Received: 25 September 2021 | Accepted: 18 November 2021

DOI: 10.1002/ueg2.12183

ORIGINAL ARTICLE

ueg journal WILEY

Composite outcomes in observational studies of ulcerative colitis: A systematic review and meta-analysis

Fernando Magro[1,2,3] | Catarina Alves[4] | Mafalda Santiago[5] | Paula Ministro[6] | Paula Lago[7] | Luís Correia[8] | Raquel Gonçalves[9] | Diana Carvalho[10] | Francisco Portela[11] | Cláudia Camila Dias[5,12] | Axel Dignass[13] | Silvio Danese[14,15] | Laurent Peyrin-Biroulet[16] | Maria Manuela Estevinho[2,17] | Paula Moreira[3] | on behalf GEDII (Portuguese IBD Group)

[1]Department of Gastroenterology, São João University Hospital Center (CHUSJ), Porto, Portugal

[2]Department of Biomedicine, Unit of Pharmacology and Therapeutics, Faculty of Medicine, University of Porto, Porto, Portugal

[3]Department of Clinical Pharmacology, São João University Hospital Center (CHUSJ), Porto, Portugal

[4]Faculty of Medicine, University of Porto, Porto, Portugal

[5]Center for Health Technology and Services Research (CINTESIS), Porto, Portugal

[6]Department of Gastroenterology, Tondela-Viseu Hospital Centre, Viseu, Portugal

[7]Department of Gastroenterology, Santo António University Hospital Center (CHUPorto), Porto, Portugal

[8]Department of Gastroenterology, Lisbon North Hospital Centre, Santa Maria Hospital, Lisbon, Portugal

[9]Gastroenterology Department, Hospital de Braga, Braga, Portugal

[10]Department of Gastroenterology, Santo António dos Capuchos Hospital at Centro Hospitalar Lisboa Central, Lisbon, Portugal

[11]Department of Gastroenterology, University Hospital Centre of Coimbra, Coimbra, Portugal

[12]Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal

Abstract

Background: Ulcerative colitis (UC) has been the focus of numerous observational studies over the years and a common strategy employed in their design is the use of composite and aggregate outcomes.

Objective: This systematic review and meta-analysis aims to identify composite and aggregate outcomes of observational studies in UC and to evaluate how the number and type of variables included and the length of follow-up affect the frequency of patients that achieve these outcomes.

Methods: A systematic literature search was carried out using MEDLINE [via PubMed], Scopus, and Web of Science online databases. Observational studies that included UC patients and reported composite or aggregate outcomes were identified. A set of variables considered to be representative of progressive or disabling UC was defined, the proportion of patients attaining the outcomes was determined and a random-effects meta-analysis was performed by dividing the identified studies into subgroups according to different criteria of interest.

Results: A total of 10,264 records were identified in the systematic search, of which 33 were retained for qualitative analysis and 20 were included in the meta-analysis. The mean frequency for composite outcomes was 0.363 [95% confidence interval (CI) 0.323-0.403]. The frequency of composite outcome for the subgroup of studies that included the variable "Biologics" was significantly higher than for those in which this variable was not reported [0.410; 95% CI 0.364-0.457 versus 0.298; 95% CI 0.232-0.364; p = 0.006]. Composite outcomes were also more frequent as the follow-up duration increased.

Conclusion: The frequency of composite outcomes in observational studies of UC is dependent on the specific identity of the variables being reported. Moreover, longer

Fernando Magro and Catarina Alves shared co-first authorship

[13]Agaplesion Markus Hospital, Department of Medicine I, Frankfurt, Germany

[14]Department of Biomedical Sciences, Humanitas University, Milan, Italy

[15]Inflammatory Bowel Disease (IBD) Center, Department of Gastroenterology, Humanitas Clinical and Research Center (IRCCS), Milan, Italy

[16]Department of Gastroenterology and Inserm NGERE U1256, University Hospital of Nancy, University of Lorraine, Vandoeuvre-lès-Nancy, France

[17]Department of Gastroenterology, Centro Hospitalar Vila Nova de Gaia/Espinho, Vila Nova de Gaia, Portugal

Correspondence
Fernando Magro, Unit of Pharmacology and Therapeutics, Department of Biomedicine, Faculty of Medicine, University of Porto, Alameda Prof. Hernâni Monteiro, 4200-319 Porto, Portugal.
Email: fm@med.up.pt

Funding information
GEDII (Portuguese Group of Studies in Inflammatory Bowel Disease)

follow-up periods are associated with higher frequencies of composite outcomes. The evidence provided here is useful for the design of future observational studies of UC that aim to maximize the frequency of patients that achieve composite outcomes.

KEYWORDS
aggregate outcomes, composite outcomes, meta-analysis, observational studies, systematic review, ulcerative colitis

Key summary

1. Summarize the established knowledge on this subject
   - Observational studies have become a valuable source of information but also present a remarkable heterogeneity.
   - The inconsistency on the reported outcomes between individual studies and the potential for reporting biases has led to calls for the development of core outcome sets (COS).
2. What are the significant and/or new findings of this study?
   - This is the first systematic review and meta-analysis of outcomes in observational studies of Ulcerative colitis (UC).
   - The frequency of composite outcomes in observational studies of UC is dependent on the specific identity of the variables being reported and on the follow-up duration.
   - Reporting of the variable “Biologics” significantly increased the frequency of composite outcomes.
   - These findings may be useful for the design of future observational studies of UC.

INTRODUCTION

Ulcerative colitis is a chronic inflammatory disease restricted to the colon, with onset usually occurring in early adulthood and a high chance of relapse episodes during the lifetime of the patient.[1] Along with Crohn's Disease (CD), it constitutes the main component of Inflammatory Bowel Disease (IBD). Treatment targets for UC have become more diversified in recent years, incorporating objective measures of inflammation such as endoscopic procedures, histology and biomarkers.[2]

Given the considerable costs and complex logistics associated with Randomized Clinical Trials (RCTs), observational studies have become a valuable source of information in the development of novel therapeutical approaches for chronic diseases.[3] However, because they are not subjected to the strict methodological regulations that govern RCTs, observational studies present a remarkable heterogeneity in terms of their basic design, number of patients enrolled, duration of the monitoring and number and type of endpoint variables being reported. In particular, the inconsistency on the reported outcomes between individual studies and the potential for reporting biases has led to calls for the development of COS to be included in all studies related to a specific clinical area.[4]

The present systematic review and subsequent meta-analysis were performed with the aim of evaluating composite and aggregate outcomes reported in observational studies on UC. These outcomes are particularly appropriate to maximize statistical power and therefore compensate for a potentially small patient population in a given study.[5] We focused specifically in estimating the frequency of patients achieving composite and aggregate outcomes and in determining if and how this value was affected by the number and type of variables included in the study, as well as by its total duration.

MATERIALS AND METHODS

Search strategy

This study was conducted following the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Guidelines[6] and the Cochrane Collaboration Guidelines for reporting meta-analyses.[7] Literature search was performed from inception to 14 July 2020 using three electronic databases: MEDLINE (via PubMed, https://pubmed.ncbi.nlm.nih.gov), Scopus (https://www.elsevier.com/solutions/scopus) and Web of Science (http://www.isiwebofknowledge). The words or medical subject heading terms used were: “((('aggressive disease') OR ('disabling disease') OR ('disabling outcome') OR ('disabling outcomes') OR ('composite outcome') OR ('composite outcomes') OR ('composite event') OR ('composite events') OR ('composite endpoint') OR ('composite endpoints') OR ('composite') OR (composit*) OR ('progressive disease')) AND (('Colitis, Ulcerative' [MeSH Terms]) OR (UC) OR ('Crohn Disease' [MeSH Terms]) OR (crohn's disease) OR ('Inflammatory Bowel Diseases' [MeSH Terms])))." This query was used for PubMed search and adjusted for the other databases used. To ensure that all relevant articles were included, the reference lists of the systematic reviews selected from the databases were manually reviewed. This study's protocol was not registered in the PROSPERO database.

Eligibility criteria

Studies enrolling both adults and children previously diagnosed with UC using clinical, endoscopic, and/or pathological features were considered eligible for inclusion in this systematic review. The inclusion criteria were: [i] cohort, case-control, and cross-sectional studies with UC patients; [ii] studies evaluating composite or aggregate outcomes; and [iii] outcomes representing UC progression. There were no publication year restrictions but only articles written in English were included. The exclusion criteria were: [i] randomized controlled trials and *post-hoc* analysis, systematic reviews and meta-analysis, review articles, descriptive and diagnostic studies, animal and in vitro studies, study protocols, guidelines, editorials, and only abstracts available; [ii] studies selecting patients with diseases other than UC; studies evaluating only CD patients [iii] studies that did not define a composite or aggregate outcome of interest; [iv] studies reporting an improvement outcome; [v] studies that did not differentiate between UC and CD patients.

Study selection and data collection process

Two reviewers (CA and MS) independently screened the titles and abstracts according to the eligibility criteria and if a particular study failed to meet these criteria, it was excluded. In the second phase, the full text of all the remaining potentially relevant studies was analyzed, and the eligibility criteria were used again to discard non-relevant studies. Disagreement was resolved via consensus between the two reviewers. The following information was collected from the selected studies: authors' names; publication year; country of origin; study design; observation period; number of patients (discriminated between UC and CD, if applicable); UC extent; subgroups (if applicable); outcome name, definition and reported variables (among the ones chosen for analysis). The number of patients achieving the defined outcome and its corresponding proportion was calculated if not explicitly stated. The proportion of patients achieving each variable of the composite outcome was not assessed. The observation period refers to the mean or median time of follow-up or the time of

