Investigation of Established Genetic Risk Variants for Glioma in Prediagnostic Samples from a Population-Based Nested Case–Control Study

Carl Wibom1,2, Florentin Späth1, Anna M. Dahlin1,2, Hilde Langseth3, Eivind Hovig4,5, Preetha Rajaraman6, Tom Børge Johannesen7, Ulrika Andersson1, and Beatrice Melin1

Abstract

Background: Although glioma etiology is poorly understood in general, growing evidence indicates a genetic component. Four large genome-wide association studies (GWAS) have linked common genetic variants with an increased glioma risk. However, to date, studies are based largely on a case-control design, where cases have been recruited at the time of or after diagnosis. They may therefore suffer from a degree of survival bias, introduced when rapidly fatal cases are not included.

Methods: To confirm glioma risk variants in a prospective setting, we have analyzed 11 previously identified risk variants in a set of prediagnostic serum samples with 598 cases and 595 matched controls. Serum samples were acquired from The Janus Serum Bank, a Norwegian population-based biobank reserved for cancer research.

Results: We confirmed the association with glioma risk for variants within five genomic regions: 8q24.21 (CCDC26), 9p21.3 (CDKN2B-AS1), 11q23.3 (PHLDB1), 17p13.1 (TP53), and 20q13.33 (RTEL1). However, previously identified risk variants within the 7p11.2 (EGFR) region were not confirmed by this study.

Conclusions: Our results indicate that the risk variants that were confirmed by this study are truly associated with glioma risk and may, consequently, affect gliomagenesis. Though the lack of positive confirmation of EGFR risk variants may be attributable to relatively limited statistical power, it nevertheless raises the question whether they truly are risk variants or markers for glioma prognosis.

Impact: Our findings indicate the need for further studies to clarify the role of glioma risk loci with respect to prolonged survival versus etiology. Cancer Epidemiol Biomarkers Prev; 24(5): 810–6. ©2015 AACR.

Introduction

The prognosis for patients with glioma is poor. Given standard care (including surgery, radiation, and chemotherapy), patients with glioblastoma (GBM), the most common glioma subtype, have a 5-year survival rate of less than 10% (1) and the median survival is 12 to 15 months (2–4). One obstacle in the development of treatments and surveillance strategies is that the etiology of the disease is generally poorly understood. The only established environmental risk factor associated with glioma is exposure to moderate to high doses of ionizing radiation (5–7). It has also been shown that glioma etiology contains a genetic component, in part through observations of an elevated glioma risk among individuals with a family history of glioma (8–10). An improved understanding of glioma etiology has the potential to facilitate treatment development, and thereby to improve the outcome of the disease.

The genetic component in glioma etiology has been further substantiated through four genome-wide association studies (GWAS), identifying eight genetic variants within six genomic regions associated with glioma risk (refs. 11–14; see corresponding references in Table 1). The two variants at 9p21.3 (rs4977756 and rs1412829) were found independently, but are in linkage disequilibrium (D’ = 0.76), whereas the two variants at 7p11.2 (rs11979158 and rs2252586) were found through the same study, but are less frequently co-inherited (D’ = 0.56). Variants in 20q13.3 (RTEL1), 5p15.33 (TERT), and 9p21.3 (CDKN2B-AS1) are largely shown to be primarily associated with higher grade tumors (11, 13, 15, 16), though Jenkins and colleagues found rs2736100 (CDKN2B) in 7p11.2 to display a similar association with all glioma subtypes (16) and Simon and colleagues found rs4977756 (CDKN2B) to not be correlated with tumor grade (15). Conversely, variants in 11q23.3 (PHLDB1) and 8q24.21 (CCDC26) are primarily associated with lower grade tumors and oligodendroglioma (11, 13, 15–17).

Further genetic variants that are shown to be associated with glioma risk include a functional variant annotating the 3’ untranslated region of the TP53 gene (rs78378222) (18), and two additional EGFR variants (rs4947979 and rs4947986) that were identified by a candidate gene approach (19).
The presently known glioma risk variants mentioned above (Table 1) have been primarily identified through studies using a case-control design for both the discovery and the replication phases. Case-control studies typically recruit cases to the study at the time of, or subsequent to, diagnosis. This strategy incorporates a risk that cases that rapidly succumb to the disease will not be included in the study, i.e., survival bias. This risk is particularly pronounced when studying diseases with very short survival, such as glioma. For example in pancreatic cancer, known to also have a very poor prognosis, the rs1533 (CLPTM1L/TERT) genotype rs41681 is shown to display an association with risk based on postdiagnostic samples but not based on prediagnostic samples (20).

