Intra and inter-observer agreement on polyp detection in colon capsule endoscopy evaluations

Buijs, Maria Magdalena; Kroijer, Rasmus; Kobaek-Larsen, Morten; Spada, Cristiano; Fernandez-Urien, Ignacio; Steele, Robert J.C.

Published in:
United European Gastroenterology Journal

DOI:
10.1177/2050640618798182

Publication date:
2018

Document Version
Peer reviewed version

Reproducibility and accuracy of visual estimation of polyp size in large colorectal polyps

Authors: Maria Magdalena Buijs (1,2), Robert James Campbell Steele (3), Niels Buch (1), Thomas Kolbro (1), Erik Zimmermann-Nielsen (1), Morten Kobaek-Larsen (2), Gunnar Baatrup (1,2)

Institutions:

1. Department of Surgery, Odense University Hospital, Svendborg, Denmark
2. Department of Clinical Research, University of Southern Denmark, Odense, Denmark
3. Department of Surgery, Ninewells Hospital and Medical School, Dundee, Scotland

Abstract

Background: Previous studies indicate that visual size estimation of polyp size (in-situ) tends to differ from post fixation measurements, which effects allocation to surveillance intervals. Little is known about interobserver variation of in-situ measurements.

Aim: The primary objective was to assess interobserver variation of in-situ measurements of colorectal polyps. Secondary objectives were the agreement of in-situ measurements with post-fixation measurements, and the agreement on detection of polyps over 2 cm between these measurements.

Methods: Interobserver variability of in-situ polyp size measurements was assessed between a primary diagnostic colonoscopy and the secondary therapeutic colonoscopy by dedicated endoscopists, in patients that were referred for an advanced polypectomy. After excision pre- and post-fixation polyp sizes were measured with a ruler in three dimensions.

Results: A total of 40 patients, with 45 polyps, were included in the study. The difference between the two in-situ measurements was 2.4 mm (95% limits of agreement (LA): -14.2 – 19.0). The differences between the second in-situ and pre-fixation measurement in comparison to post-fixation measurements were 0.1 mm (95% LA: -8.9 – 9.1) and 1.0 mm (95% LA: -5.6 – 7.6). Cohen's Kappa on detection of ≥ 20 mm polyps in agreement with post-fixation measurements was 0.65 in the primary and 0.88 in the secondary in-situ measurements.

Conclusions: This study shows a large variation between in-situ size measurements, which has consequences for referral for advanced polypectomies and surveillance. Dedicated endoscopists using instruments to assess polyp sizes had a smaller variation of difference to post-fixation sizes than non-dedicated endoscopists as well as a significant higher agreement on detection of ≥ 20 mm polyps compared to post-fixation measurements, indicating that dedication and measuring polyps with an instrument of known diameter is likely to improve the quality of in-situ measurements.

Background

The majority of colorectal cancers are thought to arise in benign adenomatous polyps.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** In screening and routine clinical practice the number and size of colorectal polyps and histology determines treatment and surveillance interval.**{** ADDIN EN.CITE <EndNote><Cite><Author>Danske Regioners tværregionale implementeringsgruppe vedr. tarmkræftscreening</Author><Year>2014</Year><RecNum>103</RecNum><DisplayText><style face="superscript">3</style></DisplayText><record><rec-number>103</rec-number><foreign-keys><key app="EN" db-id="rtxeaavscat9d8er95dvxatif22d99axs0x5" timestamp="1511866323">103</key></foreign-keys><ref-type name="Report">27</ref-type><contributors><authors><author>Danske Regioners tværregionale implementeringsgruppe vedr. tarmkræftscreening,</author></authors></contributors><titles><title>Screenings- or adenomankontrol program for tyk- og endetarmskræft - Guidelines for koloskopi og patologi</title></titles><dates><year>2014</year></dates><pub-location>Copenhagen</pub-location><urls></urls></record></Cite></EndNote>**}** Endoscopic mucosal resection (EMR) or endoscopic submucosal dissection (ESD) is the recommended treatment for polyps with a basis larger than 1.5 cm in our center. Previous studies assessing the quality of in-situ measurements found that these measurements are generally larger than post-fixation measurements.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** In a large retrospective study 46 % of the polyps with an in-situ size of over 1 cm had a size of less than 1 cm in post-fixation measurements.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** Eichenseer et al. investigated the consequences of these inconsistent measurements with respect to the timing of surveillance colonoscopy in polyps with in-situ sizes between 10 and 25 mm.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** In 63 % of these polyps a size-difference of at least 33 % was detected between in-situ and post-fixation measurements, leading to inappropriate surveillance recommendations in over a third of the cases, regardless of histology and the number of detected polyps. Multiple studies show that in-situ measurements differ significantly from post-fixation measurements, but there are few studies assessing interobserver variation of in-situ measurements.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** In this study we assess the interobserver variation by comparing in-situ measurements in two consecutive colonoscopies.

