



Original Article

# Development of Clinical Prediction Models for Surgery and Complications in Crohn's Disease

Leonardo Guizzetti,<sup>c</sup> Guangyong Zou,<sup>a,c</sup> Reena Khanna,<sup>b,c</sup>  
Parambir S. Dulai,<sup>d</sup> William J. Sandborn,<sup>c,d</sup> Vipul Jairath,<sup>a,b,c</sup>  
Brian G. Feagan<sup>a,b,c</sup>

<sup>a</sup>Department of Epidemiology and Biostatistics, University of Western Ontario, London, ON, Canada <sup>b</sup>Department of Medicine, Division of Gastroenterology, University of Western Ontario, London, ON, Canada <sup>c</sup>Robarts Clinical Trials Inc., University of Western Ontario, London, ON, Canada <sup>d</sup>Division of Gastroenterology, University of California San Diego, La Jolla, CA, USA

Corresponding author: Brian G. Feagan, MD, Robarts Clinical Trials Inc., 100 Dundas Street, Suite 200, London, ON, Canada N6A 5B6. Tel.: 226-270-7675; fax: 519-931-5278; email: [brian.feagan@robartsinc.com](mailto:brian.feagan@robartsinc.com)

## Abstract

**Background and Aims:** Crohn's disease-related complications account for a substantial proportion of inflammatory bowel disease-associated health care expenditure. Identifying patients at risk for complications may allow for targeted use of early therapeutic interventions to offset this natural course. We aimed to develop risk prediction models for Crohn's disease-related surgery and complications.

**Methods:** Using data from the Randomised Evaluation of an Algorithm for Crohn's Disease cluster-randomised clinical Trial [REACT], which involved 1898 patients from 40 community practices, separate prediction models were derived and internally validated for predicting Crohn's disease-related surgery and disease-related complications [defined as the first disease-related surgery, hospitalisation, or complication within 24 months]. Model performance was assessed in terms of discrimination and calibration, decision curves, and net benefit analyses.

**Results:** There were 130 [6.8%] disease-related surgeries and 504 [26.6%] complications during the 24-month follow-up period. Selected baseline predictors of surgery included age, gender, disease location, Harvey-Bradshaw Index [HBI] score, stool frequency, antimetabolite or 5-aminosalicylate use, and the presence of a fistula, abscess, or abdominal mass. Selected predictors of complications included those same factors for surgery, plus corticosteroid or anti-tumour necrosis factor use, but excluded 5-aminosalicylate use. Discrimination ability, as measured by validated c-statistics, was 0.70 and 0.62 for the surgery and complication models, respectively. Score charts and nomograms were developed to facilitate future risk score calculation.

**Conclusions:** Separate risk prediction models for Crohn's disease-related surgery and complications were developed using clinical trial data involving community gastroenterology practices. These models could be used to guide Crohn's disease management. External validation is warranted.

**Key Words:** Clinical trials; clinical risk prediction; prognostic model

## 1. Introduction

Crohn's disease [CD] is an idiopathic disorder of the gastrointestinal tract, characterised by chronic, segmental inflammation, which

typically has a relapsing and remitting course. Despite the use of immunosuppressive maintenance therapies, subclinical transmural inflammation persists in many patients,<sup>1</sup> which predisposes them to

such complications as strictures and fistulas.<sup>2</sup> Recent studies estimate the long-term risk of surgery to be approximately 60–80%,<sup>2–5</sup> with the greatest risk in the first few years following diagnosis.

In the past two decades, the advent of biologic therapies and refinement of treatment paradigms have revolutionised the medical management of CD. Specific advances include the introduction of tumour necrosis factor [TNF] antagonists<sup>6–13</sup> and integrin inhibitors,<sup>14–18</sup> the use of combination therapy,<sup>19–22</sup> therapeutic drug monitoring,<sup>23,24</sup> and earlier initiation of effective therapies in high-risk patients.<sup>12,19–21,25–29</sup> However, one of the greatest challenges in implementing these strategies is determining which patients are most at risk for complicated disease and thus candidates for more intensive monitoring and treatment. This decision is usually based on clinical judgement and heavily weighted by the patient's disease activity as assessed by symptoms. This approach is practical, but it does not routinely incorporate prognostic factors that drive disease-related complications. Accordingly, identification of patients at highest risk of complications and disease progression, who have the greatest chance of benefiting from early initiation of effective therapy, is an aspirational goal.

In this regard, the approach to therapy in CD has changed dramatically. Formerly, 'step-care', in which drugs are used sequentially to attain symptomatic remission, was the preferred paradigm. Whereas this approach is attractive because it avoids over-treating low-risk patients, step-care delays initiation of effective therapy in patients most at risk for adverse outcomes. Recently, attention has turned to early introduction of combination immunosuppression therapy in high-risk patients, to promote mucosal healing and minimise exposure to corticosteroids.<sup>19,20</sup> This 'top-down' approach requires accurate identification of high-risk patients, to minimise over-treatment in lower-risk patients, with the associated risk of treatment-related adverse events and costs. Conversely, misclassification of high-risk patients delays administration of effective therapies and potentially results in increased risk of complications. Therefore, the ability to accurately risk-stratify patients has garnered considerable interest.<sup>30,31</sup>

Retrospective analyses of single-centre and population-based cohorts have identified several prognostic factors in CD, including younger age at diagnosis, ileal disease location, perianal disease, stricturing or penetrating phenotype, current smoking, treatment with corticosteroids at diagnosis or corticosteroid dependence, and extensive disease involvement.<sup>1,3–5,31–40</sup>

Existing clinical prediction models have several limitations. Notably, these models were developed from single centres, referral centres, retrospective samples, small development samples, or cohorts assembled before the widespread use of TNF antagonists, and they often predict excessively long-term risk estimates of outcomes [5–10 years].<sup>32,33,41–44</sup> However, to fully utilise this information, prediction models must be established.

Data arising from randomised trials offer a unique ability in developing clinical prediction models, because of their scope, size, multicentre participation and prospective follow-up, thus overcoming some of the aforementioned design limitations. We developed and internally validated clinical prediction models for CD using data from a large cluster randomised trial conducted throughout community practices, comparing two treatment algorithms for CD.<sup>21</sup> Our specific aim was to develop separate models for predicting risks of two binary endpoints within 24 months of follow up: [1] occurrence of CD-related surgery; and [2] occurrence of a composite of CD-related surgery, complications or hospitalisation [this composite outcome is henceforth referred to as disease-related complications].

## 2. Methods

### 2.1. Data source

The study population was participants from the Randomised Evaluation of an Algorithm for Crohn's Disease [REACT] trial [NCT01030809].<sup>21</sup> Briefly, REACT was a large, cluster-randomised, controlled trial of two distinct algorithms for the management of CD. Forty Canadian and Belgian community-based gastroenterology clinics were randomised, in a 1:1 ratio, to either early combined immunosuppression or step-care. In each clinic, consecutive adult patients with CD were enrolled, regardless of disease activity or concurrent therapy, and followed up to 24 months. The dataset was used to develop separate models to predict the risk over 24 months of having CD-related surgery and disease-related complications.

