
Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
10.3389/fgene.2018.00525

Link to publication record in Explore Bristol Research
PDF-document

This is the final published version of the article (version of record). It first appeared online via Frontiers at https://www.frontiersin.org/articles/10.3389/fgene.2018.00525/full. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/about/ebr-terms
Use of Mendelian Randomization for Identifying Risk Factors for Brain Tumors

Amy Elizabeth Howell*, Jie Zheng, Philip C. Haycock, Alexandra McAleenan, Caroline Relton, Richard M. Martin and Kathreena M. Kurian

1 Brain Tumour Research Centre, Institute of Clinical Neurosciences, University of Bristol, Bristol, United Kingdom, 2 MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom, 3 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

Gliomas are a group of primary brain tumors, the most common and aggressive subtype of which is glioblastoma. Glioblastoma has a median survival of just 15 months after diagnosis. Only previous exposure to ionizing radiation and particular inherited genetic syndromes are accepted risk factors for glioma; the vast majority of cases are thought to occur spontaneously. Previous observational studies have described associations between several risk factors and glioma, but studies are often conflicting and whether these associations reflect true causal relationships is unclear because observational studies may be susceptible to confounding, measurement error and reverse causation. Mendelian randomization (MR) is a form of instrumental variable analysis that can be used to provide supporting evidence for causal relationships between exposures (e.g., risk factors) and outcomes (e.g., disease onset). MR utilizes genetic variants, such as single nucleotide polymorphisms (SNPs), that are robustly associated with an exposure to determine whether there is a causal effect of the exposure on the outcome. MR is less susceptible to confounding, reverse causation and measurement errors as it is based on the random inheritance during conception of genetic variants that can be relatively accurately measured. In previous studies, MR has implicated a genetically predicted increase in telomere length with an increased risk of glioma, and found little evidence that obesity related factors, vitamin D or atopy are causal in glioma risk. In this review, we describe MR and its potential use to discover and validate novel risk factors, mechanistic factors, and therapeutic targets in glioma.

Keywords: Mendelian randomization, glioma, risk factors, genetic variant, causal inference, SNP, causal association

THE PUBLIC HEALTH BURDEN OF GLIOMA

Malignant gliomas are responsible for approximately 80% of all malignant brain tumors, with glioblastoma being the most prevalent histological subtype (Ostrom et al., 2014) (~45% of all gliomas Ostrom et al., 2014; Visser et al., 2015). Although glioma is a relatively rare cancer, with ~9,200 cases diagnosed each year in the United Kingdom (Cancer Research United Kingdom, 2015), the disease poses a serious health burden owing to its poor prognosis. The heterogeneous nature of the tumor cells makes the vast majority of gliomas surgically incurable (Kelly, 2010).
Additionally, difficulty is faced as therapeutic agents need to penetrate the blood brain barrier (Azad et al., 2015). As a result the median survival rate of grade III gliomas is two to 5 years (Wen and Kesari, 2008) and just 15 months for glioblastoma (WHO grade IV) (Stupp et al., 2009). The 5-year survival for glioma varies, from approximately 58% for ependymoma patients to approximately 5% for glioblastoma patients (Ostrom et al., 2014; Visser et al., 2015; Cancer Research United Kingdom, 2016).

**RISK FACTORS FOR GLIOMA**

**Accepted Risk Factors for Glioma**

The only environmental factor consistently associated with glioma risk is moderate to high exposure to ionizing radiation, accounting for only a small proportion of cases (Bondy et al., 2008; Braganza et al., 2012; Urbanska et al., 2014). Evidence was first provided from the Israeli Tinea Capitus cohort of children who had undergone radiation therapy for a benign medical condition (Sadetzki et al., 2005). This was supported by data from the Childhood Cancer Survivor Study that followed-up 14,361 children and adolescents (aged < 21 at initial diagnosis) who had survived for 5 years (Neglia et al., 2006). During follow-up, 40 gliomas were diagnosed, compared to an anticipated incidence of 4.62 (standardized incidence ratios (SIR) = 8.66, 95% confidence interval (CI) 6.24–11.6). These gliomas arose at a median of 9 years after original diagnosis. In a case-control analysis (with 4 controls per case, matched on age at diagnosis, sex and time since diagnosis, and the analysis adjusted for original cancer diagnosis) the odds ratio (OR) for glioma amongst children who underwent radiation therapy vs. those who did not was 6.78 (95% CI 1.54–29.7) (Neglia et al., 2006). The authors found that the risk of glioma per Gray of radiation was greatest among children who received radiation therapy at less than 5 years of age. After adjustment for radiation dose, neither original cancer diagnosis nor chemotherapy was associated with risk (Neglia et al., 2006). Taylor et al. (2010) carried out a study of 17,980 participants who had survived at least 5 years after diagnosis of childhood cancer.

In this study the risk of glioma increased linearly with dose of radiation (Taylor et al., 2010).

Rarely, glioma occurs in more than one family member, indicating a genetic susceptibility. This susceptibility is most often described within cases where inherited tumor syndromes are present, such as Li-Fraumeni syndrome, Turcot syndrome and neurofibromatosis type 1 (Louis et al., 2016). Kinnersley et al. (2018) reviewed glioma genome wide association study (GWAS) and summarized reported associations at the 27 glioma-risk SNPs (Kinnersley et al., 2018); genetic susceptibility loci are summarized in Table 1. These risk variants contribute to an increase in glioma risk; however, additional somatic mutations are a requisite for tumorigenesis in individuals with these germline variants or familial syndromes (Rice et al., 2016).

