1. Title

Name of the study: ColoReg: A register for findings and videos of endoscopies in gastroenterology

Name of the study (English): ColoReg: A Register for Gastrointestinal Endoscopy Reports and Examination Videos

Short name (if applicable): ColoReg
Version number of the protocol: 0.4_RBK
Date of the minutes: 10.03.2024

2. Responsibilities

Study director: Dr. med. Thomas J. Lux
Institution: Medical Clinic II, University Hospital Würzburg
Participating institutions and departments:
Gastroenterology, Medical Clinic II, University Hospital Würzburg

3. Signatures to confirm the minutes

Clinic director or head of an institute (if not study director):

_____	_____
Place, date	Prof. Dr. Hermann Einsele

Head of Gastroenterology:

_____	_____
Place, date	Prof. Dr. Alexander

Director of Studies:

_____ _____

Place, date Dr. Thomas J. Lux

4. Rationale

Background

Endoscopic procedures such as gastroscopies and colonoscopies are a fundamental part of diagnostics and treatment in gastroenterology. Gastroscopy can reliably identify or rule out serious causes of abdominal pain and colonoscopy is used worldwide as part of colorectal cancer screening. Endoscopic procedures can be used therapeutically to treat pathological changes, e.g. precancerous lesions or cancer, curatively or in a symptom-oriented manner.

Technical and scientific progress

The options available for gastroenterological endoscopy are constantly increasing. Improvements in camera systems have increased the resolution and color depth of videos. In addition, digital post-processing of the image signal or changes to the light source ("*narrow band imaging*") have developed alternatives to previously used color solutions.

Significant progress has also been made in the field of therapy. While benign, larger polyps were still primarily removed surgically in the 2000s, early forms of stomach and bowel cancer are now also primarily removed endoscopically.

Clinical studies vs. real world data

As pleasing as the rapid progress in the field of endoscopy is, this progress also brings with it problems. Prospective clinical studies naturally require a high degree of planning and, in particular, only reflect the application in non-academic centers to a limited extent. As the conduct of a clinical trial itself can have a significant effect on treatment, findings should always be checked against "real world" data. For example, a study by Vavricka et al.[1] showed that third parties the proportion of colonoscopies for screening

[1] Vavricka, S. R., Sulz, M. C., Degen, L., Rechner, R., Manz, M., Biedermann, L., Beglinger, C., Peter, S., Safroneeva, E., Rogler, G., & Schoepfer, A. M. (2016). Monitoring colonoscopy withdrawal time

with detected precancerous lesions from 21% to 36% after the examiners were informed that the duration of the procedure would be measured.

Training and further education

As described above, the images seen during an endoscopy and the therapies carried out are subject to constant change. New aspects of the findings - both of healthy and pathological origin - become visible through higher image quality. Large quantities of normal and pathological findings are required to train good endoscopists.

Few illustrative materials (e.g. individual video clips) are available, particularly for novel therapies.

State of research

To our knowledge, there is no comprehensive register for the processing of video and diagnostic texts of clinical endoscopies, neither in Germany nor internationally.

Justification of the study to be carried out

The need for structured and standardized recording of findings and interventions in gastroenterological endoscopy is evident. The planned study aims to record the structuring of the findings and compare them with current recommendations for reporting. This will continuously monitor the objectivity and reproducibility of the findings and enable us to identify discrepancies between theory and practice. In addition, the study will comprehensively record the interventions performed, including the indication, the technique used, possible immediate complications and the macro- and microscopic outcome. This is particularly relevant in the context of the application of novel forms of therapy.

Another key element of the study is the creation of a database for diagnostic images and video clips of interventions performed. These databases will not only advance research in the field of endoscopy, but will also provide valuable resources for the education and training of specialists. In addition, the publication of the established databases will make an important contribution to the validation of the increasing number of AI applications in the field of endoscopy. By providing a comprehensive and high-quality database, AI models can be trained more efficiently and their performance can be better evaluated. This is particularly important in light of the rapid technological developments and the increasing integration of AI into clinical practice.

---

significantly improves the adenoma detection rate and the performance of endoscopists. Endoscopy, 48(3), 256-262. https://doi.org/10.1055/s-0035-1569674

Risk-benefit assessment

The ethical dimension of the study requires careful consideration of the risks and benefits. The primary risk of the study lies in the handling of personal data. To minimize this risk, all data will be de-identified before being transferred to the registry. In addition, the study participants are informed in detail about the type of data collection, de-identification, further processing and publication. This procedure ensures that the privacy of the participants is protected and ethical standards are adhered to.