The reporting of this study follows the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis [TRIPOD] statement [Supplementary Table 1, available at *ECCO-JCC* online].<sup>45</sup>

### 2.2. Clinical outcomes and definitions

Binary outcomes were defined within 2 years of follow up. The first was the occurrence of CD-related surgery. The second was a CD-related complication. Surgery was considered separately since it is a readily measured event of clinical importance. All events in the REACT trial used in this study were evaluated by an adjudication committee who were blinded to treatment assignment. Disease-related surgeries included resective bowel surgery [ileal resection, ileocaecal resection, proctocolectomy, colectomy, enterectomy, ostomy formation and repair, anastomosis/reanastomosis], strictureplasty, and fistula repair [incision and drainage of abscess, seton placement, fistulotomy, fistulectomy]. Disease-related complications were defined as a composite of disease-related surgery [as defined above], complications [including development of penetrating or stricturing disease, worsening abdominal pain, increased stool frequency, extra-intestinal manifestations, severe perianal disease, fistula, or abscess<sup>21</sup>] or hospitalisation.

### 2.3. Selection of predictors

Candidate prognostic factors were selected from the demographic and disease-related variables, using standardised clinical definitions, collected at baseline and before treatment allocation. A search of the literature identified potential prognostic factors of interest that were augmented by clinical judgement. These variables were age at enrolment, age at diagnosis, gender, smoking status, disease location, perianal disease, previous surgical resection for CD, use of each medication at baseline [including 5-aminosalicylates, corticosteroids, immunosuppressives, and TNF antagonists], abdominal pain, abdominal mass, extra-intestinal manifestations, strictures, fistulas, and stool frequency. No laboratory parameters, biomarkers, or cross-sectional imaging items were included. Since the purpose of these models is to predict risks using baseline factors before treatment allocation, treatment assignment was not considered as a predictor in the models. Moreover, coefficients for treatment effect were relatively small compared with other prognostic factors.

### 2.4. Missing data

No participants had missing outcome data. However, 4.24% [ $n = 84$ ] of patients had at least one missing variable at baseline, of which it was most common to be missing components of the HBI score. No patterns were observed in the missing values. Only participants with complete baseline data were used for model development.

## 2.5. Model development

Exploratory univariate data analysis was first conducted to assess adequate event frequency between each outcome and the candidate prognostic factors. Univariable associations between candidate predictors and the outcomes were assessed graphically and using logistic regression. Each model was then constructed using multivariable logistic regression analysis. Predictors with univariate associations  $p < 0.20$  entered the multivariable model. From this full model, unnecessary variables were removed based on a difference in Akaike information criteria [AIC;  $< 2.0$ ]. In terms of the relative performance of models as estimated from bootstrap replicates, the removal of these variables had a negligible impact on discrimination, the Brier score [ $< 1\%$  difference in mean difference], and calibration. Although the data arose from a cluster-randomised trial, regression coefficient estimates are not much affected when the degree of clustering is low as in this trial [intraclass correlation coefficient  $< 0.02$ ].<sup>46</sup> Since the focus in prediction studies is on the absolute risk estimate from the combined predictor effects, these analyses were conducted at the patient level without accounting for clustering.

## 2.6. Predictive performance and model validation

The model development and validation process adhered to recommended guidelines.<sup>47,48</sup> Model performance was characterised by the discrimination ability and calibration: the discrimination ability of a model to distinguish between patients with the outcome and patients without the outcome. In the present context, discrimination is measured using the c-statistic, which is identical to the area under the receiver operating characteristic curve for binary outcomes.<sup>49</sup> A value of 0.50 for the c-statistic represents the prognostic ability of a coin

flip and will therefore correctly differentiate 50% of cases. There are no universal guidelines for describing the quality of discrimination ability, but generally discrimination values below 60% are unacceptable, 60% to 70% may be acceptable and values from 70% to  $> 90\%$  range from good to excellent. The interpretation of these values is context-dependent. Calibration refers to the agreement between predicted and observed risks.<sup>47</sup> This can be assessed by plotting the predicted risks and the observed percentages of patients who had the outcome.

Since model development tends to over-fit the data, commonly referred to as optimism, it is desirable to reduce optimism using internal validation and/or external validation.<sup>50</sup> Internal validation of the models was conducted using the bootstrap method with 500 replicates.<sup>51</sup> Performance characteristics were calculated for the original [index] and validated models following correction for optimism. Additional details regarding statistical methodology and performance measures are provided in the Supplementary Appendix, available at ECCO-JCC online.

## 2.7. Sample size

Formal sample size calculation was not conducted. Of the 1982 participants in the REACT trial, 504 had disease-related complications and 130 had surgeries over 2 years of follow-up. According to the guideline of 10 events per variable,<sup>52</sup> this dataset is large enough to construct prediction models with at least ten predictors.

## 2.8. General statistical methods

Exploratory analysis was performed using Stata [version 14.1/IC; StataCorp, College Station, TX]. Model building and evaluation

**Table 1.** Overall patient characteristics at baseline [N = 1898].

Baseline characteristic	Overall N = 1898	Surgery group N = 130 6.8%	No surgery group N = 1768 93.1%	Complication group N = 504 26.6%	No complication group N = 1394 73.5%
<b>Demographics</b>					
Age, years [mean $\pm$ SD]	44.0 $\pm$ 14.6	42.7 $\pm$ 14.0	44.1 $\pm$ 14.6	42.1 $\pm$ 14.6	44.7 $\pm$ 14.5
Female	1097 [57.78%]	68 [52.3%]	1029 [58.2%]	311 [61.7%]	786 [56.4%]
<b>Smoking status</b>					
Non-smoker	939 [49.5%]	58 [44.6%]	881 [49.9%]	241 [47.9%]	698 [50.1%]
Ex-smoker	550 [29.0%]	36 [27.7%]	514 [29.1%]	405 [29.1%]	145 [28.8%]
Current smoker	407 [21.44%]	36 [27.7%]	371 [21.0%]	290 [20.8%]	233 [11.7%]
<b>Disease characteristics</b>					
Duration, months (mean [median])	148.8 [119]	144.8 [111.1]	149.1 [119.2]	147.0 [114.0]	149.5 [120.9]
HBI score [median; mean $\pm$ SD]	3; 4.1 $\pm$ 1.1	5; 6.0 $\pm$ 3.5	3; 4.0 $\pm$ 1.1	4; 5.0 $\pm$ 4.9	3; 3.8 $\pm$ 3.8
Steroid-free remission [HBI $\leq$ 4]	1,065 [56.1%]	51 [39.2%]	1014 [57.4%]	236 [46.8%]	829 [59.5%]
Extra-intestinal manifestations at baseline (n [%])	594 [31.3%]	55 [42.3%]	539 [30.5%]	185 [36.7%]	409 [29.3%]
Previous history of disease-related surgery	864 [45.5%]	56 [43.1%]	808 [45.7%]	244 [48.4%]	620 [44.5%]
<b>Involved intestinal areas</b>					
Colon	417 [22.0%]	17 [13.1%]	400 [22.6%]	92 [18.3%]	325 [23.3%]
Small bowel	654 [34.5%]	52 [40.0%]	602 [34.1%]	169 [33.5%]	485 [34.8%]
Colon and small bowel	827 [43.6%]	61 [46.9%]	766 [43.3%]	243 [48.2%]	584 [41.9%]
<b>Concurrent medications use</b>					
Aminosalicylates	539 [28.4%]	22 [16.9%]	517 [29.2%]	116 [23.0%]	423 [30.3%]
Corticosteroids	348 [18.3%]	31 [23.9%]	317 [17.9%]	113 [22.4%]	235 [16.9%]
Anti-metabolites	826 [43.5%]	48 [36.9%]	778 [44.0%]	204 [40.5%]	622 [44.6%]
TNF antagonists	622 [32.8%]	53 [40.8%]	569 [32.2%]	201 [39.9%]	421 [30.2%]

Figures may vary slightly from those published in Khanna *et al.* [2015], due to exclusion of 84 patients with at least one missing predictor variable. SD, standard deviation; HBI, Harvey-Bradshaw Index; TNF, tumour necrosis factor.

were conducted using package *rms*<sup>47</sup> in R [version 3.3.1; Linux; R Core Team<sup>53</sup>] through RStudio [version 0.99; RStudio Team<sup>54</sup>]. Summary statistics are presented as mean  $\pm$  standard deviation [SD], median, or frequencies and proportions as appropriate.