Survival bias may ultimately result in erroneous conclusions, as it is not possible to distinguish between associations with risk and associations with prognosis. To accurately confirm previously reported glioma risk variants, and ascertain whether they are associated with risk or with prolonged survival, we conducted an investigation of these variants in a prediagnostic set of serum samples collected from 17 of Norway's 19 counties. Donators to the biobank were recruited both from Red Cross Blood Bank donors (1973–1991, men and women, ages 18–65) and various health examination studies, including The Oslo Study (1972–1973, only men, ages 40–49), population-based national health screening programs (1974–1991, men and women, predominantly individuals in their forties), and an investigation of health and life conditions in Troms county (2002–2004). Samples included in this study were collected between 1972 and 2004 from individuals between 18.8 and 74.3 years of age at the time of blood draw (50% of the samples were donated by individuals between 39.9 and 42.8 years of age).

The Cancer Registry of Norway was used for identifying cases. The samples were collected at various time points before diagnosis of cases, where the average time between sample collection and diagnosis was 15 years (Table 2). Cases and controls were matched for sex, year of birth (within 15 months), county, and date of sample collection (within 4 months). All samples were analyzed for 11 single-nucleotide polymorphisms, as detailed in Table 1, using cycling temperature capillary electrophoresis, as described by Bjørheim and colleagues (23).

The study was approved by the ethical review board at the University of Oslo, Norway.

### Table 1. Included genetic variants, reported to be associated with glioma risk

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cytoband</th>
<th>Gene</th>
<th>Alleles (major/minor)</th>
<th>MAF (cases/controls)</th>
<th>Call rate (%)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2373600</td>
<td>5p15.33</td>
<td>TERT</td>
<td>G/T</td>
<td>0.44 (0.48)</td>
<td>96.4</td>
<td>(14)</td>
</tr>
<tr>
<td>rs4947979</td>
<td>7p11.2</td>
<td>EGFR</td>
<td>A/G</td>
<td>0.21 (0.19)</td>
<td>92.3</td>
<td>(19)</td>
</tr>
<tr>
<td>rs4947926</td>
<td>7p11.2</td>
<td>EGFR</td>
<td>C/T</td>
<td>0.24 (0.28)</td>
<td>99.7</td>
<td>(19)</td>
</tr>
<tr>
<td>rs2252586</td>
<td>7p11.2</td>
<td>EGFR (107 kb telomeric)</td>
<td>C/T</td>
<td>0.26 (0.26)</td>
<td>100.0</td>
<td>(15)</td>
</tr>
<tr>
<td>rs7939585</td>
<td>7p11.2</td>
<td>EGFR</td>
<td>T/C</td>
<td>0.18 (0.19)</td>
<td>99.7</td>
<td>(13)</td>
</tr>
<tr>
<td>rs4925627</td>
<td>8q24.21</td>
<td>CCDC26 (320 kb telomeric)</td>
<td>T/G</td>
<td>0.22 (0.16)</td>
<td>99.2</td>
<td>(14)</td>
</tr>
<tr>
<td>rs4977756</td>
<td>9p21.3</td>
<td>CDKN2B-AS1</td>
<td>A/G</td>
<td>0.51 (0.43)</td>
<td>99.7</td>
<td>(14)</td>
</tr>
<tr>
<td>rs1412829</td>
<td>9p21.3</td>
<td>CDKN2B-AS1</td>
<td>T/C</td>
<td>0.52 (0.43)</td>
<td>99.7</td>
<td>(12)</td>
</tr>
<tr>
<td>rs498872</td>
<td>1q32.3</td>
<td>PHD6</td>
<td>G/A</td>
<td>0.37 (0.31)</td>
<td>100.0</td>
<td>(14)</td>
</tr>
<tr>
<td>rs17737222</td>
<td>17p13.1</td>
<td>TP53</td>
<td>A/C</td>
<td>0.04 (0.02)</td>
<td>99.6</td>
<td>(18)</td>
</tr>
<tr>
<td>rs6010620</td>
<td>20q13.33</td>
<td>RET1</td>
<td>G/A</td>
<td>0.20 (0.27)</td>
<td>99.7</td>
<td>(12, 14)</td>
</tr>
</tbody>
</table>

*Minor allele frequency.

*Reference to the report that originally identified the association with risk.

