AP&T Alimentary Pharmacology and Therapeutics

Systematic review with meta-analysis: endoscopic balloon dilatation for Crohn's disease strictures

P. S. Morar*,†, O. Faiz*,†, J. Warusavitarne*,†, S. Brown‡, R. Cohen§, D. Hind¶, J. Abercrombie**, K. Ragunath**, D. S. Sanders‡, I. Arnott††, G. Wilson††, S. Bloom§, N. Arebi*,† & on behalf of the Crohn's Stricture Study (CroSS) Group

*Department of Surgery and Cancer, Imperial College, London, UK.
†St Mark's Hospital, London, UK.
‡Sheffield Teaching Hospitals, Sheffield, UK.
§University College Hospital, London, UK.
¶Clinical Trials Research Unit, University of Sheffield, Sheffield, UK.
**NIHR Nottingham Digestive Disease Biomedical Research Unit, Queen's Medical Centre Campus, Nottingham University Hospitals NHS Trust, Nottingham, UK.
††Departments of Gastroenterology and Colorectal Surgery, Western General Hospital, Edinburgh, UK.

Correspondence to:
Dr N. Arebi, St Mark's Hospital, Watford Road, Harrow London HA1 3UJ, UK.
E-mail: naila.arebi@imperial.ac.uk

Publication data
Submitted 23 March 2015
First decision 16 April 2015
Resubmitted 5 August 2015
Resubmitted 7 August 2015
Resubmitted 10 August 2015
Accepted 11 August 2015
EV Pub Online 11 September 2015

As part of AP&T's peer-review process, a technical check of this meta-analysis was performed by Dr Y. Yuan. This article was accepted for publication after full peer-review.

SUMMARY

Background
Endoscopic balloon dilatation (EBD) is recognised treatment for symptomatic Crohn's strictures. Several case series report its efficacy. A systematic analysis for overall efficacy can inform the design of future studies.

Aim
To examine symptomatic (SR) and technical response (TR) and adverse events (AE) of EBD. Stricture characteristics were also explored.

Methods
A systematic search strategy of COCHRANE, MEDLINE and EMBASE was performed. All original studies reporting outcomes of EBD for Crohn's strictures were included. SR was defined as obstructive symptom-free outcome at the end of follow-up, TR as post-dilatation passage of the endoscope through a stricture, and adverse event as the presence of complication (perforation and/or bleeding). Pooled event rates across studies were expressed with summative statistics.

Results
Twenty-five studies included 1089 patients and 2664 dilatations. Pooled event rates for SR, TR, complications and perforations were 70.2% (95% CI: 60–78.8%), 90.6% (95% CI: 87.8–92.8%), 6.4% (95% CI: 5.0–8.2) and 3% (95% CI: 2.2–4.0%) respectively. Cumulative surgery rate at 5 year follow-up was 75%. Pooled unweighted TR, SR, complication, perforation and surgery rates were 84%, 45%, 15%, 9% and 21% for *de novo* and 84%, 58%, 22%, 5% and 32% for anastomotic strictures. Outcomes between two stricture types were no different on subgroup meta-analysis.

Conclusions
Efficacy and complication rates for endoscopic balloon dilatation were higher than previously reported. From the few studies with 5 year follow-up the majority required surgery. Future studies are needed to determine whether endoscopic balloon dilatation has significant long-term benefits.

Aliment Pharmacol Ther 2015; **42:** 1137–1148

doi:10.1111/apt.13388

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 ischaemia, leaking or infection, which drives healing by tissue fibrosis.[9, 10]

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.[11] In clinical practice, the therapeutic choice is determined by stricture characteristics: accessible, short and anastomotic strictures are best considered for endoscopic balloon dilatation (EBD) whereas endoscopically inaccessible, multiple and >5 cm 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.[12, 13] 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 the short-term inability to completely dilate, perforation and bleeding are recognised risks, whereas long-term risks are related to disease recurrence which may warrant further dilatation or surgery.[14]

Several studies report outcomes of EBD in Crohn's disease strictures.[15–39] These outcomes were collated in two systematic literature reviews published in 2007[14] and 2010.[40] The lack of pooled analysis of events rates and nonconformity with PRISMA guidelines, are discernible weaknesses of both reviews.[14] Moreover, one of the reviews included outcomes of double balloon enteroscopy as well as strictureplasty.[40]

Objectives

This systematic review was performed to describe the outcomes of EBD 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 EBD 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 October 2014, reporting outcomes of EBD 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 10 were all included. Case reports, studies reporting on multiple diagnoses and conference proceedings 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: COCHRANE, MEDLINE and EMBASE.

