Epi25 Study Protocol

Version 1.5 July 10, 2017

Table of Contents

1. Contacts

1.1. Executive Committee

Full Name	Email
Sam Berkovic	s.berkovic@unimelb.edu.au
David Goldstein	dg2875@cumc.columbia.edu
Holger Lerche	holger.lerche@uni-tuebingen.de
Dan Lowenstein	Lowenstein@medsch.ucsf.edu

1.2. Broad Institute

Full Name	Email
Felecia Cerrato	fcerrato@broadinstitute.org
Ben Neale	bneale@broadinstitute.org
Mark Daly	mjdaly@atgu.mgh.harvard.edu

1.3. Project Management

Full Name	Email
Catharine Freyer (admin)	catharine.freyer@ucsf.edu
Brigid Regan (phenotyping)	bregan@unimelb.edu.au
Susie Bellows (phenotyping)	Susannah.bellows@unimelb.edu.au

1.4. Informatics Core

Full Name	Email
Roland Krause	roland.krause@uni.lu
Kevin McKenna	kevin.mckenna@ucsf.edu
Tony Marson	A.G.Marson@liverpool.ac.uk
Sam Berkovic	s.berkovic@unimelb.edu.au
Brigid Regan	bregan@unimelb.edu.au
Susie Bellows	Susannah.bellows@unimelb.edu.au
Catharine Freyer	catharine.freyer@ucsf.edu

1.5. Strategy Committee

Full Name	Email
Sam Berkovic	s.berkovic@unimelb.edu.au
Susie Bellows	Susannah.bellows@unimelb.edu.au
Gianpiero Cavalleri	gcavalleri@rcsi.ie
Patrick Cossette	patrick.cossette@umontreal.ca
Chris Cotsaspas	cotsapas@gmail.com
Peter Dejonghe	peter.dejonghe@molgen.vib-ua.be
Chantal Depondt	Chantal.Depondt@erasme.ulb.ac.be
Tracy Dixon-Salazar	Tracy@cureepilepsy.org
Catharine Freyer	catharine.freyer@ucsf.edu
David Goldstein	dg2875@cumc.columbia.edu
Renzo Guerrini	r.guerrini@meyer.it
Hakon Hakonarson	hakonarson@email.chop.edu
Erin Heinzen	eh2682@cumc.columbia.edu
Ingo Helbig	i.helbig@pedneuro.uni-kiel.de
Roland Krause	roland.krause@uni.lu
Patrick Kwan	patrick.kwan@unimelb.edu.au
Holger Lerche	holger.lerche@uni-tuebingen.de

Dan Lowenstein	Lowenstein@medsch.ucsf.edu
Tony Marson	A.G.Marson@liverpool.ac.uk
Kevin McKenna	kevin.mckenna@ucsf.edu
Slave Petrovski	slavep@unimelb.edu.au
Brigid Regan	bregan@unimelb.edu.au
Ingrid Scheffer	i.scheffer@unimelb.edu.au
Sanjay Sisodiya	s.sisodiya@ucl.ac.uk
Randy Stewart	StewartR@ninds.nih.gov
Pasquale Striano	strianop@gmail.com
Sarah Weckhuysen	sarahweck@hotmail.com

1.6. Forms Design Committee

Full Name	Email
Sam Berkovic	s.berkovic@unimelb.edu.au
Brigid Regan	bregan@unimelb.edu.au
Susie Bellows	Susannah.bellows@unimelb.edu.au
Dennis Dlugos	DLUGOS@email.chop.edu
Catharine Freyer	catharine.freyer@ucsf.edu
Roland Krause	roland.krause@uni.lu
Dan Lowenstein	Lowenstein@medsch.ucsf.edu
Tony Marson	A.G.Marson@liverpool.ac.uk
Kevin McKenna	kevin.mckenna@ucsf.edu
Ingrid Scheffer	i.scheffer@unimelb.edu.au
Pasquale Striano	strianop@gmail.com

1.7. Phenotyping and Vetting Committee

Full Name	Email
Sam Berkovic	s.berkovic@unimelb.edu.au
Brigid Regan	bregan@unimelb.edu.au
Catharine Freyer	catharine.freyer@ucsf.edu
Dan Lowenstein	Lowenstein@medsch.ucsf.edu
Dennis Dlugos	DLUGOS@email.chop.edu
Dorothee Kastelejn	d.kasteleijn@umcutrecht.nl
Heinz Krestel	Heinz.Krestel@insel.ch
Ingrid Scheffer	i.scheffer@unimelb.edu.au
Karl Martin Klein	klein.km@staff.uni-marburg.de
Kevin McKenna	kevin.mckenna@ucsf.edu
Michael Sperling	Michael.Sperling@jefferson.edu
Orrin Devinsky	od4@nyu.edu
Pasquale Striano	strianop@gmail.com
Roland Krause	roland.krause@uni.lu
Ruth Ottman	ro6@cumc.columbia.edu
Terry O'Brien	obrientj@unimelb.edu.au
Tony Marson	A.G.Marson@liverpool.ac.uk
Warren Lo	warren.lo@nationwidechildrens.org

1.8. Cohorts Participating in Epi25

Site	Collab PI	Site ID
Australia: Melbourne	Berkovic, Scheffer	AUSAUS
Australia: Royal Melbourne	O'Brien	AUSRMB
Austria: Vienna	Zimprich	AUTMUV
Belgium: Antwerp	Weckhuysen, DeJonghe	BELATW
Belgium: Brussels	Depondt	BELULB
Canada: Andermann	Andermann	CANMNI
Canada: Andrade	Andrade	CANUTN
Canada: CENet *Data	Cossette	CANCEN
Croatia	Barisic	HRVUZG
Cyprus	Kousiappa, Papacostas	CYPCYP
Czech Republic: Prague	Sterbova	CZEMTH
Denmark	Costapas	DENCOS
Denmark: Dianalund	Steensbierre Moller	DENFDH
Finland: Helsinki	Lehejoski	FINUVH
Finland: Kalviainen	Kalviainen	FINKPH
France: Lyon	Ryvlin	FRALYU
France: Paris, Auvin	Auvin	FRARDU
France: Paris, Baulac	Baulac, Leguern	FRAPMC
GenE	Koeleman	NLDUMU
Germany: Bonn	Kunz	DEUUKB
Germany: Frankfurt/Marburg	Klein, Rosenow	DEUPUM
Germany: Giessen	Neubauer	DEUUGS
Germany: Kiel	Helbig, Muhle, von Spiczak	DEUUKL
Germany: Leipzig	Lemke, Syrbe	DEUULG
Germany: Tubingen	Lerche, Weber	DEUUTB
Ghana: Kintampo		GHAKNT
Hong Kong	Kwan	HKGHKK
Ireland: Dublin	Delanty, Cavalleri	IRLRCI
Italy: Bologna	Bisulli	ITAUBG
Italy: Catanzaro	Gambardella	ITAUMC
Italy: Florence	Guerrini	ITAUMR
Italy: Genova	Striano, Zara	ITAIGI
Italy: Milan	Canafoglia, Franschetti	ITAICB
Japan: Fukuoka	Hirose	JPNFKA
Japan: RIKEN Institute	Yamakawa	JPNRKI
Kenya: Kilifi	Newton	KENKIL
Lebanon: Beirut	Beydoun	LEBABM
Lithuania	Tumiene	LTUUHK
Macedonia	Cvetkovksa	MKDCMU
Malaysia	Haerian	MYSUMY
Netherlands	Koeleman	NTHUCT
Norway	Selmer	NOROUH