**Other Postulated Risk Factors**

There have been several risk factors that have been linked to the occurrence of glioma, though results from these investigations may be spurious because of the biases that pervade observational studies (Louis et al., 2016). A recently published systematic review presents risk factors for glioma onset that are shown to increase, decrease or have a null association with glioma risk (Quach et al., 2017).

Observational studies suggest that allergies (asthma, eczema, hay fever) are associated with lower glioma risk (Wigertz et al., 2007; Berg-Beckhoff et al., 2009; Amirian et al., 2016; Wang et al., 2016) and, consistent with this, asthma-susceptibility genotypes are associated with a reduced risk of glioma (Schwartzbaum et al., 2005). Short term use of anti-inflammatory medicine has also been reported to reduce glioma risk (Scheurer et al., 2011); although other studies have found conflicting results (Daugherty et al., 2011; Gaist et al., 2013). The possible role of allergies in decreasing the risk of glioma, including glioblastoma, may be due to an increase in immune surveillance, which in turn may destroy damaged, pre-cancerous cells earlier (Scheurer et al., 2011; Safaeian et al., 2013; Zhao et al., 2014). This hypothesis is supported by reports of a higher occurrence of glioma in HIV and AIDS patients (Blumenthal et al., 1999; Jukich et al., 2001; Hall and Short, 2009); as this is based on the result from a small number of studies with small sample sizes the estimate may be biased.

Brain tumors are observed to occur more often in Europeans compared with individuals of an African or Asian origin (McLendon et al., 1985; Kuratsu et al., 2001; Daresky and Dubrow, 2009; Ostrom et al., 2013), an observation that has also been reported within children. Robertson et al. (2002) investigated ethnic variation in the incidence of adult brain cancer in 994,725 individuals over 10.5 years of follow-up. The authors identified 373 people who developed brain cancer (232 glioblastomas, 106 astrocytomas and 35 oligodendrogliomas) of whom 50 were of African ancestry and 323 of European ancestry. Age adjusted incidence rates (per 100,000 race specific) were 0.11 and 0.46 (p = 0.003) in the African and European populations, respectively. The authors report a significant difference in incidence rates for the three most common gliomas and suggest that glioma is more common in individuals of European ancestry than in individuals of African ancestry (Robertson et al., 2002). Other studies have reported that glioma occurs 3.5 times more often in Europeans compared to African Americans (Davis et al., 1999). The explanation for this observed ethnic discrepancy remains unclear and while it is possible that a genetic difference exists between the two groups (Mochizuki et al., 1999; Chen et al., 2001; Das et al., 2002), detection bias cannot be ruled out (Dubrow and Daresky, 2011).

