Introduction

Historically, the small bowel has been difficult to visualize with conventional endoscopic techniques. With the advent of balloon-assisted enteroscopy (BAE), both single-balloon enteroscopy (SBE) and double-balloon enteroscopy (DBE) have proven to be effective tools to safely visualize the small bowel [1]. SBE and DBE tend to be longer procedures (95 minutes and 105 minutes, respectively) when compared to other forms of endoscopy, thus they utilize larger volumes of air [1]. Room air insufflation is most commonly used to distend the lumen to achieve ideal visualization of the bowel. However, the use of air is not optimal as large fractions of air remain trapped within the bowel which must pass through the remaining gastrointestinal tract in order to escape [2]. Pain and discomfort commonly occur following these lengthy procedures and are often attributed to the remaining air that continues to distend the bowel [3, 4].

Unlike room air, carbon dioxide (CO$_2$) is highly diffusible, incombustible, and rapidly absorbed through the bowel wall, qualifying it as an ideal alternative to insufflate the bowel lumen [5]. Dozens of studies have explored the potential advantages and disadvantages of utilizing CO$_2$ in various types of endoscopic procedures. Individual randomized control trials (RCTs) have generated mixed results on outcomes such as procedure times, intubation depth, and abdominal pain and discomfort [6, 7]. Most recently, Wang et al. performed a meta-analysis of the effect of CO$_2$ use on multiple endoscopic procedures but were unable to definitively conclude its potential impact on BAE [5]. Therefore we performed a systematic review and meta-analysis to evaluate the safety and efficacy of CO$_2$ insufflation as compared to room air during BAE.

Background and study aims

Carbon dioxide (CO$_2$) insufflation has been suggested to be an ideal alternative to room air insufflation to reduce trapped air within the bowel lumen after balloon assisted enteroscopy (BAE). We performed a systematic review and meta-analysis to assess the safety and efficacy of utilizing CO$_2$ insufflation as compared to room air during BAE.

Patients and methods

The primary outcome is mean change in visual analog scale (VAS; 10cm) at 1, 3, and 6 hours to assess pain. Secondary outcomes include insertion depth (anterograde or retrograde), adverse events, total enteroscopy rate, diagnostic yield, mean anesthetic dosage, and PaCO$_2$ at procedure completion. We searched MEDLINE and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception until May 2015. Multiple independent extractions were performed, the process was executed as per the standards of the Cochrane collaboration.

Results

Four randomized controlled trials (RCTs) were included in the meta-analysis. VAS at 6 hours favored CO$_2$ over room air (MD 0.13; 95% CI 0.01, 0.25; p = 0.03). Anterograde insertion depth (cm) was improved in the CO$_2$ group (MD, 58.2; 95% CI 17.17, 99.23; p = 0.005), with an improvement in total enteroscopy rate in the CO$_2$ group (RR 1.91; 95% CI 1.20, 3.06; p =0.007). Mean dose of propofol (mg) favored CO$_2$ compared to air (MD, –70.53; 95% CI –115.07, –25.98; P = 0.002). There were no differences in adverse events in either group.

Conclusions

Despite the ability of CO$_2$ to improve insertion depth and decrease amount of anesthesia required, further randomized control trials are needed to determine the agent of choice for insufflation in balloon assisted enteroscopy.
Patients and methods

Selection criteria

Any randomized controlled trial evaluating the efficacy of CO₂ compared to room air in patients undergoing SBE or DBE regardless of publication status (e.g., abstracts, unpublished studies etc.) were eligible for inclusion in the systematic review. Studies that were not a randomized controlled trial, did not have a control, or included specialized treatment groups were excluded. There was no restriction on patient ethnic group or gender.

Five authors (AS, SL, AL, AR, AK) independently extracted data on outcomes from all studies. Data were extracted using a standardized data abstraction form. The same five authors independently reviewed all titles/abstracts and selected studies for inclusion. We included all references that reported results of RCTs of CO₂ versus room air in patients undergoing SBE or DBE in this review.

