The Prognostic Power of the NOD2 Genotype for Complicated Crohn’s Disease: A Meta-Analysis

Jeremy Adler, MD, MSc1, Sujal C. Rangwalla, DO1, Ben A. Dwamena, MB, ChB2 and Peter D.R. Higgins, MD, PhD, MSc3

OBJECTIVES: Crohn’s disease is often purely inflammatory at presentation, but most patients develop strictures and fistulae over time (complicated disease). Many studies have suggested that nucleotide-binding oligomerization domain 2 (NOD2) mutations are associated with a varying but increased risk of complicated disease. An accurate and sufficiently powerful predictor of complicated disease could justify the early use of biological therapy in high-risk individuals. We performed a systematic review and meta-analysis to obtain accurate estimates of the predictive power of the identified mutations (such as p.R702W, P.G908R, and p.Leu1007fsX1008) in NOD2 for the risk of complicated disease.

METHODS: An electronic search of MEDLINE, Embase, and Web of Science identified 917 relevant papers. Inclusion required specification of genetic mutations at the individual level and disease phenotypes by Vienna classification (inflammatory (B1), stricturing (B2), and fistulizing (B3)). A total of 49 studies met these criteria, which included 8,893 subjects, 2,897 of whom had NOD2 mutations. Studies were weighted by median disease duration. Studies not providing duration data were weighted at the level of the study with the shortest disease duration (3.9 years).

RESULTS: The relative risk (RR) of the presence of any NOD2 mutant allele for complicated disease (B2 or B3) was 1.17 (95% confidence interval (95% CI) 1.10–1.24; P<0.001). P.G908R was associated with an RR of complicated disease of 1.33 (95% CI 1.11–1.60; P=0.002). NOD2 did not predict perianal disease (P=0.4). The RR of surgery was 1.58 (95% CI 1.38–1.80; P<0.001). There was substantial heterogeneity across all studies (I² = 66.7%). On the basis of logistic regression of these data, the sensitivity of any mutation in predicting complicated disease was 36% and specificity was 73%, with the area under the receiver operating characteristic curve 0.56.

CONCLUSIONS: The presence of a single NOD2 mutation predicted an 8% increase in the risk for complicated disease (B2 or B3), and a 41% increase with 2 mutations. Surgery risk is increased by 58% with any NOD2 mutation, whereas perianal disease was unchanged. The predictive power associated with a single NOD2 mutation is weak. The RR of any NOD2 mutations for complicated disease was only 17% across 36 studies. However, the presence of two NOD2 mutations had 98% specificity for complicated disease. These data provide insufficient evidence to support top-down therapy based solely on single NOD2 mutations, but suggest that targeted early-intensive therapy for high-risk patients with two NOD2 mutations might be beneficial, if prospective trials can demonstrate changes in the natural history in this subset of patients.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at http://www.nature.com/ajg

INTRODUCTION
Crohn’s disease is a chronic immune-mediated inflammatory condition of the gastrointestinal tract. Complex interactions between the immune system and the intestinal luminal flora trigger chronic inflammatory responses, which results in the development of intestinal fibrosis over time. Initially, the majority...
of patients develop a purely inflammatory presentation of disease (1,2). However, most patients eventually develop intestinal fibrosis, which leads to the development of severe complications, such as intestinal strictures and fistulae.

The rate at which individuals develop stricturing or penetrating complications varies widely. The rate of progression seems to be unpredictable, although many attempts have been made to define an "aggressive phenotype" (1,3–6). The goal of predicting disease phenotype is to identify patients at risk of more severe disease who may benefit from early aggressive therapy. There are two requisite conditions that must be fulfilled for a predictive strategy to be successful. First, the predictive test must be sensitive and specific for complicated disease. Second, the group predicted to have complicated disease must have a proven response to a cost-effective change in therapy that results in a significant change in the natural history of these patients.

Susceptibility genes for Crohn's disease have been recently elucidated. The first and most promising identified mutations are in the nucleotide-binding oligomerization domain 2 (NOD2), also known as the caspase recruitment domain 15 (CARD15) gene, located on chromosome 16q12. NOD2/CARD15 has been identified as a disease-susceptibility gene, which confers a risk for the development of Crohn's disease (7–12). Many studies have suggested that NOD2 mutations are associated with a varying but increased risk of complicated disease (10,13–15). Certain NOD2 mutations have been postulated to lead to more severe defects and thus predict more aggressive disease behavior (14,16–18). However, published data to date are conflicting as to the prognostic value of NOD2 genotyping (19–21).

