GeCIP Detailed Research Plan Form

August 2015

Background

The Genomics England Clinical Interpretation Partnership (GeCIP) brings together researchers, clinicians and trainees from both academia and the NHS to analyse, refine and make new discoveries from the data from the 100,000 Genomes Project.

The aims of the partnerships are:

1. To optimise:

- clinical data and sample collection
- clinical reporting
- data validation and interpretation.

2. To improve understanding of the implications of genomic findings and improve the accuracy and reliability of information fed back to patients. To add to knowledge of the genetic basis of disease.

3. To provide a sustainable thriving training environment.

The initial wave of GeCIP domains was announced in June 2015 following a first round of applications in January 2015. On the 18th June 2015 we invited the inaugurated GeCIP domains to develop more detailed research plans working closely with Genomics England. These will be used to ensure that the plans are complimentary and add real value across the GeCIP portfolio and address the aims and objectives of the 100,000 Genomes Project. They will be shared with the MRC, Wellcome Trust, NIHR and Cancer Research UK as existing members of the GeCIP Board to give advance warning and manage funding requests to maximise the funds available to each domain. However, formal applications will then be needed to individual funders. They will allow Genomics England to plan shared core analyses and the required research and computing infrastructure to support the proposed research. They will also form the basis of assessment by the Project's Access Review Committee, to permit access to data. Some of you have requested a template for the research plan which we now provide herewith.

We are only expecting one research plan per domain and have designed this form to contain common features with funder application systems to minimise duplication of effort. Please do not hesitate to contact us if you need help or advice.

Domain leads are asked to complete all relevant sections of the GeCIP Detailed Research Plan Form, ensuring that you provide names of domain members involved in each aspect so we or funders can see who to approach if there are specific questions or feedback and that you provide details if your plan relies on a third party or commercial entity. You may also attach additional supporting documents including:

- a cover letter (optional)
- CV(s) from any new domain members which you have not already supplied (required)
- other supporting documents as relevant (optional)

Genomics England Clinical Interpretation Partnership (GeCIP) Detailed Research Plan Form

Application Summary	
GeCIP domain name	**Childhood Cancer**
Project title *(max 150 characters)*	**Childhood Tumour Research in the 100,000 Genomes Project**

Objectives. *Set out the key objectives of your research. (max 200 words)*

The aims of the domain are:

1. To improve genetic diagnosis for childhood tumours. This will allow better predictions about outcomes for such children and will identify targets for novel treatments.
2. To improve germline diagnosis for families of children with childhood tumours who may carry a predisposition to tumours.
3. To incorporate routine genomic testing in clinical practice for childhood tumours.

To meet these aims, we will pursue the following specific objectives:

1. To collect and sequence tumour and germline DNA from over 100 childhood tumours with the intention of identifying key pathogenic mutations in childhood tumours and to determine novel germline predispositions.
2. To establish routine clinical whole genome sequencing (WGS) across all GMCs both as part of the final phase of the 100,000 genomes project and in preparation for the NHSE reorganisation.
3. To form collaborative links between the many genetically driven trials in childhood tumours, the 100,000 genomes project and the NHSE reorganisation. To do this we will use the SMPaed programme as an exemplar.
4. To compare the diagnostic impact for childhood tumours of different 'omic platforms by comparing the WGS data from the 100,000 Genomes Project with standard-of-care diagnostic testing and the multiple platforms in SMPaed. This will provide an evidence base for future clinical testing.
5. To determine the genetic drivers of relapse in childhood tumours.

Lay summary. *Information from this summary may be displayed on a public facing website. Provide a brief lay summary of your planned research. (max 200 words)*

Cancer is the number one cause of death from disease in children. Each year, there are 1,600 new cases of childhood cancer and 260 children under the age of 15 die of the disease. With modern treatment, nearly 80% of children survive their tumour. However, it is clear that life expectancy and quality of life is reduced in survivors. Treatment has adverse effects on the child's development, particularly that of the nervous system. This indicates that we need novel therapies both for high-risk patients to provide cure and for low-risk patients to reduce toxicity. There are several challenges that the research in the childhood cancer domain will address:

1) identification of the genetic abnormalities underlying rare childhood tumours
2) development of diagnostic approaches to childhood tumours that allow prediction of outcome and predict response to novel treatments
3) identification of children with underlying inherited disposition to tumours

Technical summary. *Information from this summary may be displayed on a public facing website. Please include plans for methodology, including experimental design and expected outputs of the research. (max 500 words)*

Cancer is the number one cause of death from disease in children. Each year, there are 1,600 new cases of childhood cancer and 260 children under the age of 15 die of the disease. With modern treatment, nearly 80% of children survive their tumour. However, it is clear that life expectancy and quality of life is reduced in survivors. Treatment has adverse effects on the child's development, particularly that of the nervous system. Furthermore, some types of tumours have very poor outcomes, even with modern treatments.

