# Approach to Phenotyping for Diverticular Disease Subphenotypes

### Overview

Colonic diverticular disease represents a spectrum from an asymptomatic precursor of diverticulosis to complicated diverticulitis through development of an abscess, fistula, stricture, macroperforation, peritonitis, or obstruction. Thus far, excellent work has been done by Pacheco & Thompson1 as published in Joo et al (2023)2 in defining a reliable algorithm for differentiating cases of any diverticular disease from controls of no diverticular disease. There is also a need for categorizing disease severities. The introduction of ICD-10 codes provided a mechanism via sub-codes to identify diverticular disease complications, but performance has been highly variable.3 An alternative framework is available from De Roo et al (2023)4 who approximate the severity spectrum through the lens of operative disease, recurrent inpatient admissions, or outpatient encounters.

This work intended to unify existing rule-based strategies for categorizing diverticular disease in electronic medical record cohorts and apply them to colonic diverticular disease.

### Algorithm Description

A flowchart of the algorithm is shown in Figure 1 with accompanying descriptions in Table 1. Of note, this algorithm captures colonic diverticular disease, and extensions to small bowel diverticular disease would need to revise the inclusion code list. A list of diagnostic and procedure codes can be found in the attached Data Dictionary excel file. To separate cases of diverticular disease from controls with no diverticular disease, we adopt Pacheco & Thompson’s non-NLP algorithm with the addition of ICD-10 codes. Further division of the cases into sub-phenotypes is completed in a manner analogous to the efforts from De Roo et al (2023)4. The algorithm was implemented in a deidentified institutional medical records repository, Vanderbilt’s Synthetic Derivative along with plausible variations. Specifically, we considered requirement of a diverticular disease antibiotic prescription for the assignment of mild diverticulitis as well as broadening the range of days between diverticular disease ICD9/10 code and CT scan / lower GI scope code from 7 to 30 days.

Figure 1



|  |
| --- |
| *Table 1* |
| **Case-control Assignment** | **Sub-phenotype Assignment** | **Individual Criteria** | **Criteria relationship for inclusion** |
| Control | Control | 1. One or more lower GI scope procedures
2. No ICD9/10 codes for diverticulosis or diverticulitis
 | 1 AND 2 |
| Case | Any diverticular disease | 1. One or more lower GI scope procedures
2. One or more ICD9/10 codes for diverticulosis or diverticulitis
3. ICD9/10 code for diverticulosis or diverticulitis applied within 0-7 days after an abdominal CT scan or lower GI scope
 | 1 AND 2 AND 3 |
| Case | Operative or Recurrent Inpatient Diverticulitis | 1. Qualifies as a case of any diverticular disease
2. Two or more inpatient admissions associated with an ICD9/10 code for diverticulitis
3. At least one diverticulitis procedure during the same encounter as an ICD9/10 code for diverticulitis
 | 1 AND (2 OR 3) |
| Case | Mild Diverticulitis | 1. Qualifies as a case of any diverticular disease
2. At least one ICD9/10 codes for diverticulitis
3. No more than 1 inpatient admission for diverticulitis
 | 1 AND 2 AND 3 |
| Case | Diverticulosis | 1. Qualifies as a case of any diverticular disease
2. At least one ICD9/10 codes for diverticulosis
3. No ICD9/10 codes for diverticulitis
 | 1 AND 2 AND 3 |
| ICD: International Classification of Diseases. GI: Gastrointestinal. CT: computed tomography |

### References

1. Pacheco JenniferA, Thompson W. Diverticulosis and Diverticulitis. PheKB. Northwestern University The Feinberg School of Medicine. Accessed January 8, 2023. https://phekb.org/phenotype/diverticulosis-and-diverticulitis

2. Joo YY, Pacheco JA, Thompson WK, et al. Multi-ancestry genome- and phenome-wide association studies of diverticular disease in electronic health records with natural language processing enriched phenotyping algorithm. *PLoS One*. 2023;18(5):e0283553. doi:10.1371/journal.pone.0283553

3. Erichsen R, Strate L, Sørensen HT, Baron JA. Positive predictive values of the International Classification of Disease, 10th edition diagnoses codes for diverticular disease in the Danish National Registry of Patients. *Clin Exp Gastroenterol*. 2010;3:139-142. doi:10.2147/CEG.S13293

4. De Roo AC, Chen Y, Du X, et al. Polygenic Risk Prediction in Diverticulitis. *Ann Surg*. 2023;277(6):e1262-e1268. doi:10.1097/SLA.0000000000005623