

Amalgamation of cloud-based colonoscopy videos with patient-level metadata to facilitate large-scale machine learning



Authors

Rajesh N. Keswani¹, Daniel Byrd², Florencia Garcia Vicente², J. Alex Heller², Matthew Klug², Nikhilesh R. Mazumder¹, Jordan Wood¹, Anthony D. Yang³, Mozziyar Etemadi^{2,4}

Institutions

- 1 Digestive Health Center, Northwestern Medicine, Chicago, Illinois, United States
- 2 Department of Anesthesiology, Northwestern Medicine, Chicago, Illinois, United States
- 3 Surgical Outcomes and Quality Improvement Center, Department of Surgery, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States
- 4 Department of Biomedical Engineering, McCormick School of Engineering, Chicago, Illinois, United States

submitted 10.7.2020

accepted after revision 29.10.2020

Bibliography

Endoscopy International Open 2021; 09: E233–E238

DOI 10.1055/a-1326-1289

ISSN 2364-3722

© 2021. The Author(s).

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14,
70469 Stuttgart, Germany

Corresponding author

Rajesh N. Keswani MD MS, Digestive Health Center,
Northwestern Medicine, 676 N. St. Clair, Suite 1400, Chicago,
IL 60611, USA
raj-keswani@northwestern.edu

ABSTRACT

Background and study aims Storage of full-length endoscopic procedures is becoming increasingly popular. To facilitate large-scale machine learning (ML) focused on clinical outcomes, these videos must be merged with the patient-level data in the electronic health record (EHR). Our aim was to present a method of accurately linking patient-level EHR data with cloud stored colonoscopy videos.

Methods This study was conducted at a single academic medical center. Most procedure videos are automatically uploaded to the cloud server but are identified only by procedure time and procedure room. We developed and then tested an algorithm to match recorded videos with corresponding exams in the EHR based upon procedure time and room and subsequently extract frames of interest.

Results Among 28,611 total colonoscopies performed over the study period, 21,170 colonoscopy videos in 20,420 unique patients (54.2% male, median age 58) were matched to EHR data. Of 100 randomly sampled videos, appropriate matching was manually confirmed in all. In total, these videos represented 489,721 minutes of colonoscopy performed by 50 endoscopists (median 214 colonoscopies per endoscopist). The most common procedure indications were polyp screening (47.3%), surveillance (28.9%) and inflammatory bowel disease (9.4%). From these videos, we extracted procedure highlights (identified by image capture; mean 8.5 per colonoscopy) and surrounding frames.

Conclusions We report the successful merging of a large database of endoscopy videos stored with limited identifiers to rich patient-level data in a highly accurate manner. This technique facilitates the development of ML algorithms based upon relevant patient outcomes.

Introduction

Storage of videos of full-length endoscopic procedures is becoming increasingly popular for educational, research, and quality improvement purposes [1]. Due to the length of the individual endoscopic procedures and the associated large size of the data files, videos are often stored in a network of remote servers, known more commonly as the “cloud” with limited ac-

cess in the form of application program interfaces. However, to facilitate high quality large-scale machine learning (ML) focused on endoscopic clinical outcomes, these videos must be merged with the patient-level data that only exists in the electronic health record (EHR).

There is growing literature supporting the role of ML in gastrointestinal endoscopy. Specifically, in colonoscopy, development of ML techniques has facilitated improvements in polyp

detection [2–6] and classification of polyps based on their predicted histology [7–9]. However, additional ML techniques would benefit from a tight linkage of colonoscopy videos with patient-level data. For example, there is a growing interest in the predictive value of ML assessment of inflammatory bowel disease activity during colonoscopy with patient outcomes including need for therapy escalation and surgery [10]. This research requires a large video library of patients with diverse disease activity and outcomes linked to detailed patient-level data.

Although merging large databases of endoscopic videos with EHR data could significantly accelerate the development of ML, accurate methods to perform this amalgamation have not been previously described. We hypothesized that videos stored in the cloud could be merged with patient-level data in a highly accurate fashion. Thus, our primary aim was to present a method of successfully linking patient-level EHR data with cloud stored videos. Our secondary aim was to determine the feasibility of utilizing this linked video library to rapidly generate frames of interest and develop a large ML training dataset.

Methods

Setting

This study was conducted at a single academic medical center (Chicago, Illinois, United States). All endoscopic procedures were performed at one of two locations (16 total procedure rooms) by 50 endoscopists over the study period. A waiver of informed consent was obtained by the institutional IRB (STU00211291). Videos stored from August 18, 2018 to March 14, 2020 were analyzed.

