cogent
medicine

GASTROENTEROLOGY & HEPATOLOGY | RESEARCH ARTICLE

Nomogram predicting macroscopic finding with limited or no clinical implication in 19175 patients referred to esophagogastroduodenoscopy

Bo Bendvold[1], Ragnar Weberg[2], Anders Husby[1], Arne Refsum[1] and Kristoffer Watten Brudvik[1,3]*

*Corresponding author: Kristoffer Watten Brudvik, Department of Surgery, Diakonhjemmet Hospital, Oslo, Norway; Department of Gastrointestinal Surgery, Oslo University Hospital, Rikshospitalet, Oslo, Norway
E-mail: kwbrudvik@gmail.com

Abstract: *Objective:* The diagnostic yield of esophagogastroduodenoscopy (EGD) depends on appropriate patient selection. The aim of the current study was to create a nomogram predicting findings with limited or no clinical implication in patients referred to EGD. *Patients and methods:* Indications and findings were registered prospectively in patients who underwent first-time EGD from 1994 to 2013. All findings were classified as "finding with limited or no clinical implication" or "finding with or possibly with clinical implication," and used to create a predicting nomogram. *Results:* A total of 19175 patients were included (female: 59.0% [$n = 11,312$], male: 41.0% [$n = 7,863$]) with 30821 combinations of indications and findings, of which 20,512 (66.6%) constituted findings with limited or no clinical implication. The median age was 61 (58.9 ± 20.6 [mean ± standard deviation]) years and age was the strongest factor associated with normal EGD. Risk relationships were determined for age and indications by sex and used to create a nomogram. Receiver operating characteristics analysis was used to validate the nomogram-calculated probability of a finding with limited or no clinical implication (area under curve: 0.721; $p < 0.001$).

ABOUT THE AUTHORS

The primary focus of the Surgical research group, Department of Surgery, Diakonhjemmet Hospital, is benign and malignant pathology of the upper and lower gastrointestinal tract. Based on the fact that certain macroscopic malignancy of the stomach and the duodenum was described in less than 2% of the institutionally performed esophagogastroduodenoscopies, we set out to investigate and describe the frequency of findings without or with limited clinical relevance in our institution's prospectively maintained database.

PUBLIC INTEREST STATEMENT

Based on the data of the current study, two out of the three endoscopies of the upper gastrointestinal tract may be unnecessary. The association between age, gender, and indication and a finding of clinical importance is apparent. Esophagogastroduodenoscopy or endoscopy of the upper gastrointestinal tract is an excellent and widely used procedure to detect diseases of the esophagus, stomach, and the duodenum. A major concern is the fact that we now perform more esophagogastroduodenoscopies with fewer findings than previously. Furthermore, health care policy-makers and hospital owners have intensified measures to cut the increasing health care cost including low-yield procedures without certain documentation for clinical implication. Based on 19,175 esophagogastroduodenoscopies, a nomogram was created to determine the probability of a finding with limited or no clinical implication. Improved and a common understanding of these challenges by referring physicians, patients, and endoscopists may increase the yield of esophagogastroduodenoscopy.

cogent··medicine

Conclusion: The present nomogram may clarify the EGD yield challenges to referring physicians and patients. Early use of the nomogram may prevent unnecessary EDGs and assist to increase the diagnostic yield of the procedure.

Subjects: Gastroenterology; Gastrointestinal & Abdominal Surgery; General Medicine; Medical Statistics

Keywords: nomogram; esophagogastroduodenoscopy; normal; negative; yield

1. Introduction

Esophagogastroduodenoscopy (EGD) is a safe and excellent procedure to diagnose and treat disease in the upper gastrointestinal tract (Reed, Kilkenny, Dias, & Wexner, 2004; McLernon, Donnan, Crozier, Dillon, & Mowat, 2007). Individuals referred to EGD constitute a heterogeneous cohort with a wide range of symptoms or discomforts. Benign findings, such as a normal finding, reflux esophagitis, hiatus hernia, and gastritis, are common (Broe, Barry, Patchett, & Hill, 2013). In contrast, the finding of cancer is rare and the diagnostic yield of the procedure depends on proper selection criteria (Aljebreen, Alswat, & Almadi, 2013; Buri et al., 2010; Canga & Vakil, 2002). Indication guidelines have been developed to maintain the value of EGD (Adang et al., 1995; Anonymous, 2000; Buri et al., 2010, 2014; Canga & Vakil, 2002; Minoli et al., 1995; Wilcox, Cunniffe, & Tremlett, 1995); however, in our practice, we find these difficult to implement as the referring physician or the patient often express that the value of a normal EGD extends the risk and cost of the procedure. Subsequently, EGD is at risk of becoming a procedure where the endoscopist performs the procedure without questioning the indication. There is a gap between the referring physician/patient and the endoscopist regarding the yield challenges of EGD.