**Table 2. Basic characteristics describing the samples included in the separate glioma subtypes analyzed**

<table>
<thead>
<tr>
<th></th>
<th>Cases (n)</th>
<th>Age at diagnosis¹</th>
<th>Age at sampling²</th>
<th>Sampling before diagnosis (y)³</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases</td>
<td>594</td>
<td>56.3 (27.7–84.1)</td>
<td>41.4 (18.9–74.5)</td>
<td>14.7 (0.2–35.1)</td>
</tr>
<tr>
<td>Males</td>
<td>396</td>
<td>56.0 (27.7–84.1)</td>
<td>41.6 (18.9–74.5)</td>
<td>14.3 (0.2–35.1)</td>
</tr>
<tr>
<td>Females</td>
<td>208</td>
<td>56.1 (27.7–84.1)</td>
<td>41.0 (22.8–70.7)</td>
<td>15.2 (0.2–34.7)</td>
</tr>
<tr>
<td>All glioma⁴</td>
<td>566</td>
<td>56.4 (27.7–84.1)</td>
<td>41.4 (18.9–74.5)</td>
<td>14.8 (0.2–35.1)</td>
</tr>
<tr>
<td>Males</td>
<td>378</td>
<td>56.1 (27.7–84.1)</td>
<td>41.6 (18.9–74.5)</td>
<td>14.5 (0.2–35.1)</td>
</tr>
<tr>
<td>Females</td>
<td>188</td>
<td>56.5 (32.5–83.2)</td>
<td>41.0 (22.8–70.7)</td>
<td>15.2 (0.2–34.7)</td>
</tr>
<tr>
<td>GBM</td>
<td>376</td>
<td>57.6 (34.7–84.3)</td>
<td>41.7 (19.2–67.2)</td>
<td>15.7 (0.4–35.0)</td>
</tr>
<tr>
<td>Males</td>
<td>268</td>
<td>57.5 (34.7–84.1)</td>
<td>41.9 (19.2–67.2)</td>
<td>15.7 (0.4–35.0)</td>
</tr>
<tr>
<td>Females</td>
<td>108</td>
<td>57.6 (36.8–83.2)</td>
<td>41.1 (27.6–66.3)</td>
<td>15.8 (0.4–34.7)</td>
</tr>
<tr>
<td>Oligo</td>
<td>33</td>
<td>51.8 (27.7–78.3)</td>
<td>41.1 (18.9–66.4)</td>
<td>12.0 (0.2–23.1)</td>
</tr>
<tr>
<td>Males</td>
<td>24</td>
<td>51.1 (27.7–78.3)</td>
<td>41.1 (18.9–66.4)</td>
<td>9.9 (0.2–23.0)</td>
</tr>
<tr>
<td>Females</td>
<td>9</td>
<td>54.1 (44.2–64.5)</td>
<td>41.4 (27.4–62.3)</td>
<td>15.2 (0.2–25.1)</td>
</tr>
<tr>
<td>Ependymoma</td>
<td>28</td>
<td>54.2 (43.3–70.5)</td>
<td>41.4 (28.5–55.3)</td>
<td>13.0 (3.4–24.6)</td>
</tr>
<tr>
<td>Males</td>
<td>18</td>
<td>53.2 (45.5–70.5)</td>
<td>42.0 (39.9–55.3)</td>
<td>9.7 (3.4–24.6)</td>
</tr>
<tr>
<td>Females</td>
<td>10</td>
<td>57.0 (43.3–67.2)</td>
<td>40.7 (28.5–51.1)</td>
<td>15.6 (4.2–20.1)</td>
</tr>
<tr>
<td>Astrocytoma</td>
<td>93</td>
<td>51.1 (27.7–78.9)</td>
<td>41.1 (22.4–74.5)</td>
<td>10.9 (0.2–35.1)</td>
</tr>
<tr>
<td>Males</td>
<td>56</td>
<td>51.8 (27.7–78.9)</td>
<td>41.2 (22.4–74.5)</td>
<td>11.1 (0.7–35.1)</td>
</tr>
<tr>
<td>Females</td>
<td>37</td>
<td>48.2 (33.0–72.4)</td>
<td>41.1 (22.8–70.7)</td>
<td>10.4 (0.2–32.8)</td>
</tr>
</tbody>
</table>

¹Median (min – max).
²Not including ependymoma.
Figure 1.
Associations between published genetic risk variants and risk of disease, calculated by conditional logistic regression. Gray boxes denote associations that are opposite in direction of those of the original report. Purple confidence intervals denote that both OR and CI have been inverted in the plot, for easier comparison. ORs and \( P \) values represent the results of examining associations between genotypes and risk of disease by conditional logistic regression under a log-additive genetic model (accounting for sample matching).

