diagnostics

MDPI

Systematic Review

Inter/Intra-Observer Agreement in Video-Capsule Endoscopy: Are We Getting It All Wrong? A Systematic Review and Meta-Analysis

Pablo Cortegoso Valdivia [1,*,†], Ulrik Deding [2,3,†], Thomas Bjørsum-Meyer [2,3], Gunnar Baatrup [2,3], Ignacio Fernández-Urién [4], Xavier Dray [5], Pedro Boal-Carvalho [6], Pierre Ellul [7], Ervin Toth [8], Emanuele Rondonotti [9], Lasse Kaalby [2,3], Marco Pennazio [10] and Anastasios Koulaouzidis [2,11,12,13] on behalf of the International CApsule endoscopy REsearch (I-CARE) Group

1 Gastroenterology and Endoscopy Unit, University Hospital of Parma, University of Parma, 43126 Parma, Italy
2 Department of Clinical Research, University of Southern Denmark, 5230 Odense, Denmark
3 Department of Surgery, Odense University Hospital, 5000 Odense, Denmark
4 Department of Gastroenterology, University Hospital of Navarra, 31008 Pamplona, Spain
5 Center for Digestive Endoscopy, Sorbonne University, Saint Antoine Hospital, APHP, 75012 Paris, France
6 Gastroenterology Department, Hospital da Senhora da Oliveira, Creixomil, 4835 Guimarães, Portugal
7 Division of Gastroenterology, Mater Dei Hospital, 2090 Msida, Malta
8 Department of Gastroenterology, Skåne University Hospital, Lund University, 20502 Malmö, Sweden
9 Gastroenterology Unit, Valduce Hospital, 22100 Como, Italy
10 University Division of Gastroenterology, City of Health and Science University Hospital, University of Turin, 10126 Turin, Italy
11 Department of Medicine, OUH Svendborg Sygehus, 5700 Svendborg, Denmark
12 Surgical Research Unit, OUH, 5000 Odense, Denmark
13 Department of Social Medicine and Public Health, Pomeranian Medical University, 70-204 Szczecin, Poland
* Correspondence: cortegosopablo@yahoo.it
† These authors contributed equally to this work.

check for updates

Citation: Cortegoso Valdivia, P.; Deding, U.; Bjørsum-Meyer, T.; Baatrup, G.; Fernández-Urién, I.; Dray, X.; Boal-Carvalho, P.; Ellul, P.; Toth, E.; Rondonotti, E.; et al. Inter/Intra-Observer Agreement in Video-Capsule Endoscopy: Are We Getting It All Wrong? A Systematic Review and Meta-Analysis. *Diagnostics* **2022**, *12*, 2400. https://doi.org/10.3390/diagnostics12102400

Academic Editor: Giorgio Treglia

Received: 16 September 2022
Accepted: 29 September 2022
Published: 2 October 2022

Abstract: Video-capsule endoscopy (VCE) reading is a time- and energy-consuming task. Agreement on findings between readers (either different or the same) is a crucial point for increasing performance and providing valid reports. The aim of this systematic review with meta-analysis is to provide an evaluation of inter/intra-observer agreement in VCE reading. A systematic literature search in PubMed, Embase and Web of Science was performed throughout September 2022. The degree of observer agreement, expressed with different test statistics, was extracted. As different statistics are not directly comparable, our analyses were stratified by type of test statistics, dividing them in groups of "None/Poor/Minimal", "Moderate/Weak/Fair", "Good/Excellent/Strong" and "Perfect/Almost perfect" to report the proportions of each. In total, 60 studies were included in the analysis, with a total of 579 comparisons. The quality of included studies, assessed with the MINORS score, was sufficient in 52/60 studies. The most common test statistics were the Kappa statistics for categorical outcomes (424 comparisons) and the intra-class correlation coefficient (ICC) for continuous outcomes (73 comparisons). In the overall comparison of inter-observer agreement, only 23% were evaluated as "good" or "perfect"; for intra-observer agreement, this was the case in 36%. Sources of heterogeneity (high, I^2 81.8–98.1%) were investigated with meta-regressions, showing a possible role of country, capsule type and year of publication in Kappa inter-observer agreement. VCE reading suffers from substantial heterogeneity and sub-optimal agreement in both inter- and intra-observer evaluation. Artificial-intelligence-based tools and the adoption of a unified terminology may progressively enhance levels of agreement in VCE reading.

Keywords: capsule endoscopy; video reading; agreement; small bowel; colon

1. Introduction

Video-capsule endoscopy (VCE) entered clinical use in 2001 [1]. Since then, several post-market technological advancements followed, making capsule endoscopes the prime diagnostic choice for several clinical indications, i.e., obscure gastrointestinal bleeding (OGIB), iron-deficiency anemia (IDA), Crohn's disease (diagnosis and monitoring) and tumor diagnosis. Recently, the European Society of Gastrointestinal Endoscopy (ESGE) endorsed colon capsule endoscopy (CCE) as an alternative diagnostic tool in patients with incomplete conventional colonoscopy or contraindication for it, when sufficient expertise for performing CCE is available [2]. Furthermore, the COVID-19 pandemic has bolstered CCE (and double-headed capsules) in clinical practice as the test can be completed in the patient's home with minimal contact with healthcare professionals and other patients [3,4].

The diagnostic yield of VCE depends on several factors, such as the reader's performance, experience [5] and accumulating fatigue (especially with long studies) [6]. Although credentialing guidelines for VCE exist, there are no formal recommendations and only limited data to guide capsule endoscopists on how to read the many images collected in each VCE [7,8]. Furthermore, there is no guidance on how to increase performance and obtain a consistent level of high-quality reporting [9]. With accumulating data on inter/intra-observer variability in VCE reading (i.e., degree of concordance between multiple readers/multiple reading sessions of the same reader), we embarked on a comprehensive systematic review of the contemporary literature and aimed to estimate the inter- and intra-observer agreement of VCE through a meta-analysis.

2. Materials and Methods

2.1. Data sources and Search Strategy

We conducted a systematic literature search in PubMed, Embase and Web of Science in order to identify all relevant studies in which inter- and/or intra-observer agreement in VCE reading was evaluated. The primary outcome was the evaluation of inter- and intra-observer agreement in VCE examinations. The last literature search was performed on 26 September 2022. The complete search strings are available in Table S1. This review was registered at the PROSPERO international register of systematic reviews (ID 307267).

2.2. Inclusion and Exclusion Criteria

The inclusion criteria were: (i) full text articles; (ii) articles reporting either inter- or intra-observer agreement values (or both) of VCE reading; (iii) articles in English/Italian/Danish/Spanish/French language. Exclusion criteria were: article types such as reviews, case reports, conference papers or abstracts.

2.3. Screening of References

After exclusion of duplicates, references were independently screened by six authors (P.C.V., U.D., T.B.-M., X.D., P.B.-C., P.E.). Each author screened one fourth of the references (title and abstract), according to the inclusion and exclusion criteria. In case of discrepancy, the reference was included for full text evaluation. This approach was then repeated on included references with an evaluation of the full text by three authors (P.C.V., U.D., T.B.-M.). In case of discrepancy in the full-text evaluation, the third author would also evaluate the reference and a consensus discussion between all three would determine the outcome.

2.4. Data Extraction

Data were extracted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [10]. We extracted data on patients' demographics, indication for the procedure, the setting for the intervention, the type of VCE and its completion rate, and the type of test statistics.

2.5. Study Assessment and Risk of Bias

Included studies underwent an assessment of methodological quality by three independent reviewers (P.C.V., U.D., T.B.-M.) through the Methodological Index for Non-Randomized Studies (MINORS) assessment tool [11].

Items 7, 9, 10, 11 and 12 were omitted, as they were not applicable to the included studies; therefore, since the global ideal score for non-comparative studies, in MINORS, is at least two thirds of the total score (n = 24), we applied the same proportion to the maximum score with omitted items (n = 14) obtaining the arbitrary cut-off value of 10.

2.6. Statistics

In the included studies, different test statistics were used when reporting the degree of observer agreement. The most common ones are the Kappa statistics for categorical outcomes and the intra-class correlation coefficient (ICC) for continuous outcomes. Kappa and ICC are not directly comparable and our analyses were therefore stratified by type of test statistics.

The Kappa statistics estimates the degree of agreement between two or more readers, while taking into account the chance agreement that would occur if the readers guessed at random. Cohen's Kappa was introduced in order to improve the previously common used percent agreement [12].

The ICC is a measure of the degree of correlation and agreement between measurements and is a modification of the Pearson correlation coefficient, which measures the magnitude of correlation between variables (or readers) but, in addition, ICC takes readers' bias into account [13,14].

