Use of magnetically controlled capsule endoscopy for the diagnosis of gastric diseases in adults: a systematic review and meta-analysis

Hao Zhang¹, Jingyao Chen², Jianfeng Li², Chumei Huang², Mingzhe Li², Wenhui Wu², Jianlong Jiang²

¹Hepatobiliary and Pancreatic Surgery, the Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, China; ²Digestive Diseases Center, the Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, China

Contributions: (I) Conception and design: J Jiang; (II) Administrative support: M Li, W Wu; (III) Provision of study materials or patients: J Jiang, H Zhang; (IV) Collection and assembly of data: H Zhang; (V) Data analysis and interpretation: J Chen, J Li, C Huang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jianlong Jiang. Digestive Diseases Center, the Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen 518000, China.
Email: jiangjlong@mail.sysu.edu.cn.

Background: To establish a systematic review and meta-analysis of all eligible studies to assess the diagnostic value of magnetically controlled capsule endoscopy (MCCE) in gastric diseases.

Methods: All published studies identified by search in Medline, EMBASE, The Cochrane Library databases and Web of Science were meta-analyzed. Conventional gastroduodenoscopy (CG) regarded as the golden standard of diagnosis. A pooled sensitivity, specificity, diagnostic odds ratio (DOR) and summary receiver operating characteristic curve (SROC) and corresponding 95% confidence region were calculated to estimate the overall accuracy by using the bivariate mixed-effects binary regression models. Bias of involved studies was evaluated according to the quality assessment of diagnostic accuracy studies (QUADAS-2).

Results: Of 832 studies, four studies from 2013 to 2019 meeting the inclusion criteria provided sufficient data (n=612). The pooled sensitivity, specificity and DOR were 91% (95% CI, 0.87–0.93), 90% (95% CI, 0.75–0.96) and 87.20 (95% CI, 27.38–277.74), respectively. The area under curve of SROC was 91% (95% CI, 0.88–0.94). Only few mild adverse events were observed.

Conclusions: MCCE has a high sensitivity (91%) and specificity (90%) in diagnosis of gastric disease, The AUC and DOR of MCCE was 0.91 and 87.2. In conclusion, MCCE shows acceptable sensitivity and specificity for the diagnosis of gastric diseases in adults.

Keywords: Magnetically controlled capsule endoscopy (MCCE); conventional gastroduodenoscopy (CG); sensitivity and specificity; meta-analysis

Introduction
Capsule endoscopy has become a very important tool in diagnosing small bowel diseases since it was firstly introduced in 2000 (1,2). Since then many attempts have been made to spread the role of capsule endoscopy to the upper gastrointestinal tract (3,4). To observe the gastric mucosal surface clearly and accurately, capsule endoscopy must overcome several difficulties such as the capacious and complicated anatomy of the stomach, peristaltic waves and a folded stomach, which are different from conventional gastroscopy (5,6). In order to solve these problems, magnetic manipulation of capsule endoscope was firstly manufactured in 2006 (7). With the development of technology, the magnetically controlled capsule endoscopy (MCCE) now becomes more feasible and safe (5,8). What is more, MCCE was greatly welcomed by patients for
it can represent a more patient-friendly substitute, and can be utilized without notable discomfort (9). To date, several researches have shown promising information to support the potentiality of MCCE (10). Swain et al. (11) suggested that it was easy to control the capsule back from the pylorus to the cardia esophageal junction and hold the capsule at any position in the stomach and did not cause discomfort. Rey et al. (12) also believed that magnetically navigated video capsule endoscopy was feasible and sufficiently accurate for gastric examination. However, Keller et al. (13), on the other hand, showed that MCCE cannot hold the capsule endoscopy against peristalsis when it approached the gastroesophageal junction because of its weak magnetic forces. Although there are several studies that have directly compared MCCE with CG, most of the available researches are limited by their small sample sizes, and no meta-analysis has been established about it. And we still don't know whether MCCE will exceed conventional gastroduodenoscopy (CG), which remains the gold standard diagnostic method in the upper gastrointestinal tract. Therefore, we here formulated the first systematic review and meta-analysis to clarify the sensitivity and specificity of MCCE and CG. We present the following article in accordance with the PRISMA reporting checklist (available at http://dx.doi.org/10.21037/dmr-20-141).

Methods

Search and strategy

PubMed, Embase, Cochrane Library and Web of Science were systematic searched independently by two authors (Jianlong Jiang and Hao Zhang) from established to December 2019. Language was not limited during the search. The terms “magnetic-controlled capsule endoscopy” “magnetically controlled capsule endoscopy” “magnetically guided capsule” “magnetic maneuverable capsule” “magnetic assisted capsule endoscopy” “magnetic maneuverable capsule” “conventional gastroscopy” “gastroscopy” “Standard video endoscope” “Traditional passive VCE” “Videendoscope” “VCE” “sensitivity and specificity” “diagnosis” were used. Regardless of language, year of publication, type of articles, or publication status, all related documents were taken into account. We manually searched the references lists of studies of interest for further potentially relevant studies. Wherever there were conflicts, they were discussed by two authors. The whole process was made according to the PRISMA guidelines.