Most studies to date have been focused on the risk of disease development, and do not specifically address disease severity. Relatively few studies have directly investigated the question of disease severity, although some of the large disease-susceptibility studies contain information on disease phenotype and severity. Published studies of disease phenotype have demonstrated varying degrees of risk attributable to NOD2 mutations (19,20). Some papers have focused on differences between specific NOD2 mutations, whereas others have focused on numbers of mutant alleles (10,22–24).

Predictive results vary widely across these studies. For genotyping patients to be useful in practice, an estimation of the risk of severe disease associated with a genotype would have to be clearly defined. If a method of predicting disease severity was accurate, the benefits to patient care would be tremendous. An accurate and sufficiently powerful predictor of complicated disease, followed by targeted, cost-effective intervention for these patients could potentially change the natural history of Crohn's disease, and justify early use of expensive biological therapies in high-risk individuals.

Three major polymorphisms in the NOD2 gene have been reported to be associated with Crohn's disease. These single-nucleotide polymorphisms include p.R702W, p.G908R, and p.Leu1007fsX1008. We performed a systematic review of the literature with meta-analysis to obtain accurate estimates of the predictive power of these identified mutations (p.R702W, p.G908R, and p.Leu1007fsX1008) in NOD2 for the risk of complicated disease. Other minor alleles have been reported, and where sufficient data exist, we attempted to include these data as well. We defined complicated disease as stricturing and/or fistulizing Crohn's disease for the purposes of this study.

METHODS

Search strategies

This systematic review of the literature and meta-analysis was performed and reported according to standard guidelines for systematic review and meta-analysis of diagnostic studies (25–28). We conducted a computer-aided literature search of MEDLINE database (Ovid Technologies, New York, NY) from inception to 31 December 2009. The search strategy used free-text words and MeSH terms ("genotype" or "genotype.mp" or nod2 or card15 or "Nod2 Signaling Adaptor Protein/" and ("exp Crohn Disease/" or "exp Inflammatory Bowel Diseases/") and ("Phenotype/" or "phenotype.mp" or "exp Intestinal Obstruction/" or "Constriction, Pathologic/" or "stenosis.mp" or "stricture.mp" or "fibrostenosing.mp" or "stricturing.mp" or "exp Fistula/" or "exp Intestinal Fistula/" or "exp Digestive System Fistula/" or "exp Fistula/" or "exp Gastric Fistula/" or "exp Rectovaginal Fistula/" or "exp Rectal Fistula/" or "penetrating.mp")).

We conducted an additional computer-aided literature search of the Embase database (Elsevier, New York, NY) from inception to 31 December 2009. The search strategy was as follows: (nod2 or card15) and (stricture or strictureting or fistula/exp or fistula or fistulating or penetrating or 'phenotype' or phenotype or 'stenosis' or 'stenosis' or 'stricturing' or 'exp Fistula/').

The search was not restricted to English-language publications. In addition, after study selection, we reviewed the bibliographies of the studies to identify any additional studies relevant to this topic. Additional searches through ISI Web of Science (Thomson Reuters, New York, NY) and Google Scholar (Google, Mountain View, CA) were performed to identify any additional publications that may have been missed with the aforementioned searches.

Study selection

Two independent reviewers (J.A. and S.C.R.) read the titles of all candidate articles. For those articles that were not excluded based solely on the title, we reviewed the abstract. We then retrieved the full text of published manuscripts that could not be excluded based on the title and abstract alone. The articles were read and checked for inclusion criteria independently. All discrepancies were resolved by consensus. The final agreement between reviewers was 100% for study selection.

Study inclusion/exclusion criteria

The selection criteria for inclusion of a study were: (i) clear statement of disease definition; (ii) clear statement of genotyping and which mutations were studied; (iii) classification of disease phenotype by Vienna or Montreal classification with non-overlapping phenotypes (6,30); (iv) documentation of genotype by phenotype on individual level data (not population gene frequency data); and (v) if raw numbers of each genotype–phenotype combination could be found or clearly calculated from the published data.
Exclusion criteria were: (i) duplicate publications of a primary study that contained all or a subset of the original data; (ii) studies presenting only gene frequency data; and (iii) studies involving patients with ulcerative colitis or indeterminate colitis. In cases of potential or suspected overlapping subject populations (based either on location or authorship), the latest publication or the study with the largest numbers of patients was included in the meta-analysis. All others were excluded.