Site	Collab PI	Site ID
New Zealand	Sadleir	NZLUTO
South Africa: Agincourt		ZAFAGN
South Africa: Cape Town	Wilmshurst	ZAFCTW
Spain	Serratosa	SPAHJD
Switzerland: Geneve	Korff	CHEHDG
Switzerland: Bern	Krestel	CHEUBB
Taiwan	Tsai	TAICGM
Tanzania: Ifakara		TZAIFK
Turkey: Bogazici	Caglayan	TURBZU
Turkey: Istanbul	Bebek	TURIBU
Uganda: Iganga		UGAIGN
UK: Liverpool, Imperial	Johnson, Marson, Sills	GBRUNL
UK: UCL, London	Sisodiya, Leu	GBRUCL
USA: Baylor	Goldman	USABLC
USA: Boston	Poduri	USABCH
USA: Boston Brownstein	Brownstein	USABCB
USA: Cincinnati	Holland	USACCM
USA: Cincinnati / CAE	Glauser	USACCH
USA: Columbia (Now Duke)	Goldstein, Heinzen	USACOL
USA: Cooper/Rowan	Buono	USACRW
USA: EPGP	Lowenstein, Kuzniecky	USAEGP
USA: FEBSTAT	Shinnar	USAFEB
USA: Human Epilepsy Project	Kuzniecky, French, Lowenstein	USAHEP
USA: Latino Cohort	Kuzniecky	USANYU
USA: New York: Mount Sinai	Pinto	USAMSS
USA: NYU	Devinsky	USANYU
USA: Penn/CHOP Cohort	Helbig, Ellis	USAUPN
USA: Philadelphia	Buono, Hakonarson	USACHP
Wales: Cardiff	Thomas, Hamandi	GBRCFU
Wales: Swansea	Rees, Pickrell, Chung	GBRSWU

2. Epi25 Précis

2.1. Objectives and Mission

The **Epi25 Collaborative** is a partnership between the **Broad Institute**, **over 60 epilepsy research groups** from around the world and USA government agencies; the National Human Genome Research Institute (**NHGRI**) and the National Institute for Neurological Disorders and Stroke (**NINDS**). More than 200 collaborators are working together to centralize DNA and key clinical information from **39,000 research participants** with epilepsy. Large-scale Whole Exome and Whole Genome Sequencing will be performed to identify **genes** (the molecules in our bodies that help make us who we are) associated with **epilepsy**.

Mission: Epilepsy genetics research is at an exciting stage where it is now feasible, with the power of a large cohort, to understand the more complex genetic components of epilepsy. The driving principle behind the Epi25 Collaborative is collaboration and synergy - it is only possible to achieve this type of research by working together.

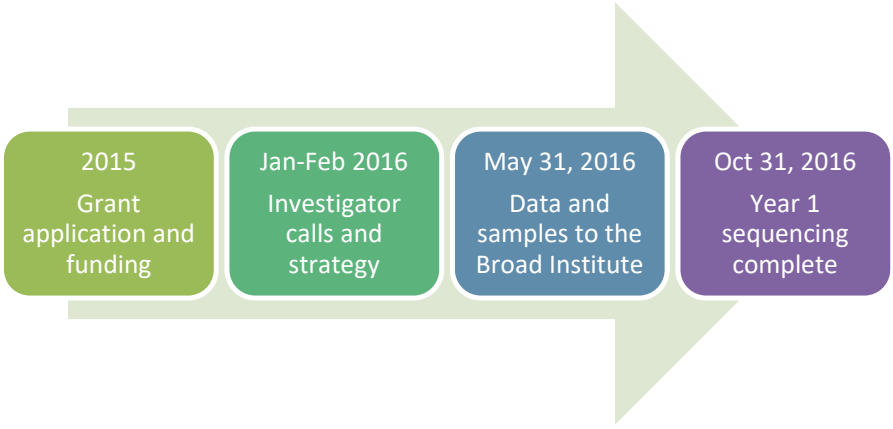
Vision: The Epi25 Collaborative is working toward a future of **precision medicine** in epilepsy in which the biological basis of each patient's epilepsy is well-understood and a targeted treatment can be identified.

2.2. Design and Outcomes

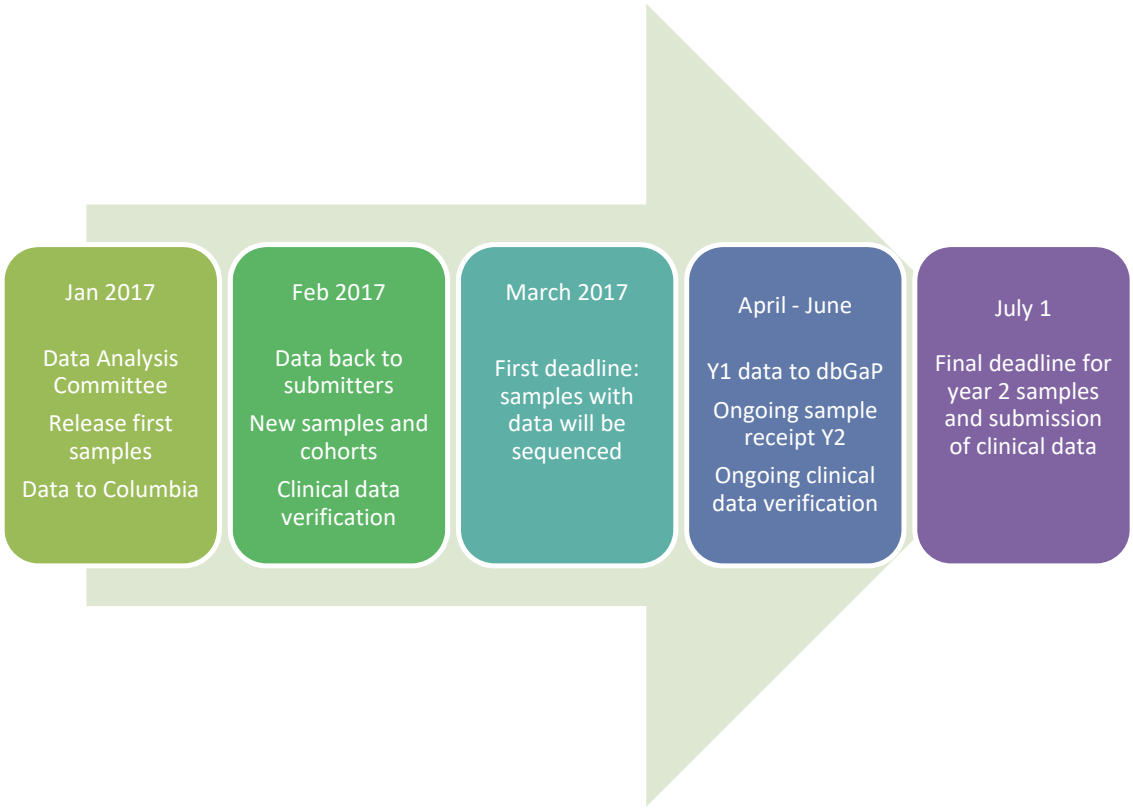
The driving principle behind the Epi25 Collaborative for Large-Scale Whole Exome and Genome Sequencing in Epilepsy (Epi25 Collaborative) is that through collaboration and synergy, collectively we will make more rapid progress towards fully understanding the inherited components of epilepsy than can be achieved individually.

The sequencing data to be collected in the centralized Epi25 Data Repository will be analyzed iteratively to identify genes and variants that explain idiopathic generalized epilepsy (IGE/GGE), non-acquired focal epilepsy (NAFE), and the epileptic encephalopathies (EEs). Summary clinical data compiled in the Epi25 clinical database will allow retrospective phenotypic analyses of groups of subjects with common genomic variant(s).

2.3. Study Timeline Year 1



2.4. Study Timeline Year 2



3. Background and Rationale for Study

3.1. Societal burden of epilepsy

Epilepsy is estimated to affect one percent of the population and 1 in 200 children (Cowan 2002). This prevalence makes it the fourth most common neurological condition following migraine, stroke, and Alzheimer's disease (World Health Organization 2012). In 2012, epilepsy was estimated to cost the United States over $15.5 billion per year [http://www.cdc.gov/chronicdisease/resources/publications/AAG/epilepsy.htm].

Epilepsy is a highly heterogeneous condition that includes many unique epilepsy syndromes and nonsyndromic subtypes. Epilepsy can be broadly divided into severe forms often with early onset seizures vs. more common, less severe forms that affect individuals of all ages. In infants and young children with severe early onset forms of epilepsy—such as the epileptic encephalopathy infantile spasms—the seizures are not controlled by medication in many patients. In addition, these young children experience developmental delay and often regression, in most cases even if the seizures are controlled medically.