Certain occupations are reported to be linked with a higher risk of glioma, including physicians (Carozza et al., 2000; Krishnan et al., 2003; Pukkala et al., 2009), firefighters (Carozza et al., 2000; Krishnan et al., 2003) and farmers (Khuder et al., 1998; Zheng et al., 2001). Occupational exposure to metals such as arsenic and lead has attracted attention with respect to brain tumors as they are able to penetrate the blood brain barrier (Sunderman, 2001; Wang and Du, 2013; Liao et al., 2016). Exposure to lead has been associated with glioma risk.
TABLE 1 | Summary of the genetic susceptibility loci identified by GWAS in Europeans.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Alleles</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TERT</td>
<td>rs2736100</td>
<td>T/G</td>
<td>1.27 (1.19–1.37)</td>
</tr>
<tr>
<td>CCDC26</td>
<td>rs4295627</td>
<td>G/T</td>
<td>1.36 (1.29–1.43)</td>
</tr>
<tr>
<td>CCDC26</td>
<td>rs891835</td>
<td>G/T</td>
<td>1.24 (1.17–1.30)</td>
</tr>
<tr>
<td>CDKN2A/B</td>
<td>rs4977756</td>
<td>A/G</td>
<td>1.24 (1.19–1.30)</td>
</tr>
<tr>
<td>PHLDL1</td>
<td>rs498872</td>
<td>C/T</td>
<td>1.18 (1.13–1.24)</td>
</tr>
<tr>
<td>RTEL1</td>
<td>rs6010620</td>
<td>G/A</td>
<td>1.28 (1.21–1.35)</td>
</tr>
<tr>
<td>TP53</td>
<td>rs78378222</td>
<td>T/G</td>
<td>2.35 (1.61–3.44)</td>
</tr>
<tr>
<td>CCDC26</td>
<td>rs55705857</td>
<td>A/G</td>
<td>6.3 (4.6–8.8)</td>
</tr>
<tr>
<td>Near TERC</td>
<td>rs1920116</td>
<td>G/A</td>
<td>1.30 (1.19–1.42)</td>
</tr>
<tr>
<td>VTI1A</td>
<td>rs11196067</td>
<td>A/T</td>
<td>1.19 (1.12–1.27)</td>
</tr>
<tr>
<td>ZBTB16</td>
<td>rs648044</td>
<td>C/T</td>
<td>1.25 (1.17–1.34)</td>
</tr>
<tr>
<td>Intergenic</td>
<td>rs12230172</td>
<td>G/A</td>
<td>1.23 (1.16–1.32)</td>
</tr>
<tr>
<td>POLR3B</td>
<td>rs3851634</td>
<td>T/C</td>
<td>1.23 (1.15–1.32)</td>
</tr>
<tr>
<td>ETFA</td>
<td>rs180159</td>
<td>G/A</td>
<td>1.36 (1.23–1.51)</td>
</tr>
<tr>
<td>JAK1</td>
<td>rs12752552</td>
<td>T/C</td>
<td>1.22 (1.15–1.31)</td>
</tr>
<tr>
<td>MDM4</td>
<td>rs4252707</td>
<td>G/A</td>
<td>1.19 (1.12–1.26)</td>
</tr>
<tr>
<td>AKT3</td>
<td>rs12076373</td>
<td>G/C</td>
<td>1.23 (1.16–1.32)</td>
</tr>
<tr>
<td>Near 10H1</td>
<td>rs7572263</td>
<td>A/G</td>
<td>1.20 (1.13–1.26)</td>
</tr>
<tr>
<td>LRG1</td>
<td>rs11709632</td>
<td>A/C</td>
<td>1.15 (1.09–1.20)</td>
</tr>
<tr>
<td>OBFC1</td>
<td>rs11598018</td>
<td>C/A</td>
<td>1.14 (1.09–1.20)</td>
</tr>
<tr>
<td>Intergenic</td>
<td>rs11233250</td>
<td>G/T</td>
<td>1.24 (1.16–1.33)</td>
</tr>
<tr>
<td>MAML2</td>
<td>rs7107785</td>
<td>T/C</td>
<td>1.16 (1.11–1.21)</td>
</tr>
<tr>
<td>AKAP6</td>
<td>rs10131032</td>
<td>G/A</td>
<td>1.33 (1.22–1.44)</td>
</tr>
<tr>
<td>Near MPG</td>
<td>rs2562152</td>
<td>A/T</td>
<td>1.21 (1.13–1.29)</td>
</tr>
<tr>
<td>LMFI</td>
<td>rs3751667</td>
<td>C/T</td>
<td>1.18 (1.12–1.25)</td>
</tr>
<tr>
<td>HEATR3</td>
<td>rs10852606</td>
<td>T/C</td>
<td>1.18 (1.13–1.24)</td>
</tr>
<tr>
<td>SLC16A8</td>
<td>rs2235573</td>
<td>G/A</td>
<td>1.15 (1.10–1.20)</td>
</tr>
<tr>
<td>Near TERC</td>
<td>rs3772190</td>
<td>G/A</td>
<td>1.11 (1.06–1.15)</td>
</tr>
<tr>
<td>TERT</td>
<td>rs10069690</td>
<td>C/T</td>
<td>1.61 (1.53–1.69)</td>
</tr>
<tr>
<td>EGFR</td>
<td>rs75061358</td>
<td>T/G</td>
<td>1.63 (1.50–1.76)</td>
</tr>
<tr>
<td>EGFR</td>
<td>rs723527</td>
<td>A/G</td>
<td>1.25 (1.20–1.31)</td>
</tr>
<tr>
<td>CCDC26</td>
<td>rs55705857</td>
<td>G/A</td>
<td>3.39 (3.09–3.71)</td>
</tr>
<tr>
<td>CDKN2A/B</td>
<td>rs634537</td>
<td>T/G</td>
<td>1.37 (1.31–1.43)</td>
</tr>
<tr>
<td>VTI1A</td>
<td>rs11599075</td>
<td>G/A</td>
<td>1.16 (1.10–1.22)</td>
</tr>
<tr>
<td>ZBTB16</td>
<td>rs648044</td>
<td>A/G</td>
<td>1.19 (1.13–1.25)</td>
</tr>
<tr>
<td>PHLDL1</td>
<td>rs12803321</td>
<td>G/C</td>
<td>1.42 (1.35–1.49)</td>
</tr>
<tr>
<td>Intergenic</td>
<td>rs1275600</td>
<td>T/A</td>
<td>1.16 (1.10–1.21)</td>
</tr>
<tr>
<td>RFX4</td>
<td>rs12227783</td>
<td>A/T</td>
<td>1.16 (1.08–1.24)</td>
</tr>
<tr>
<td>ETFA</td>
<td>rs77633900</td>
<td>G/C</td>
<td>1.35 (1.25–1.46)</td>
</tr>
<tr>
<td>TP53</td>
<td>rs78378222</td>
<td>T/G</td>
<td>2.53 (2.19–2.91)</td>
</tr>
<tr>
<td>RTEL1</td>
<td>rs2297440</td>
<td>T/C</td>
<td>1.48 (1.40–1.56)</td>
</tr>
</tbody>
</table>

Table 1 is a modified version of Table 1 in Kinnersley et al. (2018). The table describes the gene, the single nucleotide polymorphism (SNP), the allele and the odds ratio (OR) and corresponding 95% confidence interval (95% CI). ORs are reported with respect to the risk allele, highlighted in bold.