Methods

Study design

This interobserver study assesses the reproducibility of in-situ polyp size measurements as well as the accuracy in comparison to post-fixation measurements of polyp size. Endoscopists of the primary colonoscopy were not informed of this study during inclusion. The endoscopists performing the EMR and ESD were three dedicated endoscopists (NB, TK and GB) with over 10 years of experience with colonoscopy. They all used instruments with known size to compare with the polyps in-situ, usually a 25 mm snare.

Study population

All patients referred to Odense University Hospital for advanced polypectomies (EMR or ESD), were invited to participate in this study. Their primary colonoscopies were performed in our colonoscopy unit in either Nyborg or Svendborg, Denmark. In case of multiple polyps, the two largest were included for each patient. Exclusion criteria were piecemeal resection with over 3 pieces and multiple polyps at the same location.

Size measurements

In-situ measurements during the first colonoscopy were performed without the help of instruments, by estimating the largest diameter. In the second colonoscopy all endoscopists reported which instrument they used to compare polyp size with, as well as the size in three dimensions. Pre-fixation measurements were carried out by pinning the polyps on Styrofoam and subsequently measuring size in three dimensions with a ruler. Post-fixation measurements were carried out by a pathologist who measured all polyps in three dimensions.

Data collection

Data on patient demographics and the first colonoscopy was retrospectively collected from the electronic patient journal including age, sex, indication for colonoscopy, number of detected polyps, localization and morphology of the polyps as well as in-situ size measurements. Polyp localization, size measurements and morphology of the second colonoscopy and subsequent histologic assessments were collected prospectively.

Statistics

This study was planned as a pilot-study, and aimed to include 30 patients. Due to the number of patients with incomplete data and/or piecemeal resection 40 patients were included. Basic characteristics are presented as means with standard deviations for continuous variables and percentages for binary and ordinal variables. Interobserver agreement between the different size measurements was determined by Bland-Altman graphs, determining the 95% limits of agreement of each comparison. Significance of agreement between the size measurements was calculated with a paired T-test. Agreement on detection of ≥ 20 mm polyps between the size measurements was determined by Cohen's Kappa. Statistical analyses were performed using Stata IC 15.0.

Ethics

The Ethics committee was informed of this study, but not required to give their permission due to the noninvasive nature of this study. The data protection agency gave their consent (case nr: 16/14701). All patients received written information on the study and were given the opportunity to talk to the primary investigator. All patients signed an informed consent form before entering the study and could withdraw from the study at any time without consequences.

Results

Fifty patients gave consent, of which 10 were excluded due to multi-piecemeal resections or multiple polyps in the same location. Forty patients with a total of 45 polyps were enrolled between July 1st 2016 and December 31st 2016. The primary colonoscopy was performed by trainees (20 %), surgeons (62.5 %) and colorectal surgeons (17.5 %). Data of the primary colonoscopy were missing on size (n = 9) and morphology (n = 4). Nine piecemeal resections were performed and the pathologist received a total of 11 polyps in multiple pieces. All available size measurements were used in the analysis (Table 1).

