

80269_Auto_Edited-check.docx

Name of Journal: *World Journal of Hepatology*

Manuscript NO: 80269

Manuscript Type: ORIGINAL ARTICLE

Observational Study

Detection of colorectal adenomas by using artificial intelligence models in patients with chronic hepatitis C

Singh Y *et al.* AI tool to detect colorectal adenoma

Abstract

BACKGROUND

Hepatitis C virus is known for its oncogenic potential, especially in hepatocellular carcinoma and non-Hodgkin lymphoma. Several studies have shown that chronic hepatitis C (CHC) has an increased risk of development of colorectal cancer (CRC).

AIM

To analyze this positive relationship and developed an artificial intelligence (AI)-based tool using machine learning (ML) algorithms to stratify these patient populations into risk groups for CRC/adenoma detection.

METHODS

To develop the AI automated calculator, we applied ML to train models to predict the probability and the number of adenomas detected on colonoscopy. Data sets were split into 70:30 ratios for training and internal validation. Scikit-learn standard scaler was used to scale values of continuous variables. Colonoscopy findings were used as the gold standard and deep learning architecture was used to train six ML models for prediction. A Flask (customizable python framework) application programming interface (API) was used to deploy the trained ML model with the highest accuracy as a web application. Finally, Heroku was used for the deployment of the web-based API to <https://adenomadetection.herokuapp.com>.

RESULTS

Out of the 415 patients, 206 patients had colonoscopy results. On internal validation, the Bernoulli naive Bayes model predicted the probability of adenoma detection with the highest accuracy of 56%, precision of 55%, recall of 55%, and F1 measure of 54%. Support vector regressor predicted the number of adenomas with the least mean absolute error of 0.905.

CONCLUSION

Our AI-based tool can help providers stratify patients with CHC for early referral for screening colonoscopy. Along with giving a numerical percentage, the calculator can also comment on the number of adenomatous polyps a gastroenterologist can expect, prompting a higher adenoma detection rate.

Key Words: Machine learning; Calculator; Artificial intelligence; Hepatitis C; Screening

Singh Y, Gogtay M, Yekula A, Soni A, Mishra AK, Tripathi K, Abraham GM. Detection of colorectal adenomas by using artificial intelligence models in patients with chronic hepatitis C. *World J Hepatol* 2022; In press

Core Tip: Hepatitis C is associated with a wide array of extra-hepatic manifestations. In this study, we evaluated the incidence of colorectal adenomas and adenoma detection rates in hepatitis C patients. We developed an artificial intelligence-based tool to guide physicians in the detection and diagnosis of pre-malignant and malignant colorectal pathologies in these patient populations.

INTRODUCTION

Hepatitis C is the most common blood-borne infection worldwide despite being gravely under-diagnosed^[1,2]. Data from the National Health and Nutrition Examination Survey 2013-2016 and four other high-risk populations; homeless, incarcerated, active-duty military, and nursing homes estimates that approximately 4.1 million persons in the United States (approximately 1.7% of the population) were hepatitis C virus (HCV) antibody-positive indicating past exposure and 2.4 million persons (approximately 1% of the population) were HCV RNA positive indicating an active infection^[3].

Although HCV is thought of as a primary disease of the liver, it has a wide array of extrahepatic manifestations, including skin, blood, lymphatic and intestinal pathologies^[4,5]. United States chronic hepatitis cohort study showed an increased

incidence of rectal cancer with increased mortality in chronic HCV, but not colon cancers^[4-6]. Colorectal cancer (CRC) is the third most commonly diagnosed cancer and the second most common cause of cancer-related death worldwide^[3,7].

The current gold standard for diagnosing CRC is colonoscopy. United States Preventive Services Task Force, American College of Gastroenterology, and American Cancer Society recommend screening for CRC starting at age 45 years in average-risk individuals^[8-10]. There are no specific screening guidelines for CRC in patients with hepatitis C, despite there being an increased association.

Importance of artificial intelligence

Artificial intelligence (AI) is the new silk road. It is a technique that allows machines to store large amounts of data coming from various sources, process it accurately, reason, and even simulate human intelligence to provide a plan for clinical treatment. AI has augmented medical research on an enormous scale in recent years and continues to do so. It has significantly helped reduce the workload on the clinician and healthcare staff^[11].