We now perform more EGDs in younger patients with fewer findings of clinical importance than previously. Therefore, we aimed to create a nomogram predicting the probability of a finding with limited or no clinical implication upon EGD based on age and indication by sex. The nomogram could be used by referral physicians together with patients to create a common understanding of the EGD yield challenges and thereby possibly prevent unnecessary referrals and EGDs.

2. Methods

2.1. Data collection

Database approval was obtained from the data protection officer. Indications and findings in patients who underwent EGD at Diakonhjemmet Hospital from 1994 to 2013 were prospectively registered in the database. Anonymous data consisting of the following variables were extracted for analyses: age, sex, year of the EGD, name of the attending physician, referral indication(s), and EGD finding(s). All consecutive in and outpatient EGDs were included. Failed/incomplete and control EGDs were excluded before performing analyses.

2.2. Patient management

Our hospital serves as a primary EGD referral center for a population of 128,992 (2012, Statistics Norway). Almost all referred patients were accepted for EGD; only a few were refused due to insignificant symptoms in young patients or if they presented with co-morbidities that contraindicated EGD. All patients provided informed consent before the procedure was performed. The attending physicians were gastroenterologists or gastrointestinal surgeons with assistance from trained nurses. All patients received local anesthesia with lidocaine orally, unless contraindicated by allergy. Additional anesthesia was only administered to 5–10% of the patients and constituted intravenous benzodiazepine (or similar). Procedures requiring general anesthesia were not included in the prospective database registration. We used standard EGD equipment (Olympus, Tokyo, Japan, or Pentax, Tokyo, Japan). A procedure was not considered complete unless the pylorus was intubated and the duodenal bulb inspected. Macroscopic findings were systematically and prospectively recorded in the database after each EGD. Patients were grouped into two groups based on the

macroscopic finding: (1) finding with limited or no clinical implication or (2) finding with or possibly with clinical implication.

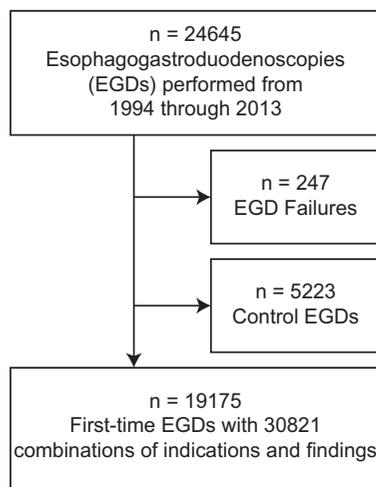### 2.3. Statistics, nomogram preparation, and nomogram validation

Continuous variables were compared with independent t-test. Categorical variables were compared with chi-square test. Pearson's *r* (r) was estimated to assess correlation between the number of EGDs over time and the rate of EGDs in patients less than 40 years of age over time as well as the nomogram-calculated rate for each subgroup and actual observed rate. A *p*-value of < 0.05 was considered significant. To create the nomogram, age was recorded into 10 years intervals: 11–24 years: 20 years, 25–34 years: 30 years; 35–44 years: 40 years; 45–54 years: 50 years; 55–64 years: 60 years; 65–74 years: 70 years; 75–84 years: 80 years; and 85–104 years: 90 years. Chi-square analysis was used to determine the rate of findings at each age interval. The age interval with the highest rate of findings with or possibly with clinical implication (90 years) was attributed zero points, and the inverse rate differences were used to attribute points to the remaining age intervals accordingly. The rate differences between the sexes were assessed individually for each indication and determined by chi-square analyses before the range difference was applied to the nomogram. The impact of age was assessed for all age groups and a corresponding score was applied to the nomogram. The impact and interrelationship of sex and indication were determined with chi-square analysis and a corresponding score used in the nomogram. The nomogram was validated by determining the predicted score for all 30,821 combinations of indications and findings and tested in receiver operator characteristics (ROC) analysis. Furthermore, the nomogram-calculated rate for each nomogram subgroup and the actual observed rate for the subgroup were compared in correlation analysis and presented in a scatter plot. Microsoft Access 2.0–2003 (Microsoft, Redmond, Washington, USA) was used for the prospective database registration and IBM SPSS Statistics (IBM, New York, USA) was used for data analyses and figure preparation.