Less commonly reported were the Spearman rank correlation [15], Kendall's coefficient and the Kolmogorov–Smirnov test. First, we evaluated each comparison using guidelines for the specific test statistics (Table 1) and divided them into groups of "None/Poor/Minimal", "Moderate/Weak/Fair", "Good/Excellent/Strong" and "Perfect/Almost perfect" to report the proportions of each, stratified by inter/intra-observer agreement evaluations.

Table 1. Evaluation guideline.

Kappa		**Intra-Class Correlation**		**Spearman Rank Correlation**	
Value	**Evaluation**	**Value**	**Evaluation**	**Value**	**Evaluation**
>0.90	Almost perfect	>0.9	Excellent	±1	Perfect
0.80–0.90	Strong	0.75–0.9	Good	±0.8–0.9	Very strong
0.60–0.79	Moderate	0.5–<0.75	Moderate	±0.6–0.7	Moderate
0.40–0.59	Weak			±0.3–0.5	Fair
0.21–0.39	Minimal	<0.5	Poor	±0.1–0.2	Poor
<0.20	None			0	None

As no guidelines were identified for the Kendall's coefficient and the Kolmogorov–Smirnov test, we adopted the guidelines used for Kappa as the scales were similar. The mean value was estimated stratified by test statistic. The significance level was set at 5%, and 95% confidence intervals (CIs) were calculated. All pooled estimates were calculated in random effects models stratified into four categories; inter-observer Kappa, intra-observer Kappa, inter-observer ICC and intra-observer ICC. To investigate publication bias and small study effects, Egger's tests were performed and illustrated by funnel plots. Individual study data were extracted and compiled in spreadsheets for pooled analyses. Data management was conducted in SAS (SAS Institute Inc. SAS 9.4. Cary, NC, USA), while analyses and plots were performed in R (R Development Core Team, Boston, MA, USA) using the metafor and tidyverse packages [16,17].

3. Results

Overall, 483 references were identified from the databases. After the removal of duplicates, 269 were screened, leading to 95 references for full-text reading. One additional reference was retrieved via snowballing. Sixty (n = 60) studies were eventually included, 37 of which had reported information on variance for their agreement measures, enabling them to be included for pooled estimates (Figure 1). MINORS scores ranged from 7 to 14, with the majority of references scoring 10 or above (n = 52) (Table 2).

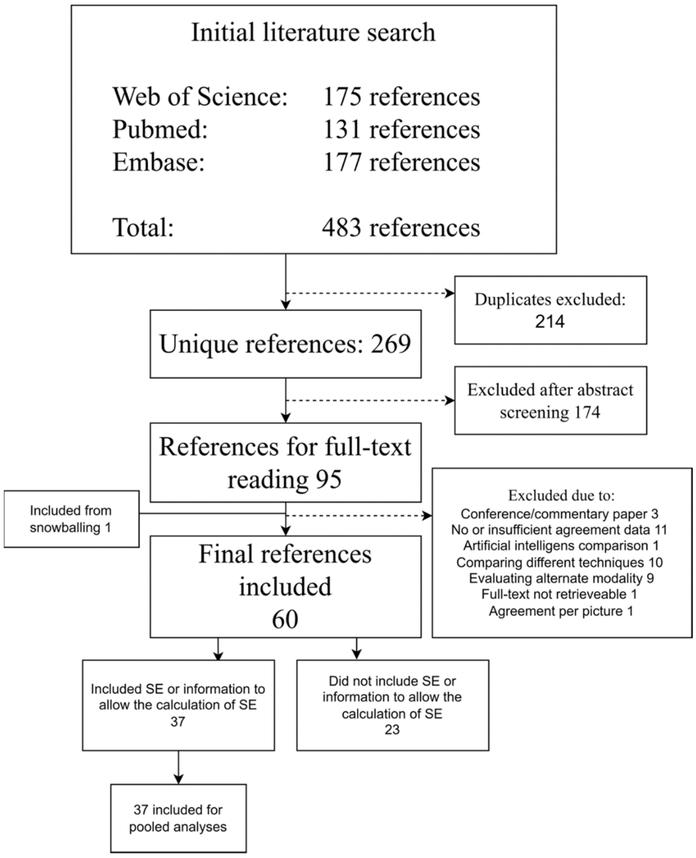

Figure 1. Flow diagram of the study. Abbreviations: SE, standard error.

Table 2. Characteristics of included studies, including methodological quality assessment.

Reference (Year)	Single or Multi Center Study	n Included for Review (Total)	Indication	Finding Group(s)	MINORS Score (0–14)
Adler DG (2004) [18]	Single	20 (20)	GI bleeding	Blood; Erosions/Ulcerations	11
Alageeli M (2020) [19]	Multi	25 (25)	GI bleeding, CD, screening for HPS	Cleanliness	11
Albert J (2004) [20]	Single	36 (36)	OGIB, suspected CD, suspected SB tumor, refractory sprue, FAP	Cleanliness	12
Arieira C (2019) [21]	Single	22 (22)	Known CD	IBD	8
Biagi F (2006) [22]	Multi	21 (32)	CeD, IBS, known CD	Villous atrophy	10
Blanco-Velasco G (2021) [23]	Single	100 (100)	IDA, GI bleeding, known CD, SB tumors, diarrhea	Blood; IBD; Blended outcomes	11

Table 2. *Cont.*

Reference (Year)	Single or Multi Center Study	n Included for Review (Total)	Indication	Finding Group(s)	MINORS Score (0–14)
Bossa F (2006) [24]	Single	39 (41)	OGIB, HPS, known CD, CeD, diarrhea	Blood; Blended outcomes; Other lesions; Polyps; Erosions/Ulcerations; Angiodysplasias	8
Bourreille A (2006) [25]	Multi	32 (32)	Illeocolonic resection	Blended outcomes; Other lesions; Villous atrophy; Erosions/Ulcerations	12
Brotz C (2009) [26]	Single	40 (541)	GI bleeding, abdominal pain, diarrhea, anemia, follow-up of prior findings	Cleanliness	10
Buijs MM (2018) [27]	Single	42 (136)	CRC screening	Blended outcomes; Polyps; Cleanliness	13
Chavalitdhamrong D (2012) [28]	Multi	65 (65)	Portal hypertension	Other lesions	12
Chetcuti Zammit S (2021) [29]	Multi	300 (300)	CeD, seronegative villous atrophy	IBD; Villous atrophy; Erosions/Ulcerations; Blended outcomes	13
Christodoulou D (2007) [30]	Single	20 (20)	GI bleeding	Other lesions; Angiodysplasias; Polyps; Blood	11
Cotter J (2015) [31]	Single	70 (70)	Known CD	IBD	12
De Leusse A (2005) [32]	Single	30 (64)	GI bleeding	Blood; Angiodysplasias; Other lesions; Erosions/Ulcerations; Blended outcomes;	12
de Sousa Magalhães R (2021) [33]	Single	58 (58)	Incomplete colonoscopy	Cleanliness	11
Delvaux M (2008) [34]	Multi	96 (98)	Known or suspected esophageal disease	Blended outcomes	13
D'Haens G (2015) [35]	Multi	20 (40)	Known CD	IBD	11
Dray X (2021) [36]	Multi	155 (637)	OGIB	Cleanliness	12
Duque G (2012) [37]	Single	20 (20)	GI bleeding	Blended outcomes	11
Eliakim R (2020) [38]	Single	54 (54)	Known CD	IBD	11
Esaki M (2009) [39]	Single	75 (102)	OGIB, FAP, GI lymphoma, PJS, GIST, carcinoid tumor	Cleanliness	12
Esaki M (2019) [40]	Multi	50 (108)	Suspected CD	Other lesions; Erosions/Ulcerations	10
Ewertsen C (2006) [41]	Single	33 (34)	OGIB, carcinoid tumors, angiodysplasias, diarrhea, immune deficiency, diverticular disease	Blended outcomes	8
Gal E (2008) [42]	Single	20 (20)	Known CD	IBD	7

