BMJ Open

Use of external control arms in immune-mediated inflammatory diseases: a systematic review

Alexa Zayadi,[1] Robert Edge,[1] Claire E Parker,[1] John K Macdonald,[1] Blue Neustifter,[1] Joshua Chang,[1] Guowei Zhong,[1] Siddharth Singh,[2,3] Brian G Feagan,[1,4,5] Christopher Ma,[1,6,7] Vipul Jairath [1,4,5]

To cite: Zayadi A, Edge R, Parker CE, *et al*. Use of external control arms in immune-mediated inflammatory diseases: a systematic review. *BMJ Open* 2023;**13**:e076677. doi:10.1136/bmjopen-2023-076677

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2023-076677).

Received 14 June 2023
Accepted 31 October 2023

For numbered affiliations see end of article.

Correspondence to
Dr Vipul Jairath;
vjairath@uwo.ca

ABSTRACT

Objectives External control arms (ECAs) provide useful comparisons in clinical trials when randomised control arms are limited or not feasible. We conducted a systematic review to summarise applications of ECAs in trials of immune-mediated inflammatory diseases (IMIDs).

Design Systematic review with an appraisal of ECA source quality rated across five domains (data collection, study populations, outcome definitions, reliability and comprehensiveness of the dataset, and other potential limitations) as high, low or unclear quality.

Data sources Embase, Medline and Cochrane Central Register of Controlled Trial were searched through to 12 September 2023.

Eligibility criteria Eligible studies were single-arm or randomised controlled trials (RCTs) of inflammatory bowel disease, pouchitis, rheumatoid arthritis, juvenile idiopathic arthritis, ankylosing spondylitis, psoriatic arthritis, psoriasis and atopic dermatitis in which an ECA was used as the comparator.

Data extraction and synthesis Two authors independently screened the search results in duplicate. The characteristics of included studies, external data source(s), outcomes and statistical methods were recorded, and the quality of the ECA data source was assessed by two independent authors.

Results Forty-three studies met the inclusion criteria (inflammatory bowel disease: 16, pouchitis: 1, rheumatoid arthritis: 12, juvenile idiopathic arthritis: 1, ankylosing spondylitis: 5, psoriasis: 3, multiple indications: 4). The majority of these trials were single-arm (33/43) and enrolled adult patients (34/43). All included studies used a historical control rather than a contemporaneous ECA. In RCTs, ECAs were most often derived from the placebo arm of another RCT (6/10). In single-arm trials, historical case series were the most common ECA source (19/33). Most studies (31/43) did not employ a statistical approach to generate the ECA from historical data.

Conclusions Standardised ECA methodology and reporting conventions are lacking for IMIDs trials. The establishment of ECA reporting guidelines may enhance the rigour and transparency of future research.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This systematic review used a comprehensive predefined approach to identify published, interventional immune-mediated inflammatory disease studies in which external data have been used as the comparator.
⇒ Although we devised a sensitive search strategy comprising keywords and controlled vocabulary, the language used to describe external control arm (ECA) methodology lacks standardisation and ECA-specific MeSH terms do not exist.
⇒ Our assessment of ECA data source quality was conducted with an instrument that has not been validated to appraise ECA methodology, as no such instrument has been developed for use in the context of systematic reviews.

BACKGROUND

Randomised controlled trials (RCTs) are the gold standard for evaluating the efficacy and safety of medical interventions.[1] The strengths of randomisation include minimisation of differences of known and unknown confounders among experimental groups, reduced patient selection bias and facilitation of robust statistical analyses. The randomised controlled design also permits blinding, which mitigates potential bias introduced by study personnel, participants and outcome assessors. Despite these advantages, RCTs can present practical, logistical and ethical challenges.[1] For example, it can be difficult to enrol a sufficient number of participants when the disease is rare; there are multiple competing drug development programmes, or trial avoidance where patients are concerned about being allocated to placebo. Furthermore, placebo-controlled RCTs may be unethical if there is a lack of clinical equipoise, or an existing well-established standard of care therapy available.

As noted by the US Food and Drug Administration (FDA)[2] and the European Medicines Agency (EMA),[3] there are instances where it may be acceptable to use external data to form a control arm, instead of a concurrently

randomised internal control group. External control arms (ECAs) comprise participants who are not part of the same study as the group receiving the investigational agent.[4] Both regulatory agencies have indicated that a well-characterised disease course, large anticipated treatment effect, sufficiently similar treatment and external populations, and the use of objective outcome measures are important factors in externally controlled designs.[2 3] Patient data used to form the ECA can be collected during the same or similar time period (contemporaneous control) or derived from a previously treated patient population where retrospective or retrospectively analysed data are used as a comparator (historical control).[2–7] These data are typically sourced from prior clinical trials, disease registries, cohort studies, electronic health records and claims databases. External data can also be used to augment the sample size of an internal control arm. The use of an ECA can reduce sample size requirements, avoid ethical pitfalls of using placebo and provides an opportunity to use historical control arm data which may be both valid and large.[2 4–6] Yet, potential bias introduced by differences in study populations, trial design and outcome assessment along with the inability to control for unknown confounders and lack of blinding necessitates careful consideration of data source selection, analytical methods and missing data handling.

To date, ECAs have most frequently been used in clinical trials of oncology and rare diseases.[8] Both the FDA and EMA have published guidance documents outlining recommendations for using real-world data and real-world evidence in regulatory decision-making,[9–13] and key recommendations for the design and conduct of externally controlled clinical trials.[2] Study designs involving comparisons with ECAs are increasingly being submitted to regulatory agencies, and explored in novel therapeutic areas.[14 15] Between 2005 and 2017, 43 product applications for indications in haematology and rare diseases were submitted for regulatory approval, of which 98% and 79% were approved by the FDA and EMA, respectively.[14] In an analysis of 45 FDA approvals of non-oncology products that relied on external controls, 20% of approved products were for the treatment of a common disease (ie, affecting more than 200 000 patients in the USA) and 91% were recognised as addressing an unmet medical need.[15] Given the willingness of regulators to consider ECA data in clinical development programmes and practical issues inherent to RCT design, there is growing interest in exploring how to best adopt this methodology within novel contexts.

Conducting RCTs for immune-mediated inflammatory diseases (IMIDs), such as inflammatory bowel disease, can be challenging due to slow and competitive patient recruitment. Contributing factors include a large clinical development pipeline, the availability of several approved treatment options in routine clinical care, and increasingly stringent eligibility criteria, particularly with respect to the use of prior and concomitant therapies.[16] In addition, the chance of exposure to placebo can be a deterrent for participation in clinical trials for both patients and investigators. The incorporation of ECA data into IMID trial design may provide an efficient alternative or supplement to a traditional randomised control arm. Currently, it is unclear how often and in what setting ECAs have been incorporated into IMID trials, and what data sources and statistical techniques have been used.

This systematic review aimed to identify and describe trials of IMIDs (ie, Crohn's disease, ulcerative colitis, chronic refractory pouchitis, rheumatoid arthritis, juvenile idiopathic arthritis, ankylosing spondylitis, psoriatic arthritis, psoriasis and atopic dermatitis) that have used an ECA, and assess the methodological quality of these studies.

METHODS

This systematic review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement[17] and was conducted following an a priori developed protocol (available on request).

Search strategy

Embase, Medline and the Cochrane Central Register of Controlled Trials were searched without language restriction from database inception to 12 September 2023, using predefined strategies (online supplemental appendix 1). A semiautomated recursive search of the bibliographies of relevant publications, including review articles and meta-analyses, was also performed using the Systematic Review Accelerator SpiderCite tool (Bond University, Gold Coast, Qld, Australia).

Eligibility criteria

Eligible studies: (1) compared data from a randomised or single-arm clinical trial with a control arm from an external source (ie, previous clinical trials, disease registries, commercial insurance or national health insurance claims databases, or electronic health records); (2) enrolled adults or children with Crohn's disease, ulcerative colitis, chronic refractory pouchitis, rheumatoid arthritis/juvenile idiopathic arthritis, ankylosing spondylitis, psoriatic arthritis, psoriasis, or atopic dermatitis; and (3) assessed a medical intervention. The ECA comprised patients from placebo, standard of care, or active comparator groups. Trials were eligible for inclusion regardless of whether the ECA was prospectively or retrospectively integrated into the study design. Protocols for upcoming trials were considered, in addition to studies published in full text or abstract form. Studies were excluded if they were non-interventional or if the active arm data were not collected as part of a clinical trial. Additionally, we excluded studies that did not report the ECA dataset source.