Search

The final search terms were ("Crohn's Disease" OR Crohn's) AND (strictu* 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, unblinded, check by the second reviewer. Extracted data were 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, 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 S1).

Summary measures

Symptomatic response was defined as patients with an obstructive symptom-free outcome at the end of follow-up, technical response by the passage of the scope following EBD 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 unweighted proportions of the size of the population studied.

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 and expressed as crude, unweighted 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., Englewood, NJ, USA) programme was used.

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). To compare the effect of the proportion of patients within each group (e.g. stricture activity and stricture type) on outcome, subgroup meta-analyses were performed and pooled group differences were summarised by the pooled relative risk and corresponding confidence intervals.

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).

They included 10 prospective studies ($N = 10$), 14 retrospective studies ($N = 14$) and 1 randomised control trial ($N = 1$).[15–39] Publication dates ranged from 1991 to October 2014 and originated from European,[17–24, 26, 28, 30–34, 36–39] (six studies from the UK),[17, 20, 23, 24, 38, 39] North American,[16, 27, 35] Japanese,[25, 29] and Australian,[15] institutions.

Study characteristics

The cumulative data for the 25 studies[15–39] included 1089 patients, 790 strictures and 2664 dilatations. Fifty-one per cent (557/1089) were females, 43% (466/1089) were males

