NERTHUS: A Bowel Preparation Quality Video Dataset

Konstantin Pogorelov
Simula Research Laboratory, Norway
University of Oslo, Norway

Kristin Ranheim Randel
Cancer Registry of Norway
University of Oslo, Norway

Thomas de Lange
Bærum Hospital, Norway
Cancer Registry of Norway

Sigrun Losada Eskeland
Bærum Hospital, Norway

Carsten Griwodz
Simula Research Laboratory, Norway
University of Oslo, Norway

Dag Johansen
UiT-The Arctic University of Norway

Concetto Spampinato
University of Catania, Italy

Mario Taschwer
University of Klagenfurt, Austria

Mathias Lux
University of Klagenfurt, Austria

Peter Thelin Schmidt
Karolinska Institutet, Solna, Sweden
Karolinska Hospital, Sweden

Michael Riegler
Simula Research Laboratory, Norway
University of Oslo, Norway

Pål Halvorsen
Simula Research Laboratory, Norway
University of Oslo, Norway

ABSTRACT

Bowel preparation (cleansing) is considered to be a key precondition for successful colonoscopy (endoscopic examination of the bowel). The degree of bowel cleansing directly affects the possibility to detect diseases and may influence decisions on screening and follow-up examination intervals. An accurate assessment of bowel preparation quality is therefore important. Despite the use of reliable and validated bowel preparation scales, the grading may vary from one doctor to another. An objective and automated assessment of bowel cleansing would contribute to reduce such inequalities and optimize use of medical resources. This would also be a valuable feature for automatic endoscopy reporting in the future. In this paper, we present NERTHUS, a dataset containing videos from inside the gastrointestinal (GI) tract, showing different degrees of bowel cleansing. By providing this dataset, we invite multimedia researchers to contribute in the medical field by making systems automatically evaluate the quality of bowel cleansing for colonoscopy. Such innovations would probably contribute to improve the medical field of GI endoscopy.

ACM Reference format:

Konstantin Pogorelov, Kristin Ranheim Randel, Thomas de Lange, Sigrun Losada Eskeland, Carsten Griwodz, Dag Johansen, Concetto Spampinato, Mario Taschwer, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. 2017. NERTHUS: A Bowel Preparation Quality Video Dataset. In *Proceedings of MMSys '17, Taipei, Taiwan, June 20–23, 2017*, 5 pages. <https://doi.org/http://dx.doi.org/10.1145/3083187.3083216>

This work is founded by the Norwegian FRINATEK project "EONS" (#231687). Contact author: Konstantin Pogorelov, email: konstantin@simula.no. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). *MMSys '17, June 20–23, 2017, Taipei, Taiwan* © 2017 Copyright held by the owner/author(s). ACM ISBN 978-1-4503-5002-0/17/06. <https://doi.org/http://dx.doi.org/10.1145/3083187.3083216>

1 INTRODUCTION

The large bowel (figure 1), also named colon and large intestine, is the lower part of the human gastrointestinal (GI) tract. It may be affected by severe diseases including cancer and chronic inflammations. Bowel cancer (colorectal cancer) is currently the third most common cancer worldwide, accounting for nearly 1.4 million new cases and 700 000 cancer deaths in 2012 [9, 27]. The current gold standard for diagnostic and screening investigations of the large bowel is *colonoscopy*. This is a real-time video examination of the inside of the large bowel by use of a digital high definition endoscope. Such endoscopic examinations are resource demanding and require both expensive technical equipment and trained personnel.

Figure 1: The large bowel (image: shutterstock.com).

Furthermore, the efficiency of colonoscopy depends on sufficient bowel cleansing to visualize the gastric mucosa (a membrane that lines the GI tract), achieved by use of oral laxatives (substances that loosen stools and increase bowel movements) administered prior to the procedure. The quality of bowel preparation has shown to influence both the colonoscopy completion rate and detection of possible precursors of cancer (e.g., adenomas, which are the benign tumor of epithelial tissue) [10, 26]. Adenoma detection rate (ADR), that is inversely associated with a patient's risk of developing colorectal cancer, has been proven to be dependent on quality of bowel preparation [7, 13]. Therefore, the degree of bowel preparation is considered to be a reliable quality measure for colonoscopy [16].