Screening and data extraction

Search results were independently screened in duplicate by two authors (AZ and JC) and data extraction was

completed by two independent authors (AZ and RE) using a standardised spreadsheet. Data collected from each study included: the publication year, study design, trial phase, indication, study arms, number of participants, inclusion/exclusion criteria, interventions, source of ECA data, statistical methods, outcomes assessed, outcome type (dichotomous vs continuous) and rationale for using an ECA. Disagreements encountered during screening and data extraction were resolved by consensus or, if consensus could not be reached, through arbitration by a third author.

Quality assessment

The methodological quality of the ECA data source was assessed by two authors (AZ and RE) using a checklist adapted from Thorlund *et al*.[18] The checklist comprised 5 key domains: (1) quality of data source; (2) similarly of the study populations; (3) similarities of the outcome definitions; (4) the reliability and comprehensiveness of the dataset; and (5) other potential limitations. Key judgement criteria are reported in online supplemental table 1. Each domain was classified as being of high (all criteria are met), low (no criteria are met) or unclear quality (some criteria are met and/or insufficient information to ascertain whether all criteria are met).

Patient involvement

No patients were involved in the study concept and design, acquisition and interpretation of data, or drafting of the manuscript.

RESULTS

Search results

A total of 2192 records were identified from database and recursive searching, from which 610 duplicates were removed. Based on the information provided in the titles and abstracts of the remaining 1582 records, 1442 were excluded as non-applicable. Full-text review was performed for 140 records, of which 97 were excluded with reasons (online supplemental table 2). Forty-three studies met the eligibility criteria and were included (figure 1).

Randomised controlled trials

Characteristics of included studies

Ten of the 43 included studies (23%) compared the active arm(s) of an RCT to an ECA (table 1).[19–28] Five of the included RCTs (5/10; 50%) prospectively leveraged external control data as part of the study design.[19–23] Four of these studies (4/5; 80%) randomised participants to 1 of 2 active arms and used an ECA to completely replace

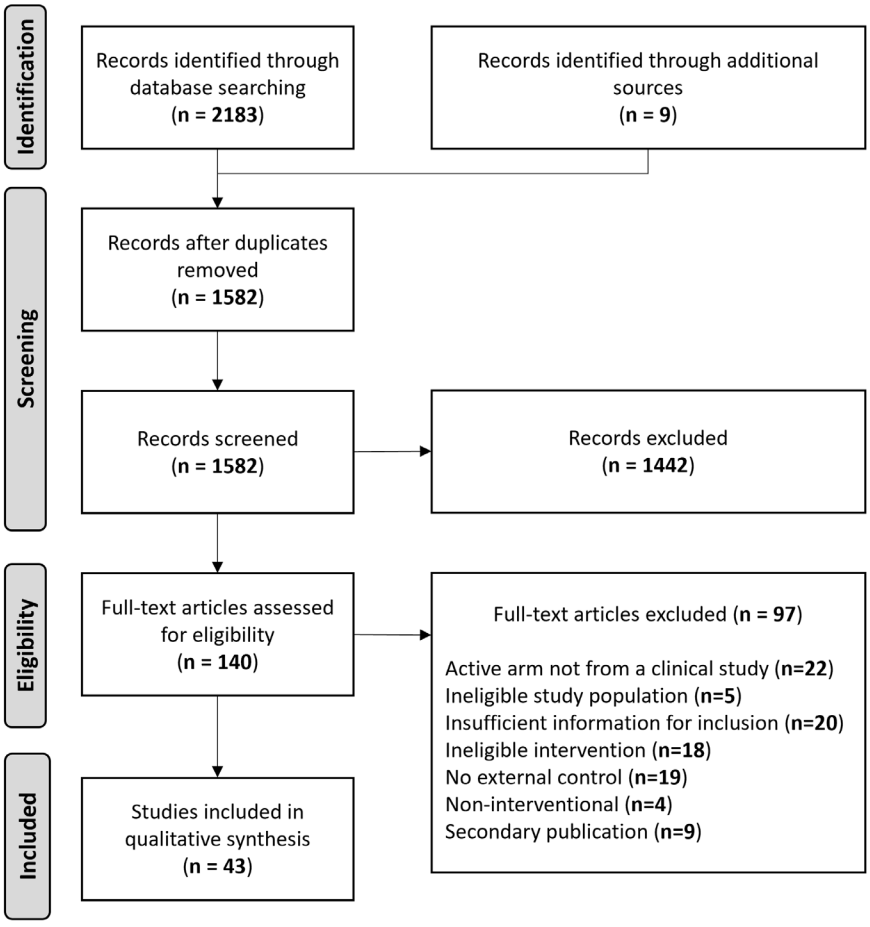

Figure 1 Study flow diagram of study inclusion. Adapted from Page *et al*.[17]

Table 1 Characteristics of included randomised controlled trials

Study ID	Indication	Population	Active arm			ECA		
			N	Phase	Intervention	N	Source	Control
Hueber *et al*, 2012[19]	CD	Adult	39	2	Secukinumab	671	6 RCTs (phase 2–3; published)	Placebo
Hyams *et al*, 2012[20]	UC	Paediatric	60	3	Infliximab	244	RCT (phase 3; published)	Placebo
Hodkinson *et al*, 2013[21]	RA	Adult	87	NR	DMARDs+tight control strategy	145	Historical case series* (unpublished)	DMARDs
Croft *et al*, 2021[22]	UC	Paediatric	93	3	Adalimumab	NR	4 RCTs (phase 3; published)	Placebo
Hupf *et al*, 2021†[25]	AS	Adult	23	2	Secukinumab	513	8 RCTs (phase 2–3; published)	Placebo
Reed *et al*, 2021†[24]	RA	Adult	NR	1–3	Tofacitinib	NR	Disease registry (published)	DMARDs
Rudrapatna *et al*, 2021†[26]	CD	Adult	73	2b	Tesnatilimab	56	EHR data (unpublished)	Ustekinumab
Wang *et al*, 2021†[27]	RA	Adult	54	2	Upadacitinib	159	RCT (phase 3; published)	Placebo
Magnolo *et al*, 2022[23]	PsO	Paediatric	84	3	Secukinumab	812	5 RCTs (phase 3; published)	Placebo
Norvang *et al*, 2022†[28]	RA	Adult	188	NR	DMARDs+stringent T2T strategy	328	Prospective observational study (published)	DMARDs

*Historical control comprised of patients who have previously received treatment at the same site(s) as active arm study participants.
†Primary study objective was to examine methodological approaches for complementing prior or concurrent RCT data with external control data.
AS, ankylosing spondylitis; CD, Crohn's disease; DMARD, disease-modifying antirheumatic drug; ECA, external control arm; EHR, electronic health record; NR, not reported; PsO, psoriasis; RA, rheumatoid arthritis; RCT, randomised controlled trial; T2T, treat-to-target; UC, ulcerative colitis.

the function of a placebo arm, while 1 (1/5; 20%) was a placebo-controlled RCT that incorporated external data to augment the sample size of the placebo group. Among the five RCTs that prospectively used external data were three phase 3[20 22 23] (60%) and one phase 2[19] (20%) trials. The remaining five RCTs (5/10; 50%) retrospectively compared data from the active arm(s) of prior (4/5; 80%) or ongoing (1/5; 20%) phase 1–3 trials with an ECA.[24–28]

Most of the included RCTs enrolled adults (7/10; 70%), with 3 of the 10 trials (30%) conducted in paediatric populations. Study populations comprised patients with inflammatory bowel disease (4/10; 40% (ulcerative colitis: 2/10; 20%, Crohn's disease: 2/10; 20%)), rheumatoid arthritis (4/10; 40%), psoriasis (1/10; 10%), and ankylosing spondylitis (1/10; 10%). Among these studies, biologic therapies (6/10; 60%) were the most common intervention under investigation, including secukinumab (3/10; 30%), infliximab (1/10; 10%), adalimumab (1/10; 10%) and tesnatilimab (1/10; 10%). Two out of 10 studies (20%) evaluated small molecules (tofacitinib (1/10; 10%) and upadacitinib (1/10; 10%)). Treatment strategies were assessed in two RCTs (tight control strategy (1/10; 10%) and stringent treat-to-target strategy (1/10; 10%)).

