Differences in SMN1 allele frequencies among ethnic groups within North America

B C Hendrickson,1 C Donohoe,1 V R Akmaev,1 E A Sugarman,1 P Labrousse,1 L Boguslavskiy,1 K Flynn,1 E M Rohls,1 A Walker,1 B Allitto,1 C Sears,2,3 T Scholl1

ABSTRACT

Background: Spinal muscular atrophy (SMA) is the most common inherited lethal disease of children. Various genetic deletions involving the bi-allelic loss of SMN1 exon 7 are reported to account for 94% of affected individuals. Published literature places the carrier frequency for SMN1 mutations between 1 in 25 and 1 in 50 in the general population. Although SMA is considered to be a pan-ethnic disease, carrier frequencies for many ethnicities, including most ethnic groups in North America, are unknown.

Objectives and methods: To provide an accurate assessment of SMN1 mutation carrier frequencies in African American, Ashkenazi Jewish, Asian, Caucasian, and Hispanic populations, more than 1000 specimens in each ethnic group were tested using a clinically validated, quantitative real-time polymerase chain reaction (PCR) assay that measures exon 7 copy number.

Results: The observed one-copy genotype frequency was 1 in 37 (2.7%) in Caucasian, 1 in 46 (2.2%) in Ashkenazi Jew, 1 in 56 (1.8%) in Asian, 1 in 91 (1.1%) in African American, and 1 in 125 (0.8%) in Hispanic specimens. Additionally, an unusually high frequency of alleles with multiple copies of SMN1 was identified in the African American group (27% compared to 3.3–8.1%). This latter finding has clinical implications for providing accurate adjusted genetic risk assessments to the African American population.

Conclusions: Differences in the frequency of SMA carriers were significant among several ethnic groups. This study provides an accurate assessment of allele frequencies and estimates of adjusted genetic risk that were previously unavailable to clinicians and patients considering testing.

With an incidence of 1 in 6000 to 1 in 10 000 live births, spinal muscular atrophy (SMA) is the most common lethal genetic disease of children.1 SMA is a neuromuscular disorder that leads to progressive, proximal muscle weakness and atrophy. Sixty per cent of individuals affected with SMA have the most severe form, type I, also known as Werdnig–Hoffman disease (Online Mendelian Inheritance in Man (OMIM) 253300), with clinical onset usually occurring before 6 months and death by respiratory failure before the age of 2 years. Clinical onset of type II SMA (OMIM 253550) usually occurs by the age of 15 months with most cases surviving beyond 10 years of age. Type II patients may learn to sit, but never gain the ability to walk. Type III SMA (OMIM 253400) is a milder form with onset typically occurring between 2 and 17 years of age. An adult form of SMA (type IV, OMIM 271150) has also been described.

SMA is caused by mutations in the Survival Motor Neuron 1 (SMA) gene.2 The SMA1 gene resides in a duplicated region of chromosome 5q13 telomeric to the near identical homologue SMN2. More than 90% of the functional SMN protein is contributed by SMN1. Comparably SMN2 produces only a small fraction of SMN due to a point mutation in exon 7 that disrupts an exon splice enhancer site, preventing normal post-transcription processing.4 In the general population the copy number of SMN1 and SMN2 genes varies. Published information suggests that 85–95% of the population has one copy of the SMN1 gene per allele; however, individuals with more than one copy per allele have been identified.5–10 Ninety-four per cent of affected individuals possess bi-allelic mutations involving various deletion mutations that share the loss of SMN1 exon 7. Of the remaining 6% of SMA cases, most are attributed to a variety of rare molecular lesions distributed throughout the gene, including point mutations, small insertions, and deletions. A minority of cases may involve genes other than SMN1.11 Additionally, there is an inverse correlation between SMN2 copy number and disease severity, such that affected individuals with three or more copies of SMN2 typically have milder forms of the disease.12 This mutation spectrum permits sensitive detection of carriers through the analysis of SMN1 exon 7 copy number by quantitative polymerase chain reaction (PCR).13–14 In calculating the mean from the results of their own and other published research, Ogino et al reported a frequency for one-copy carriers at approximately 1 in 40 individuals in the general population.15 In addition to the low prevalence mutations distributed across the gene, another small subset of carriers cannot be identified by exon 7 copy measurement since they have one allele with two SMN1 copies paired with a zero-copy allele (hereafter referred to as “2+0” genotype). By exon 7 quantification these individuals are indistinguishable from wild-type (or “1+1” genotype). It has been estimated that approximately 1 in 600 apparent two-copy specimens may be “2+0” carriers.16

Recently published guidelines from the American College of Medical Genetics (ACMG) recommend universal carrier screening for SMA.17 Before this guidance, carrier screening in the USA was offered primarily to family members of individuals diagnosed with SMA. SMA is considered to be pan-ethnic; however, published studies to date have involved small and heterogeneous sample sets.4–5 16 19 Even the few studies addressing specific races provide insufficient information to
support the calculation of carrier frequencies and risk assessments with the precision warranted by widespread carrier screening.\textsuperscript{14, 16–20}

To address this need for accurate allele frequency data, more than 1000 samples from five major ethnic groups, African American, Ashkenazi Jewish, Asian, Caucasian and Hispanic, were tested for \textit{SMN1} copy number. These ethnicities are relevant since they comprise $>95\%$ of the North American population at large and the North American patients electing to participate in genetic testing (2006 US Census estimates).