Table 2. *Cont.*

Reference (Year)	Single or Multi Center Study	n Included for Review (Total)	Indication	Finding Group(s)	MINORS Score (0–14)
Galmiche JP (2008) [43]	Multi	77 (89)	GERD symptoms	Other lesions	12
Garcia-Compean D (2021) [44]	Single	22 (22)	SB angiodysplasias	Angiodysplasias; Blended outcomes	12
Ge ZZ (2006) [45]	Single	56 (56)	OGIB, suspected CD, abdominal pain, suspected SB tumor, FAP, diarrhea, sprue	Cleanliness	12
Girelli CM (2011) [46]	Single	25 (35)	Suspected submucosal lesion	Other lesions	12
Goyal J (2014) [47]	Single	34 (34)	NA	Cleanliness	11
Gupta A (2010) [48]	Single	20 (20)	PJS	Polyps	11
Gupta T (2011) [49]	Single	60 (60)	OGIB	Other lesions	12
Hong-Bin C (2013) [50]	Single	63 (63)	GI bleeding, abdominal pain, chronic diarrhea	Cleanliness	11
Jang BI (2010) [51]	Multi	56 (56)	NA	Blended outcomes	10
Jensen MD (2010) [52]	Single	30 (30)	Known or suspected CD	Other lesions; IBD; Blended outcomes	11
Lai LH (2006) [53]	Single	58 (58)	OGIB, known CD, abdominal pain	Blended outcomes	10
Lapalus MG (2009) [54]	Multi	107 (120)	Portal hypertension	Other lesions	11
Laurain A (2014) [55]	Multi	77 (80)	Portal hypertension	Other lesions	12
Laursen EL (2009) [56]	Single	30 (30)	NA	Blended outcomes	12
Leighton JA (2011) [57]	Multi	40 (40)	Healthy volunteers	Cleanliness	13
Murray JA (2008) [58]	Single	37 (40)	CeD	IBD; Villous atrophy	12
Niv Y (2005) [59]	Single	50 (50)	IDA, abdominal pain, known CD, CeD, GI lymphoma, SB transplant	Blended outcomes	11
Niv Y (2012) [60]	Multi	50 (54)	Known CD	IBD	13
Oliva S (2014) [61]	Single	29 (29)	UC	IBD	14
Oliva S (2014) [62]	Single	198 (204)	Suspected IBD, OGIB, other symptoms	Cleanliness	12
Omori T (2020) [63]	Single	20 (196)	Known CD	IBD	8
Park SC (2010) [64]	Single	20 (20)	GI bleeding, IDA, abdominal pain, diarrhea	Cleanliness; Blended outcomes	8
Petroniene R (2005) [65]	Single	20 (20)	CeD, villous atrophy	Villous atrophy	12

Table 2. *Cont.*

Reference (Year)	Single or Multi Center Study	n Included for Review (Total)	Indication	Finding Group(s)	MINORS Score (0–14)
Pezzoli A (2011) [66]	Multi	75 (75)	NA	Blood; Blended outcomes	12
Pons Beltrán V (2011) [67]	Multi	31 (273)	GI bleeding, suspected CD	Cleanliness	14
Qureshi WA (2008) [68]	Single	18 (20)	BE	Other lesions	11
Ravi S (2022) [69]	Single	10 (22)	GI bleeding	Other lesions	14
Rimbaş M (2016) [70]	Single	64 (64)	SB ulcerations	IBD	12
Rondonotti E (2014) [71]	Multi	32 (32)	NA	Other lesions	11
Sciberras M (2022) [72]	Multi	100 (182)	Suspected submucosal lesion	Other lesions	10
Shi HY (2017) [73]	Single	30 (150)	UC	IBD; Blood; Erosions/Ulcerations	14
Triantafyllou K (2007) [74]	Multi	87 (87)	Diabetes mellitus	Cleanliness; Blended outcomes	11
Usui S (2014) [75]	Single	20 (20)	UC	IBD	9
Wong RF (2006) [76]	Single	19 (32)	FAP	Polyps	13
Zakaria MS (2009) [77]	Single	57 (57)	OGIB	Blended outcomes	9

Abbreviations: BE, Barrett's esophagus; CD, Crohn's disease; CeD, celiac disease; CRC, colorectal cancer; FAP, familial adenomatous polyposis; GERD, gastroesophageal reflux disease; GI, gastrointestinal; GIST, gastrointestinal stromal tumor; HPS; hereditary polyposis syndrome; IBS, irritable bowel syndrome; IDA, iron-deficiency anemia; NA, not available; OGIB, obscure gastrointestinal bleeding; PJS, Peutz–Jeghers syndrome; SB, small bowel; UC, ulcerative colitis.

Regarding the type of statistics used in the 60 included studies, 46 reported Kappa statistics (424 comparisons), 11 reported ICC (73 comparisons), 5 reported Spearman rank correlations (60 comparisons), 2 reported Kendall's coefficients (20 comparisons) and 1 reported Kolmogorov–Smirnov tests (2 comparisons).

The analysis of combined inter/intra-observer values (overall means) per type of statistics revealed a weak agreement for the comparisons measured by Kappa statistics (0.53, CI 95% 0.51; 0.55), good for ICC (0.81, CI 95% 0.78; 0.84) and moderate for Spearman rank correlation (0.73, CI 95% 0.68; 0.78). For Kendall's coefficient and Kolmogorov–Smirnov tests, too few studies were identified to make an overall evaluation (Table 3).

Table 3. Overall means combined inter/intra-observer statistics values.

Test Statistic	Mean	CI 95%	Range	Comparisons, n (Inter/Intra)	Studies, n	Evaluation
Kappa	0.53	0.51; 0.55	−0.33; 1.0	424 (383/41)	46	Weak
ICC	0.81	0.78; 0.84	0.51; 1.0	73 (41/32)	11	Good
Spearman Rank	0.73	0.68; 0.78	0.30; 1.0	60 (60/0)	5	Moderate
Kendall's coefficient	0.89	0.86; 0.92	0.77; 1.0	20 (18/2)	2	n too small
Kolmogorov–Smirnov	0.99	-	0.98; 1.0	2 (2/0)	1	n too small

The distribution of evaluations, stratified by inter/intra-observer agreements, was analyzed by combining all specific comparisons regardless of the type of statistics models (Kappa alone was considered in 25 inter-observer comparisons, whenever more than one model was applied for the same outcome): in 479 inter-observer comparisons, a "good" or "perfect" agreement was obtained in only 23% of the cases; in 75 intra-observer comparisons, this was the case in 36% of the cases (Figure 2).

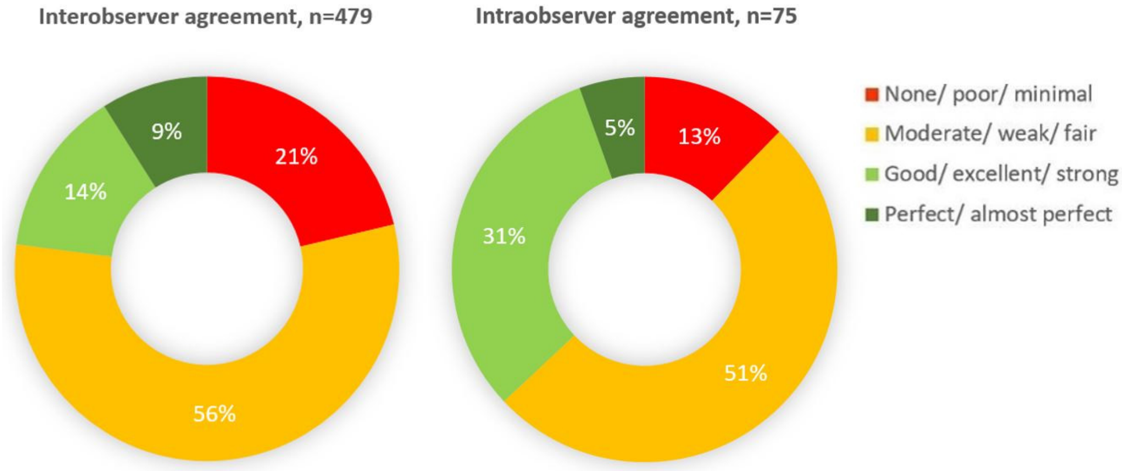

Figure 2. Distribution of agreement evaluations stratified by inter/intra-observer agreements.

For the pooled random effects models stratified by inter/intra-observer and test statistic, the overall estimates of agreement ranged from 0.46 to 0.84, although a substantial degree of heterogeneity was present in all four models (Figures 3 and 4). The I^2 statistic ranged from 81.8% to 98.1% (Figure 4). Meta-regressions investigating the possible sources of heterogeneity found no significance of any variable for ICC inter-observer agreement, but for Kappa inter-observer agreement, country, capsule type and year of publication may have contributed to the heterogeneity.

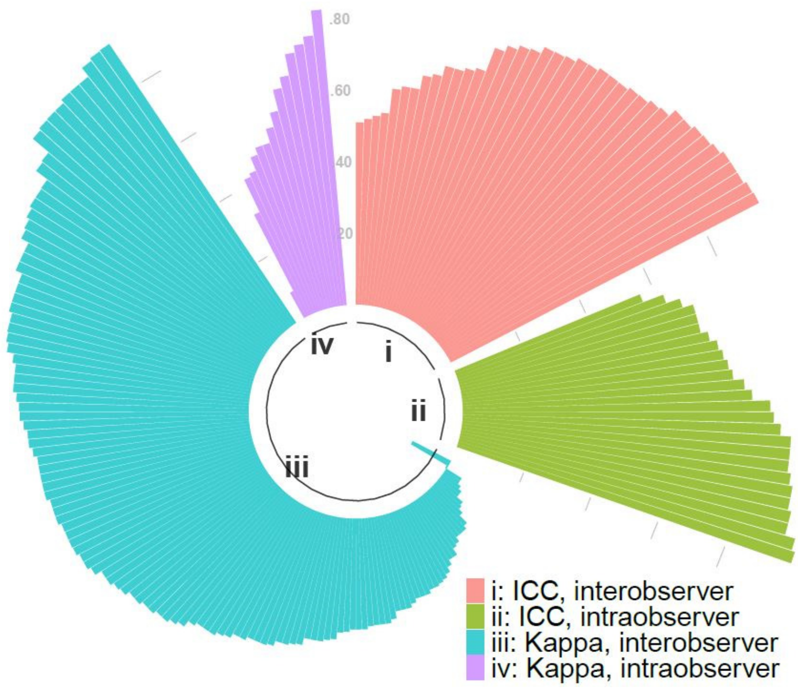

Figure 3. Circle bar chart visualizing the distribution of Kappa statistics and ICC values for every comparison.