The majority of studies (6/10; 60%) made comparisons with placebo data derived from 1 or more previously conducted RCTs. An external active control group was used by 2 out of 10 studies (20%), with data sourced from a disease registry (1/10; 10%) or electronic health records (1/10; 10%). Lastly, 2 studies (2/10; 20%) formed a standard of care control group using data from a historical case series (1/10; 10%) or a prior prospective cohort study (1/10; 10%).

External control methodology

All 10 studies used a historical ECA; we did not identify any studies that employed contemporaneously collected external data. Various methods were used to generate ECAs among the 10 studies that compared the active arm(s) of an RCT to external control data (table 2). Statistical methods used to adjust the external dataset included propensity score matching (2/10; 20%), propensity score weighting (1/10; 10%) and Bayesian dynamic borrowing (2/10; 20%). One study did not specify the weighting method employed (1/10; 10%). Four studies (4/10; 40%) did not make any statistical adjustments to ECA data to balance characteristics between active and external arms. Nine out of 10 studies (90%) made comparisons between the active arm and ECA for at least one efficacy outcome. In all cases, efficacy outcomes were defined by an indication-specific disease activity index and were assessed as dichotomous (9/9; 100%) or continuous (1/9; 11.1%) endpoints. Two studies (2/10; 20%) used external data to contextualise safety outcomes (ie, adverse events and serious adverse events).

Table 2 External control arm methodology in randomised controlled trials

Study ID	Indication	Statistical methods for generating ECA	Outcome(s) compared with ECA	Outcome timepoint (weeks)	Evidence of treatment benefit compared with ECA*	Rationale for using an ECA
Hueber *et al*, 2012[19]	CD	Weighting	Efficacy (CDAI response rate)	6	No	Small study sample size
Hyams *et al*, 2012[20]	UC	Unadjusted	Efficacy (MCS response rate)	8	Yes	Unethical to use internal placebo arm
Hodkinson *et al*, 2013[21]	RA	Unadjusted	Efficacy (DAS 28 response and remission rates)	26	Yes	NR
Croft *et al*, 2021[22]	UC	Unadjusted	Efficacy (MCS remission rate)	8, 52	Yes	Unethical to use internal placebo arm advised by regulatory organisations; low recruitment
Hupf *et al*, 2021†[25]	AS	Bayesian dynamic borrowing	Efficacy (ASAS 20 response rate)	6	NA	ECA methodology study
Reed *et al*, 2021†[24]	RA	Propensity score matching	Safety (SAEs)	NR	NA	ECA methodology study
Rudrapatna *et al*, 2021†[26]	CD	Unadjusted	Efficacy (mean change in CDAI, steroid-free CDAI remission rate)	12, 24	NA	ECA methodology study
Wang *et al*, 2021†[27]	RA	Propensity score matching	Efficacy (ACR 20 response rate)	12	NA	ECA methodology study
Magnolo *et al*, 2022[23]	PsO	Bayesian dynamic borrowing	Efficacy (PASI 75, PASI 90 & IGA 0/1 response rates) Safety (AEs)	12	Yes	Unethical to use internal placebo arm; advised by regulatory organisations
Norvang *et al*, 2022†[28]	RA	Propensity score weighting	Efficacy (DAS 28 remission rate)	24, 52, 104	NA	ECA methodology study

*As defined by the included study.
†Primary study objective was to examine methodological approaches for complementing prior or concurrent RCT data with external control data.
ACR, American College of Rheumatology; AE, adverse event; AS, ankylosing spondylitis; ASAS, Assessment of SpondyloArthritis International Society; CD, Crohn's disease; CDAI, Crohn's Disease Activity Index; DAS, Disease Activity Score; ECA, external control arm; IGA, Investigator's Global Assessment; MCS, Mayo Clinical Score; NA, not applicable; NR, not reported; PASI, Psoriasis Area Severity Index; PsO, psoriasis; RA, rheumatoid arthritis; SAE, serious adverse event; UC, ulcerative colitis.

Indications for use of an ECA

Three of the included RCTs (3/10; 30%) reported that comparisons to external data were made due to ethical issues associated with randomising participants to a placebo arm (table 2).[20 22 23] All were phase 3 trials in paediatric populations, with 2 of these studies (2/3; 66.7%) adopting a study design that included an ECA following consultations with regulatory agencies.[22 23] One study (1/10; 10%) additionally cited poor study recruitment due to non-acceptance of an internal control in the paediatric setting as a reason for incorporating external data.[22 23] Augmenting the small sample size of the internal placebo group was provided as the rationale for introducing external placebo data into the primary analysis in 1 of the included RCTs (1/10; 10%).[19]

Single-arm studies

Characteristics of included studies

The majority of the included studies (33/43; 77%) were single-arm trials that relied on external data to provide context to treatment outcomes (table 3).[29–61] Among these were one phase 2a trial (1/33; 3%), one phase 2b trial (1/33; 3%), one phase 2–3 trial (1/33; 3%), four phase 3 long-term RCT extensions (4/33; 12%) and 4 post-marketing surveillance studies (4/33; 12%).

Most of the included single-arm trials enrolled adult participants (27/33; 81%), with some studies conducted in paediatric (4/33; 12%) and mixed-age (2/33; 6%) populations. Therapies for the treatment of inflammatory bowel disease (13/33; 40% (ulcerative colitis: 5/33; 15%, Crohn's disease: 1/33; 3%)), rheumatoid arthritis

Table 3 Characteristics of included single-arm studies

Study ID	Indication	Population	Active arm		ECA		
			N	Intervention	N	Source	Control
Baker and Jewell, 1989[29]	IBD	Adult and paediatric	24	SOC+ciclosporin	406	Multiple* (published)	SOC
Rosenberg *et al*, 1990[30]	UC	Adult	20	Methylprednisolone (high dose)	175	Multiple† (published)	Prednisolone or methylprednisolone (standard dose)
Stenger *et al*, 1998[31]	RA	Adult	139	Sulfasalazine±methotrexate	89	Historical case series‡ (unpublished)	NSAIDs
Marotte *et al*, 2007[32]	RA	Adult	90	Infliximab	99	Historical case series‡ (unpublished)	Methotrexate
Shen *et al*, 2007[34]	Pouchitis	Adult	16	Ciprofloxacin+tinidazole	10	two historical case series (published)	Mesalamine
Moul *et al*, 2007[33]	PsO	Adult	16	Alefacept (high dose)	166	RCT (phase 3; published)	Alefacept (standard dose)
van der Heijde *et al*, 2008§[36]	AS	Adult	201	Infliximab	192	Disease registry (unpublished)	SOC (anti-TNF naïve)
Boyle *et al*, 2008[35]	IBD	Adult	14	Laparoscopic resection	16	Historical case series‡ (unpublished)	Open colon resection
van der Heijde *et al*, 2009§[37]	AS	Adult	307	Adalimumab	169	Disease registry (unpublished)	SOC (anti-TNF naïve)
Garg *et al*, 2011[38]	IBD	Adult	100	Rapid iron polymaltose infusion	486	two historical case series (published)	Slow iron polymaltose infusion
Yong *et al*, 2011¶[39]	PsO	Adult	NR	Etanercept	NR	Administrative claims database (unpublished)	SOC (non-biologics)
Carter *et al*, 2012**[40]	RA	Adult	50	Rituximab+prednisone	NR	RCTs (package insert; published)	Rituximab+methylprednisolone
Kwok and Leung, 2012[41]	RA	Adult	20	SOC+T2T strategy	20	Historical case series‡ (unpublished)	SOC
Baraliakos *et al*, 2014[43]	AS	Adult	22	Infliximab	34	Historical case series‡ (unpublished)	SOC (anti-TNF naïve)
Horneff *et al*, 2014††[44]	Multiple‡‡	Paediatric	127	Etanercept	323	7 RCTs (published)	Placebo
Lichtenstein *et al*, 2014§[45]	UC	Adult	388	Mesalamine	83	RCTs (phase 3+package insert; published)	Placebo
Gardenbroek *et al*, 2015[42]	UC	Adult	15	Early surgical closure	29	Historical case series‡ (unpublished)	SOC
Mahajan *et al*, 2018[48]	UC	Adult	41	Faecal microbiota transplantation	38	Historical case series‡ (unpublished)	Azathioprine
Rozette *et al*, 2018[49]	Multiple§§	Adult and paediatric	66	Infliximab (rapid infusion)	50	Historical case series‡ (unpublished)	Infliximab (standard infusion)
Tweehuysen *et al*, 2018[50]	Multiple¶¶	Adult	625	Etanercept biosimilar (SB4)	600	Historical case series‡ (unpublished)	Etanercept
Horneff *et al*, 2018***[46]	JIA	Paediatric	18	Golimumab	3620	Historical case series (unpublished)	Anti-TNF or methotrexate
Inui *et al*, 2018[47]	RA	Adult	13	Etanercept+on-demand dosing regimen	16	Historical case series‡ (unpublished)	Etanercept