### Glioma OVERALL

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>OR</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>0.82</td>
<td>0.07883</td>
</tr>
<tr>
<td>rs4947979</td>
<td>EGFR</td>
<td>1.24</td>
<td>0.09119</td>
</tr>
<tr>
<td>rs2252586</td>
<td>CDKN2B</td>
<td>0.84</td>
<td>0.11319</td>
</tr>
<tr>
<td>rs11979158</td>
<td>EGFR</td>
<td>0.95</td>
<td>0.67165</td>
</tr>
<tr>
<td>rs295627</td>
<td>CCDC26</td>
<td>1.69</td>
<td>0.23039</td>
</tr>
<tr>
<td>rs4977756</td>
<td>CDKN2B</td>
<td>0.84</td>
<td>0.61328</td>
</tr>
<tr>
<td>rs498872</td>
<td>PHLDB1</td>
<td>1.47</td>
<td>0.29489</td>
</tr>
<tr>
<td>rs6010620</td>
<td>TP53</td>
<td>1.73</td>
<td>0.10121</td>
</tr>
</tbody>
</table>

### GBM

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>OR</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>1.47</td>
<td>0.34381</td>
</tr>
<tr>
<td>rs4947979</td>
<td>EGFR</td>
<td>1.17</td>
<td>0.78191</td>
</tr>
<tr>
<td>rs2252586</td>
<td>CDKN2B</td>
<td>0.75</td>
<td>0.40304</td>
</tr>
<tr>
<td>rs11979158</td>
<td>EGFR</td>
<td>1.78</td>
<td>0.21273</td>
</tr>
<tr>
<td>rs295627</td>
<td>CCDC26</td>
<td>2.85</td>
<td>0.02137</td>
</tr>
<tr>
<td>rs4977756</td>
<td>CDKN2B</td>
<td>0.84</td>
<td>0.61328</td>
</tr>
<tr>
<td>rs498872</td>
<td>PHLDB1</td>
<td>1.47</td>
<td>0.29489</td>
</tr>
<tr>
<td>rs6010620</td>
<td>TP53</td>
<td>1.70</td>
<td>0.49124</td>
</tr>
</tbody>
</table>

### OLGIO.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>OR</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>0.82</td>
<td>0.34381</td>
</tr>
<tr>
<td>rs4947979</td>
<td>EGFR</td>
<td>1.17</td>
<td>0.78191</td>
</tr>
<tr>
<td>rs2252586</td>
<td>CDKN2B</td>
<td>0.75</td>
<td>0.40304</td>
</tr>
<tr>
<td>rs11979158</td>
<td>EGFR</td>
<td>1.78</td>
<td>0.21273</td>
</tr>
<tr>
<td>rs295627</td>
<td>CCDC26</td>
<td>2.85</td>
<td>0.02137</td>
</tr>
<tr>
<td>rs4977756</td>
<td>CDKN2B</td>
<td>0.84</td>
<td>0.61328</td>
</tr>
<tr>
<td>rs498872</td>
<td>PHLDB1</td>
<td>1.47</td>
<td>0.29489</td>
</tr>
<tr>
<td>rs6010620</td>
<td>TP53</td>
<td>1.70</td>
<td>0.49124</td>
</tr>
</tbody>
</table>

### Astrocytoma

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>OR</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>0.88</td>
<td>0.53244</td>
</tr>
<tr>
<td>rs4947979</td>
<td>EGFR</td>
<td>0.92</td>
<td>0.77741</td>
</tr>
<tr>
<td>rs2252586</td>
<td>CDKN2B</td>
<td>0.77</td>
<td>0.28139</td>
</tr>
<tr>
<td>rs11979158</td>
<td>EGFR</td>
<td>1.07</td>
<td>0.79632</td>
</tr>
<tr>
<td>rs295627</td>
<td>CCDC26</td>
<td>1.19</td>
<td>0.47448</td>
</tr>
<tr>
<td>rs4977756</td>
<td>CDKN2B</td>
<td>1.55</td>
<td>0.08892</td>
</tr>
<tr>
<td>rs4142829</td>
<td>CDKN2B</td>
<td>1.59</td>
<td>0.03488</td>
</tr>
<tr>
<td>rs498872</td>
<td>PHLDB1</td>
<td>1.80</td>
<td>0.01131</td>
</tr>
<tr>
<td>rs6010620</td>
<td>TP53</td>
<td>2.67</td>
<td>0.14740</td>
</tr>
</tbody>
</table>

### Ependymoma

<table>
<thead>
<tr>
<th>Locus</th>
<th>Gene</th>
<th>OR</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>0.94</td>
<td>0.86583</td>
</tr>
<tr>
<td>rs4947979</td>
<td>EGFR</td>
<td>0.50</td>
<td>0.25767</td>
</tr>
<tr>
<td>rs494796</td>
<td>EGFR</td>
<td>0.66</td>
<td>0.37816</td>
</tr>
<tr>
<td>rs2252586</td>
<td>CDKN2B</td>
<td>1.30</td>
<td>0.53462</td>
</tr>
<tr>
<td>rs11979158</td>
<td>EGFR</td>
<td>0.92</td>
<td>0.83493</td>
</tr>
<tr>
<td>rs295627</td>
<td>CCDC26</td>
<td>0.54</td>
<td>0.23673</td>
</tr>
<tr>
<td>rs4977756</td>
<td>CDKN2B</td>
<td>0.86</td>
<td>0.69557</td>
</tr>
<tr>
<td>rs4142829</td>
<td>CDKN2B</td>
<td>0.86</td>
<td>0.69557</td>
</tr>
<tr>
<td>rs498872</td>
<td>PHLDB1</td>
<td>1.41</td>
<td>0.41968</td>
</tr>
<tr>
<td>rs6010620</td>
<td>TP53</td>
<td>2.00</td>
<td>0.42349</td>
</tr>
</tbody>
</table>

