Population-based description of familial clustering of Chiari malformation Type I

Diana Abbott, PhD,1 Douglas Brockmeyer, MD,2 Deborah W. Neklason, PhD,1 Craig Teerlink, PhD,1 and Lisa A. Cannon-Albright, PhD1,3

1Division of Genetic Epidemiology, Department of Internal Medicine, and 2Department of Neurosurgery, Clinical Neurosciences Center, University of Utah School of Medicine; and 3George E. Wahlen Department of Veterans Affairs Medical Center, Salt Lake City, Utah

OBJECTIVE A population-based genealogical resource with linked medical data was used to define the observed familial clustering of Chiari malformation Type I (CM-I).

METHODS All patients with CM-I were identified from the 2 largest health care providers in Utah; those patients with linked genealogical data were used to test hypotheses regarding familial clustering. Relative risks (RRs) in first-, second-, and third-degree relatives were estimated using internal cohort-specific CM-I rates; the Genealogical Index of Familiality (GIF) test was used to test for an excess of relationships between all patients with CM-I compared with the expected distribution of relationships for matched control sets randomly selected from the resource. Pedigrees with significantly more patients with CM-I than expected (p < 0.05) based on internal rates were identified.

RESULTS A total of 2871 patients with CM-I with at least 3 generations of genealogical data were identified. Significantly increased RRs were observed for first- and third-degree relatives (RR 4.54, p < 0.001, and RR 1.36, p < 0.001, respectively); the RR for second-degree relatives was elevated, but not significantly (RR 1.20, p = 0.13). Significant excess pairwise relatedness was observed among the patients with CM-I (p < 0.001), and borderline significant excess pairwise relatedness was observed when all relationships closer than first cousins were ignored (p = 0.051). Multiple extended high-risk CM-I pedigrees with closely and distantly related members were identified.

CONCLUSIONS This population-based description of the familial clustering of 2871 patients with CM-I provided strong evidence for a genetic contribution to a predisposition to CM-I.

https://thejns.org/doi/abs/10.3171/2016.9.JNS161274

KEY WORDS Chiari malformation; relative risk; UPDB; familiality; skull base

Chiari malformation Type I (CM-I) is a neurological and spinal disorder characterized by herniation of the cerebellar tonsils below the foramen magnum. The underlying cause of the disorder is unknown. CM-I frequently presents with other significant radiographic findings, including caudal migration of the brainstem, odontoid retroflexion, basilar invagination, occipitalization of the atlas, syringomyelia, and scoliosis, all observed either alone or in combination with simple tonsillar herniation. CM-I may cause a wide variety of signs and symptoms, most of which are believed to be the result of abnormalities in CSF flow and/or brainstem compression.

Estimates suggest that approximately 215,000 individuals in the US may be affected with CM-I,18 and approximately 12% of these patients have at least 1 close relative with the disorder.14 In a population-based retrospective cohort study, Aitken et al.1 demonstrated that the prevalence of CM-I in 5248 asymptomatic and symptomatic patients under 20 years of age was 0.97%, thus underlining the frequency of the disorder. However, the true prevalence of CM-I is difficult to determine because of several factors, most notably the presence of asymptomatic individuals within the population, the discovery of incidental findings on imaging, and a “detection effect” in which patients might undergo MRI screening because a family member carries the CM-I diagnosis. As a result of these confound-
Familial clustering of Chiari malformation Type I

J Neurosurg Volume 128 • February 2018

Many other studies of CM-I have reported familial clustering, co-occurrence in twins, and an overlap with a subset of known genetic syndromes, suggesting a strong genetic predisposition to this disorder. However, the heritable basis of CM-I has never been investigated using a state-wide population-based data set with sophisticated epidemiological techniques. The information gained from such an analysis would inform conversations with patients and families about the inherited risk of CM-I and strengthen our understanding of the genetic basis of the disorder.