Continued

Table 3 Continued

Study ID	Indication	Population	Active arm		ECA		
			N	Intervention	N	Source	Control
Braun *et al*, 2019§[51]	AS	Adult	168	Secukinumab	69	RCT (phase 4; published)	NSAIDs
Kawashiri *et al*, 2020***‡‡‡[52]	RA	Adult	NR	Infliximab biosimilar (CT-P13)	NR	2 cohort studies (published)	SOC (anti-TNF)
Muskens *et al*, 2020[53]	Multiple¶¶	Adult	84	Etanercept biosimilar	89	Historical case series‡ (unpublished)	Etanercept
Sanchez-Hernandez *et al*, 2020[54]	IBD	Adult	81	Infliximab+proactive TDM	76	Historical case series‡ (unpublished)	Infliximab
Sagami *et al*, 2021***[57]	IBD	Adult	700	CT-P13	44 274	Administrative claims database (unpublished)	Infliximab
Gasparetto *et al*, 2021[56]	IBD	Paediatric	30	Infliximab+proactive TDM	37	Historical case series‡ (unpublished)	Infliximab
Danese *et al*, 2021[55]	UC	Adult	50	RVT-3101	234	2 RCTs (phase 3; published)	Placebo
Holland *et al*, 2022[59]	CD	Paediatric	37	Infliximab+proactive TDM	37	Historical case series‡ (unpublished)	Infliximab
Huang *et al*, 2022[60]	RA	Adult	111	Multiple†††+tight control T2T strategy	389	Historical case series‡ (unpublished)	Multiple†††+T2T strategy
Truong *et al*, 2022[58]	IBD	Adult	65	Transanal approach restorative proctocolectomy	100	Historical case series‡ (unpublished)	Open approach restorative proctocolectomy
Takei *et al*, 2022[61]	RA	Adult	60	Hydroxychloroquine+DMARDs	276	Cohort study (published)	SOC (DMARDs)

*Control data sourced from two cohort studies, one historical case series and one RCT.
†Control data sourced from two historical case series and one RCT.
‡Historical control comprised patients who have previously received treatment at the same site(s) as active arm study participants.
§Phase 3 RCT extension.
¶Primary study objective was to examine methodological approaches for complementing prior or concurrent RCT data with external control data.
**Phase 2/3 trial.
††Phase 3b trial.
‡‡JIA, enthesitis-related arthritis, and psoriatic arthritis.
§§CD, UC, and JIA.
¶¶RA, psoriatic arthritis, and AS.
***Post-marketing surveillance study.
†††Methotrexate, leflunomide, or a combination of DMARD and corticosteroid therapy.
‡‡‡Protocol for upcoming single-arm study.
AS, ankylosing spondylitis; CD, Crohn's disease; DMARD, disease-modifying antirheumatic drug; ECA, external control arm; IBD, inflammatory bowel disease; JIA, juvenile idiopathic arthritis; NR, not reported; NSAID, non-steroidal anti-inflammatory drug; PsO, psoriasis; RA, rheumatoid arthritis; RCT, randomised controlled trial; SOC, standard of care; T2T, treat-to-target; TDM, therapeutic drug monitoring; TNF, tumour necrosis factor; UC, ulcerative colitis.

and juvenile idiopathic arthritis (9/33; 27%), ankylosing spondylitis (4/33; 12%), psoriasis (2/33; 6%), and chronic refractory pouchitis (1/33; 3%) were evaluated. Four of the 33 single-arm trials (12%) administered treatment to a study population comprising participants with differing IMIDs. Pharmacological interventions were most frequently assessed (21/33; 64%). These included biologics (9/33; 31%), biosimilars (4/33; 12%), corticosteroids (2/33; 6%), immunosuppressives (3/33; 9%), antibiotics (1/33; 3%) and 5-aminosalicylic acid (1/33; 3%). The remaining 12 single-arm trials investigated a dosing strategy (4/33; 12%), treatment algorithm (2/33; 6%), surgical procedure (3/33; 9%), rapid infusion protocol (2/33; 6%), or faecal microbiota transplantation (1/33; 3%).

Unpublished historical case series of patients previously treated at the same site(s) as active arm participants was the most common ECA source (17/33; 52%) (table 3). External data were also derived from 1 or more prior phase 3–4 RCTs (6/33; 18%), product package inserts (2/33; 6%), published case series (2/33; 6%), disease registries (2/33; 6%), administrative claims databases (2/33; 6%), and previously conducted cohort studies (2/33; 6%). Two studies (2/33; 6%) pooled participant data from several different types of sources. The ECAs formed in single-arm trials comprised participants who had received standard of care (20/33; 61%), an active control therapy (10/33; 30%), or placebo (3/33; 9%).

External control methodology

All 33 included single-arm trials used a historical control, rather than a contemporaneous ECA. In most studies, external datasets were not statistically adjusted for known covariates to improve the comparability of external and target study populations (26/33; 79%) (table 4). Five of the 33 studies applied a matching methodology (15%), including propensity score matching (2/33; 6%), and 1 study employed a propensity score weighting (3%) approach to improve the comparability of active and external groups. Twenty-four of the included single-arm trials (73%) evaluated one or more efficacy outcomes. These studies assessed dichotomous (20/33; 60%) and continuous (7/33; 21%) efficacy endpoints based on an indication-specific disease activity index (18/33; 54.5%). Safety outcomes were compared with external data in 15 studies (45.4%). These included the incidence of adverse events (10/33; 30%), adverse infusion reactions (3/33; 9%), postsurgical complications (1/33; 3%) and serious adverse events (1/33; 3%).

Indication for use of an ECA

In 9 out of 33 single-arm trials (27%),[32 36 37 43–46 51 52] it was considered unethical to randomise participants to an internal control arm due to the availability of approved therapies with established efficacy and safety (7/33; 21%), long-term assessment of treatment outcomes (6/33; 18%), or paediatric study population (2/33; 6%) (table 4). Six trials (6/33; 18%)[33 35 40 42 55 57] cited practical challenges as the rationale for using an ECA. In five of these cases (5/33; 15%), an adequately powered RCT would reportedly not have been feasible due to the relatively small patient population available for study participation.

Quality assessment

The results of the methodological quality assessment are reported in online supplemental file 1. Of the 210 (43 studies × 5 domains) ratings, 99 (46%) were high quality, 113 (52%) were unclear quality and 3 (1%) were low quality with respect to the source used to generate the ECA. The domains that were rated for each study included the quality of the data source (40% high quality, 60% unclear quality), similarity of study populations (43% high quality, 57% unclear quality), similarity of outcome definitions (72% high quality, 21% unclear quality, 7% low quality), and reliability and comprehensiveness of the external dataset (21% high quality, 79% unclear quality).

DISCUSSION

In recent years, there has been increased interest among researchers, regulators and industry partners in using external control data to address the practical limitations associated with traditional RCT designs. This systematic review aimed to summarise how ECAs are currently being integrated into studies of Crohn's disease, ulcerative colitis, chronic refractory pouchitis, rheumatoid arthritis, juvenile idiopathic arthritis, ankylosing spondylitis, psoriatic arthritis, psoriasis and atopic dermatitis, and examine the methodological quality of this research. We identified 43 trials in which an ECA was used to contextualise the efficacy or safety of a medical therapy. No trials of atopic dermatitis were identified. Most of the included clinical trials were single-arm studies, and all used historical control data to form an ECA. Among included RCTs, external data were most often sourced data from multiple previously conducted RCTs. Conversely, external controls comprised participants who previously received treatment at the same site as the active arm population was the most frequent source of the ECA when the clinical trial was a single-arm design. In most cases, no adjustment methods were used to balance baseline characteristics between the treatment arm and historical data used to form the ECA. We did not identify any specific characteristics of data sources, outcome measure selection, analysis method or ECA type that were associated with a positive or negative study result.