A significant proportion of children develop tumours due to an inherited predisposition. Such a predisposition has major implications. These implications include not only the risk to that child or their family of developing another tumour, but also includes risks from their immediate treatment. This is because the genetic predisposition may mean that certain treatments (e.g. radiotherapy) carry extra risk.

The aims of the childhood cancer GeCIP domain are to improve genetic diagnosis of childhood tumours, identify the best ways to deliver genomic diagnosis, and to identify genetic abnormalities that drive relapse.

There is a tremendous opportunity in childhood cancer because in addition to the 100,000 Genomes Project, there is well-established genetic testing within the NHS for certain childhood tumours and many children are offered treatment on trials which include genomic analysis. This means we can exploit the synergies between these different testing platforms and results. To achieve this, the GeCIP will work closely with the existing molecular trials and diagnostic services. In particular, we will build links with the SMPaed study which is offering multi-omic profiling for childhood tumours, particularly at relapse, for trial screening.

The Childhood Cancer GeCIP will undertake the following workstreams:

Workstream 1: Collation and analysis of WGS data for patients recruited to the 100,000 Genomes project. We will analyse existing data to determine novel mutations and structural changes associated with childhood tumours and identify germline predisposition.
Workstream 2: Accelerated recruitment and transition to NHS provision. We will work with childhood primary treatment centres, the paediatric pathology group and the genomic centres to increase recruitment of childhood cases before the close of the project and optimise the transition to testing in the NHS.
Workstream 3: Collaboration with trial profiling to determine optimal diagnostic testing. We will work between Genomics England and the SMPaed study, as an exemplar of trial-based genomics, and integrate data from both studies. This offers an opportunity to determine the optimal testing strategy for childhood tumours.
Workstream 4: Genomic analysis of relapsed tumours. Using the combination of data from the SMPaed study and the 100,000 genomes project we will identify genetic changes emerging at relapse in childhood tumours.
Workstream 5: Assessment of shaken biopsy. We will assess the value of novel tissue handling protocols (such as shaken biopsy analysis) to improve the availability of tissue for genomic analysis in childhood tumours.

Expected start date	**Q4 2018**
Expected end date	**Q4 2021**

Lead Applicant(s)	
Name	Thomas S Jacques
Post	Professor of Paediatric Neuropathology and Honorary Consultant
Department	Great Ormond Street Institute of Child Health
Institution	University College London
Current commercial links	None

Administrative Support	
Name	Janette Gardener
Email	Janette.Gardener@gosh.nhs.uk
Telephone	020 7762 6108

Subdomain leads		
Name	**Subdomain**	**Institution**
Tom Jacques	Pathology	UCL
Dan Williamson/Bissan Al-Lazikani	Bioinformatics	Newcastle/ Institute of Cancer Research
Mike Hubank	Genomics	Institute of Cancer Research
Tom Jacques/Steve Clifford	Brain Tumours	UCL/Newcastle
Deb Tweddle/Lou Chesler	Non-CNS Tumours	Newcastle/Institute of Cancer Research

Detailed research plan

Full proposal (total max 1500 words per subdomain)	
Title *(max 150 characters)*	**Childhood Tumour Research in the 100,000 Genomes Project**

Importance. *Explain the need for research in this area, and the rationale for the research planned. Give sufficient details of other past and current research to show that the aims are scientifically justified. Please refer to the 100,000 Genomes Project acceptable use(s) that apply to the proposal (page 6).*

Cancer is the number one cause of death from disease in children. Each year, there are 1,600 new cases of childhood cancer and 260 children under the age of 15 die of the disease. With modern treatment, nearly 80% of children survive their tumour. However, it is clear that life expectancy and quality of life is reduced in survivors. Treatment has adverse effects on the child's development, particularly that of the nervous system. Furthermore, some types of tumours have very poor outcomes, even with modern treatments.

These factors mean that there is a great need to diagnose tumours accurately so that children are not under- or over-treated. Furthermore, there is a need to identify targetable mutations and pathways to develop novel treatments that can be used to treat the children with high risk tumours and to reduce treatment toxicity in other children.

There are factors that make genomics of childhood tumours distinct from that in adult tumours. Most studies indicate that childhood tumours carry fewer mutations at diagnosis than adult type tumours (and many of these may be in unique, developmentally related genes not yet targeted for adult cancer). It is also significant that many childhood tumours occur at very early ages with little time to gain additional mutations through tumour evolution. Also, there is a high risk of germline mutations predisposing to tumours.

Many traditionally recognised types of childhood tumour are being redefined, altered or eliminated by genetic studies. In contrast to the common adult tumours, many childhood tumour entities are now dependent on molecular testing for diagnosis. We anticipate that many children will have their diagnosis changed as more extensive genomic investigations become available.

Research plans. *Give details of the analyses and experimental approaches, study designs and techniques that will be used and timelines for your analysis. Describe the major challenges of the research and the steps required to mitigate these.*

The goals of the research plan are to improve genomic diagnosis for childhood tumours, both by identifying novel changes in childhood tumours, redefining diagnostic entities based on tumour genetics, identifying germline predisposition and understanding the biology of relapsed disease. There is an existing network of genomics studies in childhood tumours and our approach is to take account of existing national and international studies in order to maximise the potential synergies of these studies working with the 100,000 genomes project.