Downloaded from cebp.aacrjournals.org on August 14, 2017. © 2015 American Association for Cancer Research.
DNA amplification

Five microliter aliquots of serum was transferred into 96-well microplates (Asygen; VWR) and subjected to enzymatic amplification, as described by Ekstrom and colleagues (24). Serum aliquots were denatured at 95°C for 8 minutes. Then, 40 μL of PCR master mix was added to each well, and the plate was subjected to temperature cycling as described below.

The PCR mix contained 4.5 μL of 10× buffer (Applied Biosystems), 0.8 μL of 10 mmol/L dNTP mix (Amersham Biosciences), and a varying concentration of MgCl2 (Supplementary Table S1). Primers were added to a final volume of 45 μL (Supplementary Table S1). A combination of Taq and Pfu polymerase (both produced in-house) was added to the reaction in varying concentrations (Supplementary Table S1).

Amplification was performed on a DNA Engine Tetrad 2 (Biorad), using the following cycling conditions: denaturation for 3 minutes at 95°C, followed by 44 or 45 cycles of 30 seconds at 95°C, 30 seconds at the annealing temperature, and 60 seconds at 72°C. The cycling was followed by 10 minutes at 72°C for completion of polymerization, 5 minutes at 95°C for double-strand denaturation, 30 minutes at 65°C for heteroduplex formation, and lastly a stepwise temperature decrease of 1°C per minute until a final hold at 12°C (Supplementary Table S1). Primers used for the analysis are shown in Supplementary Table S1. In each primer set, one of the primers was labelled with 6-carboxyfluorescein (6-FAM). All primers were obtained from Integrated DNA Technologies.

Variant detection

Amplified 6-FAM-labelled PCR products were analyzed by denaturant capillary electrophoresis on a MegaBACE 1000 DNA Analysis System (GE Healthcare Bio-Sciences AB). The base variants were separated by cycling temperature capillary electrophoresis (CTCE), with separating temperatures listed in Supplementary Table S1, cycled 20 times. The variants were identified by co-analysis with a mutated internal standard, essentially as described by Bjørheim and colleagues (23). The assay was run in a 96-well format, where a minimum of two wells per microplate were used for controls, i.e., one serum control (pool of sera from 5 healthy individuals) and at least one negative control without serum template.

Quality control

To investigate potential genotyping errors, we calculated call rates for both samples and genotypes. Samples displaying a call rate of <80% (i.e., where the genotyping had failed at >2 SNPs) were removed from further analyses. In addition, the genotype frequency distribution for each SNP among controls was calculated and tested against the Hardy-Weinberg equilibrium. P values of <0.005 were considered indicative of genotyping errors (all SNPs displayed HWE P > 0.005 in this study). All 11 genetic variants included in the study displayed a call rate of >90% (Table 1). Four samples displayed a call rate of <80% (i.e., >2 failed genotypes) and were consequently removed from downstream analyses, along with their matched counterparts.

Statistical analyses

Associations between genotypes and risk of disease were examined using conditional logistic regression under a log-additive genetic model (accounting for sample matching). Separate analyses, comparing cases with controls, were performed for the following histologic groups: (i) astrocytoma, (ii) ependymoma, (iii) oligodendroglioma, (iv) GBM, and (v) all gliomas (i.e., including all histologic subgroups except ependymoma). ICDO3 codes included in each histologic group are listed in Supplementary Table S2. All analyses were performed using the R software for statistical computing (www.R-project.org).

Results

A total of 594 cases and 591 controls were eligible for inclusion in the study. The characteristics of the study population are summarized in Table 2. With reference to all included cases, the median age at sampling was 41.4 years (range, 19–75), and the median age of diagnosis was 56.3 years (range, 28–84). The complete results of this investigation are summarized in Fig. 1. Genetic variants in five of the investigated seven regions were confirmed by this study as associated with overall risk of glioma (i.e., all glioma diagnoses, not including ependymoma). The five regions were 8q24.21 (CCDC26; rs4295627), 9p21.3 (CDKN2B-AS1; rs4977756 and rs1412829), 11q23.3 (PHLD1; rs498872), 17p13.1 (TP53; rs7837822), and 20q13.33 (RTEL1; rs6010620). Within the five confirmed regions, the associations we observed were consistent with previous reports in terms of the directions of the odds ratios (OR), and all associations displayed P < 0.05.