3.2. Genetic architecture of the epilepsies

Genetics clearly plays an important role in the risk of epilepsy. This is evidenced by higher concordance rates in monozygotic compared to dizygotic twins (Berkovic et al 1998, Corey et al 1991) and an estimated two- to four-fold epilepsy risk in first-degree relatives of patients with recurrent seizures (Annegers et al 1982, Ottman et al 1996a, Ottman et al 1996b). Indeed, the classification system proposed by the International League Against Epilepsy now reflects the prominent role of genetics in epilepsy (Berg et al 2010). While it is clear from these observations that epilepsy is highly heritable, the genetic architecture underlying the epilepsies has not yet been fully elucidated; however, some aspects of the architecture have been illuminated through recent genetic studies of epilepsy.

1. Genome-wide association studies in epilepsy failed to identify common variants of large effect. Recently a meta-analysis of 8,696 cases and 26,157 controls identified only three genome-wide significant signals when considering either the entire cohort, or separate analyses of two cohorts of broadly defined phenotypes, genetic generalized epilepsy or non-lesional focal epilepsy. All three associated loci were found to have very low odds ratios (0.89-1.23) (International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address 2014).
2. Perhaps the most significant insight into the genetic architecture of epilepsy is the role for *de novo* mutations in sporadic forms like epileptic encephalopathies. Recently *de novo* mutations have been shown to explain at least 16% [12% explained by single nucleotide (SNV) or small insertion deletion mutations (indels) (Euro Epinomics- R. E. S. Consortium et al 2014), and 4% explained by copy number variants (Mefford et al 2011, Olson et al 2014)] of cases of epileptic encephalopathy patients.
3. When considering more common forms of epilepsy, such as non-lesional focal epilepsy, early indications from cases-control studies suggest slightly lower genetic heterogeneity in epilepsy compared to schizophrenia. In a moderately sized case-control study of 458 cases with non-lesional focal epilepsy with a family history of epilepsy and 3204 controls, three genes (*DEPDC5*, *LGI1*, and *PCDH19*), including two previously known non-lesional focal epilepsy genes (Dibbens et al 2013, Ottman et al 2004, Scheffer et al 2014), were significantly enriched for rare, likely protein altering mutations.

Collectively, these and other results (see references) suggest a distinctive genetic architecture of epilepsy that is less influenced by common variation than other complex diseases. Based on the epilepsy genes identified to date, rare mutations are significant

contributors to epilepsy risk. And finally, there are early indications suggest that genetic heterogeneity may be slightly reduced in epilepsy compared to other neuropsychiatric diseases. Given these observations, extremely large genome-wide sequencing studies will likely unlock much of the missing heritability associated with the epilepsies.

3.3. Role of high quality phenotyping

Detailed phenotyping and determination of epilepsy syndrome diagnoses are essential to the design of a genetic study to reduce genetic heterogeneity to the best of our current knowledge and to allow for identification of a well-matched replication cohort when association signals are detected. Furthermore, detailed phenotypic characterization is highly useful following gene discovery in defining and clarifying the phenotypic spectrum associated with a gene. This information is critical to clinical translation of genetic discovery.

3.4. Translating genetic discovery to treatments in epilepsy

Given the recent explosion of gene discoveries in epilepsy, the research community is fully aware of the growing need to rapidly translate discoveries into mechanisms and treatments. There are multiple concentrated efforts to develop functional models that can be used to screen mutations for their likelihood to cause epilepsy, and detailed investigations of epilepsy genes to try to understand how mutations give rise to disease.

4. Epi25 Organization

The *Collaborative* consists of four parts, as described in the Epi25 charter, signed by all members of the *Collaborative*. The full Epi25 charter is embedded here.

Epi25 Charter Draft 3.6.15.doc

4.1. Strategy Committee

The group within which the strategic decisions of the *Collaborative* will be made, as well as oversight of Epi25 operations. The full charter (embedded above) describes in great detail the Strategy Committee, and highlights are included here:

- All Strategy Committee investigators are **equal contributors** and are encouraged to participate in the regular teleconference meetings. Failing to actively participate in the majority of meeting and teleconferences may result in expulsion from the Strategy Committee.
- All Strategy Committee members agree that they will be held to a high degree of **credibility and reliability** within the Collaborative and acknowledge that they will be proactive in declaring conflicts of interests, entanglement in competing projects, or commercial interests. The Strategy Committee members acknowledge that their responsibility for the Collaborative extends beyond the sphere of their own group, institution, or national funding framework and they will be actively working towards transparent communication within the Collaborative.
- The composition of the Strategy Committee and other working groups will reflect the **international nature** of the Epi25 initiative and will aim to include representatives from Europe, Asia, Oceania, Africa and the Americas.
- Meets ~ monthly or bi-monthly by telephone and occasionally in person.
- Reaches decisions by **informal consensus** of those participating in calls, not by formal vote.
- **Term limit** of three years, renewable.
- Disputes or impasses are settled through **consensus**. None of the members of the Strategy Committee have a formal "veto", but in case of disagreement are expected to respect and represent the consensus opinion.
- Led by a **Chair** who is responsible for effective administration of the organizational activities of the *Collaborative*, including communication internally and externally, scheduling of meetings, documentation of *Collaborative* deliberations, decisions, and other related activities. The Chair will be a member of the Executive Committee and will have a term of three years, renewable.
- **Subcommittees** of the Strategy Committee may include:
    - Phenotype Subcommittee
    - Data Analysis Subcommittee
    - Publication Subcommittee
    - Informatics Subcommittee

4.2. Executive Committee

- A small group (4 members) that is responsible for day-to-day, operational matters of Epi25 and serves as the point of contact and primary source for Epi25-related communications.
- The composition of the Executive Committee will aim to maximize the opportunity for effective interactions with NIH and other funding agencies, the ILAE, and other relevant entities.
- Meets monthly by telephone and occasionally in person.
- Reaches decisions by informal consensus of those participating in calls, not by formal vote.

- Term limit of three years, renewable.
- A member of the Executive Committee will serve as the Chair of the Strategy Committee

4.3. Working Groups

- Temporary *ad hoc* groups established by the Strategy Committee to pursue specific and time limited projects that are better handled by smaller, specialist groups.
- A chair is appointed for each Working Group.
- Working Groups report progress and findings to the Strategy Committee.
- Working Groups meet by teleconference according to a timetable established by each Working Group chair.

4.4. Regular Members

- A large group with no strictly defined membership criteria, not only limited to principal investigators, but all individuals providing intellectual contribution to Epi25 activities.
- Regular Members can be added upon request and need to be confirmed by the Strategy Committee, who will evaluate whether the intended contribution of a new member will add to the aims of the *Collaborative*.
- New Collaborative members will sign the **Epi25 Charter** document indicating that they agree with the policies listed here and that they have obtained the required ethical clearances to participate or send data.
- Regular Members are added to the email listserv.
- Regular Members can listen in and participate in teleconference calls, and can join Working Groups, if approved by their Strategy Committee representative.

4.5. Junior Investigators

All Principal Investigators within the Collaborative are encouraged to include their colleagues, including junior MDs, PhD students and postdoctoral fellows, within the Epi25 Collaborative. The amount of members per site should be fair and balanced and the contribution of each partner should be evident to the Collaborative and indicated upon request. For additional members added after the initial conception of Epi25, the Strategy Committee will decide whether an applicant is admitted to Epi25 and will communicate this decision in a transparent manner.