### METHODS

Samples were collected from residual material following routine clinical testing of individuals presumed to have no family history of SMA. All specimens were made completely anonymous before testing in accordance with approved institutional protocols. Ethnic assignment relied upon patient reported data that were not collected as part of complete family histories. However, these ethnic assignments, which reflect clinical practice, are therefore highly representative of the anticipated ethnic distribution of SMA carriers in the general population. Ethnic assignment relied upon patient reported data and are therefore highly representative of the anticipated ethnic distribution of SMA carriers in the general population.

### RESULTS

Significant differences in the frequencies of \textit{SMN1} genotypes were observed among several ethnicities (table 1). The highest one-copy carrier rate was identified in specimens from the Caucasian group with a frequency of 1 in 37 samples (2.7%). This result agrees closely with those previously reporting on European populations.\textsuperscript{67} The 1.8% carrier frequency detected in the African American group is 3.4 to 8.4 times more prevalent when compared to the other ethnic groups. The preponderance of the two-copy allele in the African American group also suggests a much higher frequency of alleles with two or more \textit{SMN1} copies than the other four ethnic groups. These unexpected results in this group were confirmed by retesting 10\% of the samples using multiplexed ligation dependant probe amplification (MLPA, MRC Holland, The Netherlands) as an alternate technique. All MLPA results were concordant with the original real-time PCR data.

Using the observed genotype data and assuming Hardy–Weinberg equilibrium, maximum likelihood estimation was employed to determine frequencies for alleles and allele pairings within all sample groups (tables 2 and 3). These calculations reveal that the expected frequency of alleles with two or more copies of \textit{SMN1} in the African American group is 3.4 to 8.4 times more prevalent when compared to the other ethnic groups. The preponderance of the two-copy allele in the African American group also suggests a much higher frequency of individuals with the SMA carrier “2+0” genotype compared to other ancestries.

Adjusted carrier risk assessment estimates have been previously published based on aggregate results of several studies.\textsuperscript{15, 16} These estimates calculate the probability of being an SMA carrier when an individual without a family history of

\begin{table}[h]
\centering
\caption{Frequency of \textit{SMN1} copy number across various ethnicities}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Ethnicity & 1 copy & 2 copies & 3 copies & Total \n
\hline
Caucasian & 28 & 0.027 (0.019 to 0.039) & 936 & 0.910 (0.89 to 0.93) & 65 & 0.063 (0.05 to 0.08) & 1028 \\
Ashkenazi Jewish & 22 & 0.022 (0.015 to 0.033) & 827 & 0.825 (0.80 to 0.85) & 153 & 0.153 (0.13 to 0.18) & 1002 \\
Asian & 18 & 0.018 (0.011 to 0.028) & 876 & 0.873 (0.85 to 0.89) & 112 & 0.109 (0.09 to 0.13) & 1027 \\
African American & 11 & 0.011 (0.006 to 0.019) & 529 & 0.521 (0.49 to 0.55) & 475 & 0.468 (0.44 to 0.50) & 1015 \\
Hispanic & 8 & 0.008 (0.004 to 0.015) & 670 & 0.845 (0.82 to 0.87) & 152 & 0.148 (0.13 to 0.17) & 1030 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\caption{Frequency of \textit{SMN1} allele pairings}
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
Allele pairings & Caucasian & Ashkenazi Jewish & Asian & African American & Hispanic \n
\hline
2+2 & 0.0011 & 0.0065 & 0.0032 & 0.0748 & 0.0059 \\
2+1* & 0.0621 & 0.1462 & 0.0158 & 0.3931 & 0.1416 \\
1+1* & 0.0981 & 0.8230 & 0.8720 & 0.5169 & 0.8439 \\
Total & 0.9713 & 0.9756 & 0.9810 & 0.9848 & 0.9914 \\
\hline
2+1** & 0.0007 & 0.0019 & 0.0011 & 0.0041 & 0.0007 \\
2+0** & 0.0005 & 0.0004 & 0.0003 & 0.0002 & 0.0001 \\
1+1** & 0.0272 & 0.0220 & 0.0175 & 0.0108 & 0.0078 \\
Total & 0.0286 & 0.0243 & 0.0189 & 0.0152 & 0.0086 \\
\hline
1*/1** & 0.0008 & 0.0048 & 0.0048 & 0.0048 & 0.0048 \\
1*/0** & 0.0008 & 0.0048 & 0.0048 & 0.0048 & 0.0048 \\
0+0** & 0.0008 & 0.0048 & 0.0048 & 0.0048 & 0.0048 \\
Total & 0.0024 & 0.0144 & 0.0144 & 0.0144 & 0.0144 \\
\hline
\end{tabular}
\end{table}
SMA receives a test result showing two or more SMN1 copies. These calculations account for rare mutations detectable by the method described here, and model the “2+0” carrier genotype. Adjusted carrier risk was calculated for each ethnic group based on its unique allele frequencies (Table 4). The result for the Caucasian population (1:652) is similar to that previously reported by Smith et al (1:648). In that report calculations were based on data from a compilation of studies from countries in Europe, USA, and Australia. The adjusted risk estimate in the Ashkenazi Jewish population (1:350) is approximately two times higher than the Caucasian group because of the higher frequency of two-copy alleles. As expected, the unusually high frequency of two-copy alleles in the African American population produces an adjusted risk factor more than five times greater than that of Caucasians or Asians. Additionally, a “3+0” genotype carrier has been previously inferred but not been formally demonstrated by molecular techniques. The adjusted carrier risk calculations for the three-copy genotype are presented in the right hand column of Table 4. Finally, using these calculated frequencies, the negative predictive value and detection rate were determined for real-time PCR quantification of exon 7 as a screening method for each ethnic group (Table 5).