<u>National context of UK paediatric oncology research</u>
The plans for research in our domain are critically dependent on the interactions with the extensive national and international networks of tumour biology research that exist in UK paediatric oncology. This provides a huge opportunity for our research plans but also presents some specific logistical issues our research plans need to address.

Many children are recruited to trials in which there is up-front real time molecular and pathological analysis, often prior to randomisation. Furthermore, the SMPaed study is about to launch across the UK. SMPaed will provide real-time multi-omic profiling of childhood solid tumours and in particular, it will be used as a pre-screening platform for trials of novel agents for relapsed and treatment-refractory tumours.

In addition, national tissue banking for research projects is well established in the UK under the auspices of the Children's Cancer and Leukaemia Group (CCLG) tissue bank with all primary treatment centres for children's tumours routinely banking samples including frozen tumour, formalin fixed paraffin embedded tumours and fluids (bloods, plasma, CSF etc.).

There is a network of existing molecular testing as standard of care for some childhood tumours. For example, in an audit of the Great Ormond Street Hospital practice (representing approximately 45% of childhood brain tumours in the UK), over 85% of brain tumour cases had some form of molecular testing, often including DNA methylation profiling. There are also clinical diagnostic routes that have been established for molecular testing e.g. there is a funded service to provide molecular pathological investigation of medulloblastoma, offered via two hubs at Great Ormond Street Hospital and Newcastle.

The interaction between these trial-driven initiatives, the CCLG bank, the 100,000 Genomes project and the NHSE genomic reorganisation is a major opportunity for paediatric oncology. However, historically this possibility has not been achieved which has been much to the detriment

of all the projects but particularly, it has meant that paediatric recruitment to 100,000 Genomes has only taken place this year.

However, we are in excellent position to leverage these synergies as the SMPaed study is about to launch and most of the key investigators of SMPaed are also the core membership of the childhood GeCIP (Jacques, Chesler, Hargrave, Hubank, Clifford, Al-Lazikani) and wish to promote interactions between the projects. We have already begun discussion between Genomics England (Clare Craig), the CCLG tissue bank (Jacques) and SMPaed (Jacques, Hubank, Al-Lazikani) about how the projects can interact to both optimise using of existing samples/genomes and moving forward in the NHSE reorganisation.

Workstream 1: Collation and analysis of WGS data for patients recruited to the 100,000 Genomes project

Recruitment to the childhood cancer domain has really only taken place in earnest from the beginning of 2018. However, recruitment has now taken place in most GMCs and there are several (such as North Thames) where recruitment is taking place routinely across all clinical cases. At the time of writing, the childhood cancer domain includes 168 genomes across a range of paediatric tumour types. The first workstream will determine the frequency and nature of mutations both in the tumours and the germline in this patient cohort. These will include patients with known mutations from standard of care testing but also there are a significant number of paediatric tumours whose diagnostic classification is uncertain and pathogenic mutations are not recognised. Furthermore, the frequency of germline mutations is recognised to be high in childhood tumours but the extent and nature of these is only partially understood. Therefore, there is a considerable potential to identify novel mutations, particularly in many of the rare tumour types seen in children.

Workstream 2: Accelerated recruitment and transition to NHS provision

In order to facilitate accelerated recruitment prior to the close of the project and to obtain a larger cohort from which we can generate more meaningful data, we will contact the 21 Primary Treatment Centres for childhood tumours to identify bottlenecks to recruitment and assist in identifying archival cases suitable for inclusion. In particular, there are networks (under the paediatric ECMC and CCLG) of oncology centres which we can use to contact key individuals in these centres. In addition, we have developed a network of paediatric and neuropathologists, with representation from every childhood primary treatment centre which we can use to leverage recruitment. This will be important not only to improve case recruitment for workstream 1 but also will be important to optimise pathways for WGS for paediatric tumours following the NHSE reorganisation.

Workstream 3: Collaboration with trial profiling to determine optimal diagnostic testing

Studies of genomic profiling in paediatric tumours have mostly relied on single platforms (often capture panels and methylation profiling) and there is little data as to the optimal method for clinical diagnostic work. We can address this in this workstream both by comparing the existing data from 100,000 Genomes to standard-of-care diagnostic testing but, we can also compare diagnostic yields from the trial data to the 100,000 Genomes data. This is particularly true going forward as case submitted to SMPaed will have extensive multi-omic profiling (including centralised pathology review, methylation profiling, paediatric tumour-specific capture panel and RNAseq). In addition, the original SMPaed plan included WES and low density WGS but this would be unnecessary duplication where cases have been submitted to the 100,000 genomes project or had WGS under the NHSE reorganisation. We are currently in discussions between Genomics England and the SMPaed genomics (Hubank) and informatics leads (Al-Lazikani) to determine the best way to import and compare data from the studies. However, this provides an opportunity to

determine the relative contribution of different profiling modalities to a) determining the final diagnosis and b) identifying targetable mutations/pathways. It will also us to validate the interpretation of structure variants in the WGS data which are very important for many paediatric tumours. Finally, the SMPaed programme will be developing novel reporting algorithms for genomic data in clinical practice.

