
Article:
Morar, P.S., Faiz, O., Warusavitarne, J. et al. (2015) Systematic review with meta-analysis: Endoscopic balloon dilatation for Crohn's disease strictures. Alimentary Pharmacology and Therapeutics, 42 (10). 1137 - 1148. ISSN: 0269-2813

https://doi.org/10.1111/apt.13388

Alimentary Pharmacology & Therapeutics

A SYSTEMATIC REVIEW AND META-ANALYSIS - ENDOSCOPIC BALLOON DILATATION FOR CROHN'S DISEASE STRICTURES

Journal:	*Alimentary Pharmacology & Therapeutics*
Manuscript ID:	Draft
Wiley - Manuscript type:	Systematic Review with Meta-analysis
Date Submitted by the Author:	n/a
Complete List of Authors:	Morar, Pritesh; St Mark's Hospital, Colorectal Surgery and IBD; Imperial College, Surgery Faiz, Omar; St Mark's Hospital, Colorectal Surgery and IBD; Imperial College, Surgery Warusavitarne, Janindra; St Mark's Hospital, Colorectal Surgery and IBD; Imperial College, Surgery Brown, Steven; Royal Hallamshire Hospital, Cohen, Richard; University College Hospital, Dept of Surgery Hind, Daniel; University of Sheffield, Abercrombie, John; Queens Medical Centre, Ragunath, Krish; Queens Medical Centre, Sanders, David; University of Sheffield, Arnott, Ian; Western General Hospital, Wilson, Graeme; Western General Hospital, Bloom, Stuart; University College London Hospitals NHS Trust, Gastroenterology Arebi, Naila; St Mark's Hospital, Colorectal Surgery and IBD; St Mark's Hospital, Gastroenterology; Imperial College, Surgery
Keywords:	Crohn's disease < Disease-based, Inflammatory bowel disease < Disease-based, Colonoscopy < Topics, Endoscopy < Topics

A SYSTEMATIC REVIEW AND META-ANALYSIS - ENDOSCOPIC BALLOON DILATATION FOR CROHN'S DISEASE STRICTURES

Pritesh Morar [1 2], Omar Faiz [1 2], Janindra Warusavitarne [1 2], Steve Brown [3], Richard Cohen [4], David Hind [3], John Abercrombie [5], Krish Ragunath [5], David S. Sanders [3], Ian Arnott [6], Graem Wilson [6], Stuart Bloom [4], Naila Arebi [1 2] on behalf of Crohn's Stricture Study (CroSS) Group

1. Department of Surgery and Cancer, Imperial College, London, UK
2. St Mark's Hospital, London, UK
3. Royal Hallamshire Hospital, Sheffield, UK
4. University College Hospital, London, UK
5. Queen's Medical Centre, Nottingham, UK
6. Western General Hospital, Edinburgh, UK

Corresponding author: Dr Naila Arebi, St Mark's Hospital, Watford Road, Harrow London HA1 3UJ

naila.arebi@imperial.ac.uk; Tel: 020 8235 4084

Disclosures – None to declare

Keywords – "Crohn's disease", "strictures", "fibrosis", "endoscopic balloon dilatation"

Word count: 4216

ABSTRACT

Background

Endoscopic balloon dilatation is a recognised treatment for symptomatic Crohn's strictures. Several case studies report its short term and long term efficacy. A systematic analysis of the current literature is needed to define its overall efficacy and inform the design of future studies.

Aim

The primary objective was to examine symptomatic response, technical response and adverse events of endoscopic balloon dilatation. Stricture characteristics that may impact on outcome were also explored.

Methods

A systematic search strategy of COCHRANE, MEDLINE, EMBASE and OVID was performed. All original studies reporting outcomes of endoscopic balloon dilatation for Crohn's strictures in the adult population were included. Pooled event rates across studies were expressed with summative statistics. Heterogeneity across studies was assessed numerically.

Results

25 studies were included capturing 1089 patients and 2664 dilatations. The pooled event rates for symptomatic and technical response was 74.8% (95% CI: 69.9-79.3%; I2: 0%) and 90.6% (95% CI: 87.8-92.8%; I2:11.7%), respectively. The pooled event rates for complications and perforations was 6.4% (95% CI: 5.0-8.2; I2:4.0%) and 3% (95% CI: 2.2-4.0%; I2:0%), respectively. Inflammatory activity and anastomotic strictures may be associated with lower symptomatic response and inflammatory activity with higher perforation rate.

Conclusion

The efficacy and complication rates of endoscopic balloon dilatation treatment for symptomatic Crohn's strictures was higher than previously reported. Efficacy may be affected by active inflammation, previous

surgery and choice of outcome measure. Future studies should examine differential effects on stricture types using a clinically relevant outcome measure.

Prospero Registration Number: CRD42015015758

INTRODUCTION

Rationale

Strictures in Crohn's disease usually develop during the course of the disease [1 2] but in 5-27% of cases they are the presenting feature.[2-5] Strictures may also arise following surgery.[6] Both types may either be inflammatory or fibrotic or include both elements. The majority (98.8%) are found in the colon, ileo-colonic region, and ileum after 10 years of disease. [6]

The understanding of the pathogenesis of fibrosis in Crohn's disease is evolving. Chronic inflammation leads to thickening of the mucosa and narrowing of the gut lumen. [6-8] Thereafter, disruption in the normal extracellular matrix and irregular activity of fibroblasts contribute to an imbalance of collagen deposition. Anastomotic strictures on the other hand, develop through a combination of local and technical factors, such as bacterial stasis from postoperative narrowing of the lumen, high intraluminal pressures, or vascular compromise resulting in tissue ischemia, leaking, or infection which drive healing by tissue fibrosis. [9 10]

Environmental, genetic and serological factors are also implicated in the evolution of strictures. Smoking was associated with an increased rate of progression from inflammatory to stricturing disease in one study.[11] In separate studies, mutations in the NOD2 gene were associated with small bowel fibro-stenosing Crohn's Disease,[12] whereas the NOD2/CARD15 genotype was an independent risk factor for early surgical intervention due to strictures.[13] Antimicrobial antibodies are linked to complications in Crohn's disease but are not restricted to stricturing disease.[14] Further studies using animal models of intestinal fibrosis may offer further insights into the pathogenesis such as factors promoting of stricture progression, markers of early diagnosis and modulation of fibrosis pathway to arrest or reverse the process.[8 15]

Small bowel strictures have a greater impact on individuals than colonic strictures because of the narrower lumen and loss of absorptive surface. Prompt therapy and preservation of small bowel are key factors in the management of small bowel CD strictures. The ECCO consensus on management of Crohn's disease recommends resection, strictureplasty or balloon dilatation as alternatives after initial medical treatment for localised small bowel or ileo-colic disease.[16] In clinical practice, the therapeutic choice is determined by stricture characteristics: accessible, short and anastomotic strictures are best considered for endoscopic balloon dilatation whereas endoscopically inaccessible, multiple and >5cm in length are suited to surgical approaches. Strictureplasty and bowel resection, carry short term risks of anastomotic leak, wound complications and the possibility of stoma formation and long term risks from recurrent disease, reoperation and short bowel.[17 18] Balloon dilatation offers a more attractive option because of its ease of administration and low costs. There are risks associated with the dilatation procedure. In short term risks, inability to completely dilate, perforation and bleeding, whereas long term risks are related to disease recurrence which may warrant further dilatation or surgery.[19]

Several studies report outcomes of endoscopic balloon dilatation in Crohn's disease strictures.[20-44] These outcomes were collated in a systematic literature review published in 2007. The lack of pooled analysis of events rates and non-conformity with PRISMA guidelines are discernible weaknesses of the review.[19] Moreover, since then a further 12 studies have been published.

Objectives

This systematic review was performed to describe the outcomes of endoscopic balloon dilatation for Crohn's disease strictures to include additional studies. The primary aim was to examine the pooled incidence of clinical response, technical response and adverse events following endoscopic balloon dilatation for Crohn's strictures in adults. The secondary aim was to explore the impact of stricture characteristics on outcomes.

METHODOLOGY

Protocol and registration

The protocol for this study was registered on PROSPERO (CRD42015015758).

Eligibility Criteria

All original studies, from 1991 to 2014, reporting outcomes of endoscopic balloon dilatation for Crohn's disease intestinal strictures in the adult population (age ≥ 18) were included in the review. Randomised controlled trials, observational reports and case series with sample size more than 5 were all included. Case reports and studies reporting on multiple diagnoses were excluded from the review. Patients undergoing double balloon dilatation for deep seated intestinal strictures and children (age < 18) were more likely to require a general anaesthesia for the required intervention. Studies reporting exclusively on these were also excluded.

Information sources

A three step search strategy was employed. Initially a limited search was performed using PUBMED to identify keywords and index terms contained in the title or abstract. The second step involved an extensive search using all identified keywords and extensive terms. Studies were identified by searching the following databases: COCHRAINE, MEDLINE, EMBASE & OVID.

Search

The final search terms were ("Crohn's Disease" OR Crohn's OR stricture OR "Montreal B2") AND (endoscopy OR endoscopic OR ileocolonoscopy OR ileoscopic OR colonoscopy OR colonoscopic) AND ("balloon therapy" OR "balloon dilatation" OR balloon dilation" OR dilatation OR "balloon strictureplasty"). The final step was a hand search of reference lists and bibliographies from previously retrieved studies to identify further relevant trials.

Data collection process

The first reviewer (PM) screened the titles and abstracts that were identified in the search strategy. The papers were then evaluated by two reviewers (PM and NA) according to the eligibility criteria outlined above. Discrepancies were resolved by consensus between the two reviewers. Data from selected studies were extracted by the first reviewer and this was followed by a further, unblended, check by the second reviewer. Extracted data was entered into an Excel (Microsoft® software) database.

Data items

The following variables were extracted: study demographics (year and country of publication, study design, and sample size), nature of the stricture (stricture characteristics including location, activity as active or quiescent fibrotic, type as de novo or anastomotic, length and diameter), preoperative radiographic assessment, intervention technique (dilatation time, balloon dilator size and endoscopic accessibility), follow-up time period and outcome measures (symptomatic response, technical response, overall complication and perforation rates).

Risk of bias in individual studies

The quality of studies was assessed by using the Newcastle-Ottawa Scale. The quality of studies was evaluated by examining three items: patient selection, comparability and outcome (Table 1 Supplementary material).

Summary measures

Symptomatic response was defined by the resolution of symptoms, technical response by the passage of the scope following endoscopic balloon dilatation and adverse events by the proportion of patients who develop complications. Outcomes are expressed as pooled event rates (with 95% confidence interval limits), or as a proportion of the size of the population studied (patients), stricture numbers (strictures) and/or number of dilatation procedures performed (interventions).