There are multiple facets of AI used in gastroenterology like convolutional neural network (CNN), deep learning (DL), machine learning (ML), and computer-aided design (CAD)^[12]. ML is a subset of AI, DL is a subset of ML and neural networks make up the backbone of DL algorithms. Multiple AI-assisted systems like EndoBRAIN CAD by Kudo *et al*^[13] and a CNN-based auxiliary model by Ding *et al*^[14] have been used to detect and diagnose bowel pathologies including colonic adenomas and neoplasms with higher sensitivity and specificity as compared to trained endoscopic experts^[13,14].

ML is one of the subsets of AI where algorithms are trained for specific tasks. This is being increasingly used in modern medicine for analyzing large volumes and complex data. ML can be applied in any clinical setting to include mathematical and statistical assumptions that are unfamiliar to most clinicians. It can be done in a two-step process. Firstly, the clinical question. Second, which are in the clinical question/process can be optimized. Data management is another aspect of ML that is crucial. Availability of

high-quality data that can train the ML algorithms and avoid data bias and errors that can erroneously skew the results.

The use of AI has shown an increase in both polyp detection rate and adenoma detection rate (ADR) which are the primary colonoscopy quality indicators. Each 1% increase in ADR leads to an approximately 3% decrease in the future risk of cancer^[15]. Since timely intervention in the form of screening colonoscopy can help detect pre-neoplastic or early stages of CRC^[15], it is imperative to develop a tool that can help clinicians distinguish which patients with chronic hepatitis C (CHC) infection require early referral for colonoscopy. Rustagi *et al*^[16] in a case-control study showed a significantly higher number of adenomas detected on screening colonoscopy in the CHC group. In addition, building an ML model which can help predict the expected number of adenomas can help in the reduction of missed adenomas, decreasing subsequent cancer and the overall burden of CRC on healthcare. This is the aim of our research in developing an AI/ML-based tool to help.

MATERIALS AND METHODS

Study design and setting

This observational study with cross-sectional data collection and analysis was conducted at a community hospital in Massachusetts. The institutional review board approved the study. The results were tabulated and statistically analyzed using computer software (SPSS version 25 for Windows, SPSS Inc., Chicago, IL). Descriptive statistics for continuous variables were calculated with the Mann-Whitney *U* test, and categorical variables were calculated with the Chi-square test. The level of significance was set at $P < 0.05$.

Model training: Classification

Several ML models were trained and tested, and the model that could predict percentage probability with the highest accuracy was saved for the deployment stage.

Training and testing dataset characteristics

The dataset used to train and test the ML algorithms were collected manually and stored in a comma-separated file format. The dataset contained several attribute vectors from 415 patients [*i.e.*, sex, age, body mass index (BMI), obesity class, oral contraceptive use, significant alcohol use, hypothyroidism, intravenous drug use, diabetes mellitus, human immunodeficiency virus, concomitant statin use, controlled attenuation parameter (CAP) grade, HCV status, genotype, aspartate aminotransferase, alanine aminotransferase, platelet count, hemoglobin A1c, and triglycerides]. Some patients had missing data, which were replaced by zeros for training and testing. The task was to predict the percentage probability of adenoma occurrence in colonoscopy, followed by calculating the number of polyps. For this purpose, two ML models, *i.e.*, classification and regression models, were trained, and the best models were saved from both ML categories.

Preprocessing

The dataset was loaded onto a pandas DataFrame. The output label was colorectal adenomatous polyps. Numerical feature vectors were replaced with categorical variables. Categorical attributes like HCV genotype and gene polymorphisms underwent label encoding. The variables with numerical values were scaled using the sci-kit-learn StandardScaler function to scale down to the desired range of 0-1. The complete dataset was split into a 70:30 ratio, in which 70% of the total dataset was used for training. The remaining 30% of the entire dataset was used to test internal validity and select the ML model with the highest predictive value.

