This is the published version of a paper published in Scientific Reports.

Citation for the original published paper (version of record):

Sex-specific glioma genome-wide association study identifies new risk locus at 3p21.31
in females, and finds sex-differences in risk at 8q24.21
Scientific Reports, 8: 7352
https://doi.org/10.1038/s41598-018-24580-z

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-150697
Sex-specific glioma genome-wide association study identifies new risk locus at 3p21.31 in females, and finds sex-differences in risk at 8q24.21

Quinn T. Ostrom1,2,3, Ben Kinnersley4, Margaret R. Wrensch5, Jeanette E. Eckel-Passow6, Georgina Armstrong1, Terri Rice1, Yanwen Chen1, John K. Wiencke5, Helen M. Hansen5, Christopher I. Amos7, Jonine L. Bernstein8, Elizabeth B. Claus9,10, Dora Il'yasova11,12,13, Christoffer Johansen14,15, Daniel H. Lachance16, Rose K. Lai17,18, Ryan T. Merrell19, Sara H. Olson20, Siegal Sadetzki21,22, Ilana S. Schwartz23,24, Justin D. Lathia25, Michael E. Berens26,27, Ulrika Andersson28, Preetha Rajaraman29, Stephen J. Chanock29,30, Martha S. Linet29,30, Zhaoming Wang29,30, Meredith Yeager29,30, GliomaScan consortium*, Richard S. Houlston4, Robert B. Jenkins31, Beatrice Melin29, Melissa L. Bondy1 & Jill. S. Barnholtz-Sloan2

1Department of Medicine, Section of Epidemiology and Population Sciences, Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas, United States of America. 2Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio, United States of America. 3Department of Population and Quantitative Health Sciences, Case Western Reserve University School of Medicine, Cleveland, Ohio, United States of America. 4Division of Genetics and Epidemiology, The Institute of Cancer Research, Sutton, Surrey, United Kingdom. 5Department of Neurological Surgery and Institute of Human Genetics, School of Medicine, University of California, San Francisco, San Francisco, California, United States of America. 6Division of Biomedical Statistics and Informatics, Mayo Clinic College of Medicine, Rochester, Minnesota, United States of America. 7Institute for Clinical and Translational Research, Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, Texas, United States of America. 8Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York, United States of America. 9School of Public Health, Yale University, New Haven, Connecticut, United States of America. 10Department of Neurosurgery, Brigham and Women’s Hospital, Boston, Massachusetts, United States of America. 11Department of Epidemiology and Biostatistics, School of Public Health, Georgia State University, Atlanta, Georgia, United States of America. 12Cancer Control and Prevention Program, Department of Community and Family Medicine, Duke University Medical Center, Durham, North Carolina, United States of America. 13Duke Cancer Institute, Duke University Medical Center, Durham, North Carolina, United States of America. 14Oncology clinic, Finsen Center, Rigshospitalet, Copenhagen, Denmark. 15Survivorship Research Unit, The Danish Cancer Society Research Center, Copenhagen, Denmark. 16Department of Neurology, Mayo Clinic Comprehensive Cancer Center, Mayo Clinic, Rochester, Minnesota, United States of America. 17Department of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. 18Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, United States of America. 19Department of Neurology, NorthShore University HealthSystem, Evanston, Illinois, United States of America. 20Cancer and Radiation Epidemiology Unit, Gertner Institute, Chaim Sheba Medical Center, Tel Hashomer, Israel. 21Department of Epidemiology and Preventive Medicine, School of Public Health, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel. 22Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America. 23Department of Biostatistics, University of Virginia School of Medicine, Charlottesville, Virginia, United States of America. 24Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America. 25Department of Biostatistics, University of Virginia School of Medicine, Charlottesville, Virginia, United States of America. 26Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America. 27Department of Neuroscience, Washington University School of Medicine, St. Louis, Missouri, United States of America. 28Department of Biostatistics, University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America. 29Department of Stem Cell Biology and Regenerative Medicine, Cleveland Clinic Foundation, Cleveland,
Incidence of glioma is approximately 50% higher in males. Previous analyses have examined exposures related to sex hormones in women as potential protective factors for these tumors, with inconsistent results. Previous glioma genome-wide association studies (GWAS) have not stratified by sex. Potential sex-specific genetic effects were assessed in autosomal SNPs and sex chromosome variants for all glioma, GBM and non-GBM patients using data from four previous glioma GWAS. Datasets were analyzed using sex-stratified logistic regression models and combined using meta-analysis. There were 4,831 male cases, 5,216 male controls, 3,206 female cases and 5,470 female controls. A significant association was detected at rs11979158 (7p11.2) in males only. Association at rs55705857 (8q24.21) was stronger in females than in males. A large region on 3p21.31 was identified with significant association in females only. The identified differences in effect of risk variants do not fully explain the observed incidence difference in glioma by sex.

Glioma is the most common type of primary malignant brain tumor in the United States (US), with an average annual age-adjusted incidence rate of 6.0/100,000. Glioma can be broadly classified into glioblastoma (GBM, 61.9% of gliomas in adults 18+ in the US) and lower-grade glioma (non-GBM glioma, 24.2% of adult gliomas) with tumors such as ependymoma (6.3%), unclassified malignant gliomas (5.1%), and pilocytic astrocytoma (1.9%) making up the majority of other cases. Many environmental exposures have been investigated as sources of glioma risk, but the only validated risk factors for these tumors are ionizing radiation (which increases risk), and history of allergies or other atopic disease (which decreases risk). These tumors are significantly more common in people of European ancestry, in males and in older adults. The contribution of common low-penetration SNPs to the heritability of sporadic glioma in persons with no documented family history is estimated to be ~25%. A recent glioma genome-wide association study (GWAS) meta-analysis validated 12 previously reported risk loci, and identified 13 new risk loci. These 25 loci in total are estimated to account for ~30% of heritable glioma risk. This suggests that there are both undiscovered environmental risk (which accounts for ~75% of incidence variance) and genetic risk factors (accounting for ~70% of heritable risk).