**DISCUSSION**

The genetic basis for the unusually high SMN1 copy number in the African American population is unknown. Previously published data have shown that individuals with higher copy numbers of SMN1 tend to have fewer copies of SMN2. It has been suggested that this correlation indicates that SMN2 may have converted to SMN1; however, the inverse may also be true. Alleles with multiple copies of SMN1 may be the ancestral form of the duplication, with conversion to SMN2 a more recent mutation event. An assessment of SMN2 copy number is currently underway to understand the SMN1/SMN2 correlation in this sample set.

Prediction of disease phenotype in SMA is complicated by the modifying effects of SMN2 copy number, as well as other genes such as plastin 3. Additionally, there is not a completely predictive correlation between phenotype and modifiers. Similar to other autosomal diseases with variable phenotypes, SMA carrier screening results cannot predict the disease phenotype for offspring. Although this study has demonstrated that significant differences exist in SMA carrier frequency between several ethnic groups, an assessment of pregnancy outcomes data for these populations will be needed to determine if disease frequencies or phenotype incidences also vary. More widespread carrier screening will aid in identifying couples at risk of having SMA offspring and allow for appropriate follow-up to answer these questions.

The recent ACMG guideline for SMA carrier screening recommends that since SMA is present in all populations, carrier testing should be offered to couples regardless of race or ethnicity who are pregnant or considering pregnancy. The results from this study provide precise estimates of allele frequencies and genotypes that vary significantly among ethnic groups. The implications of this work are most significant for Hispanic patients who carry lower risks of SMN1 mutations and for African American patients who bear increased frequencies of two-copy alleles. These data will facilitate the accurate interpretation of clinical testing results and provide additional information for genetic counseling.

**Competing interests:** All authors, except C S, declare the potential for duality of interest because they are employed by, and receive salaries from, Genzyme Genetics, the institution that funded this research. Genzyme is engaged in clinical testing of SMN1 for SMA carrier screening.

**Patient consent:** Not required.

**Provenance and peer review:** Not commissioned; externally peer reviewed.

### REFERENCES


### Table 4 Prior and adjusted risk for spinal muscular atrophy carrier status when two or three SMN1 copies are detected by quantitative assays

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Prior risk</th>
<th>2 copies</th>
<th>3 copies*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian</td>
<td>1:35</td>
<td>1:632</td>
<td>1:3500</td>
</tr>
<tr>
<td>Ashkenazi Jewish</td>
<td>1:41</td>
<td>1:350</td>
<td>1:4000</td>
</tr>
<tr>
<td>Asian</td>
<td>1:53</td>
<td>1:628</td>
<td>1:5000</td>
</tr>
<tr>
<td>African American</td>
<td>1:66</td>
<td>1:121</td>
<td>1:3000</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1:117</td>
<td>1:1061</td>
<td>1:11000</td>
</tr>
</tbody>
</table>

*The risk estimates were rounded to two significant digits due to approximations in their calculations
1Prior risk is the probability of having any carrier genotype (1+0, 1+1*, 2+0 and 2+1*).

### Table 5 Negative predictive value and detection rate by quantitative testing

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Negative predictive value (%)</th>
<th>Detection rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caucasian</td>
<td>99.85</td>
<td>94.92</td>
</tr>
<tr>
<td>Ashkenazi Jewish</td>
<td>99.76</td>
<td>90.16</td>
</tr>
<tr>
<td>Asian</td>
<td>99.86</td>
<td>92.55</td>
</tr>
<tr>
<td>African American</td>
<td>99.56</td>
<td>71.12</td>
</tr>
<tr>
<td>Hispanic</td>
<td>99.92</td>
<td>90.57</td>
</tr>
</tbody>
</table>