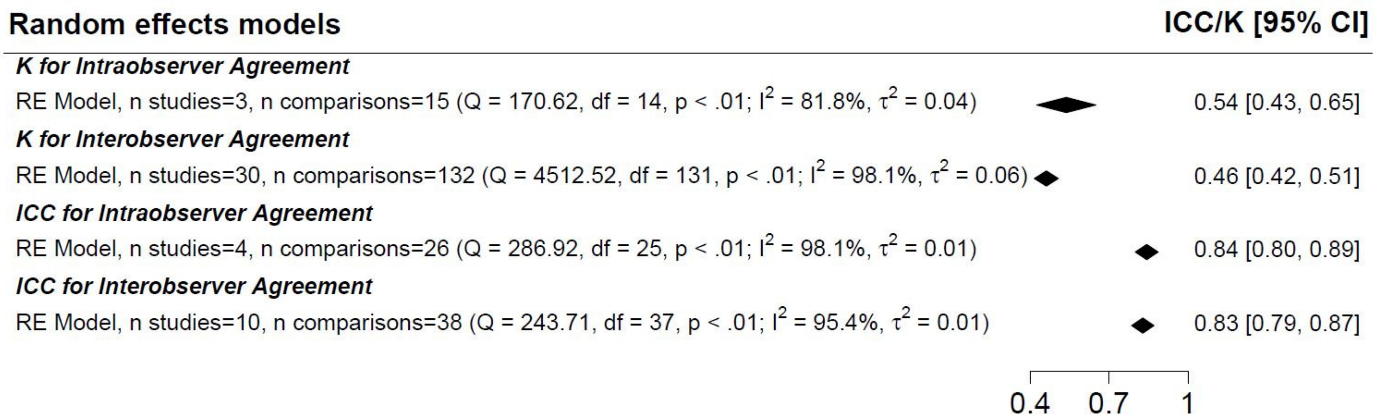

Figure 4. Pooled random effects model for inter/intra-observer agreement by studies reporting Kappa statistics or inter/intra-class correlation coefficient.

For the random effects models of the overall inter/intra-observer agreements, the Eggers tests resulted in p-values < 0.01 for inter/intra-observer ICC models, 0.78 for Kappa inter-observer and 0.20 for Kappa intra-observer (Figure 5).

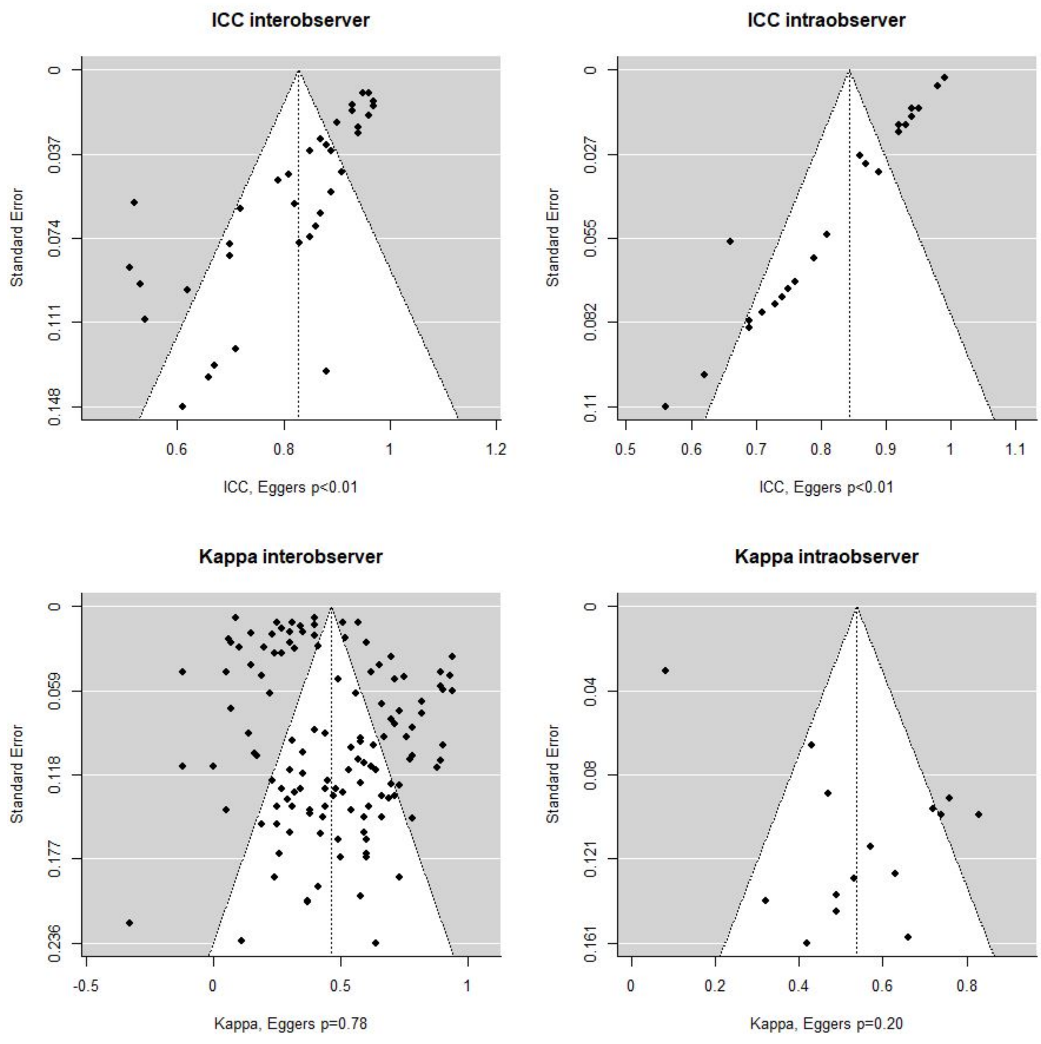

Figure 5. Eggers tests for inter/intra-observer agreements (ICC and Kappa models).

4. Discussion

Reading VCE videos is a laborious and time-consuming task. Previous work has showed that the inter-observer agreement and the detection rate of significant findings are low, regardless of the reader's experience [5,78]. Moreover, attempts to improve performance by a constructed upskilling training program did not significantly impact readers with different experience levels [78]. Fatigue has been blamed as a significant determinant of missed lesions: a recent study demonstrates that reader accuracy declines after reading just one VCE video, and that neither subjective nor objective measures of fatigue were sufficient to predict the onset of the effects of fatigue [6]. Recently, strides were made in establishing a guide for evaluating the relevance of small-bowel VCE findings [79]. Above all, artificial intelligence (AI)-supported VCE can identify abnormalities in VCE images with higher sensitivity and significantly shorter reading times than conventional analysis by gastroenterologists [80,81]. AI has, of course, no issues with inter-observer agreement and is poised to become an integral part of VCE reading in the years to come. AI develops on the background of human-based 'ground truth' (usually subjective expert opinion) [82]. So, how do we as human readers get it so wrong?

The results of our study show that the overall pooled estimate for "perfect" or "good" inter- and intra-observer agreement was only 23% and 37%, respectively (Figure 2). Although significant heterogeneity was noted in both Kappa statistic and ICC-based studies, the overall combined inter/intra-observer agreement for Kappa-evaluated outcomes was weak (0.46 and 0.54, respectively), while for ICC-evaluated outcomes the agreement was good (0.83 and 0.84, respectively).

A possible explanation to this apparent discrepancy is that ICC outcomes are more easily quantifiable, therefore providing a higher degree of unified understanding on how to evaluate, whereas categorical outcomes in Kappa statistics may be prone to a more subjective evaluation; for instance, substantial heterogeneity may be caused by pooling observations without unified definition of the outcome variables (e.g., cleansing scale, per segment or patient, categorical subgroups differences).