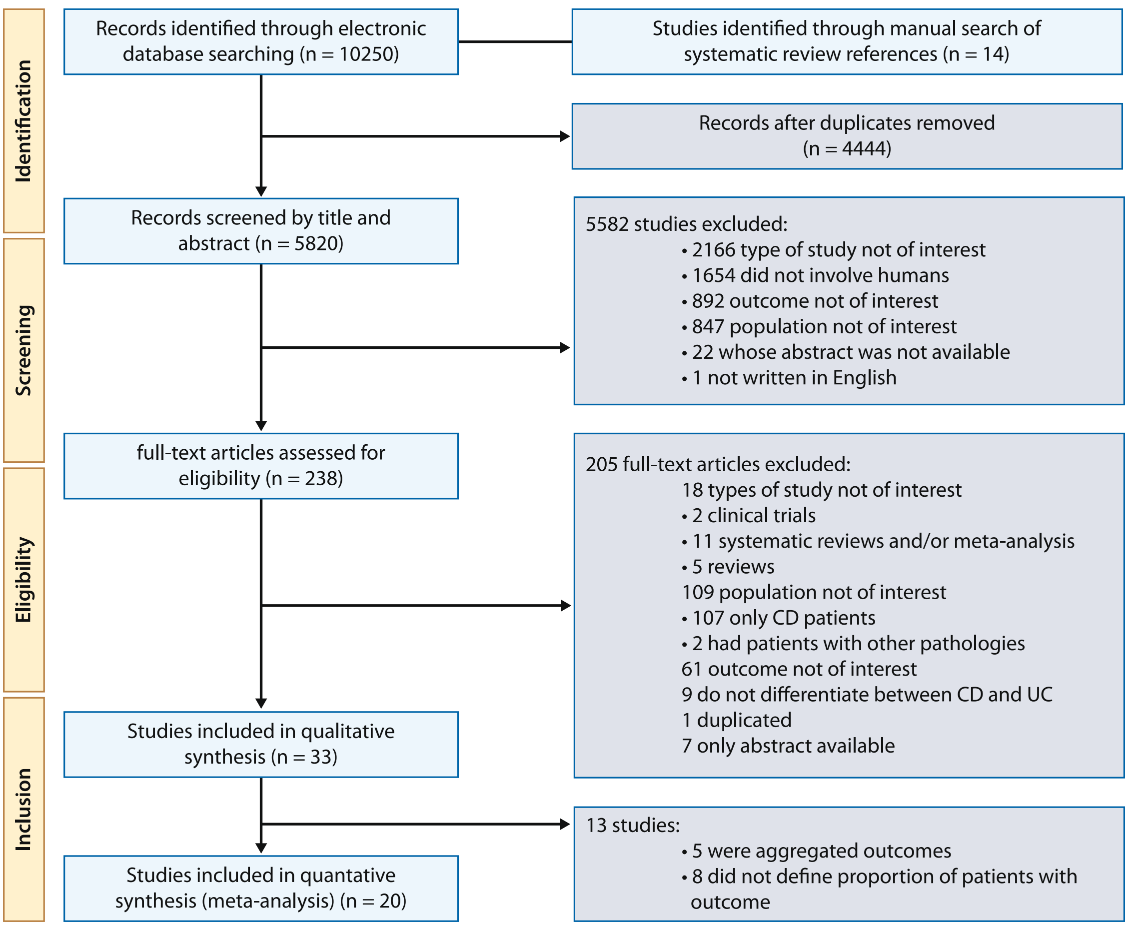

FIGURE 1 Flow diagram of study selection and data collection process

TABLE 1 Characteristics of the studies included in the systematic review

Study	Country	Study design	Observation period	Number of patients (N)	Disease extension	Subgroups	Outcome (C/A)	Outcome definition	Variables[a]	No of patients with outcome/N	%[b]
Acosta *et al.*, 2016	Spain	Unicentric prospective cohort	12 months	187	E1[e]: 28.8% E2[e]: 41.2% E3[e]: 30%	NA	Clinical relapse (A)	Rectal bleeding and any of: the need for any remission induction treatment; any treatment escalation; and the need for hospitalization or colectomy	Hospitalization Surgery Steroids Immunosuppressors Biologics Clinical assessment Therapy modification	49/187	26%
Ahmad *et al.*, 2019	USA	Unicentric retrospective cohort	30 days	268 CD[c]: 154 UC: 114	CSE, GPPs Pancolitis: 51%, 60% Left-sided: 31%, 27% Proctitis: 18%, 13%	GPPs CSE	Disease course (C)	IBD-related hospitalization, surgery, or emergency department visit	Hospitalization Surgery Steroids Immunosuppressors Biologics	19/114	17%
			NA				Treatment modification (C)	Change in corticosteroids, immunomodulators and/or biologic use		51/114	45%
Ananthakrishnan *et al.*, 2017	USA	Multicentric prospective cohort	365 ± 45 days[f]	871, adults CD[c]: 707 UC: 164	E[e]: CT, MT E1[e]: 5%, 10% E2[e]: 27%, 31% E3[e]: 68%, 59%	Combination therapy (anti-TNF and immunomodulator) Monotherapy (anti-TNF)	Complications (C)	New IBD-related surgery, hospitalizations, penetrating complications (new abscess or fistula), need for corticosteroids or new biologic	Hospitalization Surgery Steroids Biologics Clinical assessment	NA	NA
Ardizzone et al., 2011	Italy	Unicentric prospective cohort	8 (1–59) months[g]	157	Proctosigmoiditis: 15.3% Left-sided colitis: 34.4% Extensive colitis: 50.3%	NA	Aggressive clinical course (C)	Hospitalizations and/or immunosuppressive treatment (IV cyclosporin A and/or oral azathioprine/ mercaptopurine) and/or colectomy (proctocolectomy and ileal pouch–anal anastomosis)	Hospitalization Surgery Immunosuppressors	74/157	47%
Biasci *et al.*, 2019	UK	Unicentric prospective cohort	IBD1: 5.6 (3.6–7.1) years[g] IBD2: 5.5 (2.4–8.4) years[g]	118, adults CD[c]: 66 UC: 51	E[e]: IBD1, IBD2 E1[e]: 20.8%, 28.6% E2[e]: 37.5%, 39.3% E3[e]: 41.7%, 32.1%	IBD1 IBD2	Aggressive disease (C)	Treatment escalation with immunomodulators, biological therapies or surgery	Surgery Immunosuppressors Biologics	27/52	52%
Bitton *et al.*, 2001	USA and Canada	Multicentric prospective cohort	12 months (or until relapse)	74	Proctitis: 7% Proctosigmoiditis: 36%	NA	Clinical relapse (A)	Worsening of bowel function and rectal bleeding with endoscopic riley > 1	Clinical assessment Endoscopic assessment	27/74	36%

(Continues)

TABLE 1 (Continued)

Study	Country	Study design	Observation period	Number of patients (N)	Disease extension	Subgroups	Outcome (C/A)	Outcome definition	Variables[a]	No of patients with outcome/N	%[b]
					Left-sided colitis: 34% Pancolitis: 23%						
Buda *et al*., 2014	Italy	Unicentric prospective cohort	12 months	19	NA	NA	Relapse (A)	CAI ≥ 3 and abnormal mucosa at endoscopy	Clinical assessment Endoscopic assessment	7/19	37%
Carvalho *et al*., 2015	Portugal	Unicentric retrospective cohort	12 months	138	E1[e]: 46.4% E2[e]: 27.5% E3[e]: 26.1%	NA	Clinical relapse (C)	Intensification or modification of medication (steroids, immunosuppressors, biologics and others non specified), UC-related hospitalization or surgery	Hospitalization Surgery Steroids Immunosuppressors Biologics Therapy modification	28/138	20%
Christensen *et al*., 2017	USA	Unicentric retrospective case-control	22 (14–34) months[g]	310	E1[e]: 11.0% E2[e]: 28.7% E3[e]: 60.3%	Histologic normalization No histologic normalization	Clinical relapse (C)	SCCAI > 2, subscore > 1 for stool frequency or rectal bleeding, or medication escalation for symptoms, hospitalization for UC relapse, or colectomy for refractory UC	Hospitalization Surgery Clinical assessment Therapy modification	77/310	25%
Cushing *et al*., 2020	USA	Unicentric prospective cohort	2 years	83	Pancolitis: 55%	NA	Clinical relapse (C)	Change in UC therapy (i.e., dose escalation, class change, and/or need for systemic corticosteroids), UC-related hospitalization, or UC-related surgery	Hospitalization Surgery Steroids Immunosuppressors Biologics Therapy modification	26/83	31%
Fabian *et al*., 2019	Czech republic	Unicentric retrospective cohort	12 months (minimum)	41	NA	NA	Complicated disease course (C)	Acute severe colitis (PUCAI > 65), necessity of colectomy or infliximab initiation	Surgery Biologics Clinical assessment	13/41	32%
Frieri *et al*., 2017	Italy	Unicentric prospective cohort	36 months	52	Pancolitis: 23.1% Left colitis: 69.2% Proctitis: 7.7%	NA	Clinical relapse (C)	Need of steroids, immunomodulators and/or biological drugs	Steroids Immunosuppressors Biologics	NA	NA
Fries *et al*., 2017	Italy	Multicentric retrospective cohort	3 years	1091	E[e]: Elderly (>65), adults (40–64), young (<40)	NA	Clinical relapse (C)	Need for more intensive medical therapy or surgery	Surgery	411/1091	38%

TABLE 1 (Continued)

Study	Country	Study design	Observation period	Number of patients (N)	Disease extension	Subgroups	Outcome (C/A)	Outcome definition	Variables[a]	No of patients with outcome/N	%[b]
					E1[e]: 13%, 11%, 12% E2[e]: 56%, 60%, 52% E3[e]: 31%, 29%, 36%				Therapy modification		
Fuentes *et al.*, 2017	The Netherlands	Unicentric prospective cohort	35 (21.5–37) months[g]	12	E2[e]: 50% E3[e]: 50%	NA	Clinical relapse (C)	SCCAI ⩾5 and/or need for rescue therapy	Clinical assessment Therapy modification	6/12	50%
Gubatan *et al.*, 2017	USA	Unicentric prospective cohort	12 months	70	Left-sided colitis: 38.6% Extenstive colitis: 52.9%	NA	Clinical relapse (C)	SCCAI > 2, medication intensification (increase in dose of current regimen, addition of another medication, or change in class of medication due to symptom relapse), or UC-related hospitalization	Hospitalization Clinical assessment Therapy modification	20/70	29%
Höie *et al.*, 2007	Several European countries	Multicentric prospective cohort	10 years	771	NA	NA	Relapse (C)	Dose increase of ongoing medical treatment or introduction of new medication, or surgery	Surgery Therapy modification	NA	NA
Jairath *et al.*, 2019	Canada	Multicentric prospective cohort	DC: 32 weeks VC: 6 weeks	DC: 179 VC: 146	NA	NA	Composite UC-100 score (A)	(1 + 16 × Mayo clinic stool frequency subscore [0 to 3] + 6 × Mayo clinic endoscopic subscore [0 to 3] + 1 × Robarts histopathology index score [0 to 33])	Clinical assessment Endoscopic assessment	NA	NA
Kochar *et al.*, 2018	USA	Multicentric prospective cohort	24 ± 10 months[f]	4314, adults CD[c]: 2798 UC: 1516	E1[e]: 11% E2[e]: 30% E3[e]: 59%	Depression No depression	Aggressive disease (C)	Surgery or hospitalization	Hospitalization Surgery	NA	NA
Kurnool *et al.*, 2018[d]	USA	Unicentric retrospective cohort	2 years[g]	160	Pancolitis: 61%	NA	Treatment failure (C)	IBD-related surgery, hospitalization, or treatment modification (including index biologic dose escalation, drug discontinuation, or addition/continuation of corticosteroids ≥3 months after starting index biologic therapy)	Hospitalization Surgery Steroids Biologics Therapy modification	NA	NA
Lee *et al.*, 2011	UK	Unicentric prospective cohort	IBD1: 575 (435–685) days[g] IBD2: 562 (475–643) days[g]	67 CD[c]: 35 UC: 32	E[e]: IBD1, IBD2 E1[e]: 23%, 16% E2[e]: 31%, 47% E3[e]: 46%, 37%	IBD1 IBD2	Disease course (C)	Treatment escalation with immunomodulators, biologics or surgery	Surgery Immunosuppressors Biologics	16/32	50%