The SMPaed programme will provide an exemplar for this kind of interaction which we will then extend to other molecularly driven national studies.

Workstream 4: Genomic analysis of relapsed tumours

A major focus of the SMPaed programme is genomic profiling in relapsed patients. It is anticipated that patients who are submitted to the SMPaed programme in the coming years will include patients that have had WGS as part of the 100,000 Genomes project at presentation or previous relapse. This provides the chance to compare genomic profiles (from Genomics England) to relapsed profiles (from SMPaed and NHSE WGS) and determine the patterns of tumour evolution

Workstream 5: Assessment of shaken biopsy

A major problem for genomic analysis is access to sufficient tissue for histology and genomic diagnosis. Recent work from Louise Jones with Genomics England has indicated that fluid samples from breast core biopsies shaken in buffered saline can release sufficient tumour cells in a high proportion of cases for sequencing. We have started recruiting paediatric samples at Great Ormond Street Hospital in a similar manner. We are testing the genomic material against conventional frozen tissue and developing assays for tumour content. We will extend these shaken fluid samples across other primary treatment centres and assess the genomic data from these samples.

Collaborations including with other GeCIPs. *Outline your major planned academic, healthcare, patient and industrial collaborations. This should include collaborations and data sharing with other GeCIPs. Please attach letters of support.*

A significant collaboration will be with trial investigators, most importantly of the SMPaed programme. The SMPaed programme is a project funded by Cancer Research UK and Children with Cancer UK to offer multi-omic profiling of childhood solid tumours and includes primary and relapsed tumours, but has a particular aim of pre-screening children for targeted therapies. The investigators of the GeCIP and the SMPaed project are almost identical and are fully supportive of forming such a collaboration between the two projects:

SMPaeds Principle Investigators:
Lou Chesler, Institute of Cancer Research
Darren Hargrave, University College London
Tom Jacques, University College London
Mike Hubank, Institute of Cancer Research
Bissan Al-Lazikani, Institute of Cancer Research
Pam Kearns, Birmingham University

We have a collaboration with Prof. Louise Jones (QMUL) to assess the value of shaken biopsy samples.

We will work closely with GeCIP for head and neck tumours (Terry Jones) and sarcoma (Adrienne Flanagan) to analyse childhood tumours falling within their remits. Prof. Flanagan is part of the childhood domain and Prof. Jacques is part of the sarcoma GeCIP.

Training. *Describe the planned involvement of trainees in the research and any specific training that will form part of your plan.*

We will focus training on at least three groups: trainees in pathology, established pathologists developing molecular training, and junior scientists. Through Lou Chesler, we are training pathologists via an MSc course at ICR and many of the GeCIP have trained pathologist in our laboratories. We will involve trainee pathologists through this MSc scheme and will involve those in our labs. We also recognise the need to involve pathologists who have already qualified in order to improve their knowledge of genomic interpretation. We have begun this in the London Genomic Tumour Board (Sam Levine, Ash Merve) and will encourage involvement through the paediatric pathology group (see above). Finally, all the members of the GeCIP have post-doctoral scientists many developing bioinformatics experience whom we will involve in delivering the research project.

People and track record. *Explain why the group is well qualified to do this research, how the investigators would work together.*

The GeCIP domain has brought together the key leaders in clinical, genetic and pathological aspects of childhood tumours, including those currently delivering molecularly-driven trials in childhood tumours. For example, the GeCIP includes most of the principal investigators from many of the major current trials and molecular tumours studies e.g. PNET5 for medulloblastoma (Clifford, Jacques), SMPaed for upfront molecular profiling (Chesler, Hargrave, Hubank, Jacques, Al-Lazikani, Kearns, Clifford), INSTINCT for high risk paediatric tumours (Clifford, Williamson, Chesler, Jones, Jacques, Hargrave), the EVEREST study for low grade brain tumours (Jacques, Hargrave). In addition, we have included the leads for the children Clinical Trials Unit (Kearns), the national CCLG bank (Tweddle) and the sarcoma GeCIP (Flanagan) reflecting the important links with these organisations.

Tom Jacques is the lead for the childhood cancer domain GeCIP. He is Professor of Paediatric Neuropathology at the UCL GOS Institute of Child Health, the only such chair in the UK and provides the only specialist paediatric neuropathology service in the UK from Great Ormond Street. He is the lead pathologist on most of the key molecular studies for childhood tumours including the SMPaed study and has been successful application on over £20 millions of grant funding. He has published over 140 papers, mostly relating to molecular pathology. He is the editor in chief of Neuropathology and Applied Neurobiology and is a former chair of the national CCLG children's tumour bank. He was awarded the Cavanagh prize of the British Neuropathological Society, a prize open to international candidates, is "to recognise a young neuroscientist whose studies have made a significant contribution to the understanding of the Neuropathology of Human or Veterinary Neurological Disease".