3. Results

A total of 24,645 EGDs were performed between 1994 and 2013 by 35 endoscopists. After exclusion of failed and control EGDs (*n* = 247 and *n* = 5,223, respectively), 19,175 remained for analyses (Figure 1). Two or more indications and/or findings were recorded in 4,017 patients and the total number of combinations of indications and findings was 30,821, of which 20,512 (66.6%) constituted findings with limited or no clinical implication. Frequencies of all findings in both groups are presented in Table 1.

The female to male ratio in the cohort was 1.44 (females: 59.0% [*n* = 11,312], males: 41.0% [7,863]). The median age was 61 (range 11–104) years for all patients. Figure 2 depicts the age distribution stratified by finding group for females (a) and males (b). The number of annual EGDs

Figure 1. Flow chart of all patients who underwent EGD in the study period and selection of study population.

n = 24645
Esophagogastroduodenoscopies (EGDs) performed from 1994 through 2013

n = 247
EGD Failures

n = 5223
Control EGDs

n = 19175
First-time EGDs with 30821 combinations of indications and findings

cogent ·· medicine

Table 1. Frequency of findings in 30,821 pairs of indications and findings in 19,175 EGDs		
Finding	***n***	**%**
Finding with limited or no clinical implication	**20,512**	**66.6**
Normal	6,543	21.2
Hiatus hernia	4,351	14.1
GERD	4,313	14.0
Gastritis	2,609	8.5
Other minor or unspecific	2,000	6.5
Unspecified duodenitis	696	2.3
Finding with or possibly with implication	**10,309**	**33.4**
Cancer and metaplasia		
Polyp (stomach)	835	2.7
Barret's esophagus	308	1.0
Polyp (duodenum)	214	0.7
Stomach cancer	158	0.5
Esophageal cancer	137	0.4
Suspected stomach cancer	116	0.4
Ectopic gastric mucosa	83	0.3
Polyp (esophagus)	74	0.2
Suspected esophagus cancer	63	0.2
Suspected malignancy duodenum or pancreas	25	0.1
Inflammation		
Gastroduodenitis	644	2.1
Erosive gastritis	524	1.7
Atrophic gastritis	522	1.7
Esophageal candidiasis	409	1.3
Erosive duodenitis	323	1.0
Ulcer or vasculatory		
Duodenal ulcer	1,027	3.3
Stomach ulcer	983	3.2
Stomach erosion	724	2.3
Esophageal ulcer	353	1.1
Duodenal ulcer with bleeding	192	0.6
Vascular anomaly stomach	171	0.6
Mallory–Weiss lesion	139	0.5
Pylori Ulcer	123	0.4
Vascular anomaly duodenum	121	0.4
Stomach ulcer (with bleeding)	107	0.3
Anatomical		
Esophageal stricture	490	1.6
Stomach retention	332	1.1
Deformation bulbi duodeni	287	0.9
Esophageal varices	251	0.8
Esophageal diverticulum	79	0.3
Foreign body	76	0.2
Fundus varices	69	0.2
Pyloric stenosis	66	0.2

