Prediagnostic biomarkers for early detection of glioma—using case–control studies from cohorts as study approach

Wendy Yi-Ying Wu, Anna M. Dahlin, Carl Wibom, Benny Björkblom, and Beatrice Melin

Department of Radiation Sciences, Oncology, Umeå University, Umeå, Sweden (W.Y.-Y.W., A.M.D., C.W., B.M.); Department of Chemistry, Umeå University, Umeå, Sweden (B.B.)

Corresponding Author: Beatrice Melin, MD, PhD, Department of Radiation Sciences, Umeå University, SE-90187 Umeå, Sweden (Beatrice.melin@umu.se).

Abstract
Background. Understanding the trajectory and development of disease is important and the knowledge can be used to find novel targets for therapy and new diagnostic tools for early diagnosis.
Methods. Large cohorts from different parts of the world are unique assets for research as they have systematically collected plasma and DNA over long-time periods in healthy individuals, sometimes even with repeated samples. Over time, the population in the cohort are diagnosed with many different diseases, including brain tumors.
Results. Recent studies have detected genetic variants that are associated with increased risk of glioblastoma and lower grade gliomas specifically. The impact for genetic markers to predict disease in a healthy population has been deemed low, and a relevant question is if the genetic variants for glioma are associated with risk of disease or partly consist of genes associated to survival. Both metabolite and protein spectra are currently being explored for early detection of cancer.
Conclusions. We here present a focused review of studies of genetic variants, metabolomics, and proteomics studied in prediagnostic glioma samples and discuss their potential in early diagnostics.

Keywords
genetic variants | glioblastoma | metabolites | prediagnostic sample | proteins.

Several reviews have highlighted that a liquid biopsy in the form of a blood sample could potentially be useful in glioma diagnostics either by screening of high-risk individuals or at certain symptoms.1–3 The development of cancer has been suggested to start 7–8 years before diagnosis in breast and colorectal cancer, and recent studies have suggested a similar trajectory for glioma, which provides a strong rationale to study biomarkers that can be used for detection of these diseases at an earlier stage.4,5 Worldwide there are several cohorts with individuals that have given blood samples while healthy. These cohorts are followed over time and linkage to diagnostic registries allows the identification of individuals that later develop glioma, providing a unique opportunity to investigate the potential role of biomarkers for early detection of glioma. Studies nested within the cohorts are efficient, especially for investigating the useful biomarkers. Controls can be randomly chosen from the cohort who have not been diagnosed at the same time point as the corresponding case, that is, nested case–control study design. In addition to matching on the time of diagnosis, controls can also be matched on several factors such as age at blood sampling, gender, and time point of blood sampling to reduce confounding. An alternative study design, case–cohort design, is using a case set for different diseases and comparing with cohort controls, that is, a subcohort that represents the whole cohort.6