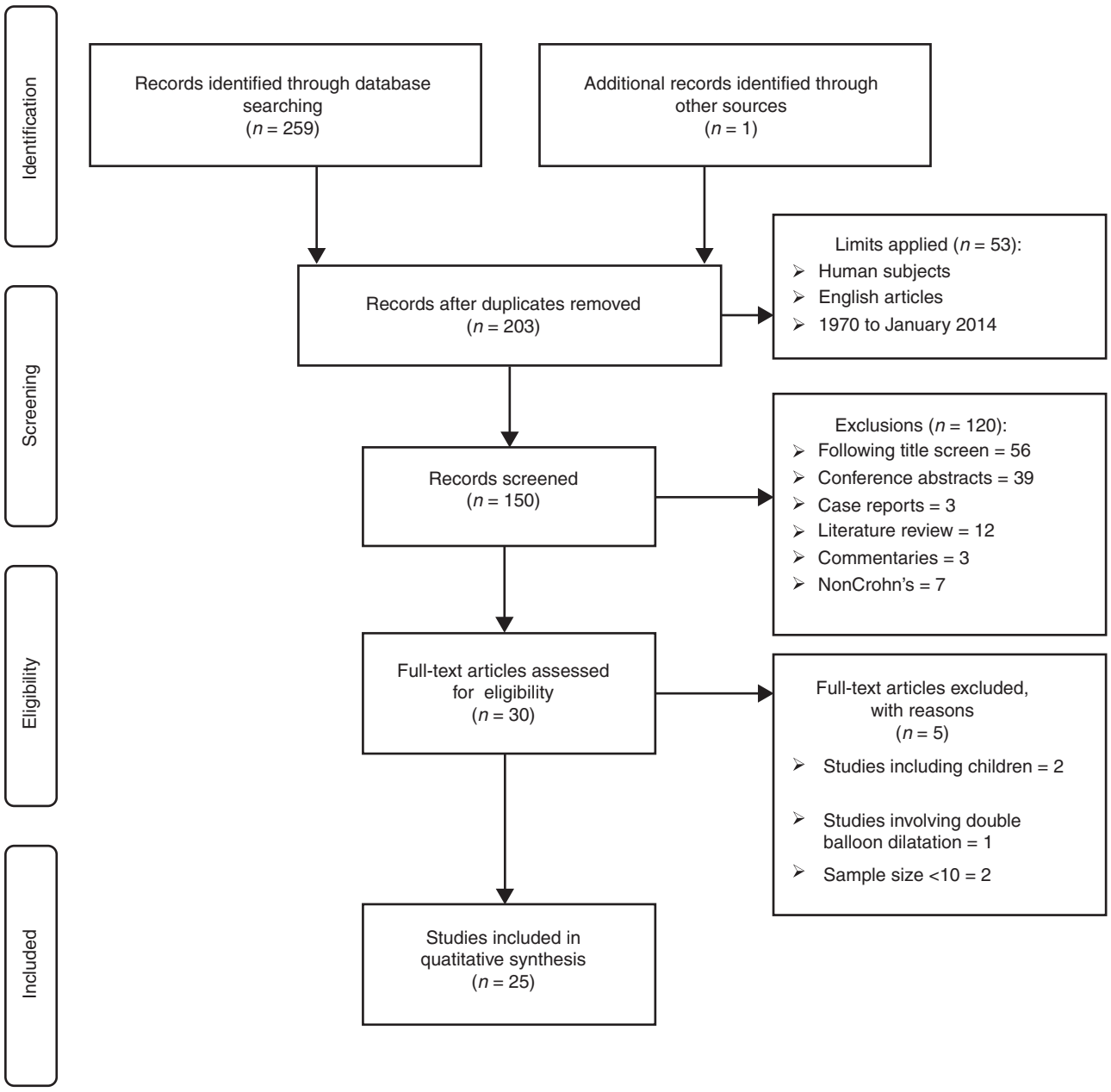

Figure 1 | Flow chart demonstrating the search strategy in accordance with PRISMA. Two hundred and three (n = 203) records were identified following duplicate removal. Fifty three (n = 53) records were removed after limits were applied. One hundred and fifty (n = 150) records underwent screening and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

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).[16, 18–23, 25, 27, 29, 31–37] Symptomatic response was reported as the outcome measure in five studies,[18, 20, 23, 27, 32] technical response in eight studies,[15, 16, 24, 25, 28, 29, 36, 39] and both symptomatic and technical response in 11 studies (Table 1).[17, 19, 21, 22, 30, 31, 33–35, 37, 38]

Follow-up duration was reported in 24 studies with wide variation.[15–19, 21–30] The median maximum follow-up time period was 83.5 months (range 12–172). The

Table 1 | Study demographics, frequency of population size, strictures and dilatations

Author	Country	Study design	Population size	Number of strictures (s)	Number of dilatations (d)	Symptomatic response occurrence, S	Technical response occurrence, E	Complication occurrence, C	Study outcome
Ajlouni[15]	Australia	R	37	83	113	NS	75 s	1	E, C
Atreja[16]	USA	R	128	169	430	NS	154 s	4	E, C
Bhalme[17]	UK	R	79	93	191	61	75	3	E, S, C
Blomberg[18]	Sweden	P	27	NS	NS	22	NS	3	S, C
Breysem[19]	Belgium	P	18	20	24	8	16	0	E, S, C
Brooker[20]	UK	R	14	14	26	11	NS	0	S, C
Couckuyt[21]	Belgium	P	55	59	78	34	70 d	6	E, S, C
De Angelis[22]	France	R	26	27	46	24	46 d		E, S
East[24]	UK	RCT	13	NS	NS	NS	12		E
Endo[25]	Japan	P	30	47	83	NS	154 s	5 s	E, C
Ferlitsch[26]	Austria	P	46	NS	73	NS	NS	3 d	C
Foster[27]	USA	R	24	29	71	22	NS	2	S, C
Gustavsson[28]	Sweden	R	125	NS	594	NS	533 d	41 d	E, C
Honzawa[29]	Japan	R	25	29	83	NS	22		E
Dear[23]	UK	R	22	NS	71	16	NS	0	S, C
Mueller[30]	Germany	P	55	74	93	42	52	1	E, S, C
Nanda[31]	Ireland	P	31	NS	55	14	55 d	0	E, S, C
Ramboer[32]	Belgium	P	13	15	53	11	NS	0	S, C
Sabate[33]	France	R	38	41	53	47d	32		E, S
Scimea[34]	Italy	P	37	39	72	30	31	0	E, S, C
Singh[35]	USA	R	17	20	29	13	28 d	4	E, S, C
Stienecker[36]	Germany	P	25	31	50	NS	24		E
Thomas-Gibson[38]	UK	R	59	NS	124	24	101 d	8 d	E, S, C
Van Assche[37]	Belgium	R	138	NS	237	61	134	12 d	E, S, C
Williams[39]	UK	R	7	NS	15	NS	5	1	E, C
Total			1089	790	2664				
Pooled unweighted event rate (%)						393/615 (63.9)	403/435 (93)	25/564 (4)	

RCT, randomised control trials; R = Retrospective; P = Prospective; NS = not specified Study Outcome Measures, S = symptomatic response, E = Endoscopic/Technical response, C = Complication, s = number of strictures, d = number of 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. Studies described both E and S outcomes [E = 8, S = 6 and E, S = 11]. Fifteen studies (n = 615) reported symptomatic outcome per patient and one study per dilatation procedure (n = 53 dilatations, n = 38 patients). 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 (Sabate *et al.*[33]) reported response rate for the number of dilatations 89%, not number of patients receiving dilatation and was excluded from cumulative analysis. Ten studies reported technical response per patient. The number of patients reporting a technical response is shown. A total of 403 (93%) reported technical response with dilatation. Fifteen studies reported the occurrence of a complication following dilatation. The number of patients reporting complications in each study is shown. A total of 25 (4%) patients reported perforation following dilatation. For all cumulative analyses, studies reporting on the number of strictures dilated (s) and the number of dilatation performed (d) were excluded.

