

# Artificial intelligence in the colonoscopy: Improving medical diagnostic of the colorectal cancer

Stefanus Bernard<sup>1</sup>, Arli Aditya Parikesit<sup>1\*</sup>

<sup>1</sup>Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia

Article Info	ABSTRACT
<b>Article type:</b> Review	<b>Introduction:</b> Colorectal cancer (CRC) is a development of abnormal cells either in the colon or rectum. CRC is the 3rd leading cause of death in 2018. It first arises during pre-cancerous stages called polyps. The detection and removal of a polyp are important to increase the survival rate of the patient. Although the various method of polyp detection is available, colonoscopy remains the standard in detection and removal of polyps. Several studies showed how Artificial Intelligence (AI) used in colonoscopy such as in detecting polyps, assessing physicians and predicting patients with a high risk of CRC. This study will describe the involvement of AI in colonoscopy and its role in improving the survival rates of patients with CRC.
<b>Article History:</b> Received: 2019-11-22 Accepted: 2019-12-18 Published: 2020-03-04	<b>Material and Methods:</b> Search for research articles conducted from various resources including PubMed and Google Scholar. The keywords of 'Artificial Intelligence' and 'Colonoscopy' were used. 6 research articles about the use of AI in colonoscopy and were published in the interval time of 2017 – 2019 were selected. Such interval time was chosen due to the recent emergence of AI in colonoscopy.
<b>* Corresponding author:</b> Arli Aditya Parikesit  Department of Bioinformatics, School of Life Sciences, Indonesia International Institute for Life Sciences, Jakarta, Indonesia  Email: arli.parikesit@i3l.ac.id	<b>Results:</b> Studies of AI in colonoscopy showed how it improves medical diagnostic of CRC in several ways, including in improving adenoma detection rate (ADR), finding physicians with a high Adenoma Detection Rate (ADR) and predicting patients with high risk of CRC. However, the use of AI also associated with limitations derived either from the model, datasets or study design.
<b>Keywords:</b> Colorectal Cancer Colonoscopy AI Machine Learning Deep Learning	<b>Conclusion:</b> A Combination of AI and colonoscopy has the potential to improve the diagnostic accuracy and survival rate of patients with CRC. Further study would be required to find the best possible cases for model, datasets and study design in order to overcome the limitations and eventually achieve the best possible results.

## Cite this paper as:

Bernard S, Parikesit AA. Artificial Intelligence in the Colonoscopy: Improving Medical Diagnostic of the Colorectal Cancer. Front Health Inform. 2020; 9: 27. DOI: [10.30699/fhi.v9i1.209](https://doi.org/10.30699/fhi.v9i1.209)

## INTRODUCTION

Colorectal cancer (CRC) or colon cancer is a development of abnormal cells that starts in the colon or rectum. CRC considered the 3rd leading cause of death in 2018 with up to 2 million cases identified with 1 million estimated fatalities worldwide in 2018 [1]. CRC first arises during pre-cancerous stages called polyps. It has a pivotal role in the survival of patients as during this stage it can be removed. Therefore, the detection of polyps in patients with CRC is important to increase the survival rate of patients. Currently available polyp detection or screening available including colonoscopy, sigmoidoscopy and fecal occult blood

testing [2]. Colonoscopy considered as the standard screening test to identify and remove the polyp particularly adenomas. Despite the fact by utilizing colonoscopies and removal of polyps will prevent the CRCs, 7-9% of CRCs still occur and eventually create what is called 'interval cancers'. Interval cancers happened due to the polyps that aren't completely removed during the colonoscopy and therefore causing cancer to reoccur although the patient conducted regular colonoscopy and removal of polyps [3].