The analysis reported here uses the Utah Population Database (UPDB), a genealogy database linked to state-wide medical data, to define the observed familial clustering for the CM-I phenotype using a large cohort of patients with CM-I identified from the 2 major health care providers for the state of Utah.

Methods

UPDB Genealogy

The UPDB is a population-based resource linking Utah health and hospital record information to a computerized genealogy of the original Utah pioneers and their descendants. The initial genealogy data for 1.6 million individuals was provided by the Church of Jesus Christ of Latter-Day Saints Family History Library in the 1970s. Over the years, the genealogy has expanded via regular updates from Utah state vital records (e.g., father, mother, and child from birth certificates) to include a large part of the Utah population. The UPDB now includes over 2.8 million individuals with at least 3 generations of genealogical data that connect to the original genealogy data. Multiple Utah data repositories have been linked to the Utah genealogy data in the UPDB, including the electronic data warehouses for the 2 largest health care providers in Utah, which cover 60%–70% of Utah individuals. After we obtained institutional review board approval, clinical data for this study were obtained from the University of Utah Health Sciences Center and Intermountain Healthcare. Data from 1994 for these providers was independently record-linked to the UPDB genealogy; 767,594 linked patients from the University of Utah Health Sciences Center and 1,750,935 linked patients from Intermountain Healthcare have at least 3 generations of genealogical data.

Clinical Classification

The ICD-9 diagnosis code 348.4 for compression of brain and/or the CPT code 61343 for Chiari decompression were used to identify patients.

Estimation of Relative Risk for CM-I in Relatives

Estimation of relative risks (RRs) in relatives is a common method to test for evidence of a genetic contribution for a phenotype. RRs were estimated using all patients identified in either of the 2 health care systems. The observed number of patients with CM-I for a group of relatives was obtained by counting (without duplication) the number of relatives diagnosed with CM-I across all cohorts. The expected number of affected relatives was estimated using cohort-specific population disease rates for CM-I estimated from the UPDB. Each individual in the UPDB with genealogical data was assigned to 1 of 136 cohorts based on sex, birth state (Utah or outside of Utah), and birth year (5-year cohorts). The cohort-specific disease rate for CM-I was estimated as the number of patients with CM-I in each cohort divided by the total number of UPDB individuals in the cohort. The expected number of patients with CM-I for a group of relatives was estimated by multiplying the total number of relatives in each cohort by the cohort-specific CM-I rate, then summing across all cohorts. The ratio of observed to expected number of affected relatives is an unbiased estimator of the RR. Significantly elevated RRs in first-degree relatives may represent either shared genetics or shared environment; significantly elevated RRs in more distant relatives, however, strongly suggests a genetic contribution to risk, rather than simply shared environment. RRs were estimated independently for first-, second-, and third-degree relatives. First-degree relatives are parents, offspring, and siblings; second-degree relatives are the first-degree relatives of all first-degree relatives (e.g., father’s brother, or uncle); third-degree relatives are the first-degree relatives of all second-degree relatives (e.g., father’s brother’s daughter, or cousin).

Genealogical Index of Familiality

The Genealogical Index of Familiality (GIF) method was used to test for excess relatedness in individuals with CM-I. This method was developed to study familial aggregation of disease within a genealogy and involves an observed/expected comparison of familial clustering using the coefficient of kinship. The coefficient of kinship is a measure of relatedness commonly used to define pairs of individuals; it represents the probability that 2 alleles, sampled at random from each individual, are identical by descent from a common ancestor. By using the coefficient of kinship, the relatedness was calculated for all possible pairs of patients with CM-I. For most pairs of individuals considered, no relationship was observed. For the GIF test, the average relatedness measure for all of the pairs considered, no relationship was observed. For the GIF test, the average relatedness measure for all of the pairs of patients with CM-I was compared with the expected average relatedness estimated from 1000 matched-control sets randomly selected from the UPDB. Each control set was formed by randomly selecting a cohort-matched control for each patient with CM-I. The empirical significance of the GIF test for excess relatedness is based on comparison of the mean pairwise relatedness of the patients with CM-I to the distribution of the mean pairwise relatedness for the 1000 matched-control sets.