5. Inclusion and Exclusion Criteria

Sites are encouraged to seek guidance on eligibility of participants from the Phenotyping Core by emailing Brigid Regan bregan@unimelb.edu.au

5.1. Subjects with IGE/GGE

Include:

- Convincing history of generalized seizure types (GTCS, Absence, Myoclonus) and generalized epileptiform on EEG
- Normal neuroimaging if performed (not required)

Exclude:

- History of focal seizures
- Moderate to severe intellectual disability

Priority given to:

- Those with phenotyping data uploaded/entered
- Lowest priority given to those with Epilepsy with GTCS Alone (https://www.epilepsydiagnosis.org/syndrome/egtcsa-overview.html) as the diagnosis is less certain

5.2. Subjects with NAFE

Include:

- Convincing history of focal seizure types
- Focal epileptiform or normal EEG
- Neuroimaging is normal or shows hippocampal sclerosis (MRI preferred but CT is accepted)

Exclude:

- History of generalized seizures
- Moderate to severe intellectual disability

Priority given to:

- Those with phenotyping data uploaded/entered
- First or second degree relative with history of non-acquired epilepsy (febrile seizures excluded)

5.3. Subjects with Lesional Focal Epilepsy

Include:

- Definitive focal lesion on MRI
- Seizure types and other clinical features consistent with lesion type and location
- Lack of progression of tumor on brain imaging required for benign tumors
- Lesional cases with seizures to be included:
    - Malformation of cortical development:
        - Focal cortical dysplasia
        - Lissencephaly
        - Subcortical band heterotopia
        - Grey matter heterotopia
        - Polymicrogyria
        - Hypothalamic Hamartoma
        - Hemimegalencephaly
        - Schizencephaly
    - Traumatic brain injury
    - Stroke
    - Hypoxic ischemic stroke
    - Benign tumor (DNET, ganglioglioma)
    - Cerebral angioma (cavernoma)

Exclude:

- Tuberous sclerosis

5.4. Subjects with Epileptic Encephalopathies

Include:

- Severe refractory epilepsy of unknown aetiology with developmental plateauing or regression
- Normal MRI
- Epileptiform on EEG (although it is acknowledged that subjects enrolled earlier on in disease may not have this)

Priority given to:

- Those with phenotyping data uploaded/entered
- Those with a negative epilepsy gene panel already completed
- Those with parental DNA available

6. Subject Phenotyping and Study Design

Phenotypic data describing the participants' medical history, seizure types and syndrome are collected either via entry in an online case report form, or via web upload from partner epilepsy cohorts.

The data dictionaries are embedded here:


Epi25 GGE Glossary_v1.0.0.docx


Epi25 EE Glossary_v1.0.0.docx


Epi25 Focal Glossary_v1.0.0.docx

7. Epi25 Phenotyping Data Repository

The Epi25 Data Repository will serve as the central location for all non-PHI data collected as part of the Epi25 study. This phenotype data will be entered:

- For existing databases, phenotype data should be merged directly into the Epi25 database.
- For samples not in databases that can be merged with the Epi25 database, phenotype data must be entered directly into Redcap.

7.1. Epi25 Data Uploads

Users merging their data directly into Epi25 from their database should use the dictionary and template embedded here.

- Larger collections should import their data **directly** into the Epi25 database.
- Fields should be mapped to the tabbed Excel file "Study to Epi25 20160307.xslx" embedded below.
- The Epi25 data dictionary is embedded below, "epi25 dd 20160307.xslx"
- Please send any fields mapping questions to Roland Krause at roland.krause@uni.lu.


epi25_dd_20170314.xlsx

7.2. Epi25 Red Cap Entry and Access

Manual, Redcap forms should be used for new data input and samples **not in a database** that can be merged.

Users requiring RedCap access should use the form embedded here to request user IDs . This form should be emailed to Roland Krause roland.krause@uni.lu . Please always use institutional e-mail addresses and obtain a login for every user accessing the database.

Epi25 Red Cap UserImportTemplate.c

Users requiring RedCap should use the instructions embedded here:

Epi25 REDCap instructions_v1.0.0.do

7.3. Epi25 phenotyping data review

Verified phenotyping data is necessary to maximise the potential for meaningful discoveries and allow more specific sub-phenotypic genomic analyses to occur. Samples sequenced in year 1 (n=6,242) were sequenced without verification of clinical data. This is occurring retrospectively to aid genomic analyses. Samples to be sequenced in year 2 now require verification of their clinical data before they are eligible for sequencing.

7.3.1. Source of data as measure of quality

Manual data entry provides the highest quality of phenotyping data but, due to limited resources and time, this is not possible for all sites. There are three levels of phenotyping data that exist within the Epi25 phenotyping dataset and these are graded based on the source of information.

- GRADE I: Sub-phenotype provided to the Broad Institute with the sample, no additional clinical information exists in Epi25 phenotyping database.
- GRADE II: Phenotyping information exists in the Epi25 database and was imported from an external clinical database.
- GRADE III:Phenotyping information exists in the Epi25 database and was provided manually by completion of REDCap phenotyping forms.

7.3.2. Review of inclusion/exclusion criteria (GRADE II & GRADE III)

As illustrated in section 5 of this protocol, three sub-phenotypic groups are included in Epi25 and each group has a specific set of inclusion/exclusion criteria. The data quality review process involves checking the Epi25 phenotyping dataset to confirm that the cases included in the study meet these criteria.

Specific fields within each sub-phenotype are assessed to verify each inclusion/exclusion criteria. Depending on the data entered in that field, the case will either be **passed**, **require manual review** or, if there is **missing data** in that field, require follow up with the site to request the information be provided. Note that to pass overall, a case must pass each specific field that is assessed.

Responses that require manual review are first checked by the phenotyping coordinator and then by a clinician if further input is required. The manual review process requires reviewing all the available information about a case and determining if there is enough evidence to support their assignment to a specific sub-phenotype. If the data is concerning or further clarification is needed, the phenotyping coordinator will reach out to the site PI.

Genetic Generalized Epilepsy (GGE)

Inclusion/exclusion criterion	Specific field(s) assessed	Responses that indicate PASS	Responses that require MANUAL REVIEW	Responses that indicate MISSING DATA
Convincing history of generalized seizure types (GTCS, Absence, Myoclonus)	- GGE Syndrome - Epilepsy syndrome comments	- Childhood Absence Epilepsy - Epilepsy with Myoclonic Absences - Childhood Absence Epilepsy/Juvenile Absence Epilepsy overlap - Juvenile Absence Epilepsy - Juvenile Myoclonic Epilepsy - Epilepsy with Generalized Tonic-Clonic Seizures Alone - Other: Late Onset GGE - Other: Early Onset Absence Epilepsy - Other: Epilepsy with Eyelid Myoclonia	*Any of the below responses to* Syndrome *AND text entered in* Epilepsy syndrome comments: - Other, please specify[1] - Generalized unspecified[2]	- Other, please specify *AND no text entered in* Epilepsy syndrome comments - *Field is blank/not completed*
Generalized epileptiform on EEG	- EEG Finding 1 - EEG Finding 2 - EEG Finding 3	*PASS requires any of the below responses in at least one of the 3 EEG Finding fields AND no MANUAL REVIEW responses entered:* - Generalized spike and wave, specify frequency - Generalized polyspike and wave - Generalized epileptiform unspecified - Photo-paroxysmal response	*Any of the below responses would initiate a manual review, even if a PASS response is entered in another EEG Finding field:* - Generalized paroxysmal fast activity (GPFA)[3] - Epileptiform unspecified[4] - Focal or multi-focal epileptiform, specify location[5] - Other, please specify[6]	*The below responses are only relevant if no PASS response is entered in any other EEG Finding field:* - Normal - Focal slowing - Generalized slowing - Unknown - Not done - *Field is blank/not completed*
Normal neuroimaging if performed (not required)	- Neuroimaging Findings	- Normal - Unknown - *Field is blank/not completed*	- Abnormal[7] - Non-specific abnormality, please specify[8] - Other, please specify[9]	N/A as data not required
Exclude: moderate to severe intellectual disability	- Intellectual Disability - Degree of intellectual disability	*PASS requires Intellectual Disability field response:* - No - Unknown *OR Intellectual Disability field response:* - Yes *AND Degree of intellectual disability field response:* - Mild	*Any of the below responses to the Degree of intellectual disability field would initiate a manual review:* - Moderate - Severe[10] - Profound - Cannot classify	N/A as data not required

Non-Acquired Focal Epilepsy (NAFE)