Lou Chesler is Professor of Childhood Cancer Biology and group leader in Paediatric Tumour Biology and Therapeutics at the Institute of Child Research. He is the lead for the SMPaed study that is providing multi-omic profiling across relapsed and primary childhood tumours (funded by CR UK-£2.5 million and Children with Cancer UK £1.5 million). He has developed numerous mouse models of paediatric tumours including the first *MYCN*-driven genetically engineered and fully regulatable mouse model of the high-risk brain tumour, medulloblastoma (Genes and Dev. 2010 May 15; 24(10):1059-72) and the first genetic modelling and therapeutic targeting of ALK overexpression in paediatric cancer, leading to identification of effective ALK therapeutics (Cancer Cell, 2012, and DMM 2014). He has used these models to develop therapeutic approaches including the first approach to drug *MYC* oncoprotein stability via inhibition of MycBox binding

proteins. Two subsequent clinical trials of Alisertib were opened (Small Molecule Inhibitors of Aurora-A Induce Proteasomal Degradation of N-Myc in Childhood Neuroblastoma. Cancer Cell 2013, Cancer Cell 2015, PNAS 2016) and he defined the *in vitro* and in vivo enhancer and dependency landscape of *MYCN* activation in neuroblastoma (Nature Genetics, 2018). Subsequent work to identify drugs targeting *MYC* transcription (in review) and a clinical trial of an orally bioavailable MYC transcriptional inhibitor CYC065 (anticipated 2019).

Steve Clifford is Professor of Molecular Paediatric Oncology & Director of the Northern Institute for Cancer Research in Newcastle. Over the last 17 years, he has built an internationally-leading bench-to-bedside programme from scratch, which is now central to worldwide advances in childhood brain tumour research and treatment. Focussing on medulloblastoma, his group has led transformative research, with the vision of improving our fundamental understanding of disease biology through advanced genomics, bioinformatics and functional studies, and its clinical application in improved disease sub-classification, molecular diagnostics, novel therapeutics and clinical trials. As a result, they have pioneered biomarker discovery for the disease worldwide and are leading international clinical trials of personalised therapy based on their work. As he is the biology lead for all three SIOP-Europe trials of medulloblastoma, the UK leads the world in offering next-generation diagnostics in real-time for all medulloblastoma patients.

Darren Hargrave is Professor of Paediatric Neuro-oncology at University College London and Honorary Consultant Paediatric Oncologist at Great Ormond Street Hospital, London UK. He is a leading international clinical research leader having been the chief investigator of over 15 clinical trials from “first in child” to large randomised international phase III studies. He has led many national and international clinical research groups/ committees including being the current chair of the SIOP-Europe Brain Tumour Groups. He has also established strong links with basic and translational researchers e.g. INSTINCT, EVEREST, UCL/Edinburgh CRUK Brain Tumour Centre of Excellence, which has allowed him to develop trials for novel targeted therapies based on underlying biology (e.g. BRAF and MEK inhibitors in paediatric glioma). He has also been the joint lead (with Lou Chesler, ICR) in developing a national molecular platform SMPaed funded by CRUK to allow Paediatric “Precision” Medicine in the UK for relapsed brain and solid tumours. He has excellent links with pharmaceutical industry partners, which is vital for novel therapeutic development.

Chris Jones is Professor of Childhood Brain Tumour Biology at the Institute of Cancer Research. He was responsible for the earliest integrated description of the paediatric glioblastoma and DIPG genomes (Paugh *et al.* J Clin Oncol, 2010; Bax *et al.*, Clin Cancer Res 2010), with subsequent landmark genomic characterisation of >1000 paediatric HGG / DIPG (Mackay *et al.*, Cancer Cell 2017), detailing the extensive biological diversity of the disease and identification of novel subgroup-specific biomarkers and drug targets (Bax *et al.*, Clin Cancer Res 2009; Bielen *et al.*, 2011; Paugh *et al.*, Cancer Res 2013; Castel *et al.*, Acta Neuropath 2015; Bender *et al.*, Nature Med 2016). He identified somatic mutations in the novel cancer gene *ACVR1* in DIPG (Taylor *et al.*, Nature Genet 2014; Buczkowicz *et al.*, Nature Genet 2014; Wu *et al.*, Nature Genet 2014) and developed subsequent pre-clinical targeting with novel inhibitors in patient-derived DIPG models. He has unravelled the mechanism by which histone H3.3 G34 mutations cause cerebral hemispheric paediatric glioblastoma via upregulation of *MYCN* (Bjerke *et al.* Cancer Discov 2013), and subsequent targeting by distinct classes of inhibitors in appropriate models. He has produced detailed mapping and phenotypic correlates of the widespread genetic intratumoral heterogeneity present in glioblastoma samples (Little *et al.*, Cancer Res 2012), and the first description of functional subclonal co-operativity in paediatric HGG and DIPG (Vinci *et al.*, Nature Med 2018). He is a member of the steering group and leader for correlative biological studies of the largest randomised clinical trial in non-brainstem paediatric HGG (HERBY) (Grill *et al.*, J Clin

Oncol 2018; Mackay *et al.*, Cancer Cell 2018] and the UK arm of the largest stratified trial ever undertaken in DIPG (BIOMEDE).