ML algorithm internal testing

Several ML algorithms were trained to predict hepatic steatosis and CAP grades using the above-listed vectors. These models included: Support vector classifier, random forest, Bernoulli naïve Bayes (BNB), Gradient boosting classifier (GBC), logistic regression, and stochastic gradient descent classifier. All models were trained using

70% of the dataset. After training, each model was tested using the remaining 30% of the data. The model with the highest testing accuracy, the GBC model, was chosen for external validity testing.

GBC model

The GBC model is a set of ML algorithms that additively combines multiple weaker ML models to produce a final predictive model. Our model assigned a binary classification to datasets and used multiple regressions along several decision trees to refine its attempts to predict the steatosis classification correctly. The model graded each attempt on a loss function which evaluates the extent to which the previous tree was inaccurate.

RESULTS

Patients were divided into cases (HCV) and controls (non-HCV). Data was tabulated as described in Table 1. Figure 1 is a strobe diagram with an overview of the methodology and results. The performance of different ML algorithms for training and testing to detect colorectal adenomatous polyps is shown in Table 2.

Using the colonoscopy results determined by the pathology of the biopsied polyp as the output, we applied a DL architecture using the above variables to train and test several ML models using a 415-patient dataset. As shown in Table 2, our results demonstrated that the BNB model had the highest testing accuracy (56%), precision (55%), recall (55%), and F1 score (54%). The distribution of actual and predicted labels during internal testing can be seen in Figure 2. As depicted in Table 3, our results demonstrated that the support vector regressor model had the lowest mean absolute error - 0.905, indicating it was the most suitable ML model to calculate an approximate number of polyps.

Application development phase

A flask-based web app was developed using the model with the highest accuracy for the application phase. A flash ² application programming interface (API) was used to

deploy the trained ML model as a web application. The web-based API was then deployed into a web server using Heroku, a cloud application platform (<https://adenomadetection.herokuapp.com/>).

DISCUSSION

The initial analyses of the variables have demonstrated that median age of 62, with higher BMI, smoking, alcohol use, and concomitant aspirin use are related to adenoma detection with statistical significance among the data collected from patients with CHC that underwent colonoscopies. Using these initial data sets, DL architecture was used to train six ML models, which were prepared and validated using a 70:30 ratio of the data set. Of the multiple models, BNB model predicted the probability of adenoma detection with the highest accuracy, precision, and recall. Flask API was used to deploy the ML model, and Heroku web API was used to develop the AI tool. The model training, performance can be seen in Tables 2-4.

The AI tool may be handy in the clinical setting to triage patients with hepatitis C. They have not received a formal CRC screening to stratify patients into high-risk and low-risk groups. This can guide the gastroenterologist during the colonoscopy to help increase the ADRs. This knowledge of the increased risk of CRC incidence in hepatitis C and isolating patients that are high risk can prompt physicians to start CRC screening early and more frequently in these specific subsets.

Although multiple review articles have been published, there are very few clinical studies. To our knowledge, this is one of the two retrospective analyses performed comparing CRC in hepatitis C and non-hepatitis C individuals. One of the main limiting factors was the study was done in the patient population that visited the primary care clinic at a community hospital in a single city in the northeast United States; it is limited by the geographical and socioeconomic aspects of the patient population. Analysis and validation were done on this limited sample size. Another drawback was that the study was primarily a retrospective analysis. Timing of the colonoscopy and patient having

active hepatitis infection might not always be documented, and timing was not specific in many patients, which limited the sample size that was analyzed.

AI application to healthcare has been increasingly pursued in the past decade, changing how we deliver advanced patient care^[17,18]. AI and ML ensure the automation of many time-consuming activities and provide better insight into patient data that are not evident to providers^[17]. Although the hype behind AI is promising, various barriers prevent the real-world application of these new AI systems. AI has helped in increased interoperability of extensive data, reasonable data-driven decisions involving evidence-based medicine, and hence a higher quality of care.

CRC is a commonly diagnosed malignancy in men and women with increasing mortality^[19]. Many epidemiological studies have shown a positive association of CHC infection with extra-hepatic malignancies, especially gastrointestinal malignancies^[16,19]. Our study focused on this positive association between CHC and CRC and on using AI and ML models to further ease the diagnosis in these patient populations.