Types of participants

We included studies that enrolled patients aged 18 years or older who were scheduled for diagnostic and/or therapeutic balloon assisted enteroscopy.

Outcomes

Primary outcomes included pain, for which mean change in visual analog scale (VAS; 10 cm) at 1, 3, and 6 hours post procedure was used to quantify pain experienced by the patient. Secondary outcomes included insertion depth in cm (anterograde or retrograde), adverse events, total enteroscopy rate, diagnostic yield, mean anesthetic dosage, and PaCO₂ at procedure completion.

Search methods

An electronic search of Cochrane Central Register of Controlled Trials (CENTRAL) and MEDLINE using a combination of MeSH and free text from inception to May 10, 2015 was performed. No language or age limits were used. The following search strategy was utilized: ("single balloon"[MeSH Terms] OR "single"[All Fields] AND "balloon"[All Fields]) OR "single balloon"[All Fields] OR "single"[All Fields] AND balloon[All Fields] AND enteroscopy[All Fields]) OR ("double-balloon enteroscopy"[MeSH Terms] OR "double-balloon"[All Fields] AND "enteroscopy"[All Fields]) OR "double-balloon enteroscopy"[All Fields] OR ("double"[All Fields] AND "balloon"[All Fields] AND "enteroscopy"[All Fields]) OR ("double balloon enteroscopy"[All Fields]).

To identify any recently completed studies that have not yet been published in full, we searched conference abstracts from the last 3 meetings (2013–2015) of the American College of Gastroenterology and Digestive Disease Week. We also hand-searched references of all included review articles and included these studies. Finally, in order to identify unpublished or ongoing studies, we searched ClinicalTrials.gov, Roche clinical trial protocol registry (www.roche-trials.com), Novartis clinical trials database (www.novctrld.com), Australian New Zealand Clinical Trials Registry (ANZCTR), and the metaRegister of Controlled Trials.

Data collection and analysis

Five authors (AS, AL, SL, AR, AK) reviewed all titles, abstracts, and full-text reports independently. Any disagreements between authors during the study selection were resolved by consensus.

Data extraction and management

Broadly, we extracted data on author names, location and setting, specific intervention and comparison details, outcomes and participants.

The same 5 authors independently extracted data according to Chapter 7 of the Cochrane Handbook for Systematic Reviews of Interventions using a standardized data extraction form containing the following items (The Cochrane Collaboration. Higgins JPT, Green S. Cochrane, Handbook for Systematic Reviews of Interventions Version 5.1.0. March 2011. In Internet: http://handbook.cochrane.org; 05/01/2014):

- General information: study title, authors, sources
- Study characteristics: study design, setting, duration of follow-up
- Patient characteristics: number of patients enrolled, number of patients included in the analysis
- Interventions: CO₂ vs. air for balloon assisted enteroscopy
- Outcomes: Pain, measured by mean change in visual analog scales (VAS; 10 cm) at 1, 3, and 6 hours post-procedure, insertion depths in cm (anterograde or retrograde), adverse events, total enteroscopy rate, diagnostic yield, mean anesthetic dosage, and PaCO₂ after procedure.

For studies with multiple publications, we used the publication with longest follow-up for extracting data. Earlier publications were used to extract data on methodology and baseline characteristics. In cases where the method of analysis was not specified by the investigators and only the number of events was reported, we used the number randomized as the denominator, i.e. we recorded results according to intention-to-treat (ITT) analysis.

Assessment of risk of bias

Five authors (AS, AL, SL, AR, AK) independently assessed the risk of bias in the included studies using The Cochrane Collaboration’s tool for assessing the risk of bias as outlined in the Cochrane Handbook for Systematic Reviews of Interventions based on extracted information (The Cochrane Collaboration. Higgins JPT, Green S. Cochrane, Handbook for Systematic Reviews of Interventions Version 5.1.0. March 2011. In Internet: http://handbook.cochrane.org; 05/01/2014). Any disagreements in data extraction were resolved by the senior author (PB). In addition to risk of bias, we evaluated the risk of random error by extracting data on the investigator’s pre-determined effect difference, alpha, power, and sample size.