Synthesis of results

Continuous numerical data is expressed as means (with standard deviations) or as medians (with range values). A per patient analysis was used to determine the cumulative proportion of patients within a group, per stricture analysis was used to determine the cumulative proportion of strictures within a group and a per dilatation analysis was used to determine the cumulative proportion of dilatations within a group. All three analyses were expressed as proportions and percentages. A per study analysis was used to assess pooled event rates across studies. The random effects model was used and results were expressed with forest plots and summative statistics.

Risk of bias across studies

Heterogeneity across studies was assessed visually with forest plots and numerically ($I^2 < 25\%$ indicates low heterogeneity). Evidence of publication bias was assessed visually using funnel plots. Comprehensive

Meta-analysis (CMA, Biostat, Inc.) programme was used. To assess the relationship of continuous variables on outcome, the pooled mean event rate for each outcome was transformed into a dichotomous form (less than or greater than the stated pooled mean event rate).

Additional analyses

To determine association between stricture characteristics and outcome subgroup analyses were performed. The pooled event rates and 95% confidence interval were expressed per outcome for each categorical variable (e.g. balloon diameter, duration of inflation, geography and pre-interventional imaging). The mean value of pooled outcomes was used to create two groups, < or ≥ the pooled mean event rate, to compare the effect of the proportion of patients within each group (e.g. stricture activity and stricture type) on outcome.

RESULTS

Study selection

Figure 1 details the study selection flow chart. Two hundred and three studies were identified following both the initial and secondary database search. Studies were screened according to the above eligibility criteria and 30 studies were included as part of a full text review. A total of 25 studies were included in the final review (Table 1).

Author	Country	Study Design	Population Size	Number Of Strictures	Number Of Dilatations	Number Of Females	Study Outcome
Ajlouni Y,[20] 2006	Australia	R	37	83	113	22	E
Atreja,[21] 2014	USA	R	128	169	430	73	E
Bahlme,[22] 2013	United Kingdom	R	79	93	191	47	E,S
Blomberg,[23] 1991	Sweden	P	27	ns	ns	16	S
Breysem,[24] 1992	Belgium	P	18	20	24	13	E,S
Brooker,[25] 2003	United Kingdom	R	14	14	26	6	S
Couckuyt,[26] 1995	Belgium	P	55	59	78	35	E,S
De Angelis,[27] 2013	France	R	26	27	46	15	E,S
East J.E,[29] 2007	United Kingdom	RCT	13	ns	ns	5	E
Endo,[30] 2013	Japan	P	30	47	83	8	E
Ferlitsch,[31] 2006	Austria	P	46	ns	73	26	Su
Foster,[32] 2008	USA	R	24	29	71	18	S
Gustavsson,[33] 2012	Sweden	R	125	ns	594	59	E
Honzawa,[34] 2013	Japan	R	25	29	83	6	E
Hunter,[28], 2001	United Kingdom	R	22	ns	71	16	S
Mueller,[35] 2010	Germany	P	55	74	93	34	E,S
Nanda,[36] 2013	Ireland	P	31	ns	55	14	E,S
Ramboer,[37] 1995	Belgium	P	13	15	53	5	S
Sabate,[38] 2003	France	R	38	41	53	18	E,S
Scimea,[39] 2011	Italy	P	37	39	72	14	E,S
Singh,[40] 2005	USA	R	17	20	29	10	E,S
Stienecker,[41] 2009	Germany	P	25	31	50	20	E
Thomas-Gibson,[43] 2003	United Kingdom	R	59	ns	124	ns	E,S
Van Assche,[42] 2010	Belgium	R	138	ns	237	77	E,S
Williams A.J.K,[44] 1991	United Kingdom	R	7	ns	15	ns	E
TOTAL			**1089**	**790**	**2664**	**557**	

RCT = Randomised Control Trials; R = Retrospective; P = Prospective; ns = not specified;
Study Outcome Measures: S = Symptomatic response, E = Endoscopic / Technical response, Su = Surgery

Table 1 - Study demographics, frequency of population size, strictures and dilatations - *Description of studies included in the literature review. A total number of 1089 subjects were described in the literature with 790 strictures undergoing 2664 interventions. Most were females (557/1023 [54%]) and most studies described both E and S outcomes [E = 8, S = 6 and E, S = 11]*

They included 10 prospective studies (N=10), 14 retrospective studies (N=14) and 1 randomised control trial (N=1).[20-44] Publication dates ranged from 1991 to 2014 and originated from European,[22-29 31 33 35-39 41-44] (6 studies from the United Kingdom),[22 25 28 29 43 44] North American,[21 32 40] Japanese,[30 34] and Australian,[20] institutions.

Study Characteristics

The cumulative data for the 25 studies [20-44] included 1089 patients, 790 strictures and 2664 dilatations. Fifty-one percent (557/1089) were females, 43% (466/1089) were males and for the remaining 8% (66/1089) gender was unspecified. The median age at first dilatation reported across 17 studies (determined from the mean age at first dilatation per study) was 41.1 (range = 32.5 - 50).[21 23-28 30 32 34 36-42] Symptomatic response was reported as the outcome measure in 16 studies,[23 25 28 32 37] technical response in 8 studies,[20 21 29 30 33 34 41 44] and both symptomatic and technical response in 11 studies (Table 1).[22 24 26 27 35 36 38-40 42 43]

Follow-up duration was reported in 24 studies with wide variation.[20-24 26-35] The median maximum follow-up time period was 83.5 months (range 12-172). The median minimum follow-up time period was 4 months (range 0 – 84).

Nine studies (N=468) did not report number of strictures per patient. All but 14 of the remaining 621 cases from 16 studies, had >1 stricture documented.[20-22 24-27 30 32 34 35 37-41] Fifteen studies examined lower gastrointestinal strictures only[23 28 29 31 33 36 42 43] while ten included both upper and lower GI strictures. The proportion of upper GI strictures was 3.8% (though due to missing data in 9 studies this was an estimation).

Stricture activity

The proportion of patients with active[22-26 37 38] and quiescent[22 24-26 30 31 37 38] strictures was 44.9% (155/345) and 47.2% (151/320) respectively. In the majority of cases 82.2% (447/544) data relating to stricture activity was not reported.[22 24 25 28 33-36 39 42]

Stricture type

Across thirteen studies (N=565)[22-26 28 31 36-38 41-43], most patients (79.1%; 447/565) had anastomotic strictures and only (19.6%; 111/565) had de novo strictures.[22-26 28 31 36-38 41-43]

Intervention technique

A maximum balloon diameter of 18mm,[24 27 28 35-37 41-43] 20mm,[20-22 25 29-32 34 39 40 44] and 25mm[23 26 33 38] was reported across all 25 studies. There was variation in the maximum inflation time across 23 studies with maximal inflation periods of one,[22 34 35] two,[20 24 26 27 32 37-44], 3,[28-30] 4, [23 33 36] and 5 minutes [31].

Imaging

The majority of studies describe pre-interventional imaging (18/20; 90%)[21 22 25-31 33 35 36 38-43] and only two studies did not (2/20; 10%).[20 32] The median maximum length of strictures reported across 21 studies was 7cm (range 3 - 25cm).

Synthesis of Results

Symptomatic response

Fifteen studies reported on symptomatic response for patient numbers (N = 615)[22-28 31 32 35-37 39 40 42 43] which was 63.9% (393/615) (Table 2). Fourteen of these were from Europe (612/653 [93.7%]) spanning 7 different countries (four from the UK)[22-28 35-39 42 43] and 2 were North American studies (41/653 [6.3%]).[32 40] The proportion of females and males was 51.8% (338/653) and 39.2% (256/653) respectively. One study did not report on gender proportions (n = 59).[43] The median age at first dilatation reported across 13 studies reporting on symptomatic response (determined from the mean age at first dilatation per study) was 42.6 (range 33.7-50).[23-28 32 36-40 42]

Author, Year	Number of patients (dilatations*) reporting symptomatic response (n) / sample size	Percentage of patients (or dilatations*) (%)
Blomberg,[23] 1991	22 / 27	81
Breysem,[24] 1992	8 / 18	44
Ramboer,[37] 1995	11 / 13	85
Couckuyt,[26] 1995	34 / 55	62
Hunter,[28], 2001	16 / 22	73
Brooker,[25] 2003	11 / 14	79
Thomas-Gibson,[43] 2003	24 / 59	41
Sabate,[38] 2003	47 / 53 *	89 *
Singh,[40] 2005	13 / 17	76
Foster,[32] 2008	22 / 24	92
Mueller,[35] 2010	42 / 55	76
Van Assche,[42] 2010	61 / 138	44
Scimea,[39] 2011	30 / 37	81
Nanda,[36] 2013	14 / 31	45
De Angelis,[27] 2013	24 / 26	92
Bahlme,[22] 2013	61 / 79	77
Total ¥	393 / 615	63.9

** Sabate et al[38] reported symptomatic response from the number of dilatations performed, not the number of patients receiving dilatation.*

¥ Sabate et al[38] excluded from cumulative analysis.

Table 2. Proportions of reported symptomatic response outcomes following endoscopic balloon dilatation - *Fifteen studies (n=615) reported symptomatic outcome per patient and one study per dilatation*

procedure. The number of patients reporting improvement in each study is shown. A total of 393 (63.9%) patients reported symptomatic response with dilatation. One study reported response rate for the number of dilatations 89%.

Analysis of pooled study outcomes demonstrated a symptomatic response rate of 70.2% [95% CI: 60-78.8%] with evidence of moderate to high heterogeneity between studies [I^2: 63.8%] (Figure 2). On exclusion of the six outlier publications,[24 27 32 36 42 43] the symptomatic response rate was 74.8% (95% CI: 69.6-79.3%; I^2: 0.0%) (Figure 3). The symptomatic response rate in one study was measured according to the number of dilatations done and was reported as 89% (47/53).[38]

A comparison between symptomatic response rates and other variables is shown in Table 3.

Variable		Symptomatic response	Technical response	Perforation
Categorical variables represented as pooled event rates % (95% Confidence Intervals)				
Balloon diameter (millimetres)	**18**	61.9 (47.4 – 74.4)	94.7 (87-97.9)	3.1 (1.9-5.1)
	20	79.5 (72.6-85)	90.6 (87.4-93.0)	3.8 (2.2-6.5)
	25	71 (48 – 86.6)	89.4 (86.9-91.5)	2.2 (1.3-3.6)
Duration of inflation (minutes)	**2**	70.6 (58.1-80.6)	92.2 (86.6-95.5)	3.1 (1.8-5.3)
	5	67.1 (41.9 – 85.2)	92.6 (84.2-96.7)	3.6 (1.6-8.2)
Geography	**European**	67.9 (57-77.2)	90.6 (86.7-93.5)	2.3 (1.6-3.3)
	North American	84.5 (62.6 – 94.7)	91.6 (86.8-94.8)	5.0 (1.3-17.7)
	Japanese		91.3 (81.9-96.0)	
Pre-interventional imaging	**Reported**	70.4 (58.8 – 79.8)	92.3 (85.8-95.9)	2.7 (1.8-4.0)

Table 3. The relationship between response rates (symptomatic and technical) and adverse event rates (perforation) with categorical variables – *The categorical variables (balloon diameter, duration of inflation, geography and pre-interventional imaging) are presented as pooled event rates across studies alongside their 95% confidence intervals.*

The proportion of patients demonstrating a higher than average compared with lower than average symptomatic response rate for active strictures was 60% and 73%, for anastomotic strictures was 75% and 85% and for de novo strictures was 25% and 15% respectively. (Table 2 of supplementary material).