AI is integrated into our everyday life to such an extent that we barely remember what life was before it. Face recognition to unlock our phones, self-driven cars, chatbots for almost every business, and even ML-based financial fraud detection. AI-driven models are increasingly utilized to screen, diagnose and monitor multiple clinical conditions. Many AI algorithms have been used in CRC detection. Hu *et al*^[20] performed ML simulations using S-Kohonen, Backpropagation and SVM neural networks, showing the S-Kohonen method's effectiveness for colon cancer classification^[20].

Zhang *et al*^[21] derived a cost-effective and sensitive method for detecting BRAF mutations in CRC using a counter propagation artificial neural network to distinguish mutant BRAF *vs* wild type^[20,21]. Many such ML models are increasingly being utilized in precision oncology to precisely guide diagnosis and management decisions in CRC patients^[19-21].

AI tools will potentially transform our practice by leveraging massive amounts of data to personalize care to the right patient, in the right amount, at the right time. These

novel tools assist physicians in the detection and early diagnosis of pre malignant and malignant lesions in general and high-risk populations.

CONCLUSION

Our AI tool can be further modified based on the treatment of hepatitis C with the new direct-acting antivirals, and how treated and cured hepatitis C alters the incidence of CRC in these groups. Long Term prospective studies, including a subgroup analysis between patients cured of hepatitis C, who had a relapse of the disease, and who refused or were untreated and how it affected the CRC detection, would help guide diagnostics. Further validation with randomized control trials with multicenter participation will ensure replicability and repeatability of the results for the smooth incorporation of such AI-based tools into clinical practice^[19-22].

ARTICLE HIGHLIGHTS

Research background

Several studies have shown that chronic hepatitis C virus (HCV) increases the risk of developing colorectal cancer (CRC). We conducted a study to analyze this positive relationship. We developed an artificial intelligence (AI) based tool using machine learning (ML) algorithms to stratify these patient populations into risk groups for CRC/adenoma detection.

Research motivation

We acknowledge the increased applications of AI with ML in medicine. Gastroenterology and hepatology have immense potential for AI integration. Hence, to develop the AI automated calculator, we applied ML to train models to predict the probability and the number of adenomas detected on colonoscopy.

Research objectives

Our objective was the create a readily available AI tool in the form of a calculator that healthcare providers throughout the globe can access to predict the prevalence of adenoma/CRC.

Research methods

We used colonoscopy findings as the gold standard and applied a deep learning architecture to train ML models for prediction. The institutional review board approved the study.

Research results

Data was collected for 415 patients. We discovered a higher incidence of adenoma/CRC in patients with chronic HCV untreated patient population. On internal validation, Bernoulli naive Bayes ML model showed the highest predictive accuracy and recall for adenoma detection rates.

Research conclusions

Our AI-based tool shows an association between HCV and colorectal adenomas. This tool can help providers stratify their patients at increased risk of CRC and prompt early referral for colonoscopy.

Research perspectives

In the future, we would like to see this calculator being used in clinical practice as a preventative measure to increase early diagnosis of high-risk adenomas/CRC.

4%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|--|-----------------|
| 1 | link.springer.com
<small>Internet</small> | 28 words — 1% |
| <hr/> | | |
| 2 | Stephanie Maldonado, Joshua Grab, Connie W. Wang, Heather Huddleston, Marcelle I. Cedars, Monika Sarkar. "Su1406: POLCYSTIC OVARY SYNDROME (PCOS) INCREASES RISK OF NASH AMONG REPRODUCTIVE-AGED WOMEN WITH NAFLD: A US POPULATION-BASED STUDY", <i>Gastroenterology</i> , 2022
<small>Crossref</small> | 26 words — 1% |
| <hr/> | | |
| 3 | cancerbiomedcentral.com
<small>Internet</small> | 20 words — 1% |
| <hr/> | | |
| 4 | digestivehealth.com
<small>Internet</small> | 16 words — 1% |
| <hr/> | | |
| 5 | Dennis L Shung, Joseph J Y Sung. "Challenges of developing artificial intelligence - assisted tools for clinical medicine", <i>Journal of Gastroenterology and Hepatology</i> , 2021
<small>Crossref</small> | 12 words — < 1% |

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES

EXCLUDE MATCHES

< 12 WORDS

< 12 WORDS