cogent ••medicine

Figure 2. Age distribution of findings with limited or no implication (solid lines) and findings with or possibly with implication (dashed lines) in females (a) and males (b).
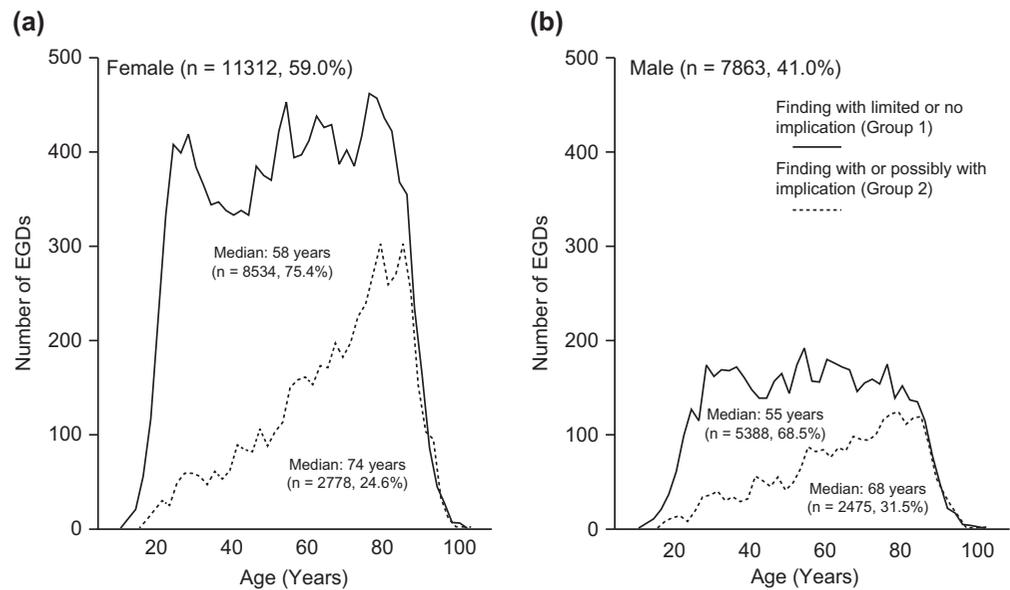

increased by 27.1% in the study period (1994–1996: 1,065 average per year; 2011–2013: 1,354 average per year; *r*, 0.866; *p* < 0.001).

The rate of EGDs performed in patients younger than 40 years of age increased from 10.8% to 27.7% (*r*, 0.907; *p* < 0.001) and 13.5% to 21.9% (*r*, 0.608, *p* = 0.003) in females and males, respectively. Figure 3 shows the five-year average rates of patients less than 40 years (females: 1993–1997, 13.6%; 1998–2002, 20.2%; 2003–2007, 22.7%, and 2008–2013, 25.1%; males: 1993–1997, 16.8%; 1998–2002, 23.8%; 2003–2007, 22.6%, and 2008–2013, 22.9%).

The five most common indications were abdominal pain (*n* = 5,746), dyspepsia (*n* = 4,392), symptom(s) of gastroesophageal reflux disease (*n* = 3,188), nausea/vomiting (*n* = 2,160), and dysphagia (*n* = 2,151). A total of 599 patients were recorded with the indication "assessment of possible

Figure 3. Rate of patients who underwent EGD in different five-year periods (females, solid line; males, dashed line).
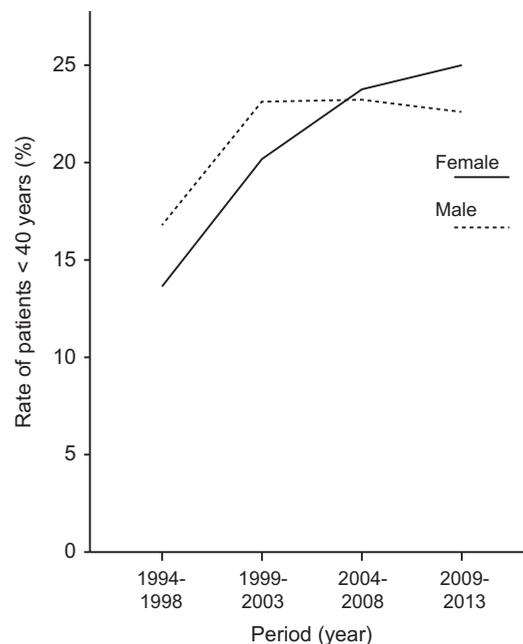
Table 2. Findings in EGDs with the indication of suspicious malignancy (*n* = 599)		
Finding	*n*	**%**
Cancer (macroscopic)	**79**	**13.2**
Stomach cancer (certain)	29	4.8
Stomach cancer (uncertain)	27	4.5
Esophageal cancer (certain)	9	1.5
Esophageal cancer (uncertain)	9	1.5
Malignancy duodenum/pancreas	5	0.8
Metaplasia	**51**	**8.5**
Stomach polyp	29	4.8
Barrett's esophagus	9	1.5
Duodenal polyp	6	1.0
Ectopic stomach mucosa	4	0.7
Esophageal polyp	3	0.5
Other (most commonly listed)	**469**	**78.3**
Normal EGD	76	12.7
Gastritis	63	10.5
Hiatus hernia	52	8.7
GERD	37	6.2
Stomach ulcer	32	5.3
Atrophic gastritis	31	5.2