Technical Response

A total of 19 studies reported on technical response as an outcome measure. Analysis for pooled study results demonstrated an event rate of 90.6% (95% CI: 87.8-92.8%) with low heterogeneity between studies (I^2: 11.7%) (Figure 1 – Appendix: Supplementary Figures). Fourteen studies originated from Europe across seven different countries. The median age at first dilatation reported across 12 studies was 40.1 years (range 32-49). The proportion of females and males were 49.8% (470/943) and 43.2% (407/943) respectively.

Reported outcomes were expressed either for patients (N = 435),[22 24 29 34 35 38 39 41 42 44] stricture (N = 299),[20 21 30] and intervention/dilatation (N = 926) (Supplementary table 3).[26 27 33 36 43 45]. The proportion of patients demonstrating technical response was 92.6% (403/435). The technical response rate for strictures was 91.3% (273/299) and for balloon dilatation interventions was 90% (833/926). The pooled event analysis results were similar due to low heterogeneity (Figure 1 -Appendix: supplementary figures).

The relationship between technical response rates and variables is shown in Table 3. The proportion of patients demonstrating a higher than average compared with lower than average technical response rate for active strictures, was 51% and 52%, for quiescent strictures, was equal at 45%, for anastomotic strictures 75% and 83% and for *de novo* strictures was 25% and 17% respectively. (Table 2 in supplementary material).

Complications and Perforations

Fifteen studies reported complications according to the number of patients (N = 564)[20-22 25 26 28 32 35-37 39 40 44] and four studies reported complications based on the number of dilatations (N = 1228).[31 33 42 43] The proportion of patients and dilatations with complications was 4.4% (25/564) and 5.2% (64/1228). Analysis of pooled study data demonstrated an overall complication rate of 6.4% (95% CI: 5.0 – 8.2; I^2: 4.0%).

Eighteen studies reported on perforation for patients (N = 654)[20-28 34-41 44] and four studies for number of dilatations (N = 1281).[31-33 42 43] The proportion of patients and dilatations that were followed by perforation was 2.4% (16/654) and 1.8% (23/1281) respectively. Study data analysis showed no heterogeneity across studies with a pooled mean perforation rate of 3% (95% CI: 2.2- 4.0%; I^2: 0%) (Supplementary figure 2).

The relationship between perforation rates and variables is shown in Table 3. Balloon inflation diameters of 18mm,[24 27 28 34-37 41-43] 20mm,[20-22 25 31 34 39 40 44] and 25mm[23 26 33 37] demonstrated pooled mean perforation rates of 3.1% (95% CI: 1.9-5.1%; I^2 0%), 3.8% (95% CI: 2.2-6.5%; I^2 0%), and 2.2% (95% CI: 1.3-3.6%; I^2 0%) respectively. An inflation time of up to 2 minutes[22 24 26-28 32 35-37 39 40 42 43] and 5 minutes[23 30 33 35] demonstrated a pooled mean perforation rate of 3.1% (95% CI: 1.8-5.3%; I^2 0%) and 3.6% (95% CI: 1.6-8.2%; I^2 0%) respectively. The mean perforation rate across 18 European studies that reported perforation according to the number of patients was 2.3% (95% CI: 1.6-3.3%; I^2 0%).[22-28 31 33 35-39 41-44] The mean perforation rate across three North American studies was 5.0% (95% CI: 1.3-17.7%; I^2 0%) (Table 5).[21 32 40] The use of pre-interventional imaging was described across 14 studies[21 22 26 27 30 33 35 36 38-43] where the pooled mean perforation rate was 2.7% (95% CI: 1.8-4.0%; I^2 0%). The perforation rate in one study that did not use pre-interventional imaging was 1.3% (95% CI: 0.1-17.8%).[20] The median maximum stricture length reported across 20 studies was 7cm (range 2-25cm).

The proportion of patients with higher than average compared with lower than average perforation rates for active strictures was 62% and 36%, for quiescent strictures was 38% and 55%, for anastomotic strictures

was 81% and 7% and for de novo strictures was 19% and 21% respectively. (Table 2 Supplementary material)

Surgery

Surgery was required for one or more of the following two events: (i) inaccessible strictures during endoscopy and (ii) persistent or recurrent symptoms i.e. failed repeated dilatation.

(i) *Endoscopic inaccessibility*

Seven studies reported 12.9% of cases (33/256) where endoscopic balloon dilatation could not be completed during endoscopy either because the stricture was too narrow or there was acute angulation.[24 26 31 35 38 39 44] Two studies (N = 130) reported this event 6.9% (9/130) by number of strictures.[20 30]

(ii) Symptomatic disease

Twenty one studies reported surgical outcomes for ongoing recurrent disease despite repeated balloon dilatation according to the number of patients (N = 849).[20-28 31-41 44] The proportion of patients who underwent surgery was 21.1% (179/849). Pooled data demonstrated a mean surgical event rate of 23.3% (95% CI: 20-26.5), with low heterogeneity demonstrated across studies (I^2: 13.2%) (Supplementary figure 3). One study reported the requirement of surgery according to the number of strictures (N = 47).[30] The proportion of strictures requiring surgery in this study was 29.8% (14/47).

DISCUSSION

This is the most comprehensive systematic review of endoscopic balloon dilatation for the management of Crohn's strictures to date.

Summary of evidence

This review offers insights into endoscopic balloon dilatation for Crohn's strictures on which to build future, more robust, study designs to measure efficacy. The pooled event rate for symptomatic response of 74.8% excludes 6 studies identified as outliers from the funnel plot. However, the distribution of the studies within the plot also suggests publication bias may be over-estimating the effect size. Incomplete data on the number of cases where access failed during an endoscopic procedure may contribute to the over-estimation of the effect size. In this review it was only reported in 24% of the studied population with a failure rate of 13%. Failure of endoscopic access is particularly relevant as the majority of strictures undergoing dilatation are likely to be anastomotic and associated with adhesions and fibrosis. In contrast, a previous systematic review on endoscopic balloon dilation indicated a 58 % response comparable to our 63.9% expressed as the proportion of patients but a lower value than the pooled effect.[19] Since the pooled summative effect addresses weight and heterogeneity between studies, it reflects a more accurate measure of efficacy albeit subject to bias.

The review reveals other relevant observations pertaining to the intervention. Where the studies examined both symptomatic and technical response, the former was consistently less than the technical rate of 90.6%.

The inferences is that passage of the endoscope through the stricture is an inadequate outcome for patients. There was a wide variation in dilatation techniques suggesting an overriding need for standardisation of endoscopic procedures. A balloon diameter of 20mm seems to be commonest, and the most effective size limit, consistent with the internal small bowel diameter of 25mm. Two minutes of dilatation is the commonest duration used and may be associated with better outcomes. The most effective dilatation technique is a three step increase in diameter with regular repeat procedures until resolution of symptoms on a normal diet.

The study focused on an adult population with a mean age at first dilatation of 41 years, which reflects their aetiology as a complication of the disease or surgery. Additional analyses to assess factors that might influence outcomes were undertaken by using the mean value of pooled outcomes to create two groups, < or ≥ the pooled mean event rate, in order to compare the effect of different variables on outcomes. There were more patients with stricture inflammation, in the below average than above average symptomatic response group; in parallel more non-inflamed strictures were associated with above average symptomatic response. For anastomotic strictures, more patients were in the below average than above average symptomatic and technical response groups. This observation is counter-intuitive as anastomotic strictures tend to be shorter but perhaps it shows that there is greater resistance of fibrotic strictures to dilatation. With respect to dilatation technique maximum balloon diameter and duration of inflation did not seem to show different outcomes, except that 20mm size was accompanied by higher symptomatic response rate than 18mm (80% vs 52%). There was no evidence of higher perforation rate with dilatation diameters of 25mm. North America symptomatic response rates were higher than Europe but were associated with higher complication rates. More inflammatory were in the above average perforation group and more quiescent strictures were in the below average perforation group.

Overall pooled complication rate was 6.4% [95% CI: 5.0 – 8.2], much higher than complication rates of 2% reported by Hassan.[19] In contrast, the perforation rate, which represents the most significant complication, was 3% [95% CI: 2.2- 4.0%] for pooled analysis and similar to that expressed as proportion of patient in this study (2.6%). The previous review by Hassan did not report perforation and a separate event.[19]

Whilst the strength of this review lies in the systematic manner in which it was conducted in accordance with PRISMA guidelines and methods for narrative reviews, it is beset by several limitations [46].

Limitations

Firstly, the absence of control groups meant that we were unable to compare the impact of variables on outcome measures and the meta-analysis focused on summative effect in relation to heterogeneity. We used an average effect to gauge whether an event occurred above or below mean rate so the relationships shown in these analyses should be interpreted with caution.

The second limitation was the diversity of the populations studied in terms of stricture characteristics, techniques and expression of results according to sample size, stricture numbers or number of interventions. This made comparisons across the studies difficult particularly for outcomes which expressed results according to sample size, stricture numbers and/or number of interventions. We used population size for primary outcomes of this review. Thirdly incomplete and variable reporting of some population and interventional characteristics mean analyses were conducted on data that were available and may not be generalizable to other studies or populations. This limitation explains why the number of strictures was less than the sample number of the review (790 and 1089). Fourthly, most studies were reported by gastroenterologists, with a bias towards showing endoscopic benefit through both performance and reporting bias. Lastly, none of the studies mention dietary restrictions on follow-up: low fibre diet will be associated with better and sustained response than a resumption to a full diet at the expense of quality of life.

Conclusions

This review measures the efficacy of endoscopic balloon dilation for treatment of Crohn's strictures: 74.8% response rate may be an overestimate due to publication bias and yet a more accurate estimate of the previous reports of 58% which did not use a pooled event rate. Whilst there is a suggestions that some strictures may respond better than others this is far from conclusive due to the lack of a control group but

merely indicates that further studies should take into account the extent and severity of inflammation in strictures. It draws attention to variation in intervention techniques between studies, inadequacy of outcome measure and deficiency in pre-assessment for suitability of dilatation.

There is a discernible absence of randomised controlled trial of endoscopic balloon dilatation for Crohn's strictures. This reflects the difficulty of a comparable control and the ethical dilemma of using sham intervention or surgery. Surgery is usually reserved for longer strictures and less so for non-accessible strictures because double balloon enteroscopy increased access to small bowel strictures. Anti-TNF therapies have a role in the treatment of strictures through anti-inflammatory effects that increase the diameter of the bowel lumen and also reduce TNF-induced fibrosis.[47 48] Drug therapy as a control arm may be a more acceptable option to address the question of the optimal small bowel preserving treatment for Crohn's strictures. The challenge of a control arm is not the only barrier to conclusive results. The other is an optimal and comparable outcome measure that capture relief of obstructive symptoms, resumption of normal dietary intake, quality of life and monitoring for repeat stenosis or fibrosis. Further exploratory studies on this aspect are warranted.