Inclusion/exclusion criterion	Specific field(s) assessed	Responses that indicate PASS	Responses that require MANUAL REVIEW	Responses that indicate MISSING DATA
Convincing history of focal seizure types	- Focal syndromes - Epilepsy Syndrome comments	*PASS requires one of the following* Focal syndromes *AND no text entered in* Epilepsy Syndrome comments: - "Benign" Childhood Epilepsies: Childhood Epilepsy with Centrotemporal Spikes - "Benign" Childhood Epilepsies: Atypical Childhood Epilepsy with Centrotemporal Spikes - "Benign" Childhood Epilepsies: Benign Occipital Epilepsy (Panayiotopoulos) - "Benign" Childhood Epilepsies: Benign Occipital Epilepsy (Gastaut) - "Benign" Childhood Epilepsies: Idiopathic photosensitive occipital lobe epilepsy - Other Non-lesional Focal Epilepsies: Frontal - Other Non-lesional Focal Epilepsies: Frontotemporal - Other Non-lesional Focal Epilepsies: Temporal - Other Non-lesional Focal Epilepsies: Occipital - Other Non-lesional Focal Epilepsies: Temporoccipital - Other Non-lesional Focal Epilepsies: Parietal - Other Non-lesional Focal Epilepsies: Multifocal - Other Non-lesional Focal Epilepsies: Unspecified	*Any text provided in:* - Epilepsy Syndrome comments[1]	- *Field is blank/not completed*
Focal epileptiform or normal EEG	- EEG Finding 1 - EEG Finding 2 - EEG Finding 3	*PASS requires any of the below responses in at least one of the 3 EEG Finding fields AND no MANUAL REVIEW responses entered:* - Normal - Epileptiform unspecified - Focal or multi-focal epileptiform, specify location - Focal slowing - Generalized slowing	*Any of the below responses would initiate a manual review, even if a PASS response is entered in another EEG Finding field:* - Generalized spike and wave, specify frequency[2] - Generalized polyspike and wave[3] - Generalized epileptiform unspecified[4] - Photo-paroxysmal response[5]	*The below responses are only relevant if no PASS response is entered in any other EEG Finding field:* - Unknown - Not done - *Field is blank/not completed*
Neuroimaging is normal or shows hippocampal sclerosis	- Neuroimaging Findings	- Normal - Hippocampal sclerosis	- Abnormal[6] - Non-specific abnormality, please specify[7] - Other, please specify	- Unknown - *Field is blank/not completed*
Exclude: moderate to severe intellectual disability	- Intellectual Disability - Degree of intellectual disability	*PASS requires Intellectual Disability field response:* - No - Unknown *OR Intellectual Disability field response:* - Yes *AND Degree of intellectual disability field response:* - Mild	*Any of the below responses to the Degree of intellectual disability field would initiate a manual review:* - Moderate - Severe - Profound - Cannot classify	N/A as data not required

Epileptic Encephalopathy (EE)

Inclusion/exclusion criterion	Specific field(s) assessed	Responses that indicate PASS	Responses that require MANUAL REVIEW	Responses that indicate MISSING DATA
Severe refractory epilepsy	- Syndrome - Epilepsy syndrome comments	*PASS requires one of the following* Syndrome *AND no text entered in* Epilepsy syndrome comments: - Neonatal onset: Ohtahara syndrome - Neonatal onset: Early myoclonic encephalopathy - Early onset epileptic encephalopathy - Infantile onset epileptic encephalopathy (not otherwise specified) - Epilepsy of infancy with migrating focal seizures - West syndrome/infantile spasms - Late-onset epileptic spasms - Lennox-Gastaut syndrome - Epilepsy with myoclonic-atonic seizures - Dravet syndrome - Landau-Kleffner syndrome (LKS) - Epileptic encephalopathy with continuous spike-and-wave during sleep (CSWS) - Febrile Infected Related Epilepsy Syndrome (FIRES) - Hemiconvulsion-Hemiplegia-Epilepsy - Nonsyndromic epileptic encephalopathy with focal seizures - Nonsyndromic epileptic encephalopathy with generalized seizures - Nonsyndromic epileptic encephalopathy with mixed or unclassified seizures	*Any text provided in:* - Epilepsy Syndrome comments[1]	- *Field is blank/not completed*
…of unknown aetiology…	- Head trauma with skull fracture, intracranial bleeding - CNS infection	- No - Unknown - *Field is blank/not completed*	- Yes[2]	N/A as data not required
	- Normal neonatal period (other than seizures)	- Yes - Unknown - *Field is blank/not completed*	- No[3]	N/A as data not required
	- Conventional karyotype - Copy number analysis - Gene panel results - Individual gene testing - Metabolic testing	- Normal - Unknown - Not done - *Field is blank/not completed*	- Abnormal, please specify[4] - Finding of unknown significance, please specify[5]	N/A as data not required
…with developmental plateauing or regression	- Developmental delay prior to seizure onset - Regression/plateau - Intellectual disability	*PASS requires* Yes *to at least one of these fields*	No *entered for each of these fields*[6]	Unknown *or field is blank/not completed for each of these fields*

Inclusion/exclusion criterion	Specific field(s) assessed	Responses that indicate PASS	Responses that require MANUAL REVIEW	Responses that indicate MISSING DATA
Epileptiform on EEG (although it is acknowledged that subjects enrolled earlier on in disease may not have this)	- EEG Finding 1 - EEG Finding 2 - EEG Finding 3	*PASS requires any of the below responses in at least one of the 3 EEG Finding fields* - Normal - Burst suppression - Classic hypsarrhythmia - Hypsarrhythmia variant - Generalized spike and wave, specify frequency - Generalized polyspike and wave - Generalized paroxysmal fast activity (GPFA) - Continuous Spike and Wave in slow-wave Sleep (CSWS) - Generalized epileptiform unspecified - Epileptiform unspecified - Focal or multi-focal epileptiform, specify location - Focal slowing - Generalized slowing - Photo-paroxysmal response Other, please specify	N/A – any EEG result allowed	*The below responses are only relevant if no PASS response is entered in any other EEG Finding field:* - Unknown - Not done *Field is blank/not completed*
Normal MRI	- Neuroimaging Findings - Additional Neuroimaging abnormality 1 - Additional Neuroimaging abnormality 2	*PASS requires any of the below responses in at least one of the 3 Neuroimaging fields AND no MANUAL REVIEW responses entered:* - Normal - Other: Hippocampal Sclerosis - Other: Atrophy	*Any of the below responses would initiate a manual review, even if a PASS response is entered in another Neuroimaging Finding field:* - Malformations: Focal Cortical Dysplasia[7] - Malformations: Heterotopia - Malformations: Peri-ventricular nodular heterotopia - Malformations: Polymicrogyria - Malformations: Pachygyria - Malformations: Hemimegalencephaly - Malformations: Schizencephaly - Malformations: Lissencephaly - Malformations: Double Cortex - Malformations: Holoprosencephaly - Malformations: Corpus callosum agenesis/dysplasia - Malformations: Septo-optic dysplasia - Malformations: other - Vascular and/or ischemic abnormalities: hypoxic ischemic injury - Vascular and/or ischemic abnormalities: Periventricular leukomalacia - Vascular and/or ischemic abnormalities: haemorrhage - Other: Porencephaly - Other: Hydrocephalus - Other, please specify - Non-specific abnormality, please specify Other, please specify	- Unknown - *Field is blank/not completed*

7.3.3. Manual review examples

GGE

[1] Syndrome = 'Other, please specify', Syndrome comments = 'Uncertain diagnosis of epilepsy'
OUTCOME = contact site to confirm eligibility as GGE

[2] Syndrome = 'Generalized unspecified', Syndrome comments = 'Single febrile seizure'
OUTCOME = contact site to confirm eligibility as GGE

[3] EEG Finding = 'Generalized paroxysmal fast activity', data manually reviewed and patient has 'Generalized epilepsy with spasms and tonic seizures'
OUTCOME = contact site to suggest transfer to EE

[4] EEG Finding = 'Epileptiform unspecified', data manually reviewed and patient has GTCS from 15 years and myoclonus from 16 years, 'Juvenile myoclonic epilepsy' entered as syndrome
OUTCOME = pass on manual review

[5] EEG Finding = 'Focal or multi-focal epilepsy, specify location', data manually reviewed and patient has GTCS from 12 years and myoclonic jerks of unknown onset, 'Juvenile myoclonic epilepsy' entered as syndrome
OUTCOME = contact site to confirm if generalized epileptiform ever captured on EEG