Dan Williamson is Lecturer in Paediatric Neuro-Oncology, Northern Institute for Cancer Research, Newcastle University. He has taken the lead role in bioinformatics/genomics in several large research programmes including CRUK programme grants, BTC/GOSH/CwC Network programme grants, SMPaed, 100K Genomes (paediatrics) and contributed to several major biomarker/genomics discoveries. He led the first national (CCLG) biological study of Malignant Rhabdoid Tumours (MRT). His team has created a first catalogue of therapeutic targets and dependencies in this disease and contributes through international collaborations and working groups to ongoing efforts to find new therapies, biomarkers and plan upcoming clinical trials.

Mike Hubank is Head of Clinical Genomics (Research) at the Royal Marsden NHS Foundation Trust and is the lead for genomics in the SMPaed study, providing multi-omic data from childhood tumours (including RNAseq, WGS, WES) including developing a capture panel being used in the NHSE reorganisation for childhood tumours. He has developed an international reputation in genomics, having published over 100 papers in the field.

Adrienne Flanagan is lead of the sarcoma GeCIP and is professor of musculoskeletal pathology at UCL. She leads the sarcoma unit at the Royal National Orthopaedic Hospital and has a strong track record in sarcoma and bone tumour genomics being a founding partner with the ICGC bone tumour project at the Wellcome Trust Sanger Institute that has already performed whole genome and whole exome sequences of 400 tumours. The ICGC project has been successful and resulted in being able to provide more accurate diagnosis for a number of primary bone tumour types because of the identification of a number of specific genetic diagnostic biomarkers viz, diagnostic assays being developed for *IDH1/2* and *H3.3* mutations for chondrosarcomas and giant cell tumours of bone respectively.

Clinical interpretation. *(Where relevant to your GeCIP) Describe your plans to ensure patient benefit through clinical interpretation relevant to your domain. This should specifically address variant interpretation and feedback and your interaction with the cross-cutting Validation and Feedback domain.*

The London Genomic Tumour Boards are well established and are led by members of the GeCIP (Jacques, Hargrave, Hubank, Flanagan). There are genomic tumour boards for children tumours starting across the other regions and we are, and will be, working closely with colleagues at these centres. In addition, there is a national tumour board led by Prof. Hargrave and involving numerous members of the GeCIP (Jacques, Hubank, Al-Lazikani, Chesler) to support trial genomics (under SMPaed). Finally, through Bissan Al-Lazikani, we anticipate the development of novel reporting algorithms that will support genetic clinical interpretation.

Beneficiaries. *How will the research benefit patients and healthcare institutions including the NHS, other researchers in the field? Are there other likely beneficiaries?*

We hope that the main beneficiaries will be children with tumours and their families because they will be offered more accurate diagnosis, genetic counselling and accessibility to targeted therapies. In addition, due to the links we are building with existing research programmes, these studies (and the 100,000 genomes project) will have better genomic data. Finally, we will provide an evidence base for genomic testing in the NHS for children's tumours.

Commercial exploitation. *(Where relevant to your GeCIP) Genomics England has a very explicit intellectual property policy. We and other funders need to know if the proposed research likely to generate commercially exploitable results. Do you have commercial partners in place?*

We don't anticipate commercial outcomes from the direct work. Some members of the GeCIP (e.g. Hargrave) work closely with Pharma-based companies in the development of early phase trials many of which require confirmation of specific mutations prior to entry.