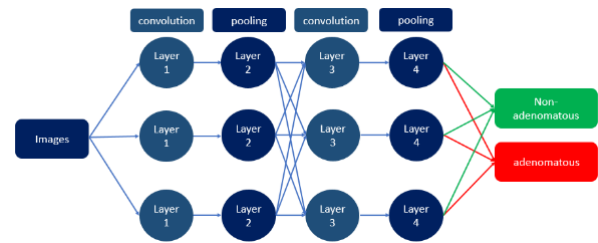
Current development of Artificial Intelligence (AI) creates abundant opportunities in many fields including in medical diagnostics of various diseases

including cancers. AI as the subfield of computer science has an aim to develop systems with advanced predictive or analytical capabilities [4]. Machine learning is the most known and successful branch of AI with a solid history of applications in medical diagnostics of various diseases. Machine learning's main focus is on how computers learn from data. It derived from both statistics and computer science intending to develop efficient computing algorithms and learn relationships from data [5]. Types of learning divided into two: supervised learning and unsupervised learning. Supervised learning predicts the output based on the datasets. Compared to supervised learning, unsupervised learning has no outputs to predict and instead users try to find any patterns within the data [5]. In precision oncology, machine learning currently being applied to a variety of diagnostic procedures, prognostic and other tasks that can be predicted based upon the data.

Deep learning is a sub-branch of machine learning consists of wide computational models of data processing layers for feature extraction and pattern recognition [4]. The application of deep learning is extended in many areas in science and technology such as natural language processing, speech recognition, computer vision and biology [3]. The methods used in deep learning allow the machine to scan large quantities of data and discover the patterns required for classification [6]. Furthermore, it able to create the abstraction model derived from the data by implementing multi-layered deep neural networks (DNNs) [7]. The development of deep learning field create variations in the models created such as Fully Convolutional Neural Network (FCNN), Convolutional Neural Network (CNN), Deep Learning System (DLS) and Artificial Neural Networks (ANN) [8].

Standard neural network application in deep learning consisting of simply connected nodes called neurons. Each neuron generates a sequence of real-valued activations. Input neurons activated through sensors while others activated through connections from previously active neurons [9]. The most established algorithm among the models is CNN that is a known method in computer vision tasks [10]. Therefore, the usage of CNN in medical diagnostic oftentimes related to real-time analysis methods such as in colonoscopy. The model of CNN used in real-time colonoscopy may vary among the study that is previously has been conducted. The model used depends on the objective of the study and available computational resources. In a real-time colonoscopy, CNN obtains an image as an input and the output is the labels corresponding to the

images (Fig 1). Convolution and pooling are applied to process the information. Convolution works by extracting features from images in the previous layer while pooling selects the strongest activated value for a feature that is extracted in convolution. In the final layer, the number of neurons is equivalent to the number of labels to be recognized and represented as a probability percentage [11].



**Fig 1: Application of convolutional neural network (CNN) in detecting precancerous polyps in CRC**

The study of AI in improving the medical diagnostic of diseases has previously been conducted. As an example in the study by Parikesit et al. showing various types of deep learning, their advantages, disadvantages and most importantly its role in improving the medical diagnosis of breast cancer. The main goal of this study is to discover the types and basic principles of AI algorithms used in colonoscopy and how it able to improve the medical diagnostic of CRC. Therefore, research articles related to AI in colonoscopy were collected from various resources including PubMed and Google scholar and were compared to observe the advantages, limitations, and prospects of using AI in colonoscopy.

## MATERIAL AND METHODS

Search for research articles conducted with keywords 'Artificial Intelligence' and 'Colonoscopy'. 6 research articles selected from resources including NCBI PubMed portal and Google scholar. Research articles selected were published in the interval time of 2017 - 2019 due to emerging of AI during the selected time range brought a various impact in biomedical research including in colonoscopy.

## RESULTS

Different kinds of AI identified in the 6 selected research articles are as follows: two of whom implements natural language processing and machine learning while the other four implement convolutional neural networks (Table 1). All of the identified AI approach and their role in colonoscopy will be discussed in this study.