(Continues)

TABLE 1 (Continued)

Study	Country	Study design	Observation period	Number of patients (N)	Disease extension	Subgroups	Outcome (C/A)	Outcome definition	Variables[a]	No of patients with outcome/N	%[b]
Lobatón *et al.*, 2017	Belgium and Spain	Multicentric prospective cohort	12 months	96	E1[e]: 15% E2[e]: 47% E3[e]: 38%	NA	Clinical relapse (C)	Clinical Mayo partial score ≥3 and/or need of introducing steroids or any other treatment escalation	Steroids Clinical assessment Therapy modification	22/96	23%
Lund *et al.*, 2011	USA	Multicentric retrospective cohort	NA	610	NA	NA	Flare (C)	Oral steroids or infliximab (de novo prescription), or oral/rectal salicylates, or need for colectomy	Surgery Steroids Biologics	NA	NA
Magro *et al.*, 2018	Portugal	Multicentric retrospective cohort	12 (7–19) years[g]	1481 DC: 1210 VC: 271	E[e]: DC, VC E1[e]: 41%, 44% E2[e]: 22%, 19% E3[e]: 36%, 37%	NA	Progressive disease (C)	One or more of the following events: Fibrous or mucosal bridges, stenosis, pseudopolyps, lead pipe or shortening or haustral markings, colectomy, 2 or more hospitalizations, 2 or more corticosteroid course requirements/year, corticosteroid dependency or refractoriness, need to switch the immunosuppressive drug (AZA or MTX) or the anti-TNF drug [IFX or ADA] or add immunosuppressive drugs to anti-TNF treated patients.	Hospitalization Surgery Steroids Immunosuppressors Biologics Endoscopic assessment	445/1210	37%
Meyer *et al.*, 2019	France	Multicentric equivalence cohort	RP: 423 (189-757) days[g] CT-P13: 286 (168-466) days[g]	3112, adults RP: 1434 CT-P13: 1678	NA	CT-P13 RP	Effectiveness (C)	Death, UC-related surgery (colectomy and/or rectal resection and intestinal stoma), all-cause hospitalization (except childbirth) or reimbursement of another anti-TNF (adalimumab/golimumab) or vedolizumab	Hospitalization Surgery Biologics	1453/3112	47%
Nguyen *et al.*, 2020	USA	Unicentric retrospective cohort	24 (14.8–34.7) months[g]	160	Extensive colitis: 61.3% Left-sided colitis: 38.1% Proctitis: 0.6%	NA	Treatment failure (C)	IBD-related surgery, hospitalization, or treatment modification (including index biologic dose escalation, drug discontinuation, or addition/continuation of corticosteroids ≥3 months after starting index biologic therapy)	Hospitalization Surgery Steroids Biologics Therapy modification	110/160	69%
Niewiadomski *et al.*, 2015	Australia	Multicentric retrospective and prospective cohort	18 (12–82) months[g]	252 CD[c]: 146 UC: 96 Undifferentiated: 10	Proctitis: 32% Left-sided colitis: 31% Pancolitis: 36%	NA	Disabling UC (C)	More than two courses of steroids, further hospitalization after diagnosis, ongoing active disease, intestinal resection.	Hospitalization Surgery Steroids Clinical assessment	24/96	25%

TABLE 1 (Continued)

Study	Country	Study design	Observation period	Number of patients (N)	Disease extension	Subgroups	Outcome (C/A)	Outcome definition	Variables[a]	No of patients with outcome/N	%[b]
Ozaki *et al.*, 2018	Japan	Unicentric retrospective cohort	608.6 days[f]	194	Proctitis: 19.1% Left-sided colitis: 33.5% Pancolitis: 44.8% Segmental colitis: 2.1% Unknown: 0.5%	NA	Clinical relapse (C)	Partial Mayo Score ≥ 3 or modification of induction treatment [dose escalation, addition of steroids, tacrolimus, topical formulations or biologics]	Surgery Immunosuppressors Biologics Clinical assessment Therapy modification	67/194	35%
Pereira *et al.*, 2020	Portugal and Belgium	Multicentric retrospective cohort	17 (11–25) years[g]	931	Proctitis: 21.4% Left-sided: 30.7% Pancolitis: 43.6% Missing: 4.3%	Discovery cohort Validation cohort	Complicated disease (C)	Need hospitalization, corticodependent or corticoresistant or no responder to immunossupressors or need biologics or colectomy	Hospitalization Surgery Steroids Biologics	NA	NA
Riley *et al.*,1991	UK	Unicentric prospective cohort	12 months	82	Proctitis: 23% Proctosigmoiditis: 34% Left-sided colitis: 22% Pancolitis: 21%	NA	Relapse (A)	Symptomatic deterioration with hemorrhagic in endoscopy	Clinical assessment Endoscopic assessment	27/82	33%
Stallmach *et al.*, 2014	German	Multicentric retrospective cohort	60 months	262	Proctitis: 16.4% Proctosigmoiditis: 17.9% Left-sided colitis: 19.1% Pancolitis: 43.5% Unknown: 3.0%	NA	Progressive disease (C)	Need for immunosuppressive therapy: thiopurines or methotrexate or anti-TNF-α antibodies or cyclosporine A or tacrolimus	Immunosuppressors Biologics	104/262	40%
Thomsen *et al.*, 2020	Denmark	Multicentric retrospective cohort	Thiopurines and allopurinol: 83 person-years Thiopurines: 12,388 person-years	10,367, adults CD[c]: 5484 UC: 4883	NA	Thiopurines and allopurinol Thiopurines alone	Clinical relapse (C)	Need for anti-TNFα treatment or IBD-related hospitalization, IBD-related major surgery or death	Hospitalization Surgery Biologics	1960/4883	40%
Ungaro *et al.*, 2019	USA	Multicentric retrospective cohort	155 to 287 days[h]	3589	NA	NA	Adverse clinical events (C)	New corticosteroid use, UC-related hospitalization or surgery	Hospitalization Surgery Steroids	NA	NA

(Continues)

TABLE 1 (Continued)

Study	Country	Study design	Observation period	Number of patients (N)	Disease extension	Subgroups	Outcome (C/A)	Outcome definition	Variables[a]	No of patients with outcome/N	%[b]
Yamamoto *et al*., 2018	Japan	Unicentric prospective cohort	12 months	164	Left-sided colitis: 76% Extensive colitis: 24%	NA	Clinical relapse (C)	Worsening of stool frequency and/or RB with the MES > 1	Clinical assessment Endoscopic assessment	46/164	28%

Note: When a scientific article included more than one cohort, characteristics were discriminated by cohort.

Abbreviations: ADA, adalimumab; AZA, azathioprine; C/A, composite/aggregate outcome; CAI, Clinical Activity Index; CD, Crohn's Disease; CSE, conventional stool evaluation; CT, Combination therapy; CT-P13, Biosimilar infliximab; DC, Derivation Cohort; GPPs, Gastrointestinal pathogen panels; IBD, Inflammatory Bowel Disease; IFX, infliximab; IV, intravenous; MES, Mayo Endoscopic Subscore; MT, Monotherapy; MTX, methotrexate; NA, Not Applicable; PUCAI, Pediatric Ulcerative Colitis Activity Index; RB, rectal bleeding; RP, reference product (Infliximab); SCCAI, Simple Clinical Colitis Activity Index; UC, ulcerative colitis; VC, Validation Cohort.

[a]Variables among the ones analyzed (Hospitalization, Surgery, Steroids, Immunosuppressors, Biologics, Clinical Assessment, Endoscopic Assessment, Therapy Modification).

[b]Only the composite (C) outcomes with % available were included in the meta-analysis. The remaining took part only in the qualitative synthesis.

[c]CD patients were not included in the analysis.

[d]This study refers to the same population as Nguyen et al., 2020 and for that reason it was not included in the meta-analysis.

[e]Montreal Classification.

[f]Mean.

[g]Median.

[h]For more precise values consult the original article.

occurrence of the outcome, when available. The variables considered were selected as the most clinically relevant parameters in IBD assessment and approach, but did not necessarily include all the aspects referred to in the reported outcomes.

Endpoints under analysis

A composite outcome was defined as an outcome composed by two or more variables and to achieve the outcome the patient needed to present at least one variable.[8] An aggregate outcome was defined as the simultaneous presence of all the parameters considered. The outcomes represented disease progression and included the following variables: hospitalization, surgery, steroids, immunosuppressors, biologics, clinical assessment, endoscopic assessment and therapy modification. Hospitalization was defined as one or more inpatient stays for any UC-related cause. Surgery was defined as at least one surgery for any UC-related cause. Steroids was defined as reported de novo use, dose increase, change, dependency or refractoriness to corticosteroids. Immunosuppressors was defined as reported de novo use, switch, dose increase or unspecified immunosuppressive therapy. Biologics was defined as reported as de novo use, switch, dose increase or treatment frequency increase of any therapeutic agent targeting tumor necrosis factor-α or other pro-inflammatory mediators. Clinical Assessment was defined as reported UC clinical symptoms or manifestations, extra intestinal manifestations, imagiological disease activity evaluation or UC clinical scores modification towards worsening disease. Endoscopic Assessment was defined as reported endoscopic scores or any other endoscopic evaluation. Therapy modification was defined as medication adjustments for UC-related symptoms or increase in UC activity when the drug was not identified.

Quality assessment

The methodological quality was assessed by using the validated Critical Appraisal Skills Programme checklist,[9] which systematically assesses the validity, results, and relevance for each study included in the analysis. Each item of the checklist was evaluated using a color scheme: [i] green if the study met all the parameters included in each item; [ii] yellow if the study met the parameters partially or if it did not have enough information; [iii] red if the study did not meet the parameters included in each item.

Statistical analysis

The main data analyzed in this meta-analysis were the proportions of patients achieving a composite outcome. The proportion of patients achieving the outcome was calculated and compared according to the predefined variables reported in the study. The following comparisons between subgroups were performed:

frequency of composite outcome according to [i] total number of variables; [ii] presence of each predefined variable; [iii] presence of a combination of two different predefined variables; [iv] total length of study follow-up; [v] number of patients included in the studies. In addition, the entire population of outcomes was divided in subgroups reporting a specific predefined “core” variable, and differences within these subgroups according to the presence or absence of the remaining predefined variables were statistically tested. Comparisons were not performed when the subgroups defined by the presence or absence of the considered variable were composed by only one or two outcomes.

To perform the meta-analysis, the “metaprop” function from the “meta” package of the R statistical programming language was used and the “PRAW” summary measure was employed for the pooling of studies. A random-effects model was adopted taking into consideration the differences observed across studies.