A viable solution to the poor inter-/intra-observer agreement on VCE reading could be represented by AI-based tools. AI offers the opportunity of a standardized observer-independent evaluation of pictures and videos relieving reviewers' workload, but are we ready to rely on non-human assessment of diagnostic examinations to decide for subsequent investigations or treatments? Several algorithms reported with high accuracy have been proposed for VCE analysis. The main deep learning algorithm for image analysis has become convolutional neural networks (CNN) as they have shown excellent performances for detecting esophageal, gastric and colonic lesions [83–85]. However, some important shortcomings need to be overcome before CNNs are ready for implementation in clinical practice. The generalization and performance of CNNs in real-life settings are determined by the quality of data used to train the algorithm. Hence, large amounts of high-quality training data are needed with external algorithm validation, which necessitates collaboration between international centers. A high sensitivity from AI should be prioritized even at the cost of the specificity as AI findings should always be reviewed by human professionals.

This study shows several limitations. As VCE is used for numerous indications and for all parts of the GI tract, an inherent weakness is the natural heterogeneity of the included studies, which is evident in the pooled analyses (I^2 statistics > 80% in all strata). The meta-regressions indicated that country, capsule type and year of publication may have contributed to the heterogeneity for Kappa inter-observer agreement, whereas no sources were identified in ICC analyses; furthermore, the Eggers' tests indicated publication bias in ICC analyses but not in Kappa analyses. Therefore, there is a risk that specific pooled estimates may be inaccurate, but the heterogeneity may also be the result of very different ways of interpreting videos or definitions of outcomes between sites and trials. No matter these substantial weaknesses to the results of the pooled analyses, the proportions of agreements and the great variance in agreements are clear. In more than 70% of the published comparisons, the agreement between readers is moderate or worse, as for intra-observer agreement.

Data regarding the reader's experience were originally extracted but omitted in the final analysis because of heterogeneity of the terminology and of the lack of a unified experience scale. This should not be considered as a problem, as most studies fail to confirm a significant lesion detection rate difference between experienced and expert readers, physician readers and nurses [86,87], while some of them point to possible equalization of any difference between novices and experienced even only after one VCE reading due to fatigue [6].

Moreover, we decided not to perform any subgroup analysis based on possible a priori clustering of findings (e.g., bleeding lesions, ulcers, polyps, etc.); the reason for this choice is related, once again, to the extreme variability of encountered definitions and the lack of a uniform terminology.

5. Conclusions

As of today, the results of our study show that VCE reading suffers from a sub-optimal inter/intra-observer agreement.

For future meta-analyses, more studies are needed enabling strata of subgroups specific to the outcome and indication, which may limit the heterogeneity. The heterogeneity may also be reduced by stratifying analyses based on the experience level of the readers or the number of them in comparisons, as this will most likely affect the agreement. The progressive implementation of AI-based tools will possibly enhance the agreement in VCE reading between observers, not only reducing the "human bias" but also relieving the significant burden in workload.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/diagnostics12102400/s1, Table S1: search strings with PICO questions.

Author Contributions: Planning of the study: U.D., T.B.-M. and G.B.; conducting the study: P.C.V., U.D. and T.B.-M.; data collection: P.C.V., U.D., T.B.-M., X.D., P.B.-C. and P.E.; statistical analysis: U.D. and L.K.; data interpretation: P.C.V., U.D., T.B.-M. and A.K.; critical revision: I.F.-U., X.D., P.E., E.T., E.R. and M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors are members of the International CApsule endoscopy REsearch (I-CARE) Group, an independent international research group promoting multicenter studies on capsule endoscopy.

Conflicts of Interest: Potential competing interest. I.F.U.: consultancy fees (Medtronic); X.D.: co-founder and shareholder (Augmented Endoscopy); lecture fees (Bouchara Recordati, Fujifilm, Medtronic, MSD and Pfizer); consultancy (Alfasigma, Boston Scientific, Norgine, Pentax); E.T.: consultancy and lecture fees (Medtronic and Olympus); research material (ANX Robotica); research and travel support (Norgine); E.R.: speaker honoraria (Fujifilm); consultancy agreement (Medtronic); M.P.: lecture fees (Medtronic and Olympus); A.K.: co-founder and shareholder of AJM Medicaps; co-director and shareholder of iCERV Ltd.; consultancy fees (Jinshan Ltd.); travel support (Jinshan, Aquilant and Falk Pharma); research support (grant) from ESGE/Given Imaging Ltd. and (material) IntroMedic/SynMed; honoraria (Falk Pharma UK, Ferring, Jinshan, Medtronic). Member of Advisory board meetings (Falk Pharma UK, Tillots, ANKON).

References

1. Meron, G.D. The development of the swallowable video capsule (M2A). *Gastrointest. Endosc.* **2000**, *52*, 817–819. [CrossRef]
2. Spada, C.; Hassan, C.; Bellini, D.; Burling, D.; Cappello, G.; Carretero, C.; Dekker, E.; Eliakim, R.; de Haan, M.; Kaminski, M.F.; et al. Imaging alternatives to colonoscopy: CT colonography and colon capsule. European Society of Gastrointestinal Endoscopy (ESGE) and European Society of Gastrointestinal and Abdominal Radiology (ESGAR) Guideline-Update 2020. *Endoscopy* **2020**, *52*, 1127–1141. [CrossRef] [PubMed]
3. MacLeod, C.; Wilson, P.; Watson, A.J.M. Colon capsule endoscopy: An innovative method for detecting colorectal pathology during the COVID-19 pandemic? *Colorectal. Dis.* **2020**, *22*, 621–624. [CrossRef]
4. White, E.; Koulaouzidis, A.; Patience, L.; Wenzek, H. How a managed service for colon capsule endoscopy works in an overstretched healthcare system. *Scand. J. Gastroenterol.* **2022**, *57*, 359–363. [CrossRef]
5. Zheng, Y.; Hawkins, L.; Wolff, J.; Goloubeva, O.; Goldberg, E. Detection of lesions during capsule endoscopy: Physician performance is disappointing. *Am. J. Gastroenterol.* **2012**, *107*, 554–560. [CrossRef] [PubMed]
6. Beg, S.; Card, T.; Sidhu, R.; Wronska, E.; Ragunath, K.; UK capsule endoscopy users' group. The impact of reader fatigue on the accuracy of capsule endoscopy interpretation. *Dig. Liver Dis.* **2021**, *53*, 1028–1033. [CrossRef] [PubMed]
7. Rondonotti, E.; Pennazio, M.; Toth, E.; Koulaouzidis, A. How to read small bowel capsule endoscopy: A practical guide for everyday use. *Endosc. Int. Open* **2020**, *8*, E1220–E1224. [CrossRef] [PubMed]
8. Koulaouzidis, A.; Dabos, K.; Philipper, M.; Toth, E.; Keuchel, M. How should we do colon capsule endoscopy reading: A practical guide. *Adv. Gastrointest. Endosc.* **2021**, *14*, 26317745211001984. [CrossRef] [PubMed]
9. Spada, C.; McNamara, D.; Despott, E.J.; Adler, S.; Cash, B.D.; Fernández-Urién, I.; Ivekovic, H.; Keuchel, M.; McAlindon, M.; Saurin, J.C.; et al. Performance measures for small-bowel endoscopy: A European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *United Eur. Gastroenterol. J.* **2019**, *7*, 614–641. [CrossRef] [PubMed]
10. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.C.; Ioannidis, J.P.A.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. *BMJ* **2009**, *339*, b2700. [CrossRef] [PubMed]
11. Slim, K.; Nini, E.; Forestier, D.; Kwiatowski, F.; Panis, Y.; Chipponi, J. Methodological index for non-randomized studies (MINORS): Development and validation of a new instrument. *ANZ J. Surg.* **2003**, *73*, 712–716. [CrossRef] [PubMed]
12. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Med.* **2012**, *22*, 276–282. [CrossRef]
13. Liu, J.; Tang, W.; Chen, G.; Lu, Y.; Feng, C.; Tu, X.M. Correlation and agreement: Overview and clarification of competing concepts and measures. *Shanghai Arch. Psychiatry* **2016**, *28*, 115–120. [PubMed]
14. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–613. [CrossRef] [PubMed]
15. Chan, Y.H. Biostatistics 104: Correlational analysis. *Singap. Med. J.* **2003**, *44*, 614–619.
16. Viechtbauer, W. Conducting Meta-Analyses in R with the metafor Package. *J. Stat. Softw.* **2010**, *36*, 1–48. [CrossRef]
17. Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; D'Agostino McGowan, L.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J.; et al. Welcome to the tidyverse. *J. Open Source Softw.* **2019**, *4*, 1686. [CrossRef]
18. Adler, D.G.; Knipschield, M.; Gostout, C. A prospective comparison of capsule endoscopy and push enteroscopy in patients with GI bleeding of obscure origin. *Gastrointest. Endosc.* **2004**, *59*, 492–498. [CrossRef]
19. Alageeli, M.; Yan, B.; Alshankiti, S.; Al-Zahrani, M.; Bahreini, Z.; Dang, T.T.; Friendland, J.; Gilani, S.; Homenauth, R.; Houle, J.; et al. KODA score: An updated and validated bowel preparation scale for patients undergoing small bowel capsule endoscopy. *Endosc. Int. Open* **2020**, *8*, E1011–E1017. [CrossRef] [PubMed]
20. Albert, J.; Göbel, C.M.; Lesske, J.; Lotterer, E.; Nietsch, H.; Fleig, W.E. Simethicone for small bowel preparation for capsule endoscopy: A systematic, single-blinded, controlled study. *Gastrointest. Endosc.* **2004**, *59*, 487–491. [CrossRef]
21. Arieira, C.; Magalhães, R.; Dias de Castro, F.; Carvalho, P.B.; Rosa, B.; Moreira, M.J.; Cotter, J. CECDAIic-a new useful tool in pan-intestinal evaluation of Crohn's disease patients in the era of mucosal healing. *Scand. J. Gastroenterol.* **2019**, *54*, 1326–1330. [CrossRef] [PubMed]
22. Biagi, F.; Rondonotti, E.; Campanella, J.; Villa, F.; Bianchi, P.I.; Klersy, C.; De Franchis, R.; Corazza, G.R. Video capsule endoscopy and histology for small-bowel mucosa evaluation: A comparison performed by blinded observers. *Clin. Gastroenterol. Hepatol.* **2006**, *4*, 998–1003. [CrossRef] [PubMed]
23. Blanco-Velasco, G.; Pinho, R.; Solórzano-Pineda, O.M.; Martínez-Camacho, C.; García-Contreras, L.F.; Murcio-Pérez, E.; Hernández-Mondragón, O.V. Assessment of the Role of a Second Evaluation of Capsule Endoscopy Recordings to Improve Diagnostic Yield and Patient Management. *GE Port. J. Gastroenterol.* **2022**, *29*, 106–110. [CrossRef] [PubMed]
24. Bossa, F.; Cocomazzi, G.; Valvano, M.R.; Andriulli, A.; Annese, V. Detection of abnormal lesions recorded by capsule endoscopy. A prospective study comparing endoscopist's and nurse's accuracy. *Dig. Liver Dis.* **2006**, *38*, 599–602. [CrossRef] [PubMed]
25. Bourreille, A. Wireless capsule endoscopy versus ileocolonoscopy for the diagnosis of postoperative recurrence of Crohn's disease: A prospective study. *Gut* **2006**, *55*, 978–983. [CrossRef]
26. Brotz, C.; Nandi, N.; Conn, M.; Daskalakis, C.; DiMarino, M.; Infantolino, A.; Katz, L.C.; Schroeder, T.; Kastenberg, D. A validation study of 3 grading systems to evaluate small-bowel cleansing for wireless capsule endoscopy: A quantitative index, a qualitative evaluation, and an overall adequacy assessment. *Gastrointest. Endosc.* **2009**, *69*, 262–270. [CrossRef]