Study election and data extraction

The inclusion criteria were: (I) participants should be healthy volunteers or have upper abdominal complaint; (II) participants must be given both MCCE and CG; (III) true positive (TP), false positive (FP), true negative (TN), false negative (FN) could be extracted; (IV) the golden standard was CG. The exclusion criteria were as follows: (I) case reports; (II) animal experiments; (III) meta-analyses and reviews; (IV) studies with insufficient information; (V) repeated studies were eliminated expect the most recent one; (VI) patients who suspected dysphagia or symptoms of gastric outlet obstruction, had a history of upper gastrointestinal surgery or abdominal surgery altering gastrointestinal anatomy, congestive heart failure, renal failure, etc. The references were screened and the data were extracted independently by two authors (Jianlong Jiang and Hao Zhang), when conflicts were encountered, the third author was consulted for the final decisions.

We collected the baseline characteristics of included studies, including: the first author, year of publication, countries, participants, age and sex distribution, the type of magnetic manipulation, the time cost by MCCE and CG operation and sample size. The primary outcomes were TP, FP, TN, FN; the secondary outcomes were the diagnostic consistencies, visualization of major stomach landmarks and numbers of side effects or complications.

Definition of outcomes

CG was regarded as the golden diagnostic standard. Both major lesions that would require endoscopic therapy or biopsy (ulceration, polyps, important bile reflux, fundic varicose, bleeding, hiatal hernia) and minor lesions (do not require subsequent gastroscopy) could be diagnosed as positive. The primary outcome measurements in our study were the diagnostic accuracy, sensitivity, and specificity of MCCE in identifying gastric diseases in healthy volunteer or patients with upper abdominal pain.

Index test

The index text was the applying of MCCE with studies describing evidence of gastric diseases in our analysis. In the worldwide, there are three types of MCCE, which mainly differs in the way of magnetic manipulation and the
amount of imaging sensor, as follows: robotic-arm magnetic manipulation (ANKON) with one imaging sensor, hand-held magnetic manipulation (Intromedic) with imaging sensor and robot-assisted magnetic manipulation (Olympus and Siemens) with two imaging sensors.

**Quality assessment of included studies**

Two authors independently evaluated each included literature according to QUADAS-2, and consensus was reached on all items. QUADAS-2 is composed of fourteen items and have already been structured so that four major issues are ranked for risk of bias and concerns regarding applicability to the research question was used to measure the included studies. Each key issue has a set of signaling questions to assess bias and applicability (14,15). We utilized tabular and graphical displays in Review Manager 5 (RevMan 5.3) to summarize the QUADAS-2 assessments.

**Statistical analysis**

To collect TP, FP, TN, FN, we formed a fourfold table for primary outcomes before we investigated each study. Then we performed the synthesis of the primary data within the bivariate mixed-effects binary regression modeling framework basing on the Cochrane DTA Working Group methodology. After obtaining all the data, we used Stata statistical software version 13.0 (Stata, College Station, TX) to pool the data and MetaDiSc statistical software (Meta-DiSc, version 1.4, Madrid, Spain) to estimate the threshold effect and meta-regression. A P value of less than 0.05 was considered statistically significant. We also calculated the positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR) and the summary receiver operating characteristics curve (SROC). Then we depicted a Fagan plot to evaluate meaningfulness or clinical efficacy. The Fagan nomogram is a very useful graphical instrument for assessing how much the effect of a diagnostic test adjusts the probability that a patient indeed has a disease (16). At last, we rated heterogeneity using forest plots and I^2 statistics. Larger values of I^2 prove increasing heterogeneity, with values of 25%, 50%, and 75% revealing low, moderate, and high degrees of heterogeneity, respectively. If there was significant heterogeneity, meta-regression was implemented to explain it. If not, then subgroup analysis was utilized.

**Results**

**Characteristics of included studies**

Four studies were finally selected (17-20). A total of 612 patients were assessed. A PRISMA flow chart of the search is shown in (Figure 1). At first, we collected a total of 832
documents through searching databases. Then we drop out duplications, case reports, animal experiments, reviews and articles that do not meet the included criteria. Finally, 4 studies were included. Furthermore, we collected baseline characteristics of included studies including the type of magnetic manipulation, time of MCCE and CG operation, distribution of gender and age, country, publication year and so on in Table 1. Three of the studies included only patients with upper abdominal discomfort and one study included only healthy volunteers. There were two kinds of MCCE as illustrated in (Table 1). Hand-held magnets are simple to use and portable. However, robotic control magnets are better at movement precision than hand-held magnets in porcine models (21). Both male and female patients were included. As mentioned in a study (22), the time of MCCE operation was greatly longer than that of CG, which leads to finding more minor lesions. Besides, the sequence of examinations could also cause bias when CG was given before MCCE, since CG can injury gastric mucosa. All participants in this meta-analysis were firstly given MCCE. In two study (16,18), the visualization of six gastric landmarks was recorded compared MCCE with CG.