Specifically, for assessment of risk of bias, we graded each component of methodological quality as low, high, or unclear.

We evaluated selection bias by assessing the investigators’ de-
scription of method of randomization and allocation concealment. See appendix A for further description of grading.

Unit of analysis issues
The unit of analysis for this review was individual study. In the case of repeated follow-up (e.g. reporting of results at 3 months and 6 months), we used the longest follow-up from each study. We considered recurring events (e.g. serious adverse events) as a single event that occurred in 1 patient (e.g. 4 instances of pneumonia in 1 patient were counted as 1 patient with pneumonia).

Missing data
As suggested in the Cochrane Handbook for Systematic Reviews of Interventions in the case of missing outcome data, we made an attempt to contact the principal investigator, corresponding author (or both) of the study (The Cochrane Collaboration. Higgins JPT, Green S. Cochrane, Handbook for Systematic Reviews of Interventions Version 5.1.0. March 2011. In Internet: http://handbook.cochrane.org; 05/01/2014). If the corresponding author was unable to provide the missing data for an outcome, the study was still included in the systematic review but excluded from the meta-analysis for the outcome with missing data. No imputation of missing individual patient data was undertaken.

Assessment of heterogeneity and reporting biases
To evaluate heterogeneity between pooled studies, we calculated $\chi^2$ and $I^2$ statistics. We considered an $I^2 >50\%$ to indicate substantial heterogeneity or a $\chi^2$ test, with the significance level set at $P<0.1$ to indicate statistically significant heterogeneity.

We planned to assess publication bias using a funnel plot if more than 10 studies were included in the review [8]. We evaluated selective reporting of outcomes within studies by comparing outcomes reported with outcomes specified in protocols, when available.

Data synthesis
We summarized dichotomous data as risk ratio (RR) along with 95% confidence intervals (CI) (i.e. clinical and histologic response) and continuous data (i.e. insertion depth, cm) as mean difference (MD) and standard error along with 95% CI using RevMan 5 software (version 5.1.6). We employed a random-effects model using the Der Simonian-Laird approach to pool studies for all analyses [9].

We constructed a Summary of Findings table using the most relevant and patient-relevant outcomes using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) guidelines [10–14]. These outcomes included mean change in visual analog scales (VAS; 10 cm) at 1, 3, and 6 hours post-procedure, insertion depths in cm (anterograde or retrograde), adverse events, total enteroscopy rate, diagnostic yield, mean anesthetic dosage, and PaCO2 after procedure. Additionally, we evaluated and summarized the quality of evidence for each outcome according to Grading of Recommendations Assessment, Development and Evaluation (GRADE) guidelines, which classify evidence as either very low, low, moderate, or high [10–14]. The systematic review has been performed and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [15].

Results
Studies
The initial search retrieved 1892 references that were screened by title and abstract. Among those 103 were selected for full text review. After the final screening, 4 published studies met the pre-determined inclusion criteria (Fig.1). No abstracts or unpublished studies met our inclusion criteria.

Methodologic quality of included studies
Overall methodologic quality of included studies ranged from moderate to very low.

Effects of interventions
Our analysis included 4 trials with 461 patients. Overall results for all outcomes and the quality of evidence for the comparison of treatment versus control are summarized in the summary of findings table (Table 1). The study design and conclusion of each study are described in (Table 2).

Benefits
Mean VAS score at 1 hour, CO2 vs. air
Data on mean VAS score at 1 hour were available from 4 studies (n = 461) [6, 7, 16, 17] The pooled analysis showed no statistical advantage between CO2 and air (MD, 0.10; 95% CI to 0.14, 0.34; $P=0.43$); See figure 2a. There was no substantial heterogeneity detected ($P=0.78$, $I^2=0\%$).

