Penetrance of \textit{NOD2/CARD15} genetic variants in the general population

Shiva Yazdanyar MD, Pia R. Kamstrup MD PhD, Anne Tybjærg-Hansen MD DMSc, Børge G. Nordestgaard MD DMSc

Previously published at www.cmaj.ca

\textbf{ABSTRACT}

\textbf{Background:} In case–control studies of Europeans, heterozygosity for \textit{Arg702Trp(rs2066844)}, \textit{Gly908Arg(rs2066845)} and \textit{Leu1007fsinsC(rs5743293)} on the \textit{NOD2/CARD15} gene is associated with a 2-fold greater risk of Crohn disease, whereas homozygosity or compound heterozygosity is associated with a 17-fold greater risk. However, the importance of these genetic variants if identified in particular individuals within the general population is unknown. We undertook this study to estimate the penetrance of these variants in the general population.

\textbf{Methods:} We genotyped 43,596 individuals from the Danish general population followed between January 1976 and July 2007. Using a logistic regression model, we estimated the risk of Crohn disease in relation to variants of the \textit{NOD2/CARD15} gene in the general population. Penetrance was calculated as the fraction of participants in whom Crohn disease was diagnosed before age 50 years.

\textbf{Results:} In the general population, 89% of participants were noncarriers of the genetic variants of interest (\(n = 38,594\)), 11% were heterozygotes (\(n = 4,838\)), and 0.4% were compound heterozygotes or homozygotes (\(n = 164\)). For Crohn disease, multifactorially adjusted odds ratios were 1.2 (95% confidence interval [CI] 0.8–1.9) for heterozygotes and 3.3 (95% CI 0.8–13.6) for compound heterozygotes and homozygotes combined, relative to noncarriers. Only 2 compound heterozygotes received a diagnosis of Crohn disease, and this disease was not diagnosed in any of the homozygotes. The penetrance at age 50 years of \textit{NOD2/CARD15} genetic variants of Crohn disease was 0.30% (95% CI 0.29%–0.31%) for heterozygotes and 1.5% (95% CI 1.4%–1.6%) for compound heterozygotes and homozygotes.

\textbf{Interpretation:} The penetrance of \textit{NOD2/CARD15} genetic variants in relation to risk of Crohn disease for this Danish population was lower than might have been expected from previous European case–control studies. This should be considered when advising healthy individuals in whom these genetic variants are identified.

\textbf{C}rohn disease and ulcerative colitis are closely related inflammatory bowel diseases characterized by dysregulation of the mucosal immune response of the gut.\textsuperscript{1} An estimated 2.2 million people in Europe and 1.4 million people in North America have these diseases. Two single-nucleotide polymorphisms (\textit{Arg702Trp(rs2066844)} and \textit{Gly908Arg(rs2066845)}) and a frameshift polymorphism (\textit{Leu1007fsinsC(rs5743293)}) on the \textit{NOD2} (nucleotide-binding oligomerization domain containing 2) gene, also known as the \textit{CARD15} (caspase recruitment domain 15) gene, have been associated with Crohn disease but not ulcerative colitis.\textsuperscript{2,3} \textit{NOD2/CARD15} encodes an intracellular receptor-sensing bacterial product that facilitates the innate immune response.\textsuperscript{4} In a recent meta-analysis of case–control studies involving Europeans, the aggregated odds ratios for Crohn disease were 2.4 (95% confidence interval [CI] 2.0–2.8) for heterozygotes relative to noncarriers and 17 (95% CI 11–27) for homozygotes or compound heterozygotes relative to noncarriers.\textsuperscript{5} However, the penetrance, the fraction of individuals with these genetic variants in whom Crohn disease develops by age 50 years in the general population, is unknown. More than 30 other genetic variants have now been identified as being associated with Crohn disease using the genome-wide association approach in case–control studies.\textsuperscript{6} However, many of these new variants have smaller effect sizes for risk of Crohn disease\textsuperscript{6} than those reported for \textit{NOD2/CARD15}.

We hypothesized that genetic variants of \textit{NOD2/CARD15} influence the risk of Crohn disease in the general population. To our knowledge, the penetrance of the \textit{NOD2/CARD15} genetic variants has never been examined in a general population study in which genotyping is performed before the participant’s disease status is known. We undertook this study to estimate this penetrance.