Endoscopic balloon dilatation plays an important role in the management of Crohn's strictures. The risk benefit profile depends on several stricture factors. Some strictures may be more effectively managed with surgery. Others may be treated with anti-TNF therapy to reduce fibrosis and inflammation with or without balloon dilatation to break down the collagen fibres. Future studies exploring this should also examine how imaging and biochemical markers, may guide treatment decisions. Reproducible outcome measures with scores to represent inflammatory vs fibrotic components as end-points will allow for comparison across studies. There is already some research in this field which is demonstrating promising results [49]. Without these studies an evidence-based management pathway that reduce variation and set standards in care for Crohn's strictures cannot be developed.

REFERENCES

1. Cosnes J, Cattan S, Blain A, et al. Long-term evolution of disease behavior of Crohn's disease. Inflamm Bowel Dis 2002;**8**(4):244-50
2. Louis E, Collard A, Oger AF, et al. Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. Gut 2001;**49**(6):777-82
3. Solberg IC, Vatn MH, Hoie O, et al. Clinical course in Crohn's disease: results of a Norwegian population-based ten-year follow-up study. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2007;**5**(12):1430-8 doi: 10.1016/j.cgh.2007.09.002[published Online First: Epub Date]|.
4. Thia KT, Sandborn WJ, Harmsen WS, et al. Risk factors associated with progression to intestinal complications of Crohn's disease in a population-based cohort. Gastroenterology 2010;**139**(4):1147-55 doi: 10.1053/j.gastro.2010.06.070[published Online First: Epub Date]|.
5. Wolters FL, Russel MG, Sijbrandij J, et al. Phenotype at diagnosis predicts recurrence rates in Crohn's disease. Gut 2006;**55**(8):1124-30 doi: 10.1136/gut.2005.084061[published Online First: Epub Date]|.
6. Rieder F, Zimmermann EM, Remzi FH, et al. Crohn's disease complicated by strictures: a systematic review. Gut 2013;**62**(7):1072-84 doi: 10.1136/gutjnl-2012-304353[published Online First: Epub Date]|.
7. Graham MF, Diegelmann RF, Elson CO, et al. Collagen content and types in the intestinal strictures of Crohn's disease. Gastroenterology 1988;**94**(2):257-65

8. Rieder F, Karrasch T, Ben-Horin S, et al. Results of the 2nd scientific workshop of the ECCO (III): basic mechanisms of intestinal healing. J Crohns Colitis 2012;**6**(3):373-85 doi: 10.1016/j.crohns.2011.11.009[published Online First: Epub Date]|.

9. Scarpa M, Angriman I, Barollo M, et al. Role of stapled and hand-sewn anastomoses in recurrence of Crohn's disease. Hepato-gastroenterology 2004;**51**(58):1053-7

10. Yamamoto T, Allan RN, Keighley MR. Strategy for surgical management of ileocolonic anastomotic recurrence in Crohn's disease. World journal of surgery 1999;**23**(10):1055-60; discussion 60-1

11. Louis E, Michel V, Hugot JP, et al. Early development of stricturing or penetrating pattern in Crohn's disease is influenced by disease location, number of flares, and smoking but not by NOD2/CARD15 genotype. Gut 2003;**52**(4):552-7

12. Abreu MT, Taylor KD, Lin YC, et al. Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease. Gastroenterology 2002;**123**(3):679-88

13. Alvarez-Lobos M, Arostegui JI, Sans M, et al. Crohn's disease patients carrying Nod2/CARD15 gene variants have an increased and early need for first surgery due to stricturing disease and higher rate of surgical recurrence. Annals of surgery 2005;**242**(5):693-700

14. Rieder F, Schleder S, Wolf A, et al. Serum anti-glycan antibodies predict complicated Crohn's disease behavior: a cohort study. Inflamm Bowel Dis 2010;**16**(8):1367-75 doi: 10.1002/ibd.21179[published Online First: Epub Date]|.

15. Rigby RJ, Hunt MR, Scull BP, et al. A new animal model of postsurgical bowel inflammation and fibrosis: the effect of commensal microflora. Gut 2009;**58**(8):1104-12 doi: 10.1136/gut.2008.157636[published Online First: Epub Date]|.

16. Dignass A, Van Assche G, Lindsay JO, et al. The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Current management. J Crohns Colitis 2010;**4**(1):28-62 doi: 10.1016/j.crohns.2009.12.002[published Online First: Epub Date]|.

17. Shental O, Tulchinsky H, Greenberg R, et al. Positive histological inflammatory margins are associated with increased risk for intra-abdominal septic complications in patients undergoing ileocolic

resection for Crohn's disease. Diseases of the colon and rectum 2012;**55**(11):1125-30 doi: 10.1097/DCR.0b013e318267c74c[published Online First: Epub Date]|.

18. Alves A, Panis Y, Bouhnik Y, et al. Factors that predict conversion in 69 consecutive patients undergoing laparoscopic ileocecal resection for Crohn's disease: a prospective study. Diseases of the colon and rectum 2005;**48**(12):2302-8 doi: 10.1007/s10350-005-0190-x[published Online First: Epub Date]|.
19. Hassan C, Zullo A, De Francesco V, et al. Systematic review: Endoscopic dilatation in Crohn's disease. Aliment Pharmacol Ther 2007;**26**(11-12):1457-64 doi: 10.1111/j.1365-2036.2007.03532.x[published Online First: Epub Date]|.
20. Ajlouni Y, Iser JH, Gibson PR. Endoscopic balloon dilatation of intestinal strictures in Crohn's disease: safe alternative to surgery. Journal of gastroenterology and hepatology 2007;**22**(4):486-90 doi: 10.1111/j.1440-1746.2006.04764.x[published Online First: Epub Date]|.
21. Atreja A, Aggarwal A, Dwivedi S, et al. Safety and efficacy of endoscopic dilation for primary and anastomotic Crohn's disease strictures. J Crohns Colitis 2014;**8**(5):392-400 doi: 10.1016/j.crohns.2013.10.001[published Online First: Epub Date]|.
22. Bhalme M, Sarkar S, Lal S, et al. Endoscopic balloon dilatation of Crohn's disease strictures: results from a large United kingdom series. Inflamm Bowel Dis 2014;**20**(2):265-70 doi: 10.1097/01.MIB.0000439067.76964.53[published Online First: Epub Date]|.
23. Blomberg B, Rolny P, Jarnerot G. Endoscopic treatment of anastomotic strictures in Crohn's disease. Endoscopy 1991;**23**(4):195-8 doi: 10.1055/s-2007-1010654[published Online First: Epub Date]|.
24. Breysem Y, Janssens JF, Coremans G, et al. Endoscopic balloon dilation of colonic and ileo-colonic Crohn's strictures: long-term results. Gastrointest Endosc 1992;**38**(2):142-7
25. Brooker JC, Beckett CG, Saunders BP, et al. Long-acting steroid injection after endoscopic dilation of anastomotic Crohn's strictures may improve the outcome: a retrospective case series. Endoscopy 2003;**35**(4):333-7 doi: 10.1055/s-2003-38145[published Online First: Epub Date]|.
26. Couckuyt H, Gevers AM, Coremans G, et al. Efficacy and safety of hydrostatic balloon dilatation of ileocolonic Crohn's strictures: a prospective longterm analysis. Gut 1995;**36**(4):577-80

27. de'Angelis N, Carra MC, Borrelli O, et al. Short- and long-term efficacy of endoscopic balloon dilation in Crohn's disease strictures. World J Gastroenterol 2013;**19**(17):2660-7 doi: 10.3748/wjg.v19.i17.2660[published Online First: Epub Date]|.

28. Dear KL, Hunter JO. Colonoscopic hydrostatic balloon dilatation of Crohn's strictures. J Clin Gastroenterol 2001;**33**(4):315-8

29. East JE, Brooker JC, Rutter MD, et al. A pilot study of intrastricture steroid versus placebo injection after balloon dilatation of Crohn's strictures. Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association 2007;**5**(9):1065-9 doi: 10.1016/j.cgh.2007.04.013[published Online First: Epub Date]|.

30. Endo K, Takahashi S, Shiga H, et al. Short and long-term outcomes of endoscopic balloon dilatation for Crohn's disease strictures. World J Gastroenterol 2013;**19**(1):86-91 doi: 10.3748/wjg.v19.i1.86[published Online First: Epub Date]|.

31. Ferlitsch A, Reinisch W, Puspok A, et al. Safety and efficacy of endoscopic balloon dilation for treatment of Crohn's disease strictures. Endoscopy 2006;**38**(5):483-7 doi: 10.1055/s-2006-924999[published Online First: Epub Date]|.

32. Foster EN, Quiros JA, Prindiville TP. Long-term follow-up of the endoscopic treatment of strictures in pediatric and adult patients with inflammatory bowel disease. J Clin Gastroenterol 2008;**42**(8):880-5 doi: 10.1097/MCG.0b013e3181354440[published Online First: Epub Date]|.

33. Gustavsson A, Magnuson A, Blomberg B, et al. Endoscopic dilation is an efficacious and safe treatment of intestinal strictures in Crohn's disease. Aliment Pharmacol Ther 2012;**36**(2):151-8 doi: 10.1111/j.1365-2036.2012.05146.x[published Online First: Epub Date]|.

34. Honzawa Y, Nakase H, Matsuura M, et al. Prior use of immunomodulatory drugs improves the clinical outcome of endoscopic balloon dilation for intestinal stricture in patients with Crohn's disease. Digestive endoscopy : official journal of the Japan Gastroenterological Endoscopy Society 2013;**25**(5):535-43 doi: 10.1111/den.12029[published Online First: Epub Date]|.

35. Mueller T, Rieder B, Bechtner G, et al. The response of Crohn's strictures to endoscopic balloon dilation. Aliment Pharmacol Ther 2010;**31**(6):634-9 doi: 10.1111/j.1365-2036.2009.04225.x[published Online First: Epub Date]|.

36. Nanda K, Courtney W, Keegan D, et al. Prolonged avoidance of repeat surgery with endoscopic balloon dilatation of anastomotic strictures in Crohn's disease. J Crohns Colitis 2013;**7**(6):474-80 doi: 10.1016/j.crohns.2012.07.019[published Online First: Epub Date]|.