Mean VAS score at 3 hours, CO2 vs. air
Data on mean VAS score at 3 hours were available from 4 studies (n = 461) [6, 7, 16, 17]. The pooled analysis showed no statistical advantage when comparing CO2 to room air (MD, $−0.06$; 95% CI $−0.41$ to 0.29; $P=0.74$); See Fig.2b. There was no substantial heterogeneity detected ($P=0.22$, $I^2=33\%$).
Table 1  Summary of Findings.

<table>
<thead>
<tr>
<th>Quality assessment</th>
<th>No of studies</th>
<th>Design</th>
<th>Risk of bias</th>
<th>Inconsistency</th>
<th>In-directness</th>
<th>Imprecision</th>
<th>Other considerations</th>
<th>No of patients</th>
<th>Effect</th>
<th>Relative (95% CI)</th>
<th>Absolute</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean VAS at 1 hour (Better indicated by lower values)</td>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
<td>none</td>
<td>230</td>
<td>231</td>
<td>–</td>
<td>MD 0.1 points higher (0.14 to 0.34)</td>
<td>⊕⊕⊕Ο</td>
</tr>
<tr>
<td>Mean VAS at 3 hours (Better indicated by lower values)</td>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
<td>none</td>
<td>230</td>
<td>231</td>
<td>–</td>
<td>MD 0.06 points lower (0.41 to 0.29)</td>
<td>⊕⊕⊕Ο</td>
</tr>
<tr>
<td>Mean VAS at 6 hours (Better indicated by lower values)</td>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
<td>none</td>
<td>230</td>
<td>231</td>
<td>–</td>
<td>MD 0.13 points higher (0.01 to 0.25)</td>
<td>⊕⊕⊕Ο</td>
</tr>
<tr>
<td>Mean VAS at 24 hours (Better indicated by lower values)</td>
<td>3</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
<td>none</td>
<td>124</td>
<td>123</td>
<td>–</td>
<td>MD 0.11 points higher (0.03 to 0.24)</td>
<td>⊕⊕⊕Ο</td>
</tr>
<tr>
<td>Insertion Depth – Anterograde (Better indicated by lower values)</td>
<td>3</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>serious</td>
<td>none</td>
<td>210</td>
<td>211</td>
<td>–</td>
<td>MD 58.20 cm higher (17.17 to 99.23)</td>
<td>⊕ΟΟΟ</td>
</tr>
<tr>
<td>Insertion Depth – Retrograde (Better indicated by lower values)</td>
<td>3</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>very serious</td>
<td>none</td>
<td>210</td>
<td>211</td>
<td>–</td>
<td>MD 22.54 cm higher (49.08 to 94.16)</td>
<td>⊕ΟΟΟ</td>
<td>VERY LOW</td>
</tr>
<tr>
<td>Insertion Depth Overall (Better indicated by lower values)</td>
<td>3</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>serious</td>
<td>none</td>
<td>120</td>
<td>127</td>
<td>–</td>
<td>MD 22.96 cm higher (8.82 to 54.74)</td>
<td>⊕ΟΟΟ</td>
<td>VERY LOW</td>
</tr>
<tr>
<td>Any adverse Events</td>
<td>4</td>
<td>randomized trials</td>
<td>serious</td>
<td>no serious inconsistency</td>
<td>no serious indirectness</td>
<td>no serious imprecision</td>
<td>none</td>
<td>1/230 (0.43%)</td>
<td>2/231 (0.87%)</td>
<td>RR 0.63 (0.08 to 4.98)</td>
<td>3 fewer events per 1000 (8 to 34)</td>
<td>⊕⊕⊕Ο</td>
</tr>
</tbody>
</table>
Mean VAS score at 6 hours, CO2 vs. air
Data on mean VAS score at 6 hours were available from 4 studies (n = 461) [6, 7, 16, 17]. The pooled analysis favored CO2 over air (MD 0.13; 95% CI 0.01 to 0.25; \( P = 0.03 \)); See Fig. 2c. There was no substantial heterogeneity detected (\( P = 0.53; I^2 = 0\% \)).