Quality of bowel preparation may also influence decisions on screening and follow-up intervals, since low-quality bowel preparation requires repeated colonoscopy [6]. An accurate description of the bowel cleanliness is therefore needed. Despite the use of reliable and validated bowel preparation scales, the grading may vary from

one doctor to another. An objective and automated assessment of bowel cleansing may contribute to reduce such inequalities and optimize use of medical resources. Since endoscopic examinations are real-time investigations, both normal and abnormal findings have to be recorded and documented within written reports. Thus, automatic report generation will probably contribute to reduce doctors' time required for paperwork and thereby increase time to patient care. To our knowledge, a standardized and automatic reporting system that ensures high quality endoscopy reports does not exist. Assessment of bowel cleanliness would be a valuable feature for such automatic endoscopy reporting in the future.

In this area of research, people start to see the synergies between multimedia and medical systems [22]. The development of real-time classification systems is a perfect match in the intersection of medicine, multimedia systems and image/video retrieval. Prototypes like the EIR [23] system targeting analysis of medical videos for detection of abnormalities may be an initial starting point. However, such systems require a lot of data for development, training and testing. To the best of our knowledge, no medical dataset exists for this type of data. In this paper, we therefore present the NERTHUS dataset¹. It contains 21 videos with a total number of 5,525 frames annotated and verified by medical doctors (experienced endoscopists). The videos are divided into four classes of predefined bowel-preparation qualities.

Our initial experiments indicate potential for improvement. In many cases, we are able to detect the annotated bowel cleansing quality. However, to deliver an automatic and reliable system for the endoscopy units, more work is needed. By providing the NERTHUS dataset, we invite researchers to contribute in order to improve important systems for automatic assessment and reporting for GI endoscopy.

2 BOWEL PREPARATION QUALITY

Traditionally, the bowel preparation quality has been categorized as *poor*, *adequate* or *good*. Such classification of bowel cleanliness often lacks clear definitions, and the judgement on quality tends to be subjective. This may result in significant inter-observer variation. In addition, such a traditional categorization relies on a global assessment of bowel cleanliness, which does not account for differences in cleansing between bowel segments. Poor quality of preparation in one segment may then result in low overall grading, despite an otherwise perfectly cleaned bowel. To minimize the inter-endoscopist variation, new score-based methods of assessing bowel cleanliness have been introduced during the last decade.

State-of-the-art scoring systems include the BBPS [3, 15] and the Ottawa bowel preparation scale (OBPS) [24]. Both these scales divide the bowel into three sections (right, middle and left) and score the bowel cleansing within each section according to a defined numeric scale. OBPS uses segmental scores ranges from 0 to 4 in addition to a global three-score fluid-quantity rating, which requires estimation of residual liquid. In contrast, the Boston bowel preparation scale (BBPS) [3, 15] uses only a four-point scoring system (ranges from 0 to 3). Figure 2 illustrates the segmental division of the large bowel used for bowel preparation assessment according to BBPS and OBPS.

Figure 2: The segmental division of the large bowel: the right side includes the cecum and ascending colon, the transverse section of the colon includes the hepatic and splenic flexures, and the left side includes the descending colon, sigmoid colon, and rectum (image: WikiJournal of Medicine [2]).

In this paper, we use BBPS as this is probably best validated and most frequently used scoring system in both routine clinic and screening settings today [19]. The BBPS scale is tested and validated to assess the cleanliness at withdrawal. It does not take into account whether the endoscopist has performed any additional cleansing maneuvers, which reflects the actual practice of colonoscopy. The definition of the BBPS segmental scores are described in table 1. The segmental scores ranges from 0 to 3, where 0 is worse and 3 is the best quality of the bowel preparation [3, 15]. Examples for the different categories are shown in figure 4.

