits): constipation (17), diarrhea (9), dysmotility (19), incontinence (10), reflux (17), and SIBO (16). Interestingly, despite a similar total number of distinct drugs and diagnoses identified, most of the hits were for drugs (56), followed by diagnoses (26), followed by ANAs (6). Three of the 6 ANA hits were for ACA, associated with diarrhea, incontinence, and SIBO. This fits with prior work demonstrating a higher burden of SSc-GI disease with ACA.[4]

Altogether, 29 of the 88 hits were confirmed using the GIT 2.0 patient-reported symptom scores; certain hits were significantly associated with multiple domains of the GIT 2.0. For PheWAS hits with a matching domain of the GIT 2.0 (constipation, diarrhea, incontinence, and reflux), we were able to confirm these associations with a greater degree of certainty.

Bringing together related PheWAS hits, RP, digital ulcers, iloprost, ACA, and migraine were all linked to increased GI disease. This may suggest that, between the canonical disease mechanisms in SSc, vasculopathy is a prominent driver of GI disease. Among these hits linked to vasculopathy, RP, iloprost, and ACA were confirmed to increase GI symptoms on the GIT 2.0.

Hits were discovered linking reflux with aspiration and ILD, which may reflect a causal pathway leading from reflux, to aspiration, to ILD. Although the most significant drivers of ILD might be inflammation and fibrosis, reflux is known to contribute to the progression of ILD.[14] As such, this association serves to emphasize the importance of controlling reflux, especially in those at risk of ILD progression.

ANA negativity was confirmed to be associated with constipation in the GIT 2.0. Prior work has identified a link between ANA-negative subtype and increased lower GI disease, specifically malabsorption.[14]

Downstream of SSc-GI disease, several likely consequences of GI dysfunction were identified. For example, SIBO was linked to osteoporosis and anemia in the PheWAS, highlighting the consequences of GI disease. Malnutrition in SSc is multifactorial,[15] including decreased appetite and upper GI dysmotility, but SIBO is certainly a driver of malabsorption. As such, treatments for SIBO not only improve the symptom burden (bloating, flatulence, discomfort) but may also improve nutritional status.[16]

The associations between drugs and SSc-GI should be considered in 2 groups. First, drugs prescribed for symptom control (appropriate prescribing)[17] and, second, drugs prescribed for another indication (eg, opioids for pain) with an off-target effect of GI disease.

The PheWAS hit linking amitriptyline and constipation was confirmed by the GIT 2.0 constipation symptom score. Amitriptyline has significant anticholinergic activity, and side effects include dry mouth, constipation, and nausea; paralytic ileus is a rare side effect. Autoantibodies to the muscarinic-3 receptor autoantibodies have been identified in those with severe SSc-GI involvement.[18] Additionally, there is observational evidence that cholinesterase inhibitor pyridostigmine is effective for SSc-GI symptoms, in particular constipation.[6] Taken together, this may prompt clinicians to review patients with a high burden of anticholinergic medication.

Oxycodone, initially a hit for dysmotility, was confirmed to be associated with reflux, bloating, constipation, and the total GI symptom scores. Opioids have well-recognized GI side effects and previous work has identified opioids as a risk factor for intestinal pseudo-obstruction in SSc.[19] Additionally, fentanyl was confirmed to increase the total GI symptom score, after it was determined to be a PheWAS hit for constipation, supporting the hypothesis.

Fexofenadine is a highly specific histamine H1 receptor reverse agonist, binding to and stabilizing the inactive form of the receptor. It is reported to have low off-target effects. The only GI side effect reported for the product is nausea, which is common. In this study, fexofenadine was a hit for constipation in the PheWAS, which was confirmed by increased constipation symptom score in the GIT 2.0 model. As such, there is a higher degree of confidence that this is a real association. However, it is possible that causally, the indication for fexofenadine (allergy/atopic conditions) is a confounder of the potential causal effect, and thus further investigation is warranted.