[t/Table 1]

The average polyp sizes were 23.9 ± 9.3 mm and 21.1 ± 8.5 mm, in the primary (OC1) and secondary (OC2) colonoscopy. The pre- and post-fixation polyp sizes were 21.7 ± 8.5 mm and 19.3 ± 8.5 mm, respectively. The agreement between the different measurements expressed by limits of agreements is presented in Figure 1 and 2 and Table 2. ***[t/Figure1,2 + Table 2]***

A sub-analysis was performed to assess the agreement on large polyps (≥ 20 mm), using Cohen's kappa (Table 3). ***[t/Table 3]*** The agreement on detection of ≥ 20 mm polyps between the dedicated endoscopists and post-fixation measurement was good (Cohen's Kappa: 0.88), and significantly higher than the agreement between the primary colonoscopy and post-fixation measurement.

Discussion

This study shows a large variation between the in-situ polyp size measurements of endoscopists in two consecutive colonoscopies. This variation of in-situ measurements compared to post-fixation measurements is larger in non-dedicated endoscopists as compared to dedicated endoscopists. The reduced variation in the second colonoscopy is probably due to the combination of more experience as well as the use of instruments with known size when measuring in-situ polyp size. The significant higher agreement on detection of ≥ 20 mm by dedicated endoscopists in comparison to post-fixation size is relevant, since the Danish national guidelines regard polyps ≥ 2 cm as high-risk polyps that warrant a surveillance colonoscopy after 1 year instead of 3 years for ≥ 1 cm polyps.**{** ADDIN EN.CITE <EndNote><Cite><Author>Rasmussen</Author><Year>2014</Year><RecNum>361</RecNum><DisplayText><style face="superscript">11</style></DisplayText><record><rec-number>361</rec-number><foreign-keys><key app="EN" db-id="rtxeaavscat9d8er95dvxatif22d99axs0x5" timestamp="1528364966">361</key></foreign-keys><ref-type name="Report">27</ref-type><contributors><authors><author>Rasmussen, M.</author><author>Ingeholm, P.</author><author>Linnemann, D.</author><author>Larsen, O.</author><author>Bang, S.</author></authors></contributors><titles><title>Screenings- og adenomkontrol program for tyk- og endetarmskræft [Screening and adenoma control program for colorectal cancer]</title></titles><dates><year>2014</year></dates><pub-location>Copenhagen</pub-location><publisher>Danish Regions</publisher><urls></urls></record></Cite></EndNote>**}**
There was also some variation between pre- and postfixation sizes, without a clear correlation depending on polyp size. The differences between pre- and postfixation sizes are larger than in the study of Turner et al.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}**, in which a difference of 0.3 mm (95% LA: -3.4 – 2.8) was shown in 107 polyps. This smaller difference could be explained by the inclusion of smaller polyps as well as a larger sample size in Turner's study. Although the variation in our study is large, the mean polyp sizes are similar in both colonoscopies, as well as pre- and post-fixation measurements. A systematic difference between the measurements is therefore unlikely.

Both a limitation and strength of this study is that we retrieved the in-situ size measurement from the first colonoscopy from the patient file, without informing the primary endoscopists of this study. Our goal was to make a comparison of the sizes that resembles daily clinical routine. Not all primary endoscopists specified the polyp size on referral for advanced polypectomy, which caused some data to be missing. This did not affect the results of this study significantly (data not presented).

Another limitation of this study is the small number of polyps, but even in this small population a large variation in size estimations can be found, which cannot be explained by a lack of experience since the majority of endoscopists were certified surgeons (80%).