References. *Provide key references related to the research you set out.*

1. Capper, D., Jones, D.T.W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D.E., Kratz, A., Wefers, A.K., Huang, K., Pajtler, K.W., Schweizer, L., Stichel, D., Olar, A., Engel, N. W., Lindenberg, K., Harter, P.N., Braczynski, A., Plate, K.H., Dohmen, H., Garvalov, B.K., Coras, R., Hölsken, A., Hewer, E., Bewerunge-Hudler, M., Schick, M., Fischer, R., Beschorner, R., Schittenhelm, J., Staszewski, O., Wani, K., Varlet, P, Pages, M., Temming, P., Lohmann, D., Selt, F., Witt, H., Milde, T., Witt, O., Aronica, E., Giangaspero, F., Rushing, E., Scheurlen, W., Geisenberger, C., Rodriguez, F.J., Becker, A., Preusser, M., Haberler, C., Bjerkvig, R., Cryan, J., Farrell, M., Deckert, M., Hench, J., Frank, S., Serrano, J., Kannan, K., Tsirigos, A., Brück, W., Hofer, S., Brehmer, S., Seiz-Rosenhagen, M., Hänggi, D., Hans, V., Rozsnoki, S., Hansford, J.R., Kohlhof, P., Kristensen, B.J., Lechner, M., Lopes, B., Mawrin, C., Ketter, R., Kulozik, A., Khatib, Z., Heppner, F., Koch, A., Jouvet, A., Keohane, C., Mühleisen, H., Mueller, W., Pohl, W., Prinz, M., Benner, A., Zapatka, M., Gottardo, N.G., Hernáiz Driever, P., Kramm, C.M., Müller, H.L., Rutkowski, S., von Hoff, K., Frühwald, M.C., Gnekow, A., Fleischhack, G., Tippelt, S., Calaminus, G., Monoranu, C-M., Perry, A., **Jones, C., Jacques, T.S.**, Radlwimmer, B., Gessi, M., Pietsch, T., Schramm, J., Schackert, G., Westphal, M., Reifenberger, G., Wesseling, P., Weller, M., Collins, V.P., Blümcke, I., Bendszus, M., Debus, J., Huang, A., Jabado, N., Northcott, P.A., Paulus, W., Gajjar, A., Robinson, G., Taylor, M.D., Jaunmuktane, Z., Ryzhova, M., Platten, M.,Unterberg, A., Wick, W., Karajannis, M.A., Mittelbronn, M., Acker, T., Hartmann, C., Aldape, K., Schüller, W., Buslei, R., Lichter, P., Kool, M., Herold-Mende, C., Ellison, D.W., Hasselblatt, M., Snuderl, M., Brandner, S., Korshunov, A., von Deimling, A. and Pfister, S.M. (2018) DNA methylation-based classification of central nervous system tumours *Nature* 555(7697):469-474
2. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA *et al.* (2018) The landscape of genomic alterations across childhood cancers. Nature 555(7696):321-327
3. Waszak SM, Northcott PA, Buchhalter I, Robinson GW, Sutter C, Groebner S *et al.* (2018) Spectrum and prevalence of genetic predisposition in medulloblastoma: a retrospective genetic study and prospective validation in a clinical trial cohort. Lancet Oncol. 2018 Jun;19(6):785-798
4. Mackay, A., Burford, A,, Molinari, V,, Jones D.T.W., Izquierdo Delgado, E., Giangaspero, F., Haberler, C., Pietsch, T., **Jacques, T.S**., Figarella-Branger, D., Rodriguez, D., Morgan, P., Raman, P., Waanders, A.J., Resnick, A., Massimino, M., Garre, M.L., Smith, H., Capper, D., Pfister, S.M., Würdinger, T., Tam, R., Garcia, J., Das Thakur, M., Vassal, G., Grill, J., Jaspan, T., Varlet, P., **Jones, C**. (2018) Integrated molecular and pathological characterisation of non-brainstem paediatric high grade glioma from the HERBY phase II randomised trial *Cancer Cell* 14;33(5):829-842.e5
5. Izquierdo, E., Yuan, L., George, S., **Hubank, M., Jones, C.,** Proszek, P., Shipley, J., Gatz, S.A., Stinson, C., Moore, A.S., **Clifford, S.C.,** Hicks, D., Lindsey, J., Hill, R., **Jacques, T.S.,** Chalker, J., Thway, K., O'Conner, S., Marshall, L., Moreno, L., Pearson, A., **Chesler, L**.,

Walker, B.A., Gonzalez De Castro, D (2017). Development of a targeted sequencing approach to identify prognostic, predictive and diagnostic markers in paediatric solid tumours. *Oncotarget 8(67):112036-112050*

6. Schwalbe, E., Lindsey, J., Nakjang, S., Crosier, S., Smith, A., Hicks, D., Rafiee, G., Hill, R.M., Iliasova, A., Stone, T., Pizer, B., Michalski, A., Joshi, A., Wharton, S.B., **Jacques, T.S.,** Bailey, S., **Williamson, D., Clifford, S.C.** (2017). Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *The Lancet Oncology,* 18(7):958-971*;* doi:10.1016/S1470-2045(17)30243-7
7. Hill, R.M., Kuijper, S., Lindsey, J.C., Schwalbe, E.C., Barker, K., Boult, J.K.R., Williamson, D., Ahmad, Z., Hallsworth, A., Ryan, S.L., Poon, E., Robinson, S.P., Ruddle, R., Raynaud, F.I., Howell, L., Kwok, C., Joshi, A., Nicholson, S.L., Crosier, S., Ellison, D.W., Wharton, S.B., Robson, K., Michalski, A., Hargrave, D., **Jacques, T.S.,** Pizer, B., Bailey, S., Swartling, F.J., Weiss, W.A., **Chesler, L., Clifford, S.C.** (2014). Combined MYC and TP53 defects emerge at medulloblastoma relapse and define rapidly progressive, therapeutically targetable disease. *Cancer Cell*, 27 (1), 72-84. doi:10.1016/j.ccell.2014.11.002

Data requirements

Data scope. *Describe the groups of participants on whom you require data and the form in which you plan to analyse the data (e.g. phenotype data, filtered variant lists, VCF, BAM). Where participants fall outside the disorders within your GeCIP domain, please confirm whether you have agreement from the relevant GeCIP domain. (max 200 words)*

We will need data from children with childhood solid tumours. The most important data will be phenotype data, filtered variant lists and structure variants.