[6] EEG Finding = 'Other, please specify', with comment 'Single spike of right frontopolar origin was documented in this study correlating with the history of contralateral clinical manifestations', data manually reviewed and patient has GTCS and absences with seizure onset 2 years, syndrome diagnosis is 'Juvenile Absence Epilepsy'
OUTCOME = contact site to confirm if generalized epileptiform ever captured on EEG or story more consistent with focal epilepsy

[7] Neuroimaging Finding = 'Abnormal', with comment 'MRI small lesion next to the left insular cortex probable DNET', data manually reviewed and no seizure types listed, syndrome is 'Generalized unspecified' and EEG Finding 'Epileptiform unspecified'
OUTCOME = contact site to clarify relevance of lesion and confirm if enough evidence to classify as GGE

[8] Neuroimaging Finding = 'Non-specific abnormality, please specify' with comment 'small encephalocele on the right temporal lobe', data manually reviewed and patient has classical febrile seizures and GTCS and absences from 17 years, syndrome is 'Generalized unspecified' and EEG Finding is 'Generalized epileptiform unspecified'
OUTCOME = pass on manual review (SFB)

[9] Neuroimaging Finding = 'Other, please specify' with comment 'there is quite extensive ischaemic damage', data manually reviewed and patient has GTCS and absence of unknown onset, syndrome is 'Generalized unspecified' and EEG Finding is 'Generalized spike and wave'
OUTCOME = contact site to confirm the ischaemic damage is not relevant to clinical presentation

[10] Degree of intellectual disability = 'Severe', data manually reviewed and patient has absences, atonic with epilepsy onset of 4 years, neurological examination notes 'Marked degree of mental retardation, some degree of microcephaly, a right esotropia', syndrome is 'Childhood Absence Epilepsy' and EEG Finding is 'Generalized polyspike and wave'
OUTCOME = contact site to suggest transfer to EE

NAFE

[1] Syndrome comments = 'as a result of limbic encephalitis', data manually reviewed and patient has focal dyscognitive seizures and 'secondary generalized tonic-clonic seizures', syndrome is 'Other non-lesional focal epilepsies: Temporal', EEG Finding and Neuroimaging Finding are not provided.
OUTCOME = contact site to confirm eligibility as non-acquired focal epilepsy, clarify certainty of limbic encephalitis and request missing data

[2] EEG Finding = 'Generalized spike and wave, specify frequency', data manually reviewed and patient has focal dyscognitive seizures and 'single tonic clonic', syndrome is 'Other non-lesional focal epilepsies: Unspecified', Neuroimaging Finding is not provided.
OUTCOME = contact site to confirm seizure semiology is consistent with focal epilepsy and question whether patient should be transferred to GGE, request missing data

[3] EEG Finding = 'Generalized polyspike and wave', data manually reviewed and patient has aura from 18 years and bilateral convulsive seizures, syndrome is 'Other non-lesional focal epilepsies: Temporal', MRI is normal.
OUTCOME = contact site to confirm the seizure semiology is convincing for focal epilepsy and they are comfortable with the patient belonging to NAFE despite the generalized epileptiform

[4] EEG Finding 'Generalized epileptiform unspecified', data manually reviewed and patient has aura, focal dyscognitive and bilateral convulsive seizures, epilepsy onset is 19 years, syndrome is 'Other non-lesional focal epilepsies: Unspecified' with comment 'brief aura (vertigo-like), she can call, loss of contact, she tends to fall (possibly due to lapse of tone at lower limbs). Seizures partially responded to topiramate', Neuroimaging is normal.
OUTCOME = pass on manual review SFB

[5] EEG Finding 'Photo-paroxysmal response', type of PPR is 'Generalized', data manually reviewed and patient has '"Benign" Childhood Epilepsies: Idiopathic photosensitive occipital lobe epilepsy'
OUTCOME = pass on manual review

[6] Neuroimaging Finding 'Abnormal', with comment 'ventricular asymmetry'
OUTCOME = pass on manual review

[7] Neuroimaging Finding 'Non-specific abnormality, please specify', with comment 'extensive haemmorhagic contusion – head injury', data manually reviewed and patient has aura, focal dyscognitive and bilateral convulsive seizures from 20 years, epilepsy syndrome is 'Other non-lesional focal epilepsies: Parietal' with comment 'Localization-related – Symptomatic – Other – Multifocal', EEG Finding is 'Epileptiform unspecified'
OUTCOME = contact site to confirm if head injury preceded seizure onset and clarify its relevance

EE

[1] Epilepsy syndrome comments = 'Aicardi syndrome', data manually reviewed and patient has infantile spasms from 1 month and focal seizures from 3 months, unknown if DD prior to seizure onset, no regression noted, profound ID, syndrome is 'West syndrome/ Infantile spasms', EEG Finding is 'hypsarrhythmia variant' and Neuroimaging Finding is 'Malformations: corpus callosum agenesis/dysplasia'
OUTCOME = contact site to confirm patient should be removed from analysis based on MRI result

[2] CNS infection = 'yes', with comment 'Meningitis', data manually reviewed and patient has atypical absence from 1 month, unknown if DD prior to seizure onset, no regression noted, ID of unknown degree, syndrome is 'Early onset epileptic encephalopathy (< 3 months)', EEG Finding is 'normal' and no Neuroimaging Finding is provided
OUTCOME = contact site to clarify history of meningitis – was pathogen isolated? Did infection correlate with seizure onset? Are there any Neuroimaging abnormalities?

[3] Normal neonatal period, apart from seizures = 'no', with comment 'gastroesophageal reflux'
OUTCOME = pass on manual review

[4] Copy number analysis = 'abnormal', with comment '5q12.3 deletion de novo'
OUTCOME = contact site to request more information about pathogenicity of deletion (size, gene content etc.) and confirm if the CNV is thought to be causative

[5] Individual gene testing ' finding of unknown significance, please specify', with comment 'CDKL5 c.145+17A>G mutation'
OUTCOME = contact site to request more information about pathogenicity of CDKL5 variant (segregation, software predictions, frequency in controls etc.) and confirm if the variant is thought to be causative

[6] Developmental delay prior to seizure onset = 'no' and regression/plateau = 'no' and intellectual disability = 'no', data manually reviewed and patient had neonatal seizures, failure to thrive, hypotonia, triangular facies, micrognathia, dimple on chin, shrunken fontanelle overlapped cranial, hyponatremia, hypokalaemia, hypercarbia, low total and free carnitine, tonic and myoclonic seizures from 2 weeks, syndrome is 'Early onset epileptic encephalopathy (< 3 months)', EEG Finding 'Burst suppression', Neuroimaging is 'normal'
OUTCOME = contact site to confirm no evidence of an encephalopathy such as developmental delay or regression

[7] Neuroimaging findings = 'Malformations: Focal Cortical Dysplasia', with comment 'R temporal FCD'
OUTCOME = contact site to confirm presence of malformation and that patient should be removed from analysis

8. DNA Sample Requirements for Broad Institute

Cohorts should not send any PHI to the Broad or Epi25.

8.1. Cohort IDs

All cohorts will create an Epi25 ID by adding their Epi25 site identifier to the front of their local ID.

Cohort IDs are available in section 1 of this protocol, Cohorts Participating in Epi25 (click link to jump).

As an example, if participant 00001 were sent from Austin Hospital (code AUSAUS), the new Epi25 ID will be AUSAUS00001.

8.2. IRB Approval and Consent Requirements

Prior to sending DNA to the Broad Institute for any processing, copies of informed consent(s) used for sample collection, along with notification of IRB approval, should be sent to the appropriate Project Manager at the Broad Institute.

For samples collected after January 25, 2015, consent forms should include language per the NIH Genomic Data Sharing policy (http://gds.nih.gov/03policy2.html). In general, the GDS policy requires that consent forms - used to collect samples after January 25, 2015 - reference future use, broad sharing, and whether individual-level data will be shared through an unrestricted or controlled access repository (such as dbGaP). Please note that since Epi25 data will be deposited into dbGaP, a data use limitation letter will also be required. The Broad Project Manager will follow up with each cohort to obtain this letter.