malignancy" and constitute a highly heterogeneous group of patients ranging from very high to very low chance of cancer finding. On the basis of this, we excluded this indication from the nomogram, but rather present the findings in patients referred with this indication in Table 2. The total numbers of reported cancers of the cohort were 158 (0.8%) and 153 (0.8%) for those assessed as macroscopic certain and 116 (0.6%) and 63 (0.3%) for those assessed as suspicious cancer, for stomach and esophageal cancer, respectively.

A nomogram was created for all other indications (Figure 4). By adding the age points and indication by sex points of a patient to a combined score, the probability of a finding with limited or no implication can be determined by the nomogram. ROC analysis (Figure 5(a)) was used to validate the nomogram-calculated probability of a finding with limited or no clinical implication (area under curve (AUC): 0.721; *p* < 0.001). The nomogram-calculated rate for each nomogram subgroup correlated with the actual observed rate (*r*: 0.824; *p* < 0.001) and presented in scatter plot (Figure 5(b)).

4. Discussion

Based on patient age and indication by sex, we created a simple nomogram to predict a finding with limited or no clinical implication. ROC analysis and correlation analysis showed strong correlation between the predicted and the observed findings of the cohort.

Low-yield diagnostic procedures and screening programs with poorly documented benefit may become targets to cut health care cost and the cost–benefit of a large proportion of EGDs has been questioned (Broe et al., 2013; Dhroove, Chogle, & Saps, 2010; Jweinat, 2010; Pohl, Robertson, & Welch, 2014). In 2012, gastroesophageal reflux disease was part of a list published by Centers for Medicare and Medicaid Services of where services may be overused, misused, or provide limited health care benefit. The European and US cost estimates of a single EGD range from $600 to $3,000 (http://cms.hhs.org) (Broe et al., 2013). We classified two-thirds of the EGD findings with limited or

