Journal of Pediatric Surgery xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Pediatric Surgery

journal homepage: www.elsevier.com/locate/jpedsurg

A population-based description of familial clustering of Hirschsprung's disease

Craig C. Teerlink [a,*], Ryan Bernhisel [b], Lisa A. Cannon-Albright [a,c], Michael D. Rollins [d]

[a] Department of Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT 84108, USA
[b] Myriad Genetics, Salt Lake City, UT 84108, USA
[c] George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, UT 84148, USA
[d] Division of Pediatric Surgery, Primary Children's Hospital, University of Utah, Salt Lake City, UT 84113, USA

ARTICLE INFO

Article history:
Received 6 April 2017
Received in revised form 2 August 2017
Accepted 23 August 2017
Available online xxxx

Key words:
Hirschsprung's disease
Familial aggregation
Genealogy

ABSTRACT

Background: Familial recurrence of Hirschsprung's disease (HSCR) is well documented, and risk estimates for relatives have been reported from various populations. We describe the familial clustering of HSCR cases using well-established unbiased familial aggregation techniques within the context of a population genealogy.
Methods: Patients included 264 HSCR cases identified using ICD-9 diagnosis coding from the two largest healthcare providers in Utah who also had linked genealogy data. The GIF statistic was used to identify excess familial clustering by comparing average relatedness of cases to matched controls. In addition, relative risks (RRs) of HSCR in relatives of cases were estimated using age-, sex- and birthplace-matched disease rates, and for several diseases frequently associated with HSCR (Down syndrome, multiple endocrine neoplasia IIa, central hypoventilation syndrome, Bardet–Biedl syndrome, ventricular and atrial septal defect).
Results: Significant excess relatedness was observed for all HSCRs ($p < 1e^{-3}$). Significant RRs for HSCR were observed for first-, second-, and fourth-degree relatives of cases (RR = 12.0, 10.0, and 4.6, respectively). Significant elevated risks of Down syndrome, Bardet–Biedl syndrome, and atrial and ventricular septal defects were observed for HSCR cases.
Conclusion: This population-based survey of HSCR provides confirmation of a genetic contribution to HSCR disease and presents unbiased risk estimates that may have clinical value in predicting recurrence.
Level of evidence rating: Prognosis study, level II.

Hirschsprung's disease (HSCR) is a rare disorder affecting 1 in 5000 births (0.02% population risk), and results from the failure of neural crest cells to fully colonize the intestinal tract during the 4th to 12th weeks of human gestation (aganglionosis) [1]. HSCR has variable presentation which is generally categorized as short-segment disease (aganglionosis of the rectum and sigmoid, about 80% of cases) long-segment disease (aganglionosis beyond the sigmoid, about 10% of cases [1], and total colon aganglionosis, about 10% of cases [2]. HSCR can occur either by itself (70% of cases) or as a manifestation with another birth defect (syndromic form, 30% of cases) [3]. Previously reported syndromes associated with HSCR include Down syndrome, multiple endocrine neoplasia type IIa, central hypoventilation syndrome, Bardet–Biedl syndrome [4], and ventricular and atrial septal defects [5].

A genetic etiology for HSCR has long been suspected, and several genes contributing to the disease have been noted [6]. Known mutations in the RET gene are responsible for about 50% of familial cases and 20% of sporadic cases [7]. Several studies conducted in different populations have characterized the disease in terms of familial recurrence (8%) [8–12]. These types of studies are typically conducted on selected patients and without complete genealogical information, resulting in some bias. Here, we used well-established methods applied to a population-based resource combining genealogy with health care data for the majority of the Utah population to provide an unbiased description of observed familial clustering for HSCR. These observations may guide future genetic investigations of HSCR and enhance clinical interpretation of risk.

Abbreviations: CI, confidence interval; dGIF, distant Genealogic Index of Familiality; HSCR, Hirschsprung's disease; GIF, Genealogic Index of Familiality; ICD, International Classification of Diseases; IM, Intermountain Healthcare; RR, relative risk; MEN-IIa, multiple endocrine neoplasia type IIa; UPDB, Utah Population Database; UUHSC, University of Utah Health Sciences Center.

* Corresponding author.
E-mail addresses: craig.teerlink@utah.edu (C.C. Teerlink), shakabra09@gmail.com (R. Bernhisel), lisa.albright@utah.edu (L.A. Cannon-Albright), Michael.Rollins@imail2.org (M.D. Rollins).

http://dx.doi.org/10.1016/j.jpedsurg.2017.08.024
0022-3468/

1. Materials and methods

1.1. Utah Population Database (UPDB) genealogy data

The UPDB includes a genealogy representing the pioneers who founded Utah and their descendants to modern day. This genealogy data have been record-linked to other Utah data sources, including Utah census data, the Utah Cancer Registry, Utah death certificates from 1904, Utah birth certificates, Utah drivers' licenses, and the electronic data warehouse data for the 2 major health care providers in Utah since 1994 (The University of Utah and Intermountain Healthcare).

