Appendix

Table 1. Quality Assessment using JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies

Parameters	Geiss et al., 2018	Melchior et al., 2017	Mules et al., 2022	Chojnacki et al., 2022	Iordache et al., 2023
1. **Were the criteria for inclusion in the sample clearly defined?**	1	1	1	1	1
2. **Were the study subjects and the setting described in detail?**	1	1	1	1	1
3. **Was the exposure measured in a valid and reliable way?**	0	1	1	1	1
4. **Were objective, standard criteria used for measurement of the condition?**	1	1	1	1	1
5. **Were confounding factors identified?**	1	1	1	1	1
6. **Were strategies to deal with confounding factors stated?**	1	1	1	2	1
7. **Were the outcomes measured in a valid and reliable way?**	1	1	1	1	1
8. **Was appropriate statistical analysis used?**	1	1	1	1	1

Note. 0 indicates “Not Applicable”; 1 indicates “Yes”; 2 indicates “No”; 3 indicates “Unclear”.

Table 2. Quality Assessment using JBI Critical Appraisal Checklist for Cohort Studies

Parameters	Liskiewicz et al., 2021	Asscher et al., 2022
1. **Were the two groups similar and recruited from the same population?**	1	1
2. **Were the exposures measured similarly to assign people to both exposed and unexposed groups?**	1	1
3. **Was the exposure measured in a valid and reliable way?**	1	1
4. **Were confounding factors identified?**	1	1
5. **Were strategies to deal with confounding factors stated?**	1	1
6. **Were the groups/participants free of the outcome at the start of the study (or at the moment of exposure)?**	0	0
7. **Were the outcomes measured in a valid and reliable way?**	1	1
8. **Was the follow up time reported and sufficient to be long enough for outcomes to occur?**	1	0
9. **Was follow up complete, and if not, were the reasons to loss to follow up described and explored?**	1	0
10. **Were strategies to address incomplete follow up utilized?**	0	0
11. **Was appropriate statistical analysis used?**	1	1

Note. 0 indicates “Not Applicable”; 1 indicates “Yes”; 2 indicates “No”; 3 indicates “Unclear”.

Table 3. Characteristics of Studies Included in the Review

Study	Inclusion criteria	Exclusion criteria	Depression Scale	Number of participants	Male (%)	Mean age (years)	Time of between FC test and depression rating	Relationship between depression scores and FC
Geiss et al., 2018	Aged 16-80; established diagnosis of CD or UC; available SIBDQ data; one FC or lactoferrin measurement within 30 days of outpatient clinic visit	Diagnosis of IBD-U; previous ileostomy, colostomy, or colectomy; known bacterial or viral infection or the GI tract	PHQ-9	CD: 228	47.4	38 (27-50)*	Any time frame (n=193)**	Positive correlation between PHQ-9 score and FC levels (r = 0.142, P =.042).
							+/- 30 days (n=150)**	No significant difference in PHQ-9 scores between CD patients with FC levels <200µg/g vs ≥200µg/g (P =.124).
							+/- 3 days (n=62)**	Significant difference in PHQ-9 scores between CD patients with FC levels <200µg/g vs ≥200µg/g (P =.021).
				UC: 120	49.2	30 (22-38)*	Any time frame (n=107)	Non-significant correlation between PHQ-9 score and FC levels (r = 0.028, P =.776).
							+/- 30 days (n=107)**	No significant difference in PHQ-9 scores between UC patients with FC levels <200µg/g vs ≥200µg/g (P =.55).
							+/- 3 days (n=21)**	No significant difference in PHQ-9 scores between UC patients with FC levels <200µg/g vs ≥200µg/g (P =.199).
Melchior et al., 2017	Rome III criteria IBS diagnosis	IBD diagnosis (determined by minimum 1-year clinical follow-up)	HADS-D	93	26.9	41	Unspecified	Non-significant correlation between FC and HADS-D scores (r = -0.00, P =.98).