(Anthill et al., 1996; van Wijngaarden and Dosemeci, 2006) and brain cancer mortality (Cocco et al., 1998; van Wijngaarden and Dosemeci, 2006). In a cohort study of 1,779,646 men and 1,066,346 women aged 25–64 years at baseline and subsequently followed for 19 years, an increased glioma risk was observed amongst men exposed to arsenic, mercury, and petroleum products (Navas-Acien et al., 2002). However, no relationship of lead, cadmium, nickel, chromium and iron with glioma risk was reported in a study of 1856 cases and 5189 controls (Parent et al., 2017). Other studies investigating the relationship between glioma and occupational exposure to metal (Samkange-Zeeb et al., 2010) or lead (Rajaraman et al., 2006; Bhatti et al., 2009), and between brain cancer more generally and lead (Lam et al., 2007) reported no strong evidence of a causal association.

There has been speculation that certain lifestyle choices, including alcohol intake, the use of drugs, or dietary exposure to nitrous compounds affect the risk of glioma; however, to date the evidence is inconclusive (Giles et al., 1994; Michaud et al., 2009; Kyritsis et al., 2011; Little et al., 2013; Shao et al., 2016; Tamimi and Juweid, 2017).

Mobile phone use has been speculated to be associated with brain tumor risk (Schüz et al., 2006). However, conflicting finding have also been reported (Frei et al., 2011). In a nationwide study involving Danish citizens aged 30 years or older (born after 1925), there was no evidence that mobile phone use increased brain tumor risk (Frei et al., 2011).

Other risk factors that are not discussed here have been investigated in relation to glioma risk, including but not limited to: Type 1 and type 2 diabetes, body mass index, birth weight, hypertension, height, birth weight, menarche (age at onset), menopause (age at onset), coffee/caffeine consumption, low-density lipoprotein cholesterol, insulin-like growth factor 1, insulin-like growth factor binding protein, triglycerides, high-density lipoprotein cholesterol, pesticide exposure, extremely low frequency magnetic fields, vitamin E, A and C levels (Preston-Martin and Mack, 1991; Kaplan et al., 1997; Houben et al., 2004; Linos et al., 2007; Holick et al., 2010; Kabat et al., 2011; Little et al., 2013; Malerba et al., 2013; Lee et al., 2014; Andersen et al., 2015; Lee et al., 2013; Juweid, 2017).
A confounder is a factor that is a common cause of both the disease under consideration and the exposure of interest. Importantly, a confounder is not on the causal pathway between the exposure and outcome (Hammer et al., 2009). For instance, in 2002 an association had been established between alcohol intake and the incidence of 3.6% of all cancers (Boffetta et al., 2006; Testino, 2011) but it is still uncertain whether an association exists between any class of glioma and alcohol intake (Braganza et al., 2014; Qi et al., 2014). An observed association between glioma incidence and alcohol intake could be because individuals who consume more alcohol are more likely to smoke (Hart et al., 2010) and to adhere to an unhealthy life-style; (Sayon-Orea et al., 2011; Bendsen et al., 2013) thus, it could be these other factors that influence the risk of glioma rather than alcohol consumption per se (Sergentanis et al., 2015).

Reverse causation occurs when the disease outcome precedes, and leads to, the exposure rather than being a consequence of the exposure (Flegal et al., 2011). For example a higher level of blood glucose has been reported to be protective against glioma (Kitahara et al., 2014); however, an alternative explanation is that tumors take-up glucose, leading to low glucose levels (Schwartzbaum et al., 2017).

Mendelian Randomization Analogous to Randomized Control Trials (Figure 1)

Randomized controlled trials are considered the gold standard study design for inferring causality, as successful randomization, adequately blinded implementation of the intervention, high rates of follow-up and intention-to-treat analysis should yield results that are relatively free from the biases afflicting observational studies (Iturrieta-Zuazo and Walter, 2015). On the other hand, RCTs often reflect short-term exposures at one time point in life, with limited follow-up, and participants are usually not representative of general populations, a particularly important issue if the priority is to identify primary prevention targets (Meyer, 2010). Additionally, due to ethical, practical and financial reasons, it is not feasible to randomize people to every risk factor (Bochud and Rousson, 2010): e.g., exposure to power lines, mobile phone use or breastfeeding.

One method to appraise causality within observational epidemiology is the use of Mendelian randomization (MR). MR is a type of "instrumental variable" analysis that utilizes genetic variants, such as SNPs, that are robustly associated with an exposure as proxies for the risk factor of interest. The aim of MR is to strengthen causal inference in observational studies of exposure as proxies for the risk factor of interest. The aim of MR is to strengthen causal inference in observational studies of exposure as proxies for the risk factor of interest (Greenland, 2000; Lawlor et al., 2008). Within the framework of MR, research has, therefore, focused primarily on causal effects of environmental exposures on disease risk (Walsh et al., 2015; Haycock et al., 2017; Disney-Hogg et al., 2018a; Takahashi et al., 2018), as opposed to survival. There are some instances where factors are involved in both disease incidence and progression, such as low-density lipoprotein cholesterol levels for heart disease risk and recurrence (Ference et al., 2017), although such instances may be exceptional. Cases do exist where a risk factor for a disease is not implicated in progression, as has been proposed for the relationship between folate consumption and colon cancer (Kim, 2003). Thus, current case-control GWAS of glioma risk have the potential to inform on the underlying causal mechanisms of disease onset but (at the current time) may be less informative for discovering drug targets to improve glioma survival (Paternoster et al., 2017). The latter requires case-only GWAS that examine genetic variation in relation to disease progression, but such studies are currently rare (Melin et al., 2017). The most probable explanation for this is due to a research focus to determine mechanisms that cause disease incidence and because of the challenges inherent in collecting progression data (see section “Future of MR in research” below). At present, a few MR studies have been conducted that investigate progression of disease (Brunner et al., 2017) but none in glioma progression, which is required for the discovery of targets for improving glioma survival (Paternoster et al., 2017).