Cochran's Q test and the I^2 statistic were used to assess statistical heterogeneity.[10] In addition, Egger's test was used to detect potential publication bias.[11] A sensitivity analysis was performed to assess the influence of any individual study on the overall results. A Venn diagram and Upset plot were generated using the “UpsetR” and “nVennR” packages included within the R software, to graphically illustrate the distribution of the predefined variables among the individual studies included in the meta-analysis.

All analyses and charts were executed using R software version 4.1.0 and a p-value lower than 0.05 was considered statistically significant.

RESULTS

Literature search and study selection

The selection strategy followed is summarized in Figure 1. The electronic database search yielded 10,250 records (1885 in PubMed, 4323 in Scopus, and 4042 in Web of Science); the manual search identified 14 additional studies. Following the removal of duplicates (n = 4444), 5820 records remained, of which 5582 were excluded on the screening phase. The remaining 238 records were evaluated for eligibility. Following full-text assessment, 205 articles were excluded, and the remaining 33 articles were selected for the qualitative analysis. Twenty of those 33 were included in the meta-analysis (Figure 1).

Quality assessment

The results of the methodological quality assessment are summarized in Supplementary Table 1. All the selected studies clearly stated the issue being evaluated, but nearly all of them exhibited some problem related to the recruitment of the patient cohorts. In addition, some of the studies failed to take confounding factors into consideration and adapt their design and/or analysis accordingly.[12–17] The follow-up was considered complete enough and of suitable duration in most of the cases, and the results were deemed to be believable in general.

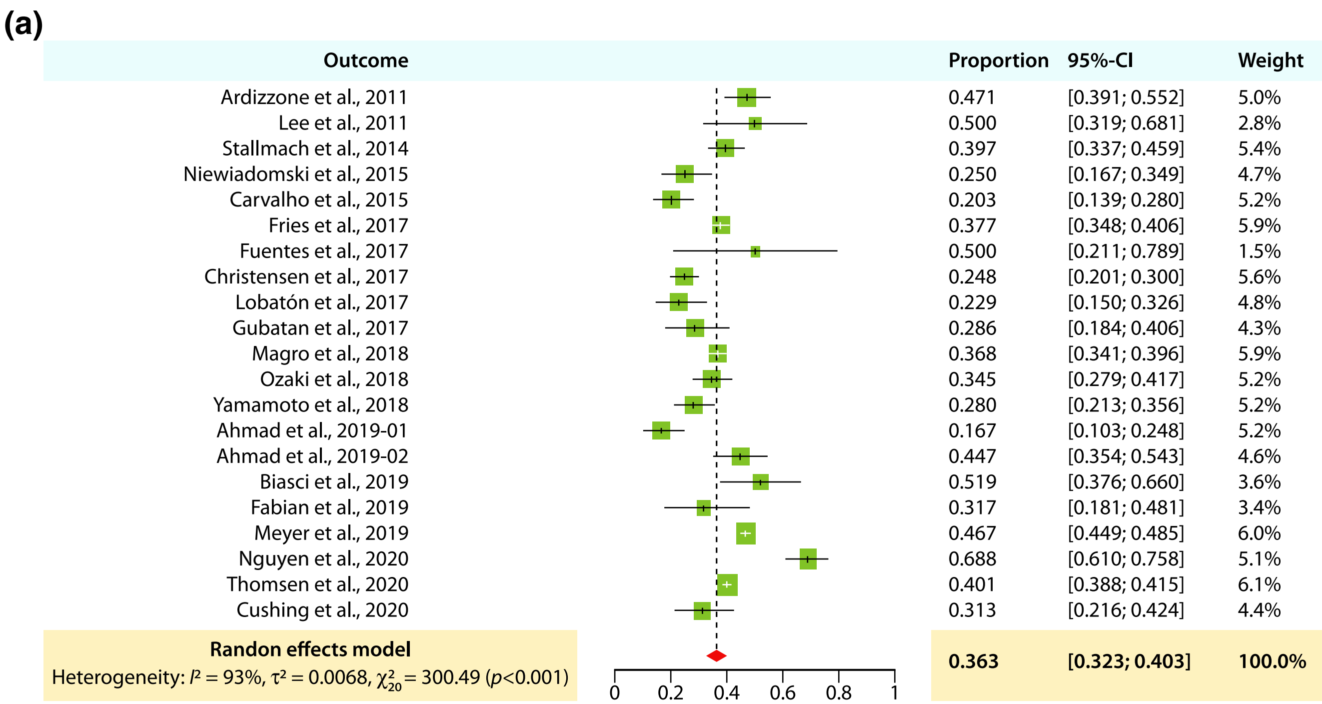

FIGURE 2 (a) Frequency of composite outcomes, n = 21 (b) Frequency of composite outcomes according to the number of variables reported in the study. n = 21

Characteristics of included studies

Study characteristics are summarized in Table 1. Of all the articles selected for this systematic review, 26 included only UC patients[12,14,16,18–40] while the remaining seven[13,15,17,41–44] also included CD patients. Both the number of patients and the observation period of individual studies showed wide variations, ranging from twelve[38] to 10,367 patients[44] and from 30 days[13] to a median of 17 years,[33] respectively. One of the selected articles[13] included two cohorts of patients with the outcomes discriminated by cohort. In this case, each cohort was considered as an independent outcome for the purposes of the analysis. Twenty-nine composite outcomes were registered from a total of 28 studies[12–15,17–23,27–30,32–44] and five aggregate outcomes from five studies.[16,24–26,31] The outcomes where heterogeneous in terms of the reporting of predefined variables considered here: the variable Hospitalization was reported in 17 outcomes,[13,15,17,18,20,23,26,30,32,33,35,36,39,41,44] Surgery in 22 outcomes,[13–15,17,20–23,26–30,32,33,35,36,39,41–44] Steroids in 15 outcomes,[13,15,19,20,22,23,26,28–30,32,33,35,40,41] Immunosuppressors in 11 outcomes,[13,19,20,22,26,29,34–36,42,43] Biologics in 18 outcomes,[13,14,19–23,26,28–30,33–35,41–44] Clinical Assessment in 14 outcomes,[12,14–16,18,19,24–26,31,38–41] Endoscopic Assessment in 6 outcomes[11,15,23,35,39,40] and Therapy Modification in 12 outcomes[18,19,22,23,26,27,29,30,35,38–40](Supplementary Figure 1). The total

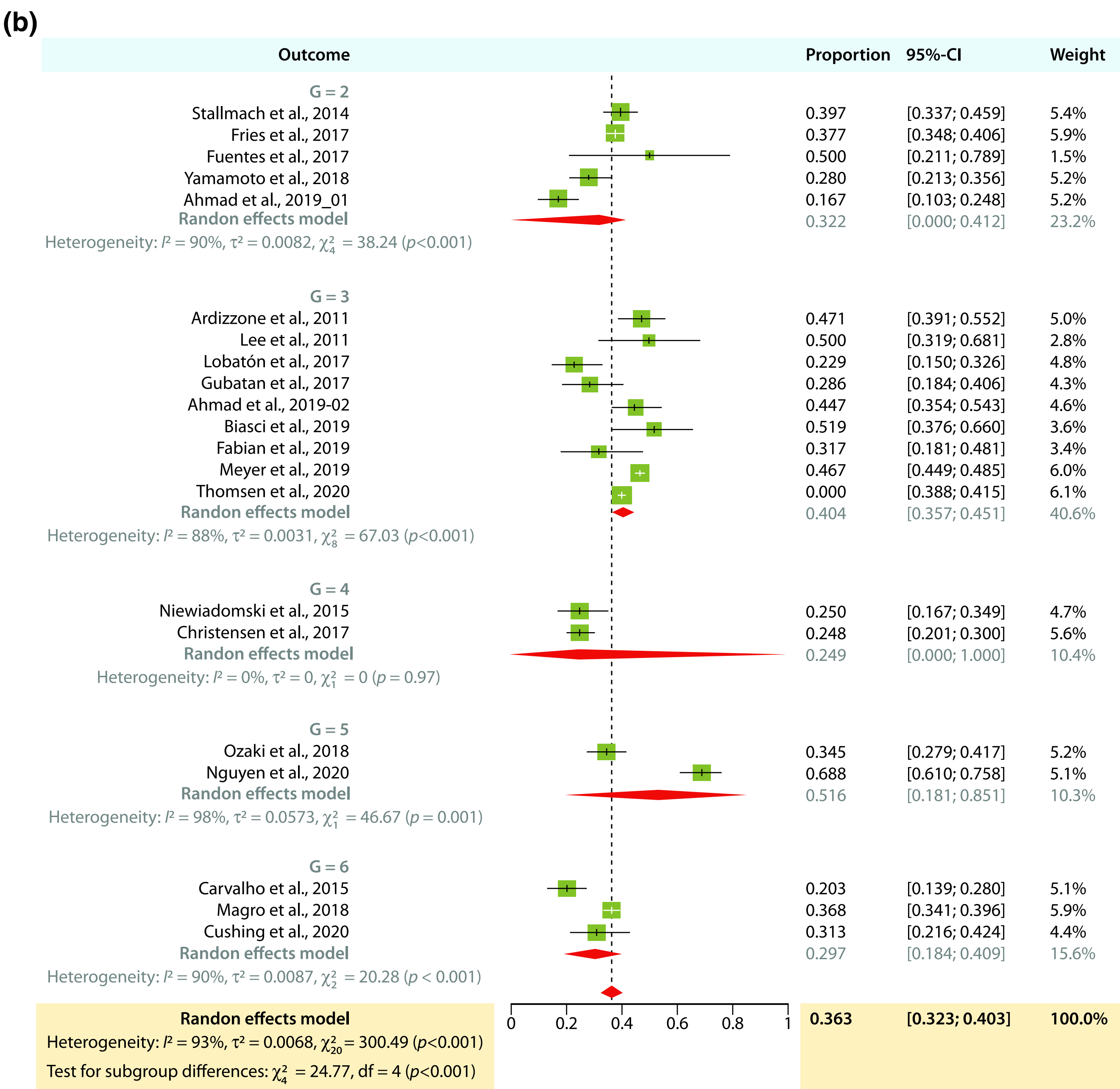

FIGURE 2 (Continued)

number of predefined variables that were reported ranged from 2 in 11 different outcomes[12,13,16,17,24,25,27,31,34,37,38] to seven in a single outcome[26] (Supplementary Figure 2).

Composite and aggregate outcomes

Twenty one composite outcomes[12–15,18–23,34–40,42–44] were included in the meta-analysis. The mean frequency for composite outcomes was 0.363 [95% confidence interval (CI) 0.323-0.403] (Figure 2a).