Ambrisentan was a hit for diarrhea; this was a recognized side effect in clinical trials.[20] Use of ambrisentan may also be a surrogate marker for severe vasculopathy, which in itself might be linked to more severe GI disease.

Tamsulosin was a hit for dysmotility. As an alpha-1 receptor antagonist, it leads to the relaxation of smooth muscle. Constipation and diarrhea are reported to be uncommon side effects; however, the plausibility of the mechanism supports this hypothesis.

Iloprost was a hit for dysmotility, increased reflux, and the total GI scores in the GIT 2.0 analysis. This relationship could be confounded by vasculopathy, making a direct effect of iloprost on GI dysfunction less likely.

The hit for mycophenolate suggested a protective effect for incontinence (OR 0.27, $P = 0.01$). Although this could be a treatment effect targeting inflammatory-driven GI dysfunction, it is possible that—as mycophenolate prescription is linked with diffuse SSc, and ACA/limited SSc is linked to GI disease, especially incontinence—the relationship is confounded by skin subset, which was not adjusted for in the PheWAS analysis. An observational study of a US healthcare insurance claims database suggested that the real-world prescribing of immunomodulatory therapy for SSc was not enriched for certain organ involvement, including GI disease.[21] However, in that study, the prevalence of GI involvement at 1 year was 22%, which is low compared to cohort studies.[4]

In this study, we mapped the links between comorbidities and SSc-GI disease. As well as demonstrating expected mechanistic associations between diagnoses (eg, reflux and aspiration), our approach yielded links that may suggest novel disease mechanisms (eg, migraine being linked to constipation in SSc). Regarding drugs, we highlight their adverse effects. In addition, mapping drug effects in the context of SSc may shed light on SSc disease mechanisms. Real-world data including drugs prescribed for an orthogonal indication to SSc-GI disease may demonstrate the disease-specific effects of drug target perturbation, which could be explored further with the aim of drug repurposing.

There are several strengths to the present study. These include a long duration of follow-up of the cohort, and the granularity of detail for individual clinical records available for phenotyping. The use of a PROM, the GIT 2.0, to confirm initial PheWAS hit, strengthens our conclusions. Many of the PheWAS hits and those confirmed with the GIT 2.0 replicate well-known associations between drugs and diagnoses in SSc, particularly for the most common treatments for SSc-GI manifestations, including the following: PPIs, H2-receptor antagonists, antibiotics for SIBO, laxatives, prucalopride, antiemetics, and promotility agents. This supports the validity of the novel associations uncovered.

There are also some clear limitations. This study used real-world data, and as such there are possible biases. This includes potential single-center bias and that our first-line immunosuppressive treatment is usually mycophenolate mofetil. In addition, it is possible that as a large tertiary referral center, we see more severe SSc cases than in general rheumatology practice. This may limit generalizability of our findings. The documented clinical assessments are written in a semistructured format, and we would expect a degree of variation between clinicians and over time; we included individuals with at least 3 separate documented clinical assessments across time to mitigate this variation. The GIT 2.0 subcohort was cross-sectional, at a timepoint toward the end of the SMART observational cohort window. Although this was adequate for the purpose of confirming PheWAS hits, longitudinal GIT 2.0 data would allow the examination of the relationships between risk factor and SSc-GI over time. We could not look in detail at temporal association and we recognize this limitation. The associations we report do not imply a causal direction of effect, and as such, the order of events does not invalidate the associations from the PheWAS analysis. We interpret the discovered set of associations and the subset validated by patient-reported outcome scores in the context of prior knowledge, including pharmacodynamic and adverse event profiles.

In conclusion, we have used a novel analytical approach in a large single-center observational cohort to explore the association of drug treatments and disease characteristics with significant GI manifestations in SSc and the associated symptom burden. Our findings have face validity and reflect previous studies, but they also highlight the relevance of treatments for non-GI complications. These findings highlight the importance of careful and integrated multidisciplinary care for SSc, including specialist pharmacist input and routine assessment of GI symptom severity using validated PROMs. Future work using real-world data covering drugs, diagnoses, and disease-related outcomes might look at polypharmacy in complex autoimmune diseases, especially the co-occurrence of drug-disease pairs and appropriate, insufficient, and problematic polypharmacy.