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.[15–17, 19–22, 25, 27, 29, 30, 32–36] Fifteen studies examined lower gastrointestinal strictures only[18, 23, 24, 26, 28, 31, 37, 38] while 10 included both upper and lower GI strictures. The proportion of upper GI strictures was 3.8% (though due to missing data in nine studies, this was a estimation).

Stricture activity. The proportion of patients with active[17–21, 32, 33] and quiescent[17–21, 25, 26, 32, 33] 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.[17, 19, 20, 23, 28–31, 34, 37]

Stricture type. Across thirteen studies (N = 565),[17–21, 23, 26, 31–33, 36–38] most patients (79.1%; 447/565) had anastomotic strictures and 19.6% (111/565) had *de novo* strictures.[17–21, 23, 26, 31–33, 36–38]

Intervention technique. A maximum balloon diameter was reported across all 25 studies: 18 mm,[19, 22, 23, 30–32, 36–38] 20 mm,[15–17, 20, 24–27, 29, 34, 35, 39] and 25 mm.[18, 21, 28, 33] There was variation in the maximum inflation time across 23 studies with maximal inflation periods of 1,[17, 29, 30] 2,[15, 19, 21, 22, 27, 32–39], 3,[23–25] 4,[18, 28, 31] and 5 min.[26]

Imaging. The majority of studies describe pre-interventional imaging (18/20; 90%)[16, 17, 20–26, 28, 30, 31, 33–38] and two studies did not (2/20; 10%).[15, 27] The median maximum length of strictures reported across 21 studies was 7 cm (range 3–25 cm).

Synthesis of results

Symptomatic response. The population sample size across all 16 studies reporting on symptomatic response was 653. Fifteen studies reported on symptomatic response for patient numbers ($N = 615$)[17–23, 27, 30–32, 34, 35, 37, 38] which was 63.9% (393/615) (Table 1). Fourteen studies were from Europe [612/653 (93.7%)] spanning seven different countries (four from the UK)[17–23, 30–34, 37, 38] and two were North American studies [41/653 (6.3%)].[27, 35] 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 ($n = 59$).[38] 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).[18–23, 27, 31–35, 37] The proportion of patients who require further dilatation at 1, 2 and 5 year follow-up was 31.6% (160/506), 25.9% (117/451) and 1.7% (5/299). The cumulative proportion of patients requiring further dilatation over 5 years was 80.6% (241/299).

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).

The relationship between symptomatic response and variables is shown in Table 2, and Tables S2 and S3. Three studies,[19, 20, 35] reported subgroup data on stricture activity for symptomatic response with no evidence of heterogeneity across studies [I^2 0%]. Subgroup meta-analysis demonstrated no evidence of a difference of active vs. quiescent (reference) strictures (RR 0.91, 95% CI: 0.6–1.3, $P = 0.7$) (Table S2). Ten studies,[16, 19, 25, 27, 32, 34–36, 38, 39] reported subgroup data on stricture type for symptomatic response with no evidence