Data extraction
Two independent reviewers (J.A. and S.C.R.) extracted the following data from the selected studies onto data abstraction forms: study characteristics (such as design, country, year of publication, setting/region, sample size, phenotype data, genotype data, age at diagnosis, disease duration), population (mean age, sex, number of patients with each disease phenotype, criteria for diagnosis, genotyping technique, criteria for definition of phenotype), and outcomes (numbers of each NOD2 allele mutation and genotype–phenotype combinations). Other variables that were sought included methods for handling indeterminate or missing data. Discrepancies were resolved by discussion and consensus with the senior author (P.D.R.H.). We included two Spanish-language papers because one author (J.A.) is literate in Spanish. We excluded one Polish, two Hungarian, and one German language paper as we had no available resources to assist in translation.

After completion of manuscript review and data extraction, all included studies were reviewed for assessment of study quality. Manuscripts were assessed according to a modified subset of the QUADAS (Quality Assessment Tool for Diagnostic Accuracy) guidelines (31).

Statistical analysis
The primary outcome for this analysis was the diagnostic test performance of NOD2 genotyping for predicting Crohn’s disease phenotype. We applied the bivariate random-effects regression model for treatment trial meta-analysis and modified for synthesis of diagnostic test data assuming a binomial error distribution for sensitivity and specificity (32,33). Sensitivity, specificity, and diagnostic odds ratios with 95% confidence intervals (95% CIs) were calculated. The diagnostic odds ratio is defined as the odds of having a positive test result (mutant allele present) in patients with disease finding (fistulizing disease, etc.) compared with the odds of a positive test result in patients without the disease (no fistulizing disease). The diagnostic odds ratio is a single indicator of test accuracy that combines data from sensitivity and specificity (34).

The heterogeneity of the results between the studies was assessed graphically by forest plots, and statistical assessment was performed with the $\chi^2$ test of homogeneity and the inconsistency index ($I^2$) and Cochran’s $Q$-test for non-combinability. The $I^2$ index describes the percentage of total variation across studies that are due to heterogeneity rather than chance. A value between 50 and 75 represents moderate heterogeneity (35). A non-significant Cochran $Q$ represents minimal heterogeneity, indicating that the studies can be combined.

The overall performance of NOD2 was assessed with summary receiver operating characteristic (ROC) curve. We analyzed the data using the midas (Meta-analytical Integration of Diagnostic Accuracy Studies) command (36,37). All statistical analyses were performed using Stata 11.0 for Mac (StataCorp LP, College Station, TX).

RESULTS

Literature search
The results of the literature search are summarized in Figure 1. Two independent reviewers (J.A. and S.C.R.) identified abstracts and titles of 917 studies for initial review based on the described search strategies. A total of 357 studies were chosen for detailed review, and 49 studies that met the inclusion and exclusion criteria were included in the meta-analysis. There was 100% agreement between reviewers on the selection of studies.

Excluded studies
Upon review of the title and abstract, 560 studies were excluded for not meeting the pre-specified inclusion criteria. An additional 308 studies were excluded after full review of the published manuscript. Upon full review of the manuscripts, 96 were identified as not being studies of Crohn’s disease. A total of 84 studies investigated other genetic mutations than those of NOD2/CARD15. In all, 34 studies were excluded for lack of phenotypic information. An additional 95 studies were excluded for inadequate documentation of genotype–phenotype combinations. Overall, 58 studies were excluded for not providing original data, being individual case reports, review articles, commentaries, or small meta-analyses. Six manuscripts contained incomplete or inconsistent data. An additional 18 studies included only gene frequency or otherwise did not include patient level data. One paper was excluded as it was a single patient case report.

Included studies
A total of 49 studies included in this meta-analysis comprised 8,893 patients with Crohn’s disease, of whom 2,897 had NOD2 mutations (Table 1). Of the 33 studies documenting disease duration, these data included 64,955 patient-years of disease duration. Studies were weighted by median disease duration. The 16 studies not providing disease duration data were weighted at the level of the study with the shortest documented median disease duration (3.9 years).