Quality assessment of included studies

The quality of the eligible studies was evaluated by QUADAS-2 criteria and is illustrated (Figure 2). In all included studies, a high risk of bias concerning the selection of patients was discovered, because both healthy volunteers and patients with upper abdominal discomforts should be recruited in clinical practice. There was a low risk of bias issues of the index text, reference standard, flow and timing.

Pooled sensitivity and specificity

The pooled measurements of sensitivity and specificity of MCCE in the diagnosis of gastric diseases were computed as (Figure 3). Four total studies assessing gastric diseases demonstrated pooled of sensitivity and specificity of 91% (95% CI, 0.87–0.93) and 90% (95% CI, 0.75–0.96), respectively. The diagnostic accuracy of MCCE was 91% (95% CI, 0.88–0.94), Area under the curve (AUC) for MCCE was 0.91 (0.88–0.94) by ROC analysis (Figure 4), the PLR and NLR of MCCE were 9.14 (95% CI, 3.44–24.29) and 0.10 (95% CI, 0.07–15), respectively (Figure 5). The DOR was 87.2 (95% CI, 27.38–277.74) (Figure 6). There was no threshold effect (Spearman correlation coefficient =0.4, P=0.60), but high heterogeneity was found when pooling of specificity, we did not find the reason through meta-regression. Moreover, we did not observe significant publication bias among the studies as assessed according to the DEEK’s test (P=0.456) and Deeks plot (Figure 7).

The associated Fagan Nomogram is illustrated in (Figure 8). With a low pre-test probability (20%) of gastric diseases, if MCCE demonstrated the presence of gastric diseases, the post-test probability that the patient truly has gastric diseases would be approximately 70%. Alternatively, if the patient tests negative (i.e., no gastric diseases are seen on MCCE), the post-test probability that the patient indeed has gastric diseases would be lower than 3%.

Adverse events

Two studies reported minor discomforts due to the process of drinking too much water and swallowing the capsule (3 patients had abdominal distension and nausea, 1 patient had headache and vomiting, 1 patient had foreign body sensations and 1 patient had chronic diarrhea).

Discussion

This review was intended to confirm the sensitivity and specificity of MCCE in the diagnosis of gastric diseases in adults. We demonstrate that MCCE has a high sensitivity (91%) and specificity (90%) in diagnosis of gastric disease. In addition, we also found that AUC and DOR of MCCE was 0.91 and 87.2, which indicate MCCE has good capability at distinguishing patients from healthy. Collectively, our data proved that MCCE is a suitable alternative for gastric diseases that do not require biopsy.