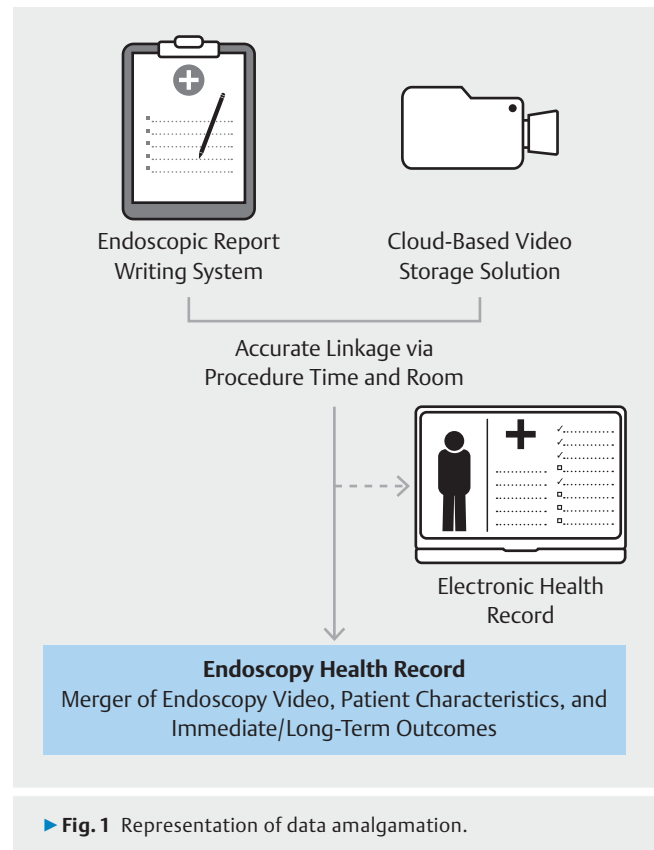
Electronic health record and video storage data sources

All endoscopic reports were written in a single endoscopic reporting system (Provation, Minneapolis, Minnesota, United States) and all EHR data was stored in a separate system (Epic, Madison, Wisconsin, United States). All videos were stored via a commercial gastrointestinal endoscopy cloud storage company (Virgo Surgical Video Solutions, San Francisco, California, United States). Procedure videos are automatically uploaded to the cloud server and are identified only by the time of the procedure and the procedure room where the procedure was performed. However, automated video recording was not enabled in all procedure rooms for the duration of the study resulting in the loss of some procedure videos.

This method of identification results in incomplete linkage of individual files between the EHR to the endoscopic report writing system and the cloud system, a common scenario in many medical centers that rely on software packages from disparate vendors that do not interoperate.

Video linkage

To match colonoscopy videos against the corresponding EHR exam record, we initially identified all exams in the endoscopy report writing system where the endoscope insertion time occurred within 3 minutes of the cloud-storage video start time.



► Fig. 1 Representation of data amalgamation.

We subsequently determined if the room corresponding with the procedure matches the room identified in the cloud-storage video metadata to confirm a match. Thus, this linkage can result in a complete “endoscopy health record” which combines the entire endoscopic procedure with patient characteristics (including demographics and disease status) with immediate and long-term outcomes (► Fig. 1).

To compute this connection, we first built a map of all exams in the EHR, which linked rooms to specific exam times. We also created a map of these exams with their start times incremented by up to three minutes so that we could map the cloud provider videos to any video in this range. Then we searched over the cloud provider video set, and mapped each video to its corresponding room, and then to a time which was attached to that room as a possible start time. By this process, a complete map was considered valid if a video only linked to one EHR exam. Without this mapping, we calculated finding the intersection between the cloud provider and the Provation health system would have taken ~12 hours. With this mapping we can compute and re-compute this intersection in ~1 second (Supplemental Fig.).

Extraction of potential frames of points of interest from videos

Photo documentation of points of interest during gastrointestinal endoscopy is standard of care at our institution. Capturing an image during endoscopy results in a unique “picture in picture” event—when the endoscopy image that is captured appears as a small inset on the larger video frame. Because our

video storage solution already recognizes these events, we developed a process to automatically extract frames surrounding these endoscopist-identified points of interest and export these frames to standalone image files. Using downloaded videos, frames can be extracted at a framerate of 30 frames per second. By utilizing a range of 4 seconds (2 seconds before and after the highlight time) and extracting every fifth frame within that range, 15 potential frames of interest around each video highlight are extracted.