Quality assessment
The selection criteria were clearly documented in 10 out of 49 studies. The description of phenotype was clearly documented in 12 studies. The description of the method for genotyping was present in all but six papers. The spectrum of patients was generally reflective of clinical practice, except in four papers. Explicit documentation of how missing data were handled was present in only two papers, whereas eight papers clearly had missing data that were unaccounted for. Only 11 papers included documentation that those assessing phenotypes were blinded to the results of genotyping. A summary of the QUADAS criteria is depicted in Supplementary Figure 1 online.
All studies included were assessed for quality. Specific items assessed included the presence of a clear description of selection criteria. The description of phenotyping was recorded as the reference standard.

**Risk of complicated disease**

The risk of complicated disease (stricturing or fistulizing) was compared between patients with various predetermined combinations of NOD2 mutations (Table 2). The pooled relative risk (RR) of complicated disease with the presence of any (one or more) NOD2 mutant allele was 1.17 (95% CI 1.10–1.24; P < 0.001; Figure 2a) across 36 studies with a total of 6,115 patients. There was moderate evidence of heterogeneity between studies (I² = 66.7%; Cochran’s Q 59.7, P = 0.003). Study-specific sensitivities are shown in Figure 2b, with a pooled sensitivity of 0.36 (95% CI 0.31–0.40) and a specificity of 0.73 (95% CI 0.69–0.77).

A dose effect was seen with the presence of one NOD2 mutant allele (see Supplementary Figure 2 online) associated with an RR of complicated disease of 1.08 (95% CI 0.96–1.21; P = 0.2; F = 44.6%; Q = 18.4, P = 0.05) across 13 studies, whereas the presence of 2 NOD2 mutant alleles (see Supplementary Figure 3 online) had an RR of 1.41 (95% CI 1.26–1.57; P < 0.001; F = 55.0%; Q = 12.5, P = 0.2) with a sensitivity of 0.11 (95% CI 0.06–0.21) and a specificity of 0.98 (95% CI 0.92–0.99) for complicated disease. The presence of homozygous mutations of NOD2 was associated with an RR of complicated disease of 1.45 (95% CI 1.31–1.61; P < 0.001; F = 93.8%; Q = 4.5, P = 0.2), with a sensitivity of 0.05 (95% CI 0.01–0.16) and a specificity of 1.00 (95% CI 0.94–1.00) across 7 studies (see Supplementary Figure 4 online).

When specific individual mutant alleles were assessed, p.G908R was the only allele found to have a significant association with complicated disease. In 10 studies with sufficient data to assess its effect, p.R702W was found to confer an RR of complicated disease of 1.11 (95% CI 0.92–1.36; P = 0.3; F = 67.2%; Q = 18.3, P = 0.02), with a sensitivity of 0.28 (95% CI 0.09–0.62) and a specificity of 0.86 (95% CI 0.79–0.91). In 10 studies, p.G908R was found to be associated with an RR of complicated disease of 1.33 (95% CI 1.11–1.60; P = 0.002; F = 34.4%; Q = 9.0, P = 0.3) with a sensitivity of 0.1 (95% CI 0.07–0.13) and a specificity of 0.93 (95% CI 0.88–0.96). In 13 studies, p.Leu1007fsX1008 was associated with an RR of complicated disease of 1.07 (95% CI 0.94–1.21; P = 0.3; F = 66.4%; Q = 25.0, P = 0.009) with a sensitivity of 0.31 (95% CI 0.11–0.63) and a specificity of 0.89 (95% CI 0.84–0.92) (see Supplementary Figure 5 online).

When homozygous mutations were assessed for individual mutant alleles, p.Leu1007fsX1008 was the only allele found to have a significant association with complicated disease. In 3 studies, p.R702W was found to confer an RR of complicated disease of 1.18 (95% CI 0.31–4.57; P = 0.81; F = 0.0%; Q = 0.02, P = 0.88). In 3 studies, p.G908R was found to confer an RR of complicated disease of 1.70 (95% CI 0.77–3.76; P = 0.19; F = 0.0%; Q = 0.09, P = 0.76). In 5 studies, p.Leu1007fsX1008 was found to be associated with an RR of complicated disease of 1.30 (95% CI 1.10–1.54; P = 0.002; F = 17.6%; Q = 3.21, P = 0.36) with a sensitivity of 0.03 (95% CI 0.01–0.11) and a specificity of 0.99 (95% CI 0.96–100) (23,38–41).
Table 1. Study characteristics