Liskiewicz et al., 2021	Aged 16-85; experienced a depressive episode of at least moderate severity according to the ICD-10	Comorbid psychotic features or severe mental health disorders or treatment resistant depression; psychoactive substance use; cancer or GI disease; diabetes and thyroid dysregulation; antibiotic, non-steroidal anti-inflammatory drug, corticosteroid, immunomodulating drug, or proton pump inhibitor use for ≥3 months before baseline assessment	HDRS24	16	50.0	42.9	Unspecified, but must have been within at least 7 days due to the study design	Non-significant correlation between baseline FC levels and baseline HDRS24 scores. Positive correlation between change in HDRS24 and change in FC level after 6-week Escitalopram administration procedure (r = 0.67, P =.009).
Asscher et al. 2022	Aged 65+; confirmed clinical, endoscopic, and/or histologic diagnosis of CD, UC, or IBD-U	Language barriers (no Dutch or English)	GDS	No deficits***: 213 of 405 total	61.5	69.0 (67.0-72.0)*	+/-3 months	No significant difference in proportion of abnormal geriatric depression scores between patients with FC levels <250µg/g vs ≥250µg/g. No significant difference in proportion of abnormal geriatric depression scores between patients with FC levels <50µg/g vs ≥50µg/g.
				Moderate deficits***: 160 of 405 total	49.4	71.0 (68.0-75.0)*		
				Severe deficits***: 32 of 405 total	21.9	72.5 (70.3-79.8)*		
Mules et al., 2022	Aged 16+; confirmed endoscopic, histological or radiological diagnosis of CD	Unable to understand written English	PHQ-9	CD: 107	44.9	45	Unspecified, must have been within 7 days due to study design	Non-significant correlation between PHQ-9 and FC levels in patients with CD (spearman r = 0.04, P >.05).

	or UC for ≥3 months			UC: 65	49.2	50	Unspecified, must have been within 7 days due to study design	Non-significant correlation between PHQ-9 and FC levels in patients with UC (spearman r = 0.22, P >.05).
Chojnacki et al., 2022	Aged 24-60; healthy controls (negative LHBT); small intestinal bacterial overgrowth (SIBO) group (positive LHBT)	H-pylori-induced gastritis; Lymphocytic and ulcerative colitis; CD; allergy and food intolerance; liver or kidney diseases; diabetes; use of antibiotics, probiotics, and psychotropic drugs in month prior to enrollment	HAM-D	80	40.0	45.0	Unspecified	Positive correlation between HAM-D score and FC levels (spearman r = 0.33, P =.0105).
Iordache et al., 2023	Aged 18+; radiological, histological, or endoscopic diagnosis of CD or UC	Severe active IBD, severe psychiatric comorbidities (schizophrenia, dementia); PHQ-9 scores over 19	PHQ-9	30 (CD: 12, UC: 18)	50.0	50 (40-60)*	Unspecified	Positive correlation between PHQ-9 score and FC (spearman r = 0.416, P =.022). FC level of 131 µg/g or higher predicted depression with a sensitivity of 82%, a specificity of 61%, and an accuracy of 70%.

Abbreviations: CD, Crohn's Disease; FC, fecal calprotectin; GDS, Geriatric Depression Scale; GI, gastrointestinal; HADS-D, Hospital Anxiety and Depression Scale depression subscale; HAM-D, Hamilton Depression Rating Scale; HDRS24, Hamilton Depression Rating Scale 24-item version; IBD, inflammatory bowel disease; IBD-U, inflammatory bowel disease unclassified; ICD-10, International Classification of Diseases 10th Revision; PHQ-9, Patient Health Questionnaire 9-item version; LHBT, lactulose hydrogen breath test; SIBDQ, Short Inflammatory Bowel Disease Questionnaire; UC, ulcerative colitis.*Median participant age (IQR) was reported instead of mean age. **In Geiss et al (2018), FC levels within 30 days or within 3 days were not collected for all participants. ***In Asscher et al (2022), median age, sex proportion, and sample size was reported separately for participants with no, moderate, and severe deficits in geriatric assessment