

NBT
National Brain Tumour Study
In association with Gliogene
PROTOCOL
Version 7

Chief Investigator

Richard S Houlston MD PhD FRCP FRCPATH FMedSci
Professor in Molecular and Population Genetics

Division of Genetics and Epidemiology
Institute of Cancer Research,
Brookes-Lawley Building
Cotswold Road
Sutton
Surrey, SM2 5NG

Tel: +44(0)208 722 4175
Fax: +44(0)208 722 4059
e-mail: richard.houlston@icr.ac.uk

Protocol last amended: June 2014
Final copy: please destroy all previous drafts

The study is funded by the National Cancer Institute

1. AIMS

- To establish a DNA resource to enable identification of genetic variants conferring an elevated risk of gliomas.
- To identify genetic variants influencing glioma risk through genome-wide scans of Single Nucleotide Polymorphisms (SNPs) and other association-based analyses.
- To ascertain and collect families segregating gliomas and using these identify novel susceptibility alleles through genetic linkage.

2. BACKGROUND

- Primary tumours of the central nervous system (CNS) in adults are the third most common tumour in men and sixth most common tumour in women between the ages of 35 and 49 years. Meningiomas and gliomas are the principal primary brain tumours in adults. The prognosis associated with gliomas especially is dismal, with poor survival and high morbidity.
- Epidemiological case-control and cohort studies have consistently shown cancer risk to be elevated 2 to 3-fold in relatives of patients with glioma and other primary brain tumours (PBTs)[1-4]. The familial aggregation of tumours could solely be due to shared environmental exposure but to date no general lifestyle exposures have been consistently linked to an increased risk of any of the tumour types. The only environmental factor that has been established to induce PBTs is high cranial doses of ionising radiation [5]. Such exposures are rare, however, accounting for few cases of PBT, most being meningiomas.
- Evidence for an inherited predisposition to PBTs is convincingly provided by the Li-Fraumeni syndrome (MIM151623), neurofibromatosis (NF) type 1 and type 2 (MIM162200 & MIM101000), tuberose sclerosis (MIM191100) and Turcot's syndrome (MIM276300). While these inherited syndromes confer substantive risks of PBTs to those with the susceptibility alleles, collectively their contribution to the overall incidence of PBTs is small and they are therefore will not account for much of the familial risk.
- Outside the context of the Mendelian cancer family syndromes that confer an increased risk of PBTs, the underlying genetic basis of the observed familial risk is unclear. Clustering of disease in families is not confined to a specific histology, with aggregation of gliomas and meningiomas in pedigrees. This, together with the observation of familial associations between PBTs and breast cancer and melanoma, implies that predisposition is in part mediated by alleles with pleiotropic effects.
- Two segregation analyses of non-syndromic glioma have provided evidence to support the involvement of additional loci in susceptibility, one providing support for a major gene[6, 7]. Such an assertion is supported by linkage data[8]. Collectively these data provide evidence for an inherited predisposition to glioma outside the context of the rare cancer prone families and a clear rationale for seeking to identify a disease causing locus by genetic linkage analysis. The feasibility of this strategy will be contingent on the ascertainment and collection of a large series of families, from which germline mutations in the known susceptibility genes have been excluded. This has been the motivation for developing the GLIOGENE Consortium Study which aims to ascertain and collect families segregating two or more gliomas in which involvement of known genes has been excluded. Twelve centres from the US, Sweden, Denmark, UK, and Israel are currently participating in GLIOGENE.
- Whilst part of the residual familial risk may be mediated through highly penetrant alleles detectable by genetic linkage analyses it is highly probable that a component of the familial risk of PBTs is polygenic, mediated through the action of low penetrance alleles, some of which may be common. The identification of alleles conferring relative risks (RRs) of <2.0 is contingent on association studies, because risks of this magnitude rarely cause multiple-case families and are therefore difficult, if not impossible, to identify through linkage searches[9]. The identification of low penetrance alleles through association analyses is, however, contingent on the ascertainment and collection of a large series of cases and controls.