The current review highlights some of the existing studies performed in prediagnostic samples with different omics techniques.
| Year | Author            | Country | Study design                  | Sample size | Sample collection time prior to diagnosis | Biomarkers | Measurement | Findings                                                                                                                                                                                                 |
|------|-------------------|---------|--------------------------------|-------------|------------------------------------------|------------|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2012 | Rajaraman et al.  | Worldwide| Meta-analysis (14 cohorts, 3 case-control studies, and 1 population-based case-only study) | 556 cases from cohort studies and 3649 controls | Genome-wide variants | Illumina 550K, 660W, or 610Q | The consistent findings were shown in cohort and case-control study designs. Stronger associations were found for loci rs6010620 (20q,13.33; RTEL) and rs2736100 (5p15.33, TERT) in cohort studies. |
| 2015 | Wibom et al.      | Norway  | Nested case-control study      | 598 cases and 595 matched controls | Median: 14.7 y, range: 0.2–35.1 y | 1 single nucleotide polymorphisms | Amplified 6-FAM-labeled PCR | The study confirmed the genetic variants within 5 genomic regions: 8q24.21 (CCDC26), 9p21.3 (CDKN2B-AS1), 1q23.3 (PHLDB1), 17p13.1 (TP53), and 20q13.33 (RTEL1). |
| 2016 | Björkblom et al.  | Norway  | Nested case-control study      | 110 cases and 110 matched controls | Mean: 12.6 y, SD: 5.1 y | 180 known metabolites | GCxGC-TOFMS | 9 metabolites (γ-tocopherol, α-tocopherol, erythritol, erythronic acid, myo-inositol, cysteine, 2-keto-3-gluconic acid, hypoxanthine and xanthine) were involved in antioxidant metabolism. |
| 2017 | Huang et al.      | Finland | Nested case-control study      | 64 cases and 64 matched controls | Median: 9.0 y, interdecile range: 3–20 y | 750 known molecules | LC/MS–MS | 43 metabolites were associated with glioma before multiple adjustment. 2-Oxyarginine, cysteine, alpha-ketoglutarate, chenodeoxycholate and arginine were inversely associated with overall glioma risk. 7 xanthine metabolites related to caffeine metabolism were higher in cases than in controls. Ascorbate/aldarate and steroid hormone metabolites associated with high-grade glioma. |
| 2020 | Jonsson et al.    | Sweden  | Nested case-control study      | 132 cases and 132 matched controls | Mean: 11.7 y, SD: 5.6 (baseline); mean 4.2 y, SD: 3 y (repeated); mean: 8 y, SD: 5.4 y (single) | 142 metabolites | GC–MS | 15 significantly metabolites associated with glioma progression were identified. Higher plasma levels of myo-inositol, cysteine, N-acetylg glucosamine, creatinine, glycine, proline, erythronic, 4-hydroxyphenylacetic-, uric-, and aceturic acid were observed in glioma cases. |
| 2007 | Lönn et al.       | Finland | Nested case-control study      | 22 cases and 400 unmatched controls | More than 5 y | IGF-I and IGFBP-3 | ELISA | IGF-I was inversely associated with glioma but not IGFBP-3. |
| 2011 | Schlehofer et al. | EPIC    | Nested case-control study      | 275 cases and 528 matched controls | Median: 8.24 y, range: 29–4981 days | Specific IgE | ImmunoCAP specific IgE test | Allergic sensitization was inversely associated with glioma, especially pronounced in women. |
| Year | Author | Country | Study design | Sample size | Sample collection time prior to diagnosis | Biomarkers | Measurement | Findings |
|------|--------|---------|--------------|-------------|------------------------------------------|------------|-------------|----------|
| 2011 | Calboli et al. | USA | Nested case–control study | 169 cases and 520 matched controls | Total IgE, food, and respiratory allergen-specific IgE | UniCAP and ImmunoCAP fluorescent assays | Borderline elevated total IgE levels (25–100 kU/L) were inversely associated with glioma but not elevated IgE (>100 kU/L). Food allergen-specific and respiratory allergen-specific IgE levels showed no association with glioma. |
| 2011 | Rohrmann et al. | EPIC | Nested case–control study | 282 cases and 561 matched controls | IGF-I and IGFBP-3 | ELISA | Higher levels of IGF-I were positively associated with low-grade glioma risk and pronounced after adjustment for IGFBP-3, but not with glioma overall. |
| 2012 | Schwartzbaum et al. | Norway | Nested case–control study | 594 cases and 1177 matched controls | Total IgE and respiratory allergen-specific IgE | ImmunoCAP fluorescent assays | Elevated total IgE (>100 kU/L) levels were inversely associated with glioma risk and found even at least 20 years before diagnosis. Elevated allergen-specific IgE (>0.35 kU(A)/L) levels were inversely associated with glioblastoma in women but not in men. |
| 2015 | Ma et al. | Meta-analysis (5 studies in glioma) | 2461 cases and 3934 controls (2021 glioma cases and 3446 controls) | | | | Elevated total IgE levels were inversely associated with glioma risk. |
| 2015 | Schwartzbaum et al. | Norway | Nested case–control study | 487 cases and 487 matched controls | IL4, IL13, IL5, IL6, IL10, IFNG, TGFβ2, sIL4RA, sIL13RA2, FOXP3, STAT3, and STAT6 | RayBio Human Cytokine Antibody Array Kits | IL4 and sIL4RA were inversely associated with glioma and glioblastoma. TGFβ2 was inversely associated with glioblastoma. |
| 2016 | Späth et al. | Norway | Nested case–control study | 593 cases and 590 matched controls | Mean: 15.2 y, SD: 8.6 y | EGFR and ErbB2 | EGFR and ErbB2 levels were associated with glioblastoma risk. High serum ErbB2 concentration was also associated with glioma risk overall ($P = .049$; OR = 1.39, 95% CI = 1.00–1.93). |
| 2017 | Schwartzbaum et al. | Norway | Nested case–control study | 487 cases and 487 matched controls | Median: 15 y, interquartile range: 9–21 y | 277 cytokines | RayBio Human Cytokine Antibody Array Kits | sIL10RB, VEGF, IL4, and sIL4RA were associated with glioma risk. Interaction between IL4 and sIL4RA was established. sIL10RB, VEGF, beta-Catenin and CCL22 were associated with glioma risk more than 10 years before diagnosis. LIF was inversely associated with glioma within 5 years before diagnosis. |
| Year | Author          | Country | Study design          | Sample size | Sample collection time prior to diagnosis | Biomarkers     | Measurement                                      | Findings                                                                 |
|------|-----------------|---------|-----------------------|-------------|-------------------------------------------|----------------|------------------------------------------------|-------------------------------------------------------------------------|
| 2018 | Brenner et al.  | USA     | Nested case–control study | 457 cases and 457 matched controls | Median: 15 y, interquartile range: 9–21 y | 14 cytokines | Four sensitive custom V-PLEX Meso Scale Discovery kits (MSD) and 1 standard MSD kits | IL-15 and IL-16 were inversely associated with glioma risks.            |
| 2021 | Cote et al.     | UK      | Cohort                | 428,537 participants and 417 incident gliomas | Median: 3.8 y | CRP, WBC, NLR, and IGF-I | Beckman Coulter AU5800 analytical platform, LH750 instruments, and DiaSorin Liaison XL analytic platform | IGF-I was associated with glioma risk in women but not men.             |
| 2022 | Wu et al.       | Sweden  | Nested case–control study | 133 cases and 133 matched controls | Mean: 11.8 y, SD: 5.6 y (baseline); mean: 4.3 y, SD: 3.0 y (repeated); mean: 8 y, SD: 5.4 y (single) | 19 proteins | Luminex bead-based commercial assay panels and ELISA assays | sVEGFR2, sTNFR2, sIL-2Rα, and sIL-6R were associated with glioma risk. |
Methods