Data analysis plans. *Describe the approaches you will use for analysis. (max 300 words)*

To identify novel mutations and structural variants, we will review variant lists and structural variants across the domain looking for recurrent events. Cases will be analysed by traditional diagnostic groups, but it recognised for many of tumour types that some traditional diagnostic groups are of doubtful biological significance. Therefore, the data will be inspected for recurrent patterns of genetic change across the entire domain. In addition, tumours will also be classified according to molecularly defined subgroups. For example, most paediatric brain tumours will have had methylation analysis and classification as part of standard of care testing and all cases in the SMPaed study will have transcriptional analysis by RNAseq. Therefore, we will be any define novel subgroups of tumour type and then analyse the WGS variants and structural variants according to both traditional pathological subtypes and molecularly-defined subtypes.

As many patients will have additional molecular testing as part of standard of care (e.g. >85% of paediatric brain tumours have genetic testing as part of their diagnostic investigation, most having methylation profiling) or as part of other molecularly driven studies/trials (notably SMPaed), we are in a position to determine the relative diagnostic yield of each technique and possibly the diagnostic limitations of different approaches. In addition, we will be able to determine the yield of shaken biopsies in order to evaluate the viability of this approach for childhood tumours.

Finally, for cases with 'omic data (either form SMPaed or Genomics England) at both relapse and primary presentation, we will be able to develop a model of the genetic profile of paediatric tumours at relapse.

Key phenotype data. *Describe the key classes of phenotype data required for your proposed analyses to allow prioritisation and optimisation of collection of these. (max 200 words)*

We will need basic demographic data (e.g. age at diagnosis and biopsy) and the stage of treatment (i.e. is this primary or relapse and has there been previous treatment). We will need the pathological tumour type and any associated standard of care molecular data.

Alignment and calling requirements. *Please refer to the attached file (Bioinformatics for 100,000 genomes.pdf) for the existing Genomics England analysis pipeline and indicate whether your requirements differ providing explanation. (max 300 words)*

The existing pathway for variant calling will be suitable.

Tool requirements and import. *Describe any specific tools you require within the data centre with particular emphasis on those which are additional to those we will provide (see attached excel file List_of_Embassy_apps.xlsx of the planned standard tools). If these are new tools you must discuss these with us. (max 200 words)*

At this stage, we primarily anticipate using the tools available in the Research Embassy.

Data import. *Describe the data sets you would require within the analysis environment and may therefore need to be imported or accessible within the secure data environment. (max 200 words)*

We will wish to import data from the SMPaed study (and potentially other molecular studies and trails), which will include data from capture panels, methylation profiling and RNAseq. The investigators of SMPaed are currently in discussion with Genomics England about the practicalities of this.

Computing resource requirements. *Describe any analyses that would place high demand on computing resources and specific storage or processing implications. (max 200 words)*

At this stage, it is not anticipated that there would be extraordinary requirements for processing or storage, particularly as the numbers of samples across the cohort would be modest, but this may change once we have experience of the research embassy environment capability.

Omics samples

Analysis of omics samples. *Summarise any analyses that you are planning using omics samples taken as part of the Project. (max 300 words*

Please see comments above regarding imported data from the SMPaed study.

Data access and security	
GeCIP domain name	Childhood Cancer
Project title *(max 150 characters)*	Childhood Tumour Research in the 100,000 Genomes Project

Applicable Acceptable Uses. *Tick all those relevant to the request and ensure that the justification for selecting each acceptable use is supported in the 'Importance' section (page 3).*

X Clinical care
X Clinical trials feasibility
X Deeper phenotyping
X Education and training of health and public health professionals
X Hypothesis driven research and development in health and social care - observational
X Hypothesis driven research and development in health and social care - interventional
X Interpretation and validation of the Genomics England Knowledge Base
X Non hypothesis driven R&D - health
X Non hypothesis driven R&D - non health
X Other health use - clinical audit
X Public health purposes
X Tool evaluation and improvement

Information Governance

X The lead and sub-leads of this domain will read and signed the Information Governance Declaration form provided by Genomics England and will submit by e-mail signed copies to Genomics England alongside this research plan.

Any individual who wishes to access data under your embassy will be required to read and sign this for also. Access will only be granted to said individuals when a signed form has been processed and any other vetting processes detailed by Genomics England are completed.

Other attachments

Attach other documents in support of your application here including:

- a cover letter (optional)
- CV(s) from any new domain members which you have not already supplied (required)
- other supporting documents as relevant (optional)