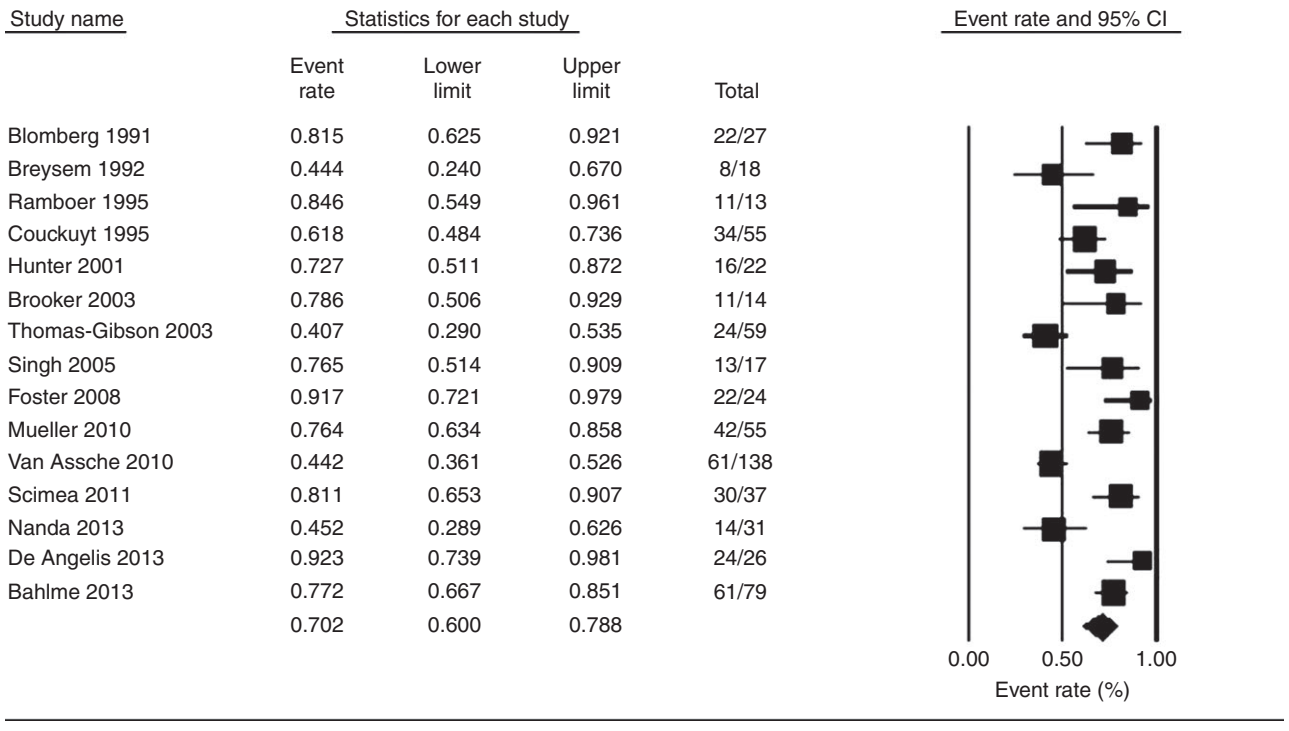

Figure 2 | Forest plot for studies reporting on symptomatic response (excluding Sabate *et al.*[33]). Random effects model demonstrating a pooled event rate for symptomatic response as 70.2% (95% CI: 60.0–78.8%; I^2 63.8%).

Table 2 | The relationship between response rates (symptomatic and technical) and adverse event rates (perforation) with categorical variables

Variable	Symptomatic response	Technical response	Perforation
Categorical variables represented as pooled event rates % (95% confidence intervals)			
Balloon diameter (mm)			
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 (min)			
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)

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.

of heterogeneity across studies [I^2 0%]. Subgroup meta-analysis demonstrated no evidence of a difference of anastomotic vs. *de novo* (reference) strictures [RR 1.1, 95% CI 0.96–1.2, $P = 0.2$] (Table S3).

Sensitivity analysis. In view of heterogeneous data, a sensitivity analysis was performed. Visual inspection of the funnel plot (Figure S1) demonstrated six outlier publications,[19, 22, 27, 31, 37, 38] and evidence of asymmetry indicating overestimation of the effect size as well as the possibility of publication bias. The symptomatic response rate in one study was measured according to the number of dilatations done and was reported as 89% (47/53).[33]

A comparison between symptomatic response rates and other variables is shown in Table 2.

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 S2). 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 ($N = 435$),[17, 19, 24, 29, 30, 33, 34, 36, 37, 39] stricture ($N = 299$),[15, 16, 25] and intervention/dilatation ($N = 926$) (Table 1).[21, 22, 28, 31 ,35, 38] The proportion of patients demonstrating technical response was 92.6% (403/435). The pooled event analysis results were similar due to low heterogeneity (Figure S2).

The relationship between technical response rates and variables is shown in Table 2, and Tables S2 and S3. Three studies[19, 20, 35] reported subgroup data for active and quiescent strictures for technical response with no evidence of heterogeneity across studies [I^2 0%]. Subgroup meta-analysis demonstrated no evidence of a difference of active vs. quiescent (reference) strictures (RR 0.98, 95% CI: 0.7–1.4, $P = 0.89$) (Table S2). Seven studies[16, 19, 22, 27, 35, 36, 39] reported subgroup data on stricture type for technical response with moderate evidence of heterogeneity across studies [I^2 43%]. Subgroup meta-analysis demonstrated no evidence of a difference of anastomotic vs. *de novo* (reference) strictures (RR 1.0, 95% CI 0.92–1.1, $P = 0.81$) (Table S3).

Complications and perforations. Fifteen studies reported complications according to the number of patients ($N = 564$)[15–21, 23, 27, 30–32, 34, 35, 39] and four studies reported complications based on the number of dilatations ($N = 1228$).[26, 28, 37, 38] The proportion of patients with complications was 4.4% (25/564) (Table 1). 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 ($N = 654$)[15–23, 29–36, 39] and four studies for number of dilatations ($N = 1281$).[26–28, 37, 38] The proportion of patients who had perforation was 2.4% (16/654). Study data analysis showed no heterogeneity across studies

with a pooled perforation rate of 3% (95% CI: 2.2–4.0%; I^2 0%) (Figure S3).