All primary phase 3 trials of IMIDs identified in this systematic review that were submitted to or relied on by regulatory agencies for decision-making were in the paediatric setting (4/43; 9%), where an internal placebo group may have been deemed unethical by stakeholders. In a phase 3 trial that supported FDA and EMA approval of infliximab in paediatric ulcerative colitis, Hyams *et al*[20] created a historical control group using pooled placebo data from prior infliximab RCTs in adults. A positive

Table 4 External control arm methodology in single-arm studies

Study ID	Indication	Statistical methods for generating ECA	Outcome(s) compared with ECA	Outcome timepoint (weeks)	Evidence of treatment benefit compared with ECA*	Rationale for using an ECA
Baker and Jewell, 1989[29]	IBD	Unadjusted	Efficacy (severe attack rate)	6	No	Well-characterised disease course
Rosenberg *et al*, 1990[30]	UC	Unadjusted	Efficacy (response and surgical rates)	>1	No	NR
Stenger *et al*, 1998[31]	RA	Unadjusted	Efficacy (cumulative CRP-AUC, median change in Sharp-score) Safety (AEs)	104	Yes	NR
Marotte *et al*, 2007[32]	RA	Matching	Efficacy (DAS28 response rate)	52	Yes	Unethical to use internal placebo arm
Shen *et al*, 2007[34]	Pouchitis	Unadjusted	Efficacy (PDAI clinical remission and response rates) Safety (AEs)	4	Yes	NR
Moul *et al*, 2007[33]	PsO	Unadjusted	Efficacy (PASI 50 and PASI 75 response rates)	12	No	Small study sample size
van der Heijde *et al*, 2008[36]	AS	Matching	Efficacy (median change in mSASSS)	104	No	Unethical to use internal placebo arm
Boyle *et al*, 2008[35]	IBD	Matching	Safety (LOS, postoperative hospitalizations)	LOS	Yes	Impractical to use internal placebo
van der Heijde *et al*, 2009[37]	AS	Unadjusted	Efficacy (median change in mSASSS)	104	No	Unethical to use internal placebo arm
Garg *et al*, 2011[38]	IBD	Unadjusted	Safety (AEs)	>1	NA	NR
Yong *et al*, 2011†[39]	PsO	NR	Safety (SAEs)	NR	NA	ECA methodology study
Carter *et al*, 2012[40]	RA	Unadjusted	Safety (acute infusion reactions)	>1	Yes	Small study sample size
Kwok and Leung, 2012[41]	RA	Unadjusted	Efficacy (DAS20 and ACR response rate)	52	Yes	NR
Baraliakos *et al*, 2014[43]	AS	Unadjusted	Efficacy (mean change in mSASSS)	8 years	No	Unethical to use internal placebo arm
Horneff *et al*, 2014[44]	Multiple‡	Unadjusted	Efficacy (JIA ACR 30 response rate)	12	Yes	Unethical to use internal placebo arm
Lichtenstein *et al*, 2014[45]	UC	Unadjusted	Safety (AEs)	104	Yes	Unethical to use internal placebo arm
Gardenbroek *et al*, 2015[42]	UC	Unadjusted	Efficacy (anastomotic healing, closure of anastomotic defect and function pouch rates)	24	Yes	Impractical to use internal placebo
Mahajan *et al*, 2018[48]	UC	Unadjusted	Efficacy (clinical, endoscopic, and steroid-free clinical remission rates) Safety (AEs)	22	Yes	NR
Rozette *et al*, 2018[49]	Multiple §	Unadjusted	Safety (adverse infusion reactions)	> 1	NA	NR
Tweehuysen *et al*, 2018[50]	Multiple¶	Unadjusted	Efficacy (CRP and DAS28-CRP response rates) Safety (AEs)	24	No	NR

Continued

Table 4 Continued

Study ID	Indication	Statistical methods for generating ECA	Outcome(s) compared with ECA	Outcome timepoint (weeks)	Evidence of treatment benefit compared with ECA*	Rationale for using an ECA
Horneff *et al*, 2018[46]	JIA	Unadjusted	Efficacy (JADAS10 response rate) Safety (AEs)	NR	NA	Unethical to use internal placebo arm
Inui *et al*, 2018[47]	RA	Unadjusted	Efficacy (DAS28-ESR and Sharp score response rates)	52, 104	No	NR
Braun *et al*, 2019[51]	AS	Unadjusted	Efficacy (mean change in mSASSS)	104	No	Unethical to use internal placebo arm
Kawashiri *et al*, 2020**[52]	RA	Unadjusted	Efficacy (DAS28-ESR clinical relapse rate)	24	NA	HESDE previously established
Muskens *et al*, 2020[53]	Multiple¶	Unadjusted	Safety (AEs)	52	NA	NR
Sanchez-Hernandez *et al*, 2020[54]	IBD	Unadjusted	Efficacy (treatment failure and IBD-related surgery and hospitalisation rates) Safety (AEs, acute infusion reactions)	156	Yes	NR
Sagami *et al*, 2021[57]	IBD	Unadjusted	Safety (AEs)	104	NA	Impractical to use internal placebo arm
Gasparetto *et al*, 2021[56]	IBD	Unadjusted	Efficacy (dose escalation rate)	52	Yes	NR
Danese *et al*, 2021[55]	UC	Propensity score weighting	Efficacy (Mayo endoscopic score)	14	Yes	Impractical to use internal placebo arm; low recruitment
Holland *et al*, 2022[59]	CD	Matching	Efficacy (mPCDAI response rate)	26, 52, 78	No	NR
Huang *et al*, 2022[60]	RA	Propensity score matching	Efficacy (SDAI, RA-CDAI and DAS28 response rates)	12, 26, 52, 104	Yes	NR
Truong *et al*, 2022[58]	IBD	Unadjusted	Safety (postoperative complications)	4	No	Lack of large single-institution studies
Takei *et al*, 2022[61]	RA	Propensity score matching	Efficacy (ACR20, DAS28-ESR, SDAI response rates)	24	Yes	NR

*As defined by the included study.
†Primary study objective was to examine methodological approaches for complementing prior or concurrent RCT data with external control data.
‡Juvenile idiopathic arthritis, enthesitis-related arthritis and psoriatic arthritis.
§Crohn's disease, ulcerative colitis and juvenile idiopathic arthritis
¶Rheumatoid arthritis, psoriatic arthritis and ankylosing spondylitis.
**Protocol for an upcoming single-arm study with an ECA.
ACR, American College of Rheumatology; AE, adverse event; AS, ankylosing spondylitis; AUC, area under the curve; CD, Crohn's disease; CRP, C reactive protein; DAS, disease activity score; ECA, external control arm; ESR, erythrocyte sedimentation rate; HESDE, historical evidence of sensitivity to drug effects; IBD, inflammatory bowel disease; JADAS, Juvenile Arthritis Disease Activity Score; JIA, juvenile idiopathic arthritis; LOS, length of stay; mPCDAI, modified Paediatric Crohn's Disease Activity Index; mSASSS, modified Stoke Ankylosing Spondylitis Spinal Score; NA, not applicable; NR, not reported; PASI, Psoriasis Area Severity Index; PDAI, Pouchitis Disease Activity Index; PsO, psoriasis; RA, rheumatoid arthritis; RA-CDAI, Rheumatoid Arthritis Clinical Disease Activity Index; SAE, serious adverse event; SDAI, Simplified Disease Activity Index; UC, ulcerative colitis.

result was determined by using the upper limit of the 95% CI for the primary endpoint in the adult trials to define the lower 95% CI threshold in the paediatric trial. Croft *et al*[22] assessed the efficacy and safety of adalimumab in a phase 3 study of paediatric ulcerative colitis. This study encountered recruitment delays due to poor acceptance of the placebo group and, following consultation with the FDA and EMA, a meta-analysis of historical adult placebo

groups was used to replace the internal control. Adalimumab subsequently received regulatory approval in this patient population.[22] In indications where historical paediatric data may be limited, leveraging prior adult placebo data may be acceptable and advantageous to shorten the time between regulatory approval of advanced therapies in adult and paediatric patients.