Population-based studies consistently demonstrate that incidence of gliomas varies significantly by sex. Most glioma histologies occur with a 30–50% higher incidence in males, and this male preponderance of glial tumors increases with age in adult glioma. Several studies have attempted to estimate the influence of lifetime estrogen and progestogen exposure on glioma risk in women. Results of these analyses have been mixed, and it is not possible to conclusively determine the impact of hormone exposure on glioma risk. Male predominance in incidence occurs broadly across multiple cancer types and is also evident in cancers that occur in pre-pubertal children and in post-menopausal adults. Together these observations suggest that other mechanisms in addition to acute sex hormone actions must be identified to account for the magnitude of sex difference in glioma incidence. Though sex differences exist in glioma incidence, sex differences have not been interrogated in previous glioma GWAS. Sex-specific analyses have the potential to reveal genetic sources of sexual dimorphism in risk, as well as to increase power for detection of loci where effect size or direction may vary by sex. The aim of this analysis is to investigate potential sex-specific sources of genetic risk for glioma that may contribute to observed sex-specific incidence differences.

Results

Study population. There were 4,831 male cases, 5,216 male controls, 3,206 female cases, and 5,470 female controls (Table 1). A slightly larger proportion of male cases were GBM (58.7% of male cases vs 52.5% of female cases). Controls were slightly older than cases. GBM cases had a higher mean age than non-GBM cases, which was consistent with known incidence patterns of these tumors. Male and female cases within histology groups had similar age at diagnosis. The proportion of non-GBM cases varied by study due to differing recruitment patterns as to increase power for detection of loci where effect size or direction may vary by sex. The aim of this analysis is to investigate potential sex-specific sources of genetic risk for glioma that may contribute to observed sex-specific incidence differences.

Previously discovered glioma risk regions. There were 5,934 SNPs within 500 kb of 26 previously discovered glioma risk loci with IMPUTE2 information score (INFO) > 0.7 and MAF > 0.01 that were previously found to have at least a nominal (p < 5 × 10^-8) association with glioma, and results were considered significant at p < 2.8 × 10^-8 level (adjusted for 6,000 tests in each of three histologies [18,000 tests], see Fig. 2A for schematic of study design). Among the 25 previously validated glioma risk loci, nine loci contained 10 SNPs with pM < 2.8 × 10^-8 and/or pF < 2.8 × 10^-8 in any histology: 1p31.3 (RAVER2), 5p15.33 (TERT), 7p11.2 (EGFR, two independent loci), 8q24.21 (intergenic region near MYO), 9p21.3 (CDKN2B-AS1), 11q23.3 (PHLDB1), 16p13.3 (RHBDF1), 17p13.1 (TP53), and 20q13.33 (RTEL1) (Table 2). ORM and ORF were similar in the majority of these loci.

For one of two independent loci at 7p11.2 (rs11979158), there was a significant association only for males all glioma (ORM = 1.33 [95% CI = 1.23–1.44], pM = 4.87 × 10^-14) and GBM (ORM = 1.40 [95% CI = 1.28–1.54], 0.7 and MAF = 0.01 that were previously found to have at least a nominal (p < 5 × 10^-8) association with glioma, and results were considered significant at p < 2.8 × 10^-8 level (adjusted for 6,000 tests in each of three histologies [18,000 tests], see Fig. 2A for schematic of study design). Among the 25 previously validated glioma risk loci, nine loci contained 10 SNPs with pM < 2.8 × 10^-8 and/or pF < 2.8 × 10^-8 in any histology: 1p31.3 (RAVER2), 5p15.33 (TERT), 7p11.2 (EGFR, two independent loci), 8q24.21 (intergenic region near MYO), 9p21.3 (CDKN2B-AS1), 11q23.3 (PHLDB1), 16p13.3 (RHBDF1), 17p13.1 (TP53), and 20q13.33 (RTEL1) (Table 2). ORM and ORF were similar in the majority of these loci.

For one of two independent loci at 7p11.2 (rs11979158), there was a significant association only for males all glioma (ORM = 1.33 [95% CI = 1.23–1.44], pM = 4.87 × 10^-14) and GBM (ORM = 1.40 [95% CI = 1.28–1.54],
pM = 1.26 × 10^{-12} but the sex differences did not meet the significance threshold (overall pD = 0.0055, and GBM pF = 0.1184) (Fig. 3, Table 2).

The previously identified SNP at 8q24.21 (rs55705857) was the most significant SNP in both males and females. Odds ratio for rs55705857 in all glioma was significantly higher in females (ORF = 2.45 [95% CI = 2.14–2.80], pF = 1.22 × 10^{-39}) as compared to males (ORM = 1.56 [95% CI = 1.40–1.75], pM = 1.09 × 10^{-14}) with pD = 3.46 × 10^{-7}. In non-GBM only, ORF (ORF = 4.71 [95% CI = 3.94–5.63], pF = 1.85 × 10^{-69}) was also elevated as compared to ORM (ORM = 2.66 [95% CI = 2.28–3.10], pM = 8.13 × 10^{-36}) with pD = 8.44 × 10^{-7} (Fig. 3, Table 2). This association was further explored in a case-only analysis, where there was a significant difference between males and females overall (p = 0.0012), and in non-GBM (p = 0.0084) (Supplemental Table 1).