Several studies (16,17) in related fields demonstrated that MCCE was able to detect gastric lesions with comparable accuracy with conventional gastroscopy and can be an alternative for screening gastric diseases without sedation. Likewise, we quantitatively pooled all the related studies, and the result confirmed that MCCE serves as a useful tool in diagnosis gastric disease. The Fagan nomogram for MCCE demonstrated that even in participants with a low pre-test probability, MCCE is an effective tool for expelling the presence of gastric diseases. However, two studies (18,19) reported that the visualization of gastric fundus and cardia in the up 1/3 stomach was lower than bottom 2/3, which caused by insufficient expansion of the stomach. In our study, the distribution of lesions almost was in the bottom 2/3 of stomach in healthy volunteers and patients with...
<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Country</th>
<th>Patient</th>
<th>Age (range)</th>
<th>Type of MCCE</th>
<th>Time of MCCE</th>
<th>Time of CG</th>
<th>Procedure</th>
<th>Positive No. cases (F/M)</th>
<th>Diagnostic accuracy</th>
<th>Visualization of gastric landmarks</th>
<th>No with side effects or complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhuan Liao</td>
<td>2016</td>
<td>China</td>
<td>Patient with upper abdominal complaint</td>
<td>46.6 [18–75]</td>
<td>Robotic arm (ANKON)</td>
<td>26.4 min</td>
<td>–</td>
<td>CG was given within 2 hours after MCCE</td>
<td>Major lesions 350 (186/164)</td>
<td>93.40%</td>
<td>Cardia 75.2%, fundus 73.2%, body 88.7%, Angulus 92.3%, antrum 96.6%, pylorus 97.4%</td>
<td>5 (3 patients had abdominal distension and nausea, 1 patient had headache and vomiting, and 1 patient had foreign body sensations)</td>
</tr>
<tr>
<td>Imdadur Rahman</td>
<td>2015</td>
<td>UK</td>
<td>Healthy volunteers</td>
<td>38 [26–45]</td>
<td>Hand held (Intromedic)</td>
<td>24 min</td>
<td>–</td>
<td>CG was given within 3 days after MCCE</td>
<td>Major lesions 26 (26/0)</td>
<td>–</td>
<td>EGJ 92%, cardia 88%, fundus 96%, body 100%, incisura 96%, antrum 96%, pylorus 100%</td>
<td>No</td>
</tr>
<tr>
<td>Wen-Bin Zou</td>
<td>2014</td>
<td>China</td>
<td>Patient with upper abdominal complaint</td>
<td>48 [24–70]</td>
<td>Robotic arm (ANKON)</td>
<td>29.1 min</td>
<td>5 min</td>
<td>CG was given within 4–24 hours after MCCE</td>
<td>Major lesions 68 (32/36)</td>
<td>91.20%</td>
<td>–</td>
<td>1 patient had chronic diarrhea</td>
</tr>
<tr>
<td>Ulrike W. Denzer</td>
<td>2013</td>
<td>France</td>
<td>Patient with upper abdominal complaint</td>
<td>53</td>
<td>Hand held (Given Imaging)</td>
<td>10.6 min</td>
<td>4 min</td>
<td>CG was given within 4–24 hours after MCCE</td>
<td>Minor lesions 168</td>
<td>90.50%</td>
<td>–</td>
<td>No</td>
</tr>
</tbody>
</table>

MCCE, magnetically controlled capsule endoscopy; CG, conventional gastroduodenoscopy; EGJ, esophagogastric junction.
upper abdominal discomforts. Studies about MCCE focus on detecting lesion in the proximal stomach were greatly needed.

There was considerable heterogeneity in the analysis of specificity between the studies. In two studies (17,19), the specificity of MCCE in detecting gastric diseases was only 78.0% and 71.0% respectively, while in two studies the specificity was 100% and 95.0% respectively (16,18).
Meta-regression was undertaken to discern the cause of the heterogeneity, however we didn’t found the culprit through statistical methods in the three expected variables that is the operation time of MCCE (P=0.3244), the kind of participants (P=0.9584), the type of MCCE (P=0.3463). We speculated some hidden variables that may arouse heterogeneity. The most possible reason for the presence of heterogeneity may be connected with the experience of the endoscopists and the number of endoscopists who read the capsule images. Despite the fact that the MCCE investigators were all blinded to CG findings, their experience was not set out in detail.

However, the limitations of our study have to be acknowledged. Firstly, because MCCE is a new device which has not been commonly used for detecting gastric diseases, the number of related studies is insufficient. Secondly, the numbers of involved patients of included studies were significantly dissimilar, ranging from 26 to 350 patients. Thirdly, only 26 (26/612, 4.24%) healthy volunteers were recruited, which is not consistent with clinical practice. Fourthly, as reported in a study (17), luminal visibility and location of the lesion can dissatisfy the diagnostic accuracy of minor lesions and have no impact on major lesions. MCCE perform better than CG at diagnosing minor lesions, however the proportion of minor lesion is not sufficient, which may confuse researchers. Further studies concerning about minor lesions were demanded. At last, lesions distributing in the upper 1/3 of the stomach were not sufficient, which often were missed by MCCE, this may lead to highly diagnostic accuracy compared to CG.

In summary, MCCE have acceptable sensitivity and specificity in diagnosing gastric diseases in adults, however there are still many problems waiting to be resolved (23). A lens-cleaning system which is available with CG, a stronger guidance system which seems to be too weak at present and requires faster speed of movement with stronger force, a biopsy system (24) and a better capability of keeping the capsule more steadily have to be settled urgently in the future. All in all, further large scale studies are required to test and verify its value in the upper gastrointestinal tract. We believe that one day in the near future capsule endoscopy will change the way we practice gastroenterology.

Figure 4 Diagnostic accuracy of MCCE for the diagnosis of gastric diseases.

Figure 5 Pooled positive (A) and negative (B) likelihood ratio of total studies included for meta-analysis.
Figure 6  Pooled diagnostic odds ratio of total studies included for meta-analysis.

Figure 7  Deek’s plot for publication bias.

Acknowledgments

Funding: This work was supported by the Science and Technology Planning Project of Guangdong Province, China (2017A020215014) and the Medical Scientific Research Foundation of Guangdong Province, China (A2017274).

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at http://dx.doi.org/10.21037/dmr-20-141

Conflict of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/dmr-20-141). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.
References


doi: 10.21037/dmr-20-141