In a phase 3 study that supported FDA and EMA approval of secukinumab in paediatric psoriasis, Magnolo *et al*[23] adopted a Bayesian dynamic borrowing approach using historical adult and paediatric placebo data from previous RCTs. Low-dose and high-dose secukinumab were found to be superior to historical placebo with respect to the coprimary and key secondary outcomes, with an estimated probability of a positive treatment effect of 100% compared with external placebo for both dosing regimens. Finally, Horneff *et al*[44] conducted a single-arm phase 3b study to expand EMA approval to three under-studied categories of juvenile idiopathic arthritis (extended oligoarticular juvenile idiopathic arthritis, enthesitis-related arthritis, and psoriatic arthritis). Given the paediatric population, and well-established efficacy and safety of etanercept in polyarticular-course juvenile idiopathic arthritis, it was deemed unethical to randomise participants to receive placebo. Although etanercept response rates were statistically superior compared with the ECA, external placebo data were primarily sourced from a meta-analysis of paediatric RCTs that had been published nearly a decade prior.

If the historical ECA and target study sample substantially differ due to population drift, changes in eligibility criteria, or evolving standards of care, treatment effects may be misestimated. The primary method for mitigating estimation bias is careful selection of the studies from which the historical ECA is drawn. Statistical methods can also be used to adjust for imbalances between historical and target study samples. Propensity score matching involves performing regression on a set of prespecified baseline covariates to determine a propensity score for each participant. This enables the creation of matched sets of historical and target study participants with similar propensity score values.[55 62] However, propensity score matching and other methods of balancing that use prespecified covariates have important limitations. First, confounders must be correctly identified and accurately measured. When matching is based on *observed* covariates, the historical ECA may not be representative of the parent population and therefore not suitable for comparison with the target study.[62] Second, propensity score matching may result in 'pruning', where individuals with extreme covariate values have no suitable matches and are systematically excluded from the ECA, which can increase imbalance and bias.[63] With Bayesian dynamic borrowing, data are weighted based on the level of discrepancy between the historical control and target study.[23 64] When correctly performed, dynamic borrowing can provide additional information and augment sample size with a sufficient degree of scepticism to reduce bias. As with propensity score matching, dynamic borrowing requires correct specification of potential confounders to determine whether the ECA and target study data are closely matched. Evidence suggests that dynamic borrowing may carry an inflated chance of type I error; however, the benefits of additional power and precision may outweigh this risk.[65]

When evaluating the data collection process, similarity of the populations, similarity of the outcome definitions, the reliability and comprehensiveness of the datasets, and other limitations, only 3 of the 215 items (1%) assessed were rated as low quality. However, half of the items (113/215; 52.5%) could not be assigned a definitive low-quality or high-quality rating due to insufficient detail reported in the study publication. The included studies frequently failed to describe in detail the eligibility criteria and baseline characteristics of the active arm and ECA, outcome definitions, covariates, whether there were missing data, and the analytical methods employed to balance the cohort.

Comprehensive reporting is therefore necessary to critically appraise the quality of ECA studies, particularly when real-world data are used. This is underscored by the external control quality checklist proposed by Thorlund *et al*[18] and in draft guidance by the FDA and EMA which state that adequate documentation during data curation and transformation is essential to increase confidence in the resultant data.[11 13] The FDA has additionally issued draft guidance outlining key considerations for designing externally-controlled clinical trials, selecting fit-for-use ECA data sources, analysing comparisons and accounting for potential bias.[2] While this document does not recommend specific statistical methodologies for making comparisons to external data, it is noted that underlying assumptions should be identified and further examined using sensitivity analyses and model diagnostics. The need to assess the impact of missing and misclassified data in the ECA, which may be more common when using real-world data sources, is emphasised.[66]

Moreover, objective and reliable measurements for the data of interest are recommended to mitigate bias due to misclassified data and lack of blinding. Within the context of IMIDs, this may involve restricting outcome measures to the most objective item(s) included in composite disease active indices. For example, centrally read endoscopic improvement, as defined by the Mayo Endoscopic subscore, was selected as the primary outcome in a statistically significant study by Danese *et al*[55] evaluating a novel biologic in a phase 2 ulcerative colitis trial. Trial patients were matched with historical placebo patients from prior RCTs using propensity score analysis, and placebo endoscopic response rates were used to determine the null hypothesis.

Regulatory submissions are evaluated on a case-by-case basis, and approvals are often conditional on additional evidence from postapproval studies.[14] Our results suggest that comprehensive ECA guidelines may be particularly beneficial for IMID research. The patient populations,

standards of care, outcome definitions and study timepoints have substantially evolved over time in IMIDs, and all studies identified in the current systematic review included historical controls. Ensuring that appropriate source data and optimal statistical methodologies are selected may therefore be of particular importance within this context.

Our study has some important strengths. We examined how external controls have been implemented in IMID trials using a predefined systematic approach. To our knowledge, this is the first systematic review on the use of ECAs in these populations. This qualitative summary may help inform ECA data source selection in future externally controlled trials of IMIDs and help in trial design where a treatment arm might be useful in the future as an ECA. However, the limitations of this study should also be acknowledged. First, while we aimed to devise a highly sensitive search strategy using text words and controlled vocabulary, the language used to describe ECA methodology is varied, and ECA-specific MeSH terms do not exist. In an effort to ensure that all eligible studies were identified, bibliographies of relevant review articles were also searched. Nine of the 2192 records screened were identified using this approach. Furthermore, only published studies were searched and included in this systematic review. Second, the quality of included studies was rated using an instrument that was not developed and validated to appraise ECA methodology, as no such instrument currently exists. Our quality assessment exclusively focused on the source of the external control data. We did not evaluate the methods used to generate external controls given that trials were eligible for inclusion regardless of whether they used trial-level or patient-level data.

In conclusion, we found that external control data have been applied in a variety of IMID settings to contextualise efficacy and safety outcomes. The reporting of the methodology used to generate and analyse ECAs has been incomplete and heterogeneous. The establishment of authoritative reporting guidelines may serve as a catalyst for transparent reporting and rigorous study design for ECA studies in IMIDs.

Author affiliations
[1]Alimentiv Inc, London, Ontario, Canada
[2]Division of Gastroenterology, University of California, San Diego, La Jolla, California, USA
[3]Division of Biomedical Informatics, University of California, San Diego, La Jolla, California, USA
[4]Department of Medicine, Division of Gastroenterology, Western University, London, Ontario, Canada
[5]Department of Epidemiology and Biostatistics, Western University, London, Ontario, Canada
[6]Division of Gastroenterology and Hepatology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[7]Department of Community Health Sciences, University of Calgary, Calgary, Alberta, Canada