In contrast to this is the considerable benefit of the study for clinical application, research and training. The systematic recording and structuring of findings and interventions will improve the quality of medical care and advance research in the field of gastroenterological endoscopy. In addition, the databases created will provide valuable resources for the education and training of specialists and enable the validation of AI applications. Against this backdrop, the risk can be considered minimal, particularly due to the data security precautions taken, while the potential benefits of the study are far-reaching and of great importance.

5. Study objectives

The primary objective of this registry study is to systematically record and analyze the efficacy and safety of various endoscopic procedures used to prevent and treat colorectal cancer and its precursors. In addition, relevant image material is to be selected so that it can be used for training purposes and further research projects.

The secondary objectives include the evaluation of the quality of findings and treatment in the context of the current guidelines. For example, questions such as: "How often was the follow-up interval selected in accordance with the guidelines? And if this was deviated from, can this be justified by the documentation of the findings?"

The tertiary goal of the study is to increasingly automate the process of data extraction and processing. Subsequently, the developed system is to be propagated to non-academically interested treatment centers in order to establish a nationwide, possibly international, register.

For patients themselves, there is no direct benefit from participating in the study. The benefit to society lies on the one hand in the rapid assessment of the success and benefits of the methods used under "real world" conditions, and on the other hand in the promotion of research and training.

Primary study objective

The following parameters define an endoscopic ablation technique:

- Technology used
    - En bloc ablation vs. fractional ablation
    - + / - Injection
    - Technology:
        - Endoscopic mucosal resection (EMR)
        - Endoscopic submucosal dissection (ESD)
        - Endoscopic full-thickness wall resection (EFTR)

- Instrument used
    - Cold sling
    - Diathermy loop (incl. current setting used

In our study, an endoscopic technique is defined as effective if one of the following results was achieved in more than 90% of the applications:

- Histologically confirmed R0 ablation
- Inconspicuous follow-up examination

In our study, an endoscopic technique is defined as safe if at least one of the following events occurs in less than 3% of applications:

- Perforation
- Endoscopically uncontrollable bleeding
- Need for an unplanned follow-up examination with evidence of interval bleeding or perforation
- Detection of a stenosis with impassable lumen narrowing in a follow-up examination

The target diagnostic reliability is 95 %.

Secondary study objectives

- The recommended follow-up interval depending on polyp number, size and histopathology should correspond to the current national guideline at the time of the examination in 90% of examinations.
- Polyps recovered in > 95% of cases
- Minor bleedings (defined as continuous bleeding > 30 seconds with subsequent clip supply)

Tertiary study goal

Video processing: The software should identify and mask the sensitive information contained in 100% of 1,000 examinations (name, date of birth, date of examination, video sections outside the human body) or mark the video as "to be processed manually".

Text processing: The software should recognize and remove the sensitive information contained in 100% of 1,000 examinations (names, date of birth, case numbers, examiners, telephone numbers) or mark the text as "to be edited manually".

6. Study design

This study is a prospective registry study in which data is collected from patients undergoing endoscopic examination for diagnosis or treatment.

Included patients will be followed up for at least 3 (no recommendation for follow-up within one year) but not more than 12 months before anonymization and transfer of data.

7. Study population

The study population includes patients undergoing endoscopic examination for diagnostic or therapeutic purposes. Study participants are recruited as part of their treatment at the participating center.

Patients are eligible to participate if they fulfill the following inclusion criteria:

- Age of at least 18 years
- Ability to provide information and consent
- Indication for an endoscopic examination

Patients are excluded if they fulfill at least one of the following exclusion criteria:

- Inability to perform the planned endoscopic examination for medical reasons

8. Individual study program

Place of the intended investigations & processing of personal data:

Robert Bosch Hospital
Auerbachstrasse 110
70376 Stuttgart

A list of all participating trial centers can be found in the patient information

Place of processing anonymized data records:

University Hospital Würzburg
Medical Clinic and Polyclinic II, Gastroenterology
Oberdürrbacher Str. 6
97080 Würzburg

Information and consent

The indication for an endoscopy is given by an internal medicine doctor. The patient is informed about the procedure and risks in writing using the information sheets used in everyday clinical practice.

After checking the inclusion and exclusion criteria, a member of staff from the Department of Internal Medicine II will provide information about the study at the earliest possible time, on the day the indication is established.