The funnel plot and sensitivity analysis are depicted in Supplementary Figures 3 and 4. The results of Egger's test on the frequency of composite outcomes were not significant (p = 0.166), indicating that the dataset was unbiased. This can also be visually appreciated by the symmetry of the corresponding funnel plot in which the standard error was plotted against the outcome frequency. No outliers were detected on the results of the sensitivity analysis, with the mean being unaffected by the sequential exclusion of each individual outcome (Supplementary Figure 4).

Subgroup analysis

Significant differences between subgroups were observed when the frequency of composite outcomes was discriminated according to the total number of predefined variables reported (Figure 2b). However, a post-hoc test was performed and although it presented only a statistically significant result, overall, it failed to show any significant trend (Supplementary Table 2).

The presence of specific variables in the study outcome had a significant effect on the frequency of composite outcomes. The frequency corresponding to the subgroup of studies that included the variable "Biologics" was significantly higher than for those in which this variable was not reported (0.410; 95% CI 0.364-0.457 vs. 0.298; 95% CI 0.232-0.364; p = 0.006; Figure 3a). On the other hand, the studies that included the variable "Clinical Assessment" exhibited a significantly lower frequency of composite outcomes compared to the subgroup where this variable was not present (0.279; 95% CI 0.244-0.314 vs. 0.402; 95% CI 0.357-0.448; p < 0.001; Figure 3b). No significant differences between subgroups were identified when the remaining predefined variables were considered (Supplementary Figure 5).

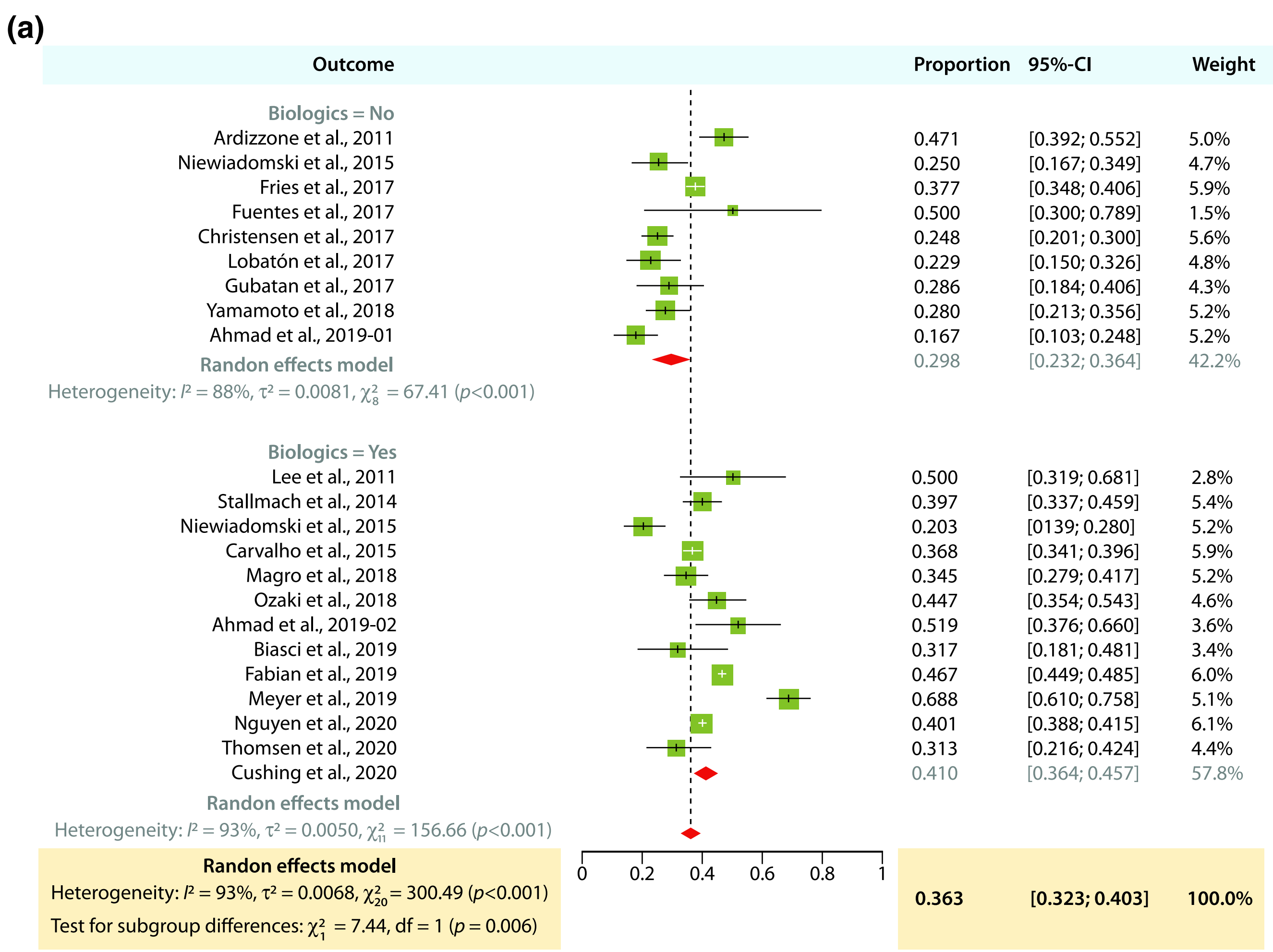

FIGURE 3 Frequency of composite outcomes according to the presence of individual predefined variables reported in the study. (a) Subgroups determined by the presence or absence of the variable "Biologics", n = 21. (b) Subgroups determined by the presence or absence of the variable "Clinical Assessment," n = 21

A further statistical analysis was performed in which the studies were assigned to subgroups based on the reporting of specific variable pairs. The presence of the pair "Biologics" and "Surgery" in the outcome definition was associated with higher frequencies of composite outcome comparing with the outcomes where this variable was not present. The complete set of statistical comparisons using paired variables and their corresponding statistical significance are summarized in Table 2.

To further characterize the impact that the reporting of individual variables has on the frequency of composite outcomes, we tested the effect that this reporting had within subgroups defined by the presence of a specific "core" variable. In the groups defined by the presence of the core variables "Hospitalization," "Surgery" and "Steroids," the reporting of the variable "Biologics" significantly increased the frequency of composite outcomes (Table 3).

When the individual studies were divided into subgroups according to the length of their respective follow-up periods, the frequency of composite outcomes for those in which it was a year long or longer was significantly higher than for those that lasted less than a year (0.398; 95% CI 0.343-0.454 vs. 0.272; 95% CI 0.000-0.359; p = 0.020; Figure 4a). Moreover, the six studies with a follow-up period equal or longer than two years exhibited a significantly higher frequency of composite outcomes compared to the rest (0.464; 95% CI 0.379-0.549 vs. 0.311; 95% CI 0.234-0.387; p = 0.008, Figure 4b).

Considering the number of patients included in the studies, we found that studies with more than 50 patients compared to the ones that include 50 or less patients had a significantly higher frequency of composite outcome achievement (0.416; 95% CI 0.372-0.461 vs. 0.295; 95% CI 0.237-0.353; p = 0.001), respectively (Supplementary Figure 6).

DISCUSSION

Assessment of disease progression in UC patients is particularly challenging because considering clinical symptoms in isolation could result in either underestimation or overestimation of actual disease activity.[45] For this reason, symptom-based scoring assessments have gradually been losing ground to what are considered as more objective measures of inflammation, namely endoscopic and histological evaluation, and the use of biomarkers.[4] In the present

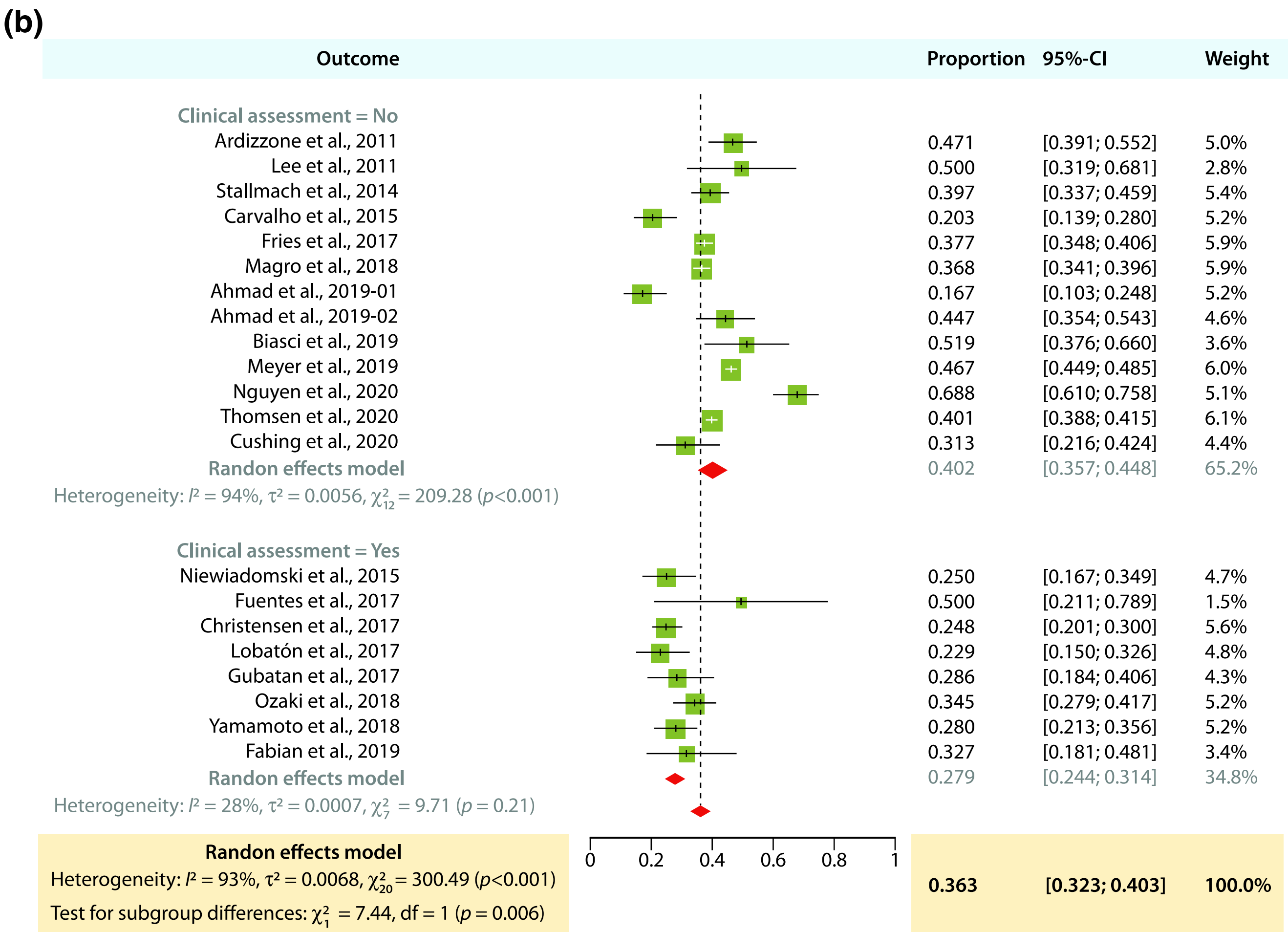

FIGURE 3 (Continued)

systematic review and meta-analysis we attempted to assess the prevalence of composite and aggregate outcomes that combine these different methods of quantifying disease progression, focusing specifically on observational studies reported in the literature.