8.3. Sample Shipment

Once IRB documents have been reviewed and approved, sample shipment can be coordinated. **All DNA samples being shipped to the Broad institute's Genomics Platform should be sent in barcoded Matrix Kits**. Matrix Kits will be requested by a Broad Project Manager to be shipped to the external Institution. Upon receipt, DNA samples should be transferred to the matrix kits and **returned by courier with dry ice to the Broad**.

Samples will be sent back to Broad using the following address:

Broad Institute
Attn: Genomics Platform – Samples Lab
320 Charles St – Lab 181
Cambridge, MA 02141
Phone: (617) 714-8952

8.4. Spreadsheet Instructions

Matrix Kits will be accompanied by an excel spreadsheet, sent via email, reflecting the bar codes and layout of your DNA kit. Once a submitter aliquots the DNA samples into the tubes and assign each sample a position in the container, the submitter must fill out the sample data (de-identified IDs and phenotypes) in the corresponding kit spreadsheets.

It is important that this information is filled out correctly so that the data is correct in Broad Institute databases. Completed spreadsheets should be emailed back to the Broad Project Manager at the time of sample shipment.

Required Sample Information for DNA Sample Shipment:

- ***"Sample Info" Tab***
  - Collaborator Participant ID
  - Collaborator Sample ID
  - Gender
  - Volume
  - Concentration
  - Collected After 1/25/15 (Yes/No)

- ***"Participant Phenotypes" Tab***
  - Primary Disease (e.g. Epilepsy, Control)
  - Race
  - Ethnicity
  - If applicable for family/trio samples, please add columns for Family ID, Maternal ID, and Paternal ID, and Family Relationship

The Epi25 ID should be entered as both the "Collaborator Participant ID" and the "Collaborator Sample ID."

Please refer to the list of identifiers that should NOT be provided with any sample before completing spreadsheets.

IDENTIFYING INFORMATION
(please do not include any of this information when returning spreadsheets)

- Names
- All geographic subdivisions smaller than a state, including street address, city, county, precinct, ZIP Code, and their equivalent geographical codes
- All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated into a single category of age 90 or older.
- Telephone numbers
- Facsimile numbers
- Electronic mail addresses
- Social security numbers
- Medical record numbers
- Health plan beneficiary numbers
- Account numbers
- Certificate/license numbers
- Vehicle identifiers and serial numbers, including license plate numbers
- Device identifiers and serial numbers
- Web universal resource locators (URLs)
- Internet protocol (IP) address numbers
- Biometric identifiers, including fingerprints and voiceprints
- Full-face photographic images and any comparable images
- Any other unique identifying number, characteristic, or code, unless otherwise permitted by the Privacy Rule for re-identification

8.5. DNA Requirements

The optimal volume and concentration to send to the Broad Institute for sequencing is 50ul at 80ng/ul per sample. More DNA is always acceptable. This will permit Broad samples and sequencing platforms to quantify, qualify and process the samples, with enough DNA left over to run a redo if needed. Sending this amount of DNA reduces the time and effort needed to process samples, especially when the need of a rerun arises. In instances where optimal volume and concentration is not available, lower inputs may be acceptable on a per case basis.

8.6. Broad Institute Data Storage and Security

The Broad Institute has substantial resources for data storage (over 10 petabytes) and databases. Sample data provided to the Stanley Center generated by the Broad's Genomics Platform is accessible to only the Stanley Center's project management and genetic analysts as well as the Broad's Genomics Platform. Data will be provided to those groups providing samples via a password-protected Genomics Platform data portal or a secured server. Alternative data distribution can be discussed on a case by case basis.

9. Access to Epi25 Data

Epi25 data will be returned to the Epi25 Consortium by the Broad Institute.

9.1. Proposals for Access to Data

Any project (large or small) using the combined DNA samples or phenotype data gathered together by the Collaborative, or other unpublished Collaborative data, must be formally proposed to the Strategy Committee Group.

Each proposal will include:

- A description of what data is required.
- A description of the analysis to be performed.
- A timetable.
- The name of the Collaborative member who will be responsible for the project and a list of others who will be involved.
- A publishing/authorship plan (see section 5.)

All proposals will be distributed to all Strategy Committee members.

Each proposal will be discussed during Strategy Committee meetings and either approved or declined by consensus.

If a project is approved, no data will be used without the express permission of the investigators who contributed that data.

In addition to the use described above by Epi25 Collaborative members, the Epi25 database will collect a rich resource of genetic and phenotypic information that will be useful for research by non-Epi25 investigators. These investigators will be required to seek IRB approval and submit a research proposal to Epi25 as described previously.

9.2. NIH Data Sharing and dbGaP

Data will become available to the wider research community after a 2-year embargo into the NIH-maintained Database of Genotypes and Phenotypes (dbGaP). All data that will be transferred to dbGaP will be de-identified. Genetic sequence data will be accompanied by clinical endpoints that will be stripped of identifying information prior to transfer. Access to the data in dbGaP is controlled by an NINDS and NHGRI Data Access Committee. More detailed information is available at http://gds.nih.gov/04po2_1DAC.html.

Additional information is embedded here:

NIH Data Use Policy
NIH_GDS_Policy.pdf

In general, although this is whole genome and not GWAS (genome wide association study) data, the principles outlined in http://gds.nih.gov/pdf/PTC_for_IRBs_and_Institutions_revised5-31-11.pdf apply.

This document is embedded here:

Model Data Use
Certification Agreemer

Investigators and institutions seeking data from the results of the EPI25K project will submit a Data Access Request and a Data Use Certification to NIH that stipulates a number of protections for participants including:

- The data will be used only for the approved research
- Data confidentiality must be protected in a way described in detail by the requestor
- Data security protocols and protections are required
- All local laws and regulations must be followed
- No attempt to identify individual participants may be made
- No data can be sold
- Data cannot be shared outside of the listed and approved recipients (ie, no third party sharing)
- Research use only
- Any violations must be reported immediately to the appropriate Data Access Committee
- Annual progress reports on all data used are required.

10. Publication of Research Findings

For any research that would rely wholly or in significant part on Epi25 Collaborative or resources, Members are required to discuss their ideas within the Epi25 Collaborative. This principle of “no surprises” extends to all new analyses, poster and oral presentations, manuscripts and other tools of the trade.

Amongst other advantages, benefits of this cooperative approach are expected to include the strengthening and improving of preliminary ideas via discussion with fellow experts, the possibility of attracting additional resources in support of proposed projects, access to larger sample and data sets, better analytical tools, and sharing of complementary skills.

The Epi25 Steering Committee reviews publication topics, drafts, and final manuscripts on behalf of the Epi25 consortium. All publications must be reviewed and approved by this committee.

11. Authorship

The Collaborative is the authoritative coordinating body for the epilepsy-related WGS efforts. When appropriate, all authors/contributors will have their specific contributions acknowledged in tabular form as an appendix to each paper.

11.1. Junior Investigators

The Collaborative acknowledges that even though much of the analytic work will be performed by junior investigators (PhD students, postdoctoral fellows), the junior investigators in the field are naturally disadvantaged in large collaborative projects due to the limited number of leadership roles in team science, and the ongoing emphasis by most academic institutions and funding-agencies on contributions represented by first authorship or senior authorship in manuscripts. The Collaborative will try to compensate for this imbalance by insuring that: (a) contributions of all junior investigators will be explicitly and fully stated in manuscripts; (b) senior investigators advocate strongly for junior investigators in the promotions and funding process by including detailed descriptions of the junior investigator’s creative contributions and related attributes in letters of reference; (c) junior investigators have the opportunity to take lead roles and assume first-authorship for spin-off projects; (d) junior investigators have the opportunity to present Epi25 results at high-profile symposia and related events; and (e) junior investigators are actively involved in the key decision making processes of the Collaborative such as those related to the Strategy

Committee. The Collaborative considers these junior scientists as future leaders in the field and feels responsible for their mentoring for leadership in the field.

Given these considerations, membership in Epi25 will not only extend to the individual group leaders or Principal Investigators of a group, but to other group members actively involved in the project. The Principal Investigator will propose which members of his/her group should be considered for membership in the Collaborative.