### 3. Results

#### 3.1. Patient characteristics and clinical outcomes

Baseline characteristics are outlined in Table 1. A total of 1097 [57.79%] patients were female. Median disease duration was

148.8 months, the mean HBI was 4.1, and almost one-third of patients were taking biologics at baseline. Over the 24-month follow-up period, 6.84% [ $n = 130$ ] of participants underwent CD-related surgery, and 26.55% [ $n = 504$ ] of participants experienced a CD-related complication. Univariate associations for each outcome are shown in Table 2.

#### 3.2. Model performance and validation

A separate logistic regression model was estimated for each outcome. Table 3 shows the estimates of regression coefficients for the

**Table 2.** Univariate associations for Crohn's disease-related complications or CD-related surgery alone.

Baseline variable	CD-related surgery		CD-related complication <sup>a</sup>	
	OR [95% CI]	<i>p</i> -value	OR [95% CI]	<i>p</i> -value
Current age	0.99 [0.98, 1.00]	0.27	0.99 [0.98, 1.00]	< 0.0001
Gender [male vs. female]	1.27 [0.89, 1.81]	0.19	0.80 [0.65, 0.99]	0.04
HBI score [except stool frequency]	1.36 [1.25, 1.48]	< 0.0001	1.18 [1.12, 1.24]	< 0.0001
Stool frequency	1.08 [1.15, 1.02]	< 0.01	1.07 [1.03, 1.11]	< 0.0001
Location of disease involvement				
Colon only [reference]	1	0.03	1	0.02
Small bowel and colon	1.87 [1.08, 3.25]		1.47 [1.12, 1.94]	
Small bowel only	2.03 [1.16, 3.57]		1.23 [0.92, 1.65]	
Presence of new fistula, abscess, or definite abdominal mass	10.29 [6.23, 16.99]	< 0.0001	3.94 [2.48, 6.26]	< 0.0001
Anti-metabolite use	0.75 [0.52, 1.08]	0.11	0.84 [0.69, 1.04]	0.10
5-aminosalicylate use	0.49 [0.31, 0.79]	< 0.01	0.69 [0.54, 0.87]	< 0.01
Corticosteroid use	1.43 [0.94, 2.18]	0.10	1.43 [1.11, 1.84]	< 0.01
TNF antagonist use	1.45 [1.01, 2.09]	0.05	1.53 [1.24, 1.89]	< 0.0001
Current smoking vs. ex-/non-smoker	1.44 [0.96, 2.15]	0.08	1.15 [0.90, 1.47]	0.26

All outcomes are defined within 24 months of follow-up.

CD, Crohn's disease; OR, odds ratio, CI, confidence interval; HBI, Harvey-Bradshaw Index; TNF, tumour necrosis factor.

<sup>a</sup>Complication refers to the first occurrence of a CD-related surgery, complication or hospitalisation.

**Table 3.** Coefficient estimates of logistic regression models for CD-related surgery and a CD-related complication.

Baseline predictor	CD-related surgery			CD-related complication <sup>a</sup>		
	Regression coefficient <sup>b</sup>	SE	<i>p</i> -value	Regression coefficient <sup>c</sup>	SE	<i>p</i> -value
Intercept	-3.607	0.327		-1.6134	0.162	
Age subtract 45 years	-0.0027	0.007	0.72	-0.0102	0.004	0.02
Male vs female	0.3946	0.194	0.06	-0.1848	0.110	0.13
HBI score [except stool frequency]	0.2425	0.065	< 0.001	0.1139	0.042	0.01
Stool frequency <sup>d</sup>	0.0741	0.060	0.25	0.0752	0.034	0.04
Location of disease involvement						
Colon only [reference]	0	–	–	0	–	–
Small bowel and colon	0.3614	0.293	0.26	0.2525	0.145	0.12
Small bowel only	0.5167	0.297	0.11	0.1739	0.153	0.31
Corticosteroid use	–	–	–	0.2465	0.138	0.11
Anti-metabolite use	-0.4519	0.200	0.04	-0.1598	0.110	0.19
5-aminosalicylate use	-0.6015	0.253	0.03	–	–	–
TNF antagonist use	–	–	–	0.3887	0.114	< 0.01
Presence of new fistula, abscess, or definite abdominal mass	1.7019	0.306	< 0.0001	1.0840	0.261	< 0.001
Interaction						
HBI score $\times$ stool frequency	-0.0182	0.014	0.22	-0.0136	0.009	0.18

SE, standard error; CD, Crohn's disease; HBI, Harvey-Bradshaw Index; TNF, tumour necrosis factor.

<sup>a</sup>Complication refers to the first occurrence of a CD-related surgery, complication, or hospitalisation.

<sup>b</sup>Regression coefficients are presented before shrinkage [shrinkage factor = 0.92].

<sup>c</sup>Regression coefficients are presented before shrinkage [shrinkage factor = 0.90].

<sup>d</sup>Stool frequency has a maximum value of 12.

CD-related surgery and CD-related complication models. The baseline predictors included in the surgery model were age, gender, disease location, HBI score, stool frequency, immunosuppressive use, 5-aminosalicylate use, and the presence of a fistula, abscess, or abdominal mass. The baseline predictors for the CD-related complication model also uniquely included the use of corticosteroids and TNF antagonists, in addition to all the variables identified for surgery.

The validated discrimination ability for the CD-related surgery model was reasonably good with a c-statistic of 0.70 (95% confidence interval [CI]: 0.69 to 0.71), whereas the discrimination ability of disease-related complication model was relatively acceptable at 0.62 [95% CI: 0.61 to 0.64] [Supplementary Table 2, available at *ECCO-JCC* online]. Both models demonstrated low degrees of optimism [shrinkage factors were 0.92 and 0.90, respectively], and excellent calibration and goodness-of-fit statistics [see Supplementary Appendix].