The GIF statistic analysis for CM-I tested for an excess of pairwise relatedness of the patients with CM-I over the expected relatedness of a group of similar individuals in the UPDB population. Significant excess relatedness (or familial clustering) could be due to shared environment, shared genes, or both. To minimize the effects of shared environment on familial clustering, an additional test, the distant GIF (dGIF), was also performed. The dGIF statistic is calculated in the same manner as the GIF, except that...
it ignores all relationships closer than first cousins. The dGIF statistic therefore tests for excess relatedness when only distant relationships are considered.

The contribution to the GIF statistic from each different type of relationship (close and distant) can be displayed by the pairwise genetic distance for all pairs of study patients and controls, allowing consideration of which relationships contributed most to any observed excess relatedness (Fig. 1). The pairwise genetic distances (x-axis) represent parent/offspring (1), siblings or grandparent/grandchild (2), uncle/niece or similar avuncular (3), first cousins or similar (4), second cousins (6), and so forth, and the contribution to the overall GIF statistic can be seen by relationship. There are more pairs of relationships of each type as the genetic distance of the pair increases, but the coefficient of kinship is smaller for more distant relationships (e.g., 0.25 for siblings, 0.0625 for first cousins, and so forth). In Fig. 1, the distribution shown for the mean GIF for the 1000 sets of matched controls shows the expected distribution of relationships of all pairs in a group of individuals from Utah who are similar to the study patients with CM-I (matched for sex, 5-year birth year range, and birth state [Utah or not]); this distribution is very similar to the control distribution observed for most studies and represents the typical relatedness structure for a random group of individuals selected from the UPDB.

High-Risk Pedigrees

High-risk CM-I pedigrees, defined as pedigrees that have a statistically significant excess of pedigree members with CM-I, were identified. All related clusters of patients with CM-I descending from a common founder pair were identified in the UPDB. Because these clusters of related patients (or pedigrees) may have occurred by chance, a test for a significant excess of patients with CM-I in each pedigree was performed. Using cohort-specific CM-I rates applied to the descendants, the expected number of persons with CM-I was estimated for each pedigree. The ratio of the observed to the expected number of persons with CM-I was calculated and used to assess whether there was a significant excess of individuals with CM-I in the pedigree. Pedigrees with a significant excess of individuals with CM-I (p < 0.05) were termed high risk.

Results

A total of 2871 patients diagnosed with CM-I with at least 3 generations of genealogical data were identified from the 2 Utah health care systems; 9 patients had only a CPT procedure code for decompression, 244 patients had both the ICD-9 diagnosis code and the procedure code, and 2618 of the identified patients had only the diagnosis code present at least once in their linked medical record.

RRs for CM-I in first-, second-, and third-degree relatives of patients are shown in Table 1. Significantly increased RRs were observed for first- and third-degree relatives (RR 4.54, p < 0.00001, and RR 4.94, p = 0.000054, respectively). The RR observed in second-degree relatives was elevated, but not significantly (RR 1.20, p = 0.13). Because second-degree relatives are primarily in different generations (for example, grandparent/grandchild or uncle/niece), the narrow window of view on diagnoses (data available from 1994) precludes observation of many such relationships and may explain these results. RRs for the 3 first-degree relative types occurring in different generations (parents, offspring, siblings) were estimated separately but did not differ significantly by type (RR for 3916 parents = 4.23, 95% confidence interval [CI] 2.81–6.11; RR for 7219 siblings = 4.94, 95% CI 3.79–6.31; RR for 4321 children = 4.00, 95% CI 2.66–5.78). Some individu-
Familial clustering of Chiari malformation Type I