Figure 4. Nomogram predicting finding with limited or no implication upon EGD. Add points for age (1) and points for indication by sex (2) to determine total points (1 + 2). The probability of finding with limited or no implication can be read by matching with the total point scale.
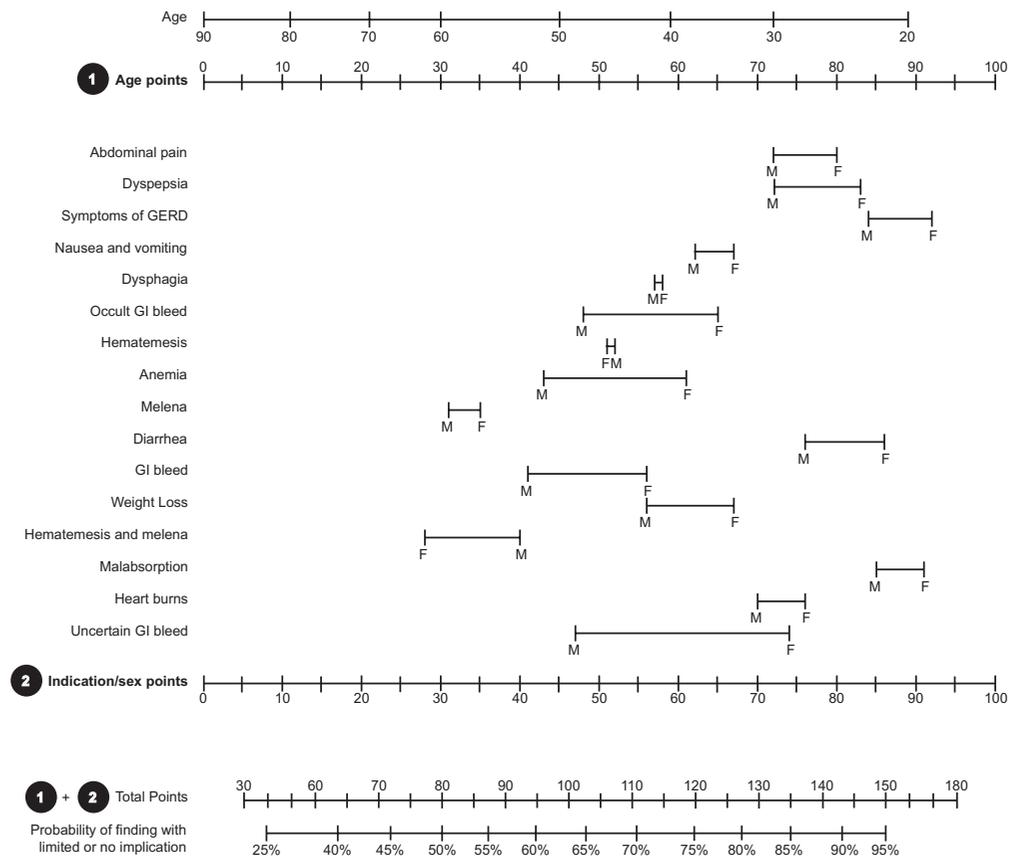
no clinical implication, indicating immense costs of unnecessary procedures. In fact, the ratio of unnecessary EGDs may be even higher: Broe et al. (2013) reported 87.6% normal EGDs in 4,262 patients who underwent EGD. The discrepancy between this study and the current study could be explained by different definitions of normal EGD. Interestingly, in the Broe et al. (2013) study, the rate of normal EGDs in patients under 40 years of age was 92.2%, which is in agreement with our findings.

As of 1 January 2012, our hospital served a population of 128,992 and 54.1% were younger than 40 years (Statistics Norway, 2012). On 1 January 1994, the population was 82,949 and 53.1% were younger than 40 years (Statistics Norway, 1994). Despite the similar age distribution during the

Figure 5. ROC analysis of the nomogram-calculated probability (a) and scatter plot (b) where each circle represents a pair of nomogram-calculated probability of finding with limited or no clinical implication and the actual finding rate within each subgroup (n = 256) of age and sex/indication combinations.
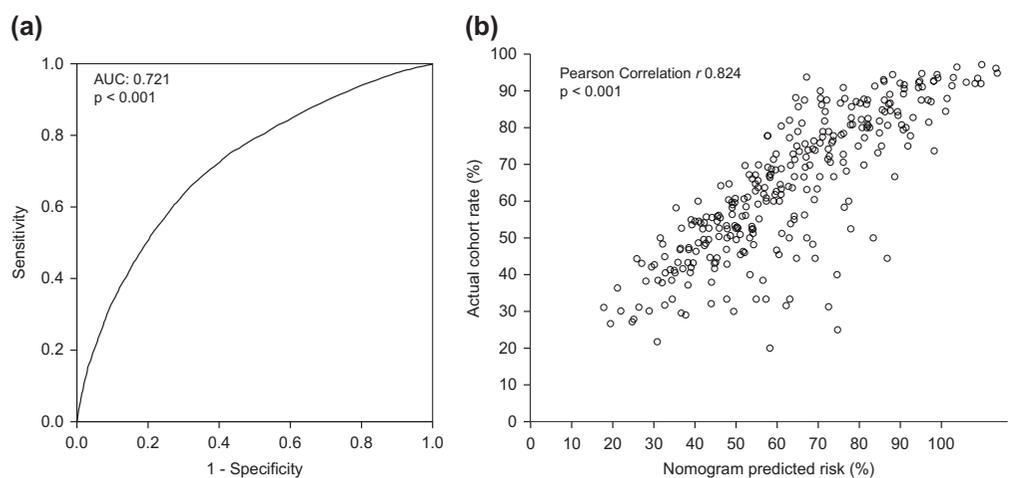
study period, the rate of patients younger than 40 years who underwent EGD increased significantly for both sexes. Furthermore, when considering all EGDs performed in the study period, the number of EGDs performed for each age interval was stable between 20 and 85 years for both sexes, while the patterns of findings were not (Figure 2).