In real colonoscopy examinations, a segmental score is applied [15] to each of the three bowel segments and summed in a total score ranging from 0 to 9. In the NERTHUS dataset, however, all videos

Table 1: The score points used in the Boston bowel preparation scale (BBPS) [3, 15] to define the degrees of cleanliness.

Score	Description
0	Unprepared colon segment with mucosa not seen because of solid stool that cannot be cleared
1	Portion of mucosa of the colon segment seen, but other areas of the colon segment are not well seen because of staining, residual stool, and/or opaque liquid
2	Minor amount of residual staining, small fragments of stool, and/or opaque liquid, but mucosa of colon segment is seen well
3	Entire mucosa of colon segment seen well, with no residual staining, small fragments of stool, or opaque liquid

¹<http://datasets.simula.no/nerthus>

are recorded in the left part of the bowel. Automatic detection of scope position or total score calculation is thereby irrelevant for this dataset and only quality of bowel preparation by segmental scores are of value here. For the future development of automated systems, detecting position and assessment of total BBPS score will be of interest.

3 DATA COLLECTION

The data is collected using equipment as shown in figure 3 at Bærum hospital, Vestre Viken Hospital Trust in Norway. Furthermore, the videos are annotated by one or more medical experts from the Cancer Registry of Norway. A selection of the videos will in addition be annotated by several medical experts from Norway, Sweden, UK, US and Canada through a web based test. These video clips will be marked as the gold standard within dataset and will be released as an addition to the NERTHUS dataset with higher quality regarding bowel preparation assessment.

Figure 3: A colonoscope.

4 DATASET DETAILS

The NERTHUS dataset consists of 21 videos with a total number of 5,525 frames, annotated and verified by medical doctors (experienced endoscopists), including 4 classes showing four-score BBPS-defined bowel-preparation quality (see figure 4). The number of videos per class varies from 1 to 10. The number of frames per class varies from 500 to 2,700. The number of videos and frames is sufficient to be used for different tasks, e.g., image retrieval, machine learning, deep learning and transfer learning, etc. [1, 8, 21]. The dataset consists of videos with resolution 720x576 and is organized by sorting the videos into separate folders named according to their BBPS-bowel preparation quality score. Most of the included videos and images have a green picture in each frame, illustrating the position and configuration of the endoscope inside the bowel. This is obtained by use of an electromagnetic imaging system (ScopeGuide, Olympus Europe) and may support the interpretation of the image. This type of information may be important for later investigations on segmental position within the bowel, but must be handled with care for the bowel preparation quality assessment.

5 SUGGESTED METRICS

When reviewing related work in medical and the computer science areas, there are a lot of different metrics used, with potentially different names when used in the medical area and the computer science (information retrieval) area. For future research, in addition to describing the dataset with respect to total number of images, total number of images in each class and total number of positives, etc., it might be good to provide as many of the *recall* (also known as sensitivity and probability of detection, REC), *precision* (also known as the positive predictive value, PREC), *specificity* (also known as the true negative rate, SPEC), *accuracy* (ACC), *Matthews correlation coefficient* (MCC) and the *F1 score* (F1) metrics [20] as possible in

order to enable a comparison with other work. In addition to the above metrics, processing speed and resource consumption are of interest. In our work, we have used the achieved frame-rate (FPS) as a metric as real-time feedback is important.

6 BASELINE PERFORMANCE

We have performed an initial multi-class detection experiment on NERTHUS as a baseline for future experiments. We have experimented using various configurations of three different main approaches: classification using global features (GF), deep learning convolutional neural networks (CNN) and transfer learning in deep learning (TFL).

For the GF approaches, we extracted several image features for classification using the latest version of the Lire open source software [17]. The extracted features are JCD, Tamura, Color Layout, Edge Histogram, Auto Color Correlogram and Pyramid Histogram of Oriented Gradients. For a first GF run (2 GF), we combined JCD and Tamura, resulting in a feature vector of 187. For a second GF run (6 GF), we combined all extracted features, resulting in a feature vector of 1186 values. The decision of these combinations was based on our previous findings and experiments in [23]. Subsequently, we performed a simple early fusion of the features. All extracted features are included in the dataset in the *arff* file format for reuse and reproducibility. We used the Random Forrest (RF) and Logistic Model Tree (LMT) classifiers provided in the Weka machine learning library [12].