As perhaps it is to be expected from a dataset composed of observational studies conducted worldwide in the course of a decade, there was a marked heterogeneity between the studies in terms of number of patients enrolled, total duration of the study and variables included in the outcome. Interestingly, the mean frequency of composite outcomes was unaffected by the total number of variables included in the outcome. Because this result seems counterintuitive, it is possible that the relatively small number of studies that reported four or more variables was simply not enough to show a clear statistical trend, and this may become evident as future or ongoing observational studies are incorporated into the available literature.

However, it was clear from the analysis of the present dataset that the reporting of the variable "Biologics" significantly increased the frequency of composite outcomes. This was evident on the overall analysis of the entire dataset, on the analysis by paired variables in combination with the variable "Surgery," and on three out of seven subgroups defined by the presence of a "core" variable. It should be noticed that this increase in the frequency of composite outcomes was not observed with the inclusion of the other predefined variables considered, indicating that merely reporting additional variables regardless of their specific identity does not necessarily increase the chances of observing a composite outcome.

TABLE 2 Subgroup analysis. Frequency of composite outcomes according to the presence of pairs of predefines variables

			Frequency of composite outcome in subgroup		
First variable	**Second variable**	**No of outcomes**	**Both variables reported**	**At least one not reported**	**Significance**
Hospitalization	Surgery	10	0.359 [0.300; 0.419]	0.362 [0.317; 0.407]	$p = 0.940$
	Steroids	5	0.365 [0.000; 0.511]	0.361 [0.319; 0.403]	$p = 0.960$
	Immunosuppressors	4	0.339 [0.000; 0.437]	0.369 [0.323; 0.415]	$p = 0.590$
	Biologics	6	0.409 [0.345; 0.473]	0.341 [0.291; 0.391]	$p = 0.100$
	Clinical assessment	3	0.254 [0.215; 0.293] *	0.380 [0.339; 0.421]	$p < 0.001$
	Endoscopic assessment	1	0.368 [0.341; 0.395]	0.363 [0.319; 0.407]	$p = 0.850$
	Therapy modification	5	0.348 [0.000; 1.000]	0.367 [0.330; 0.404]	$p = 0.840$
Surgery	Steroids	5	0.365 [0.000; 0.511]	0.361 [0.319; 0.403]	$p = 0.960$
	Immunosuppressors	6	0.383 [0.296; 0.469]	0.356 [0.308; 0.405]	$p = 0.600$
	Biologics	9	0.416 [0.360; 0.473] *	0.323 [0.270; 0.376]	$p = 0.020$
	Clinical assessment	3	0.254 [0.000; 0.295] *	0.378 [0.337; 0.419]	$p < 0.001$
	Endoscopic assessment	1	0.368 [0.341; 0.395]	0.363 [0.319; 0.407]	$p = 0.850$
	Therapy modification	5	0.365 [0.000; 0.502]	0.362 [0.322; 0.402]	$p = 0.960$
Steroids	Immunosuppressors	5	0.334 [0.263; 0.405]	0.373 [0.325; 0.421]	$p = 0.380$
	Biologics	6	0.394 [0.000; 0.509]	0.349 [0.304; 0.394]	$p = 0.480$
	Clinical assessment	3	0.279 [0.000; 1.000] *	0.378 [0.335; 0.420]	$p = 0.020$
	Endoscopic assessment	1	0.368 [0.341; 0.395]	0.363 [0.319; 0.407]	$p = 0.850$
	Therapy modification	5	0.356 [0.000; 1.000]	0.364 [0.324; 0.403]	$p = 0.940$
Immunosuppressors	Biologics	8	0.371 [0.314; 0.428]	0.356 [0.303; 0.410]	$p = 0.710$
	Clinical assessment	1	0.345 [0.278; 0.412]	0.364 [0.322; 0.405]	$p = 0.640$
	Endoscopic assessment	1	0.368 [0.341; 0.395]	0.363 [0.319; 0.407]	$p = 0.850$
	Therapy modification	3	0.286 [0.000; 1.000]	0.376 [0.334; 0.418]	$p = 0.090$
Biologics	Clinical assessment	2	0.340 [0.280; 0.401]	0.366 [0.323; 0.408]	$p = 0.500$
	Endoscopic assessment	1	0.368 [0.341; 0.395]	0.363 [0.319; 0.407]	$p = 0.850$
	Therapy modification	4	0.388 [0.000; 1.000]	0.356 [0.317; 0.395]	$p = 0.780$
Clinical assessment	Endoscopic assessment	1	0.280 [0.212; 0.349] *	0.367 [0.327; 0.408]	$p = 0.030$
	Therapy modification	5	0.285 [0.228; 0.343] *	0.382 [0.339; 0.425]	$p = 0.009$

*: statistically significant from the mean of the subgroup that does not include both variables, $p < 0.05$.

TABLE 3 Subgroup analysis. The entire population of outcomes was divided in subgroups reporting a specific predefined Core Variable, and differences within these subgroups according to the presence or absence of the remaining predefined variables were statistically tested and each square represents the test's result

		Secondary Variable							
		Hosp	Sur	Ster	Immuno	Bio	CA	EA	TM
Core Variable	Hosp (11)			p=0.800	p=0.720	p=0.040	p<0.001		p=0.890
	Sur (14)	p=0.300		p=0.890	p=0.790	p=0.040	p<0.001		p=0.870
	Ster (8)	p=0.790	p=0.790		p=0.730	p=0.020	p=0.130		p>0.990
	Immuno (9)	p=0.170	p=0.880	p=0.010					p=0.009
	Bio (12)	p=0.970	p=0.510	p=0.590	p=0.030				p=0.780
	CA (8)	p=0.190	p=0.230	p=0.820					p=0.790
	EA (2)								
	TM (9)	p=0.900	p=0.440	p=0.750	p=0.250	p=0.470	p=0.240		

The frequency of composite outcomes increases in the subgroup reporting the secondary variable

The frequency of composite outcomes decreases in the subgroup reporting the secondary variable

No significant differences were observed between the groups

The statistical test could not be performed for that particular combination or the subgroups defined by the presence or absence of the considered variable were composed by only one or two outcomes

The *p* values corresponding to each statistical test appear within the corresponding square.

Abbreviations: (n), number of outcomes in the subgroup reporting the Core Variable; Bio, Biologics; CA, Clinical Assessment; EA, Endoscopic Assessment; Hosp, Hospitalization; Immuno, Immunosuppressors; Ster, Steroids; Sur, Surgery; TM, Therapy Modification.

"Biologics" encompasses several different therapeutic agents targeting pro-inflammatory mediators. Unlike immunosuppressors, which have been assigned to a different category in the present analysis, these compounds do not suppress the entire immune system but employ a more selective mechanism of action.[46] Biological therapy has been incorporated relatively recently as a tool for the management of UC (the first therapeutic agent to be developed, Infliximab, only received approval by the FDA in 2005).[47] This fact may explain why the variable is linked to higher frequencies of composite outcomes in the present study, as it is less likely to have been selected as the sole outcome in detriment of other ways of assessing disease progression with a longer tradition in the field.[48]

Our data also highlight how the frequency of composite outcomes is directly dependent on the total duration of the follow-up period in observational studies of UC, with the mean frequency being higher for follow-up periods longer than a year and even higher for those studies that lasted two years or longer. This result is intuitive but it needs to be underlined, as the monitoring of event-free survival should be an important criterion in therapy evaluation[49] and recording the

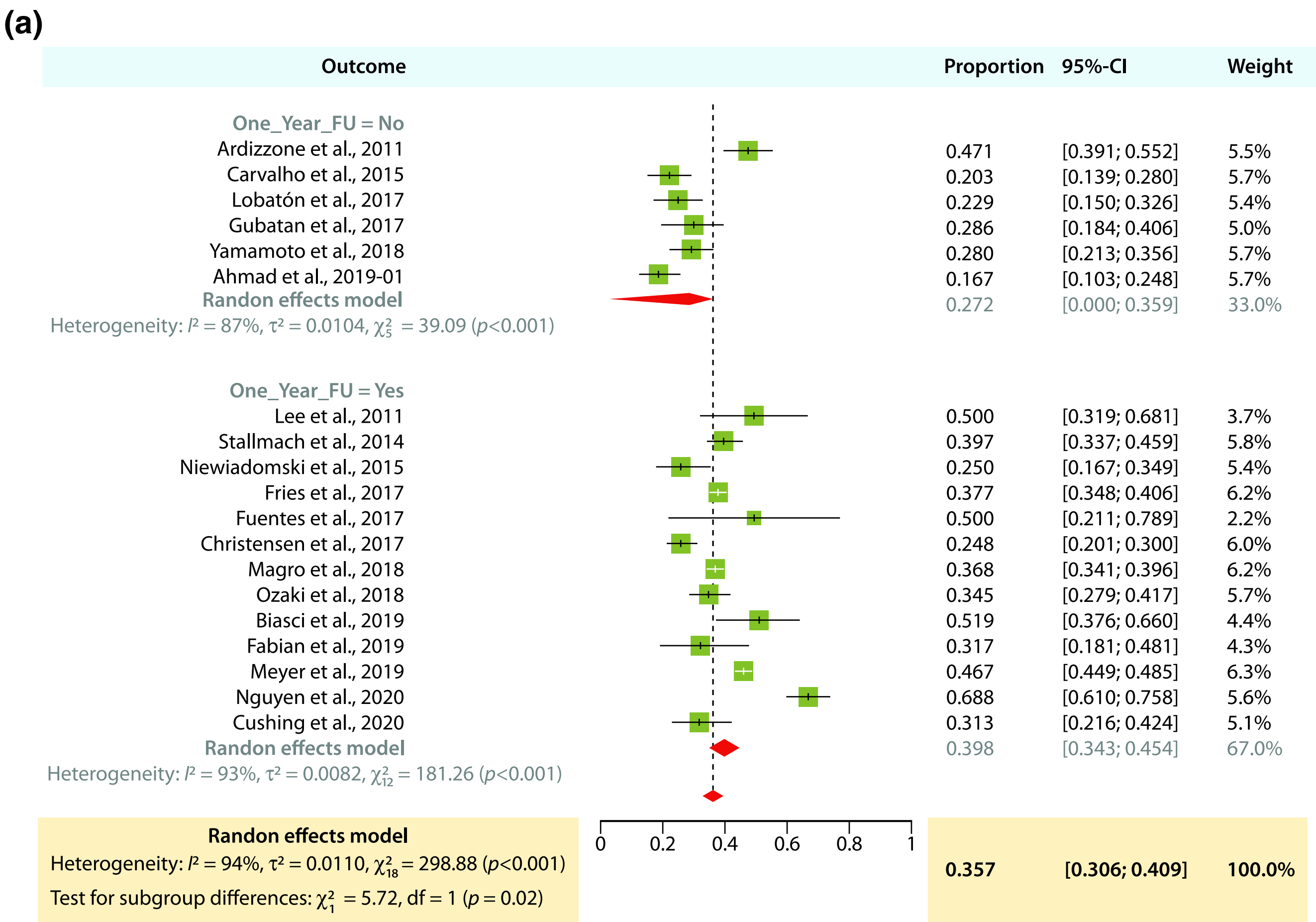

FIGURE 4 Frequency of composite outcomes according to length of the follow-up period. (a) Subgroups determined by a total duration of the follow-up above or below one year. (b) Subgroups determined by a total duration of the follow-up above or below two years. FU: follow-up

frequency of composite outcomes during an extended follow-up period would therefore be a suitable method to do this.