**Table 1: Review summary**

References	Algorithm	Datasets	Application	Accuracy	Sensitivity	Description
Hilsden et al., 2018	ML	Colonoscopies and associated pathology reports	Predict the presence of polyps at colonoscopy	N/A	N/A	Model identify individuals with high risk of having CRC at specificity of 95%
Mehrotra et al., 2018	NLP	Colonoscopies and associated pathology reports	Describe physicians associated with high ADR	87–99%	N/A	NLP determined female physicians and well trained physicians with higher ADR rate
Urban et al., 2018	CNN	8,641 hand-labelled images from 2,000 patients	Identification of adenomas or polyps	96.4%	96.9% (5% FPR) and 88.1% (1% FPR)	Pre-trained models identified all polyps and additional polyps missed during review
Komeda et al., 2019	CNN	1,200 colonoscopy images	Identification of adenomas or polyps	70%	N/A	Decision by model correct in every 7 out of 10 cases
Byrne et al., 2017	CNN	Colonoscopy video	Identification of adenomas or polyps	94%	98%	Accuracy of the model was 94 % with sensitivity of 98 % and specificity of 83 %
Wang et al., 2019	CNN	Colonoscopy video	Identification of adenomas or polyps	N/A	94.38%	CAD system increase ADR to 29.1 % and detect more polyps compared to standard colonoscopy

ML: Machine Learning

NLP: Natural Language Processing

CNN: Convolutional Neural Network

ADR: Adenoma Detection Rate

CAD: Computer-Aided Diagnosis

N/A: Not Available/not mentioned by Authors

## DISCUSSION

### ColonFlag machine learning algorithm

The study by Hilsden et al. represents the use of ColonFlag to predict the adenomatous polyps present at colonoscopy. ColonFlag principles based on machine learning and use basic medical information and complete blood cell counts (CBC) to identify individuals with an elevated risk of having CRC [12]. Additional data elements obtained including age, gender, date of procedure, indication, depth of endoscope insertion, bowel preparation quality and unique lifetime identifier.

ColonFlag model was able to identify individuals with risk of having CRC based on the data of routinely collected CBC and data such as patient's age, gender, and colonoscopy. Besides, as stated by the authors the use of ColonFlag could identify unscreened individuals at higher risk for CRC and therefore individuals may be targeted to achieve greater compliance with conventional screening tests [12].

### Natural language processing evaluate physician characteristics

Natural Language Processing (NLP) is the subfield of AI that enables computers to extract and study the human language. It is composed of many techniques grouped together and began in the 1950s as the intersection between AI and linguistics [13]. NLP was originally distinct from text information retrieval that implements statistics techniques to searches and index large volumes of text [14]. Currently, NLP is influenced by data-driven approaches due to the ambiguity, large size and unrestrictive nature of natural language. This approaches combining both machine learning and Hidden Markov Models (HMMs). Statistical analysis and machine learning allow the program to scan for patterns and make predictions through the numerical measure and iterative process. While HMMs allow variables to switch between several states and eventually generating several possible output symbols. The nature of HMMs set possible conditions and generate unique symbols may be large, however, it is known and finite [14].

The study conducted by Mehrotra et al. represents the use of NLP to analyze the colonoscopy examinations and pathology reports with the aim to evaluate the physicians performing colonoscopy examinations based on the ADR. NLP was used to extract the relevant data from both pathology and colonoscopy reports. NLP system also validates a sample of 2,127 colonoscopy and associated pathology reports. The data were analyzed both by NLP and manually abstracted. The study identified considerable variation in ADR among physicians. 56% of physicians met the ADR targets. Further analysis by logistic regression shows that adenoma has a higher probability to be identified by female endoscopists who were trained in gastroenterology and under 9 years of practice [15].

### Deep learning convolutional neural network

A study by Urban et al. implement the use of CNN architectures with the aim to increase the ADR. Different CNN models used in the study including the models previously trained to identify pictures and the models not previously trained. Datasets involved in this study including the 8,641 selected colonoscopy images obtained from 2,000 patients, another 1,330 colonoscopy images obtained from different patients and 9 colonoscopy videos. Cross-validation conducted by training the model on one dataset and tested it on another completely different dataset. The result shows the models that are previously trained using polyp and random images able to detect the polyps with an accuracy of 96.4 % and sensitivity of 96.9% (5% FPR) and 88.1% (1% FPR). In addition, the model identified all polyps found by experts and polyps missed during the review [3].