Previous studies have found a strong association between rs55705857 and oligodendrogial tumors (particularly tumors with isocitrate dehydrogenase 1/2 (IDH1/2) mutation and loss of the short arm of chromosome 1 [1p] and the long arm of chromosome 19 [19q]), so this association was further explored in the non-GBM (lower grade glioma [LGG]) histology groups (Table 3). For World Health Organization (WHO) grade II–grade III astrocytoma, effect was stronger in females (ORF = 4.64 [95% CI = 3.53–6.09], pF = 2.15 × 10^{-30}) as compared to males (ORM = 2.87 [95% CI = 2.31–3.56], pM = 1.19 × 10^{-22}) with pD = 0.0065. For WHO grade II–III oligodendrogliomas effect was stronger than observed in WHO grade II–III astrocytomas, and effect size was stronger in females (ORF = 12.15 [95% CI = 8.96–16.48], pF = 3.68 × 10^{-58}) as compared to males (ORM = 5.47 [95% CI = 4.16–7.19], pM = 5.37 × 10^{-34}) with pD = 6.60 × 10^{-4}. Oligoastrocytic tumors were not included in sub-analyses due to recent research that suggests that these tumors are not an entity that is molecularly distinct from oligodendrogliomas or astrocytomas.

**Genome-wide scan of nominally significant regions.** In a previous eight study meta-analysis, ~12,000 SNPs (INFO > 0.7, MAF > 0.01) were identified as having a nominally significant (p < 5 × 10^{-4}) association with all glioma, GBM, or non-GBM. A sex-stratified genome-wide scan was conducted within this set of SNPs and results were considered significant at pD < 1.4 × 10^{-6} (adjusted for 12,000 tests in each of three histologies [36,000 tests], see Fig. 2A for schematic of study design). Similar genome-wide peaks were observed between males and females (Fig. 4). One large region within 3p21.31 (49400kb–49600kb, ~200 kb) was identified as being significantly associated with glioma and GBM in females only (Fig. 5, Supplemental Fig. 1). There were 243 SNPs with nominally significant associations within this region in the previous eight-study meta-analysis (p < 5 × 10^{-4}), and 32 of these had nominally significant sex associations (pF < 5 × 10^{-4} or pM < 5 × 10^{-4}) in all glioma or GBM. The strongest association in females within this region was at rs9841110, in both all glioma (ORF = 1.22 [95% CI = 1.14–1.32], pF = 5.55 × 10^{-6}) and GBM only (ORM = 1.27 [95% CI = 1.16–1.38], pM = 3.86 × 10^{-5}) with pD = 6.04 × 10^{-4}, while there were no significant associations detected in males (Fig. 3). No SNPs in this region were significantly associated with non-GBM. In a case-only analysis a marginally significant difference was detected between males and females overall (p = 0.0520) and in GBM (p = 0.0428) (Supplemental Table 1).

**Agnostic scan of sex chromosome loci.** SNPs on the sex chromosomes were analyzed in GICC only. There were 245,746 SNPs with INFO > 0.7 and MAF > 0.01 on the X chromosome after quality control and imputation, and results were considered significant at p < 2 × 10^{-7} (corrected for 250,000 tests, see Fig. 2B for a schematic of study design). No SNPs met this significance threshold. After quality control procedures were complete, there were 300 SNPs remaining on the Y chromosome. There was no imputation performed of the Y chromosome data and only the 300 genotyped SNPs were evaluated. No significant signals were detected on the Y chromosome.

**Figure 1.** Average Annual Incidence of all glioma, glioblastoma and lower grade glioma by sex and age at diagnosis (CBTRUS 2010–2014).
Mean Age (SD)  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study</th>
<th>Cases</th>
<th>Controls</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Total</td>
<td>4,831</td>
<td>5,216</td>
<td>3,206</td>
<td>5,470</td>
</tr>
<tr>
<td></td>
<td>GIACC</td>
<td>2,733</td>
<td>1,868</td>
<td>1,831</td>
<td>1,397</td>
</tr>
<tr>
<td></td>
<td>SFAGS-GWAS</td>
<td>440</td>
<td>749</td>
<td>237</td>
<td>1,611</td>
</tr>
<tr>
<td></td>
<td>MDA-GWAS</td>
<td>714</td>
<td>1,094</td>
<td>429</td>
<td>1,142</td>
</tr>
<tr>
<td></td>
<td>GliomaScan</td>
<td>944</td>
<td>1,465</td>
<td>709</td>
<td>1,260</td>
</tr>
</tbody>
</table>

Mean Age (SD)  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study</th>
<th>Cases</th>
<th>Controls</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBM (% of total)</td>
<td>Total</td>
<td>52.5 (14.5)**</td>
<td>58.2 (15.2)**</td>
<td>51.8 (14.9)**</td>
<td>54.7 (14.5)**</td>
</tr>
<tr>
<td></td>
<td>GIACC</td>
<td>52.5 (14.3)</td>
<td>36.1 (13.4)</td>
<td>51.3 (14.6)</td>
<td>53.4 (14.3)</td>
</tr>
<tr>
<td></td>
<td>SFAGS-GWAS</td>
<td>53.8 (13.0)</td>
<td>50.6 (14.8)</td>
<td>53.5 (14.0)</td>
<td>49.3 (13.2)</td>
</tr>
<tr>
<td></td>
<td>MDA-GWAS</td>
<td>47.1 (13.0)</td>
<td>69.3 (12.7)</td>
<td>55.1 (15.7)</td>
<td>64.0 (15.4)</td>
</tr>
</tbody>
</table>