Regarding the other histologic subgroups included in the study, only one of the investigated variants displayed an association with oligodendroglioma risk, namely the 8q24.21 (CCDC26) variant rs4295627 (OR, 2.85; 95% CI, 1.76–6.93; P = 0.0214). Both variants annotating 9p21.3 (CDKN2B-AS1) displayed significant association with increased astrocytoma risk [rs4977756 (OR, 1.59; 95% CI, 1.03–2.44; P = 0.0349); rs1412829 (OR, 1.80; 95% CI, 1.14–2.84; P = 0.0113)]. None of the investigated
variants were found to be associated with ependymoma risk. It
should, however, be emphasized that the ependymoma sample
set was very small \( n = 28 \) cases.

**Discussion**

We have analyzed prediagnostic serum samples, with matched
cases and controls from a population-based cohort, to verify
previously reported associations between common genetic var-
iants and glioma risk. Analyses of prediagnostic samples eliminate
the risk of survival bias and thereby enable distinction between
genetic variants associated with glioma risk (i.e., etiology) and
genetic variants associated with prognosis. Our results confirmed
the previously reported associations with risk for variants in five
of seven regions \( \{8q24.21 \text{(CCDC26)}, 9p21.3 \text{(CDKN2B-AS1)},
11q23.3 \text{(PHLD1B)}, 17p13.1 \text{(TERT)}, \text{and} 20q13.33 \text{(RELI1)} \} \),
indicating that these variants are truly associated with glioma risk
and may, consequently, affect gliomagenesis.

All hitherto-presented evidence that the investigated gene var-
iants are associated with glioma risk is largely based on case–
control studies, including postdiagnostic samples (11–14, 18,
19), though the study by Rajaraman and colleagues (GliomaScan; ref. 11) included 30% incident cases from cohort studies. Post-
diagnostic sampling inherently risks excluding rapidly fatal cases,
and thereby introduce a bias toward cases with longer survival.
Therefore, based on case–control studies using postdiagnostic samples, it is impossible to know whether identified associations are
toward risk of disease or toward prognosis. This issue is parti-
cularly pronounced when a variant is detected in a disease with such
difficulties in survival as glioma, and especially GBM, which is the most
common glioma subtype. It is thus plausible that the variants that
did not replicate in this study are in fact primarily associated with
longer survival, rather than with risk of glioma. On the other hand,
it is important to keep in mind that this study is relatively small,
thus suffering from limited statistical power to detect associations,
particularly affecting low-frequency variants and variants with
small effect sizes.

In this study, variants mapping to two of seven investigated regions
were not confirmed, namely 5p15.33 (TERT) and 7p11.2
\((\text{EGFR})\). EGFR is a growth factor receptor that commonly displays
somatic aberrations in glioma, including amplification, trunca-
tion, and point mutations (25). EGFR protein overexpression has
been associated with poor prognosis (26), as have certain
gene variants (27), demonstrating that they may be important
determinants of treatment response and survival. In the present
study, we investigated four variants annotating the 7p11.2 (\(\text{EGFR}\))
region that have previously been identified as associated with
glioma risk; two variants were identified through a GWAS
\((\text{rs2252586 and rs11979158};\text{ ref. 13})\), and two variants were
identified through a candidate gene approach \((\text{rs4947979 and
rs4947986};\text{ ref. 19})\). In terms of linkage disequilibrium, each
variant is practically independent of the other variants.

In this study, using prediagnostic samples, the two \(\text{EGFR}\)
genome variants identified through the candidate gene approach,
\(\text{rs4947979 and rs4947986} \), both displayed ORs that were oppo-
site in direction of those of the original report \((19)\). The associ-
ation between \(\text{rs4947986} \) and decreased risk of overall glioma was
significant in this study \((P = 0.0456)\). However, considering that
the original report found an association with increased risk, and
that we had no hypothesis for an association with decreased risk,
we believe this to be a chance finding.

Moreover, one of the GWAS findings, \(\text{rs2252586} \) (located 107
kb telomeric from \(\text{EGFR}\), displayed no association with risk in
any of the investigated histologic groups. This, however, does not
concur with the result of GliomaScan \((11)\), where \(\text{rs2252286} \)
displayed a significant association with glioma risk when the
analysis was restricted to prediagnostic samples from incident
cohort cases \((n = 556)\). The other GWAS finding at the same locus
\((7p11.2; \text{EGFR})\), \(\text{rs11979158} \), displayed what may be considered a
marginally significant association with GBM risk \((P = 0.0670)\),
though it did not appear to be associated with glioma risk overall
\((P = 0.3737)\). The latter is in agreement with the corresponding
results from GliomaScan \((11)\).