\textbf{Methods}

\textbf{Participants}

We drew the participants in the study reported here from participants in the Copenhagen City Heart Study and the Copenhagen General Population Study. The Department of Clinical Biochemistry, Herlev Hospital, and the Copenhagen and Fred-

From the Department of Clinical Biochemistry (Yazdanyar, Kamstrup, Nordestgaard), Herlev Hospital, Herlev, Denmark; the Faculty of Health Sciences (Yazdanyar, Kamstrup, Tybjærg-Hansen, Nordestgaard), the Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark; and the Department of Clinical Biochemistry (Tybjærg-Hansen), Rigshospitalet, Copenhagen, Denmark

\textit{CMAJ} 2010. DOI:10.1503/cmaj.090684
eriksberg and the Danish Capital Region ethics committees approved all of the investigations in these two studies, and all participants gave written informed consent. There was no overlap of participants between the two studies. All of the participants were white and of Danish descent, as documented through ethnicity information in the Danish Central Population Register, which covers all persons living in Denmark.

The Copenhagen City Heart Study is a study of the Danish general population initiated in 1976–1978, with follow-up examinations in 1981–1983, 1991–1994 and 2001–2003.14 Two of us (A.T.H., B.G.N.) selected the participants for that study from the Danish Central Population Register to reflect the adult Danish general population aged 20–100 years. Blood samples for extraction of DNA were available for 10,997 of the participants who attended the 1991–1994 and/or the 2001–2003 examinations.

The Copenhagen General Population Study is another study of the Danish general population that was initiated in 2003 and that was still recruiting participants at the time of writing.15 Two of us (A.T.H., B.G.N.) selected the participants for that study from the Danish Central Population Register to reflect the adult Danish general population aged 20–100 years. For that study, we genotyped the first 32,999 participants. Data collection in that study was almost identical with that for the Copenhagen City Heart Study, as described below.

For both studies, we used the following standard data-collection methods: a self-administered questionnaire, which was completed by the participant and an investigator on the day of attendance; a physical examination; biochemical tests; collection of DNA for genotyping; and collection of registry information as described below. The age-adjusted odds ratios for Crohn disease for those attending after invitation versus those not attending was 0.8 (95% CI 0.5–1.3) in the Copenhagen City Heart Study and 1.1 (95% CI 0.7–1.7) in the Copenhagen General Population Study.

**Diagnosis of Crohn disease and ulcerative colitis**

In both studies, information on the diagnosis of Crohn disease and ulcerative colitis was based on diagnoses (World Health Organization, International Classification of Diseases, 8th edition, codes 536.01, 536.19 and 536.99, and 10th edition, codes K50 and K51) entered in the national Danish Patient Registry, the national Danish Causes of Death Registry and the national Danish Civil Registration System through July 2007. The national Danish Patient Registry contains information on all patient contacts with all clinical hospital departments in Denmark. The national Danish Causes of Death Registry contains data on the causes of all deaths in Denmark, as reported by hospitals and general practitioners. The national Danish Civil Registration System contains data for all deaths in and emigrations from Denmark. We obtained information from these three registries using each person’s unique Danish personal identification number. For both studies, we captured all prevalent (as well as incident) cases of Crohn disease and ulcerative colitis diagnosed in the entire country. Our complete registry information goes back to the year 1976, when the national Danish Patient Registry was established, and contains information on all patient contacts with all clinical hospital depart-

ments in the entire country, including outpatients. Because all patients with Crohn disease and ulcerative colitis in Denmark are followed regularly at clinical hospital departments, where they receive treatment or are under medical observation, we believe that our methods have captured all patients with these diseases, provided they had at least one contact with their respective hospitals during or after 1976.

The only exceptions would be persons in whom Crohn disease or ulcerative colitis was diagnosed before 1976 and who had complete remission and therefore no further contact with any hospital department in Denmark during or after 1976. These rare patients would not be registered as cases in our study. However, because Crohn disease and ulcerative colitis are chronic diseases entailing lifelong need for medical observation and treatment, we estimate that the magnitude of this potential problem was minimal.