The relationship between perforation rates and variables is shown in Table 2, and Table S3. Balloon inflation diameters of 18 mm,[19, 22, 23, 29–32, 36–38] 20 mm,[15–17, 20, 26, 29, 34, 35, 39] and 25 mm[18, 21, 28, 32] demonstrated pooled 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 min[17, 19, 21–23, 27, 30–32, 34, 35, 37, 38] and 5 min[18, 25, 28, 30] demonstrated a pooled 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%).[17–23, 26, 28, 30–34, 36–39] The mean perforation rate across three North American studies was 5.0% (95% CI: 1.3–17.7%; I^2 0%) (Table 2).[16, 27, 35] The use of pre-interventional imaging was described across 14 studies,[16, 17, 21, 22, 25, 28, 30, 31, 33–38] where the pooled 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%).[15] The median maximum stricture length reported across 20 studies was 7 cm (range 2–25 cm) (Table 2).

Four studies[27, 29, 35, 38] reported subgroup data on stricture type for perforation with no evidence of heterogeneity across studies [I^2 0%]. Subgroup meta-analysis demonstrated no evidence of a difference of anastomotic vs. *de novo* (reference) strictures [RR 0.87, 95% CI 0.29–2.6, *P* 0.8] (Table S3).

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.

Endoscopic inaccessibility: Seven studies reported 12.9% of cases (33/256) where EBD could not be completed during endoscopy either because the stricture was too narrow or there was acute angulation.[19, 21, 26, 30, 33, 34, 39] Two studies (N = 130) reported this event 6.9% (9/130) by number of strictures.[15, 25]

Symptomatic disease: Twenty one studies reported surgical outcomes for ongoing recurrent disease despite repeated balloon dilatation according to the number of patients (N = 849).[15–23, 26–36, 39] The mean length of follow-up from studies where surgery rates are reported was 19.7 months. The proportion of patients who underwent surgery within a 5 year follow-up was 75% (341/455). The proportion of patients who had further surgery at 1, 2 and 5 year follow-up was 21.5% (153/712), 18.7% (118/632) and 8.4% (38/455) respectively. Pooled data demonstrated a mean surgical event rate of 20.2% (95% CI: 15.7–25.6), with low heterogeneity demonstrated across studies [I^2 13.2%] (Figure S4). One study reported the requirement of surgery according to the number of strictures (N = 47).[25] The proportion of strictures requiring surgery in this study was 29.8% (14/47).

One study[19] reported subgroup data on stricture activity for surgery. Subgroup meta-analysis demonstrated no evidence of a difference of active vs. quiescent (reference) strictures (RR 2.0, 95% CI: 0.5–8.0, P = 0.33) (Table S2). Eight studies[15, 16, 19, 27, 29, 34, 35, 38] reported subgroup data on stricture type for surgery with minimal evidence of heterogeneity across studies [I^2 15%]. Subgroup meta-analysis demonstrated no evidence of a difference of anastomotic vs. *de novo* (reference) strictures [RR 1.1, 95% CI 0.77–1.5, *P* 0.61] (Table S3).

DISCUSSION

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

Summary of evidence

This review offers insights into EBD for Crohn's strictures on which to build future more robust study designs to measure efficacy. The pooled event rate for symptomatic response of 70.2% includes six 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. Furthermore, the source of heterogeneity across studies may also be secondary to differences in the intervention performed. 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 58% response comparable to our 63.9% expressed as the unweighted proportion of patients but a lower value than the pooled effect.[14] 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 are that passage of the endoscope through the stricture is an inadequate outcome for patients. There was a wide variation in dilation techniques suggesting an overriding need for standardisation of endoscopic procedures. Balloon diameter of 20 mm seems to be commonest and most effective size limit, consistent with the internal small bowel diameter of 25 mm. Two minutes of dilatation is the commonest duration used and may be associated with better outcomes. The commonest reported dilatation technique was the three step technique increase in diameter with regular repeat dilatations 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. Moreover subgroup analysis of disease activity and type of stricture did not reveal differences in outcomes for stricture characteristics. With respect to dilatation technique maximum balloon diameter and duration of inflation did not seem to show different outcomes, except that 20 mm size was accompanied by higher symptomatic response rate than 18 mm (80% vs. 52%). There was no evidence of higher perforation rate with dilatation diameters of 25 mm. North America symptomatic response rates were higher than Europe, but were associated with higher complication rates.