Study dates / timetable

Unscheduled re-endoscopies related to the primary examination within 3 months and scheduled follow-up appointments within one year are included. There are no additional appointments within the study.

Measures

As part of the endoscopic examination, the video signal during the examination and the full text of the findings with associated histology are recorded in de-identified form. In the case of ablation using a diathermy snare, the device setting used (current curve) should also be recorded.

9. Adverse events

Not applicable.

10. Biometric aspects

Our study is intended to record the endoscopic procedures currently used, the quality of the current findings and videos of the corresponding examinations in various centers. The anonymized data collected will be published and used for further research and teaching. The feasibility and usefulness of the study should be examined on the basis of the above-mentioned study objectives. The following data will be collected for this purpose:

The planned exploratory comparisons, such as subgroup comparisons, are not of a confirmatory nature. The reasons for this are the high variance in the literature data, the fluctuating methodology of the underlying sources and the number of patients required.

Primary study objectives

1. Primary successful endoscopic ablation

Objective: To estimate the rate of primary successful endoscopic ablation of polyps ≥ 20 mm for the EMR and ESD techniques.

Measure: Percentage success rate with 95% confidence interval for each technique.

Exploratory analysis: Odds ratio for primary successful ablation (ESD vs. EMR) and subgroup comparisons:

- EMR en bloc vs. EMR fractionated vs. ESD

- Influence of the diathermy setting.

2. *R0 Resection rate*

Objective: To estimate the R0 resection rate for polyps ≥ 20 mm for EMR and ESD.

Measure: Percentage R0 resection rate with 95% confidence interval for each technique.

Exploratory analysis: odds ratio for R0 resection (ESD vs. EMR) and subgroup comparisons:

- EMR en bloc vs. EMR fractionated vs. ESD

- Influence of the diathermy setting.

3. *Inconspicuous follow-up*

Objective: To estimate the rate of inconspicuous follow-up examinations for polyps ≥ 20 mm after EMR and ESD.

Measure: Percentage rate of inconspicuous follow-up with 95% confidence interval for each technique.

Exploratory analysis: Odds ratio for inconspicuous follow-up (ESD vs. EMR) and subgroup comparisons:

- EMR en bloc vs. EMR fractionated vs. ESD

- Influence of the diathermy setting.

4. *(u. 5) Major complication rate*

Objective: To estimate the rate of major complications (endoscopically uncontrollable bleeding or perforation) in polyps ≥ 20 mm for EMR and ESD.

Measure: Percentage rate of major complications with 95% confidence interval for each technique.

Exploratory analysis: odds ratio for major complications (ESD vs. EMR) and subgroup comparisons:

- EMR en bloc vs. EMR fractionated vs. ESD

- Influence of the diathermy setting.

Secondary study objectives

1. *Duration of treatment*

Objective: To estimate the average treatment duration for polyps ≥ 20 mm for EMR and ESD.

Measure: Mean and standard deviation of treatment duration (minutes) for each technique.

Explorative analysis: influence of subgroups:

- EMR en bloc vs. EMR fractionated vs. ESD
- Use of a cap (+/-)
- Influence of the diathermy setting.

2. *Unplanned re-intervention*

Objective: To estimate the rate of unplanned re-interventions within 3 months after EMR and ESD.

Measure: Percentage rate of unplanned re-interventions with 95% confidence interval for each technique.

Exploratory analysis: odds ratio for unplanned re-intervention (ESD vs. EMR) and subgroup comparisons:

- EMR en bloc vs. EMR fractionated vs. ESD
- Influence of the diathermy setting.

Case number estimation

The case number estimation is based on the **desired precision** (width of the 95% confidence intervals) for the estimation of the event rates for the primary endpoints. Despite the purely descriptive nature of the primary endpoints, a Bonferroni correction is also applied to obtain precise estimates for EMR and ESD separately

In the currently available literature, the success and complication rates for EMR and ESD vary greatly. In addition, much of the data is based on patients from academic centers. Therefore, the precise estimation of event rates in terms of our clinical endpoints provides highly relevant information for clinical and scientific purposes.

Secondary, exploratory comparisons are made separately and are not of a confirmatory nature (odds ratio calculations, subgroups).