### 3.3. Computing a risk estimate, score chart, and nomogram

The validated logistic regression model for CD-related surgery within 2 years is shown in Equation 1.

$$\begin{aligned} \text{logit}(\text{Pr}(\text{Surgery} = Y)) = & -3.511 \\ & + (-0.0025) \times (\text{Age} - 45 \text{ years}) \\ & + (0.3634) \times (1 \text{ if gender} = \text{Male}; 0 \text{ otherwise}) \\ & + (0.2234) \times (\text{HBI score} \\ & \text{less stool frequency}) \\ & + (0.0683) \times (\text{stool frequency}) \\ & + (-0.0167) \times (\text{HBI score} \\ & \text{less stool frequency}) \\ & \times (\text{stool frequency}) \\ & + (0.3328) \times (1 \text{ if small bowel and colonic} \\ & \text{disease}; 0 \text{ otherwise}) \\ & + (0.4759) \times (1 \text{ if small bowel} \\ & \text{only disease}; 0 \text{ otherwise}) \\ & + (-0.4162) \times (1 \text{ if anti - metabolite use}; \\ & 0 \text{ otherwise}) \\ & + (-0.5540) \times (1 \text{ if 5 - ASA use}; 0 \text{ otherwise}) \\ & + (1.5674) \times (1 \text{ if new fistula,} \\ & \text{abscess or abdominal mass}) \end{aligned} \tag{1}$$

The validation logistic regression model for CD-related complications is shown in Equation 2.

$$\begin{aligned} \text{logit}(\text{Pr}(\text{Complication} = Y)) = & -1.552 \\ & + (-0.0092) \times (\text{Age} - 45 \text{ years}) \\ & + (-0.1669) \times (1 \text{ if gender} = \text{Male}; 0 \text{ otherwise}) \\ & + (0.1028) \times (\text{HBI score} \\ & \text{less stool frequency}) \\ & + (0.0679) \times (\text{stool frequency}) \\ & + (-0.0123) \times (\text{HBI score} \\ & \text{less stool frequency}) \\ & \times (\text{stool frequency}) \\ & + (0.2280) \times (1 \text{ if small bowel and colonic} \\ & \text{disease}; 0 \text{ otherwise}) \\ & + (0.1570) \times (1 \text{ if small bowel only} \\ & \text{disease}; 0 \text{ otherwise}) \\ & + (0.2226) \times (1 \text{ if corticosteroid} \\ & \text{use}; 0 \text{ otherwise}) \\ & + (-0.1443) \times (1 \text{ if antimetabolite use}; \\ & 0 \text{ otherwise}) \\ & + (0.3510) \times (1 \text{ if TNF - antagonist use}; \\ & 0 \text{ otherwise}) \\ & + (0.9789) \times (1 \text{ if new fistula, abscess or} \\ & \text{abdominal mass}) \end{aligned} \tag{2}$$

To compute the risk score, multiply the value of each variable with its associated regression coefficient, and then sum all values, producing the linear predictor [lp]. The linear predictor on the log-odds scale may be converted to a risk probability [%] by the inverse log-odds formula:  $p = 100\% \times [1 + \exp(lp)]^{-1}$ . An example risk calculation is demonstrated in Table 4. A risk score may alternatively be calculated using score charts and converted to a risk estimate [Tables 5 and 6] or computed by the provided nomograms [Figure 1A and B].

For prediction of CD-related surgery, the score ranges from 0 to 233 [Table 5]. A score of 61, 119, or 171 respectively predict a 5%, 20%, or 50% chance of experiencing a disease-related surgery within 24 months of follow-up. For prediction of CD-related complications, the score ranges from 0 to 314 [Table 6]. A score of 88, 177, or 266 respectively predict a 25%, 50%, or 75% chance of experiencing a disease-related complication within 24 months of follow-up.

## 4. Discussion

We developed two clinical prediction models, one for predicting individual risk of experiencing CD-related surgery, and the other for

**Table 4.** Computing the predicted risk score.

As an example, suppose a 50-year old male patient presents to the clinic with disease confined to the small bowel. His total HBI score is 3 with a stool frequency of two per day, he is currently taking 5-ASA, and does not have a fistula, stricture, or definite abdominal mass. The clinician would now like to predict the risk of surgery within the next 2 years.

Start by plugging these values into the equation for the linear predictor.

$$\begin{aligned} \text{lp} = & -3.51 \text{ [Intercept]} \\ & + [50 - 45; \text{age, years}] \times [-2.5 \times 10^{-3}] + [1; \text{male}] \times [0.36] \\ & + [1 \text{ point; HBI subtract stools/day}] \times [0.22] \\ & + [2 \text{ stools/day}] \times [6.8 \times 10^{-2}] \\ & + [1; \text{small bowel only}] \times [0.48] \\ & + [0; \text{antimetabolite use}] \times [-0.42] \\ & + [1; 5\text{-ASA use}] \times [-0.55] \\ & + [0; \text{fistula, abscess, abdominal mass}] \times [1.57] \\ & + [1 \text{ point; HBI subtract stools/day}] \times [2 \text{ stools/day}] \times [-1.7 \times 10^{-2}] \\ = & -2.8175 \approx -2.82. \end{aligned}$$

Now convert the linear predictor into percent risk.

$$\text{Risk} = 100\% \times 1 / [1 + \exp(-\text{lp})] = 5.6\% \text{ [expected risk is approximately 5.6\%].}$$

Table 5. Score chart for risk of CD-related surgery.

CD-Related Surgery																			
HBI score (without stool frequency) where					Patient's Age (years)		Gender												
Stool frequency is																			
(times/day)	0	2	4	6															
<i>and HBI value:</i>					20 to 29	6	Male	14											
0	0	5	10	15	30 to 39	5	Female	0											
1	8	12	16	20	40 to 49	4	<b>Disease Location</b>												
2	17	19	22	24	50 to 59	3	Colon ( <i>reference</i> )	0											
3	25	26	28	29	60 to 69	2	Colon and small bowel	12											
4	33	33	34	34	70 to 79	1	Small bowel	18											
5	42	41	39	38	80 years or older	0													
6	50	48	45	43	<b>Baseline antimetabolite use</b>		<b>Baseline 5-ASA use</b>												
7	58	55	51	47	No	16	No	21											
8	67	62	57	52	Yes	0	Yes	0											
9	75	69	63	57	<b>New fistula, abscess or definite abdominal mass at baseline</b>														
10	83	76	69	61			No	0											
11	92	83	74	66			Yes	58											
12	100	90	80	70															
<b>Total the points, then match the total with the predicted risk using the below scale.</b>																			
<26	41	52	61	68	74	80	89	106	119	130	139	148	156	163	171	178	186	194	>203
<2%	3%	4%	5%	6%	7%	8%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	>70%

CD, Crohn's disease; HBI, HarveyBradshaw Index; 5-ASA, 5-aminosalicylate.

predicting CD-related complications within 24 months, using data from one of the largest clinical trials conducted in CD to date. These models have been developed using patients from the REACT trial, and are thus intended for use in a broad cross-section of patients found in community gastroenterology clinics and not solely with regard to a specific disease duration or severity. Although both models demonstrated high specificity and negative predictive values, the surgery model had relatively greater predictive ability, likely because surgery is an easily defined and more objective clinical outcome, compared with disease-related complications. Additionally, the development of a score chart using basic clinical variables which are readily calculable facilitates implementation of this score in the outpatient clinic setting, potentially allowing for enhanced decision making between patients and physicians. If validated in an external dataset, the use of this model may help identify patients at higher risk of disease-related complications and especially surgery, who may benefit most from more intensive treatment, combination therapy, and/or closer monitoring for symptoms that require planned

surgery [eg. obstructive symptoms in the setting of fibrostenotic disease]. Additionally, all identified baseline predictors in our models were prospectively defined and measured. We believe this is a distinct advantage to retrospective cohort studies in which a greater potential exists for bias during the data acquisition process.