37. Ramboer C, Verhamme M, Dhondt E, et al. Endoscopic treatment of stenosis in recurrent Crohn's disease with balloon dilation combined with local corticosteroid injection. Gastrointest Endosc 1995;**42**(3):252-5

38. Sabate JM, Villarejo J, Bouhnik Y, et al. Hydrostatic balloon dilatation of Crohn's strictures. Aliment Pharmacol Ther 2003;**18**(4):409-13

39. Scimeca D, Mocciaro F, Cottone M, et al. Efficacy and safety of endoscopic balloon dilation of symptomatic intestinal Crohn's disease strictures. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver 2011;**43**(2):121-5 doi: 10.1016/j.dld.2010.05.001[published Online First: Epub Date]|.

40. Singh VV, Draganov P, Valentine J. Efficacy and safety of endoscopic balloon dilation of symptomatic upper and lower gastrointestinal Crohn's disease strictures. J Clin Gastroenterol 2005;**39**(4):284-90

41. Stienecker K. Long-term results of endoscopic balloon dilatation of lower gastrointestinal tract strictures in Crohn's disease: A prospective study. World Journal of Gastroenterology 2009;**15**(21):2623 doi: 10.3748/wjg.15.2623[published Online First: Epub Date]|.

42. Thienpont C, D'Hoore A, Vermeire S, et al. Long-term outcome of endoscopic dilatation in patients with Crohn's disease is not affected by disease activity or medical therapy. Gut 2010;**59**(3):320-4 doi: 10.1136/gut.2009.180182[published Online First: Epub Date]|.

43. Thomas-Gibson S, Brooker JC, Hayward CM, et al. Colonoscopic balloon dilation of Crohn's strictures: a review of long-term outcomes. European journal of gastroenterology & hepatology 2003;**15**(5):485-8 doi: 10.1097/01.meg.0000059110.41030.bc[published Online First: Epub Date]|.

44. Williams AJ, Palmer KR. Endoscopic balloon dilatation as a therapeutic option in the management of intestinal strictures resulting from Crohn's disease. The British journal of surgery 1991;**78**(4):453-4

45. Singh VV, Draganov P, Valentine J. *Efficacy and safety of endoscopic balloon dilation of symptomatic upper and lower gastrointestinal Crohn's disease strictures*: Journal of Clinical Gastroenterology. 39 (4) (pp 284-290), 2005. Date of Publication: April 2005., 2005.

46. York Uo. Systematic Reviews. CRD's guidance for undertaking reviews in health care. Centre for Reviews and Dissemination, 2009.

47. Pelletier AL, Kalisazan B, Wienckiewicz J, et al. Infliximab treatment for symptomatic Crohn's disease strictures. Aliment Pharmacol Ther 2009;**29**(3):279-85 doi: 10.1111/j.1365-2036.2008.03887.x[published Online First: Epub Date]|.

48. Y. Bouhnik1, D. Laharie2, C. Stefanescu1, X. et al. UEG Week 2014 Oral Presentations - OP122 EFFICACY OF ADALIMUMAB IN PATIENTS WITH CROHN'S DISEASE AND SYMPTOMATIC SMALL BOWEL STRICTURE: A MULTICENTRE, PROSPECTIVE, OBSERVATIONAL COHORT STUDY. United European Gastroenterology Journal. Sage UK: London, England: SAGE Publications, 2014:A1-A131.

49. Van Assche G, Herrmann KA, Louis E, et al. Effects of infliximab therapy on transmural lesions as assessed by magnetic resonance enteroclysis in patients with ileal Crohn's disease. J Crohns Colitis 2013;**7**(12):950-7 doi: 10.1016/j.crohns.2013.01.011[published Online First: Epub Date]|.

STATEMENT OF AUTHORSHP

All named authors have approved the final version of the manuscript, including the authorship list.

Guarantor of article: Pritesh Morar

Specific author contributions: PM and NA performed the research, PM,

FUNDING

None

ACKNOWLEDGEMENTS

The authors would like to acknowledge Dr Ravi Misra for study quality assessment.

COMPETING INTERESTS

None to declare

Section/topic	#	Checklist item	Reported on page #
TITLE			1
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			2
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	Supplementary file structure summary
INTRODUCTION			5
Rationale	3	Describe the rationale for the review in the context of what is already known.	5 - 6
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	6
METHODS			7
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	7
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	7
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	7
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	7, Figure 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	Figure 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	8
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	8
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	8; Table 1
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	9

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	9

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	9
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	10
RESULTS			11
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	11; Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	13; Table 1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Supp. Table 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Table 2, Figure 2-3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Page 15 – 21, Figure 3, supp. Figures 1-3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	Page 15 – 21, Figure 2 – 5.
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	Page 19 – 24; Tables 3, Sup. table 3
DISCUSSION			25
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	22
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	23
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for	24

		future research.	
FUNDING			32
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	32

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Figure 1 – Flow chart demonstrating the search strategy in accordance with PRISMA – Two hundred and three (n=203) records were identified following duplicate removal. Fifty three (n=53) records were removed after limits were applied. One hundred and fifty (n=150) records underwent screening and one hundred and twenty records were excluded (n=120). Thirty records (n=30) were assessed for eligibility and twenty five (n=25) articles were included in our quantitative analysis.
183x174mm (150 x 150 DPI)

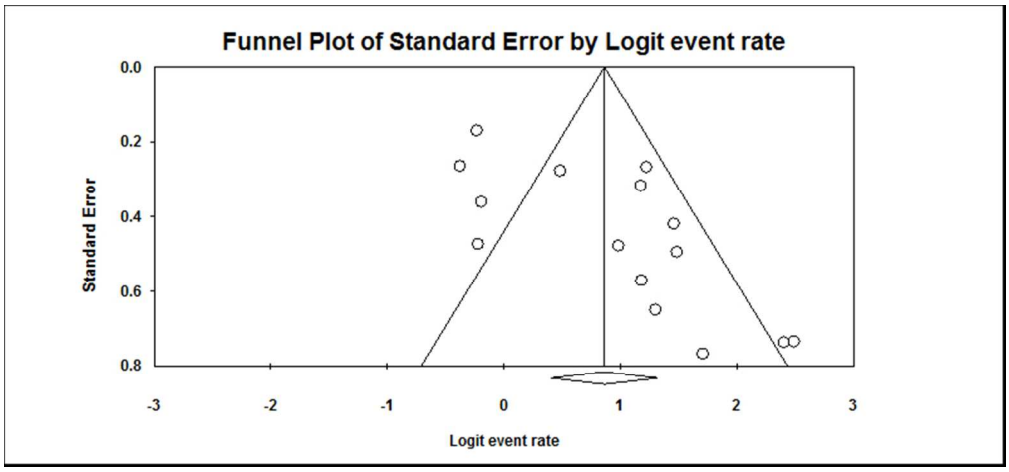


Figure 2 – Funnel plot for studies reporting on symptomatic response rate - Heterogeneity across studies was demonstrated (I2: 63.8%). Six studies[24 27 32 36 42 43] were outliers and determined as sources for publication bias on sensitivity analysis.
211x97mm (96 x 96 DPI)

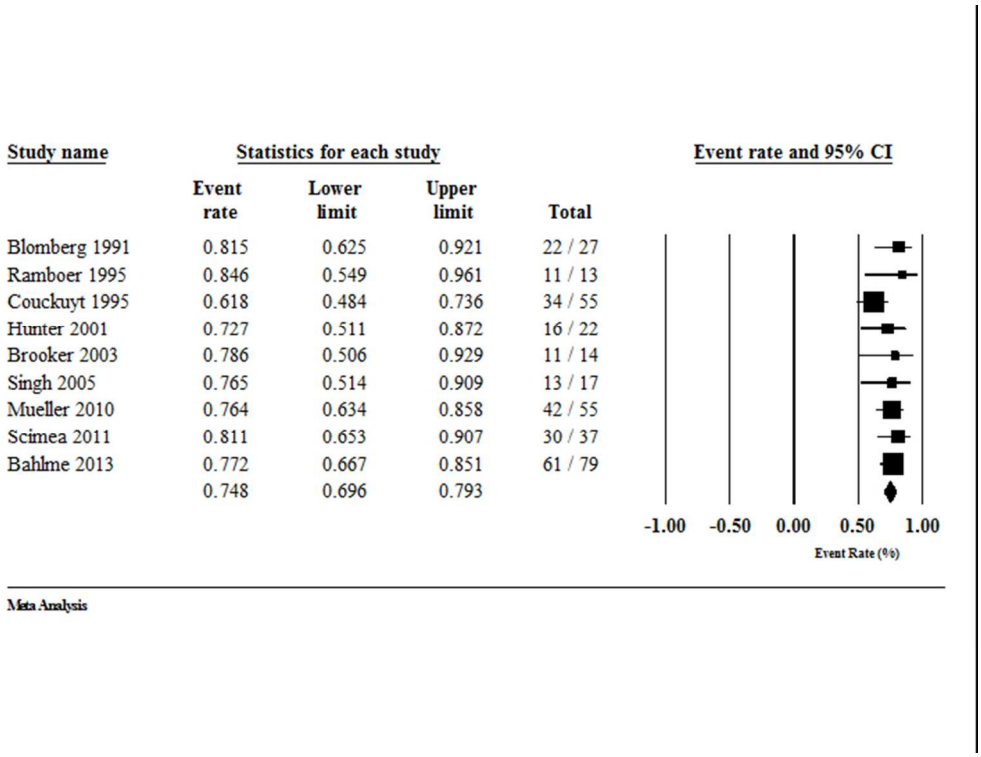


Figure 3 – Forest plot for studies reporting on symptomatic response following exclusion of outlier studies[24 27 32 36 42 43] – Random effects model demonstrating a pooled event rate for symptomatic response as 74.8% (95% CI: 69.6 – 79.3%; I2: 0%).
191x142mm (96 x 96 DPI)

Supplementary Tables

Author, year	Selection	Comparability	Outcome	Score
Ajlouni Y,[20] 2006	***	-	**	***
Atreja,[21] 2014	***	-	**	**
Bahlme,[22] 2013	***	-	**	****
Blomberg,[23] 1991	**	-	*	****
Breysem,[24] 1992	*	-	*	*****
Brooker,[25] 2003	****	-	**	******
Couckuyt,[26] 1995	***	-	*	*****
De Angelis,[27] 2013	***	-	**	****
East J.E,[29] 2007	**	-	*	****
Endo,[30] 2013	***	-	**	****
Ferlitsch,[31] 2006	***	-	*	*****
Foster,[32] 2008	***	-	*	****
Gustavsson,[33] 2012	***	-	**	****
Honzawa,[34] 2013	***	-	**	*****
Hunter,[28], 2001	***	-	**	*****
Mueller,[35] 2010	***	-	**	*****
Nanda,[36] 2013	***	-	**	*****
Ramboer,[37] 1995	**	-	**	*****
Sabate,[38] 2003	***	-	*	*****
Scimea,[39] 2011	***	-	**	*****
Singh,[40] 2005	***	-	*	*****
Stienecker,[41] 2009	**	-	**	*****
Thomas-Gibson,[43] 2003	***	-	**	***
Van Assche,[42] 2010	***	-	**	****
Williams A.J.K,[44] 1991	***	-	*	*****