Estimates from the literature

The following event rates are estimated on the basis of the existing literature (clear meta-analysis: Hassan C, Repici A, Sharma P, et al., Gut 2016;65:806-820) and new results from a systematic search. Conservative assumptions were made for certain endpoints in order to calculate the number of cases:

- **Primary successful ablation:**
  - EMR: 85% (95%-KI: 80% - 90%)
  - ESD: 95% (95%-KI: 93% - 97%)
  -
- **R0 resection:**
  - EMR: 75% (95%-KI: 70% - 80%)
  - ESD: 90% (95%-KI: 87% - 93%)

- **Inconspicuous follow-up:**
  - EMR: 86.2% (95%-KI: 85.3% - 87.1%)
  - ESD: 90% (estimated, based on similar studies).

- **Major complications (bleeding):**
  - EMR: 3.6% (95%-KI: 2.9% - 4.3%)
  - ESD: 6.5% (95%-KI: 5.9% - 7.1%)

- **Major complications (perforations):**
  - EMR: 1.95% (95%-KI: 1.5% - 2.4%)
  - ESD: 5.3% (95%-KI: 4.5% - 6.1%)

Calculation of the number of cases for primary endpoints with Bonferroni correction

Adjustment of the significance level:

- Original significance level: α=0.05
- Number of primary endpoints: k=5
- Adjusted significance level: α_{adj}=α / k= 0.05= 0.01
- The z-value for a 99% confidence interval (corresponding to α_{adj}=0.01 is: z= 2.576

The number of cases is calculated using the adjusted formula:

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2}$$

Calculations

1. major complications (bleeding):

- **EMR:** p = 0.036, target width of the CI: ±1.5%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}036 \cdot (1-0{,}036)}{(0{,}015)^2} \cong 411$$

- **ESD:** p=0.065, target width of the CI: ±2%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}065 \cdot (1-0{,}065)}{(0{,}02)^2} \cong 870$$

2. major complications (perforations):

- **EMR:** p=0.0195, target width of the CI: ±1%

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}0195 \cdot (1-0{,}0195)}{(0{,}01)^2} \cong 128$$

- **ESD:** p=0.053, target width of the CI: ±1.5%

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}053 \cdot (1-0{,}053)}{(0{,}015)^2} \cong 289$$

3. inconspicuous follow-up:

- **EMR:** p=0.862, target width of the CI: ±3%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}862 \cdot (1-0{,}862)}{(0{,}03)^2} \cong 666$$

- **ESD:** p=0.90, target width of the CI: ±3%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}9 \cdot (1-0{,}9)}{(0{,}03)^2} \cong 692$$

4. R0 resection rate:

- **EMR:** p=0.75, target width of the CI: ±4%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}75 \cdot (1-0{,}75)}{(0{,}04)^2} \cong 650$$

- **ESD:** p=0.90, target width of the CI: ±3%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}9 \cdot (1-0{,}9)}{(0{,}03)^2} \cong 692$$

5. primary Successful ablation:

- **EMR:** p=0.85, target width of the CI: ±3%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}85 \cdot (1-0{,}85)}{(0{,}03)^2} \cong 785$$

- **ESD:** p=0.95, target width of the CI: ±2%.

$$n = \frac{z^2 \cdot p \cdot (1-p)}{(Breite\ d.\ KI)^2} = \frac{2{,}576^2 \cdot 0{,}95 \cdot (1-0{,}95)}{(0{,}02)^2} \cong 385$$

Consideration of the endpoints with the highest number of cases

For the most precise estimate of the endpoints, the highest case numbers after correction for multiple testing are used:

- **EMR:** Primary successful ablation: **785 patients.**
- **ESD:** Major complications (bleeding): **870 patients.**

Summary of case number planning

- The aim is to precisely estimate the event rates taking into account the Bonferroni correction (α_{adj}= 0.01)
- Separate case numbers were calculated for EMR (785 patients) and ESD (870 patients).
- A total of **1,655 patients** are **required** for both indications
- In the sense of an observational colonoscopy registry, recruitment will continue until the number of cases for both indications has been reached

11. Data management

Responsible for data processing: Dr. Thomas J. Lux

Fundamentals

Data protection is of the utmost importance for our study. All personal data collected as part of the study will be treated with the highest level of confidentiality and security. The data is collected in accordance with the General Data Protection Regulation (GDPR) and the principles of "*privacy by default*" and "*privacy by design*". Our aim is to protect the privacy of our study participants while collecting the necessary data to answer our research questions.