The number of EGDs performed each year increased, in contrast to the decrease in combined national incidence of stomach and esophageal cancer for the same period (Norway CRo, 2012). In contrast, the incidence of colorectal cancer has been increasing during the same time period and colorectal cancer screening programs have not yet been effectuated in Norway (Berstad et al., 2014; Norway CRo, 2012). The cost–benefit of allocating resources from EGD to colonoscopy may become relevant and should be investigated further.

The present study has the following limitations. First, the incidence rates of esophageal and stomach cancer vary between countries and continents from 5 per 10,000 in low-risk areas (North America and Australia) to 70 per 10,000 in high-risk areas (parts of Asia and Africa) (Crew & Neugut, 2005). Age-adjusted incidence rates in Norway 2010 were 4.4 (men) and 1.1 (women) per 100,000 for esophageal cancer and 6.9 (men) and 3.0 (woman) per 100,000 for stomach cancer (Cancer Registry of Norway, 2010). Furthermore, the prevalence of helicobacter pylori, a major risk factor of pathology in the upper gastrointestinal tract, shows the same endemic pattern (Bai et al., 2010; Crew & Neugut, 2005). As a consequence, our findings may not be applicable in high-risk regions. Second, while findings were registered prospectively, their significance was classified retrospectively. Based on comparable studies in the literature, we acknowledge that we are likely to over classify the significance of certain findings. However, this was done intentionally to reduce the risk of under classifying significant findings in the present study. In contrast, there may be findings falsely classified as non-significant, for example, the low clinical implication of hiatus hernia and gastroesophageal reflux disease, which were recorded in 4,351 and 4,313 EGDs, respectively. During the study period, we performed 190 anti-reflux surgical procedures (183 laparoscopic and 7 open), indicating that the clinical implication for the vast majority was medical advice, which in most cases already had been given by the referring physician. Third, the findings reported in the present study were not biopsy verified as database entries were made prospectively, directly after EGD, which may have caused reporting bias. However, as the primary aim was to develop a nomogram predicting finding with limited or no clinical implication and all macroscopic cancers and suspicious cancers were excluded, we believe that the possibility of bias based in this was negligible.

In summary, the current nomogram is easy to use and includes variables that are accessible early in the evaluation of patients. Cancers do occur in young individuals with few or no early symptoms and in practice, we would not recommend the nomogram to replace current guidelines and a thorough clinical individual assessment. However, the high rates of normal EGDs have become a major problem and indicate poor use of resources. We believe that the nomogram could assist referring physicians together with patients to understand the yield challenges of EGD, and this may prevent a number of unnecessary procedures.

Author details
Bo Bendvold[1]
E-mail: bo@bendvold.com
Ragnar Weberg[2]
E-mail: ragweber@gmail.com
Anders Husby[1]
E-mail: husbyanders@gmail.com
Arne Refsum[1]
E-mail: arne.refsum@gmail.com

Kristoffer Watten Brudvik[1,3]
E-mail: kwbrudvik@gmail.com
ORCID ID: http://orcid.org/0000-0002-1503-2107
[1] Department of Surgery, Diakonhjemmet Hospital, Oslo, Norway.
[2] Department of Medicine, Diakonhjemmet Hospital, Oslo, Norway.
[3] Department of Gastrointestinal Surgery, Oslo University Hospital, Rikshospitalet, Oslo, Norway.

References

Adang, R. P., Vismans, J. F. J. F. E., Talmon, J. L., Hasman, A., Ambergen, A. W., & Stockbrügger, R. W. (1995). Appropriateness of indications for diagnostic upper gastrointestinal endoscopy: Association with relevant endoscopic disease. *Gastrointestinal Endoscopy, 42*, 390–397. http://dx.doi.org/10.1016/S0016-5107(95)70037-4

Aljebreen, A. M., Alswat, K., & Almadi, M. A. (2013). Appropriateness and diagnostic yield of upper gastrointestinal endoscopy in an open-access endoscopy system. *Saudi Journal of Gastroenterology, 19*, 219–222.

Anonymous. (2000). Appropriate use of gastrointestinal endoscopy. *Gastrointest Endosc, 52*, 831–837.