Overall pooled complication rate was 6.4% (95% CI: 5.0–8.2), much higher than complication rates of 2% reported by Hassan *et al.*,[14] and 3% reported by Wilber *et al.*[40] 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 reviews did not report perforation and a separate event.[14, 40]

While the strength of this review lies in systematic manner in which it was conducted in accordance with PRISMA guidelines and methods for narrative reviews, it is beset by several limitations.[41]

Limitations

First, the absence of control groups for comparison, we created subgroups to examine the effect of stricture characteristics on outcome. However, as the number of studies where results were consistently described for stricture types was small, the lack of effect may be related to a type 2 statistical error due to the exclusion of a large proportion of the data. The exclusion of 39 conference abstracts may have contributed to this effect, however, the lack of methodological information would have made it difficult to assess the quality of data in this format.

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. Third, incomplete and variable reporting of some population and interventions characteristics means 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). Fourth, most studies were reported by gastroenterologists, with a bias towards showing endoscopic benefit through both performance and reporting bias. Also, 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: 70.2% 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. While there is no evidence of differential response of based on stricture type and inflammatory activity this if far from conclusive due to small percentage of studies describing these results instead 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, outcome measures and in pre-assessment for suitability of dilatation.

There is a discernible absence of randomised controlled trial of EBD 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 nonaccessible stricture 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 diam-

eter of the bowel lumen and also reduce TNF-induced fibrosis.[42, 43] 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 captures 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.

The place for EBD in the management of Crohn's strictures may need to be redefined. While a symptomatic response rate of 70.2% indicates short-term benefit the longer term avoidance of surgery is not so positive – 75% of patients with a 5 year follow-up data eventually required surgery. Since this comprises 42% of the original population of 1089 patients, it may be an overestimate as it may reflect a group with unfavourable disease and hence more likely to attend followed up. It also raises the question of how to predict patients with strictures that may be more effectively managed with surgery where EBD is an accepted low risk short-term measure. Future studies on this subject should therefore 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.[44] Without these studies an evidence-based management pathway to reduce variation and set standards in care for Crohn's strictures cannot be developed.

AUTHORSHIP

Guarantor of the article: Naila Arebi.

Author contributions: PM, JW and NA performed the research. PM, and NA collected and analysed the data. PM, OF, DH, JA, NA, designed the research study and wrote the paper. SBr, RC, KR, DSS, IA, GW, SBl contributed to the design of the study. All authors have approved the final manuscript.

All named authors have met the criteria for authorship based on the International Committee of Medical Journal Editors Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals.

ACKNOWLEDGEMENTS

The authors acknowledge Dr Ravi Misra for study quality assessment, and Mr Paul Bassett for assistance with statistical analysis.

Declaration of personal interests: Naila Arebi has served as a speaker, a consultant and an advisory board member for Abbvie and MSD. Stuart Bloom has served as a speaker, a consultant and an advisory board member for MSD, Abbvie, Takeda, Hospira, Napp, Actavis, Shire, Vifor, Pharmacosmos. Stuart Bloom is an employee of UCL Hospitals NHS Foundation Trust.

Declaration of funding interests: None.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Funnel plot for studies reporting on symptomatic response rate. Heterogeneity across studies was demonstrated [I^2 63.8%]. Six studies[19, 22, 27, 31, 37, 38] were outliers and determined as sources for publication bias on sensitivity analysis. Evidence of asymmetry in the funnel plot was also present.

Figure S2. 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%) across 19 studies.

Figure S3. Forest plot reporting on perforation rates. A random effects model demonstrating a pooled perforation rate of 3% (95% CI: 2.2–4.0%; I^2 0%) across 22 studies with reported outcomes expressed according to number of patients,[20–28, 34–40, 43] balloon dilatation,[31–33, 41, 42] and strictures.[30]

Figure S4. Forest 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 20.2% (95% CI: 15.7–25.6; I^2 13.2%) across reported outcomes expressed according to the number of patients.[20–28, 31–40, 43]

Table S1. 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 studies was 5 (range 2–6).

Table S2. Meta-analysis in studies reporting on outcomes for active and quiescent strictures.

Table S3. Meta-analysis in studies reporting on outcomes for anastomotic and *de novo* strictures.