Mendelian randomization can be used to identify and investigate potential drug targets (Mokry et al., 2015;
Zheng et al., 2017). A quarter of the drugs that enter clinical development fail due to their ineffectiveness (Ashburn and Thor, 2004; Arrowsmith and Miller, 2013). Current drug targets are authenticated using in-vitro and animal models, but these can fail to predict the potential benefits (or harms) in humans (Mokry et al., 2015; Zheng et al., 2017). Nelson et al. (2015) aimed to establish whether current genetic evidence could predict drug mechanisms. The authors reported that opting for targets that are genetically supported may result in twice the success rate in clinical development (Nelson et al., 2015). MR could substantially augment these methods (Mokry et al., 2015; Zheng et al., 2017).

The theory is that specific genetic variants can be utilized to imitate the effects of targeting a protein pharmacologically. If the variant codes for a potential drug target that causes an alteration in activity of the encoded protein, the causal effect of the drug on disease can be assessed by MR (Sofat et al., 2010; Evans and Smith, 2015). Additionally, MR can be used to examine all pairwise associations between serum protein levels and disease risk (Sun et al., 2018). If a variant is identified that is robustly associated with levels of a serum protein that display a putative causal relationship with disease risk, methods can be employed to search for available drugs that cause an alteration in the
For some exposures there is a lack of genetic variants (SNPs) available for instrumentation (Smith and Ebrahim, 2004). For example, ionizing radiation emitted by mobile phones has been suggested as a risk factor for glioma (Yang et al., 2017). However, currently no genetic variants have been associated with exposure (or response) to ionizing radiation and therefore MR analysis cannot be performed for this particular risk factor (Smith, 2010).

A key limitation of MR is pleiotropy (Sheehan et al., 2008). Pleiotropy occurs when a genetic variant has more than one effect. If one or more of these effects influence the outcome through pathways other than the exposure of interest (so called horizontal pleiotropy) a core MR assumption is violated, i.e., that variants only exert their effect on the outcome via their influence on the exposure of interest (Evans et al., 2013; Burgess, 2014; Bennett and Holmes, 2017; Yarmolinsky et al., 2017). Techniques have been developed, such as MR-Egger regression, that can quantify the amount of bias caused by horizontal pleiotropy, as well as providing a valid causal estimate despite the presence of horizontal pleiotropy (Bowden et al., 2015). Another type of pleiotropy that exists is vertical pleiotropy. This is where the genetic variants have associations with biomarkers that are downstream of the biomarker of interest (Bennett and Holmes, 2017). Thus, they are on the causal pathway and should be considered as intermediates of the relationship between an exposure and an outcome, not as confounding factors.

Mendelian randomization studies typically require large sample sizes, an issue that can be compounded by the rare nature of glioma. One way to increase power is to develop genetic risk scores that contain multiple alleles to explain more of the variance in the exposure of interest. This runs the risk of including invalid variants, such as those that do not exert their effect on the outcome via the exposure of interest (Evans et al., 2013; Burgess, 2014; Yarmolinsky et al., 2017), although such potential violations of the MR assumption can be formally tested using MR-Egger regression. Power can also be increased by using a two-sample approach, where large case-control GWAS can be used even if they have not measured the exposure of interest.

Limitations of MR have been discussed in detail in several published papers (Davey Smith and Ebrahim, 2003; Smith and Ebrahim, 2004; Lawlor et al., 2008; Sheehan et al., 2008; Bochud and Rousson, 2010; Davey Smith, 2011b; VanderWeele et al., 2014).

### MR IN GLIOMA RESEARCH

#### Studies That Have Evaluated Risk Factors for Glioma Using MR

Two-sample MR is a method that can harness information from GWAS summary statistics and has been applied to the context of glioma to look at several risk factors. We discuss key studies that have used two-sample MR to investigate associations between previously reported risk factors and glioma (Table 3).

An MR study to evaluate the causal relevance of telomere length on the risk of cancer and non-neoplastic diseases found that genetically predicted longer telomeres increased the risk of glioma, while being protective for certain non-neoplastic diseases (Glynn, 2010). For some exposures there is a lack of genetic methods described are the inverse variance weighted (IVW), maximum likelihood estimation (MLE), weighted median estimate (WME), mode-based estimate (MBE) and MR-Egger.