REFERENCES

1. Volkmann ER, McMahan Z. Gastrointestinal involvement in systemic sclerosis: pathogenesis, assessment and treatment. Curr Opin Rheumatol 2022;34:328-36.
2. Nassar M, Ghernautan V, Nso N, et al. Gastrointestinal involvement in systemic sclerosis: an updated review. Medicine 2022;101:e31780.
3. Nihtyanova SI, Sari A, Harvey JC, et al. Using autoantibodies and cutaneous subset to develop outcome-based disease classification in systemic sclerosis. Arthritis Rheumatol 2020;72:465-76.
4. Ahmed F, Maclean RH, Nihtyanova SI, Ong VH, Murray CD, Denton CP. Autoantibody predictors of gastrointestinal symptoms in systemic sclerosis. Rheumatology 2022;61:781-6.
5. Cheah JX, Khanna D, McMahan ZH. Management of scleroderma gastrointestinal disease: lights and shadows. J Scleroderma Relat Disord 2022;7:85-97.
6. Ahuja NK, Mische L, Clarke JO, Wigley FM, McMahan ZH. Pyridostigmine for the treatment of gastrointestinal symptoms in systemic sclerosis. Semin Arthritis Rheum 2018;48:111-6.
7. Dein EJ, Wigley FM, McMahan ZH. Linaclotide for the treatment of refractory lower bowel manifestations of systemic sclerosis. BMC Gastroenterol 2021;21:174.
8. Bastarache L, Denny JC, Roden DM. Phenome-wide association studies. JAMA 2022;327:75-6.
9. Khanna D, Hays RD, Maranian P, et al. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum 2009;61:1257-63.
10. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737-47.
11. Schriml LM, Mitraka E, Munro J, et al. Human Disease Ontology 2018 update: classification, content and workflow expansion. Nucleic Acids Res 2019;47:D955-62.
12. Wishart DS, Knox C, Guo AC, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res 2006;34:D668-72.
13. R Core Team. R: a language and environment for statistical computing. 2022. [Internet. Accessed February 3, 2022.] Available from: https://R-project.org
14. Christmann RB, Wells AU, Capelozzi VL, Silver RM. Gastroesophageal reflux incites interstitial lung disease in systemic

sclerosis: clinical, radiologic, histopathologic, and treatment evidence. Semin Arthritis Rheum 2010;40:241-9.
15. Salazar GA, Assassi S, Wigley F, et al. Antinuclear antibody-negative systemic sclerosis. Semin Arthritis Rheum 2015;44:680-86.
16. Harrison E, Herrick AL, McLaughlin JT, Lal S. Malnutrition in systemic sclerosis. Rheumatology 2012;51:1747-56.
17. Hansi N, Thoua N, Carulli M, et al. Consensus best practice pathway of the UK scleroderma study group: gastrointestinal manifestations of systemic sclerosis. Clin Exp Rheumatol 2014;32:S214-21.
18. Kawaguchi Y, Nakamura Y, Matsumoto I, et al. Muscarinic-3 acetylcholine receptor autoantibody in patients with systemic sclerosis: contribution to severe gastrointestinal tract dysmotility. Ann Rheum Dis 2009;68:710-14.
19. Dein E, Kuo PL, Hong YS, Hummers LK, Mecoli CA, McMahan ZH. Evaluation of risk factors for pseudo-obstruction in systemic sclerosis. Semin Arthritis Rheum 2019;49:405-10.
20. Galiè N, Barberà JA, Frost AE, et al. Initial use of ambrisentan plus tadalafil in pulmonary arterial hypertension. N Engl J Med 2015;373:834-44.
21. Gale SL, Trinh H, Mathew N, Jahreis A, Lin CJF, Sarsour K. Characterizing disease manifestations and treatment patterns among adults with systemic sclerosis: a retrospective analysis of a US healthcare claims population. Rheumatol Ther 2020;7:89-99.