Statistical analysis

Descriptive statistics were used to report procedure and patient characteristics. Mean and standard deviations were reported when appropriate. To manually validate the accuracy of this linkage, a single investigator (RNK) randomly sampled 100 video and procedure pairs. A “match” was confirmed as accurate when the still images stored in the endoscopy reporting system were identified during the linked video. The adjusted Wald method was used to estimate the overall matching accuracy and calculate the 95% confidence intervals on the entire set from this random subset. Analyses were conducted using Microsoft Excel 365 (Microsoft, Redmond, Washington, United States).

Results

Over the 19-month study period, 50,188 total procedures were performed at our institution, a majority of which were colonoscopy ($n=28,611$). In total, 21,170 procedure recordings were identified which matched with colonoscopies occurring in the same procedure room at the same procedure time (74% of all colonoscopies). Because there was gradual adoption of cloud-based recording over the study period, we analyzed the match rate in four procedure rooms where continuous video recording could be confirmed during a single month. We found that 92% of all patients (423/458) undergoing colonoscopy could be matched to a video recording.

These 21,170 colonoscopies were performed in 20,420 unique patients (54.2% male, 45.8% female; median age 58, interquartile range [IQR] 51–68). In total, these 21,170 colonoscopy videos represent 489,721 minutes of colonoscopy video performed by 50 unique endoscopists (median 214 colonoscopies per endoscopist; IQR 51–758).

Of 100 manually reviewed videos, all 100 of the videos were accurately linked to the correct patient and procedure (Wald estimate of 99% accuracy, 95% confidence interval 96.8–100%).

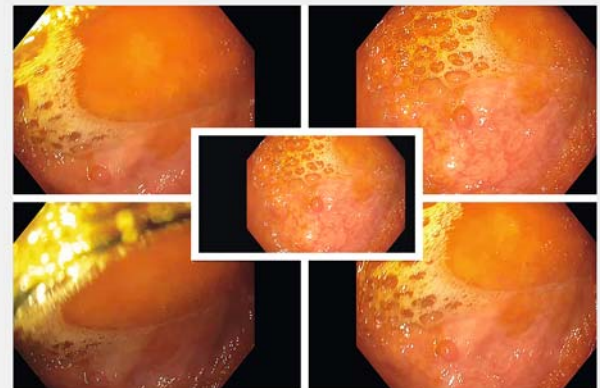
Of the 21,170 colonoscopy videos recorded, the most common procedure indication was colon polyp screening (47.3%), followed by polyp surveillance (28.9%) (► **Table 1**). Nearly 2000 (9.4%) colonoscopy procedures were performed for an indication of inflammatory bowel disease. Of colonoscopies performed for an indication inflammatory bowel disease, a similar percentage of patients underwent colonoscopy for ulcerative colitis (41%) and Crohn’s disease (47%).

Among the 21,170 matched colonoscopy videos, 179,660 total picture-in-picture events occurred with a mean of 8.5 (SD 8.8) “highlights” per video. Thus, using the automated frame

► **Table 1** Details of colonoscopy videos merged with patient-level EHR data ($n=21,170$).

Number of endoscopists performing colonoscopy	50
Number of colonoscopies per endoscopist, median (IQR)	214 (51–758)
Colonoscopy Indications ¹	
▪ Screening	10,018 (47.3%)
▪ Surveillance	6,112 (28.9%)
▪ Diagnostic	3,451 (16.3%)
▪ Inflammatory bowel disease	1,998 (9.4%)
▪ Positive stool test	312 (14.7%)
Inflammatory bowel disease indications ($n=1998$)	
▪ Crohn’s disease	930 (46.5%)
▪ Ulcerative colitis	824 (41.2%)
▪ Indeterminate/not specified	244 (12.2%)

HER, electronic health record; IQR, interquartile range.
¹ Some cases with indications fitting into ≥ 2 categories.



► **Fig. 2** Example of frame sampling around a highlighted point of interest. The central image is the picture captured by the endoscopist. Sampling frames around the highlight displays different (and sometimes more subtle) representations of the same highlight (in this case, a small polyp in the cecum).

generation algorithm, a mean of 127.5 potential frames of interest are generated from each colonoscopy (► **Fig. 2**).