One of us (S.Y.) collected and reviewed the complete medical records for 31 randomly selected cases of Crohn disease and 32 randomly selected cases of ulcerative colitis. This review confirmed the diagnosis for all 63 patients on the basis of clinical, radiologic, endoscopic and/or histologic criteria, in accordance with the criteria of Lennard-Jones.10

**Laboratory analyses**

We used TaqMan gene expression assays (Applied Biosystems, Uppsala, Sweden) to genotype the NOD2/CARD15 genetic variants. We confirmed the results of genotyping by DNA sequencing (MegaBace, GE Healthcare Life Sciences, Uppsala, Sweden) for all homozygotes and for 400 randomly selected compound heterozygotes, heterozygotes and noncarriers. We found no discrepancies between the two methods of genotyping. We used standard hospital assays to determine leukocyte counts and concentrations of hemoglobin, albumin, high-sensitive C-reactive protein and orosomucoid.

**Other covariates**

We defined smoking in terms of whether the participant was a current smoker, using information from the questionnaires. Information on whether the patient had undergone appendectomy was based on diagnoses (World Health Organization, International Classification of Diseases, 8th edition, codes 5400–5409, 5419 or 5420, and 10th edition, codes K35–K37) entered in the national Danish registries, as described above.

**Statistical analyses**

A $p$ value less than 0.05 with two-sided testing was considered significant. Pearson $\chi^2$ test was used to compare categorical variables, and Kruskal–Wallis analysis of variance was used to compare continuous variables, as a function of genotyope.

We used logistic regression to estimate odds ratios with 95% confidence intervals. We adjusted the analyses for age or multifactorially for age, sex, smoking (yes/no) and appendectomy (yes/no); changes in smoking status and appendectomy over time were not considered.

The penetrance at age 50 years was calculated as the fraction of participants, analyzed at age 50 years, with NOD2/CARD15 heterozygosity, compound heterozygosity or homozygosity in whom Crohn disease had developed before age 50 years.
Results

We obtained registry information for all individuals from January 1976 until the diagnosis of Crohn disease or ulcerative colitis, death or July 2007, whichever came first. Follow-up was complete for all of the participants.

In the general population, 89% were noncarriers (n = 38,594), 11% were heterozygotes (n = 4838), and 0.4% were compound heterozygotes or homozygotes (n = 164) for the NOD2/CARD15 genetic variants. The genotype distributions for the three NOD2/CARD15 polymorphisms in the two study populations did not differ from those expected according to Hardy–Weinberg equilibrium ($\chi^2$ test, $p = 0.06$–0.99); that is, the genotype frequencies were the same as they would have been if mating occurred at random and there was no selection, migration or new mutations occurring in the Danish population. In addition, the distribution of baseline characteristics did not differ among individuals stratified by NOD2/CARD15 genotype (Table 1).

Crohn disease and ulcerative colitis

In the Copenhagen City Heart Study, Crohn disease was diagnosed in 52 participants and ulcerative colitis in 127 participants over the period 1976 to July 2007 (Table 2). The age-standardized incidence rates were 5.3 per 100,000 for the three populations did not differ from those expected according to the Copenahgen City Heart Study. Crohn disease was diagnosed in 110 participants and ulcerative colitis in 249 participants over the period 1976 to July 2007 (Table 2). The multifactorially adjusted odds ratios for heterozygotes versus noncarriers were 1.1 (95% CI 0.6–2.1) for Crohn disease and 0.9 (95% CI 0.7–1.5) for ulcerative colitis. The corresponding odds ratio for compound heterozygotes and homozygotes combined was 3.0 (95% CI 0.4–22.1) for Crohn disease, based on the presence of only one compound heterozygote with a diagnosis of Crohn disease; no homozygotes had Crohn disease, and no compound heterozygotes or homozygotes had ulcerative colitis. The age-adjusted odds ratios were similar (Table 2).

In the combined analysis for both general population studies, with maximal statistical power, the multifactorially adjusted odds ratios for heterozygotes versus noncarriers were 1.2 (95% CI 0.8–1.9) for Crohn disease and 1.1 (95% CI 0.8–1.5) for ulcerative colitis (Table 2). For compound heterozygotes and homozygotes combined, the corresponding odds ratios were 3.3 (95% CI 0.8–13.6) for Crohn disease and 0.7 (95% CI 0.1–5.1) for ulcerative colitis.