Summary results for the GIF and dGIF tests for excess relatedness are reported in Table 2, which includes the number of patients with CM-I, the mean relatedness of the patients with CM-I (patient GIF), the mean relatedness of the 1000 control sets (mean control GIF), the empirical significance for the GIF test (GIF empirical p value), the mean relatedness of distant relationships for the patients (patient dGIF), and the empirical p value for the dGIF test (dGIF empirical p value). The overall GIF tests showed significant excess relatedness when all relationships were considered (p < 0.001). When only distant relationships were considered, the dGIF test for excess relatedness was borderline significant (p = 0.051). Figure 1 shows the contribution to the GIF statistic by pairwise genetic distance for patients compared with matched controls; a clear excess of pairwise relationships up to first cousins (genetic distance = 4) was observed for patients compared with controls. We identified 399 high-risk pedigrees (p < 0.05) with between 2 and 79 related patients with CM-I. Figure 2 shows an example of a high-risk Utah CM-I pedigree identified in the UPDB.

Discussion

This population-based description of the familial clustering of 2871 patients with CM-I observed in the Utah population provides strong evidence for the existence of a genetic contribution to predisposition to CM-I. Significantly elevated RRs were observed in both close and distant relatives. Significant evidence for overall excess relatedness (relationships) was observed, as was borderline-significant evidence for excess relatedness when only distant relationships were considered. Finally, extended pedigrees with a significant excess of patients with CM-I have been identified. This information strengthens our understanding of the familial nature of CM-I. Although familial clustering may occur because of shared environment and shared risk factors, the results presented provide strong evidence that the familial clustering observed for this disease has an inherited contribution.

Previous work has shown a strong tendency for close familial clustering in Chiari malformations. Multiple studies describe a family or families in which more than 1 affected individual is present.2–4,6,15,16,19 Inherited disorders may also be associated with CM-I; examples include Klippel-Feil syndrome, spondyloepiphyseal dysplasia tarda, Goldher syndrome, and achondroplasia, among many others.18 This close association with other disorders raises the possibility that both syndromic and nonsyndromic forms of CM-I exist.16 Improved ability to identify and observe coaggregation of clinically meaningful phenotypes will be important in answering this question.

By using genomic approaches, other groups have investigated the underlying genetics of CM-I, with preliminary, but promising, results. Boyles et al.3 performed genome-wide linkage analysis of 71 individuals in 23 families using more than 10,000 single-nucleotide polymorphisms. They found 2-point logarithm of odds (LOD) scores of > 3.0 on chromosomes 15 and 9, implicating the fibrillin-1 gene, which is reported to be important in Marfan syndrome. Markunas et al.11 performed genome-wide linkage analysis of 367 individuals in 66 families. They stratified their population into individuals positive or negative for connective tissue disease and found linkage evidence for chromosomes 8 and 12 in connective tissue disease–negative families, implicating the GDF6 and GDF3 genes, both important in Klippel-Feil syndrome.

In a set of 2 publications, Markunas et al.9,10 examined the relationship between cranial base morphometrics and whole-genome expression profiles, finding that much posterior fossa morphology was heritable. Multiple genomic regions were strongly implicated with posterior fossa morphology using ordered subset analysis, including chromosomes 1 and 22.9

The analytical methods presented here are well validated; however, this study has some limitations. These include the use of diagnostic or procedure coding in the electronic medical record to identify patients with CM-I.