- Through participation in the INTERPHONE study (a case-control study of the impact of mobile phone use on risk of PBTs) we have developed the necessary procedures for ascertaining family histories, blood and tumour samples from brain tumour patients.
- We propose to extend this by collecting samples from a series of 5,000 patients with glioma. To do this we will ascertain cases through centres specialising in the treatment of gliomas. Cases with a family history of gliomas will be recruited into the GLIOGENE/NBT Study. Using this resource we will be able to conduct association studies to identify low penetrance alleles for the disease.

3. DETAILED PLAN OF INVESTIGATION

3.1 Ascertainment of patients and families

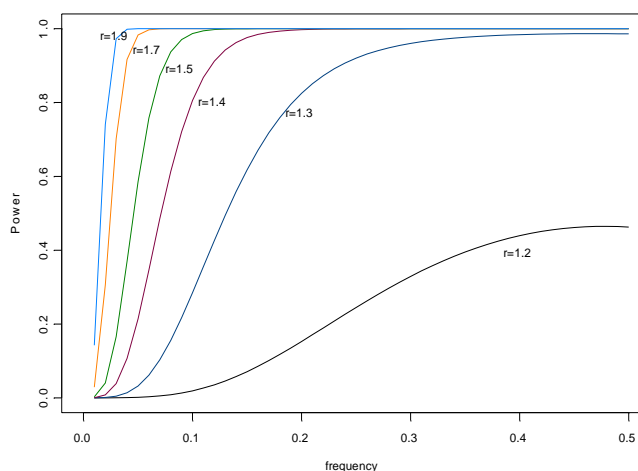
- Patients will be identified through their participation in National Trials or through centres specialising in the management of the disease.
- Centres will be recruited by writing to Clinicians at specialist centres and Clinical Genetics Centres inviting them to participate and to act as Lead Clinician at their Centre (or nominate a colleague), and also to identify a person (e.g. a NCRN nurse) to act as the local administrative coordinator of the study.
- Patients diagnosed with glioma within the last five years who remain alive will be identified from local databases, outpatient clinics and/or central records of trial participants according to local preferences.
- Eligible patients will be invited to participate in the study either directly or by a letter from their hospital consultant. If the patient is not currently under review of their clinical status, their General Practitioner (through letter or direct contact) will first be asked whether there is any reason why the individual should not be invited to participate.
- The letter (if used) to the patient will outline the study and will also provide a detailed patient information sheet and consent form. The letter will request that, if they are willing to take part in the study, they return the consent form to the Coordinating Centre at the Institute of Cancer Research in a pre-paid envelope. If the patient is ascertained directly through outpatients the letter will be omitted from the ascertainment process.
- Patients who agree to participate in the NBT study will be asked to complete a questionnaire detailing demographic information on all first degree relatives, incidences of cancer in other blood relatives and their own prior cancer history.
- In addition to ascertaining information from each patient, permission will be sought to collect a 10-20ml EDTA venous blood sample. These blood samples will be taken by centre nurses or the patient's General Practitioner/Practice Nurse after obtaining their agreement.
- Permission will be requested from all patients participating in the study to obtain confirmation of diagnosis by referral to medical records and to gain access to archival histopathology material relating to their diagnosis of glioma.
- Patients with a family history of PBTs may be asked to invite family members to join the international familial GLIOGENE study. These patients will be sent an "Invitation to Invite Relatives" letter and one information pack per affected family member, who is eligible to join the familial study.
- Should patients wish to invite relatives to join the NBT study, they would pass the information pack directly on to them, in the pre-paid envelope provided.
- If relatives receive an information pack from the index case and are willing to join the study, they can read the Information Sheet and return the completed reply slip in the Freepost envelope provided.
- Once a completed reply slip is received at the ICR, a researcher can contact the relative and send a sample collection pack to them, if they are willing to provide samples.
- Their general practitioner will take a 10-20ml EDTA venous blood and this sample will be returned to the ICR using the pre-paid envelope provided.
- Under special circumstances where an affected family member has deceased, we will request the patient to contact the deceased person's spouse and children where possible, on our behalf, to invite them to join the study and provide samples.