GBM - Mean Age (SD)  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study</th>
<th>Cases</th>
<th>Controls</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>GBM (% of total)</td>
<td>Total</td>
<td>57.3 (12.0) **</td>
<td>57.8 (12.1) **</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>GIACC</td>
<td>57.7 (11.4)</td>
<td>57.8 (11.6)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>SFAGS-GWAS</td>
<td>56.4 (11.5)</td>
<td>56.2 (12.3)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MDA-GWAS</td>
<td>52.0 (11.7)</td>
<td>53.7 (11.3)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>GliomaScan</td>
<td>60.4 (13.0)</td>
<td>61.4 (12.5)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Non-GBM - Mean Age (SD)  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study</th>
<th>Cases</th>
<th>Controls</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-GBM (% of total)</td>
<td>Total</td>
<td>1,716 (35.5)% **</td>
<td>1,320 (41.2)% **</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>GIACC</td>
<td>1,036 (37.9%)</td>
<td>862 (47.1%)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>SFAGS-GWAS</td>
<td>107 (24.3%)</td>
<td>59 (24.9%)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MDA-GWAS</td>
<td>317 (44.4%)</td>
<td>183 (42.7%)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>GliomaScan</td>
<td>256 (27.1%)</td>
<td>216 (30.5%)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Non-GBM - Mean Age (SD)  

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study</th>
<th>Cases</th>
<th>Controls</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-GBM (% of total)</td>
<td>Total</td>
<td>44.3 (14.4)% **</td>
<td>43.9 (14.3)% **</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>GIACC</td>
<td>44.7 (14.6)</td>
<td>44.6 (14.2)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>SFAGS-GWAS</td>
<td>45.7 (14.2)</td>
<td>45.4 (15.8)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>MDA-GWAS</td>
<td>41.0 (11.9)</td>
<td>39.6 (12.9)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>GliomaScan</td>
<td>46.3 (15.5)</td>
<td>44.4 (15.2)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Table 1. Population characteristics by study and sex. ‘Data from Glioma International Case-Control Study (GIACC; Melin, et al.);’ Data from San Francisco Adult Glioma Study GWAS (SFAGS-GWAS; Wrensch, et al.); ‘data from MD Anderson Cancer Center GWAS (MDA-GWAS; Shete, et al.);’ Data from the National Cancer Institute’s GliomaScan (GliomaScan; Rajaraman, et al.); ‘Data from CGEMS prostate study (Yegar et al.).’ Continuous age is not available, age distribution is as follows 50–59: 12.3%, 60–69: 56.7%, 70–79: 30.7%, 80–89: 3.3%; ‘Data from CGEMS breast study (Hunter et al.).’ Continuous age is not available, age distribution is as follows: 0–54: 4.3%, 55–59: 15.0%, 60–64: 23.6%, 65–69: 27.5%, 70–74: 19.0%, 75–99: 10.7%; ‘Information not available for all cases and frequencies may not add to 100%.’ *Significant differences between included studies at the p < 0.05 level.

Combined analysis of germline variants and somatic characterization. Due to the lack of molecular classification data included in the GIACC, MDA-GWAS, SFAGS-GWAS and GliomaScan datasets, glioma data obtained from TCGA datasets (GBM and LGG) were used to explore the potential confounding due to molecular subtype variation with histologies. There were 758 individuals from the TCGA dataset available for analysis with available germline genotyping, molecular characterization, sex, and age data (Supplemental Table 2). Overall, slightly more females (53.2%) as compared to males (47.2%) had IDH1/2 mutant glioma, but this difference was not statistically significant (p = 0.1104) (Supplemental Fig. 2). When tumors were stratified by histological type, approximately equal proportions of males and females had IDH1/2 mutations present in their tumors (GBM: 6.0% in males, and 5.2% in females; LGG: 17.9% in males, and 17.7% in females). There were also no significant differences by sex in IDH1/TER1/1p19q subtype (Supplemental Fig. 3, overall p = 0.2859), or pan-glioma methyl-
There was a nominally significant signal in the case-only meta-analysis for the proxy SNP at 7p11.2, but no significant association in the TCGA, but RAF was elevated in males as compared to females in the GBM set, as well as in all IDH1/2 wild type gliomas. There was no significant signal detected in the overall case-only meta-analysis for the proxy SNP at 8q24.21, or within the TCGA set. Among both LGG and IDH1/2 mutant, RAF was elevated in females as opposed to males.

**Sex-stratified genotypic risk scores.** In order to estimate the cumulative effects of significant variants by sex, unweighted risk scores (URS) were calculated by summing all risk alleles for each individual using the 10 SNPs (rs12752552, rs9841110, rs10069690, rs11979158, rs55705857, rs634537, rs12803321, rs3751667, rs78378222, and rs2297440) found to be significantly associated with glioma in this analysis. GBM (URS-GBM) and non-GBM (URS-NGBM) specific URS were calculated only using sets of 6 SNPs in this set that were significantly associated with these histologies (URS-GBM: rs9841110, rs10069690, rs11979158, rs634537, rs78378222, and rs2297440, and URS-NGBM: rs10069690, rs55705857, rs634537, rs12803321, rs78378222, and rs2297440). See Methods for additional information on score calculation. Median URS, URS-GBM, and URS-NGBM were significantly different \(p < 0.0001\) between cases and controls in both males and females in all histology groups (Supplemental Fig. 5). There was no significant difference in median risk scores between male and female cases for any histology group. Glioma risk increased with increasing number of alleles in both males and females for the 10 SNPs included in the overall URS, as well as the 6 SNPs in the URS-GBM and 6 SNPs in URS-NGBM (Fig. 6, Supplemental Table 4). Risk was higher in females (OR = 3.97 [95% CI = 2.42–6.80]) as compared to males (OR = 1.74 [95% CI = 1.21–2.53]) in all glioma for individuals for with 13–16 alleles, though the difference between these estimates were not statistically significant. Risk was also higher among females (OR = 2.69 [95% CI = 1.98–3.66]) as compared to males (OR = 1.79 [95% CI = 1.38–2.32]) in GBM for individuals with 8–11 risk alleles, as well as in non-GBM for individuals with 6–11 risk alleles (females: OR = 2.83 [95% CI = 2.12–3.78], males: OR = 1.70 [95% CI = 1.31–2.19]), though the difference between these estimates were not statistically significant. The estimates may underestimate actual risk due to varying effect sizes and alleles frequencies between risk variants.
Table 2. Previously identified glioma risk loci and histology-specific odds ratios (OR) and 95% confidence intervals (95% CI) stratified by sex.