For all deep learning implementations, we used Keras [5] with Google Tensorflow [1] as backend. For the two CNN runs, we trained two different CNNs from scratch, i.e., one with three convolution layers and one with six. As activation function, we used the rectified linear unit (ReLU) [11], and for pooling, we used max-pooling. In all layers, we also included a 0.5 dropout, and the final classification step was performed using two dense layers with first ReLU and then Sigmoid as activation functions. The networks were trained for 200 epochs using the Adam optimizer [14].

The TFL run is based on transfer learning [4] by re-training and fine-tuning the pre-trained Inception v3 model [25]. For the re-training, we followed a similar approach to the one presented in [8]. Firstly, we locked all the basic convolutional layers of the network and only retrained the two top dense classification layers. The dense layers were retrained for 100 epochs using the RMSprop optimizer that allows an adaptive learning rate during the training process. This was done for 100 epochs. After that, fine-tuning of a subset of the convolutional layers was performed. We decided to apply the fine-tuning on the two top convolutional layers of the re-trained model. For this training step, we used the SGD optimizer with a low learning rate (to achieve the best effect in terms of speed and accuracy) [18].

The exact configurations of the CNN and TFL approaches are included in the dataset. We did not perform any data augmentation, such as cropping, for any of the approaches for this work. For the experiments, we first split all the dataset videos into non-overlapping five-second-long video segments resulting in a new set of videos consisting of 52 files. Next, we split the new set randomly into two training and testing subsets, sized as equally as possible depending

(a) BBPS 0 (from splenic flexure)

(b) BBPS 1 (from descending colon)

(c) BBPS 2 (from sigmoid colon)

(d) BBPS 3 (from rectum)

Figure 4: Sample images for each bowel preparation ("cleanliness") score according to BBPS.

on the number of videos available for each class. The subsets contain 27 and 25 videos each. Then, we extracted all the frames from the videos, which gave us two new subsets containing 2800 and 2725 frames each, and applied our algorithms to the new frame sets. We also performed two-folded cross-validation by switching the training and testing sets and calculated the average. As baselines, we provide results of two runs: the RF GF-based classifier with the JCD feature and the random/majority-based classifier.

Table 2 gives an overview of the results, and table 3 contains the confusion matrix for the best performing approach (6 GF with LMT) for a more detailed insight into the performance. We observe that all approaches would outperform the random and majority class baseline, which is presented in the last row. Our own baseline in the second last row is only outperformed by two 6 GF-based approaches

Table 2: Classification performance in terms of weighted average (2-folded) using the metrics described above.

Method	PREC	REC	SPEC	ACC	MCC	F1	FPS
6 Layer CNN	0.856	0.852	0.952	0.854	0.805	0.854	42
3 Layer CNN	0.811	0.694	0.937	0.772	0.621	0.742	40
Inception v3 TFL	0.751	0.745	0.918	0.748	0.665	0.748	61
2 GF Random Forrest	0.792	0.774	0.849	0.847	0.647	0.769	310
2 GF Logistic Model Tree	0.744	0.737	0.862	0.825	0.594	0.737	200
6 GF Random Forrest	0.885	0.866	0.895	0.913	0.801	0.860	101
6 GF Logistic Model Tree	0.901	0.901	0.960	0.949	0.863	0.899	77
Baseline (JCD Random Forrest)	0.805	0.794	0.870	0.861	0.679	0.79	330
Baseline (Random/Majority)	0.240	0.489	0.512	0.652	0.000	0.322	-

and the 6 layer CNN approach. The best performing approach is a combination of six global features and the LMT classifier with an overall F1 score of 0.899 and 77 FPS. The 6 GF RF approach is slightly faster, but less accurate, than 6 GF LMT. The 2 GF approaches

Table 3: Confusion matrix for both cross validated folds for the 6 GF LMT experiment in table 2. The classes correspond to bowel-preparation quality in BBPS score points from zero to three.