Search Strategy and Study Selection

We performed a comprehensive literature search in PubMed database to identify relevant studies published through February 23, 2022. The searches typically included 4 key terms “glioma,” “biomarkers,” “early-detection,” and “pre-diagnostic sample.” We especially focus on the early-detection role of biomarkers. The references of the identified articles were also searched for other relevant articles. Studies included into this review have to meet the following inclusion criteria: (1) Studies used prediagnostic samples to investigate the relationship between biomarkers, including genetic variants, metabolites, and proteins, and glioma risk; (2) Studies in cohort design or nested within cohort design (i.e., nested case–control study and case–cohort study). We excluded the studies that were not written in English, not conducted on humans and published as letters, case reports, and meeting records.

Results

Table 1 summarizes the studies that were included in this review. The following section was divided into (1) genetic predisposition, (2) metabolites, and (3) proteins.

Genetic Predisposition

Glioma predisposition caused by highly penetrant gene mutations occurs at low incidence. In addition to these rare mutations, common genetic variants have been linked to the risk of developing glioma. The genetic variants may be deeply involved in the biological development of glioma, but the functional mechanisms are only known in a few cases. The fact that the genetic variants in most cases are located in or in close proximity to genes commonly somatically mutated in glioma suggests a functional importance (Figure 1).

The common genetic risk variants for glioma have been confirmed in several independent studies, of which most are using a case–control design. Since an individual’s germline genetic variants will not change during the course of disease, causation bias is not an issue in these studies. Selection bias and survival bias are however of more concern in case–control studies of a disease with high disease-related mortality such as glioma. Considering the potential survival bias, Rajaraman et al. compared the findings from 7 susceptibility regions, including TERT, EGFR, CCDC26, CDKN2B, PHLDB1, and RETL1, between cohort and case–control design. The consistent findings were shown in both designs. However, greater association was found in 2 variants (rs6010620 in RETL and rs2736100 in TERT) in cohort studies, which implied a certain extent of survival bias. Wibom et al. demonstrated the potential of using the nested case–control study design to validate the previous GWAS findings. Most genetic variants have an association in the same direction as previous studies, even if there is limited power for significant results.