Table 1. Estimated relative risks for CM-I in relatives of 2871 CM-I probands

<table>
<thead>
<tr>
<th>Relative Type</th>
<th>No. of Relatives</th>
<th>Observed</th>
<th>Expected</th>
<th>RR</th>
<th>p Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>First degree</td>
<td>15,346</td>
<td>119</td>
<td>26.2</td>
<td>4.54</td>
<td>&lt;0.00001</td>
<td>3.76–5.44</td>
</tr>
<tr>
<td>Second degree</td>
<td>36,010</td>
<td>71</td>
<td>59.1</td>
<td>1.20</td>
<td>0.13</td>
<td>0.94–1.52</td>
</tr>
<tr>
<td>Third degree</td>
<td>79,731</td>
<td>178</td>
<td>130.7</td>
<td>1.36</td>
<td>0.000054</td>
<td>1.17–1.58</td>
</tr>
</tbody>
</table>

*“Expected” refers to the expected number of relatives with a diagnosis of CM-I, whereas “Observed” refers to the observed number of relatives diagnosed with CM-I. First-degree relatives are parents, offspring, and siblings. The first-degree relatives of first-degree relatives are second-degree relatives, and the first-degree relatives of second-degree relatives are third-degree relatives.*

Table 2. GIF test for excess relatedness in 2871 patients with CM-I

<table>
<thead>
<tr>
<th>No. w/ CM-I</th>
<th>Patient GIF</th>
<th>Mean Control GIF</th>
<th>GIF Empirical p Value</th>
<th>Patient dGIF</th>
<th>Mean Control dGIF</th>
<th>dGIF Empirical p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2871</td>
<td>2.51</td>
<td>2.12</td>
<td>&lt;0.001</td>
<td>2.11</td>
<td>2.02</td>
<td>0.051</td>
</tr>
</tbody>
</table>

*Patient GIF = mean relatedness of the patients; mean control GIF = mean relatedness of the 1000 control sets; patient dGIF = mean relatedness of distant relationships to the patients; mean control dGIF = mean relatedness of distant relationships for the 1000 control sets.*
It is possible that some misclassification of patients may have occurred, but the methods used are very robust to misclassification; missing diagnoses of true instances of CM-I would simply reduce the power of the study and inclusion of false-positive instances of CM-I would similarly reduce the power to identify excess familial relationships, both of which would serve to make all findings conservative in nature.

The study could also be limited by censoring of patients whose linked hospital data did not include appropriate diagnosis or procedure coding. Other potential censoring includes missing or incorrect genealogical data, failure to link matched records, or diagnosis before 1994 or outside the 2 Utah health care systems included. Although there is no ascertainment or referral bias in the study and it is based on medical data for almost the entire Utah population, the possibility of increased frequency of screening (and thus diagnoses) in relatives of patients with CM-I must be recognized. This might affect risks in close relatives, but it is unlikely to explain increased risks in third-degree relatives and extremely unlikely to explain the extended high-risk pedigrees observed.

The Utah population has been shown to be similar to the Northern European populations from which the majority of the Utah pioneers originated. The results can be assumed to represent similar populations of Northern European extraction but should be validated for other populations.

Given the radiographic and clinical heterogeneity of CM-I, it is no surprise that conclusive evidence of a genetic basis has not been found. Clearly, clues toward a genetic basis of CM-I are beginning to emerge, but much work needs to be done. This population-based analysis of patients with CM-I has provided strong evidence for a genetic contribution to predisposition and evidence of more extended high-risk pedigrees than have previously been reported. This information is important, because it provides a basis for informative discussions with patients and families about the genetic contribution of CM-I. More detailed genomic analysis of the Utah high-risk CM-I pedigrees identified will be important for the identification of specific genetic factors.

Acknowledgments

Partial support for all data sets within the UPDB was provided by the Huntsman Cancer Institute, Huntsman Cancer Foundation, University of Utah, and the Huntsman Cancer Institute's Cancer Center Support grant (P30 CA42014) from the National Cancer Institute.

References


Disclosures
The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author Contributions
Conception and design: Cannon-Albright, Brockmeyer, Neklason. Acquisition of data: Cannon-Albright. Analysis and interpretation of data: Cannon-Albright. Drafting the article: Abbott. Critically revising the article: Cannon-Albright, Brockmeyer, Neklason. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Cannon-Albright. Statistical analysis: Cannon-Albright, Abbott, Teerlink. Study supervision: Cannon-Albright.

Correspondence