Altogether, based on this study, we could not positively confirm
any of the reported associations between common genetic var-
iants annotating 7p11.2 (\(\text{EGFR}\)) and glioma risk. This could, of
course, be due to the original reports being false positives.
Alternatively, it may be attributable to the relatively limited
statistical power of the current study. Nevertheless, the lack of
a positive confirmation of these \(\text{EGFR}\) gene variants in a pre-
diagnostic setting raises the question whether they may be pri-
marily associated with glioma prognosis.

\(\text{Sp15.33} \) \((\text{TERT})\) was the other region harboring a variant
previously associated with risk that was not positively confirmed
by this study. The \(\text{TERT}\) variant \(\text{rs2736100} \) displayed consistency
with previous reports regarding the direction of the OR and effect
size; however, the observed associations with both glioma overall
\((P = 0.0655)\) and specifically with GBM \((P = 0.0788)\) may merely
be considered marginally significant. \(\text{TERT}\) is known to be
involved in telomere regulation, and certain genetic variants
mapping to \(\text{TERT}\) \((\text{rs2736100})\) have been linked with
relative telomere length in an age-dependent manner \((28)\). This
indicates a dynamic change in functionality of \(\text{TERT}\) variants
across a person’s lifespan. The complexity is emphasized further
by the fact that there are contradictory results regarding the
association between relative telomere length and glioma risk
\((29, 30)\). In contrast with our results, providing evidence that
\(\text{rs2736100} \) \((\text{Sp15.33}, \text{TERT})\) is accurately linked with risk, are
the findings of Rajaraman and colleagues \((11)\). When comparing the
strength of \(\text{rs2736100} \) \((\text{Sp15.33}, \text{TERT})\) association with glioma
risk between incident cases from cohort studies (prediagnostic
samples; \(n = 556)\) and cases from case–control studies (post-
diagnostic samples; \(n = 1,300)\), Rajaraman and colleagues found
the association to be stronger in the cohort studies \((11)\), indicating
that the association is preferentially with high-grade, rapidly fatal
cases. The study presented here is based on a relatively limited
number of samples, which makes it difficult to draw decisive
conclusions regarding marginally significant findings with mod-
est effect sizes such as this.

As mentioned, it is important to emphasize that the study
presented here was relatively small \((n = 594)\) cases, thus limiting
the power to detect weak associations. Specifically, it is possible
that the association between \(\text{rs11979158} \) \((7p11.2; \text{EGFR})\) and risk
for GBM \((P = 0.0670)\) could also reach statistical significance in a
prediagnostic setting given a larger sample size. Similarly, the lack
of a positive confirmation of associations between \(\text{rs2736100}
\) \((\text{Sp15.33}, \text{TERT})\) and both glioma overall \((P = 0.0655)\) and GBM
\((P = 0.0788)\) is likely attributable to small sample sizes, especially
when factoring in previous results indicating an association
between \(\text{rs2736100} \) and glioma risk among incident cohort cases
\((11)\). Furthermore, the \(17p13.1 \) \((\text{TERT})\) variant \(\text{rs78378222} \) has
previously been associated with risk for both GBM and non–GBM
In this study, we found an association with risk for glioma overall (i.e., all glioma diagnoses, not including ependymoma), but it did not replicate specifically for GBM ($P = 0.1012$). This is a very rare variant (minor allele frequency $= 0.03$ in this study), and the investigated sample set only included two individuals who were homozygous for the rare allele. Hence, it is likely that the lack of confirmation of rs78378222 specifically for GBM is due to the limited statistical power of the study.

As discussed above, there are disparities comparing the findings of this study with those of the GliomaScan study (11), where associations between genetic variants and glioma risk were also investigated in samples collected before diagnoses (incident cohort cases). Both the allele frequencies and the distribution of histologic subtypes are similar between the studies, though $17\%$ of the incident cohort cases included in the latter study are of unknown histology, making the comparison somewhat uncertain. The numbers of included cases are also similar in both studies (594 for this study and 556 for incident cohort cases in GliomaScan), though GliomaScan included 5-fold more controls. The present study is based on a homogeneous collection of population-based samples, acquired from the Janus Serum Bank, with carefully matched controls, whereas the incident cohort cases in the GliomaScan study originate from 14 different cohort studies from several centers across the world (11). In addition, the time between date of blood draw and diagnosis for incident cases was shorter in GliomaScan (mean 5 years) than in this study (mean 15 years).