Genetic variants with small effects can be combined and used as polygenetic risk score to estimate the glioma risk. However, due to low absolute lifetime risk, using genetic variants as biomarkers for early detection in the population-level screening is not useful. Although the common genetic variants alone are not useful for risk prediction in the general population, the genetic variants have been associated with
different subgroups of glioma, lending some support that the variants could be included into future combined biomarker panels including other more predictive biomolecules.

**Metabolites**

Glioma can be defined as a metabolic driven disease both by in 1 subgroup detection of IDH mutations and the metabolic reprogramming of different subgroups. Prediagnostic samples are preferred as the samples taken at glioma diagnoses and at surgery might be affected by several factors that might change, such as stress, treatment with corticosteroids, seizure control drugs, or the sample been taken with arterial or venous puncture. A standard protocol such as sample collection, storage, preparation steps, and operating processes is therefore crucial to control the noise and variation in the experiments for reproducibility. Data preprocessing procedures, that is, identifying metabolites from raw data, and suitable multivariate analysis, also play important roles to find the true association. Few studies have investigated broader metabolite spectra in gliomas using prediagnostic samples. In a study of prediagnostic serum samples from JANUS biobank in Norway, we observed increased levels of several metabolites such as tocopherol, erythritol, and myo-inositol in samples from individuals that later in life have developed glioblastoma. In this study, most of the cases had been sampled 5–20 years before diagnosis. A small study of 64 cases and 64 matched controls that covered wider spectra of metabolites identified 43 associated metabolites, including arginine/proline, antioxidant, and coffee-related metabolites. In another study, 64 glioma cases from Northern Sweden Health and Disease study with repeated plasma samples and 68 single time point cases were analyzed. Tightly matched controls were used from the same sampling year, similar thawing cycles, smoking, and body mass index to be able to separate differences clearly associated with disease. Fifteen significant metabolites associated with the glioma progression were identified by comparing the 2 repeated samples between cases and matched controls. A elevated metabolic pattern in glioma cases was observed in blood plasma of several metabolites for example myo-inositol, cysteine, N-acetylg glucosamine, creatinine, glycine, and proline. This study highlighted the benefits of using repeated samples for progression pattern analysis. We combined the significant metabolites identified from 3 studies to investigate the potential metabolic pathways (Figure 2). The results suggest

---

**Figure 2.** Over representation analysis of significant metabolites found in prediagnostic glioma blood samples enriched in metabolic pathways of the Small Molecule Pathway Database (SMPDB).
that glutathione metabolism, cysteine metabolism, urea cycle, inositol phosphate metabolism, ammonia recycling, glutamate metabolism, inositol metabolism, and glycine and serine metabolism might involve in glioma tumorigenesis. A common limitation of using collected sample from retrospective cohort is that participants often have been recruited 10–20 years ago and some of glioma cases might not able to use new classification for example for IDH mutation status. Larger validation studies of the most promising metabolites are necessary to understand the potential of metabolic biomarkers in different diagnostic settings.

### Proteins

Few studies have explored proteomics in prediagnostic samples from glioma cases, likely due to the fact that the techniques still are being developed to be able to study the full proteome. A suggested theory for glioma development is that the innate immune system is of importance of glioma development. The innate immune system could release proteins before the development of symptoms that would render a radiological examination. One potential biological mechanism would be chronic inflammation as a mediator in tumor development. For example, release of inflammatory proteins such as cytokines that have been suggested also for several other cancer categories. In glioma, studies which investigated few proteins have also been done such as immunoglobulin E (IgE), insulin-like growth factor (IGF), EGFR, and ErbB2. IgE and specific IgE are allergy biomarkers which have been used in medical atopy diagnostics. There were 3 nested case–control studies that using prediagnostic samples to investigate the association between serum IgE level and glioma risk. Two of them showed significantly negative association, especially in women. A review study concluded negative association between total IgE level but not respiratory allergen-specific IgE.