Another study by Komeda et al. uses the CNN system in CAD of colon polyps. The datasets used in the study are 1,200 images of colonoscopy. Images were extracted from videos of actual endoscopic examinations. The study shows the CNN-CAD system able to distinguish between non-adenomatous from adenomatous polyps with accuracy up to 70 % (decision by CNN-CAD correct in 7 out of 10 cases). The study by Byrne et al. also applied the CNN system for real-time assessment of endoscopic video images of colorectal polyps. Narrowband imaging video frames and unaltered videos from routine exams used in model training and validation. The CNN system used in the study able to identify polyps with an accuracy of 94 %, the sensitivity of 98 % and specificity of 83 % all with 95 % CI (confidence interval) [11,16].

While the study by Wang et al. uses the CNN system intending to investigate the effect of ADR by using an automatic polyp detection system. 1,130 patients with eligibility criteria involved in the study. At the end of the study, they found the CAD colonoscopy

able to significantly increase the ADR to 29.1% compared to the standard colonoscopy ADR of 20.3% ( $p < 0.001$ ). CAD colonoscopy found 185 diminutive adenomas and 114 hyperplastic polyps. These discoveries are significantly higher compared to 102 adenomas and 52 hyperplastic polyps discovered using standard colonoscopy [17].

### Limitations

Introduction of AI in medical diagnostic of CRC specifically to colonoscopy able to improve several aspects including the ADR and identification of patients with an elevated risk of CRC. However, the utilization of AI in colonoscopy brought several limitations. Some of them either derived from the algorithm, dataset used or study design (Table 2). Hilsden et al. in their study of machine learning algorithm stated that the limitation in their study came from the dataset as further detailed information was required such as medical history of patients and his/her family regarding the CRC and besides, complete characterization of all polyps also required as the accuracy of the output depends on such data [12]. While the use of NLP in assessing the physicians performing colonoscopy by Mehrotra et al. didn't mention any limitations of NLP associated in the study. Instead, they stated the limitation of the study design and the physicians used in the study cannot stand as the representative of physicians who conducting colonoscopies outside the study [15]. Deep learning holds a pivotal role in colonoscopy. It is often associated with image classification and pattern recognition as observed in the study by Urban et al., Komeda et al., Byrne et al. and Wang et al. Urban et al. in their study able to create a CNN model with accuracy up to 96.4% and the highest sensitivity of 96.9% with 5% FPR. However, the performance may vary and depends either in surveillance or screening. The unknown effects of CNN on inspection behavior by colonoscopists is one of the limitations identified although a study suggests the use CNN during live colonoscopy reduce the number of missed polyps. Another limitation came from the de-identified nature videos which it exclude the information about histology of polyps. Polyp histology is relevant for added time and pathology costs [3]. While Komeda et al. identified the limitation of their CNN-CAD system came from the dataset. The model was able to differentiate adenomatous and nonadenomatous polyps with an accuracy of 70%. They stated the diagnostic accuracy of 70% is unsatisfactory and authors also assume that deep learning with 1,200 images is not sufficient and might be contributing to the low number of diagnostic accuracy [11]. Byrne et al. identified the limitation in their study came from the use of video recordings. The use of it rather than real-time assessments causing the model to not develop at



least 50% confidence in the diagnosis and therefore eventually generate low confidence determination by model and low confidence interpretation by colonoscopists [16]. While in the study by Wang et al. identified limitation more to the study design. Their study uses an open and non-blinded trial as patients were randomly selected whether they are assigned to standard colonoscopy or colonoscopy with the use of AI. The subsequent result shows that the exact contribution of the system is difficult to assess and they suggest to conduct double-blind studies in the future to investigate the exact contribution of the system [17].

limitations associated with using AI in colonoscopy may vary depending on the aim and the study design. However, in most cases, the use of AI often related to the dataset which is required to produce an accurate output [11]. Furthermore, additional data might be needed to support the output. Unfortunately, in some cases, the data either not accessible or available due to some reasons. As an example in the study by Hilsden et al. the database used didn't have the complete characterization of polyps before 2013 and therefore, several high-risk polyps were misclassified as a non-high risk [12].