Existing prediction models for CD-related outcomes [summarised in Supplementary Table 5, available at *ECCO-JCC* online] have been limited by selection or referral centre bias, small sample size, evolving drug treatments and management strategies, long-term time horizons, or broadly inclusive outcome definitions. Furthermore, the time frame in which these models were produced must be considered. Since the greatest risk of complications occurs within the first few years after diagnosis, long-term risk prediction models eventually become confounded or outdated with evolving therapies and treatment strategies. For example, while some models were developed in the pre-biologic era, others define the need for thiopurines as advanced disease; but currently thiopurine monotherapy represents a treatment strategy generally reserved for less severe disease rather

**Table 6.** Score chart for risk of CD-related complications.

<b>CD-Related Complications</b>																		
<b>HBI score (without stool frequency) where</b>																		
<b>Stool frequency is</b>																		
<b>(times/day)</b>	0	2	4	6	<b>Patient's Age (years)</b>		<b>Gender</b>											
<b>and HBI value:</b>					20 to 29	45	Male	0										
0	0	11	22	33	30 to 39	37	Female	14										
1	8	17	26	35	40 to 49	30	<b>Disease Location</b>		Colon ( <i>reference</i> )	0								
2	17	24	31	38	50 to 59	22	Colon and small bowel		18									
3	25	30	35	40	60 to 69	15	Small bowel		13									
4	33	36	39	42	70 to 79	7	<b>Baseline antimetabolite use</b>		<b>Baseline anti-TNF use</b>									
5	42	43	44	45	80 years or older	0	No	12	No	0								
6	50	49	48	47	<b>New fistula, abscess or definite abdominal mass at baseline</b>		Yes	0	Yes	28								
7	58	55	52	49	No	0	<b>Baseline corticosteroid use</b>		No	0								
8	67	62	57	52	Yes	79	No	0	Yes	18								
9	75	68	61	54	<b>Total the points, then match the total with the predicted risk using the below scale.</b>													
10	83	74	65	56	0	37	65	88	109	127	144	161	177	193	210	227	246	>266
11	92	81	70	59	<10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	>75%
12	100	87	74	61														

CD, Crohn's disease; HBI, Harvey-Bradshaw Index; TNF, tumour necrosis factor.

than severe disease.<sup>55,56</sup> Nevertheless, it is notable that many of the items we identified as independent predictors have previously been shown to have prognostic value in population-based cohorts with longer-term follow up,<sup>1,4,57-60</sup> supporting the validity of our results. Prediction of CD-related complications is limited by the use of non-standard or broadly inclusive outcome definitions [Supplementary Table 5], thus preventing meaningful comparisons with our model. This may also explain the lower performance of the model for prediction of CD-related complications.

In the present study, the surgery model has shown promise to accurately predict an adverse outcome which may be useful to guide management strategies and patient care. Such a model could also reasonably be used for risk stratification of patients in clinical studies.

A major strength of the current study is use of data from a large randomised controlled trial [RCT] conducted across community

centres, meaning that the endpoints were well defined, complete to 24 months of follow-up, and adjudicated by an independent panel. The pragmatic nature of the trial design meant that consecutive patients were enrolled, regardless of disease activity, duration, phenotype, or treatment, and an algorithm of care applied which is broadly representative of the patient population seen in clinical practice. Thus, we believe these models may be more generalisable to the routine clinical care of adult patients with CD, and can be applied within the community clinic, regardless of disease duration, in keeping with the design of the trial from which the model and data were derived.

It is interesting to point out that smoking is commonly regarded as a poor prognostic factor for Crohn's disease and especially for postoperative recurrence. In the present study, we found that smoking had limited predictive ability for CD-related complications. Another important consideration with smoking is that there is a

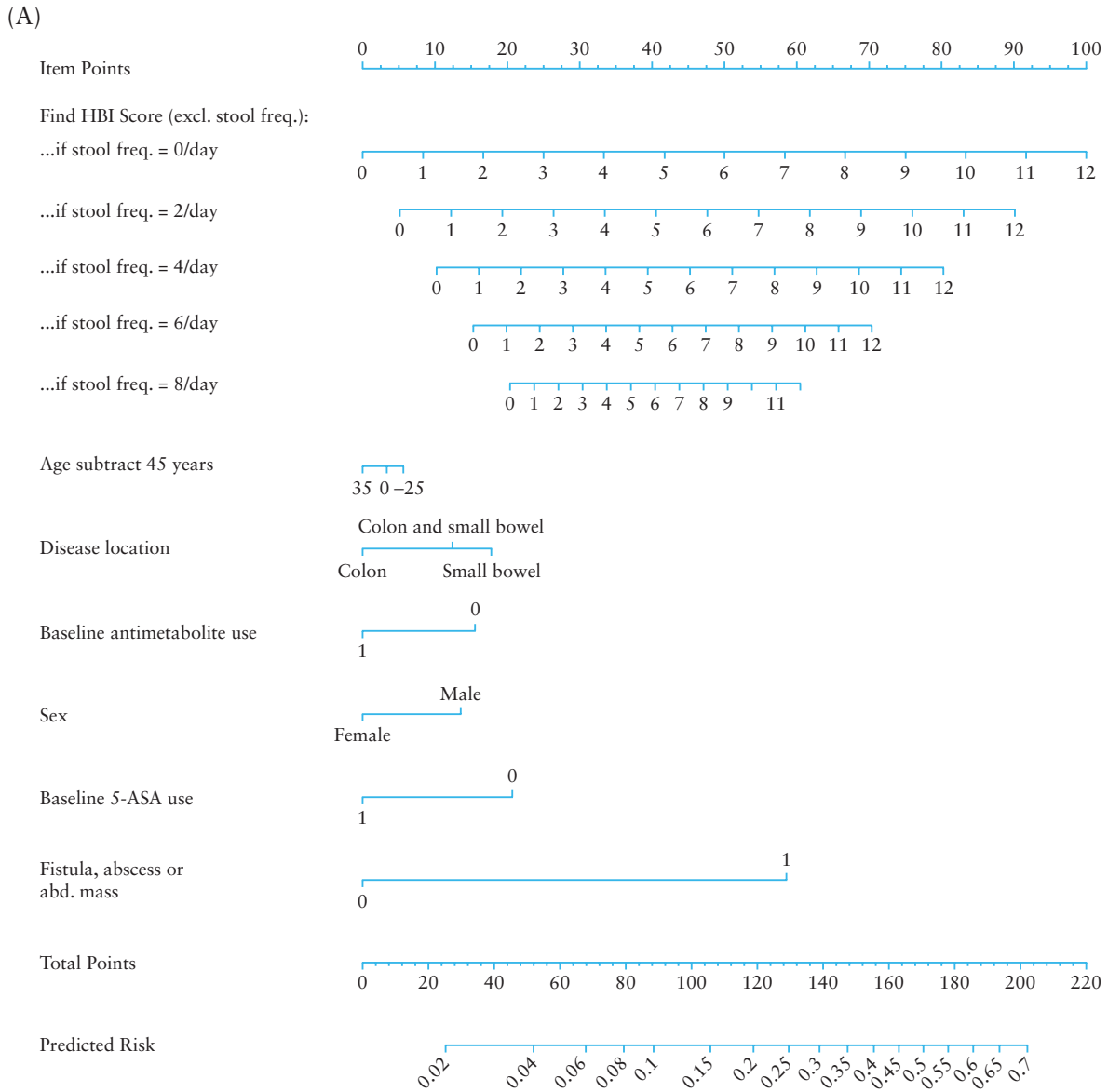


Figure 1. Nomogram for the computation of percentage risk of Crohn's disease-related surgery [A] or complication [B].

degree of subjective reporting, which can lead to inaccurate categorisation of current/recent/past smoking status as well as confounding with the quantity of smoking. Thus, we did not retain this in the final prediction models, which did not appreciably affect predictive performance.