Table 1 – Quality assessment of studies using the Newcastle-Otawa scale – *The maximum number of stars each study can receive is 9 (maximum 4 for selection, 2 for comparison and 3 for outcome. The median number of stars across studes was 5 (range: 2 – 6).*

		Symptomatic response [a]		Technical response [b]		Perforation [c]	
Pooled Mean Event Rate PMER		*74.8% (95% CI: 69.6-79.3%)*		*90.6% (95% CI: 87.8-92.8%)*		*3% (95% CI: 2.2- 4.0%)*	
Dichotomous Outcome		*< PMER*	*≥ PMER*	*< PMER*	*≥ PMER*	*< PMER*	*≥ PMER*
		Continuous variables % (N / Population size)					
Stricture activity	**Active**	73 (53/73)	60 (102/171)	52 (29/56)	51 (84/164)	36 (62/170)	62 (93/150)
	Quiescent	25 (18/73)	33 (57/171)	45 (25/56)	45 (73/164)	55 (94/170)	38 (57/150)
Stricture type	**Anastomotic**	85 (274/323)	75 (128/171)	83 (95/115)	75 (246/328)	77 (217/282)	81 (230/283)
	De novo	15 (49/323)	25 (43/171)	17 (20/115)	25 (82/328)	21 (58/282)	19 (53/283)

Table 2. The relationship of response (symptomatic and technical) and adverse outcomes expressed in dichotomous form (< or ≥ the pooled mean event rate) with continuous variables (stricture activity and stricture type). (a) *For symptomatic response in relation to (i) stricture activity, 73 patients were in the lower PMER and 171 patients were in the higher PMER groups. Comparison of the lower PMER with higher PMER group, showed more active strictures (73% vs 60%) and less quiescent strictures (25% vs 33%) (ii) stricture type, 323 patients were in the lower PMER and 171 patients were in the higher PMER groups. Comparing the lower with higher PMER groups, there were more anastomotic strictures (85% vs 75%) and less de novo strictures (15% vs 25%). (b) For technical response in relation to (i) stricture activity, 56 patients were in the lower than average PMER group and 164 patients were in the higher than average PMER group. There were no differences in stricture activity between high and low technical response groups (ii) stricture type, 115 patients were in lower than average PMER group and 328 patients in higher than average PMER. Comparing the lower PMER group with higher PMER group, there were more anastomotic strictures (83% vs 75%) and less de novo strictures (17% vs 25%). (c) For perforation, in relation to (i) stricture activity, 170 patients were in*

the lower than average perforation rate group and 150 patients were in the higher than average group. Comparing the higher perforation with lower perforation groups, there were more active strictures (62% vs 36%) and less quiescent strictures (38% vs 55%) (ii) stricture type, there a 282 patients in the lower than average perforation group and 283 patients in the higher than average perforation group. Comparison of the higher perforation with lower perforation groups, there were similar numbers of anastomotic strictures (81% vs 77%) and de novo strictures (19% vs 21%). PMER: pooled mean event rate.

Author, Year	Number of patients (n)	Technical response rate (%)
Bahlme [22], 2013	75	95
Breysem [24], 1992	16	89
East J.E [29], 2007	12	92
Honzawa [34], 2013	22	88
Mueller [35], 2010	52	95
Sabate [38], 2003	32	84
Scimea [39], 2011	31	84
Stienecker [41], 2009	24	96
VanAssche [42], 2010	134	97
Williams A.J.K [44], 1991	5	71
Total	**403**	**93**
Author, Year	**Number of strictures (n)**	**Technical response rate (%)**
Ajlouni Y [20], 2006	75	90
Atreja [21], 2014	154	91
Endo [30], 2013	44	94
Total	**273**	**91**
Author, Year	**Number of balloon dilatations (n)**	**Technical response rate (%)**
Couckuyt [26], 1995	70	90
De Angelis [27], 2013	46	100
Gustavsson [33], 2012	533	90
Nanda [36], 2013	55	100
Singh [40], 2005	28	97
Thomas-Gibson [43], 2003	101	81
Total	**833**	**61**

Table 3 - Proportions of reported technical response outcomes following endoscopic balloon dilatation - *Ten studies (n=435) reported a technical response outcome per patient, three per stricture (n=299) and six per balloon dilatation procedure (n = 1356). The number of patients, stricture and balloon dilatations reporting improvement in each study is shown. A total of 403 (92.6%) patients, 273 (91.3%) strictures and 833 (61.4%) balloon dilatations reported technical response with dilatation.*

Supplementary Material - Figures

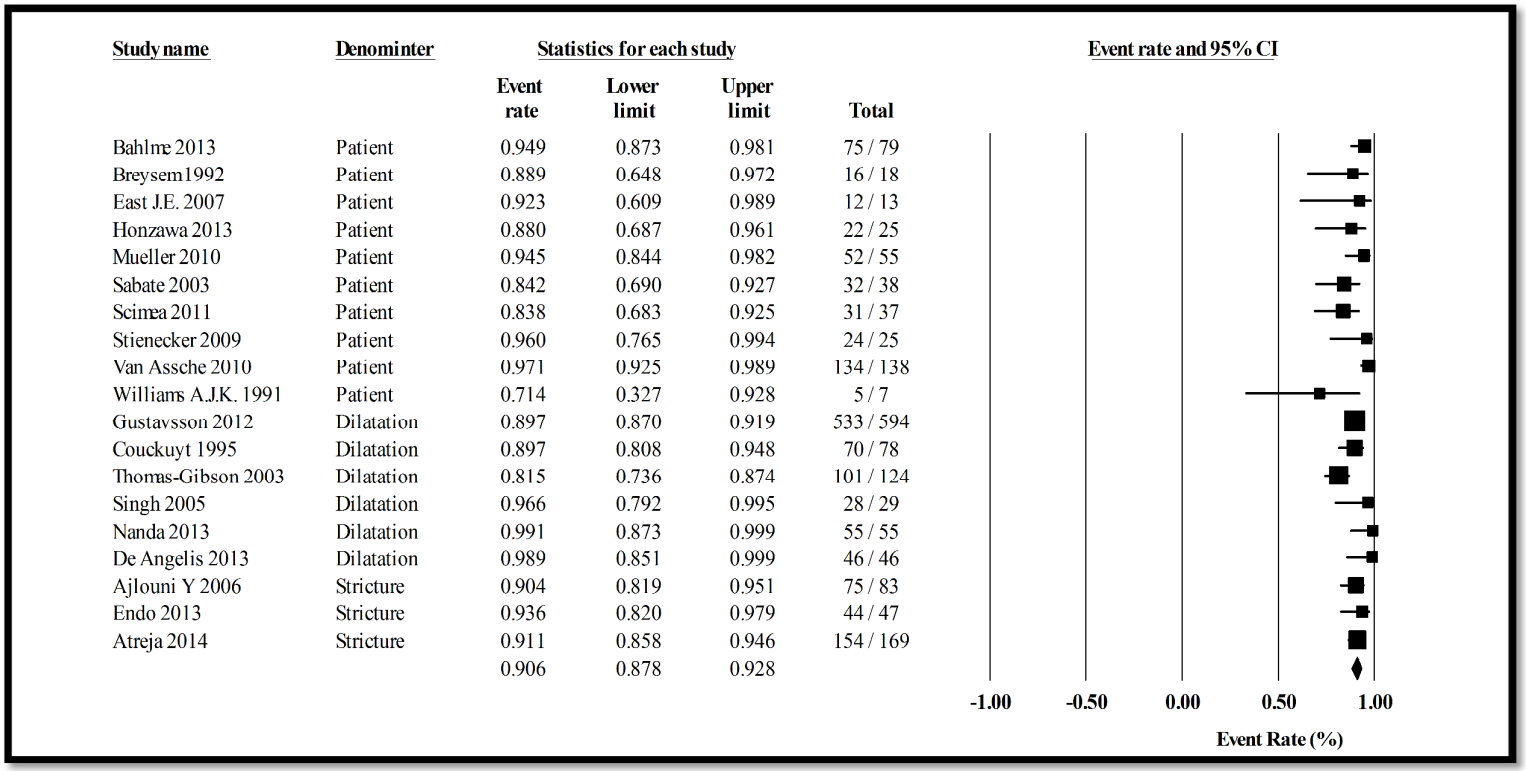
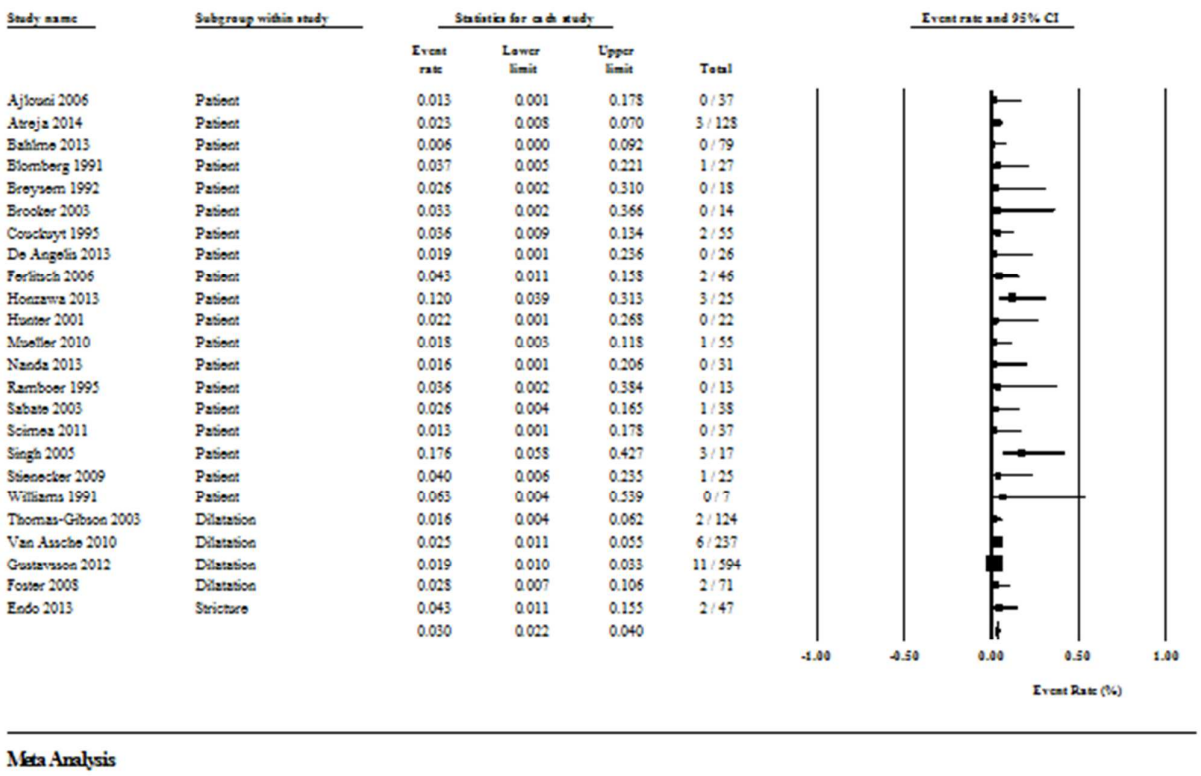
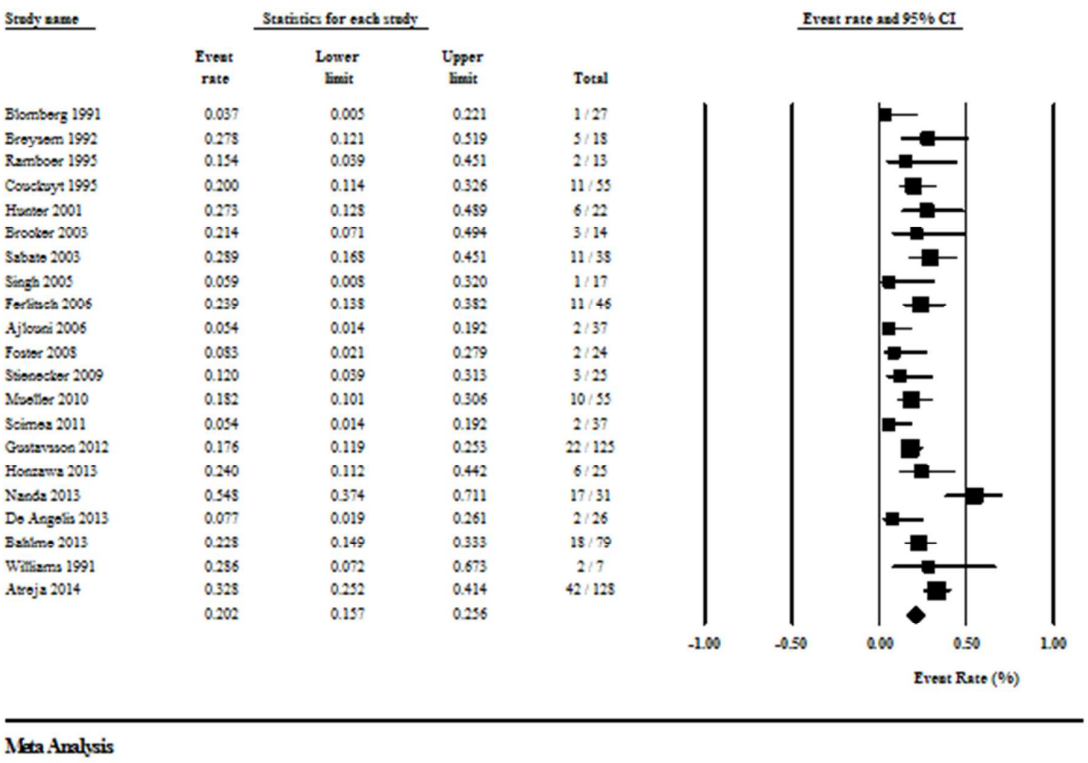