Discussion

This is the first analysis of inherited risk variants in sporadic glioma focused specifically on sex differences, and the first agnostic unbiased scan for glioma risk variants on the X and Y sex chromosomes. One SNP at the 7p11.2 locus (rs11979158) showed significant association in males only, in both all glioma and GBM (Table 2; see Supplemental Table 5 and Supplemental Fig. 6 for study-specific estimates). This variant is within one of two previously identified independent glioma risk loci located near epidermal growth factor receptor (EGFR) and is most strongly associated with risk for GBM. Through EGF is implicated in many cancer types and is a target for many anti-cancer therapies, this risk locus has not been previously associated with any other cancer type. While a murine model of gliomagenesis, where EGF treatment was transforming for male but not female astrocytes that did not express EGFR, the sex difference in risk could be explained by estrogen or an estrogen signaling pathway.

The association at 8q24.21 (rs5705857; 8q24.21) showed significant association in both sexes, with a trend in odds ratio of 1.99 (95% CI: 1.85–2.13, p = 9.53 × 10^{-7}) in all glioma, and an odds ratio of 3.39 (95% CI: 3.09–3.71, p = 7.28 × 10^{-149}) in non-GBM (see Supplemental Table 5 and Supplemental Fig. 7 for study-specific estimates). The identified SNP, rs55705857, is located in an intergenic region near coiled-coil domain containing 26 (CCDC26, a long non-coding RNA). This analysis found a stronger association in females...
than males in all glioma and non-GBM, where female odds ratio estimates are ~2 × those of males (Table 2). ORs were higher in women than men in all studies included in the analysis (see Supplemental Tables 5 and 6 for study-specific estimates and MAF). A sensitivity analysis was conducted to assess the effect of study heterogeneity on this estimate in non-GBM using only the GICC, MDA-GWAS, and GliomaScan datasets. The exclusion of SEAGS-GWAS did not substantially change the results (Main analysis pD

ORs were higher in women than men in all studies included in the analysis (see Supplemental Tables 5 and 6 for study-specific estimates and MAF). A sensitivity analysis was conducted to assess the effect of study heterogeneity on this estimate in non-GBM using only the GICC, MDA-GWAS, and GliomaScan datasets. The exclusion of SEAGS-GWAS did not substantially change the results (Main analysis pD

that the ‘true’ effect that this exposure might have

As compared to a model containing age at diagnosis and sex alone, the three SNPs (rs55705857, rs9841110 and rs11979158) identified as having sex-specific effects explain an additional 1.4% of trait variance within the GICC. The variance explained by these SNPs varies by histology (0.6% in GBM, and 3.3% in Non-GBM). The variance explained by the addition of these three SNPs was not statistically significant in any group, but a similar trend in RAF was observed in the overall GBM and non-GBM datasets. The exclusion of SEAGS-GWAS did not substantially change the results (Main analysis pD

Figure 3. Sex-specific odds ratios overall and by histology grouping, 95% CI and p values for selected previous GWAS hits and 3p21.31 (rs9841110) for all glioma, GBM, and non-GBM.
of glioma risk variants by sex by summing all risk alleles using the 10 SNPs found to be significantly associated with glioma in this analysis. GBM (URS-GBM) and non-GBM (URS-NGBM) specific URS were calculated using sets of six SNPs in this set that were associated with significantly associated with each histology. Individuals with lower numbers of risk alleles had significantly lower odds of glioma, and those with higher numbers of alleles had increased odds of glioma, with statistically significant trends in each histology group. Males and females with low risk scores had similar odds of glioma, while females had increased odds in the upper strata of scores as compared to males. Development of risk scores that weight alleles by effect size, and use sex-specific estimates for variants for which effect size varies by sex (such as 7p11.2 and 8q24.21), may lead to better predictive values.

While often not included in GWAS, sex-stratified analyses can reveal genetic sources of sexual dimorphism in risk. Sex-stratified analyses not only contribute to understanding of sources of sex difference in incidence, but may also suggest mechanisms and pathways disease development that vary by sex. Sex variation in genetic susceptibility to disease is likely not due to sex differences in actual DNA sequence, but is thought to be the result of sex-specific regulatory functions. In addition to genetic sources of difference, there are likely several additional factors acting in combination which contribute to sex differences in glioma incidence. Sex differences in disease can also be linked to in-utero development, during which time gene expression and risk phenotypes are patterned through the action of X chromosome alleles that escape inactivation and genes on the non-pseudo-autosomal component of the Y chromosome, as well as the epigenetic effects of in utero...

Figure 4. Manhattan plot of -log(p) values for all glioma in (A) males and (B) females, for GBM in (C) males and (D) females, and for non-GBM in (E) males and (F) females.
A previous analysis estimating heritability of brain and CNS tumors by sex using twins attempted to estimate sex-specific relative risks, but these analyses were limited by a small sample size. Further investigation of the inheritance patterns of familial glioma by sex may also provide additional information about sex differences in this disease.