3.2 Laboratory methods

3.2.1 Identification of low penetrance alleles through genomwide association analyses

- **Rationale for genome-wide association analysis of tagged SNPs** - Following the sequencing of the human genome, large-scale harvests of SNPs have been conducted and >10 million documented with smaller numbers of small insertion/deletion and copy number polymorphisms identified. Patterns of linkage disequilibrium (LD) between SNPs have been characterised allowing subsets of SNPs (tagging SNPs) to be selected that, through LD with other variants, capture a large proportion of the common sequence variation in the human genome. The high resolution LD maps now available and hence comprehensive sets of tagging SNPs allow whole genome-wide studies for disease associations to be conducted cost effectively. This approach is unbiased and does not depend upon prior knowledge of function or presumptive involvement of any gene in disease causation. Moreover, it avoids the possibility of missing the identification of important variants in hitherto unstudied genes.
- The ongoing development of highly efficient analytical platforms for parallel processing of SNPs makes commentary of which precise technology will be employed premature. However, it is likely that at this juncture, although it may be slightly premature, the strategies adopted may be along the following lines: Stage 1. Analysis of 2,500 cases from the United Kingdom (UK) compared to 2,500 UK controls using a set of 550,000 tagging SNPs.; Stage 2. Analysis of a further 2,500 UK cases compared with 2,500 controls using ~30,000 SNPs that show the strongest association; Stage 3. Validation and confirmation analyses - Analysis of SNPs that show association in Stages 1 and 2 at a significance level of 10^{-4} or better will be evaluated in additional series of cases from the UK, USA and Europe and regionally matched controls; Stage 4. Detailed evaluation of statistically significant loci.
- Power calculations: It is well recognised that as genome-wide association studies involve typing a vast number of markers a large number of false positive associations will inevitably be generated, with only a small number being truly associated with disease susceptibility. Hence associations need to attain a high level of statistical significance to be established beyond reasonable doubt. On this basis significance levels of $\sim 10^{-7}$ have been advocated[9]. The power of the proposed study has been computed stipulating a significance level of 10^{-4} over Stages 1 and 2 combined. Given the staged design, only ~55 SNPs would be expected to be significant at this level by chance plus a further 10+ (depending upon the number of true associations), leaving a manageable number of loci to evaluate in the follow-up stages. Since the number of samples available to us for follow-up analyses after Stage 2 should potentially be substantial, any "true" susceptibility should be readily confirmed at 10^{-7} or better. The power of the proposed study to detect a susceptibility allele is detailed in Figure 1, for a range of allele frequencies and relative risks, under a multiplicative model. To be highly conservative we have not assumed any genetic enrichment of cases as enrichment in Stage 1 will in part be offset by characteristics of samples typed in Stage 2.

Figure 1. Estimated combined power of study to identify a susceptibility locus conferring a given genotypic relative risk (r) over a range of allele frequencies ($P=10^{-4}$ after two stages) assuming a multiplicative model; Calculations were made using the following assumptions: (a). Marker and disease allele have identical frequencies; (b). Distribution of r^2 as a measure of LD between susceptibility locus and tagging SNP based on published data.



- For common alleles the power to detect a non-recessive disease locus is largely governed by the contribution the locus makes to the overall inherited genetic variance. Figure 1 shows that our study will be adequately powered to identify susceptibility loci even under relatively poor scenarios. Specifically, for alleles with frequencies >0.1 , provided that the locus accounts for $>2\%$ of the familial risk, our study will have 80% power to identify a disease locus. As expected, power diminishes for effect sizes that are smaller than this. For alleles acting in a recessive fashion, unless they are very common, far larger studies than ours will be required for their identification. The only feasible way to do this will be through multi-centre international cooperative efforts. Our study will be ideally placed to contribute to such collaborations.