27. Buijs, M.M.; Kroijer, R.; Kobaek-Larsen, M.; Spada, C.; Fernandez-Urien, I.; Steele, R.J.; Baatrup, G. Intra and inter-observer agreement on polyp detection in colon capsule endoscopy evaluations. *United Eur. Gastroenterol. J.* **2018**, *6*, 1563–1568. [CrossRef]
28. Chavalitdhamrong, D.; Jensen, D.M.; Singh, B.; Kovacs, T.O.; Han, S.H.; Durazo, F.; Saab, S.; Gornbein, J.A. Capsule Endoscopy Is Not as Accurate as Esophagogastroduodenoscopy in Screening Cirrhotic Patients for Varices. *Clin. Gastroenterol. Hepatol.* **2012**, *10*, 254–258.e1. [CrossRef]
29. Chetcuti Zammit, S.; McAlindon, M.E.; Sanders, D.S.; Sidhu, R. Assessment of disease severity on capsule endoscopy in patients with small bowel villous atrophy. *J. Gastroenterol. Hepatol.* **2021**, *36*, 1015–1021. [CrossRef]
30. Christodoulou, D.; Haber, G.; Beejay, U.; Tang, S.J.; Zanati, S.; Petroniene, R.; Cirocco, M.; Kortan, P.; Kandel, G.; Tatsioni, A.; et al. Reproducibility of Wireless Capsule Endoscopy in the Investigation of Chronic Obscure Gastrointestinal Bleeding. *Can. J. Gastroenterol.* **2007**, *21*, 707–714. [CrossRef]
31. Cotter, J.; Dias de Castro, F.; Magalhães, J.; Moreira, M.J.; Rosa, B. Validation of the Lewis score for the evaluation of small-bowel Crohn's disease activity. *Endoscopy* **2014**, *47*, 330–335. [CrossRef] [PubMed]
32. De Leusse, A.; Landi, B.; Edery, J.; Burtin, P.; Lecomte, T.; Seksik, P.; Bloch, F.; Jian, R.; Cellier, C. Video Capsule Endoscopy for Investigation of Obscure Gastrointestinal Bleeding: Feasibility, Results, and Interobserver Agreement. *Endoscopy* **2005**, *37*, 617–621. [CrossRef]
33. de Sousa Magalhães, R.; Arieira, C.; Boal Carvalho, P.; Rosa, B.; Moreira, M.J.; Cotter, J. Colon Capsule CLEansing Assessment and Report (CC-CLEAR): A new approach for evaluation of the quality of bowel preparation in capsule colonoscopy. *Gastrointest. Endosc.* **2021**, *93*, 212–223. [CrossRef] [PubMed]
34. Delvaux, M.; Papanikolaou, I.; Fassler, I.; Pohl, H.; Voderholzer, W.; Rösch, T.; Gay, G. Esophageal capsule endoscopy in patients with suspected esophageal disease: Double blinded comparison with esophagogastroduodenoscopy and assessment of interobserver variability. *Endoscopy* **2007**, *40*, 16–22. [CrossRef]
35. D'Haens, G.; Löwenberg, M.; Samaan, M.A.; Franchimont, D.; Ponsioen, D.; van den Brink, G.R.; Fockens, P.; Bossuyt, P.; Amininejad, L.; Rajamannar, G.; et al. Safety and Feasibility of Using the Second-Generation Pillcam Colon Capsule to Assess Active Colonic Crohn's Disease. *Clin. Gastroenterol. Hepatol.* **2015**, *13*, 1480–1486.e3. [CrossRef]
36. Dray, X.; Houist, G.; Le Mouel, J.P.; Saurin, J.C.; Vanbiervliet, G.; Leandri, C.; Rahmi, G.; Duburque, C.; Kirchgesner, J.; Leenhardt, R.; et al. Prospective evaluation of third-generation small bowel capsule endoscopy videos by independent readers demonstrates poor reproducibility of cleanliness classifications. *Clin. Res. Hepatol. Gastroenterol.* **2021**, *45*, 101612. [CrossRef] [PubMed]
37. Duque, G.; Almeida, N.; Figueiredo, P.; Monsanto, P.; Lopes, S.; Freire, P.; Ferreira, M.; Carvalho, R.; Gouveia, H.; Sofia, C. Virtual chromoendoscopy can be a useful software tool in capsule endoscopy. *Rev. Esp. Enferm. Dig.* **2012**, *104*, 231–236. [CrossRef]
38. Eliakim, R.; Yablecovitch, D.; Lahat, A.; Ungar, B.; Shachar, E.; Carter, D.; Selinger, L.; Neuman, S.; Ben-Horin, S.; Kopylov, U. A novel PillCam Crohn's capsule score (Eliakim score) for quantification of mucosal inflammation in Crohn's disease. *United Eur.Gastroenterol. J.* **2020**, *8*, 544–551. [CrossRef]
39. Esaki, M.; Matsumoto, T.; Kudo, T.; Yanaru-Fujisawa, R.; Nakamura, S.; Iida, M. Bowel preparations for capsule endoscopy: A comparison between simethicone and magnesium citrate. *Gastrointest. Endosc.* **2009**, *69*, 94–101. [CrossRef]
40. Esaki, M.; Matsumoto, T.; Ohmiya, N.; Washio, E.; Morishita, T.; Sakamoto, K.; Abe, H.; Yamamoto, S.; Kinjo, T.; Togashi, K.; et al. Capsule endoscopy findings for the diagnosis of Crohn's disease: A nationwide case—Control study. *J. Gastroenterol.* **2019**, *54*, 249–260. [CrossRef]
41. Ewertsen, C.; Svendsen, C.B.S.; Svendsen, L.B.; Hansen, C.P.; Gustafsen, J.H.R.; Jendresen, M.B. Is screening of wireless capsule endoscopies by non-physicians feasible? *Ugeskr. Laeger.* **2006**, *168*, 3530–3533. [PubMed]
42. Gal, E.; Geller, A.; Fraser, G.; Levi, Z.; Niv, Y. Assessment and Validation of the New Capsule Endoscopy Crohn's Disease Activity Index (CECDAI). *Dig. Dis. Sci.* **2008**, *53*, 1933–1937. [CrossRef]
43. Galmiche, J.P.; Sacher-Huvelin, S.; Coron, E.; Cholet, F.; Ben Soussan, E.; Sébille, V.; Filoche, B.; d'Abrigeon, G.; Antonietti, M.; Robaszkiewicz, M.; et al. Screening for Esophagitis and Barrett's Esophagus with Wireless Esophageal Capsule Endoscopy: A Multicenter Prospective Trial in Patients with Reflux Symptoms. *Am. J. Gastroenterol.* **2008**, *103*, 538–545. [CrossRef] [PubMed]
44. García-Compeán, D.; Del Cueto-Aguilera, Á.N.; González-González, J.A.; Jáquez-Quintana, J.O.; Borjas-Almaguer, O.D.; Jiménez-Rodríguez, A.R.; Muñoz-Ayala, J.M.; Maldonado-Garza, H.J. Evaluation and Validation of a New Score to Measure the Severity of Small Bowel Angiodysplasia on Video Capsule Endoscopy. *Dig. Dis.* **2022**, *40*, 62–67. [CrossRef] [PubMed]
45. Ge, Z.Z.; Chen, H.Y.; Gao, Y.J.; Hu, Y.B.; Xiao, S.D. The role of simeticone in small-bowel preparation for capsule endoscopy. *Endoscopy* **2006**, *38*, 836–840. [CrossRef]
46. Girelli, C.M.; Porta, P.; Colombo, E.; Lesinigo, E.; Bernasconi, G. Development of a novel index to discriminate bulge from mass on small-bowel capsule endoscopy. *Gastrointest. Endosc.* **2011**, *74*, 1067–1074. [CrossRef] [PubMed]
47. Goyal, J.; Goel, A.; McGwin, G.; Weber, F. Analysis of a grading system to assess the quality of small-bowel preparation for capsule endoscopy: In search of the Holy Grail. *Endosc. Int. Open* **2014**, *2*, E183–E186. [PubMed]
48. Gupta, A.; Postgate, A.J.; Burling, D.; Ilangovan, R.; Marshall, M.; Phillips, R.K.; Clark, S.K.; Fraser, C.H. A Prospective Study of MR Enterography Versus Capsule Endoscopy for the Surveillance of Adult Patients with Peutz-Jeghers Syndrome. *AJR Am. J. Roentgenol.* **2010**, *195*, 108–116. [CrossRef] [PubMed]
49. Gupta, T. Evaluation of Fujinon intelligent chromo endoscopy-assisted capsule endoscopy in patients with obscure gastroenterology bleeding. *World J. Gastroenterol.* **2011**, *17*, 4590. [CrossRef]