In the combined analysis of Crohn disease, we had 80% statistical power with a one-sided $p < 0.05$ to exclude an odds

| Character                     | Noncarriers | Heterozygotes | Compound heterozygotes and homozygotes | p value‡ |
|------------------------------|-------------|---------------|----------------------------------------|----------|
| Copenhagen City Heart Study  | n = 9381    | n = 1162      | n = 54                                 |          |
| Sex, female                  | 5189 (55.3) | 666 (57.3)    | 26 (48.1)                              | 0.24     |
| Age at start of follow-up, yr, median (IQR) | 41 (26–52) | 41 (26–52) | 40 (28–53) | 0.82     |
| Current smoker               | 5226 (55.7) | 636 (54.7)    | 28 (51.8)                              | 0.66     |
| Had appendectomy             | 286 (3.0)   | 40 (3.4)      | 2 (3.7)                                | 0.74     |
| Copenhagen General Population Study | n = 29213 | n = 3676      | n = 110                                |          |
| Sex, female                  | 15189 (52.0) | 1972 (53.6) | 55 (50.0)                               | 0.15     |
| Age at start of follow-up, yr, median (IQR) | 30 (20–39) | 30 (20–39) | 30 (21–39) | 0.75     |
| Current smoker               | 6329 (21.7) | 783 (21.3)    | 26 (23.6)                              | 0.76     |
| Had appendectomy             | 956 (3.3)   | 116 (3.2)     | 7 (6.4)                                | 0.18     |
| Combined studies             | n = 38594   | n = 4838      | n = 164                                |          |
| Sex, female                  | 20378 (52.8) | 2638 (54.5) | 81 (49.4)                              | 0.05     |
| Age at start of follow-up, yr, median (IQR) | 32 (21–42) | 32 (21–42) | 32 (23–44) | 0.47     |
| Current smoker               | 11555 (29.9) | 1419 (29.3) | 54 (32.9)                              | 0.66     |
| Had appendectomy             | 1242 (3.2)  | 156 (3.2)     | 9 (5.5)                                | 0.25     |

* NOD2/CARD15 genotypes (Arg702Trp, Gly908Arg, Leu1007fsinsC): noncarrier for all three NOD2/CARD15 variants, heterozygous for one of the NOD2/CARD15 variants, compound heterozygous with two NOD2/CARD15 variants or homozygous for one of the NOD2/CARD15 variants.
† Unless indicated otherwise.
‡ Pearson $\chi^2$-test and Kruskal–Wallis analysis of variance.
ratio for heterozygotes versus noncarriers of 1.8 or larger or to exclude an odds ratio for compound heterozygotes and homozygotes combined of 6.4 or larger. The corresponding odds ratios for ulcerative colitis were 1.5 (heterozygotes) and 4.1 (compound heterozygotes and homozygotes combined).

The penetrance at age 50 years of NOD2/CARD15 genetic variants for Crohn disease was 0.30% (95% CI 0.29%–0.31%) for heterozygotes and 1.5% (95% CI 1.4%–1.6%) for compound heterozygotes and homozygotes.

Biochemical markers in the general population
Leukocyte counts and concentrations of hemoglobin, albumin and C-reactive protein did not differ according to NOD2/CARD15 genotype in either of the two studies (Kruskal–Wallis analysis of variance, \( p = 0.13–0.95 \)). Also, levels of orosomucoid did not differ according to NOD2/CARD15 genotype in the Copenhagen City Heart Study (\( p = 0.56 \)).

Interpretation
Surprisingly, we found no statistically significant association between NOD2/CARD15 genetic variants and Crohn disease in either of the two general population studies that we analyzed, which suggests a low penetrance of the NOD2/CARD15 genetic variants in the European general population.

The odds ratio for Crohn disease in heterozygotes relative to noncarriers was 1.2 (95% CI 0.8–1.9), and the study had 80% statistical power to exclude an odds ratio of 1.8 or larger. This suggests that the penetrance in the general population was indeed less than would be expected from the odds ratio of 2.4 (95% CI 2.0–2.8) reported in a recent meta-analysis of case–control studies for heterozygotes versus noncarriers. For compound heterozygotes and homozygotes combined versus noncarriers, the odds ratio for Crohn disease was 3.3 (95% CI 0.8–13.6), and the study had 80% statistical power to exclude an odds ratio of 6.4 or larger. This also suggested that the penetrance of NOD2/CARD15 genetic variants in the general population was less than would be expected from the odds ratio reported in the recent meta-analysis (17; 95% CI 11–27).