The original Utah genealogy data consisted of over 180,000 three-generation family group sheets collected by the Family History Library of the Church of Jesus Christ of Latter Day Saints (LDS or Mormon) that include at least one event in Utah or along the Mormon pioneer trail from the east coast that were entered and record-linked in the early 1970s to form a genealogy of 1.6 million individuals representing up to 6 generations [13]. The genealogy data since the 1970s have been expanded using Utah State Vital Statistics Data (e.g., father, mother, and child from a Utah birth certificate). Today the Utah genealogy data extend up to 11 generations and include almost 3 million individuals who are part of at least 3 generations of genealogy that connects to the original Utah genealogy data; this subset of 3 million individuals was used for the analyses of genetic relationships described here. The methods used in this study were designed specifically for analysis of population genealogical data, which limits the number of populations where a similar approach could be used to the Icelandic genealogy (deCODE Genetics), and an emerging genealogy of the United States that is currently being created at the United States Veterans Administration (VA) and linked to the VA's nationwide medical record system [14].

1.2. HSCR cases and comorbid phenotypes

The 3 million individuals with genealogy data include 1.7 million individuals treated at an Intermountain Healthcare facility (IM) and over 760,000 individuals treated at the University of Utah Health Sciences Center (UUHSC). Diagnostic codes using International Classification of Disease (ICD) Revision 9 were used to identify individuals with HSCR from IM and UUHSC spanning 1994–2014 (ICD-9 code 751.3). ICD-9 diagnostic codes (in parentheses) were also used to identify cases with several other diseases that have been reported to be associated with HSCR, including: Down syndrome (758.0), multiple endocrine neoplasia type IIa (MEN-IIa) (258.02), central hypoventilation syndrome (327.25), Bardet–Biedl syndrome (759.89), ventricular septal defect (745.4), and atrial septal defect (745.5).

1.3. Genealogical Index of Familiality (GIF) method

The Genealogical Index of Familiality (GIF) method, a well-established statistical test, was used to test for excess relatedness (or familial clustering) among a set of individuals [15–18]. The GIF method utilizes the kinship coefficient to measure the relatedness for a pair of individuals with the specified phenotype (probands). The kinship coefficient (ψ) is defined as the probability that randomly selected homologous alleles from two individuals are identical by descent from a shared ancestor. The case GIF statistic represents the average pairwise relatedness for all pairs of probands, multiplied by 10^5 for ease of presentation. To test for excess relatedness, the case GIF statistic is compared to the expected relatedness for a group of individuals similar to the probands, randomly selected from the UPDB population. To estimate the expected relatedness for the probands a distribution of GIF statistics is created from analysis of 1000 independent sets of randomly selected, matched controls. Controls are matched to probands for 5-year birth year, sex, and place of birth (Utah or not). To test for an excess of familial clustering, the case GIF is compared to the distribution of 1000 matched control GIFs. The empirical p-value for the test is the proportion of control GIFs that are as extreme as, or more extreme than, the case GIF. The contribution to the GIF statistic for cases and controls is plotted at each genetic distance and represents the proportion of case or control pairs appearing with the specified genetic distance.

It is reasonable to assume that close relatives share more environmental exposures than distantly related cases, and excluding these close relationships removes the majority of instances where shared environmental exposures may be a factor. The GIF method has been extended to exclude first- and second-degree relationships (distant GIF or dGIF). A significant dGIF outcome is a strong indicator of a genetic component to the disease.

1.4. Relative risk in relatives (RR)

The relative risk (RR) is defined as the ratio of observed number of affected relatives of cases to the expected number of affected relatives of cases [16–19]. The observed number of cases is obtained by counting the number of affected relatives of cases, without duplication. The expected number of affected cases requires the estimation of disease rates for the phenotype of interest. To estimate disease rates all individuals are assigned to one of 137 cohorts based on 5-year birth year cohorts, sex, and place of birth (Utah or not). Cohort-specific disease rates are calculated by adding up the total number of cases in each cohort and dividing by the total number of individuals in the cohort. The expected number of affected relatives is calculated by multiplying the total number of relatives of cases in a cohort by disease rate for each cohort summing over all cohorts.

$$\text{Expected} = \sum_{i=1}^{n} R_i C_i \Big/ N_i$$

R_i number of relatives of cases
C_i number of cases in each cohort
N_i total number in each cohort

Once this ratio is obtained, the significance test for the RR and 95% confidence intervals for the RR can be calculated based on the assumption that the observed number of cases follows a Poisson distribution with mean equal to the expected value [15]. Relative risks for HSCR were estimated for siblings, brothers, sisters, children, all first-, second-, third-, and fourth-degree relatives. Relative risks for birth defects recognized to be associated with HSCR were estimated similarly for the set of HSCR cases themselves, and for first-, second-, and third-degree relatives of cases. Significance for these tests is reported without correction for the number of multiple tests.

1.5. A test for excess consanguinity

To test for excess consanguinity among cases, we calculated the coefficient of inbreeding for each case (equivalent to the pairwise kinship coefficient of their parents). We then compared the average coefficient of inbreeding for cases in distribution to the average values from 1000 sets of matched controls (age, sex and birthplace in or out of Utah). The Utah population has similar levels of inbreeding to other geographic regions of the United States [20]).

2. Results

374 HSCR cases were identified by ICD-9 code (with and without genealogy data) from electronic medical records between 1994 and 2014, 264 of which intersected the analyzed genealogy. Of the 264 analyzed HSCR cases, 161 (61%) were male and 103 (39%) were female. The number of subjects with one of the six specified birth defects previously

Please cite this article as: Teerlink CC, et al, A population-based description of familial clustering of Hirschsprung's disease, J Pediatr Surg (2017), http://dx.doi.org/10.1016/j.jpedsurg.2017.08.024