Bai, Y., Li, Z. S., Zou, D. W., Wu, R. P., Yao, Y. Z., Jin, Z. D., … Xu, G. M. (2010). Alarm features and age for predicting upper gastrointestinal malignancy in Chinese patients with dyspepsia with high background prevalence of Helicobacter pylori infection and upper gastrointestinal malignancy: An endoscopic database review of 102,665 patients from 1996 to 2006. *Gut, 59*, 722–728. http://dx.doi.org/10.1136/gut.2009.192401

Berstad, P., Løberg, M., Larsen, I. K., , Kalager, M., Holme, Ø., Botteri, E., … Hoff, G. (2014). Long-term lifestyle changes after colorectal cancer screening: Randomised controlled trial. *Gut, 64*, 1268–1276. doi:10.1136/gutjnl-2014-307376

Broe, M., Barry, M., Patchett, S., & Hill, A. D. K. (2013). Evaluating the clinical efficacy and cost effectiveness of direct access endoscopy. *The Surgeon, 11*, 304–308. http://dx.doi.org/10.1016/j.surge.2013.02.005

Buri, L., Hassan, C., Bersani, G., Anti, M., Bianco, M. A., Cipolletta, L., … Buscema, M. (2010). Appropriateness guidelines and predictive rules to select patients for upper endoscopy: A nationwide multicenter study. *The American Journal of Gastroenterology, 105*, 1327–1337. http://dx.doi.org/10.1038/ajg.2009.675

Burri, E., Manz, M., Schroeder, P., Froehlich, F., Rossi, L., Beglinger, C., & Lehmann, F. (2014). Diagnostic yield of endoscopy in patients with abdominal complaints: Incremental value of faecal calprotectin on guidelines of appropriateness. *BMC Gastroenterology, 14*, 57. http://dx.doi.org/10.1186/1471-230X-14-57

Canga, 3rd, C., & Vakil, N. (2002). Upper GI malignancy, uncomplicated dyspepsia, and the age threshold for early endoscopy. *The American Journal of Gastroenterology, 97*, 600–603. http://dx.doi.org/10.1111/j.1572-0241.2002.05536.x

Crew, K. D., & Neugut, A. I. (2005). Epidemiology of gastric cancer. *World Journal of Gastroenterology, 12*, 354–362.

Dhroove, G., Chogle, A., & Saps, M. (2010). A million-dollar work-up for abdominal pain: Is it worth it? *Journal of Pediatric Gastroenterology and Nutrition, 51*, 579–583. http://dx.doi.org/10.1097/MPG.0b013e3181de0639

Jweinat, J. J. (2010). Hospital readmissions under the spotlight. *Journal of Healthcare Management, 55*, 252–264.

McLernon, D. J., Donnan, P. T., Crozier, A., Dillon, J., & Mowat, C. (2007). A study of the safety of current gastrointestinal endoscopy (EGD). *Endoscopy, 39*, 692–700. http://dx.doi.org/10.1055/s-2007-966578

Minoli, G., Prada, A., Gambetta, G., Formenti, A., Schalling, R., Lai, L., & Pera, A. (1995). The ASGE guidelines for the appropriate use of upper gastrointestinal endoscopy in an open access system. *Gastrointestinal Endoscopy, 42*, 387–389. http://dx.doi.org/10.1016/S0016-5107(95)70036-6

Norway CRo. (2012). *Cancer in Norway 2012 - Cancer incidence, mortality, survival and prevalence in Norway*. Oslo: Institute of population based cancer research.

Pohl, H., Robertson, D., & Welch, H. G. (2014). Repeated upper endoscopy in the Medicare population: A retrospective analysis. *Annals of internal medicine, 160*, 154–160.

Reed, W. P., Kilkenny, J. W., Dias, C. E., & Wexner, S. D. (2004). A prospective analysis of 3525 esophagogastroduodenoscopies performed by surgeons. *Surgical Endoscopy, 18*, 11–21. http://dx.doi.org/10.1007/s00464-003-8913-3

Wilcox, M. H., Cunniffe, J. G., & Tremlett, C. (1995). Guidelines on appropriate indications for upper gastrointestinal endoscopy. *BMJ, 311*, 57. http://dx.doi.org/10.1136/bmj.311.6996.57b