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Number of patients</th>
<th>Median duration of disease (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmad et al.</td>
<td>2002</td>
<td>England</td>
<td>244</td>
<td>16.2</td>
</tr>
<tr>
<td>Alvarez-Lobos et al.</td>
<td>2007</td>
<td>Spain</td>
<td>170</td>
<td>6</td>
</tr>
<tr>
<td>Annese et al.</td>
<td>2005</td>
<td>Italy</td>
<td>316</td>
<td>8.0</td>
</tr>
<tr>
<td>Brant et al.</td>
<td>2003</td>
<td>USA</td>
<td>214</td>
<td>14.9</td>
</tr>
<tr>
<td>Bianchi et al.</td>
<td>2007</td>
<td>Italy</td>
<td>177</td>
<td>8.5</td>
</tr>
<tr>
<td>Büning et al.</td>
<td>2004</td>
<td>Germany</td>
<td>180</td>
<td>6.4</td>
</tr>
<tr>
<td>Canto et al.</td>
<td>2007</td>
<td>Spain</td>
<td>97</td>
<td>NR</td>
</tr>
<tr>
<td>Chua et al.</td>
<td>2009</td>
<td>Malaysia</td>
<td>45</td>
<td>NR</td>
</tr>
<tr>
<td>Crawford et al.</td>
<td>2007</td>
<td>USA</td>
<td>205</td>
<td>4.6</td>
</tr>
<tr>
<td>Cucchiara et al.</td>
<td>2007</td>
<td>Italy</td>
<td>196</td>
<td>9</td>
</tr>
<tr>
<td>Cukovic-Cavka et al.</td>
<td>2006</td>
<td>Croatia</td>
<td>136</td>
<td>10</td>
</tr>
<tr>
<td>Dassopoulos et al.</td>
<td>2009</td>
<td>USA</td>
<td>183</td>
<td>11.1</td>
</tr>
<tr>
<td>De Diego et al.</td>
<td>2006</td>
<td>Spain</td>
<td>183</td>
<td>7.4</td>
</tr>
<tr>
<td>de Ridder et al.</td>
<td>2007</td>
<td>The Netherlands</td>
<td>72</td>
<td>NR</td>
</tr>
<tr>
<td>Esters et al.</td>
<td>2004</td>
<td>Belgium</td>
<td>344</td>
<td>NR</td>
</tr>
<tr>
<td>Fernandez et al.</td>
<td>2004</td>
<td>Spain</td>
<td>210</td>
<td>9.9</td>
</tr>
<tr>
<td>Ferreira et al.</td>
<td>2005</td>
<td>Portugal</td>
<td>231</td>
<td>9.7</td>
</tr>
<tr>
<td>Fidder et al.</td>
<td>2003</td>
<td>Israel</td>
<td>170</td>
<td>NR</td>
</tr>
<tr>
<td>Giachino et al.</td>
<td>2004</td>
<td>Italy</td>
<td>184</td>
<td>NR</td>
</tr>
<tr>
<td>Heresbach et al.</td>
<td>2004</td>
<td>France</td>
<td>205</td>
<td>7.1</td>
</tr>
<tr>
<td>Hsiao et al.</td>
<td>2007</td>
<td>Taiwan</td>
<td>10</td>
<td>4.7</td>
</tr>
<tr>
<td>Laghi et al.</td>
<td>2005</td>
<td>Italy</td>
<td>239</td>
<td>10.1</td>
</tr>
<tr>
<td>Lakatos et al.</td>
<td>2005</td>
<td>Hungary</td>
<td>527</td>
<td>8.2</td>
</tr>
<tr>
<td>Lappalainen et al.</td>
<td>2008</td>
<td>Finland</td>
<td>240</td>
<td>7.0</td>
</tr>
<tr>
<td>Lesage et al.</td>
<td>2002</td>
<td>Europe</td>
<td>453</td>
<td>11.0</td>
</tr>
<tr>
<td>Linskens et al.</td>
<td>2004</td>
<td>The Netherlands</td>
<td>108</td>
<td>16</td>
</tr>
<tr>
<td>Louis et al.</td>
<td>2003</td>
<td>Belgium</td>
<td>101</td>
<td>NR</td>
</tr>
<tr>
<td>Maconi et al.</td>
<td>2009</td>
<td>Italy</td>
<td>253</td>
<td>14.4</td>
</tr>
<tr>
<td>Mardini et al.</td>
<td>2005</td>
<td>USA</td>
<td>202</td>
<td>7.1</td>
</tr>
<tr>
<td>Marrakchi et al.</td>
<td>2009</td>
<td>Tunisia</td>
<td>75</td>
<td>NR</td>
</tr>
<tr>
<td>Mendoza et al.</td>
<td>2003</td>
<td>Spain</td>
<td>204</td>
<td>8</td>
</tr>
<tr>
<td>Protic et al.</td>
<td>2008</td>
<td>Serbia</td>
<td>131</td>
<td>7</td>
</tr>
<tr>
<td>Radlmayr et al.</td>
<td>2002</td>
<td>Germany</td>
<td>97</td>
<td>NR</td>
</tr>
<tr>
<td>Rigoli et al.</td>
<td>2008</td>
<td>Italy</td>
<td>133</td>
<td>NR</td>
</tr>
<tr>
<td>Rodrigo et al.</td>
<td>2007</td>
<td>Spain</td>
<td>183</td>
<td>12</td>
</tr>
<tr>
<td>Sabate et al.</td>
<td>2008</td>
<td>France</td>
<td>239</td>
<td>9.3</td>
</tr>
<tr>
<td>Schoepfer et al.</td>
<td>2009</td>
<td>Switzerland</td>
<td>252</td>
<td>NR</td>
</tr>
<tr>
<td>Seiderer et al.</td>
<td>2006</td>
<td>Germany</td>
<td>299</td>
<td>10.5</td>
</tr>
<tr>
<td>Shaoul et al.</td>
<td>2009</td>
<td>Israel</td>
<td>125</td>
<td>3.9</td>
</tr>
<tr>
<td>Smith et al.</td>
<td>2004</td>
<td>Scotland</td>
<td>143</td>
<td>5</td>
</tr>
<tr>
<td>Sun et al.</td>
<td>2003</td>
<td>Germany</td>
<td>55</td>
<td>NR</td>
</tr>
<tr>
<td>Tekin et al.</td>
<td>2009</td>
<td>Turkey</td>
<td>45</td>
<td>NR</td>
</tr>
</tbody>
</table>