The data protection and security concept was therefore an integral part of the project planning. No data is collected that goes beyond what is necessary for treatment and further care planning. Personal data is stored at the center of the medical treatment taking place on encrypted data carriers (digital data) and is stored in an access-protected location (digital and analog data).

At the end of the follow-up period, the data is de-identified at the treating center and the original data set is destroyed. Only data that has been anonymized in accordance with the patient information and the information sheet will be transferred to the study center (Würzburg) and included in the registry.

In summary, it can be stated that participation in the study *does not* increase the number of people with access to personal data.

Data collection, source data

The video data is tapped directly at the endoscopy processor and stored pseudonymously on a computer in the treating clinic

The texts of the examination findings and the histopathological findings are also stored pseudonymously on this computer.

Data storage in the Participating Study Center

The pseudonymized data is stored in encrypted form (*Advanced Encryption Standard* (AES), key length: 256-bit).

Anonymization / pseudonymization

In order to ensure the greatest possible data security at every stage of data processing, we use a multi-layered procedure. This includes the following steps:

- Ensuring anonymity of the video material (in particular: No patient data visible in the video; all video segments show only the inside of the body)
- Ensuring anonymity of the findings texts (in particular: no remaining names, dates of birth, examination date in the text)
- During the follow-up period, the data set contains a pseudonym as the only identifiable characteristic, which contains neither the patient's name nor date of birth
- Before the data is transferred to the study center, it is irrevocably deleted

The anonymized data set created for the register contains the following components:

- Video from inside the body without names or dates of birth in the image
- Findings texts without names and dates of birth in the text

The anonymity of medical image material is an increasingly discussed topic. This discussion is based on the fact that - with the help of the original examination images - it is theoretically possible to trace the above-mentioned video back to the patient.

In accordance with recital 26.3[2] of the GDPR, the means used for identification - in this case, access to the patient's medical data - should be taken into account. For persons with access to these specific medical data of the patient, there would be no gain in information by assigning the study data set.

In addition, according to recital 26.4[3] of the GDPR, it should be determined how likely it is that appropriate means of identifying the person will be used, taking into account costs and time. Since the medical data is not for sale and the time required for identification would be considerable (attempt to find a specific video from a database using individual images), we finally refer to recital 26.5[4] and classify our data set as

[2]To determine whether a natural person is identifiable, account should be taken of all the means likely to be used by the controller or another person to identify the natural person directly or indirectly, such as singling out.

[3] In determining whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the cost of identification and the time required, taking into account the technology available at the time of processing and technological developments.

[4] The principles of data protection should therefore not apply to anonymous information, i.e. information that does not relate to an identified or identifiable natural person, or personal data that has been anonymized in such a way that the data subject cannot or can no longer be identified.

anonymized. The residual risk described here is also presented accordingly in the patient information and education.

Data transmission and data storage at the study center

The pseudonymized data is viewed and processed by employees of the treating specialist department.

The anonymized data is viewed and processed by employees of the Department of Internal Medicine II at the University Hospital of Würzburg.

In line with the study objective, the anonymized data will be made available to the public in whole or in part. This includes the cases mentioned in the patient information (online teaching and learning platform, video, image and findings data sets for further research).

12. Ethical and regulatory aspects

The investigators are responsible for ensuring that the study is conducted in accordance with the Declaration of Helsinki and the rules for Good Clinical Practice and the corresponding regulations.

Patient information and consent

It is the responsibility of the investigators and participating centers to explain the study to each patient and obtain consent signatures before any study-specific procedures are performed.

Participants have the right to withdraw their consent. If not specified in more detail, all data traceable to this study participant will be deleted as quickly as possible. At the patient's request, it is also possible to anonymize the data set immediately by terminating the follow-up interval. Data records that have already been anonymized cannot be assigned retroactively and therefore cannot be deleted.

Approval of the Ethics Committee

The study will only be conducted with the approval of the responsible ethics committee.

Amendments

Amendments to the study protocol are submitted in writing to the ethics committee.

13. Index

AI:	artificial intelligence
for example:	for example
if applicable:	if applicable
vs:	versus
EMR:	Endoscopic mucosal dissection
ESD:	Endoscopic submucosal dissection
EFTR:	Endoscopic full wall resection
incl:	inclusive
R0:	Microscopically complete ablation
H0:	Null hypothesis
n:	number
GDPR:	General Data Protection Regulation
AES:	Advanced Encryption Standard
above:	Mentioned above

14. Attachments

- Patient information
- Declaration of consent