The divergence of our results, obtained through two general population studies, from those of previous hospital- or population-based case–control studies could be due to any of several reasons. First, our population-based studies may provide more accurate estimates because population-based studies avoid some of the biases inherent in a case–control design. For example, in contrast to diseased individuals in the general population, case patients in case–control studies could have clustering of other known and unknown risk factors. Clustering of risk factors may introduce ascertainment bias and thus an overestimation of the importance of NOD2/CARD15 genetic variants in case–control studies, as we and others have previously observed for the effect of hemochromatosis geno-

| Characteristic                      | Crohn disease | Ulcerative colitis |
|------------------------------------|---------------|--------------------|
|                                    | Non-carriers  | Heterozygotes      | Compound heterozygotes and homozygotes | Non-carriers  | Heterozygotes      | Compound heterozygotes and homozygotes |
| **Copenhagen City Heart Study**    |               |                    |                                  |               |                    |                                  |
| No. of events                      | n = 9 381     | n = 1 162          | n = 54                          | n = 9 381     | n = 1 162          | n = 54                          |
| Odds ratio (95% CI)                | 1.0           | 1.3 (0.6–2.9)      | 4.0 (0.5–29.4)                  | 1.0           | 1.1 (0.6–1.9)      | 1.6 (0.2–11.3)                  |
| Age-adjusted                       | 1.0           | 1.3 (0.6–2.9)      | 4.1 (0.6–30.6)                  | 1.0           | 1.1 (0.6–1.9)      | 1.5 (0.2–11.2)                  |
| Multifactorially adjusted†         | 1.0           | 1.3 (0.6–2.9)      | 4.1 (0.6–30.6)                  | 1.0           | 1.1 (0.6–1.9)      | 1.5 (0.2–11.2)                  |
| **Copenhagen General Population Study** | n = 29 213 | n = 3 676          | n = 110                         | n = 29 213 | n = 3 676          | n = 110                         |
| Odds ratio (95% CI)                | 1.0           | 1.2 (0.7–2.0)      | 2.8 (0.4–20.5)                  | 1.0           | 1.1 (0.8–1.6)      | –                              |
| Age-adjusted                       | 1.0           | 1.2 (0.7–2.0)      | 3.0 (0.4–22.1)                  | 1.0           | 0.9 (0.7–1.5)      | –                              |
| Multifactorially adjusted†         | 1.0           | 1.2 (0.7–2.0)      | 3.0 (0.4–22.1)                  | 1.0           | 0.9 (0.7–1.5)      | –                              |
| **Combined studies**               | n = 38 594    | n = 4 838          | n = 164                         | n = 38 594    | n = 4 838          | n = 164                         |
| No. of events                      | 139           | 21                 | 2                              | 329           | 46                 | 1                              |
| Odds ratio (95% CI)                | 1.0           | 1.2 (0.8–1.9)      | 3.4 (0.8–14.0)                  | 1.0           | 1.1 (0.8–1.5)      | 0.7 (0.1–5.1)                  |
| Age-adjusted                       | 1.0           | 1.2 (0.8–1.9)      | 3.3 (0.8–13.6)                  | 1.0           | 1.1 (0.8–1.5)      | 0.7 (0.1–5.1)                  |
| Multifactorially adjusted†         | 1.0           | 1.2 (0.8–1.9)      | 3.3 (0.8–13.6)                  | 1.0           | 1.1 (0.8–1.5)      | 0.7 (0.1–5.1)                  |

Note: CI = confidence interval.