Continued on following page
The risk of stricturing disease was compared between patients with various predetermined combinations of NOD2 mutations. The pooled RR of stricturing disease with the presence of any (one or more) NOD2 mutant allele was 1.16 (95% CI 1.05–1.28; P < 0.003; Figure 3b) across 36 studies. There was little evidence of heterogeneity between studies (I² = 27.3%; Cochran’s Q 38.2, P = 0.3). Study-specific sensitivities and specificities included a sensitivity of 0.34 (95% CI 0.30–0.38) and a specificity of 0.73 (95% CI 0.69–0.76). In 9 studies, p.G908R was found to be associated with an RR of fistulizing disease of 1.42 (95% CI 1.06–1.90; P = 0.02; I² = 26.2%; Q = 8.69, P = 0.28) with a sensitivity of 0.08 (95% CI 0.05–0.12) and a specificity of 0.94 (95% CI 0.89–0.96). In 12 studies, there was no attributable risk for fistulizing disease with p.R702W (P = 0.13) or p.Leu1007fsX1008 (P = 0.66) mutations.

The risk of perianal disease was reported separately from stricturing or fistulizing phenotype in 17 manuscripts (3,836 patients). When this end point was analyzed, there was no increased or decreased risk of perianal disease associated with the presence of any NOD2 mutant alleles (RR 1.03; 95% CI 0.92–1.16; P = 0.6; I² = 52.9%; Q = 33.6, P = 0.004; Figure 4a), with a pooled sensitivity of 0.38 (95% CI 0.32–0.43) and a specificity of 0.63 (95% CI 0.59–0.66). However, when individual mutations were considered, p.G908R was found to have a slightly decreased risk of perianal disease. In 4 studies (RR 0.40; 95% CI 0.19–0.83; P = 0.014; I² = 61.7%; Q = 4.4, P = 0.1), there was no attributable risk for perianal disease with either p.R702W (P = 0.3) or p.Leu1007fsX1008 (P = 0.5) mutations (see Supplementary Figure 6 online).