<table>
<thead>
<tr>
<th>Statistical Method</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inverse-variance weighted (IVW)</td>
<td>Assumes causal estimate due to each SNP is the same (fixed effects IVW) or that their deviations are balanced (random effects IVW) (Hemani et al., 2018). Gives an unbiased estimate when there is no horizontal pleiotropy (fixed effects IVW) or when horizontal pleiotropy is balanced (random effects IVW).</td>
</tr>
<tr>
<td>Maximum likelihood estimation (MLE)</td>
<td>Assumes effect of the exposure on the outcome due to each SNP is equal (fixed effects IVW makes the same assumption). A benefit of this method is that it might give more reliable results when measurement error in the SNP-exposure effect is present (Hemani et al., 2018). Gives an unbiased estimate when there is no horizontal pleiotropy or when horizontal pleiotropy is balanced (but variance of the estimate will be underestimated in the latter scenario).</td>
</tr>
<tr>
<td>Weighted median estimate (WME)</td>
<td>Takes the median effect of all SNPs. Returns an unbiased estimate if half the SNPs are valid instruments (Hemani et al., 2018). Requires a large number of instrumental SNPs otherwise method is underpowered.</td>
</tr>
<tr>
<td>Mode-based estimate (MBE)</td>
<td>SNPs are clustered into groups determined by similarity of causal effects. Returns the causal effect estimate based on the cluster that has the greatest number of SNPs (Hemani et al., 2018). Gives an unbiased estimate if the SNPs in the largest cluster are valid, even if most SNPs are invalid instruments. Requires a large number of instrumental SNPs otherwise method is underpowered.</td>
</tr>
<tr>
<td>MR-Egger</td>
<td>Modifies the IVW analysis by permitting a non-zero intercept, permitting the net-horizontal pleiotropic effect for all SNPs to be unbalanced, or directional (Hemani et al., 2018). Gives an unbiased estimate even if all SNPs do not adhere to instrumental variable assumptions but requires the InSIDE (instrument strength independent of direct effects) assumption to be valid. Requires a large number of instrumental SNPs otherwise method is underpowered.</td>
</tr>
<tr>
<td>Wald ratio</td>
<td>This is the easiest method to estimate a causal effect. Wald ratio method is appropriate when only a single SNP is available to proxy the risk factor of interest. However, a limitation is that it is much harder to appraise MR assumptions with only a single SNP. (Wald, 1940).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Description of statistical methods used in Mendelian randomization analysis.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical Method</strong></td>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>Inverse-variance weighted (IVW)</td>
<td>Assumes causal estimate due to each SNP is the same (fixed effects IVW) or that their deviations are balanced (random effects IVW) (Hemani et al., 2018). Gives an unbiased estimate when there is no horizontal pleiotropy (fixed effects IVW) or when horizontal pleiotropy is balanced (random effects IVW).</td>
</tr>
<tr>
<td>Maximum likelihood estimation (MLE)</td>
<td>Assumes effect of the exposure on the outcome due to each SNP is equal (fixed effects IVW makes the same assumption). A benefit of this method is that it might give more reliable results when measurement error in the SNP-exposure effect is present (Hemani et al., 2018). Gives an unbiased estimate when there is no horizontal pleiotropy or when horizontal pleiotropy is balanced (but variance of the estimate will be underestimated in the latter scenario).</td>
</tr>
<tr>
<td>Weighted median estimate (WME)</td>
<td>Takes the median effect of all SNPs. Returns an unbiased estimate if half the SNPs are valid instruments (Hemani et al., 2018). Requires a large number of instrumental SNPs otherwise method is underpowered.</td>
</tr>
<tr>
<td>Mode-based estimate (MBE)</td>
<td>SNPs are clustered into groups determined by similarity of causal effects. Returns the causal effect estimate based on the cluster that has the greatest number of SNPs (Hemani et al., 2018). Gives an unbiased estimate if the SNPs in the largest cluster are valid, even if most SNPs are invalid instruments. Requires a large number of instrumental SNPs otherwise method is underpowered.</td>
</tr>
<tr>
<td>MR-Egger</td>
<td>Modifies the IVW analysis by permitting a non-zero intercept, permitting the net-horizontal pleiotropic effect for all SNPs to be unbalanced, or directional (Hemani et al., 2018). Gives an unbiased estimate even if all SNPs do not adhere to instrumental variable assumptions but requires the InSIDE (instrument strength independent of direct effects) assumption to be valid. Requires a large number of instrumental SNPs otherwise method is underpowered.</td>
</tr>
<tr>
<td>Wald ratio</td>
<td>This is the easiest method to estimate a causal effect. Wald ratio method is appropriate when only a single SNP is available to proxy the risk factor of interest. However, a limitation is that it is much harder to appraise MR assumptions with only a single SNP. (Wald, 1940).</td>
</tr>
</tbody>
</table>

The statistical methods described are the inverse variance weighted (IVW), maximum likelihood estimation (MLE), weighted median estimate (WME), mode-based estimate (MBE) and MR-Egger.

levels of that protein (Corsello et al., 2017). As discussed, only case-control GWAS exist at present for glioma which may be less informative for the discovery of drug targets to improve survival (Paternoster et al., 2017).

Table 2 provides a summary of some of the different methods used to obtain MR estimates (Hemani et al., 2018).

**Limitations of MR Pertinent to Glioma**

Mendelian randomization has widely recognized limitations (Glynn, 2010). For some exposures there is a lack of genetic
diseases, such as cardiovascular diseases (Zheng et al., 2017). The analysis employed summary genetic data for 35 cancers and 45 non-neoplastic diseases, including 1,130 glioma cases and 6,294 controls. The strongest association was for glioma (OR per SD increase in genetically predicted telomere length was 5.27; 95% CI: 3.15–8.81) (Zheng et al., 2017). A possible explanation for this observation is that telomere shortening may act as a tumor suppressor, restricting the proliferative potential of cells. Therefore, those with longer telomeres have a greater probability of obtaining somatic mutations due to an increased proliferative potential (Hanahan and Weinberg, 2011).