* NOD2/CARD15 genotypes (Arg702Trp, Gly908Arg, Leu1007InsC): noncarrier for all three NOD2/CARD15 variants, heterozygous for one of the NOD2/CARD15 variants, compound heterozygous with two NOD2/CARD15 variants or homozygous for one of the NOD2/CARD15 variants.
† Multifactorial adjustment for age, sex, smoking (yes/no) and appendectomy (yes/no).
‡ One compound heterozygote received a diagnosis of both Crohn disease and ulcerative colitis; no homozygotes had either disease.
§ One compound heterozygote received a diagnosis of Crohn disease; no homozygotes had either disease.
types on development of clinically manifest hemochromatosis\textsuperscript{11–14} and for the effect of mutations of the low-density lipoprotein receptor and apolipoprotein B on cholesterol levels.\textsuperscript{12,13,16}

Conversely, our general population studies may have generated incorrect estimates as a result of selection bias or misclassification of cases and controls. Selection bias was a possibility in the Copenhagen City Heart Study, in which participants underwent genotyping only if they attended the third examination, in 1991–1994, or the fourth examination, in 2001–2003. Selection bias might also have occurred if death or severe disease prevented certain individuals from participating in the third or fourth survey. However, the observed genotype frequencies did not differ from those predicted by Hardy–Weinberg equilibrium. Furthermore, such selection bias was much less likely in our second study, the Copenhagen General Population Study, and the results of the two studies were similar. Our results might also have been biased if the presence of Crohn disease prevented people from participating in our general population studies. However, the age-adjusted odds ratio for Crohn disease in those who attended after the invitation to participate versus those who did not attend was 0.8 (95% CI 0.5–1.3) in the Copenhagen City Heart Study and 1.1 (95% CI 0.7–1.7) in the Copenhagen General Population Study, and the results of the two studies were similar. We found no association between NO\textsubscript{2}/CARD\textsubscript{15} genetic variants and the risk of ulcerative colitis\textsuperscript{11} or levels of inflammatory biochemical markers.\textsuperscript{17}

We believe that our results for the general population do not invalidate the results obtained in previous case–control studies. Rather, they indicate that the interpretation of the importance of NO\textsubscript{2}/CARD\textsubscript{15} genetic variants differs according to whether genotyping is being conducted for patients in a hospital clinic or for individuals in the general population. Because the NO\textsubscript{2}/CARD\textsubscript{15} genetic variants are associated only with Crohn disease and not with ulcerative colitis, these genetic variants could be useful in distinguishing between these two clinical subtypes of inflammatory bowel disease. However, the low penetrance of the NO\textsubscript{2}/CARD\textsubscript{15} genetic variants in the general population documented in our study suggests that these genetic variants should not be considered as a genetic test for screening individuals in the general population. However, the low penetrance documented here should be considered when advising healthy individuals in whom these genetic variants are identified.

This article has been peer reviewed.

Competing interests: None declared.

Contributors: Shiva Yazdanyar, Anne Tybjerg-Hansen and Børge G. Nordestgaard participated in the conception and design of the study. All of the authors contributed to the analysis and interpretation of the data. Shiva Yazdanyar did the genotyping and drafted the manuscript, Shiva Yazdanyar and Pia R. Kamstrup performed statistical analyses, and Pia R. Kamstrup, Anne Tybjerg-Hansen and Børge G. Nordestgaard provided critical revisions of the manuscript. All of the authors approved the final version submitted for publication.

Acknowledgments: Shiva Yazdanyar, Pia R. Kamstrup, Anne Tybjerg-Hansen and Børge G. Nordestgaard were investigators with the Copenhagen General Population Study. Anne Tybjerg-Hansen and Børge G. Nordestgaard are investigators with the Copenhagen City Heart Study. The authors thank the staff and participants of the Copenhagen General Population Study and the Copenhagen City Heart Study for their important contributions.

Funding: This work was supported by the University of Copenhagen.

REFERENCES

1. Loftus EV Jr. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. Gastroenterology 2004;126:1504–17.

2. Beutler E, Felitti VJ, Kozol JA, et al. Penetration of 845G-->A (C282Y) HFE haemochromatosis mutation in the general population: iron overload progression rate. Lancet 2001;358:221–30.

3. Tybjerg-Hansen A, Jensen HK, Benn M, et al. Phenotype of heterozygotes for low-density lipoprotein receptor mutations identified in different background populations. Atheroscler Thromb Vasc Biol 2005;25:211–5.

Correspondence to: Dr. Børge G. Nordestgaard, Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev Ringvej 75, DK-2730 Herlev, Denmark; brno@righ.regionh.dk

CMAJ • APRIL 20, 2010 • 182(7)