**Table 2: Limitation of Each Study**

References	Algorithm	Application	Limitations
Hilsden et al., 2018	ML	Predict the presence of polyps at colonoscopy	Require detailed information on person's history of CRC and complete characterization of sessile serrated polyps
Mehrotra et al., 2018	NLP	Describe physicians associated with high ADR	Physicians involved in the study cannot stand as representative of physicians outside the study
Urban et al., 2018	CNN	Identification of adenomas or polyps	CNN performance may vary by indication (screening vs surveillance)
Komeda et al., 2019	CNN	Identification of adenomas or polyps	Diagnostic accuracy for CNN-CAD system is considered unsatisfactory
Byrne et al., 2017	CNN	Identification of adenomas or polyps	low confidence determination due to the use of video recordings
Wang et al., 2019	CNN	Identification of adenomas or polyps	Exact contribution of the system is difficult to assess

### Future of AI in colonoscopy

Colonoscopy procedure was done in real-time and therefore, it is reasonable that the current development of AI algorithms more to assist the physicians in detecting polyps to ensure all polyps are completely removed and hence, CRC will not reoccurs. To help assist the physicians in conducting colonoscopy, deep learning and particularly CNN shows its advantages by detecting polyps missed by physicians and improving ADR. All of who improve the medical diagnostic of CRC although associated with limitations in terms of data quantities, computational resources, and output accuracy. Furthermore, this study showed AI not only able to improve the medical diagnostic of CRC based on real-time colonoscopy. The use of NLP in assessing physicians associated with high ADR and machine learning to predict the patient with a high risk of CRC proves the AI also able to improve medical diagnostic of CRC based on non-real time colonoscopy.

In the future, AI has abundant opportunities in improving the medical diagnostic of CRC, especially in colonoscopy. Instead of using deep learning to identify whether there is a polyp or not in an image, further advance approach such as image classification can be used to categorize the type of

polyps detected during colonoscopy. However, it would require further study specifically in image classification pipeline and how to fine-tuning multiple parameters in order to achieve such a task [18].

### CONCLUSION

A Combination of AI and colonoscopy has the potential to improve the diagnostic accuracy and survival rate of patients with CRC. As shown in this study, the utilization of AI able to improve the ADR and reduce the number of missed polyps during a colonoscopy. However, based on the previously conducted study, some limitations either derived from the model itself, datasets or even the study design used. Therefore, further study needs to be done in order to find the best possible cases in obtaining high ADR or reducing the number of missed polyps. Subsequent analysis can be done through careful assessments of the three main pivotal points which are the model, datasets and study design. In terms of the model, careful observation of how the composition or layer inside the model created a need to be conducted in order to find the most efficient yet fast algorithm. Datasets are known as the backbone in image classification and pattern recognition. It is used to train the model and therefore the model can differentiate between polyps and non-polyps. Large datasets not always

associated with high accuracy and therefore, assessment of the content inside datasets should be done prior to research. Finally, choosing the best study design that is suitable for certain cases would also be required in order to achieve good results.

## ACKNOWLEDGEMENT

Many thanks go to Institute for Research and Community Empowerment (LPPM), Indonesia International Institute for Life Sciences for their heartfelt support toward this research. Thanks also go to Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia for providing Hibah Penelitian Dasar DIKTI/LLDIKTI III 2019 No. 1/AKM/PNT/2019.

## AUTHOR'S CONTRIBUTION

SB worked on the technical and conceptual making of the manuscript, while AA supervised the whole process. The authors agree on this final form of the manuscript, and attested that all authors contributed in the final draft of the manuscript.

## CONFLICTS OF INTEREST

The authors declare no conflicts of interest regarding the publication of this study.

## FINANCIAL DISCLOSURE

This study was funded by Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jenderal Penguatan Riset dan Pengembangan Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia dengan bantuan Hibah Penelitian Dasar DIKTI/LLDIKTI III No.1/AKM/PNT/2019.

## REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018; 68(6): 394-424. PMID: 30207593 DOI: 10.3322/caac.21492 [[PubMed](#)]
2. Vatandoost N, Ghanbari J, Mojaver M, Avan A, Ghayour-Mobarhan M, Nedaeinia R, et al. Early detection of colorectal cancer: From conventional methods to novel biomarkers. *J Cancer Res Clin Oncol.* 2016; 142(2): 341-51. PMID: 25687380 DOI: 10.1007/s00432-015-1928-z [[PubMed](#)]
3. Urban G, Tripathi P, Alkayali T, Mittal M, Jalali F, Karnes W, et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology.* 2018; 155(4): 1069-78. PMID: 29928897 DOI: 10.1053/j.gastro.2018.06.037 [[PubMed](#)]
4. Azuaje F. Artificial intelligence for precision oncology: Beyond patient stratification. *NPJ Precision Oncology.* 2019; 3: 6.
5. Deo RC. Machine learning in medicine. *Circulation* 2015; 132: 1920-30.
6. Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. *J Glob Health.* 2018 Dec; 8(2): 1-8. PMID: 30405904 DOI: 10.7189/jogh.08.020303 [[PubMed](#)]
7. Cao C, Liu F, Tan H, Song D, Shu W, Li W, et al. Deep learning and its applications in biomedicine. *Genomics, Proteomics & Bioinformatics.* 2018; 16: 17-32.
8. Parikesit AA, Ramanto KN. The usage of deep learning algorithm in medical diagnostic of breast cancer. *Malaysian Journal of Fundamental and Applied Sciences.* 2019; 15(2): 274-81.
9. Schmidhuber J. Deep learning in neural networks: An overview. *Neural Networks.* 2015; 61: 85-117.
10. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: An overview and application in radiology. *Insights into Imaging.* 2018; 9: 611-29.
11. Komeda Y, Handa H, Matsui R, Kashida H, Watanabe T, Sakurai T, et al. Computer-aided diagnosis (CAD) based on convolutional neural network (CNN) system using artificial intelligence (AI) for colorectal polyp classification. *Oncology.* 2017; 93(Suppl 1): 30-4. PMID: 29258081 DOI: 10.1159/000481227 [[PubMed](#)]
12. Hilsden RJ, Heitman SJ, Mizrahi B, Narod SA, Goshen R. Prediction of findings at screening colonoscopy using a machine learning algorithm based on complete blood counts (ColonFlag). *PLoS One.* 2018; 13(11): 1-9. PMID: 30481208 DOI: 10.1371/journal.pone.0207848 [[PubMed](#)]
13. Chary M, Parikh S, Manini A, Boyer E, Radeous M. A review of natural language processing in medical education. *West J Emerg Med.* 2019 Jan; 20(1): 78-86. PMID: 30643605 DOI: 10.5811/westjem.2018.11.39725 [[PubMed](#)]
14. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: An introduction. *J Am Med Inform Assoc.* 2011; 18(5): 544-51. PMID: 21846786 DOI: 10.1136/amiajnl-2011-000464 [[PubMed](#)]
15. Mehrotra A, Morris M, Gourevitch RA, Carrell DS, Leffler DA, Rose S, et al. Physician characteristics associated with higher adenoma detection rate. *Gastrointest Endosc.* 2018; 87(3): 778-786. PMID: 28866456 DOI: 10.1016/j.gie.2017.08.023 [[PubMed](#)]
16. Byrne MF, Chapados N, Soudan F, Oertel C, Pérez ML, Kelly R, 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(1): 94-100. PMID: 29066576 DOI: 10.1136/gutjnl-2017-314547 [[PubMed](#)]
17. Wang P, Berzin TM, Brown JRG, Bharadwaj S, Becq A, Xiao X, et al. Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: A prospective randomised controlled study. *Gut*. 2019; 68(10): 1813-9. PMID: 30814121 DOI: 10.1136/gutjnl-2018-317500 [[PubMed](#)]
18. Zhang R, Zheng Y, Mak TWC, Yu R, Wong SH, Lau JYW, et al. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. *IEEE Journal of Biomedical and Health Informatics*. 2017; 21(1): 41-7.