Our study shows that the previously reported associations between glioma risk and common gene variants annotating 8q24.21 ($\text{CCDC26}$), 9p21.3 ($\text{CDKN2B-AS1}$), 11q23.3 ($\text{PHLDB1}$), 17p13.1 ($\text{TP53}$), and 20q13.33 ($\text{RTEL1}$) also hold true in a prediagnostic setting. This indicates that they are accurately linked with risk of disease and may thus be important to help understand glioma etiology. These variants were all significantly associated with overall glioma risk ($P < 0.05$). Most of them were also found to be associated specifically with GBM risk, with the exception of rs78378222 (17p13.1; $\text{TP53}$) as mentioned above, and the association between rs498872 (11q23.3; $\text{PHLDB1}$) and GBM risk was marginally significant ($P = 0.0574$). In addition, variants annotating the two regions 9p21.3 ($\text{CDKN2B-AS1}$) and 8q24.21 ($\text{CCDC26}$) displayed associations with risk of astrocytoma and oligodendroglioma, respectively. No other variants were specifically linked with these histologic groups. The association between rs295627 (8q24.21; $\text{CCDC26}$) and oligodendroglioma was the most pronounced association observed for this variant, which is in agreement with previous findings linking it primarily with oligodendroglioma risk (11, 16).

In this study, we investigated serum samples from the Janus Serum Bank. The samples were stored at $-25^\circ$C for up to 40 years. The samples were genotyped by means of cycling temperature capillary electrophoresis (23), a PCR-based method that has successfully been applied in previous projects investigating Janus serum samples (24, 32). Because the DNA is expected to be degraded, we applied a primer design that yielded short ampli-cons ($< 100$ bases). This approach has proven to achieve a PCR success rate similar to that seen when using full-length DNA (24). In addition to using a method well suited for the samples at hand, cases and controls were matched for date of sample collection (i.e., storage time). Hence, we have no reason to believe that sample storage has influenced our results in any way.

We have analyzed prediagnostic serum samples from a well-defined, population-based cohort, with cases and controls matched on sex, county of blood draw, year of birth, and date of sample collection. We have shown that the associations between established glioma risk variants and risk of glioma also hold true in a prediagnostic setting for variants in five of the seven regions previously associated with glioma risk: 8q24.21 ($\text{CCDC26}$), 9p21.3 ($\text{CDKN2B-AS1}$), 11q23.3 ($\text{PHLDB1}$), 17p13.1 ($\text{TP53}$), and 20q13.33 ($\text{RTEL1}$). This indicates that these variants are true risk factors and may affect glioma etiology. Out of the four investigated variants annotating $\text{EGFR}$, one (rs11979158) showed a marginally significant association with GBM risk, but no association with glioma risk overall. The remaining three variants (rs2252286, rs4947979, and rs4947986) were not found to be associated with risk within any histologic group in this study, and may thus be associated primarily with longer survival. Considering the limited power of the current study, we believe that further studies are warranted to determine whether these risk loci are primarily linked with risk or prognosis.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Authors’ Contributions**

Conception and design: H. Langseth, P. Rajaraman, U. Andersson, B. Melin
Development of methodology: C. Wibom, B. Melin
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Langseth, E. Hovig, T.B. Johannessen, B. Melin
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): C. Wibom, E. Hovig, P. Rajaraman, T.B. Johannessen, B. Melin
Writing, review, and/or revision of the manuscript: C. Wibom, F. Spah, A.M. Dahlén, H. Langseth, E. Hovig, P. Rajaraman, T.B. Johannessen, U. Andersson, B. Melin
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Langseth, U. Andersson
Study supervision: C. Wibom, B. Melin
Other (quality assurance of cancer data and biological samples): H. Langseth

**Acknowledgments**

B. Melin was supported by grants from Acta Oncologica foundation through the Royal Swedish Academy of Science (salary), The Swedish Research Council, Swedish Cancer Foundation, Northern Sweden Cancer Foundation, Umeå University Young investigator reward, and Umeå Hospital Cutting Edge Grant. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 26, 2014; revised December 22, 2014; accepted January 28, 2015, published OnlineFirst February 20, 2015.

**References**


Investigation of Established Genetic Risk Variants for Glioma in Prediagnostic Samples from a Population-Based Nested Case-Control Study

Carl Wibom, Florentin Späth, Anna M. Dahlin, et al.

Cancer Epidemiol Biomarkers Prev 2015;24:810-816. Published OnlineFirst February 20, 2015.

Updated version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-14-1106

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2015/02/21/1055-9965.EPI-14-1106.DC1

Cited articles
This article cites 32 articles, 3 of which you can access for free at:
http://cebp.aacrjournals.org/content/24/5/810.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.