50. Chen, H.-B.; Huang, Y.; Chen, S.-Y.; Huang, C.; Gao, L.-H.; Deng, D.-Y.; Li, X.-J.; He, S.; Li, X.-L. Evaluation of visualized area percentage assessment of cleansing score and computed assessment of cleansing score for capsule endoscopy. *Saudi J. Gastroenterol.* **2013**, *19*, 160–164.
51. Jang, B.I.; Lee, S.H.; Moon, J.S.; Cheung, D.Y.; Lee, I.S.; Kim, J.O.; Cheon, J.H.; Park, C.H.; Byeon, J.S.; Park, Y.S.; et al. Inter-observer agreement on the interpretation of capsule endoscopy findings based on capsule endoscopy structured terminology: A multicenter study by the Korean Gut Image Study Group. *Scand. J. Gastroenterol.* **2010**, *45*, 370–374. [CrossRef]
52. Jensen, M.D.; Nathan, T.; Kjeldsen, J. Inter-observer agreement for detection of small bowel Crohn's disease with capsule endoscopy. *Scand. J. Gastroenterol.* **2010**, *45*, 878–884. [CrossRef] [PubMed]
53. Lai, L.H.; Wong, G.L.H.; Chow, D.K.L.; Lau, J.Y.; Sung, J.J.; Leung, W.K. Inter-observer variations on interpretation of capsule endoscopies. *Eur. J. Gastroenterol. Hepatol.* **2006**, *18*, 283–286. [CrossRef] [PubMed]
54. Lapalus, M.G.; Ben Soussan, E.; Gaudric, M.; Saurin, J.C.; D'Halluin, P.N.; Favre, O.; Filoche, B.; Cholet, F.; de Leusse, A.; Antonietti, M.; et al. Esophageal Capsule Endoscopy vs. EGD for the Evaluation of Portal Hypertension: A French Prospective Multicenter Comparative Study. *Am. J. Gastroenterol.* **2009**, *104*, 1112–1118. [CrossRef] [PubMed]
55. Laurain, A.; de Leusse, A.; Gincul, R.; Vanbiervliet, G.; Bramli, S.; Heyries, L.; Martane, G.; Amrani, N.; Serraj, I.; Saurin, J.C.; et al. Oesophageal capsule endoscopy versus oesophago-gastroduodenoscopy for the diagnosis of recurrent varices: A prospective multicentre study. *Dig. Liver Dis.* **2014**, *46*, 535–540. [CrossRef] [PubMed]
56. Laursen, E.L.; Ersbøll, A.K.; Rasmussen, A.M.O.; Christensen, E.H.; Holm, J.; Hansen, M.B. Intra- and interobserver variation in capsule endoscopy reviews. *Ugeskr. Laeger.* **2009**, *171*, 1929–1934. [PubMed]
57. Leighton, J.; Rex, D. A grading scale to evaluate colon cleansing for the PillCam COLON capsule: A reliability study. *Endoscopy* **2011**, *43*, 123–127. [CrossRef]
58. Murray, J.A.; Rubio–Tapia, A.; Van Dyke, C.T.; Brogan, D.L.; Knipschield, M.A.; Lahr, B.; Rumalla, A.; Zinsmeister, A.R.; Gostout, C.J. Mucosal Atrophy in Celiac Disease: Extent of Involvement, Correlation with Clinical Presentation, and Response to Treatment. *Clin. Gastroenterol. Hepatol.* **2008**, *6*, 186–193. [CrossRef] [PubMed]
59. Niv, Y.; Niv, G. Capsule Endoscopy Examination—Preliminary Review by a Nurse. *Dig. Dis. Sci.* **2005**, *50*, 2121–2124. [CrossRef] [PubMed]
60. Niv, Y.; Ilani, S.; Levi, Z.; Hershkowitz, M.; Niv, E.; Fireman, Z.; O'Donnel, S.; O'Morain, C.; Eliakim, R.; Scapa, E.; et al. Validation of the Capsule Endoscopy Crohn's Disease Activity Index (CECDAI or Niv score): A multicenter prospective study. *Endoscopy* **2012**, *44*, 21–26. [CrossRef] [PubMed]
61. Oliva, S.; Di Nardo, G.; Hassan, C.; Spada, C.; Aloi, M.; Ferrari, F.; Redler, A.; Costamagna, G.; Cucchiara, S. Second-generation colon capsule endoscopy vs. colonoscopy in pediatric ulcerative colitis: A pilot study. *Endoscopy* **2014**, *46*, 485–492. [CrossRef] [PubMed]
62. Oliva, S.; Cucchiara, S.; Spada, C.; Hassan, C.; Ferrari, F.; Civitelli, F.; Pagliaro, G.; Di Nardo, G. Small bowel cleansing for capsule endoscopy in paediatric patients: A prospective randomized single-blind study. *Dig. Liver Dis.* **2014**, *46*, 51–55. [CrossRef]
63. Omori, T.; Matsumoto, T.; Hara, T.; Kambayashi, H.; Murasugi, S.; Ito, A.; Yonezawa, M.; Nakamura, S.; Tokushige, K. A Novel Capsule Endoscopic Score for Crohn's Disease. *Crohns Colitis* **2020**, *2*, otaa040. [CrossRef]
64. Park, S.C.; Keum, B.; Hyun, J.J.; Seo, Y.S.; Kim, Y.S.; Jeen, Y.T.; Chun, H.J.; Um, S.H.; Kim, C.D.; Ryu, H.S. A novel cleansing score system for capsule endoscopy. *World J. Gastroenterol.* **2010**, *16*, 875–880. [PubMed]
65. Petroniene, R.; Dubcenco, E.; Baker, J.P.; Ottaway, C.A.; Tang, S.J.; Zanati, S.A.; Streutker, C.J.; Gardiner, G.W.; Warren, R.E.; Jeejeebhoy, K.N. Given Capsule Endoscopy in Celiac Disease: Evaluation of Diagnostic Accuracy and Interobserver Agreement. *Am. J. Gastroenterol.* **2005**, *100*, 685–694. [CrossRef] [PubMed]
66. Pezzoli, A.; Cannizzaro, R.; Pennazio, M.; Rondonotti, E.; Zancanella, L.; Fusetti, N.; Simoni, M.; Cantoni, F.; Melina, R.; Alberani, A.; et al. Interobserver agreement in describing video capsule endoscopy findings: A multicentre prospective study. *Dig. Liver Dis.* **2011**, *43*, 126–131. [CrossRef] [PubMed]
67. Pons Beltrán, V.; González Suárez, B.; González Asanza, C.; Pérez-Cuadrado, E.; Fernández Diez, S.; Fernández-Urién, I.; Mata Bilbao, A.; Espinós Pérez, J.C.; Pérez Grueso, M.J.; Argüello Viudez, L.; et al. Evaluation of Different Bowel Preparations for Small Bowel Capsule Endoscopy: A Prospective, Randomized, Controlled Study. *Dig. Dis. Sci.* **2011**, *56*, 2900–2905. [CrossRef] [PubMed]
68. Qureshi, W.A.; Wu, J.; DeMarco, D.; Abudayyeh, S.; Graham, D.Y. Capsule Endoscopy for Screening for Short-Segment Barrett's Esophagus. *Am. J. Gastroenterol.* **2008**, *103*, 533–537. [CrossRef] [PubMed]
69. Ravi, S.; Aryan, M.; Ergen, W.F.; Leal, L.; Oster, R.A.; Lin, C.P.; Weber, F.H.; Peter, S. Bedside live-view capsule endoscopy in evaluation of overt obscure gastrointestinal bleeding-a pilot point of care study. *Dig. Dis.* **2022**. [CrossRef] [PubMed]
70. Rimbaş, M.; Zahiu, D.; Voiosu, A.; Voiosu, T.A.; Zlate, A.A.; Dinu, R.; Galasso, D.; Minelli Grazioli, L.; Campanale, M.; Barbaro, F.; et al. Usefulness of virtual chromoendoscopy in the evaluation of subtle small bowel ulcerative lesions by endoscopists with no experience in videocapsule. *Endosc. Int. Open* **2016**, *4*, E508–E514. [CrossRef] [PubMed]
71. Rondonotti, E.; Koulaouzidis, A.; Karargyris, A.; Giannakou, A.; Fini, L.; Soncini, M.; Pennazio, M.; Douglas, S.; Shams, A.; Lachlan, N.; et al. Utility of 3-dimensional image reconstruction in the diagnosis of small-bowel masses in capsule endoscopy (with video). *Gastrointest. Endosc.* **2014**, *80*, 642–651. [CrossRef] [PubMed]