		Detected class				Number of images in the test set
		0	1	2	3	
Actual class	0	248/236	1/0	1/14	0/0	250/250
	1	14/9	1186/1457	23/2	2/7	1225/1475
	2	1/0	4/8	398/423	97/44	500/475
	3	0/0	104/16	132/68	514/516	750/600

gives almost the same detection, while 2 GF RF-based approach is faster and performs slightly better. The 6 layer CNN approach performs almost as good as 6 GF RF-based approach, while it is more than 2 times slower. The 3 layers CNN and TFL approaches performs almost equally in terms of detection performance, while the TFL output is slightly better balanced in terms of precision and recall. The performance of the 2 GF approaches and 3 layers CNN and TFL approaches is almost the same. The relatively low performance of the deep learning based approaches is expected since neural network parameters are not optimized and we trained over a rather small number of epochs. Nevertheless, even if we use very basic methods, our results can be a good starting point for other researchers and used as baselines to benchmark other methods applied to the dataset.

7 CONCLUSION

Adequate bowel preparation (cleansing) is required to achieve high quality colonoscopy examinations. Despite the use of reliable and validated bowel preparation assessment scales, the grading may vary from one doctor to another. By providing the NERTHUS dataset, we invite multimedia researchers to contribute in the medical field by making systems that automatically and consistently can evaluate the quality of bowel cleansing. Innovations in this area contributing with computer-aided assessment and automatic reporting may potentially improve the medical field of GI endoscopy.