Kim et al. investigated interobserver variability in 40 endoscopists (16 experts, 24 beginners) and found that the use of an open biopsy forceps increased the diagnostic accuracy in both experts and beginners.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** The intra-class correlation coefficient on agreement of visual estimation

was 0.69 in beginners and 0.84 in experts, increasing to 0.76 and 0.89 respectively when using biopsy forceps. In polyps ≥ 10 mm the difference between the true polyp size and visual estimation was 2.7 ± 0.8 mm, which is similar to our study.

The consequences of estimating in-situ size incorrectly are currently minor, due to the habit of removing all polyps seen in endoscopy. However, the in-situ size does affect those patients who are referred for an advanced polypectomy due to size and location of the polyp. More accurate size estimation could decrease the amount of secondary colonoscopies with corresponding discomfort, risks of perforation and bleeding in these patients.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** Another group of patients that are affected by incorrect polyp sizes are those in which surveillance intervals are based on in-situ size, especially because of piecemeal resections. The aforementioned study of Eichenseer showed a large variation of mis-sizing and consequential inappropriate surveillance recommendations between individual endoscopists.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** The range of mis-sizing polyps varied from 0-91% and the subsequent inappropriate surveillance recommendations from 0-67%. Comparing polyps with an instrument of known size (for example biopsy forceps or snare), is likely to improve the quality of in-situ estimations.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}**

Since the introduction of colon capsule endoscopy and CT colonography, for which criteria are developed to refer patients for a therapeutic colonoscopy, a renewed interest in the natural history of polyps has developed. Two CT colonography studies following patients for three years after detection of small polyps (< 10 mm) showed a regression of polyp size in about a third of the patients and complete disappearance in 10-14% of the polyps.**{** ADDIN EN.CITE **{** ADDIN EN.CITE.DATA **}}** In this new era we might consider to follow-up on polyps instead of resecting them, which increases the need of accurately assessing polyp sizes in-situ.

Another development is the increasing interest of artificial intelligence in endoscopy. An algorithm might be developed that can estimate polyp size closer to the post-fixation size than is humanly possible. In order to train a learning-based algorithm it is necessary to develop a reliable gold standard for polyp size, and assess the quality of current measurements. If possible, these polyp sizes will be correlated to advanced histology, in order to select those patients that are most likely to benefit from a therapeutic colonoscopy.

Conclusion

This study shows a large variation between in-situ colorectal polyp size measurements, which has consequences for referral for advanced polypectomies and surveillance. Dedicated endoscopists using instruments to assess polyp sizes had a smaller variation of difference to post-fixation sizes than non-dedicated endoscopists, indicating that dedication and/or comparison of the polyp with an instrument of known diameter is likely to improve the quality of in-situ measurements.