There are several limitations to this analysis. Individuals included in these datasets were recruited during different time periods from numerous institutions, with no central review of pathology. Molecular tumor markers were unavailable for all datasets, and as a result classifications are based on the treating pathologist using the prevailing histologic criteria at time of diagnosis. The variant at 8q24.21 has been shown to have significant association with particular molecular subtypes, and without molecular data it was not possible to determine whether the observed result is an artifact of varying molecular features by sex. Oligodendroglioma as a histology is highly enriched for IDH1/2 and 1p/19q co-deleted tumors (117/174, or ~67% within TCGA) and it is therefore likely that the analysis using only tumors classified as oligodendroglioma captured most of this molecular subtype. Males and females within histology groups have different frequencies of IDH1/2 mutation, which may have confounded the estimates for 8q24.21. The TCGA dataset was used to explore sex differences in allele frequency within molecular groups, but none of the identified SNPs were able to be directly validated within this set.

Table 4. Risk allele frequencies (RAF) Case-only odds ratios, 95% confidence intervals (95% CI), and p values for marker SNPs from four study meta-analysis and the Cancer Genome Atlas genotyping data.
however SNPs in strong LD were evaluated except for in 8q24.21. The 8q24.21 region is not well characterized on the array used for the TCGA genotyping, and as a result this region imputed poorly. No proxy SNP in strong LD with rs55705857 was able to be identified. Similar trends in RAF to those observed in the overall meta-analysis were seen in the TCGA set, though these differences were not statistically significant. Further interrogation in datasets with molecular classification where direct genotyping of these regions is warranted in order to confirm the sex-specific associations observed in this analysis.

Conclusions
Sex and other demographic differences in cancer susceptibility can provide important clues to etiology, and these differences can be leveraged for discovery in genetic association studies. This analysis identified potential sex-specific effects in 2 previous identified glioma risk loci (7p11.2, and 8q24.21), and 1 newly identified autosomal locus (3p21.31). Odds ratios for the highest strata of an unweighted risk score calculated by summing total risk alleles was higher in females as compared to males in all three histology groups. These significant differences in effect size may be a result of differing biological function of these variants by sex due to biological sex differences, or interaction between these variants and unidentified risk factors that vary in prevalence or effect by sex.

Materials and Methods
Study cohorts. This study was approved locally by the institutional review board (IRB) at University Hospitals Cleveland Medical Center and by each participating study sites IRB. Written informed consent was obtained from all participants. All research was performed in accordance with relevant guidelines and regulation. In this study, data was combined from four prior glioma GWAS: Glioma International Case-Control Study (GICC), San Francisco Adult Glioma Study GWAS (SFAGS-GWAS), MD Anderson Glioma GWAS (MDA-GWAS), and National Cancer Institute’s GliomaScan (Fig. 4A)⁴⁻¹⁴. The SFAGS-GWAS includes controls from the Illumina iControls dataset, and MDA-GWAS includes controls from Cancer Genetic Markers of Susceptibility (CGEMS) breast and prostate studies⁴⁻¹⁴. Details of data collection and classification are available in previous publications⁴⁻¹⁴.

Figure 6. Odds ratios and 95% confidence intervals for unweighted risk (URS) score in (A) all glioma, (B) GBM-specific URS (URS-G) in GBM, and (C) non-GBM-specific URS (URS-NGBM) for in non-GBM.
Genotyping and imputation of GWAS datasets. GICC cases and controls were genotyped on the Illumina Oncoarray\textsuperscript{17}. The array included 37,000 beadchips customized to include previously-identified glioma-specific candidate single nucleotide polymorphisms (SNPs). SFAGS-GWAS cases and some controls were genotyped on Illumina's HumanCNV370-Duo BeadChip, and the remaining controls were genotyped on the Illumina HumanHap300 and HumanHap550. MDA-GWAS cases were genotyped on the Illumina HumanHap610 and controls using the Illumina HumanHap550 (CGEMS breast\textsuperscript{34,39}) or HumanHap300 (CGEMS prostate\textsuperscript{35}). GliomaScan cases were genotyped on the Illumina 660 W, while controls were selected from cohort studies and were genotyped on Illumina 370D, 550 K, 610Q, or 660 W (see Rajaraman et al. for specific details of genotyping\textsuperscript{54}). Details of DNA collection and processing are available in previous publications\textsuperscript{12-14}. Individuals with a call rate (CR) < 99% were excluded, as well as all individuals who were of non-European ancestry (< 80% estimated European ancestry using the FastPop\textsuperscript{48} procedure developed by the GAMEON consortium). For all apparent first-degree relative pairs were removed (identified using estimated identity by descent [IBD] ≥ 0.5), for example, the control was removed from a case-control pair; otherwise, the individual with the lower call rate was excluded. SNPs with a call rate < 95% were excluded as those with a minor allele frequency (MAF) < 0.01, or displaying significant deviation from Hardy-Weinberg equilibrium (HWE) (p < 1 × 10\textsuperscript{−8}). Additional details of quality control procedures have been previously described in Melin et al.\textsuperscript{4}. All datasets were imputed separately using SHAPEIT v2.837 and IMPUTE v2.3.2 using a merged reference panel consisting of data from phase three of the 1,000 genomes project and the UK10K\textsuperscript{39-41}.

TCGA cases were genotyped on the Affymetrix Genomewide 6.0 array using DNA extracted from whole blood (see previous manuscript for details of DNA processing\textsuperscript{21,23}), and underwent standard GWAS QC, and duplicate and related individuals within datasets have been excluded\textsuperscript{1}. Ancestry outliers were identified in TCGA using principal components analysis in plink 1.9\textsuperscript{43}. Resulting files were imputed using Eagle 2 and Minimac3 as implemented on the Michigan imputation server (https://imputationserver.spb.umich.edu) using the HaploType Reference Consortium Version r1.1 2016 as a reference panel\textsuperscript{46-48}. Somatic characterization of TCGA cases was obtained from the final dataset used for the TCGA pan-glioma analysis\textsuperscript{22}, and classification schemes were adopted from Eckel-Passow, et al.\textsuperscript{49} and Ceccarelli, et al.\textsuperscript{21}.