### Table 1. Continued

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Number of patients</th>
<th>Median duration of disease (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Torner et al. (84)</td>
<td>2003</td>
<td>USA</td>
<td>101</td>
<td>4.1</td>
</tr>
<tr>
<td>Törkvist et al. (85)</td>
<td>2006</td>
<td>Sweden</td>
<td>179</td>
<td>16.5</td>
</tr>
<tr>
<td>van der Linde et al. (86)</td>
<td>2007</td>
<td>The Netherlands</td>
<td>114</td>
<td>NR</td>
</tr>
<tr>
<td>Vavassori et al. (23)</td>
<td>2004</td>
<td>Italy</td>
<td>165</td>
<td>12.0</td>
</tr>
<tr>
<td>Vermeire et al. (87)</td>
<td>2002</td>
<td>Quebec</td>
<td>229</td>
<td>NR</td>
</tr>
<tr>
<td>Vind et al. (88)</td>
<td>2005</td>
<td>Denmark, Portugal</td>
<td>87</td>
<td>NR</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>8,883</td>
<td></td>
</tr>
</tbody>
</table>

NR, not reported.
LR was 0.88 (95% CI 0.84–0.93). A positive LR of $>10$ and a negative LR of $<0.1$ are generally considered useful predictors. As a prognostic indicator, NOD2 is a poor predictor of the complicated disease phenotype (Figure 5a).
Figure 2. Forest plots of NOD2 and complicated disease. (a) Forest plot of NOD2 mutant alleles and complicated disease showing study relative risk (RR) and 95% confidence interval (95% CI) of complicated disease. There was moderate heterogeneity between studies ($I^2 = 56.7\%$, $P = 0.000$). Study-specific mean sensitivity (left axis) and specificity (right axis) are shown in b. FN, false negative; FP, false positive; NOD2, nucleotide-binding oligomerization domain 2; TN, true negative; TP, true positive.
Figure 3. Forest plots of NOD2 and stricturing and fistulizing disease. Forest plot of the presence of one or more NOD2 mutant alleles predicting (a) stricturing disease and (b) fistulizing disease for Crohn’s disease. B1, inflammatory; B2, stricturing; B3, fistulizing phenotypes; NOD2, nucleotide-binding oligomerization domain 2; RR, risk ratio; 95% CI, 95% confidence interval.
Effect of diagnostic threshold

The sensitivity and specificity of the presence of any \(NOD2\) mutant alleles for complicated disease in each study are displayed on the summary ROC diagram. The summary ROC curve and the summary operating point (with 95% confidence contour) are shown in relation to the individual studies. It must be noted that a coin flip has a diagnostic accuracy of 0.5. The AuROC was 0.56 for complicated disease (Figure 5b) and 0.54 for perianal disease, consistent with low diagnostic accuracy. The AuROC for strictureing disease was 0.61, the AuROC for fistulizing disease was 0.54, and the AuROC was 0.63 for surgery for Crohn’s disease all of which represent poor diagnostic accuracy (see Supplementary Figure 7 online). The AuROC was 0.98 for p.Leu1007fsX1008 homozygous mutations predicting complicated disease. This represents excellent diagnostic accuracy; however, the positive LR is 2.6 (95% CI 0.4–16.6), and the negative LR is 0.98 (95% CI 0.94–1.03) (see Supplementary Figure 8 online).

DISCUSSION

\(NOD2/CARD15\) was identified to be associated with Crohn’s disease. There is abundant evidence supporting \(NOD2\) as a susceptibility allele conferring risk for developing Crohn’s disease. As the early studies of \(NOD2\), and with the later advent of other genes associated with Crohn’s disease, interest has grown toward identifying specific mutations that may be responsible for distinct disease behavior phenotypes.
There have been conflicting reports from early in the literature as to the prognostic value of NOD2. Abreu et al. (10) in 2002 found that NOD2 mutations were specifically associated with fibrostenosing disease and not penetrating disease. That same year, Lesage et al. (24), found no association between NOD2 mutations and specific disease phenotypes. Alvarez-Lobos et al. (13) and Ince et al. (21) showed an association between NOD2 mutations and increased surgical rates in Crohn’s disease, whereas Baptista et al. (19) found no measurable associations with disease behavior could be detected at all.

There have been a number of small meta-analyses that have looked to answer these questions, but no full review of the literature to date exists (22,42,43). We undertook a full systematic review of the published literature to gain a better understanding as to how the presence of NOD2 mutations modifies disease behavior.