Discussion

In this study, we report the successful merging of a large database of endoscopy videos stored using limited identifiers with rich patient-level data. We found that this linkage was highly accurate, identifying 21,170 colonoscopy videos representing 8162 hours of colonoscopy video from 50 different colonoscopists. Moreover, matching the videos to patient-level data allowed us to stratify the videos by procedure indication, including colon polyp screening/surveillance, inflammatory bowel

disease, and diagnostic colonoscopy. Finally, we then successfully built an interface to extract multiple still frames at prespecified time intervals surrounding endoscopist-identified findings of interest to facilitate ML algorithm development.

While there has been dramatic progress in the use of ML in medicine, its impact has been most pronounced in gastrointestinal endoscopy. There are now multiple research groups that are reliably reporting the ability to identify neoplasia [3–5, 9], predict histology [7, 8], and ascribe validated inflammatory assessment scores [10, 11]. However, many of the reported outcomes are intermediate rather than representing the outcome of interest. In other words, in inflammatory bowel disease, gastrointestinal bleeding, or even colon cancer screening, the true outcome of interest is not immediately realized within the procedure, but at a future “non-endoscopic” time point. While it is an important step to develop a ML algorithm that detects colorectal polyps, it is more impactful to correlate endoscopic technique with the development of advanced neoplasia at surveillance colonoscopy. Similarly, while it is helpful to have an algorithm assess inflammatory bowel disease activity, it will be important to link the inflammation seen in these videos with patient outcomes such as the effectiveness of medical therapy escalation or the need for colectomy. This research can only be performed when large numbers of colonoscopies are recorded from a diverse set of colonoscopists and subsequently accurately merged with patient-level data.

Development of ML algorithms requires a rich labeled data set of findings of interest. This data set is often extracted through still images and then labeled by the research team. While there is certainly value to this approach, these images often represent the “clearest” representations of these findings. Obtaining still frames surrounding these images of interest may augment the training, but to be feasible, requires automated extraction of these frames from stored video. To facilitate this, our team developed an automated process to extract frames surrounding these images of interest, identified via the “picture-in-picture” events, which can then be quickly labeled by the research team.

While there are many additional applications of merging of endoscopy videos with rich patient-level data, a primary interest is that this could facilitate automatic calculation of endoscopist and center quality metrics including adenoma detection rates, colonoscopy completion rates, and bowel preparation quality using ML techniques. This data could potentially enrich and increase adoption of centralized endoscopy databases which now collect limited data and often require some manual entry. Moreover, this rich “endoscopy health record” combining a video recording, patient characteristics, and immediate/long-term outcomes will greatly enrich endoscopic research in areas as diverse as inflammatory bowel disease and colorectal cancer prevention, with or without the use of adjunctive ML techniques.

While the merger of an endoscopic videos with EHR data has clear value for quality improvement, clinical care, and research, there are potential downsides to consider. First, there are unknown medicolegal considerations; specifically, it is unclear whether a de-identified recorded endoscopic video could po-

tentially expose clinicians to further scrutiny in cases of adverse events or interval malignancy. Furthermore, storage of these endoscopic videos is only feasible using cloud-storage. While these videos are de-identified in the cloud, external storage of the videos requires increasing vigilance about data protection.

There are important limitations to consider. Although we were able to accurately match colonoscopy videos to patients in the EHR, only 74% of all colonoscopies during the study period were matched with a corresponding video. However, this occurred primarily due to the gradual adoption of automated recording during the study period, resulting in a “loss” of some videos. When looking only at procedure rooms where continuous recording during the study period could be confirmed, the match rate was markedly higher at 92%. In addition, although we expect that a similar approach can accurately match other procedure videos (upper endoscopy, endoscopic retrograde cholangiography, and endoscopic ultrasound) to the EHR, this requires further study to confirm. Finally, although we anticipate that such a large video library integrated with detailed patient-level data will facilitate ML research focused on robust outcomes, this confirmatory work is ongoing.

Conclusion

In summary, we have shown that a large cloud-based video library with limited interoperability with the corresponding local EHR can be successfully merged with robust patient-level data. Furthermore, these videos can be utilized to automatically extract additional frames of interest beyond those specifically captured by the endoscopist, potentially accelerating the development of ML techniques. Further work is ongoing to utilize this large video library to explore the relationship between colonoscopy findings and patient outcomes.

Acknowledgments

Funding was provided by the Gordon and Betty Moore Foundation.