REFERENCES

- [1] Martin Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, and others. 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467* (2016).
- [2] Blausen.com staff. 2014. Medical gallery of Blausen Medical 2014. *WikiJournal of Medicine* 1, 2 (2014).
- [3] AH Calderwood and BC Jacobson. 2010. Comprehensive validation of the Boston Bowel Preparation Scale. *Gastrointestinal endoscopy* 72, 4 (2010), 686–692.
- [4] Souad Chaabouni, Jenny Benois-Pineau, and Chokri Ben Amar. 2016. Transfer learning with deep networks for saliency prediction in natural video. In *Proc. of ICIP*. 1604–1608.
- [5] François Chollet. 2015. Keras: Deep learning library for theano and tensorflow. (2015). <https://keras.io/> Accessed: 2017-04-19.
- [6] Brian T. Clark, Tarun Rustagi, and Loren Laine. 2014. What level of bowel prep quality requires early repeat colonoscopy: systematic review and meta-analysis of the impact of preparation quality on adenoma detection rate. *Gastroenterol Report* 109, 11 (2014), 1714–23.
- [7] Douglas A. Corley, Christopher D. Jensen, Amy R. Marks, Wei K. Zhao, Jeffrey K. Lee, Chyke A. Doubeni, Ann G. Zauber, Jolanda de Boer, Bruce H. Fireman, Joanne E. Schottinger, Virginia P. Quinn, Nirupa R. Ghai, Theodore R. Levin, and Charles P. Quesenberry. 2014. Adenoma detection rate and risk of colorectal cancer and death. *New england journal of medicine* 370, 14 (2014), 1298–306.
- [8] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. 2014. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. In *Proc. of ICML*, Vol. 32. 647–655.
- [9] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. M. Parkin, D. Forman, and F. Bray. 2015. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer* 136, 5 (2015), E359–86.
- [10] Florian Froehlich, Vincent Wietlisbach, Jean-Jacques Gonvers, Bernard Burnand, and John-Paul Vader. 2005. Impact of colonic cleansing on quality and diagnostic yield of colonoscopy: the European Panel of Appropriateness of Gastrointestinal Endoscopy European multicenter study. *Gastrointestinal Endoscopy* 61, 3 (2005), 378–384.
- [11] Richard HR Hahnloser, Rahul Sarpeshkar, Misha A Mahowald, Rodney J Douglas, and H Sebastian Seung. 2000. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405, 6789 (2000), 947–951.
- [12] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter* 11, 1 (2009), 10–18.
- [13] M. F. Kaminski, J. Regula, E. Kraszewska, M. Polkowski, U. Wojciechowska, J. Didkowska, M. Zwierko, M. Rupinski, M. P. Nowacki, and E. Butruk. 2010. Quality indicators for colonoscopy and the risk of interval cancer. *New England Journal of Medicine* 362, 19 (2010), 1795–1803.
- [14] Diederik Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- [15] Edwin J Lai, Audrey H Calderwood, Gheorghe Doros, Oren K Fix, and Brian C Jacobson. 2009. The Boston bowel preparation scale: a valid and reliable instrument for colonoscopy-oriented research. *Gastrointestinal endoscopy* 69, 3 (2009), 620–625.
- [16] Benjamin Lebwohl, Fay Kastrinos, Michael Glick, Adam J. Rosenbaum, Timothy Wang, and Alfred I. Neugut. 2011. The impact of suboptimal bowel preparation on adenoma miss rates and the factors associated with early repeat colonoscopy. *Gastrointest Endoscopy* 73, 6 (2011), 1207–14.
- [17] Mathias Lux, Michael Riegler, Pål Halvorsen, Konstantin Pogorelov, and Nektarios Anagnostopoulos. 2016. LIRE: open source visual information retrieval. In *Proc. of MMSys*. Article no. 30.
- [18] Jiquan Ngiam, Adam Coates, Ahbik Lahiri, Bobby Prochnow, Quoc V Le, and Andrew Y Ng. 2011. On optimization methods for deep learning. In *Proc. of ICML*. 265–272.
- [19] R. Parmar, M. Martel, A. Rostom, and AN Barkun. 2016. Validated Scales for Colon Cleansing: A Systematic Review. *American journal of Gastroenterology* 111, 2 (2016), 197–204.
- [20] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. 2017. Kvasir: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection. In *Proc. of MMSYS*.
- [21] Ariadna Quattoni, Michael Collins, and Trevor Darrell. 2008. Transfer learning for image classification with sparse prototype representations. In *Proc. of CVPR*. 1–8.
- [22] Michael Riegler, Mathias Lux, Carsten Griwodz, Concetto Spampinato, Thomas de Lange, Sigrun L Eskeland, Konstantin Pogorelov, Wallapak Tavanapong, Peter T Schmidt, Cathal Gurrin, Dag Johansen, Håvard Johansen, and Pål Halvorsen. 2016. Multimedia and Medicine: Teammates for Better Disease Detection and Survival. In *Proc. of ACM MM*. 968–977.
- [23] Michael Riegler, Konstantin Pogorelov, Pål Halvorsen, Thomas de Lange, Carsten Griwodz, Peter Thelin Schmidt, Sigrun L. Eskeland, and Dag Johansen. 2016. EIR - Efficient Computer Aided Diagnosis Framework for Gastrointestinal Endoscopies. In *Proc. of CBMI*.
- [24] Alaa Rostom and Emilie Jolicoeur. 2004. Validation of a new scale for the assessment of bowel preparation quality. *Gastrointestinal endoscopy* 59, 4 (2004), 482–486.
- [25] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna. 2015. Rethinking the inception architecture for computer vision. *arXiv preprint arXiv:1512.00567* (2015).
- [26] t. Rai, U. Navaneethan, T. Gohel, A. Podugu, P. N. Thota, R. P. Kiran, R. Lopez, and M. R. Sanaka. 2016. Effect of quality of bowel preparation on quality indicators of adenoma detection rates and colonoscopy completion rates. *Gastroenterol Report* 4, 2 (2016), 148–53.
- [27] World Health Organization - International Agency for Research on Cancer. 2012. Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012. (2012). <https://goo.gl/IgZpV1> Accessed: 2017-04-19.