Figure 1– Forest plot for studies reporting on technical response – *A random effects model demonstrating a pooled technical response event rate of 90.6% (95% CI: 87.8 – 92.8%; I^2: 11.7%) with reported outcomes expressed for number of patients* [3, 5, 10, 15, 16, 19, 20, 22, 23, 25], *strictures* [1, 2, 11] *and interventions/dilatations* [7, 8, 14, 17, 24, 26].

Supplementary Figure 2 – Forrest plot reporting on perforation rates – A random effects model demonstrating a pooled perforation rate of 3% (95% CI: 2.2-4.0%; I2: 0%) across 22 studies with reported outcomes expressed according to number of patients,[20-28 34-41 44] balloon dilatation,[31-33 42 43] and strictures.[30]



Supplementary Figure 3 – Forrest plot reporting on the rate of surgical intervention in the event of a failed clinical outcome – A random effects model demonstrating a pooled surgical intervention rate of 23.3% (95% CI: 20-26.5; I2: 13.2%) across reported outcomes expressed according to the number of patients.[20-28 31-41 44]



Supplementary material – supplementary text

RESULTS

Description of population for symptomatic response outcomes

Stricture activity

There was variation in the reporting of stricture activity. Eight studies reported stricture activity expressed for number of patients (N = 290) [23-27, 32, 38, 39]. The proportion of patients with active and quiescent strictures was 53.4% (155/290) and 41.7% (121/290) respectively. One study reported disease activity based on number of strictures (N = 20) [41]. The proportion of strictures that showed active and no inflammation was 80% (16/20) and 20% (4/20) respectively.

Stricture type

Twelve studies described stricture types for patient numbers (N = 540) [23-27, 29, 32, 37-39, 43, 44]. The proportion of patients with anastomotic and de novo strictures was 80.5% (435/540) and 18.1% (98/540). Five studies expressed stricture types as a proportion of total stricture numbers (N = 189) [28, 33, 36, 40, 41]. The proportion of active and de novo strictures was 45.5% (86/189) and 54% (102/189).

Intervention technique

Balloons inflated up to 18mm [25, 28, 29, 36-38, 43, 44], 20mm [23, 26, 33, 40, 41], and 25mm [24, 27] demonstrated a pooled mean symptomatic response rate of 61.9% (95% CI: 47.4 – 74.4%; I^2 27%), 79.5% (95% CI: 72.6-85%; I^2 0%) and 71% (95% CI: 48 – 86.6%; I^2 0%) respectively. The mean symptomatic response rate with a balloon dilatation duration of up to 2 minutes [23, 25, 27, 28, 33, 36, 38, 40, 41, 43, 44] and up to 5 minutes [24, 29, 37] was 70.6% (95% CI: 58.1-80.6%; I^2 9.1%) and 67.1% (95% CI: 41.9 – 85.2%; I^2 0%) respectively.

Geography

Comparison of European and North American symptomatic success rate, which may reflect different patient characteristics or techniques, showed the pooled mean symptomatic success rate of 67.9% (95% CI: 57-77.2%; I^2 2%) in Europe studies [23-29, 36-38, 40, 43, 44] compared with 84.5% (95% CI: 62.6 – 94.7%; I^2 0%) in North American studies [33, 41].

Imaging

The pooled mean symptomatic response rate across twelve studies that described the use of pre-interventional imaging was 70.4% (95% CI: 58.8 – 79.8%; I^2 0%) [23, 26-29, 36, 37, 39-41, 43, 44]. The median maximum stricture length across fifteen studies was 7cm (range 3 - 25cm).

Description of population for technical response outcomes

Stricture activity

Seven studies reported on the stricture activity. Five studies assessed activity based on the number of patients (N = 220) [23, 25, 27, 31, 39]. The proportion of patients with active and quiescent strictures was 51.4% (113/220) and 44.5% (98/220) respectively. Only two studies reported on disease activity according to the stricture population (N = 189) [22, 41]. The proportion of strictures that demonstrated active and quiescent disease was 59.2 % (112/189) and 35.4% (67/189) respectively.

Stricture type

Seventeen studies reported on whether the stricture was anastomotic or de novo. Eight studies assessed this as per the number of patients (N = 443) [23, 25, 27, 37, 39, 42-44]. The proportion of patients with anastomotic and de novo strictures were 77% (341/443) and 23% (102/443). Eight studies assessed stricture type as per the number of strictures involved

(N=488) [21, 22, 28, 31, 35, 36, 40, 41]. The proportion of anastomotic and de novo strictures was 42.8% (209/488) and 57.2% (279/488) respectively. Only one study assess stricture type as per the number of balloon dilatation performed (N = 594) [34]. The proportion of anastomotic and de novo strictures in this study was 81.6% (487/597) and 17.9% (107/597) respectively.

Intervention technique

Balloons inflated up to 18mm [25, 28, 36, 37, 42-44], 20mm [21-23, 30, 31, 35, 40, 41, 45], and 25mm [27, 34, 39] reported a pooled mean technical success rate of 94.7% (95% CI: 87-97.9%; I^2 0%), 90.6% (95% CI: 87.4-93.0%; I^2 39.3%), and 89.4% (95% CI: 86.9-91.5%; I^2 0%) respectively. An inflation time of up to 2 minutes [23, 25, 27, 28, 36, 40, 41, 43, 44] and up to 5 minutes [31, 34, 37] demonstrated a pooled mean technical success rate was 92.2% (95% CI: 86.6-95.5%; I^2 0%) and 92.6% (95% CI: 84.2-96.7%; I^2 20.6%) respectively.

Geography

The pooled mean technical success rates across the 14 European [23, 25, 27, 28, 30, 34, 36, 37, 39, 40, 42-45], 2 North American [22, 41] and 2 Japanese [31, 35] studies were 90.6% (95% CI: 86.7-93.5%; I^2 18.8%), 91.2% (95% CI: 87.8-93.8%; I^2 0%) and 91.3% (95% CI: 81.9-96.0%; I^2 0%), respectively.

Imaging

The pooled mean technical success rate across 15 studies that reported the use of pre-dilatation imaging to further define the stricture was 91.4% (95% CI: 88.3-93.8%; I^2 13.8%) [23, 30, 36, 39, 40, 43]. The median maximum stricture length reported across 17 studies is 7.2cm (range 3cm – 25cm)

Description of population for perforation outcomes

Stricture Activity

There were nine studies that reported stricture activity according to the number of patients (N = 320) [23-27, 29, 31, 32, 38, 39]. The proportion of patients with active and quiescent disease was 48.4% (155/320) and 47.1% (151/320) respectively. There are two studies that reported disease activity according to the number of strictures (N = 189) [22, 41]. The proportion of active and quiescent strictures in these studies were 59.2% (112/189) and 35.4% (67/189).

Stricture type

There were 12 studies that reported the type of stricture according to the number of patients (N = 540) [23-27, 29, 32, 37-39, 43, 44]. The proportion of patients with anastomotic and de novo strictures was 80.5% (435/540) and 20.5% (111/540). Nine studies reported the type of stricture according to the number of strictures (N = 517) [21, 22, 28, 31, 33, 35, 36, 40, 41]. The proportion of anastomotic and de novo strictures was 42.7% (221/517) and 57% (295/517).

Technique

Balloon inflation diameters of 18mm [25, 28, 29, 35-38, 42-44], 20mm [21-23, 26, 32, 35, 40, 41, 45] and 25mm [24, 27, 34, 38] demonstrated pooled mean perforation rates of 3.1% (95% CI: 1.9-5.1%; I^2 0%), 3.8% (95% CI: 2.2-6.5%; I^2 0%), and 2.2% (95% CI: 1.3-3.6%; I^2 0%) respectively. An inflation time of up to 2 minutes [23, 25, 27-29, 33, 36-38, 40, 41, 43, 44] and 5 minutes [24, 31, 34, 36] demonstrated a pooled mean perforation rate of 3.1% (95% CI: 1.8-5.3%; I^2 0%) and 3.6% (95% CI: 1.6-8.2%; I^2 0%) respectively.