Anterograde insertion depth, CO2 vs. air
Data on mean anterograde insertion depth (cm) were available from 3 studies (n = 261) [6, 7, 16, 17]. The pooled analysis favored CO2 over air (MD 58.2; 95% CI 17.17 to 99.23 \( P = 0.005 \)); See Fig. 3a. There was substantial heterogeneity detected (\( P < 0.0001; I^2 = 89\% \)).

Retrograde insertion depth, CO2 vs. air
Data on mean retrograde insertion depth (cm) were available from 3 studies (n = 421) [6, 16,17]. The pooled analysis showed no statistical difference between CO2 and air (MD 22.54; 95%...
There was no substantial heterogeneity detected (P<0.001; I² = 96%).

Overall insertion depth, CO₂ vs. air
Data on mean overall insertion depth (cm) were available from 3 studies (n = 247) [6, 7, 16, 17]. The pooled analysis showed no significant difference between CO₂ and air (MD 22.96; 95% CI – 8.82 to 54.74; P = 0.24); See Fig. 3c. There was no substantial heterogeneity detected (P = 0.27; I² = 24%).

Adverse events, CO₂ vs. air
Data on adverse events were available from 4 studies (n = 461) [6, 7, 16, 17]. However, 2 studies reported zero adverse events for both groups and therefore a summary measure was not derivable [6, 17]. The pooled analysis with the 2 remaining studies showed no statistical difference between CO₂ and air (RR 1.91; 95% CI 1.20 to 3.06; P = 0.007). There was no substantial heterogeneity detected (P = 0.38; I² = 0%).

Diagnostic yield, CO₂ vs. air
Data on diagnostic yield were available from 2 studies (n = 321) [6, 7, 16, 17]. The pooled analysis showed no statistical difference between CO₂ and air (RR 0.63; 95% CI 0.48 to 0.82; P = 0.66) [7, 16]. There was no substantial heterogeneity detected (P = 0.10; I² = 0%).

Total enteroscopy rate, CO₂ vs. air
Data on total enteroscopy rates were available from 2 studies (n = 239) [7, 16]. The pooled analysis showed no statistical difference between CO₂ and air (RR 0.63; 95% CI 0.48 to 0.82; P = 0.66) [7, 16]. There was no substantial heterogeneity detected (P = 0.10; I² = 0%).

CO₂ vs. air, mean anesthetic dosage ± SD, mg
Data on mean anesthetic dosage (mg), particularly amount of required propofol (mg) during BAE was available from 2 studies (n = 207) [6, 17]. The pooled analysis favored CO₂ over air (MD = 22.96; 95% CI – 8.82 to 54.74; P = 0.24). There was no substantial heterogeneity detected (P = 0.27; I² = 24%).

CO₂ vs. air, PaCO₂ after procedure
Data on measured serum PaCO₂ after the procedure was available from 2 studies (n = 239) [7, 16]. The pooled analysis showed no statistical difference between PaCO₂ levels between CO₂ and air (MD 1.20; 95% CI – 0.25 to 2.66; P = 0.10). There was no substantial heterogeneity detected (P = 0.38; I² = 0%).

Discussion
Carbon dioxide has historically been the most common gas used to distend the abdominal cavity to create pneumoperitoneum during laparoscopic abdominal procedures to provide sufficient operating space and adequate visualization [18]. With similar goals during endoscopic procedures, the role of CO₂ insufflation in gastrointestinal endoscopy is rapidly evolving.