References

{ ADDIN EN.REFLIST }

Table 1: Demographics

	Patients (n = 40)	
Age (years)	70.6 ± 8.6	
Sex	22 males (55.0 %)	
Indication of primary colonoscopy		
National screening program	19 (47.5 %)	
Symptoms	14 (35.0 %)	
Surveillance	4 (10.0 %)	
Miscellaneous	3 (7.5 %)	
Interval between colonoscopies (days)	29.3 ± 32.4	
	Polyps (n = 45)	
Measurement technique 2nd colonoscopy		
2 cm snare	2 (5.3 %)	
2.5 cm snare	26 (68.4 %)	
3 cm snare	1 (2.6 %)	
4 cm snare	7 (18.4 %)	
ESD-instruments	1 (2.6 %)	
Endoscope (while looping)	1 (2.6 %)	
Not specified	7 (18.4 %)	
EMR / ESD	38 (84.4 %) / 7 (15.6 %)	
Polyp location	**Primary colonoscopy**	**Secondary colonoscopy**
Caecum	8 (17.8 %)	7 (15.6 %)
Ascending colon	4 (8.9 %)	8 (17.8 %)
Hepatic flexure	2 (4.4 %)	-
Transverse colon	3 (6.7 %)	4 (8.9 %)
Splenic flexure	4 (8.9 %)	2 (4.4 %)
Descending colon	1 (2.2 %)	2 (4.4 %)
Sigmoid colon	11 (24.4 %)	12 (26.7 %)
Rectum	10 (22.2 %)	10 (22.2 %)
Not specified	2 (4.4 %)	-
Polyp morphology	**Primary colonoscopy**	**Secondary colonoscopy**
Pedunculated	10 (22.2 %)	13 (28.8 %)
Sessile	22 (48.9 %)	25 (55.6 %)
Non-polypoid	9 (20.0 %)	7 (15.6 %)
Not specified	4 (8.9 %)	-
Histology		
Hyperplastic	3 (6.7 %)	
Sessile serrated adenoma	3 (6.7 %)	
Tubular adenoma	29 (64.4 %)	
Tubulovillous adenoma	8 (17.8 %)	
Villous adenoma	-	
Adenocarcinoma	2 (4.4 %)	
Dysplasia (n = 40)		
None	1 (2.5 %)	
Low grade dysplasia	36 (90.0 %)	
High grade dysplasia	3 (7.5 %)	

Figure 1: OC1 vs OC2 with limits of agreement

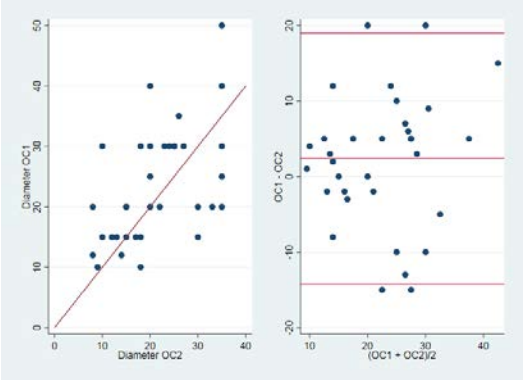

Figure 2: OC2 vs post-fixation with limits of agreement

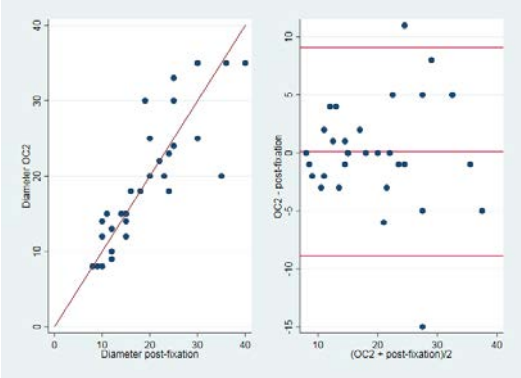

Table 2: 95% limits of agreement and average difference between size measurements

	95% Limits of agreement	Difference in mm (95% CI)	P-value
OC1 vs OC2 (n=37)	-14.2 - 19.0	2.4 (-0.4 – 5.2)	0.046
OC1 vs post-fixation (n=26)	-13.2 - 16.9	1.8 (-1.2 – 4.9)	0.12
OC2 vs post-fixation (n=33)	-8.9 - 9.1	0.1 (-1.5 – 1.8)	0.44
Pre-fixation vs post-fixation (n=33)	-5.6 - 7.6	1.0 (-0.2 – 2.2)	0.044

P-value of a one-sided paired t-test comparing the means of different size measurements.

Table 3: Agreement on ≥ 20 mm polyp sizes

	Observed agreement	Cohen's Kappa (95 % CI)
OC1 vs OC2 (n=37)	83.8 %	0.65 (0.60 – 0.70)
OC1 vs post-fixation (n=26)	81.5 %	0.63 (0.56 – 0.69)
OC2 vs post-fixation (n=33)	93.9 %	0.88 (0.76 – 0.94)
Pre-fixation vs post-fixation (n=33)	93.9 %	0.88 (0.83 – 0.93)