Geography

The mean perforation rate across 18 European studies that reported perforation according to the number of patients was 2.3% (95% CI: 1.6-3.3%; I^2 0%) [23-29, 32, 34, 36-40, 42-45]. The mean

perforation rate across three North American studies was 5.0% (95% CI: 1.3-17.7%; I^2 0%) [22, 33, 41].

Imaging

The use of pre-interventional imaging was described across 14 studies [22, 23, 27, 28, 31, 34, 36, 37, 39-44] where the pooled mean perforation rate was 2.7% (95% CI: 1.8-4.0%; I^2 0%). The perforation rate in one study that did not use pre-interventional imaging was 1.3% (95% CI: 0.1-17.8%) [21]. The median maximum stricture length reported across 20 studies was 7cm (range 2-25cm).

2. STRUCTURED SUMMARY

Background

Endoscopic balloon dilatation is a recognised treatment for symptomatic Crohn's strictures. Several case studies report its short term and long term efficacy. A systematic analysis of the current literature is needed to define its overall efficacy and inform the design of future studies.

Objectives

The aims of the review were to examine symptomatic response, technical response and adverse events of endoscopic balloon dilatation. Stricture characteristics that may impact on outcome were also explored.

Data sources

A systematic search strategy of COCHRANE, MEDLINE, EMBASE and OVID was performed in accordance with PRISMA guidelines.

Study Eligibility Criteria

Original studies (including randomised controlled trials, observational reports and case series with sample size more than 5) reporting outcomes of endoscopic balloon dilatation for Crohn's disease intestinal strictures.

Participants and Interventions

Patients with Crohn's intestinal strictures undergoing endoscopic balloon dilatation in the adult population (age > 16).

Study Appraisal and Synthesis Methods

The quality of individual studies was assessed by using the Newcastle-Ottawa Scale. A per patient analysis was used to determine the cumulative proportion of patients within a group. A per study analysis was used to assess pooled event rates across studies. The random effects model was used and results were expressed with forest plots and summative statistics. Heterogeneity across studies was assessed visually with forest plots and numerically.

Results

Twenty-five studies were included in the final review capturing 1089 patients, 790 strictures and 2664 dilatations. Overall, the proportion of patients who achieved symptomatic and technical response was 63.9% (393/615) and 92.6% (403/435) respectively. The proportion of patients who experienced perforation was 2.6% (18/700). The pooled event rates for symptomatic and technical response was 74.8% (95% CI: 69.9-79.3%) and 90.6% (95% CI: 87.8 – 92.8%) with no (I^2: 0.0%) and moderate heterogeneity (I^2: 11.7%) respectively. The pooled event rates for complications was 6.4% (95% CI: 5.0 – 8.2; I^2: 4.0%) and perforation was 3% (95% CI: 2.2-4.0%; I^2: 0%) respectively. Inflammatory activity may be associated with lower symptomatic response and higher perforation rate. More anastomotic strictures, were associated with low symptomatic and technical response rates.

Limitations

The lack of control groups precluded comparisons between the effects of variables on outcome measures directly. Selection bias is anticipated to influence the magnitude of response. Most studies were reported by gastroenterologists, with a bias towards showing endoscopic benefit as both performance and reporting biases.

Conclusions and implications of key findings

Endoscopic balloon dilatation for symptomatic Crohn's strictures shows high efficacy with a low complication rate. Efficacy may be affected by active inflammation, previous surgery

and choice of outcome measure. Future studies should identify valid, precise and relevant outcome measures and examine differential effects on stricture types.

FIGURE LEGENDS

Figure 1 – Flow chart demonstrating the search strategy in accordance with PRISMA – *Two hundred and three (n=203) records were identified following duplicate removal. Fifty three (n=53) records were removed after limits were applied. One hundred and fifty (n=150) records underwent screening and one hundred and twenty records were excluded (n=120). Thirty records (n=30) were assessed for eligibility and twenty five (n=25) articles were included in our quantitative analysis.*

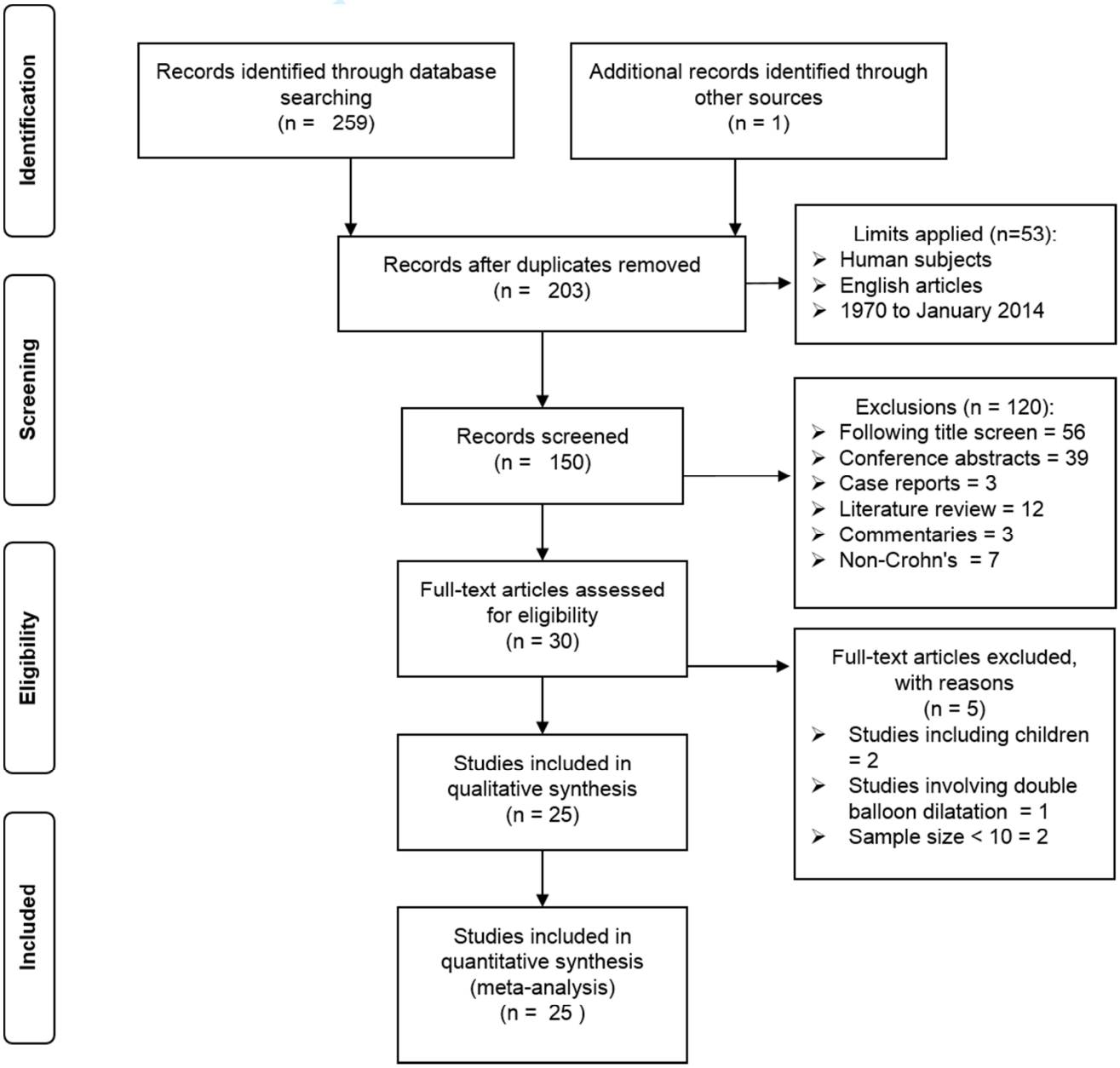


Figure 2 – Funnel plot for studies reporting on symptomatic response rate - *Heterogeneity across studies was demonstrated (I^2: 63.8%). Six studies*[24 27 32 36 42 43] *were outliers and determined as sources for publication bias on sensitivity analysis.*

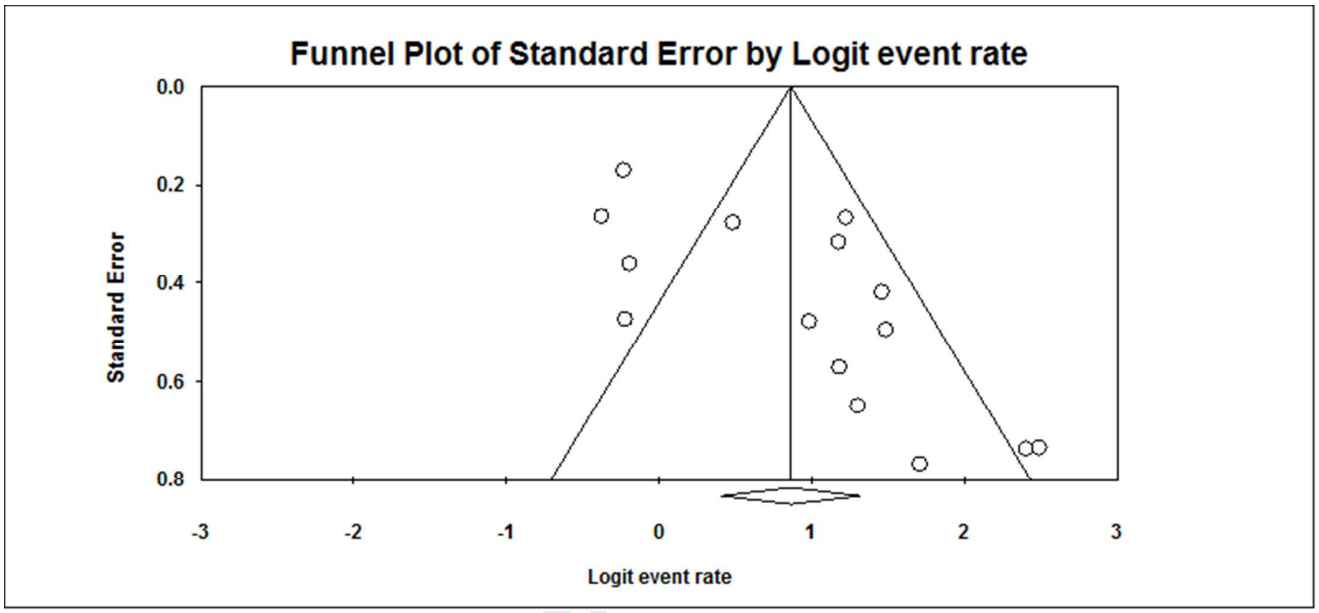


Figure 3 – Forrest plot for studies reporting on symptomatic response following exclusion of outlier studies[24 27 32 36 42 43] *– Random effects model demonstrating a pooled event rate for symptomatic response as 74.8% (95% CI: 69.6 – 79.3%; I^2: 0%).*