ADY is also supported by the National Heart, Lung, and Blood Institute (K08HL145139).

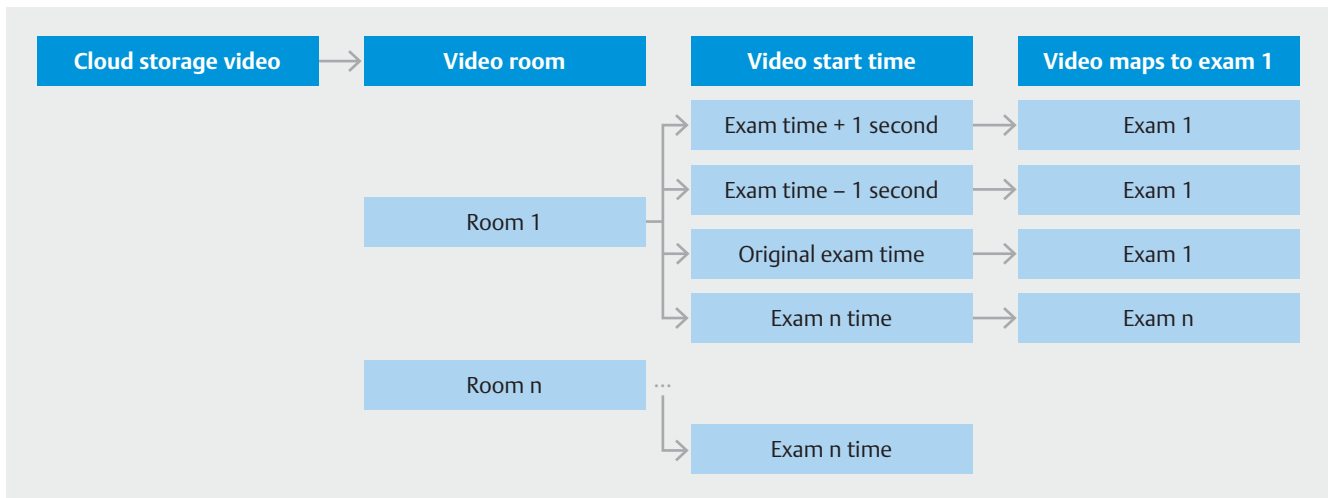
Competing interests

RNK – Consultant, Boston Scientific

References

- [1] Duloy AM, Kaltenbach TR, Wood M et al. Colon polypectomy report card improves polypectomy competency: results of a prospective quality improvement study (with video). *Gastrointest Endosc* 2019; 89: 1212–1221
- [2] Becq A, Chandnani M, Bharadwaj S et al. Effectiveness of a deep-learning polyp detection system in prospectively collected colonoscopy videos with variable bowel preparation quality. *J Clin Gastroenterol* 2019; doi:10.1097/MCG.0000000000001272

- [3] Liu WN, Zhang YY, Bian XQ et al. Study on detection rate of polyps and adenomas in artificial-intelligence-aided colonoscopy. *Saudi J Gastroenterol* 2020; 26: 13–19
- [4] Wang P, Berzin TM, Glissen Brown JR et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study. *Gut* 2019; 68: 1813–1819
- [5] Wang P, Xiao X, Glissen Brown JR et al. Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2018; 2: 741–748
- [6] Yamada M, Saito Y, Imaoka H et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Sci Rep* 2019; 9: 14465
- [7] Shahidi N, Rex DK, Kaltenbach T et al. Use of endoscopic impression, artificial intelligence, and pathologist interpretation to resolve discrepancies between endoscopy and pathology analyses of diminutive colorectal polyps. *Gastroenterology* 2020; 158: 783–785 e781
- [8] Byrne MF, Chapados N, Soudan F et al. Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut* 2019; 68: 94–100
- [9] Kudo SE, Misawa M, Mori Y et al. Artificial fsms. *Clin Gastroenterol Hepatol* 2019: doi:10.1016/j.cgh.2019.09.009
- [10] Stidham RW, Liu W, Bishu S et al. Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis. *JAMA Netw Open* 2019; 2: e193963
- [11] Takenaka K, Ohtsuka K, Fujii T et al. Development and Validation of a deep neural network for accurate evaluation of endoscopic images from patients with ulcerative colitis. *Gastroenterology* 2020: doi:10.1053/j.gastro.2020.02.012



► **Supplemental Figure** Schematic to rapidly match cloud-based colonoscopy videos with cases stored in the electronic health record.