Limitations of this study should be acknowledged. Perhaps the greatest limitation is that these models were validated in the cohort in which they were developed. Thus, independent validation in an external dataset is required before the widespread adoption of any prediction model can be considered. The results presented here encourage seeking such external validation. Participants in the REACT study had a longer average duration of disease, though approximately 15% and 28% were diagnosed within 2 and 5 years, respectively. However, the operating characteristics of these models in only newly diagnosed patients remain to be determined. Indeed, the REACT design was intended to include a broad cross-section of patients within community clinics, so that we may not be limited

to specific subsets of disease or demographic characteristics, and applies more generally to a broader patient population.

The CD-related surgery model performed better than the complication model, which may be related to using a more well-defined and objective outcome definition. Although the models use routinely available and clearly defined clinical data collected by trained gastroenterologists, inter-observer variability may influence the validity of the model. Also, the REACT study did not include biomarkers or serological or genetic factors; therefore these factors could not be incorporated into the current models. These factors may provide additional prognostic value beyond routine clinical data.<sup>41</sup>

Related to this issue of validation is discrimination performance. Prediction of CD-related complications is relatively difficult, as reflected in its low discriminative ability. However, the surgery model performs nominally as well as the model by Siegel and colleagues,<sup>43</sup> using completely different modelling approaches and relying solely on readily available clinical variables. Nevertheless, we are



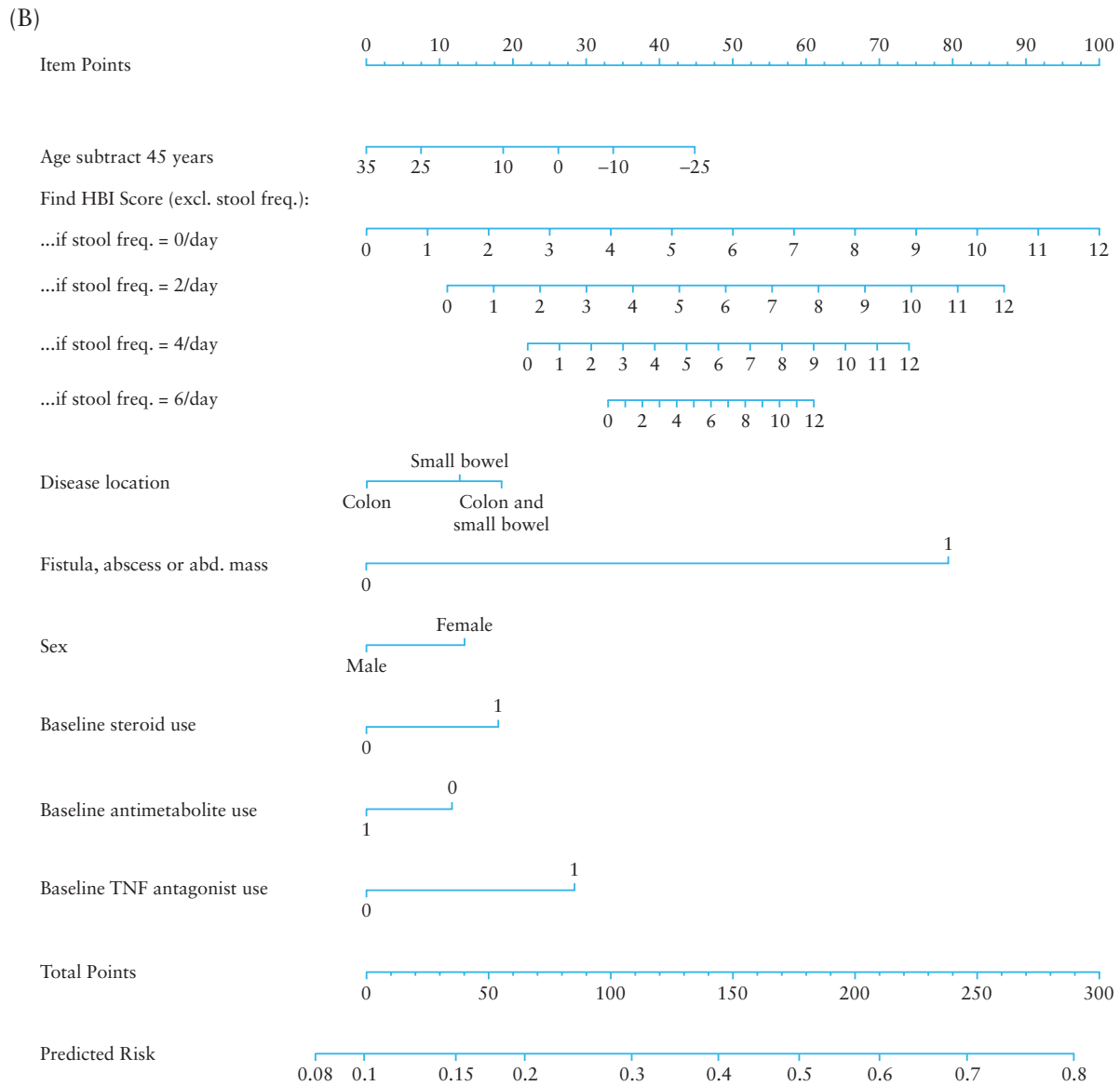


Figure 1. Continued

cautiously optimistic that the surgery model may prove useful once externally validated.

The treatment allocation was not included in our models, which some may view as a limitation. However, this decision was made a priori for two reasons. First, our primary interest was to develop prediction models using only baseline predictive factors that are available in practice. Second, the effect of treatment to predict these outcomes was found to be relatively small as compared with the factors in our prediction models. This is consistent with prediction literature in other areas, such as cardiovascular disease.<sup>61</sup>

In summary, we have developed and validated clinical prediction models for CD-related surgery and complications using data from a large randomised clinical trial with good overall performance for predicting the outcome of disease-related surgery within 2 years. We have transformed these models into scoring tools, facilitating their use in the clinic; external validation in independent cohorts with longitudinal follow up is recommended.

### Funding

This work was not specifically supported by any funding agency in the public, commercial, or not-for-profit sectors.

### Conflict of Interest

RK: honoraria from AbbVie, Janssen, Pfizer, Shire, and Takeda Pharma. PD: research support funding from Takeda and Pfizer, and a training grant through the National Institute of Diabetes and Digestive and Kidney Diseases [NIDDK T32DK007202]. WJS: grant support from Exact Sciences, the American College of Gastroenterology, and the Broad Foundation; grant support and personal fees from Receptos, Amgen, Prometheus Laboratories, AbbVie, Boehringer Ingelheim, Takeda, Atlantic Pharmaceuticals, Janssen, Bristol-Myers Squibb, Genentech, Pfizer, and Nutrition Science Partners; and personal fees from Kyowa Hakko Kirin, Millennium Pharmaceuticals, Celgene Cellular Therapeutics, Santarus, Salix Pharmaceuticals, Catabasis Pharmaceuticals, Vertex Pharmaceuticals, Warner Chilcott, Gilead Sciences, Cosmo Pharmaceuticals, Ferring Pharmaceuticals, Sigmoid Biotechnologies, Tillotts Pharma, Am Pharma BV, Dr. August Wolff, Avaxia Biologics,