Sex-stratified scan of the autosomal chromosomes. The data were analyzed using sex-stratified logistic regression models in SNPTTEST for all SNPs on autosomal chromosomes within 500 kb of previously identified risk loci, and/or those found to be nominally significant (p < 5 × 10\textsuperscript{−8}) in a previous meta-analysis (Fig. 2A)\textsuperscript{4,50}. Sex-specific betas (β\textsubscript{M} and β\textsubscript{F}), standard errors (SE\textsubscript{M} and SE\textsubscript{F}), and p-values (p\textsubscript{M} and p\textsubscript{F}) were generated using sex-stratified logistic regression models that were adjusted for number of principal components found to significant differed between cases and controls within each study in a previous meta-analysis\textsuperscript{4,50}. Genomic inflation factors were calculated After excluding SNPs with MAF < 0.05, INFO score < 0.7, and that significantly violated Hardy-Weinberg equilibrium in controls (p < 5 × 10\textsuperscript{−8}), genomic inflation factors (Males: GICC: λ\textsubscript{adjusted} = 1.04, SFAGS-GWAS: λ\textsubscript{adjusted} = 1.01 MDA-GWAS: λ\textsubscript{adjusted} = 1.02; Gliomascan: λ\textsubscript{adjusted} = 1.01. Females: GICC: λ\textsubscript{adjusted} = 1.03; SFAGS-GWAS: λ\textsubscript{adjusted} = 1.02; MDA-GWAS: λ\textsubscript{adjusted} = 1.04; Gliomascan: N\textsubscript{adjusted} = 1.01).

Estimation of sex difference and test of statistical significance. \( \beta_D = \beta_M - \beta_F \) and SE\textsubscript{D} were estimated using the sex-specific betas and standard errors separately for each dataset, as follows:

\[
\beta_D = \beta_M - \beta_F
\]

\[
SE_D = \sqrt{SE_M^2 + SE_F^2}
\]

The difference between the groups was then tested using a z test\textsuperscript{51,52}. Sex-stratified results and differences estimates from the four studies were separately combined via inverse-variance weighted fixed effects meta-analysis in META\textsuperscript{53}. See Fig. 2A for schematic of autosomal analysis methods. Case only-analesys were performed for SNPs found to be significant in agnostic analyses using sex as outcome for all glioma, GBM, and non-GBM by study and sex-stratified logistic regression models in SNPTEST for all SNPs on autosomal chromosomes within 500 kb of previously identified risk loci, and/or those found to be nominally significant (p < 5 × 10\textsuperscript{−8})

Sex chromosome analysis. X and Y chromosome data were available from GICC set only. Males and females were imputed separately for the X chromosome using the previously described merged reference panel. X chromosomes were analyzed using logistic regression model in SNPTTEST module 'newml' assuming complete inactivation of one allele in females, and males are treated as homozygous females (Fig. 2B). For prioritized SNPs in the combined model, sex-specific effect estimates were generated using stratified logistic regression models. Y chromosome data were analyzed using logistic regression in SNPTTEST (Fig. 2B)\textsuperscript{54}. Figures were generated using LocusZoom and R 3.3.2 using GenABEL, qqman, and ggplot\textsuperscript{55-59}.

Analysis of TCGA germline and somatic data. Only newly diagnosed cases from TCGA GBM and LGG with no neo-adjuvant treatment or prior cancer were used. Demographic characteristics, molecular classification and somatic alterations data was obtained from Ceccarelli, et al.\textsuperscript{21}. Chi-square tests were used to compare the frequency of somatic alterations between age groups. SNPs found to be nominally significant (p < 5 × 10\textsuperscript{−8}) in a previous 8 study meta-analysis\textsuperscript{4}, with imputation quality ≥ 0.7 were identified within the TCGA genotype data and D' and r\textsuperscript{2} values in CEU were used to select proxy SNPs\textsuperscript{16}. Using these SNPs, a case-only analysis using sex as a binary phenotype was conducted using logistic regression in SNPTEST assuming an additive model to estimate beta, standard error, and p values\textsuperscript{50}. Results were considered significant at p < 0.003 (Bonferroni correction for 15 tests, for the three assessed loci in each of five histology groups).
Calculation of unweighted genetic risk scores. In order to estimate the cumulative effects of significant variants by sex, histology-specific unweighted risk scores were calculated using the SNPs found to be significantly associated with each outcome. Data from all four studies was merged, and any imputed genotypes with genotype probability >0.8 were converted to hard calls. An overall unweighted risk score (URS) was generated using the sum of risk alleles at rs1275252, rs9841110, rs10069690, rs11979158, rs55705857, rs634537, rs12803321, rs3751667, rs78378222, and rs2297440. As risk alleles are known to have histology specific associations, histology specific scores were generated for GBM and non-GBM using only the SNPs found to have a significant association with each histology. GBM-specific URS (URS-G) was calculated by summing the number of risk alleles at rs9841110, rs10069690, rs11979158, rs634537, rs78378222, and rs2297440. Non-GBM-specific (URS-N) specific URS was calculated by summing the number of risk alleles at rs10069690, rs55705857, rs634537, rs12803321, rs78378222, and rs2297440. Unweighted risk scores (URS) were calculated by summing all risk alleles for each individual. Differences in median scores between groups using were tested using Wilcoxon rank sum tests. Scores were compared against the median score for each set (URS: ten alleles, URS-GBM: six alleles, URS-NGBM: four alleles). Odds ratios and 95% confidence intervals for each level of the score using sex-stratified logistic regression adjusted for age at diagnosis (for controls where only an age range was available, the mean value of the range was used), where each score was compared to the median score within the entire population as described in Shete et al.35.