Walsh et al. (2015) also used an MR approach to establish whether a genotypically estimated longer or shorter telomere length was linked with an increased risk of glioma and whether inheritance of SNPs associated with telomere length are indicators of glioma risk. The authors accessed differences in genotypically estimated relative telomere length in a total of 1,130 glioma patients and 6,294 controls. The average approximated telomere length was 31bp (5.7%) longer in glioma cases compared with controls in discovery analyses ($P = 7.82 \times 10^{-8}$). This finding was supported in the replication analysis as the mean telomere length was 27bp (5.0%) longer in glioma cases than controls ($1.48 \times 10^{-3}$). The authors reported that the risk of glioma increases monotonically with each increasing septime of telomere length (O.R 1.12; 95% CI: 0.90–1.62). Additionally, the authors reported that four telomere length-associated SNPs were significantly related with glioma risk in pooled analyses, including those in the telomerase component genes TERC (O.R 1.14; 95% C.I. = 1.03–1.28) and TERT (O.R 1.39; 95% C.I. = 1.27–1.52), and those in the CST complex genes OBFC1 (O.R 1.18; 95% C.I. = 1.05–1.33) and CTC1 (O.R 1.14; 95% C.I. = 1.02–1.28). The indication of risk alleles for glioma close to TERC and TERT that are also related with telomere length suggests that telomerase is important in glioma formation (Walsh et al., 2014).

Takahashi et al. (2018) used two-sample MR to investigate whether a causal relationship exists between circulating vitamin D and glioma risk, involving 12,488 glioma cases and 18,169 controls. The authors reported no strong evidence of a causal relationship between vitamin D and glioma when either the inverse-variance weighted (IVW) method (OR per SD increase 1.21, 95% CI: 0.90–1.62, $P = 0.201$) or the maximum likelihood estimation (MLE) method (OR per SD increase 1.20, 95% CI: 0.98–1.48, $P = 0.083$) was used (Takahashi et al., 2018).

Disney-Hogg et al. (2018a) used an MR approach to evaluate the observed inverse relationship between allergies and glioma risk. The instrumental variables were SNPs robustly associated with atopic dermatitis, asthma and hay fever, IgE levels, and self-reported allergy. The study involved 12,488 cases and 18,169 controls. The authors found no significant association between glioma and asthma, hay fever, IgE levels, or self-reported allergy. For atopic dermatitis an inverse association was found (OR per 2.7-fold increase in odds of the risk factor (exposure). No strong evidence of any association between glioma and asthma and hay fever (OR 0.96; 95% CI: 0.90–1.03. $P = 0.248$), IgE levels (OR 0.88; 0.69–1.13. $P = 0.319$), or self-reported allergy (OR 1.03; 95% CI: 0.95–1.11. $P = 0.834$).

Disney-Hogg et al. (2018b) carried out an MR analysis to interrogate the observed association between obesity-related factors and risk of glioma. The authors identified variants that were robustly associated with 10 key obesity-related factors: 2-h post-challenge glucose, BMI, fasting glucose, fasting insulin, HDL cholesterol, LDL cholesterol, type-2 diabetes, total cholesterol, triglycerides and waist-hip ratio. This study encompassed 12,488 cases and 18,169 controls. This study found little evidence that indicated that obesity-related factors contribute to glioma (Disney-Hogg et al., 2018b).

**Potential Application of Different MR Study Designs in Glioma Research**

There are several different design strategies for MR that have been discussed in detail by Zheng et al. (2017). The potential application of these different MR study designs in glioma research are outlined below.
Improved knowledge of signaling pathways that are causally associated with glioma incidence can be helpful to design preventative strategies and effective therapeutic targets (Wang et al., 2015). A useful MR strategy to establish whether a molecular intermediate plays a role in the causal pathway between a risk factor and disease is the use of two-step MR (Relton and Davey Smith, 2012). An improved understanding of the molecular changes that drive glioma formation will allow for opportunities to modify disease causing factors.

Bidirectional MR involves using instruments for both the exposure and the outcome to assess the direction of causality: i.e., does the exposure cause the outcome or does the outcome cause the exposure (Timpson et al., 2010). For instance, observational studies have suggested that there is an inverse association between allergies and glioma risk, but the direction and causality of the association remains uncertain: it is not clear whether allergies decrease the risk of glioma or whether the reverse association arises because of suppression of the immune system by glioma itself (Schoemaker et al., 2006).

There are cases in which genetic variants are related to numerous correlated phenotypes (Low, 2001), for example, genetic variants that associate with lipoprotein metabolism tend not to correlate with just one specific lipid fraction (Wurtz et al., 2013). As a result assessing the causal association of one specific intermediate phenotype with disease can be challenging (Davey Smith and Hemani, 2014). Multi-phenotype MR can be used in these cases (Davey Smith and Hemani, 2014; Burgess et al., 2015; Burgess and Thompson, 2015; Kemp et al., 2016). Multivariable MR can be applied to glioma research when testing the effect of lipids on glioma to identify the independent effect of each lipid subtype on glioma.

Hypothesis-driven MR has huge potential in glioma research. Hypothesis-driven MR can validate the relationship between a risk factor and glioma for which a causal association has previously been reported.