Zyngenia, Ironwood Pharmaceuticals, Index Pharmaceuticals, Nestle, Lexicon Pharmaceuticals, UCB Pharma, Orexigen, Luitpold Pharmaceuticals, Baxter Health care, Ferring Research Institute, Amgen, Novo Nordisk, Mesoblast Inc., Shire, Ardelyx Inc., Actavis, Seattle Genetics, MedImmune [AstraZeneca], Actogenix NV, Lipid Therapeutics GmbH, Eisai, Qu Biologics, Toray Industries Inc., Teva Pharmaceuticals, Eli Lilly, Chiasma, TiGenix, Adherion Therapeutics, Immune Pharmaceuticals, Celgene, Arena Pharmaceuticals, Ambrx Inc., Akros Pharma, Vascular Biogenics, Theradiag, Forward Pharma, Regeneron, Galapagos, Seres Health, Ritter Pharmaceuticals, Theravance, Palatin, Biogen, and the University of Western Ontario [owner of Robarts Clinical Trials Inc]. VJ: scientific advisory board fees from Abbvie, Sandoz, Ferring and Janssen; speaker's fees from Takeda and Ferring; and travel support for conference attendance from Vifor pharmaceuticals. BGF: grant/research support from Millennium Pharmaceuticals, Merck, Tillotts Pharma AG, AbbVie, Novartis Pharmaceuticals, Centocor Inc., Elan/Biogen, UCB Pharma, Bristol-Myers Squibb, Genentech, ActoGenix, and Wyeth Pharmaceuticals Inc.; consulting fees from Millennium Pharmaceuticals, Merck, Centocor Inc., Elan/Biogen, Janssen-Ortho, Teva Pharmaceuticals, Bristol-Myers Squibb, Celgene, UCB Pharma, AbbVie, Astra Zeneca, Serono, Genentech, Tillotts Pharma AG, Unity Pharmaceuticals, Albireo Pharma, Given Imaging Inc., Salix Pharmaceuticals, Novonordisk, GSK, Actogenix, Prometheus Therapeutics and Diagnostics, Athersys, Axcan, Gilead, Pfizer, Shire, Wyeth, Zealand Pharma, Zyngenia, GiCare Pharma Inc., and Sigmoid Pharma; and speaker fees from UCB, AbbVie, and J&J/Janssen.

## Author Contributions

LG: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; statistical analysis. GZ: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; statistical analysis; study supervision. RK: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; study supervision. PD: critical revision of the manuscript for important intellectual content. WJS: critical revision of the manuscript for important intellectual content. VJ: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; study supervision. BGF: study concept and design; acquisition of data; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; study supervision. All authors have approved the final version of this manuscript.

## Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

## References

- Henriksen M, Jahnsen J, Lygren I, et al.; Ibsen Study Group. Clinical course in Crohn's disease: results of a five-year population-based follow-up study [the IBSen study]. *Scand J Gastroenterol* 2007;42:602–10.
- Peyrin-Biroulet L, Loftus EV Jr, Colombel JF, Sandborn WJ. The natural history of adult Crohn's disease in population-based cohorts. *Am J Gastroenterol* 2010;105:289–97.
- Cosnes J, Gower-Rousseau C, Seksik P, Cortot A. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* 2011;140:1785–94.
- Peyrin-Biroulet L, Harmsen WS, Tremaine WJ, Zinsmeister AR, Sandborn WJ, Loftus EV. Surgery in a population-based cohort of Crohn's disease from Olmsted County, Minnesota [1970–2004]. *Am J Gastroenterol* 2012;107:1693–701.
- Solberg IC, Vatn MH, Høie O, et al.; IBSen Study Group. Clinical course in Crohn's disease: results of a Norwegian population-based ten-year follow-up study. *Clin Gastroenterol Hepatol* 2007;5:1430–8.
- Colombel JF, Sandborn WJ, Rutgeerts P, et al. Adalimumab for maintenance of clinical response and remission in patients with Crohn's disease: the CHARM trial. *Gastroenterology* 2007;132:52–65.
- Hanauer SB, Feagan BG, Lichtenstein GR, et al.; ACCENT I Study Group. Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. *Lancet* 2002;359:1541–9.
- Rutgeerts P, Feagan BG, Lichtenstein GR, et al. Comparison of scheduled and episodic treatment strategies of infliximab in Crohn's disease. *Gastroenterology* 2004;126:402–13.
- Sandborn WJ, Feagan BG, Stoinov S, et al.; PRECISE 1 Study Investigators. Certolizumab pegol for the treatment of Crohn's disease. *N Engl J Med* 2007;357:228–38.
- Sandborn WJ, Hanauer SB, Rutgeerts P, et al. Adalimumab for maintenance treatment of Crohn's disease: results of the CLASSIC II trial. *Gut* 2007;56:1232–9.
- Sandborn WJ, Lee SD, Randall C, et al. Long-term safety and efficacy of certolizumab pegol in the treatment of Crohn's disease: 7-year results from the PRECISE 3 study. *Aliment Pharmacol Ther* 2014;40:903–16.
- Schreiber S, Colombel JF, Bloomfield R, et al.; PRECISE 2 Study Investigators. Increased response and remission rates in short-duration Crohn's disease with subcutaneous certolizumab pegol: an analysis of PRECISE 2 randomised maintenance trial data. *Am J Gastroenterol* 2010;105:1574–82.
- Schreiber S, Khaliq-Kareemi M, Lawrance IC, et al.; PRECISE 2 Study Investigators. Maintenance therapy with certolizumab pegol for Crohn's disease. *N Engl J Med* 2007;357:239–50.
- Sandborn WJ, Colombel JF, Enns R, et al.; International Efficacy of Natalizumab as Active Crohn's Therapy [ENACT-1] Trial Group; Evaluation of Natalizumab as Continuous Therapy [ENACT-2] Trial Group. Natalizumab induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2005;353:1912–25.
- Sandborn WJ, Feagan BG, Rutgeerts P, et al.; GEMINI 2 Study Group. Vedolizumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2013;369:711–21.
- Sands BE, Feagan BG, Rutgeerts P, et al. Effects of vedolizumab induction therapy for patients with Crohn's disease in whom tumor necrosis factor antagonist treatment failed. *Gastroenterology* 2014;147:618–27.e3.
- Targan SR, Feagan BG, Fedorak RN, et al.; International Efficacy of Natalizumab in Crohn's Disease Response and Remission [ENCORE] Trial Group. Natalizumab for the treatment of active Crohn's disease: results of the ENCORE Trial. *Gastroenterology* 2007;132:1672–83.
- Vermeire S, Loftus EV, Colombel J-F, et al. Long-term efficacy of vedolizumab for Crohn's disease. *J Crohns Colitis* 2017;11:412–24.
- Colombel JF, Sandborn WJ, Reinisch W, et al.; SONIC Study Group. Infliximab, azathioprine, or combination therapy for Crohn's disease. *N Engl J Med* 2010;362:1383–95.
- D'Haens G, Baert F, van Assche G, et al.; Belgian Inflammatory Bowel Disease Research Group; North-Holland Gut Club. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet* 2008;371:660–7.
- Khanna R, Bressler B, Levesque BG, et al.; REACT Study Investigators. Early combined immunosuppression for the management of Crohn's disease [REACT]: a cluster randomised controlled trial. *Lancet* 2015;386:1825–34.
- Vermeire S, Noman M, Van Assche G, Baert F, D'Haens G, Rutgeerts P. Effectiveness of concomitant immunosuppressive therapy in suppressing the formation of antibodies to infliximab in Crohn's disease. *Gut* 2007;56:1226–31.
- Scott FI, Lichtenstein GR. Advances in therapeutic drug monitoring of biologic therapies in inflammatory bowel disease: 2015 in review. *Curr Treat Options Gastroenterol* 2016;14:91–102.
- Vande Casteele N, Ferrante M, Van Assche G, et al. Trough concentrations of infliximab guide dosing for patients with inflammatory bowel disease. *Gastroenterology* 2015;148:1320–9.e3.
- Colombel JF, Reinisch W, Mantzaris GJ, et al. Randomised clinical trial: Deep remission in biologic and immunomodulator naïve patients with Crohn's disease - a sonic post hoc analysis. *Aliment Pharmacol Ther* 2015;41:734–46.