IGF plays important role in human growth and normal brain development and it was associated with increased risk of several cancers. Two nested case–control studies and 1 cohort have investigated the association between IGF-I and IGF-binding proteins and glioma risk in prediagnostic samples. Higher circulating IGF-I seemed to increase the risk of low-grade gliomas but reverse causation bias could not be excluded. Interestingly, a Mendelian randomization study using genetic instruments for serum IgE and plasma IGF-I levels could not support their etiological roles of glioma.

### Conclusions

The development of broad omics analyses with need of limited amount of DNA and plasma have paved the way for good opportunities to explore both the etiology and early detection of disease in the era of personalized medicine and precision diagnostics. Genetic variants have been discovered that gives us understanding of the etiology of glioma but have limited contribution to risk prediction. Metabolite studies show some promising results but needs further confirmation. Most importantly, the samples need to go along with adequate health data information of the cohort individuals, and information on all preanalytical sample conditions that are corner stones for being able to detect true biological relevant biomarkers. As glioma is a rare disease, collaborative efforts with several independent validations are necessary to find robust biomarkers that can be taken forward to clinical trials.

### Funding

This study was funded by grants from the Swedish Research Council, Swedish Cancer foundation and Umeå University hospital grant and Sjöberg foundation.

### Conflict of interest statement

None declared.

### References

1. Best MG, Sol N, Zijl S, et al. Liquid biopsies in patients with diffuse glioma. Acta Neuropathol. 2015;129(6):849–865.
2. Ali H, Harting R, de Vries R, Ali M, Wurdinger T, Best MG. Blood-based biomarkers for glioma in the context of gliomagenesis: a systematic review. Front Oncol. 2021;11:665235.

3. McCarthy M, Binney E. Personalized profiles for disease risk must capture all facets of health. Nature. 2021;597(7875):175–177.

4. Zahir N, Sun R, Gallahan D, Gatenby RA, Curtis C. Characterizing the ecological and evolutionary dynamics of cancer. Nat Genet. 2020;52(8):759–767.

5. Barthel FP, Johnson KC, Varn FS, et al. Longitudinal molecular trajectories of diffuse glioma in adults. Nature. 2019;576(7785):112–120.

6. Forouhi NG, Wareham NJ. The EPIC-InterAct study: a study of the interplay between genetic and lifestyle behavioral factors on the risk of type 2 diabetes in European populations. Curr Nutr Rep. 2014;3(4):355–363.

7. Johansson G, Andersson U, Melin B. Recent developments in brain tumor predisposing syndromes. Acta Oncol. 2016;55(4):401–411.

8. Melin BS, Barnholtz-Sloan JS, Wrensch MR, et al. Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. Nat Genet. 2017;49(5):789–794.

9. Kimmersley B, Labussiere M, Holroyd A, et al. Genome-wide association study identifies multiple susceptibility loci for glioma. Nat Commun. 2015;6:8559.

10. Andersson U, Schwartzbaum J, Wulkan F, et al. A comprehensive study of the association between the EGFR and ERBB2 genes and glioma risk. Acta Oncol. 2010;49(6):767–775.

11. Shete S, Hosking FJ, Robertson LB, et al. Genome-wide association study identifies five susceptibility loci for glioma. Nat Genet. 2009;41(8):899–904.

12. Wang Z, Rajaraman P, Melin BS, et al. Further confirmation of germline glioma risk variant rs78378222 in TP53 and its implication in tumor tissues via integrative analysis of TCGA data. Hum Mutat. 2015;36(7):684–688.

13. Rajaraman P, Melin BS, Wang Z, et al. Genome-wide association study of glioma and meta-analysis. Hum Genet. 2012;131(12):1877–1888.

14. Wibom C, Spith F, Dahlin AM, et al. Investigation of established genetic risk variants for glioma in prediagnostic samples from a population-based nested case-control study. Cancer Epidemiol Biomarkers Prev. 2015;24(5):810–816.

15. Eckel-Passow JE, Decker PA, Kosel MI, et al. Using germline variants to estimate glioma and subtype risks. Neuro Oncol. 2019;21(4):451–461.

16. Wu WY, Johansson G, Wibom C, et al. The genetic architecture of gliomagenesis-genetic risk variants linked to specific molecular subtypes. Cancers (Basel). 2019;11(12):2001.

17. Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. Science. 2008;321(5897):1807–1812.

18. Xiong N, Gao X, Zhao H, et al. Using arterial-venous analysis to characterize cancer metabolic consumption in patients. Nat Commun. 2020;11(1):3169. doi:10.1038/s41467-020-16810-8.

19. Ren S, Hinzman AA, Kang EL, Szczesniak RD, Lu LJ. Computational and statistical analysis of metabolomics data. Metabolomics. 2015;11(6):1492–1513.

20. Bjorkblom B, Wibom C, Jonsson P, et al. Metabolomic screening of prediagnostic serum samples identifies association between alpha- and gamma-tocopherol and glioblastoma risk. Oncotarget. 2016;7(24):37043–37053.

21. Huang J, Weinstein SJ, Kitahara CM, et al. A prospective study of serum metabolites and glioma risk. Oncotarget. 2017;8(41):70386–70377.

22. Jonsson P, Antti H, Spith F, Melin B, Bjorkblom B. Identification of pre-diagnostic metabolic patterns for glioma using subset analysis of matched repeated time points. Cancers (Basel). 2020;12(11):3349.

23. Coussens LM, Werb Z. Inflammation and cancer. Nature. 2002;420(6917):860–867.

24. Schlehofer B, Siegmund B, Linseisen J, et al. Primary brain tumours and specific serum immunoglobulin E: a case-control study nested in the European Prospective Investigation into Cancer and Nutrition cohort. Allergy. 2011;66(11):1434–1441.

25. Schwartzbaum J, Ding B, Johanesen TB, et al. Association between prediagnostic IgE levels and risk of glioma. J Natl Cancer Inst. 2012;104(16):1251–1259.

26. Calboli FC, Cox DG, Buring JE, et al. Prediagnostic plasma IgE levels and risk of adult glioma in four prospective cohort studies. J Natl Cancer Inst. 2011;103(21):1588–1595.

27. Ma C, Cao L, Zhao J, et al. Inverse association between prediagnostic IgE levels and the risk of brain tumors: a systematic review and meta-analysis. Biomed Res Int. 2015;2015:294213.

28. Qian F, Huo D. Circulating insulin-like growth factor-I and risk of total and 19 site-specific cancers: cohort study analyses from the UK Biobank. Cancer Epidemiol Biomarkers Prev. 2020;29(11):2332–2342.

29. Rohrmann S, Linseisen J, Becker S, et al. Concentrations of IGF-I and IGFBP-3 and brain tumor risk in the European Prospective Investigation into Cancer and Nutrition. Cancer Epidemiol Biomarkers Prev. 2011;20(10):2174–2182.

30. Lonn S, Inskip PD, Pollak MN, et al. Glioma risk in relation to serum levels of insulin-like growth factors. Cancer Epidemiol Biomarkers Prev. 2007;16(4):844–846.

31. Cote DJ, Creed HD, Samanic CM, et al. A prospective study of inflammatory biomarkers and growth factors and risk of glioma in the UK Biobank. Cancer Epidemiol. 2021;75:102043.

32. Saunders CN, Cornish AJ, Kinnersley B, et al. Lack of association between modifiable exposures and glioma risk: a Mendelian randomization analysis. Neuro Oncol. 2020;22(2):207–215.

33. Spith F, Andersson U, Dahlin AM, et al. Pre-diagnostic serum levels of EGFR and ErbB2 and genetic glioma risk variants: a nested case-control study. Tumour Biol. 2016;37(8):11065–11072.

34. Brenner AV, Inskip PD, Rusiecki J, et al. Serially measured pre-diagnostic levels of serum cytokines and risk of brain cancer in active component military personnel. Br J Cancer. 2018;119(7):893–900.

35. Schwartzbaum J, Wang M, Root E, et al. A nested case-control study of 277 prediagnostic serum cytokines and glioma. PLoS One. 2017;12(6):e0178705.

36. Wu WY, Spith F, Wibom C, et al. Pre-diagnostic levels of sVEGFR2, sNFR2, sIL-2Ralpha and sIL-6R are associated with glioma risk: a nested case-control study of repeated samples. Cancer Med. 2022;11(4):1016–1025.