Contributors Study concept and design: VJ. Acquisition, analysis and interpretation of data: AZ, CP, JC, JKM, and RE. Draft of initial manuscript: AZ, BN, CP and RE. Critical revision of the manuscript for important intellectual content: AZ, BGF, BN, CP, CM, GZ, JC, JKM, RE, SS and VJ. Guarantor of article: VJ. All authors have made substantial contributions to the manuscript and approved the final version of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests AZ is an employee of Alimentiv. RE has received consulting fees from Alimentiv, CADTH, SpringWorks Therapeutics, and Turnstone Biologics. CEP is an employee of Alimentiv. JKM is an employee of Alimentiv. JC is a former employee of Alimentiv. BN is an employee of Alimentiv. GZ is an employee of Alimentiv. SS has received personal fees from Pfizer for ad hoc grant review, and his institute has received research grants from AbbVie and Janssen. BFG has received grant/research support from AbbVie, Amgen, AstraZeneca/MedImmune, Atlantic Pharmaceuticals, Boehringer-Ingelheim, Celgene Corporation, Celltech, Genentech Inc./Hoffmann-La Roche, Gilead Sciences, GSK, Janssen Research & Development, Pfizer, Receptos/Celgene International, Sanofi, Santarus, Takeda Development Center Americas, Tillotts, and UCB; consulting fees from AbbVie, Akebia Therapeutics, Allergan, Amgen, AMT, Aptevo Therapeutics, Astra Zeneca, Atlantic Pharma, Avir Pharma, Biogen Idec, BioMx Israel, Boehringer-Ingelheim, Bristol-Myers Squibb, Calypso Biotech, Celgene, Elan/Biogen, EnGene, Ferring Pharma, Roche/Genentech, Galapagos, GiCare Pharma, Gilead, Gossamer Pharma, GSK, Inception IBD, Johnson & Johnson/Janssen, Kyowa Kakko Kirin Co, Lexicon, Lilly, Lycera BioTech, Merck, Mesoblast Pharma, Millennium, Nestle, Nextbiotix, Novo Nordisk, Pfizer, Prometheus Therapeutics and Diagnostics, Progenity, Protagonist, Receptos, Salix Pharma, Shire, Sienna Biologics, Sigmoid Pharma, Sterna Biologicals, Synergy Pharma, Takeda, Teva Pharma, TiGenix, Tillotts, UCB, Vertex Pharma, Vivelix Pharma, VHsquared, and Zyngenia; speaker's bureau fees from Abbott/AbbVie, JnJ/Janssen, Lilly, Takeda, Tillotts, and UCB Pharma; is a scientific advisory board member for Abbott/AbbVie, Allergan, Amgen, Astra Zeneca, Atlantic Pharma, Avaxia Biologics, Boehringer-Ingelheim, Bristol-Myers Squibb, Celgene, Centocor, Elan/Biogen, Galapagos, Genentech/Roche, JnJ/Janssen, Merck, Nestle, Novartis, Novo Nordisk, Pfizer, Prometheus Laboratories, Protagonist, Salix Pharma, Sterna Biologicals, Takeda, Teva, TiGenix, Tillotts, and UCB; and is the Senior Scientific Officer of Alimentiv. CM has received consulting fees from AbbVie, Alimentiv, Amgen, AVIR Pharma, BioJAMP, Bristol Myers Squibb, Celltrion, Ferring, Fresenius Kabi, Janssen, McKesson, Mylan, Takeda, Pendopharm, Pfizer, Prometheus Biosciences, Roche, Sanofi; speaker's fees from AbbVie, Amgen, AVIR Pharma, Alimentiv, Bristol Myers Squibb, Ferring, Fresenius Kabi, Janssen, Organon, Pendopharm, Pfizer, Takeda; royalties from Springer Publishing; research support from Ferring, Pfizer. VJ has received has received consulting/advisory board fees from AbbVie, Alimentiv, Arena pharmaceuticals, Asahi Kasei Pharma, Asieris, Astra Zeneca, Bristol Myers Squibb, Celltrion, Eli Lilly, Ferring, Flagship Pioneering, Fresenius Kabi, Galapagos, GlaxoSmithKline, Genentech, Gilead, Janssen, Merck, Metacrine, Mylan, Pandion, Pendopharm, Pfizer, Protagonist, Prometheus, Reistone Biopharma, Roche, Sandoz, Second Genome, Sorriso pharmaceuticals, Takeda, Teva, Topivert, Ventyx, Vividion; speaker's fees from, Abbvie, Ferring, Bristol Myers Squibb, Galapagos, Janssen Pfizer Shire, Takeda, Fresenius Kabi.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD
Vipul Jairath http://orcid.org/0000-0001-8603-898X

REFERENCES

1 Bothwell LE, Greene JA, Podolsky SH, *et al*. Assessing the gold standard–lessons from the history of Rcts. *N Engl J Med* 2016;374:2175–81.
2 FDA. Considerations for the design and conduct of externally controlled trials for drug and biological products: Guidance for industry, 2023. Available: https://www.fda.gov/media/164960/download
3 EMA. Ich topic E 10. choice of control group in clinical trials, 2001. Available: https://www.ema.europa.eu/documents/scientific-guideline/ich-e-10-choice-control-group-clinical-trials-step-5_en.pdf
4 ICH. Choice of control group and related issues in clinical trials E10, 2000. Available: https://www.ema.europa.eu/documents/scientific-guideline/ich-e-10-choice-control-group-clinical-trials-step-5_en.pdf
5 EMA. Guideline on clinical trials in small populations, 2006. Available: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-trials-small-populations_en.pdf
6 FDA. Submitting documents using real-world data and real-world evidence to FDA for drugs and biologics: guidance for industry, 2019. Available: https://www.fda.gov/media/124795/download
7 FDA. Rare diseases: natural history studies for drug development guidance for industry, 2019. Available: https://www.fda.gov/media/122425/download
8 Hatswell AJ, Baio G, Berlin JA, *et al*. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014. *BMJ Open* 2016;6:e011666.
9 FDA. Real-world data: assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products draft guidance for industry, 2021. Available: https://www.fda.gov/media/152503/download
10 FDA. Data standards for drug and biological product submissions containing real-world data–draft guidance for industry, 2021. Available: https://www.fda.gov/media/153341/download
11 FDA. Considerations for the use of real-world data and real-world evidence to support regulatory decision-making for drug and biological products–draft guidance for industry, 2021. Available: https://www.fda.gov/media/154714/download
12 FDA. Real-world data: assessing registries to support regulatory decision-making for drug and biological products-guidance for industry, 2021. Available: https://www.fda.gov/media/154449/download
13 EMA. Real-World evidence framework to support EU regulatory decision-making, 2023. Available: https://www.ema.europa.eu/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained_en.pdf
14 Goring S, Taylor A, Müller K, *et al*. Characteristics of non-randomised studies using comparisons with external controls submitted for regulatory approval in the USA and Europe: a systematic review. *BMJ Open* 2019;9:e024895.
15 Jahanshahi M, Gregg K, Davis G, *et al*. The use of external controls in FDA regulatory decision making. *Ther Innov Regul Sci* 2021;55:1019–35.
16 Johnson C, Barnes EL, Zhang X, *et al*. Trends and characteristics of clinical trials participation for inflammatory bowel disease in the United States: a report from IBD partners. *Crohns Colitis 360* 2020;2.
17 Page MJ, McKenzie JE, Bossuyt PM, *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
18 Thorlund K, Dron L, Park JJH, *et al*. Synthetic and external controls in clinical trials - A primer for researchers. *Clin Epidemiol* 2020;12:457–67.
19 Hueber W, Sands BE, Lewitzky S, *et al*. Secukinumab, a human anti-IL-17A Monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. *Gut* 2012;61:1693–700.
20 Hyams J, Damaraju L, Blank M, *et al*. Induction and maintenance therapy with Infliximab for children with moderate to severe ulcerative colitis. *Clin Gastroenterol Hepatol* 2012;10:391–9.
21 Hodkinson B, Monyai N, Tikly M. Thu0128 the utility of clinical disease activity index for tight control of disease activity in South African rheumatoid arthritis patients. *Ann Rheum Dis* 2013;72:A205.
22 Croft NM, Faubion WA, Kugathasan S, *et al*. Efficacy and safety of Adalimumab in Paediatric patients with moderate-to-severe ulcerative colitis (ENVISION I): a randomised, controlled, phase 3 study. *Lancet Gastroenterol Hepatol* 2021;6:616–27.
23 Magnolo N, Kingo K, Laquer V, *et al*. A phase 3 open-label, randomized multicenter study to evaluate efficacy and safety of Secukinumab in pediatric patients with moderate to severe plaque psoriasis: 24-week results. *J Am Acad Dermatol* 2022;86:122–30.
24 Reed G, Schrader A, Madsen A. Methods and practical considerations in use of an external control to assess potential drug risk. *Pharmacoepidemiol Drug Saf* 2021;30:273.
25 Hupf B, Bunn V, Lin J, *et al*. Bayesian Semiparametric meta-analytic-predictive prior for historical control borrowing in clinical trials. *Stat Med* 2021;40:3385–99.
26 Rudrapatna VA, Cheng Y-W, Feuille C, *et al*. Creation of a Ustekinumab external control arm for Crohn's disease using electronic health records data: A pilot study. *PLOS One* 2023;18:e0282267.
27 Wang Z, Yu Z, Zhang L. Investigating synthetic controls with randomized clinical trial data in rheumatoid arthritis studies. *J Clin Trials* 2021;11:1–9.
28 Norvang V, Haavardsholm EA, Tedeschi SK, *et al*. Using observational study data as an external control group for a clinical trial: an empirical comparison of methods to account for longitudinal missing data. *BMC Med Res Methodol* 2022;22:152.
29 Baker K, Jewell DP. Cyclosporin for the treatment of severe inflammatory bowel disease. *Aliment Pharmacol Ther* 1989;3:143–9.
30 Rosenberg W, Ireland A, Jewell DP. High-dose methylprednisolone in the treatment of active ulcerative colitis. *J Clin Gastroenterol* 1990;12:40–1.
31 Stenger AA, Van Leeuwen MA, Houtman PM, *et al*. Early effective suppression of inflammation in rheumatoid arthritis reduces radiographic progression. *Br J Rheumatol* 1998;37:1157–63.
32 Marotte H, Pallot-Prades B, Grange L, *et al*. A 1-year case-control study in patients with rheumatoid arthritis indicates prevention of loss of bone mineral density in both responders and Nonresponders to Infliximab. *Arthritis Res Ther* 2007;9:R61.
33 Moul DK, Routhouska SB, Korman NJ. Open-label, single-center, safety dose escalation trial of Alefacept for the treatment of moderate to severe chronic plaque psoriasis. *J Cutan Med Surg* 2007;11:132–6.
34 Shen B, Fazio VW, Remzi FH, *et al*. Combined ciprofloxacin and tinidazole therapy in the treatment of chronic refractory pouchitis. *Dis Colon Rectum* 2007;50:498–508.
35 Boyle E, Ridgway PF, Keane FB, *et al*. Laparoscopic Colonic resection in inflammatory bowel disease: minimal surgery, minimal access and minimal hospital stay. *Colorectal Dis* 2008;10:911–5.
36 van der Heijde D, Landewé R, Baraliakos X, *et al*. Radiographic findings following two years of Infliximab therapy in patients with Ankylosing Spondylitis. *Arthritis Rheum* 2008;58:3063–70.
37 van der Heijde D, Salonen D, Weissman BN, *et al*. Assessment of radiographic progression in the spines of patients with Ankylosing Spondylitis treated with Adalimumab for up to 2 years. *Arthritis Res Ther* 2009;11:R127.
38 Garg M, Morrison G, Friedman A, *et al*. A rapid infusion protocol is safe for total dose iron Polymaltose: time for change. *Intern Med J* 2011;41:548–54.
39 Yong M, Schenfeld J, Creamer K. External comparator data for the observational post-marketing safety surveillance registry of Enbrel (etanercept) for the treatment of psoriasis-A single-arm registry study. *Pharmacoepidemiol Drug Saf* 2011:S228.
40 Carter JD, Zarabadi SA, Ricca LR, *et al*. A safety analysis of oral prednisone as a pretreatment for Rituximab in rheumatoid arthritis. *Clin Rheumatol* 2012;31:1605–10.
41 Kwok KY, Leung MH. Tight control early rheumatoid arthritis clinic in Hong Kong: a pilot study. *Hong Kong Med J* 2012;18:108–14.
42 Gardenbroek TJ, Musters GD, Tanis PJ, *et al*. Sa1138 early closure of anastomotic leakage after ileal pouch-anal anastomosis: a novel solution to an old problem. *Gastroenterology* 2013;144:S-211–S-212.
43 Baraliakos X, Haibel H, Listing J, *et al*. Continuous long-term anti-TNF therapy does not lead to an increase in the rate of new bone formation over 8 years in patients with Ankylosing Spondylitis. *Ann Rheum Dis* 2014;73:710–5.
44 Horneff G, Burgos-Vargas R, Constantin T, *et al*. Efficacy and safety of open-label Etanercept on extended Oligoarticular juvenile idiopathic arthritis, Enthesitis-related arthritis and Psoriatic arthritis: part 1 (week 12) of the CLIPPER study. *Ann Rheum Dis* 2014;73:1114–22.
45 Lichtenstein GR, Barrett AC, Bortey E, *et al*. Long-term safety and tolerability of once-daily Mesalamine granules in the maintenance of remission of ulcerative colitis. *Inflamm Bowel Dis* 2014;20:1399–406.
46 Horneff G, Klein A, on behalf of BIKER registry working group. Ab1096 observational safety study of golimumab in treatment of polyarticular juvenile idiopathic arthritis using the German biologics