72. Sciberras, M.; Conti, K.; Elli, L.; Scaramella, L.; Riccioni, M.E.; Marmo, C.; Cadoni, S.; McAlindon, M.; Sidhu, R.; O'Hara, F.; et al. Score reproducibility and reliability in differentiating small bowel subepithelial masses from innocent bulges. *Dig. Liver Dis.* **2022**, *54*, 1403–1409. [CrossRef] [PubMed]
73. Shi, H.Y.; Chan, F.K.L.; Higashimori, A.; Kyaw, M.; Ching, J.Y.L.; Chan, H.C.H.; Chan, J.C.H.; Chan, A.W.H.; Lam, K.L.Y.; Tang, R.S.Y.; et al. A prospective study on second-generation colon capsule endoscopy to detect mucosal lesions and disease activity in ulcerative colitis (with video). *Gastrointest. Endosc.* **2017**, *86*, 1139–1146.e6. [CrossRef]
74. Triantafyllou, K.; Kalantzis, C.; Papadopoulos, A.A.; Apostolopoulos, P.; Rokkas, T.; Kalantzis, N.; Ladas, S.D. Video-capsule endoscopy gastric and small bowel transit time and completeness of the examination in patients with diabetes mellitus. *Dig. Liver Dis.* **2007**, *39*, 575–580. [CrossRef]
75. Usui, S.; Hosoe, N.; Matsuoka, K.; Kobayashi, T.; Nakano, M.; Naganuma, M.; Ishibashi, Y.; Kimura, K.; Yoneno, K.; Kashiwagi, K.; et al. Modified bowel preparation regimen for use in second-generation colon capsule endoscopy in patients with ulcerative colitis: Preparation for colon capsule endoscopy. *Dig. Endosc.* **2014**, *26*, 665–672. [CrossRef] [PubMed]
76. Wong, R.F.; Tuteja, A.K.; Haslem, D.S.; Pappas, L.; Szabo, A.; Ogara, M.M.; DiSario, J.A. Video capsule endoscopy compared with standard endoscopy for the evaluation of small-bowel polyps in persons with familial adenomatous polyposis (with video). *Gastrointest. Endosc.* **2006**, *64*, 530–537. [CrossRef]
77. Zakaria, M.S.; El-Serafy, M.A.; Hamza, I.M.; Zachariah, K.S.; El-Baz, T.M.; Bures, J.; Tacheci, I.; Rejchrt, S. The role of capsule endoscopy in obscure gastrointestinal bleeding. *Arab. J. Gastroenterol.* **2009**, *10*, 57–62. [CrossRef]
78. Rondonotti, E.; Soncini, M.; Girelli, C.M.; Russo, A.; Ballardini, G.; Bianchi, G.; Cantù, P.; Centenara, L.; Cesari, P.; Cortelezzi, C.C.; et al. Can we improve the detection rate and interobserver agreement in capsule endoscopy? *Dig. Liver Dis.* **2012**, *44*, 1006–1011. [CrossRef]
79. Leenhardt, R.; Koulaouzidis, A.; McNamara, D.; Keuchel, M.; Sidhu, R.; McAlindon, M.E.; Saurin, J.C.; Eliakim, R.; Fernandez-Urien Sainz, I.; Plevris, J.N.; et al. A guide for assessing the clinical relevance of findings in small bowel capsule endoscopy: Analysis of 8064 answers of international experts to an illustrated script questionnaire. *Clin. Res. Hepatol. Gastroenterol.* **2021**, *45*, 101637. [CrossRef]
80. Ding, Z.; Shi, H.; Zhang, H.; Meng, L.; Fan, M.; Han, C.; Zhang, K.; Ming, F.; Xie, X.; Liu, H.; et al. Gastroenterologist-Level Identification of Small-Bowel Diseases and Normal Variants by Capsule Endoscopy Using a Deep-Learning Model. *Gastroenterology* **2019**, *157*, 1044–1054. [CrossRef]
81. Xie, X.; Xiao, Y.F.; Zhao, X.Y.; Li, J.J.; Yang, Q.Q.; Peng, X.; Nie, X.B.; Zhou, J.Y.; Zhao, Y.B.; Yang, H.; et al. Development and validation of an artificial intelligence model for small bowel capsule endoscopy video review. *JAMA Netw. Open* **2022**, *5*, e2221992. [CrossRef] [PubMed]
82. Dray, X.; Toth, E.; de Lange, T.; Koulaouzidis, A. Artificial intelligence, capsule endoscopy, databases, and the Sword of Damocles. *Endosc. Int. Open* **2021**, *9*, E1754–E1755. [CrossRef] [PubMed]
83. Horie, Y.; Yoshio, T.; Aoyama, K.; Yoshimizu, S.; Horiuchi, Y.; Ishiyama, A.; Hirasawa, T.; Tsuchida, T.; Ozawa, T.; Ishihara, S.; et al. Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointest. Endosc.* **2019**, *89*, 25–32. [CrossRef] [PubMed]
84. Cho, B.J.; Bang, C.S.; Park, S.W.; Yang, Y.J.; Seo, S.I.; Lim, H.; Shin, W.G.; Hong, J.T.; Yoo, Y.T.; Hong, S.H.; et al. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* **2019**, *51*, 1121–1129. [CrossRef]
85. Wang, P.; Berzin, T.M.; Glissen Brown, J.R.; Bharadwaj, S.; Becq, A.; Xiao, X.; Liu, P.; Li, L.; Song, Y.; Zhang, D.; et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: A prospective randomised controlled study. *Gut* **2019**, *68*, 1813–1819. [CrossRef]
86. Yung, D.; Fernandez-Urien, I.; Douglas, S.; Plevris, J.N.; Sidhu, R.; McAlindon, M.E.; Panter, S.; Koulaouzidis, A. Systematic review and meta-analysis of the performance of nurses in small bowel capsule endoscopy reading. *United Eur. Gastroenterol. J.* **2017**, *5*, 1061–1072. [CrossRef]
87. Handa, Y.; Nakaji, K.; Hyogo, K.; Kawakami, M.; Yamamoto, T.; Fujiwara, A.; Kanda, R.; Osawa, M.; Handa, O.; Matsumoto, H.; et al. Evaluation of performance in colon capsule endoscopy reading by endoscopy nurses. *Can. J. Gastroenterol. Hepatol.* **2021**, *2021*, 8826100. [CrossRef]