26. Colombel JF, Rutgeerts PJ, Sandborn WJ, *et al.* Adalimumab induces deep remission in patients with Crohn's disease. *Clin Gastroenterol Hepatol* 2014;**12**:414–22.e5.
27. Feagan BG, Sandborn WJ, Gasink C, *et al.* UNITI-IM-UNITI Study Group. Ustekinumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2016;**375**:1946–60.
28. Lémann M, Mary JY, Duclos B, *et al.* Groupe d'Etude Therapeutique des Affections Inflammatoires du Tube Digestif [GETAID]. Infliximab plus azathioprine for steroid-dependent Crohn's disease patients: a randomised placebo-controlled trial. *Gastroenterology* 2006;**130**:1054–61.
29. Rutgeerts P, Van Assche G, Sandborn WJ, *et al.* EXTEND Investigators. Adalimumab induces and maintains mucosal healing in patients with Crohn's disease: data from the EXTEND trial. *Gastroenterology* 2012;**142**:1102–11.e2.
30. Benitez JM, Louis E. Can we predict the high-risk patient? *Dig Dis* 2014;**32**:328–36.
31. Nasir BF, Griffiths LR, Nasir A, *et al.* An envirogenomic signature is associated with risk of IBD-related surgery in a population-based Crohn's disease cohort. *J Gastrointest Surg* 2013;**17**:1643–50.
32. Beaugerie L, Seksik P, Nion-Larmurier I, Gendre JP, Cosnes J. Predictors of Crohn's disease. *Gastroenterology* 2006;**130**:650–6.
33. Loly C, Belaiche J, Louis E. Predictors of severe Crohn's disease. *Scand J Gastroenterol* 2008;**43**:948–54.
34. Ramadas AV, Gunesh S, Thomas GAO, Williams GT, Hawthorne AB. Natural history of Crohn's disease in a population-based cohort from Cardiff [1986–2003]: A study of changes in medical treatment and surgical resection rates. *Gut* 2010;**59**:1200–6.
35. Romberg-Camps MJ, Dagnelie PC, Kester AD, *et al.* Influence of phenotype at diagnosis and of other potential prognostic factors on the course of inflammatory bowel disease. *Am J Gastroenterol* 2009;**104**:371–83.
36. Ryan JD, Silverberg MS, Xu W, *et al.* Predicting complicated Crohn's disease and surgery: phenotypes, genetics, serology and psychological characteristics of a population-based cohort. *Aliment Pharmacol Ther* 2013;**38**:274–83.
37. Tarrant KM, Barclay ML, Frampton CM, Gearty RB. Perianal disease predicts changes in Crohn's disease phenotype-results of a population-based study of inflammatory bowel disease phenotype. *Am J Gastroenterol* 2008;**103**:3082–93.
38. Thia KT, Sandborn WJ, Harmsen WS, Zinsmeister AR, Loftus EV Jr. Risk factors associated with progression to intestinal complications of Crohn's disease in a population-based cohort. *Gastroenterology* 2010;**139**:1147–55.
39. Wolters FL, Russel MG, Sijbrandij J, *et al.* Phenotype at diagnosis predicts recurrence rates in Crohn's disease. *Gut* 2006;**55**:1124–30.
40. Yang CH, Ding J, Gao Y, Chen X, Yang ZB, Xiao SD. Risk factors that predict the requirement of aggressive therapy among Chinese patients with Crohn's disease. *J Dig Dis* 2011;**12**:99–104.
41. Dubinsky MC, Kugathasan S, Kwon S, *et al.* Multidimensional prognostic risk assessment identifies association between IL12B variation and surgery in Crohn's disease. *Inflamm Bowel Dis* 2013;**19**:1662–70.
42. Lakatos PL, Sipeki N, Kovacs G, *et al.* Risk matrix for prediction of disease progression in a referral cohort of patients with Crohn's disease. *J Crohns Colitis* 2015;**9**:891–8.
43. Siegel CA, Horton H, Siegel LS, *et al.* A validated web-based tool to display individualised Crohn's disease predicted outcomes based on clinical, serologic and genetic variables. *Aliment Pharmacol Ther* 2016;**43**:262–71.
44. Solberg IC, Cvancarova M, Vatn MH, Moum B; IBSEN Study Group. Risk matrix for prediction of advanced disease in a population-based study of patients with Crohn's Disease [the IBSEN Study]. *Inflamm Bowel Dis* 2014;**20**:60–8.
45. Collins GS, Reitsma JB, Altman DG, Moons KG; TRIPOD Group. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis [TRIPOD]: the TRIPOD statement. The TRIPOD Group. *Circulation* 2015;**131**:211–9.
46. Bouwmeester W, Twisk JW, Kappen TH, van Klei WA, Moons KG, Vergouwe Y. Prediction models for clustered data: comparison of a random intercept and standard regression model. *BMC Med Res Methodol* 2013;**13**:19.
47. Harrell FE Jr. *Regression Modeling Strategies*. New York, NY: Springer International Publishing; 2015.
48. Steyerberg EW. *Clinical Prediction Models*. New York, NY: Springer; 2009.
49. Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. *Med Decis Making* 1993;**13**:49–58.
50. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;**15**:361–87.
51. Efron B. How biased is the apparent error rate of a prediction rule? *J Am Stat Assoc* 1986;**81**:461–70.
52. Wynants L, Bouwmeester W, Moons KG, *et al.* A simulation study of sample size demonstrated the importance of the number of events per variable to develop prediction models in clustered data. *J Clin Epidemiol* 2015;**68**:1406–14.
53. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing; 2016.
54. R Studio Team. *Rstudio: Integrated Development Environment for R*. Vienna: R Foundation for Statistical Computing; 2015.
55. Dignass A, Van Assche G, Lindsay JO, *et al.* European Crohn's and Colitis Organisation [ECCO]. The second European evidence-based consensus on the diagnosis and management of Crohn's disease: Current management. *J Crohns Colitis* 2010;**4**:28–62.
56. Gomollón F, Dignass A, Annese V, *et al.* ECCO. 3<sup>rd</sup> European evidence-based consensus on the diagnosis and management of Crohn's disease 2016: Part 1: diagnosis and medical management. *J Crohns Colitis* 2017;**11**:3–25.
57. Bernell O, Lapidus A, Hellers G. Risk factors for surgery and postoperative recurrence in Crohn's disease. *Ann Surg* 2000;**231**:38–45.
58. Golovics PA, Lakatos L, Mandel MD, *et al.* Prevalence and predictors of hospitalisation in Crohn's disease in a prospective population-based inception cohort from 2000–2012. *World J Gastroenterol* 2015;**21**:7272–80.
59. Lovasz BD, Lakatos L, Horvath A, *et al.* Evolution of disease phenotype in adult and pediatric onset Crohn's disease in a population-based cohort. *World J Gastroenterol* 2013;**19**:2217–26.
60. Nguyen GC, Nugent Z, Shaw S, Bernstein CN. Outcomes of patients with Crohn's disease improved from 1988 to 2008 and were associated with increased specialist care. *Gastroenterology* 2011;**141**:90–7.
61. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;**35**:1925–31.