JIA registry. Annual European Congress of Rheumatology, EULAR 2018, Amsterdam, 13–16 June 2018, 2018.
47 Inui K, Koike T, Tada M, *et al*. Clinical and Radiologic analysis of on-demand use of Etanercept for disease flares in patients with rheumatoid arthritis for 2 years: the RESUME study: A case-control study. *Medicine (Baltimore)* 2018;97:e12462.
48 Mahajan R, Midha V, Mehta V, *et al*. Iddf2018-Abs-0218 efficacy of Faecal Microbiota therapy in patients with steroid dependent active ulcerative colitis. International Digestive Disease Forum (IDDF) 2018, Hong Kong, 9–10 June 2018; 2018
49 Rozette NA, Hellauer CM, McKee C, *et al*. Evaluation of rapid vs standard Infliximab infusions in the pediatric population. *Inflamm Bowel Dis* 2018;24:2007–14.
50 Tweehuysen L, Huiskes VJB, van den Bemt BJF, *et al*. Open-label, non-mandatory Transitioning from Originator Etanercept to Biosimilar Sb4: six-month results from a controlled cohort study. *Arthritis Rheumatol* 2018;70:1408–18.
51 Braun J, Haibel H, de Hooge M, *et al*. Spinal radiographic progression over 2 years in Ankylosing Spondylitis patients treated with Secukinumab: a historical cohort comparison. *Arthritis Res Ther* 2019;21:142.
52 Kawashiri S-Y, Shimizu T, Sato S, *et al*. Switching from Originator Infliximab to Biosimilar Infliximab in Japanese patients with rheumatoid arthritis achieving clinical remission (the IFX-SIRIUS study I): study protocol for an Interventional, multicenter, open-label, single-arm and Noninferiority clinical trial with clinical, ultrasound, and biomarker assessments. *Medicine (Baltimore)* 2020;99:e21151.
53 Müskens WD, Rongen-van Dartel SAA, Teerenstra S, *et al*. One-year results after Transitioning from Etanercept Originator to Biosimilar in a setting promoting shared decision-making in rheumatology. *Rheumatol Adv Pract* 2020;4:rkaa042.
54 Sánchez-Hernández JG, Rebollo N, Martin-Suarez A, *et al*. A 3-year prospective study of a Multidisciplinary early Proactive therapeutic drug monitoring programme of Infliximab treatments in inflammatory bowel disease. *Br J Clin Pharmacol* 2020;86:1165–75.
55 Danese S, Klopocka M, Scherl EJ, *et al*. Anti-TI1A antibody PF-06480605 safety and efficacy for ulcerative colitis: A phase 2A single-arm study. *Clin Gastroenterol Hepatol* 2021;19:2324–32.
56 et alGasparetto M, Burgess N, Studart D. P07 a single-centre prospective observational study comparing proactive with reactive therapeutic drug monitoring (TDM) in two cohorts of children with inflammatory bowel disease (IBD). Abstracts of the BSPGHAN Virtual Annual Meeting, 27–29 April 2021, 2021.
57 Sagami S, Nishikawa K, Yamada F, *et al*. Post-marketing analysis for Biosimilar CT-P13 in inflammatory bowel disease compared with external data of Originator Infliximab in Japan. *J Gastroenterol Hepatol* 2021;36:2091–100.
58 Truong A, Wood T, Fleshner PR, *et al*. A single-center experience of Transanal Proctectomy with Ileal pouch-Anal anastomosis for inflammatory bowel disease. *Dis Colon Rectum* 2022;65:1121–8.
59 Holland K, Bennett WE, Slaven JE, *et al*. Proactive measurement of Infliximab drug levels in children with Crohn's disease. *Ann Gastroenterol* 2022;35:56–62.
60 Huang H, Xie W, Geng Y, *et al*. Towards a better implementation of treat-to-target strategy in rheumatoid arthritis: A comparison of two real-world cohorts. *Rheumatol Ther* 2022;9:907–17.
61 Takei H, Takanashi S, Otomo K, *et al*. Clinical and immunological effects of hydroxychloroquine in patients with active rheumatoid arthritis despite Antirheumatic treatment. *Mod Rheumatol* 2022:roac153.
62 Streiner DL, Norman GR. The pros and cons of propensity scores. *Chest* 2012;142:1380–2.
63 King G, Nielsen R. Why propensity scores should not be used for matching. *Polit. Anal*. 2019;27:435–54.
64 Dron L, Golchi S, Hsu G, *et al*. Minimizing control group allocation in randomized trials using dynamic borrowing of external control data - an application to second line therapy for non-small cell lung cancer. *Contemp Clin Trials Commun* 2019;16:100446.
65 Harun N, Liu C, Kim MO. Critical appraisal of Bayesian dynamic borrowing from an imperfectly commensurate historical control. *Pharm Stat* 2020;19:613–25.
66 Wang X, Dormont F, Lorenzato C, *et al*. Current perspectives for external control arms in oncology clinical trials: analysis of EMA approvals 2016-2021. *J Cancer Policy* 2023;35:S2213-5383(23)00020-6.