This meta-analysis showed that the presence of a single NOD2/CARD15 mutation has poor predictive ability for disease phenotype in Crohn’s disease with an AuROC of 0.56. Summary estimates showed poor sensitivity and fairly low specificity for aggressive stenosing or fistulizing phenotypes. However, the presence of two mutations in NOD2 (either homozygous or complex heterozygous) has a high degree of specificity for aggressive disease phenotype, although the sensitivity and AuROC remain poor. The presence of homozygous mutations in NOD2 for p.Leu1007fsX1008 has an even higher degree of specificity for aggressive disease phenotype, although the sensitivity remains poor. These results show that the presence of multiple mutations identifies a particularly high-risk group, although this is a very insensitive prognostic test. In contrast, the presence of single mutations in NOD2 is not a reliable predictor of disease phenotype.

These findings do not minimize the role NOD2 and other genetic mutations have in the development of Crohn’s disease. NOD2 has a clear role as a susceptibility gene for the risk of development of Crohn’s disease. However, at this point, the ability of NOD2 to predict disease severity is limited to patients with multiple mutant alleles. It is clear that genetics has a key role in the susceptibility to developing Crohn's disease, and likely has a role in disease behavior. The pathophysiology is complex, and the role genetics has is far more nuanced than the presence or absence of one mutation determining a certain disease behavior. More likely, a reliable disease phenotype prediction will require a multifactorial approach, which will include environmental factors, genetic factors, medication exposure, smoking, initial clinical presentation, etc.

The findings of this meta-analysis suggest that the presence of two mutant NOD2 alleles is quite specific for future complicated disease. Although the sensitivity is poor and the AuROC is poor, the high degree of specificity may be sufficient to recommend testing for double mutations if we can identify therapies that can truly change the outcomes for this high-risk stratum of patients. Testing for double mutations would likely miss many patients with aggressive disease. However, the presence of double mutations, phenotype prediction will require a multifactorial approach, which will include environmental factors, genetic factors, medication exposure, smoking, initial clinical presentation, etc.

The findings of this meta-analysis suggest that the presence of two mutant NOD2 alleles is quite specific for future complicated disease. Although the sensitivity is poor and the AuROC is poor, the high degree of specificity may be sufficient to recommend testing for double mutations if we can identify therapies that can truly change the outcomes for this high-risk stratum of patients. Testing for double mutations would likely miss many patients with aggressive disease. However, the presence of double mutations, phenotype prediction will require a multifactorial approach, which will include environmental factors, genetic factors, medication exposure, smoking, initial clinical presentation, etc.

The American Journal of GASTROENTEROLOGY

© 2011 by the American College of Gastroenterology
particular mutations in p.Leu1007fsX1008, does identify a high-risk stratum of reagent-grade patients that should be identified and analyzed as a distinct subgroup in clinical trials. If prospective trials can show a change in outcomes in these patients, this would be a strong justification for targeted top-down therapy in this group of high-risk patients to prevent stricturing or fistulizing complications.

The other major finding of this meta-analysis is the lack of prognostic value of the presence of one mutant allele for predicting Crohn’s disease behavior. Although we specifically studied NOD2 mutations, recent genome-wide association studies have identified numerous other genes that contribute to the susceptibility of developing Crohn’s disease (44–46). It is likely that the level at which genetics contributes to phenotypic disease expression is complex and multiple risk alleles contribute to complex interactions. Further studies addressing disease behavior will likely require large-scale, multicenter trials to obtain sufficient clinical data to accurately ascertain the cumulative risk from clinical, genetic, and environmental factors. Once additional predictive factors are identified and validated, we will be better able to study targeted therapies in high-risk patients.

**CONFLICT OF INTEREST**

Guarantor of the article: Jeremy Adler, MD, MSc.

Specific author contributions: Study and concept design, literature search, data extraction, and statistical analysis: Jeremy Adler; secondary literature search and independent data extraction: Sujal C. Rangwalla; statistical analysis: Ben A. Dwamena; concept development and study design, and third reviewer when consensus could not be reached during data extraction: Peter D.R. Higgins.

Financial support: None.

Potential competing interests: None.

**REFERENCES**

29. van der Weijden T, Jozemans CJ, Dinant GJ et al. Identifying relevant diagnostic studies in MEDLINE. The diagnostic value of the erythrocyte sedimentation rate (ESR) and dipstick as an example. Fam Pract 1997;14:204–8.