The present report represents the first systematic review and meta-analysis of outcomes in observational studies of UC, and it complements a previous report by other authors focused on UC outcomes but restricted to RCTs.[50] In agreement with the results presented here, that study identified remarkable heterogeneity in the reporting of outcomes, which further emphasizes the current need to reach a consensus on core outcomes for UC. In fact, as described previously,[51] the development of an IBD-specific COS involves four steps: (i) a systematic literature review to identify outcomes previously used in IBD RCTs; (ii) qualitative interviews conducted with patients, clinicians, researchers and other stakeholders to recognize another important outcomes; (iii) an international two-round Delphi survey (to prioritize outcomes for inclusion); (iv) a consensus meeting to accredit and disseminate the findings.

A particular strength of our study is that by focusing on observational studies we were able to include long-term studies (as long as 19 years of continuous monitoring[20]), which is not possible in the case of RCTs, and thus report on the effect that multi-year monitoring has on the frequency of composite outcomes. This is particularly relevant considering the chronic nature of the disease.

Among the limitations of our study is the relatively small number of studies that were included in the analysis. This was a direct consequence of the lack of observational studies of UC available in the literature in which the outcome is properly and unambiguously reported, but the reduced statistical power may have obscured some trends that would perhaps become evident otherwise. We did not assess the reliability of the outcome variables considered, and patient-reported outcome measures were not included because they have yet to be properly validated.[52]

In summary, the present meta-analysis illustrates the heterogeneity that is prevalent for the reporting of clinical outcomes in observational studies of UC. Furthermore, it identifies a specific variable whose inclusion impacts the frequency of composite outcomes in these studies, and provides evidence that the follow-up period is critical to maximize this frequency. Our results suggest that by monitoring treatment with the therapeutic agents included under the general category "Biologics" in addition to a standard clinical assessment, by extending the follow-up period to two years or above and by including more than 50 patients in each study, future observational studies can effectively increase the frequency of patients achieving composite outcomes in the results. Considered together with information already provided by other systematic

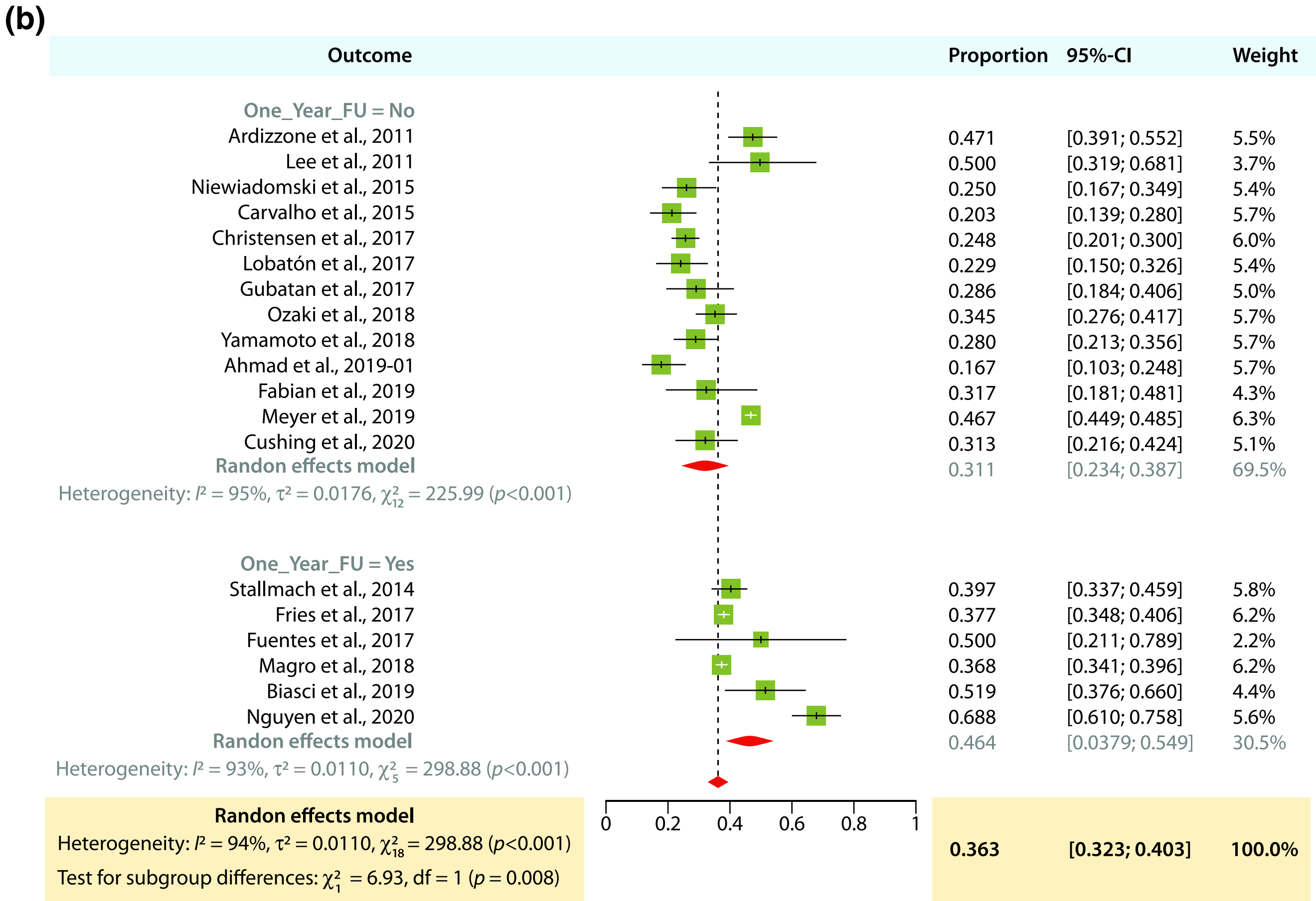

FIGURE 4 (Continued)

reviews in the area, these conclusions have relevance for the development of effective methods to optimize outcome reporting in the UC field.

ACKNOWLEDGEMENT

This work was supported by the Portuguese Group of Studies in Inflammatory Bowel Disease (GEDII).

CONFLICT OF INTERESTS

Fernando Magro received a fee for presenting from: AbbVie, Ferring, Falk, Hospira, PharmaKern, MSD, Schering, Laboratórios Vitoria, Vifor Pharma, OM Pharma. Axel Dignass reports fees for participation in clinical trials, review activities, such as data monitoring boards, statistical analysis, end point committees from Falk, Abbvie, Janssen, Gilead, Pfizer; consultancy fees from Abbvie, MSD, Ferring, Roche/Genentech, Takeda, Vifor, Pharmacosmos, Boehringer-Ingelheim, Falk, Janssen, Pfizer, Sandoz/Hexal, BMS/Celgene, Tillotts, Galapagos, Amgen and Fresenius Kabi; payment from lectures including service on speakers bureaus from Falk Foundation, Ferring, MSD, Abbvie, Vifor, Janssen, Pfizer, Tillotts, Takeda, Gilead/Galapagos; payment for manuscript preparation from Falk Foundation, Thieme, Takeda and UniMed Verlag. Silvio Danese has served as a speaker, consultant and advisory board member for Schering-Plough, AbbVie, MSD, UCB Pharma, Ferring, Cellerix, Millenium Takeda, Nycomed, Pharmacosmos, Actelion, Alphawasserman, Genentech, Grunenthal, Pfizer, Astra Zeneca, Novo Nordisk, Cosmo Pharmaceuticals, Vifor and Johnson & Johnson, Nikkiso Europe GMBH, Theravance. Laurent Peyrin-Biroulet has served as a speaker, consultant, and advisory board member for Merck, Abbvie, Janssen, Genentech, Mitsubishi, Ferring, Norgine, Tillots, Vifor, Hospira/Pfizer, Celltrion, Takeda, Biogaran, Boerhinger-Ingelheim, Lilly, HAC Pharma, Index Pharmaceuticals, Amgen, Sandoz, Forward Pharma GmbH, Celgene, Biogen, Lycera, Samsung Bioepis, Theravance.

AUTHOR CONTRIBUTIONS

Fernando Magro contributed to the study conception and design, drafting of the manuscript, data collection, data interpretation and provided an intellectual contribution. Catarina Alves, Mafalda Santiago, and Maria Manuela Estevinho contributed to the drafting of the manuscript, data collection, data interpretation and analysis. All other authors contributed to data collection and critically revised the manuscript. All authors approved the final version of the manuscript and the authorship list and take responsibility for the accuracy and integrity of any part of the work. Writing assistance was provided by

Ana Rolo and Cristian Bodo, and funded by the Portuguese Group of Studies in IBD (GEDII).

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article (also in the main text document).