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**: 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**: 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. *Clin Gastroenterol Hepatol* 2007; **5**: 1430–8.
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**: 1147–55.
5. Wolters FL, Russel MG, Sijbrandij J, *et al.* Phenotype at diagnosis predicts recurrence rates in Crohn's disease. *Gut* 2006; **55**: 1124–30.
6. Rieder F, Zimmermann EM, Remzi FH, *et al.* Crohn's disease complicated by strictures: a systematic review. *Gut* 2013; **62**: 1072–84.
7. Graham MF, Diegelmann RF, Elson CO, *et al.* Collagen content and types in the intestinal strictures of Crohn's disease. *Gastroenterology* 1988; **94**: 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**: 373–85.
9. Scarpa M, Angriman I, Barollo M, *et al.* Role of stapled and hand-sewn anastomoses in recurrence of Crohn's disease. *Hepatogastroenterology* 2004; **51**: 1053–7.
10. Yamamoto T, Allan RN, Keighley MR. Strategy for surgical management of ileocolonic anastomotic recurrence in Crohn's disease. *World J Surg* 1999; **23**: 1055–60; discussion 60-1
11. 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**: 28–62.
12. 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. *Dis Colon Rectum* 2012; **55**: 1125–30.
13. 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. *Dis Colon Rectum* 2005; **48**: 2302–8.
14. Hassan C, Zullo A, De Francesco V, *et al.* Systematic review: endoscopic dilatation in Crohn's disease. *Aliment Pharmacol Ther* 2007; **26**: 1457–64.
15. Ajlouni Y, Iser JH, Gibson PR. Endoscopic balloon dilatation of intestinal strictures in Crohn's disease: safe alternative to surgery. *J Gastroenterol Hepatol* 2007; **22**: 486–90.
16. 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**: 392–400.
17. 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**: 265–70.
18. Blomberg B, Rolny P, Jarnerot G. Endoscopic treatment of anastomotic strictures in Crohn's disease. *Endoscopy* 1991; **23**: 195–8.
19. 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**: 142–7.
20. 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**: 333–7.
21. 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**: 577–80.
22. 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**: 2660–7.
23. Dear KL, Hunter JO. Colonoscopic hydrostatic balloon dilatation of Crohn's strictures. *J Clin Gastroenterol* 2001; **33**: 315–8.
24. East JE, Brooker JC, Rutter MD, *et al.* A pilot study of intrastricture steroid versus placebo injection after balloon dilatation of Crohn's strictures. *Clin Gastroenterol Hepatol* 2007; **5**: 1065–9.
25. 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**: 86–91.
26. 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**: 483–7.
27. 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**: 880–5.
28. 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**: 151–8.
29. 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. *Dig Endosc* 2013; **25**: 535–43.
30. Mueller T, Rieder B, Bechtner G, *et al.* The response of Crohn's strictures to endoscopic balloon dilation. *Aliment Pharmacol Ther* 2010; **31**: 634–9.
31. 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**: 474–80.
32. 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**: 252–5.
33. Sabate JM, Villarejo J, Bouhnik Y, *et al.* Hydrostatic balloon dilatation of Crohn's strictures. *Aliment Pharmacol Ther* 2003; **18**: 409–13.
34. Scimeca D, Mocciaro F, Cottone M, *et al.* Efficacy and safety of endoscopic balloon dilation of symptomatic intestinal Crohn's disease strictures. *Dig Liver Dis* 2011; **43**: 121–5.
35. 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**: 284–90.
36. Stienecker K. Long-term results of endoscopic balloon dilatation of lower gastrointestinal tract strictures in Crohn's disease: a prospective study. *World J Gastroenterol* 2009; **15**: 2623.
37. 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**: 320–4.

38. Thomas-Gibson S, Brooker JC, Hayward CM, *et al.* Colonoscopic balloon dilation of Crohn's strictures: a review of long-term outcomes. *Eur J Gastroenterol Hepatol* 2003; **15**: 485–8.
39. Williams AJ, Palmer KR. Endoscopic balloon dilatation as a therapeutic option in the management of intestinal strictures resulting from Crohn's disease. *Br J Surg* 1991; **78**: 453–4.
40. Wibmer AG, Kroesen AJ, Grone J, *et al.* Comparison of strictureplasty and endoscopic balloon dilatation for stricturing Crohn's disease–review of the literature. *Int J Colorectal Dis* 2010; **25**: 1149–57.
41. University of York. *Systematic Reviews. CRD's Guidance for Undertaking Reviews in Health Care*. York, England: Centre for Reviews and Dissemination, 2009.
42. Pelletier AL, Kalisazan B, Wienckiewicz J, *et al.* Infliximab treatment for symptomatic Crohn's disease strictures. *Aliment Pharmacol Ther* 2009; **29**: 279–85.
43. Bouhnik Y, Laharie D, Stefanescu C., *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. London: Sage Publications, 2014; A1–131.
44. 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**: 950–7.