In addition, hypothesis-free MR has the potential to identify novel causal associations. Hypothesis-free MR can be used to examine causality in complex frameworks in glioma, as a well as a method to data mine high-dimensional studies (Evans and Davey Smith, 2015). Haycock et al. (2017) implemented a mixture of hypothesis-driven and hypothesis-free MR to investigate the relationship between telomere length and 22 cancers and 32 primary non-neoplastic diseases.

Mendelian randomization-Base is a tool that improves the accessibility of GWAS summary data for MR research (Hemani et al., 2016). MR-Base can assist hypothesis-free testing as it allows researchers to examine all pairwise associations to data mine for causal relationships of interest (Davey Smith, 2011a). Where novel associations are identified, these associations can then be subjected to formal and extensive hypothesis-testing studies (Evans and Smith, 2015).

Factorial MR can be used to develop therapeutic strategies to improve glioma survival. Factorial RCT is where a participant is either assigned to a group that obtains neither intervention, one of the interventions, or both (Montgomery et al., 2003). In a factorial trial the separate effects of each intervention can be considered, as well as, the benefits of obtaining both interventions together (Montgomery et al., 2003). Similarly, factorial MR can be performed by using combinations of genetic variants to attain unconfounded estimates of the effect of co-occurrence of the two drug targets on disease (Davey Smith and Hemani, 2014). In glioma research if we have two drug targets and we want to know the combined effects of these two drugs on glioma, then we can apply factorial MR. Factorial MR can assess the antitumor efficacy of drug targets on glioma by investigating the combination of different targeted drugs (Reardon and Wen, 2006).

Future of MR in Glioma Research

For GWAS and MR of glioma progression to be successful for the development of drug targets to improve glioma survival, large scale case-only studies will be required with both progression and germline genetic data. RCTs offer a potential reservoir of data for such studies; however, due to the rare nature of glioma, sample size is limited (Vuorinen et al., 2003; Gehring et al., 2009; U.S. National Library of Medicine, 2018). A limitation of progression studies is the introduction of collider bias, discussed in detail in Paternoster et al. (2017). Collider bias is problematic in MR of disease progression as a risk factor of interest that causes the disease may be correlated with other risk factors involved in incidence, and any association between the index risk factor and progression can be confounded by these correlated risk factors. If the problems of sample size and collider bias can be adequately overcome, GWAS and MR of disease progression offer a promising opportunity to identify new treatments for glioma that could enhance survival (Davey Smith et al., 2017). Additionally, an improved understanding of the molecular changes that drive glioma progression will allow for opportunities to develop targeted molecular therapies. At present, although there are some examples where targeted therapy responses have been recorded in glioma patients, no targeted therapy has been approved as an effective treatment in clinical trials (Touat et al., 2017).

Future research will involve hypothesis free MR, which make use of omics data. There is a growing body of evidence showing that epigenetic biomarkers of glioma can be used for prediction and prognosis. Notably in neuro-oncology the O6-methylguanine-DNA methyltransferase promoter methylation can act both a prognostic and predictive biomarker for glioblastoma (Esteller et al., 2000; Olson et al., 2011; Reifenberger et al., 2012; Wick et al., 2012). As genetic variants associated with DNA methylation seem to overlap with expression quantitative trait loci (eQTLs) at many loci throughout the genome (Bell et al., 2012; Shi et al., 2014), both DNA methylation and gene expression may exist on the causal pathway between genetic variation and disease. The ability to identify epigenetic and transcriptomic markers for glioma risk and progression could be important in understanding the underlying mechanisms of glioma. Using an MR approach, the causal chain between DNA methylation, gene expression and glioma onset/progression can be investigated (Relton and Davey Smith, 2012).

Given the lack of large-scale case-only studies with data on progression and germline genetic data, a priority of research
in the near term should be to identify causes of glioma onset. The findings from such studies will be informative for the design of primary and secondary prevention strategies. The latter could be particularly valuable for glioma prevention in high risk populations, such as childhood cancer survivors (who received radiation therapy), people with genetic syndromes known to increase risk of glioma and people exposed to known causal factors because of their occupations. For example, if a specific dietary factor is found to be causally associated with a decrease in glioma risk, high risk populations could be advised to increase their consumption of that specific dietary factor.

CONCLUSION

Mendelian randomization offers a promising, novel way to identify risk factors and drug targets for glioma to both inform public health policy for prevention, as well as, allowing the development of therapeutic approaches to improve prognosis.

REFERENCES


The latter will require the development of large-scale case-only studies with data on progression and germline genetic data.

AUTHOR CONTRIBUTIONS

AH contributed to the manuscript research and writing. KK, PH, RM, CR, IZ, and AM reviewed and revised the manuscript.

FUNDING

This research was supported by Brain Tumour Bank and Research Fund and Southmead Hospital Charity. Charity registration number:1055900. RM, AM, and CR are supported by a Cancer Research United Kingdom programme grant, the Integrative Cancer Epidemiology Programme (Grant no. C18281/A19169). Dr. Haycock is supported by CRUK Population Research Postdoctoral Fellowship C52724/A20138.


Van Wijngaarden, E., and Dosemeci, M. (2006). Brain cancer mortality and potential occupational exposure to lead: findings from the national longitudinal...

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Howell, Zheng, Haycock, McAleenan, Relton, Martin and Kurian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.