3.2.2 Genetic linkage analysis of glioma families

- Genetic linkage analysis of families will be conducted in concert with the GLIOGENE consortium which are major investigators. Genome-wide linkage scans will be conducted using high density SNP arrays.
- **Positional cloning-** For regions with evidence of linkage we shall undertake measures designed to identify the disease gene. We shall prioritise regions for analysis based on: 1. Maximum LOD/NPL score; Size of the supported region; Plausibility of candidate genes mapping to the region of linkage; Allelic loss data. Exons and flanking intronic sequences of genes mapping to linked regions will be screened for sequence changes. If variants of equivocal functional importance are found, we shall screen DNA from 200 healthy individuals for the presence of the variant.
- **Statistical considerations-** Reliably predicting the number of families required to identify a predisposition locus by linkage is rarely straightforward as locus heterogeneity and phenocopies erode power. We are, however optimistic that through participation in GLIOGENE we shall maximise our chances of identifying a novel predisposition locus.

4. FUNCTIONAL ANALYSES

- Once any potential disease-associated variant has been identified, further work will be needed to assess the validity of the association and to determine the mechanism by which the variant increases risk. Functional studies will then be required, as well as an assessment of the contribution of the novel gene to other types of cancer and its interactions with other genes.

5. VALUE OF STUDY

- The two major reasons why this field of research has been, and will continue to be, of importance are: 1. The ability to identify those at increased risk is of immediate clinical relevance, in terms of primary and secondary interventions; 2. The identification of susceptibility genes provides a greater understanding of the mechanisms of carcinogenesis, offering potential targets for therapeutic interventions

6. WEB ADDRESSES

INTERPHONE STUDY- <http://www.iarc.fr/ENG/Units/RCAAd.html>

GLIOGENE- <http://braintumor.epigenetic.org/>

ONLINE MENDELIAN INHERITANCE IN MAN

- <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=omim>

dbSNP- <http://www.dbsnp.org>

7. CITED REFERENCES

1. Malmer, B., R. Henriksson, and H. Gronberg, *Familial brain tumours-genetics or environment? A nationwide cohort study of cancer risk in spouses and first-degree relatives of brain tumour patients*. Int J Cancer, 2003. **106**(2): p. 260-3.
2. Wrensch, M., et al., *Familial and personal medical history of cancer and nervous system conditions among adults with glioma and controls*. Am J Epidemiol, 1997. **145**(7): p. 581-93.
3. Malmer, B., et al., *Familial aggregation of astrocytoma in northern Sweden: an epidemiological cohort study*. Int J Cancer, 1999. **81**(3): p. 366-70.
4. Hill, D.A., et al., *Cancer in first-degree relatives and risk of glioma in adults*. Cancer Epidemiol Biomarkers Prev, 2003. **12**(12): p. 1443-8.
5. Sadetzki, S., et al., *Radiation-induced meningioma: a descriptive study of 253 cases*. J Neurosurg, 2002. **97**(5): p. 1078-82.
6. Malmer, B., et al., *Genetic epidemiology of glioma*. Br J Cancer, 2001. **84**(3): p. 429-34.
7. de Andrade, M., et al., *Segregation analysis of cancer in families of glioma patients*. Genet Epidemiol, 2001. **20**(2): p. 258-70.
8. Malmer, B., et al., *Homozygosity mapping of familial glioma in Northern Sweden*. Acta Oncol, 2005. **44**(2): p. 114-9.
9. Risch, N. and K. Merikangas, *The future of genetic studies of complex human diseases*. Science, 1996. **273**(5281): p. 1516-7.
10. Sellick, G.S., et al., *Genomewide linkage searches for Mendelian disease loci can be efficiently conducted using high-density SNP genotyping arrays*. Nucleic Acids Res, 2004. **32**(20): p. e164.