A previous review by Wang et al. included just 2 RCTs exclusively studying DBE and was unable to determine the advantages of CO₂ insufflation [5]. Additional RCTs completed since the previous meta-analysis made updating these results essential. This meta-analysis nearly triples the study population since the previous analysis and is the first to include SBE, the newest form of balloon assisted enteroscopy.

This study found a significant improvement within the CO₂ group in anterograde insertion depth, with improved total enteroscopy rates in the subgroup stratification. While retrograde and total insertion depth analysis did not generate significant findings, it is important to note the majority of patients within this review underwent BAE procedures performed through the oral route. This is similar to findings throughout the literature, which not only finds anterograde enteroscopy most popular but also more effective regarding diagnostic and therapeutic yields [19].

We also found a reduction in the average dose of anesthetic, particularly propofol, required during the procedures in patients within the CO₂ group, which may clinically suggest a pain reducing effect of CO₂ during BAE. Pain assessed using the subjective VAS score was measured at 1, 3, and 6 post-procedure. A reduction in the VAS score of CO₂ patients only existed at 6 hours post-procedure. An improvement at a single in-
The interval of VAS testing is not strong enough evidence to suggest decreased pain with the use of CO₂ over air, especially as no difference occurred at either the 1- or 3-hour assessment. No differences emerged regarding the safety of CO₂ insufflation compared to air.

A few limitations within this study must be noted. While each study measures several outcomes, not all the numerical data were available to include within this meta-analysis. Secondly several cofounders such as patients with increased likelihood of stenosis, obstruction, or adhesions were not identified in each of the included studies. Depending on the outcome, the quality of evidence ranged from moderate to very low. A meta-analysis in itself has several limitations, which for this study included a limited number of outcomes to measure as studies have to measure the same outcomes in similar formats to be able to compare them using a meta-analysis. The study performed by Lentz et al. due to its size did have a greater weight on the analysis, which is not ideal due to the overall limited number of studies available. Nevertheless, for the main outcomes studied the results were consistent across studies and the quality of evidence was moderate. Lastly using studies that utilized both single balloon and double balloon technique during small bowel enteroscopy is not ideal, however, we feel due to the limited amount of data and number of quality studies in existence currently we gained more power by including both techniques for the inclusion criteria for this meta-analysis.

The inclusion of over 450 patients was a significant strength of this meta-analysis. Four robust RCTs exclusively studying the role of CO₂ insufflation during BAE generated much more statistical power compared to a single study. Lastly, the population was very diverse from multiple centers across the world and did not have significant ethnic disparity.

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>CO₂</th>
<th>Room air</th>
<th>Mean Difference</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domagk, 2007</td>
<td>0.28</td>
<td>3.6</td>
<td>0.39</td>
<td>0.48</td>
</tr>
<tr>
<td>Hirai, 2011</td>
<td>0.2</td>
<td>0.97</td>
<td>0.5</td>
<td>2.34</td>
</tr>
<tr>
<td>Lenz, 2014</td>
<td>0.68</td>
<td>0.86</td>
<td>0.49</td>
<td>0.79</td>
</tr>
<tr>
<td>Li, 2014</td>
<td>0.3</td>
<td>1.5</td>
<td>0.3</td>
<td>1.51</td>
</tr>
<tr>
<td>Total (95 % CI)</td>
<td>0.31</td>
<td>1.65</td>
<td>0.4</td>
<td>1.73</td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 0.00; Chi² = 1, df = 3 (P = 0.78); I² = 0 %
Test for overall effect: Z = 0.79 (P = 0.43)

![Fig. 2 a VAS at 1 hour. b VAS at 3 hours. c VAS at 6 hours.](image)
Conclusion

In conclusion this study determined several potential benefits of utilizing CO₂ rather than room air for insufflation during BAE such as the ability to improve insertion depth and decrease the amount of anesthesia required. However, these limited improvements are insufficient to declare CO₂ as the agent of choice over room air for insufflation in balloon assisted enteroscopy and further RCTs are needed.

Competing interests

None

References