ORCID

Fernando Magro https://orcid.org/0000-0003-2634-9668
Maria Manuela Estevinho https://orcid.org/0000-0001-7171-0139

REFERENCES

1. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. Nature. 2011;474:307–17.
2. Peyrin-Biroulet L, Sandborn W, Sands BE, Reinisch W, Bemelman W, Bryant RV, et al. Selecting therapeutic targets in inflammatory bowel disease (STRIDE): determining therapeutic goals for treat-to-target. Am J Gastroenterol. 2015;110:1324–38.
3. Goldberg R, Gore JM, Barton B, Gurwitz J. Individual and composite study endpoints: separating the wheat from the chaff. Am J Med. 2014;127:379–84.
4. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials. 2012;13:132.
5. Ferreira-González I, Permanyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. J Clin Epidemiol. 2007;60:651–7.
6. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6.
7. Higgins G. Cochrane handbook for systematic reviews of interventions | Cochrane training. John Wiley & Sons; 2011. p. 649.
8. Vetter TR, Mascha EJ. Defining the primary outcomes and justifying secondary outcomes of a study: usually, the fewer, the better. Anesth Analg. 2017;125:678–81.
9. Oxford Centre for Triple Value Healthcare. CASP checklists – CASP – critical appraisal Skills Programme; 2021. https://casp-uk.net/casp-tools-checklists/. Accessed 8 Jul 2021.
10. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Br Med J. 2003;327:557–60.
11. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629–34.
12. Yamamoto T, Shimoyama T, Umegae S, Matsumoto K. Endoscopic score vs. fecal biomarkers for predicting relapse in patients with ulcerative colitis after clinical remission and mucosal healing. Clin Transl Gastroenterol. 2018;9.
13. Ahmad W, Nguyen NH, Boland BS, Dulai PS, Pride DT, Bouland D, et al. Comparison of multiplex gastrointestinal pathogen panel and conventional stool testing for evaluation of diarrhea in patients with inflammatory bowel diseases. Dig Dis Sci. 2019;64:382–90.
14. Fabian O, Hradsky O, Lerchova T, Mikus F, Zamecnik J, Bronsky J. Limited clinical significance of tissue calprotectin levels in bowel mucosa for the prediction of complicated course of the disease in children with ulcerative colitis. Pathol Res Pract. 2019:215.
15. Niewiadomski O, Studd C, Hair C, Wilson J, Ding NS, Heerasing N, et al. Prospective population-based cohort of inflammatory bowel disease in the biologics era: disease course and predictors of severity. J Gastroenterol Hepatol. 2015;30:1346–53.
16. Riley SA, Mani V, Goodman MJ, Dutt S, Herd ME. Microscopic activity in ulcerative colitis: what does it mean? Gut. 1991;32:174–8.
17. Kochar B, Barnes EL, Long MD, Cushing KC, Galanko J, Martin CF, et al. Depression is associated with more aggressive inflammatory bowel disease. Am J Gastroenterol. 2018;113:80–5.
18. Gubatan J, Mitsuhashi S, Zenlea T, Rosenberg L, Robson S, Moss AC. Low serum vitamin D during remission increases risk of clinical relapse in patients with ulcerative colitis. Clin Gastroenterol Hepatol. 2017;15:240–6.
19. Ozaki R, Kobayashi T, Okabayashi S, Nakano M, Morinaga S, Hara A, et al. Histological risk factors to predict clinical relapse in ulcerative colitis with endoscopically normal mucosa. J Crohns Colitis. 2018;12:1288–94.
20. Magro F, Dias CC, Portela F, Miranda M, Fernandes S, Bernardo S, et al. Development and validation of risk matrices concerning ulcerative colitis outcomes-bayesian network analysis. J Crohns Colitis. 2019;13:401–9.
21. Meyer A, Rudant J, Drouin J, Weill A, Carbonnel F, Coste J. Effectiveness and safety of reference infliximab and biosimilar in Crohn disease: a French equivalence study. Ann Intern Med. 2019;170:99–107.
22. Cushing KC, Tan W, Alpers DH, Deshpande V, Ananthakrishnan AN. Complete histologic normalisation is associated with reduced risk of relapse among patients with ulcerative colitis in complete endoscopic remission. Aliment Pharmacol Therapeut. 2020;51:347–55.
23. Nguyen NH, Kurnool S, Dulai PS, Boland BS, Sandborn WJ, Singh S. Short disease duration is associated with increased risk of treatment failure in biologic-Treated patients with ulcerative colitis. Inflamm Bowel Dis. 2020;26:1429–35.
24. Bitton A, Peppercorn MA, Antonioli DA, Niles JL, Shah S, Bousvaros A, et al. Clinical, biological, and histologic parameters as predictors of relapse in ulcerative colitis. Gastroenterology. 2001;120:13–20.
25. Buda A, Hatem G, Neumann H, D' Incà R, Mescoli C, Piselli P, et al. Confocal laser endomicroscopy for prediction of disease relapse in ulcerative colitis: a pilot study. J Crohns Colitis. 2014;8:304–11.
26. Barreiro-de Acosta M, Vallejo N, De La Iglesia D, Uribarri L, Bastón I, Ferreiro-Iglesias R, et al. Evaluation of the risk of relapse in ulcerative colitis according to the degree of mucosal healing (Mayo 0 vs 1): a longitudinal cohort study. J Crohns Colitis. 2016;10:13–19.
27. Höie O, Wolters F, Riis L, Aamodt G, Solberg C, Bernklev T, et al. Ulcerative colitis: patient characteristics may predict 10-yr disease recurrence in a European-wide population-based cohort. Am J Gastroenterol. 2007;102:1692–701.
28. Lund JL, Stürmer T, Porter CQ, Sandler RS, Kappelman MD. Thiazolidinedione use and ulcerative colitis-related flares: an exploratory analysis of administrative data. Inflamm Bowel Dis. 2011;17:787–94.
29. Frieri G, Galletti B, Di Ruscio M, Tittoni R, Capannolo A, Serva D, et al. The prognostic value of histology in ulcerative colitis in clinical remission with mesalazine. Therap Adv Gastroenterol. 2017;10:749–59.
30. Kurnool S, Nguyen NH, Proudfoot J, Dulai PS, Boland BS, Vande Casteele N, et al. High body mass index is associated with increased risk of treatment failure and surgery in biologic-treated patients with ulcerative colitis. Aliment Pharmacol Therapeut. 2018;47:1472–9.
31. Jairath V, Jeyarajah J, Zou G, Parker CE, Olson A, Khanna R, et al. A composite disease activity index for early drug development in ulcerative colitis: development and validation of the UC-100 score. Lancet Gastroenterol Hepatol. 2019;4:63–70.
32. Ungaro RC, Limketkai BN, Jensen CB, Allin KH, Agrawal M, Ullman T, et al. Stopping 5-aminosalicylates in patients with ulcerative colitis starting biologic therapy does not increase the risk of adverse clinical outcomes: analysis of two nationwide population-based cohorts. Gut. 2019;68:977–84.

33. Pereira MS, Durães C, Catarino TA, Costa JL, Cleynen I, Novokmet M, et al. Genetic variants of the MGAT5 gene are functionally implicated in the modulation of T cells glycosylation and plasma IgG glycome composition in ulcerative colitis. Clin Transl Gastroenterol. 2020;11:e00166.
34. Stallmach A, Nickel L, Lehmann T, Bokemeyer B, Bürger M, Hüppe D, et al. Parameters of a severe disease course in ulcerative colitis. World J Gastroenterol. 2014;20:12574–80.
35. Carvalho PB, De Castro FD, Rosa B, Moreira MJ, Cotter J. Mucosal healing in ulcerative colitis – when zero is better. J Crohns Colitis. 2016;10:20–5.
36. Ardizzone S, Cassinotti A, Duca P, Mazzali C, Penati C, Manes G, et al. Mucosal healing predicts late outcomes after the first course of corticosteroids for newly diagnosed ulcerative colitis. Clin Gastroenterol Hepatol. 2011;9.
37. Fries W, Viola A, Manetti N, Frankovic I, Pugliese D, Monterubbianesi R, et al. Disease patterns in late-onset ulcerative colitis: results from the IG-IBD "AGED study". Dig Liver Dis. 2017;49: 17–23.
38. Fuentes S, Rossen NG, Van Der Spek MJ, Hartman JHA, Huuskonen L, Korpela K, et al. Microbial shifts and signatures of long-term remission in ulcerative colitis after faecal microbiota transplantation. ISME J. 2017;11:1877–89.
39. Christensen B, Hanauer SB, Erlich J, Kassim O, Gibson PR, Turner JR, et al. Histologic normalization occurs in ulcerative colitis and is associated with improved clinical outcomes. Clin Gastroenterol Hepatol. 2017;15:1557–64.
40. Lobatón T, Bessissow T, Ruiz-Cerulla A, De Hertogh G, Bisschops R, Guardiola J, et al. Prognostic value of histological activity in patients with ulcerative colitis in deep remission: a prospective multicenter study. United European Gastroenterol J. 2018;6: 765–72.
41. Ananthakrishnan A, Sakuraba A, Barnes EL, Pekow J, Raffals L, Long MD, et al. The benefit of combination therapy depends on disease phenotype and duration in Crohn's disease. Aliment Pharmacol Ther. 2017;46:162–8.
42. Lee JC, Lyons PA, McKinney EF, Sowerby JM, Carr EJ, Bredin F, et al. Gene expression profiling of CD8 + T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. J Clin Invest. 2011;121:4170–9.
43. Biasci DB, Lee JC, Noor NM, Pombal DR, Hou M, Lewis N, et al. A blood-based prognostic biomarker in IBD. Gut. 2019;68:1386–95.
44. Thomsen SB, Allin KH, Burisch J, Jensen CB, Hansen S, Gluud LL, et al. Outcome of concomitant treatment with thiopurines and allopurinol in patients with inflammatory bowel disease: a nationwide Danish cohort study. United European Gastroenterol J. 2020;8:68–76.
45. Walsh A, Travis S. Assessing disease activity in patients with ulcerative colitis. Gastroenterol Hepatol. 2012;8:751–4.
46. Arora Z, Shen B. Biological therapy for ulcerative colitis. Gastroenterol Rep. 2015;3:103–9.
47. Levin A, Shibolet O. Infliximab in ulcerative colitis. Biol Targets & Ther. 2008;2:379–88.
48. D'Haens G, Sandborn WJ, Feagan BG, Geboes K, Hanauer SB, Irvine EJ, et al. A review of activity indices and efficacy end points for clinical trials of medical therapy in adults with ulcerative colitis. Gastroenterology. 2007;132:763–86.
49. Lubsen J, Kirwan BA. Combined endpoints: can we use them? Stat Med. 2002;21:2959–70.
50. Ma C, Panaccione R, Fedorak RN, Parker CE, Nguyen TM, Khanna R, et al. Heterogeneity in definitions of endpoints for clinical trials of ulcerative colitis: a systematic review for development of a core outcome set. Clin Gastroenterol Hepatol. 2018;16:637–47.
51. Ma C, Panaccione R, Fedorak RN, Parker CE, Khanna R, Levesque BG, et al. Development of a core outcome set for clinical trials in inflammatory bowel disease: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. BMJ Open. 2017;7(6):e016146.
52. de Jong MJ, Huibregtse R, Masclee AAM, Jonkers DMAE, Pierik MJ. Patient-reported outcome measures for use in clinical trials and clinical practice in inflammatory bowel diseases: a systematic review. Clin Gastroenterol Hepatol. 2018;16:648–63.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Magro F, Alves C, Santiago M, Ministro P, Lago P, Correia L, et al. Composite outcomes in observational studies of ulcerative colitis: a systematic review and meta-analysis. United European Gastroenterol J. 2022; 10(1):54–72. https://doi.org/10.1002/ueg2.12183