Calculation of trait variance explained by SNPs with sex-specific effects. In order to determine whether the identified SNPs with sex-specific effects more accurate estimate odds of glioma than sex alone, logistic regression models were used to estimate odds of all glioma, GBM, and non-GBM glioma based on sex using the GICC data only. Proportion of variance in odds of glioma explained by sex-specific SNPs was calculated using R² estimated using the log likelihood of the null model (sex, age at diagnosis, and the first two principal components only) and the full model (including identified SNPs, rs9841110, rs11979158, rs55705857)29, calculated as follows:

\[
R^2 = 1 - \frac{\log(L_{null})}{\log(L_{full})}
\]  

(3)

Proportion of variance explained was also calculated separately by sex for each histology (null model adjusted for age at diagnosis, and the first two principal components only).

References


Acknowledgements
The results of this study were previously presented at the 2016 Annual Meeting of the Society for Neuro-Oncology, and the 2017 Annual Meeting of the American Association for Cancer Research. A pre-publication version of this manuscript was previously made available on bioRxiv (https://doi.org/10.1101/229112). QTO is supported by a Research Training Grant from the Cancer Prevention and Research Institute of Texas (CPRIT; RP1600977T).

The GICC was supported by grants from the National Institutes of Health, Bethesda, Maryland (R01CA139020, R01CA52689, P50CA097257, P30CA125123, P30CA008748, C. Thompson PI). Additional support was provided by the McNair Medical Institute and the Population Sciences Biorepository at Baylor College of Medicine. In Sweden work was additionally supported by Acta Oncologica through the Royal Swedish Academy of Science (BM salary) and The Swedish Research council and Swedish Cancer Foundation. We are grateful to the National clinical brain tumor group, all clinicians and research nurses throughout Sweden who identified all cases. The UCSF Adult Glioma Study was supported by the National Institutes of Health (grant numbers R01CA52689, P50CA097257, R01CA126831, and R01CA139020), the Loglio Collective, the National Brain Tumor Foundation, the Stanley D. Lewis and Virginia S. Lewis Endowed Chair in Brain Tumor Research, the Robert Magnin Newman Endowed Chair in Neuro- oncology, and by donations from families and friends of John Berardi, Helen Glaser, Elvera Olsen, Raymond E. Cooper, and William Martinussen. This project also was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI Grant Number UL1 RR024131. Its contents are solely supported by the National Institutes of Health, National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement # U86DP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the authors and do not necessarily represent the official views of the NIH. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention’s National Program of Cancer Registries, under agreement # U86DP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the author(s) and endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred. Other significant contributors for the UCSF Adult Glioma Study include: M Berger, P Bracci, S Chang, J Clarke, A Molinaro, A Perry, M Pezmecki, M Prados, I Smirnov, T Tihan, K Walsh, J Wiemels, S Zheng. UK10K data generation and access was organized by the UK10k consortium and funded by the Wellcome Trust. We are grateful to all the patients and individuals for their participation and we would also like to thank the clinicians and other hospital staff, cancer registries and study staff in respective centers who contributed to the blood sample and data collection.

Author Contributions
QTO and JSB-S conceptualized and designed the analyses. QTO wrote the main manuscript text and prepared all tables and figures. QTO, BK, YC, GA, and TR prepared the analytic files. MRW, JEE-P, JKW, LSM, HMH, CIA, JLB, EBC, DI, CJ, DHL, RKL, RTM, SHO, SS, JMS, SS, UA, PR, SJ, IC, MS, ZW, MY, LEBF, SK, DA, KV, VLS, RH, DSM, MF, AA, GGG, RM, RM-C, LLM, MS, AM, TC, GH, AZ-J, JMG, HDS, MPP, EW, UP, JB, RSH, RBJ, BM, MLB, and JSB-S provided data. JSB-S supervised all analyses. BK, MRW, JEE-P, EC, SHO, JBR, JDL, MEB, RSH, RBJ, BM, MLB, and JSB-S critically revised the manuscript for intellectual content. All authors reviewed the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-24580-z.

Competing Interests: This study was approved locally by the institutional review board (IRB) at University Hospitals Cleveland Medical Center and by each participating study site’s IRB. Written informed consent was obtained from all participants. There are no conflicts of interest to report. The GliomaScan (phs000652.v1.p1), CGEMS (phs000812.v1.p1), San Francisco Adult Glioma Study (phs001497.v1.p1), and GICC (phs001319.v1.p1) data used for these analyses are available through dbGap. Individual genotypes and phenotypes from the MD Anderson Glioma GWAS are being submitted to dbGAP, and final identifiers to these data will be provided in the final manuscript. The results here are in part based upon data generated by the TCGA Research Network: (http:// cancergenome.nih.gov/).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018
Consortia
GliomaScan consortium

32Department of Epidemiology, John Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America. 33American Cancer Society, Atlanta, Georgia, United States of America. 34Department of Oncology, Karolinska University Hospital, Stockholm, Sweden. 35Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, Massachusetts, United States of America. 36Department of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. 37Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Australia. 38Department of Public Health, John A. Burns School of Medicine, University of Hawaii at Manoa, Manoma, Hawaii, United States of America. 39Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, United States of America. 40Department of Epidemiology, Nutrition, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, United States of America. 41National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. 42Division of Surveillance, Hazard Evaluations, and Field Studies, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America. 43Department of Public Health and Clinical Medicine, Faculty of Medicine, Umeå University, Umeå, Sweden. 44Departments of Population Health and Environmental Medicine, New York University School of Medicine, New York, New York, United States of America. 45Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, United States of America. 46Department of Epidemiology, School of Public Health and Community Medicine, University of Washington, Seattle, Washington, United States of America.