11.2. Project Proposals Authorship Plans

All project proposals must include a specific publishing/authorship plan.

Typically the plan will be one of the following three options taken from the MS model:

- A full Epi25 Collaborative paper: The Epi25 Collaborative is listed as the author (along with other Collaboratives/Consortia, as appropriate), with a list of individual contributors made according to the journal's style preference. No individual Member will be designated as a corresponding author – the corresponding author will be listed as the Epi25 Collaborative with a mailing and email address.
- A partial Epi25 Collaborative paper: Individual authors listed. The Collaborative name is included in the author list, along with other Collaboratives/Consortia, as appropriate. (This option must include a list of authors.)
- A local paper: Simply acknowledges the Epi25 Collaborative.

All publications by the *Collaborative* will be open-access.

Junior scientists (PhD students, postdoctoral fellows) will be given preference in presenting data of the *Collaborative* at conferences.

Collaborative members will be reasonable and self-critical in their claims for authorship.

11.3. No Surprise Policy

All members of the *Collaborative* are bound by a "no surprise" policy with respect to Epi25 data and possibly interfering data from individual groups. The members acknowledge that "conflict of interest" can have a wide range of interpretations within the international framework of the *Collaborative*. Therefore, members are encouraged to proactively indicate conflicts if they feel that they might arise.

12. Human Subjects

12.1. Institutional Review Board (IRB) Review

This protocol and any subsequent modifications can be reviewed and approved if required by local IRBs responsible for cohorts submitting DNA samples to the Broad Institute. The Broad Institute has its own IRB approvals in place for the procedures conducted at the Broad. Although this is not a research trial, it is a secondary use of research materials, all research will be conducted in compliance with the protocol, current Good Clinical Practices (GCP), adopting the principles of the Declaration of Helsinki, and all applicable regulatory requirements.

12.2. Informed Consent

Because Epi25 is a secondary use project and not a clinical trial, it is up to the discretion of the local IRB supervising the study that collected the original DNA samples whether the samples can be shared under the previously obtained informed consent, or whether the IRB will grant a waiver to consent.

13. Risks and Benefits

13.1. Potential Benefits to Subjects

There are no direct benefits to the subject in participating in this study, but study participants will know that they have contributed to a major, long-term effort that is likely to better the lives of millions of people with epilepsy for many generations to come.

13.2. Potential Risks :Genetic Information

The collection of genetic information about participants carries the risk of a participant developing anxiety or other untoward emotional responses to the knowledge about potential disease susceptibility. Genetic information can also have damaging effects on family relationships, and can influence matters related to employment or health insurance. Any connection between a participant enrolled in Epi25 and his or her genetic information will be severed through the use of anonymized participant identifiers and password-protected access to the Epi25 Data Repository.

14. References

- Abou-Khalil B, Alldredge B, Bautista J, Berkovic S, Bluvstein J, et al. 2013. The epilepsy phenome/genome project. *Clin Trials* 10: 568-86
- Annegers JF, Hauser WA, Anderson VE, Kurland LT. 1982. The risks of seizure disorders among relatives of patients with childhood onset epilepsy. *Neurology* 32: 174-9
- Annegers JF, Hauser WA, Elveback LR. 1979. Remission of seizures and relapse in patients with epilepsy. *Epilepsia* 20: 729-37
- Bearden D, Strong A, Ehnot J, DiGiovine M, Dlugos D, Goldberg EM. 2014. Targeted treatment of migrating partial seizures of infancy with quinidine. *Ann Neurol* 76: 457-61
- Ben-Shachar S, Lanpher B, German JR, Qasaymeh M, Potocki L, et al. 2009. Microdeletion 15q13.3: a locus with incomplete penetrance for autism, mental retardation, and psychiatric disorders. *J Med Genet* 46: 382-8
- Berg AT, Berkovic SF, Brodie MJ, Buchhalter J, Cross JH, et al. 2010. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia* 51: 676-85
- Berkovic SF, Howell RA, Hay DA, Hopper JL. 1998. Epilepsies in twins: genetics of the major epilepsy syndromes. *Ann Neurol* 43: 435-45
- Berry JG, Poduri A, Bonkowsky JL, Zhou J, Graham DA, et al. 2012. Trends in resource utilization by children with neurological impairment in the United States inpatient health care system: a repeat cross-sectional study. *PLoS Med* 9: e1001158
- Carvill GL, Regan BM, Yendle SC, O'Roak BJ, Lozovaya N, et al. 2013. GRIN2A mutations cause epilepsy-aphasia spectrum disorders. *Nat Genet* 45: 1073-6
- Corey LA, Berg K, Pellock JM, Solaas MH, Nance WE, DeLorenzo RJ. 1991. The occurrence of epilepsy and febrile seizures in Virginian and Norwegian twins. *Neurology* 41: 1433-6
- Cowan LD. 2002. The epidemiology of the epilepsies in children. *Ment Retard Dev Disabil Res Rev* 8: 171-81
- de Kovel CG, Trucks H, Helbig I, Mefford HC, Baker C, et al. 2010. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 133: 23-32
- de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, et al. 2012. Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. *N Engl J Med*
- Dibbens LM, de Vries B, Donatello S, Heron SE, Hodgson BL, et al. 2013. Mutations in DEPDC5 cause familial focal epilepsy with variable foci. *Nat Genet* 45: 546-51
- Dibbens LM, Mullen S, Helbig I, Mefford HC, Bayly MA, et al. 2009. Familial and sporadic 15q13.3 microdeletions in Idiopathic Generalized Epilepsy: Precedent for Disorders with Complex Inheritance. *Hum Mol Genet*
- Endele S, Rosenberger G, Geider K, Popp B, Tamer C, et al. 2010. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* 42: 1021-6
- Epi4K Consortium and Epilepsy Phenome/Genome Project. 2013. De novo mutations in epileptic encephalopathies. *Nature* 501: 217-21
- Euro Epinomics- R. E. S. Consortium, Epilepsy Phenome/Genome Project, Epi4k Consortium. 2014. De Novo Mutations in Synaptic Transmission Genes Including DNM1 Cause Epileptic Encephalopathies. *Am J Hum Genet* 95: 360-70
- Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, et al. 2014. De novo mutations in schizophrenia implicate synaptic networks. *Nature* 506: 179-84
- Gaitatzis A, Carroll K, Majeed A, J WS. 2004. The epidemiology of the comorbidity of epilepsy in the general population. *Epilepsia* 45: 1613-22
- Gulsuner S, Walsh T, Watts AC, Lee MK, Thornton AM, et al. 2013. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* 154: 518-29
- Harkin LA, McMahon JM, Iona X, Dibbens L, Pelekanos JT, et al. 2007. The spectrum of SCN1A-related infantile epileptic encephalopathies. *Brain* 130: 843-52

- Heinzen EL, Radtke RA, Urban TJ, Cavalleri GL, Depondt C, et al. 2010. Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. *Am J Hum Genet* 86: 707-18
- Helbig I, Mefford HC, Sharp AJ, Guipponi M, Fichera M, et al. 2009. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet*
- International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address e-auea. 2014. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol* 13: 893-903
- Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, et al. 2014. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*
- Ishida S, Picard F, Rudolf G, Noe E, Achaz G, et al. 2013. Mutations of DEPDC5 cause autosomal dominant focal epilepsies. *Nat Genet* 45: 552-5
- Jiang Y, Satten GA, Han Y, Epstein MP, Heinzen EL, et al. 2014. Utilizing population controls in rare-variant case-parent association tests. *Am J Hum Genet* 94: 845-53
- Lal D, Reinthaler EM, Schubert J, Muhle H, Riesch E, et al. 2014. DEPDC5 mutations in genetic focal epilepsies of childhood. *Annals of neurology*
- Lemke JR, Lal D, Reinthaler EM, Steiner I, Nothnagel M, et al. 2013. Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet* 45: 1067-72
- Lesca G, Rudolf G, Bruneau N, Lozovaya N, Labalme A, et al. 2013. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. *Nat Genet* 45: 1061-6
- McCarthy SE, Gillis J, Kramer M, Lihm J, Yoon S, et al. 2014. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol Psychiatry* 19: 652-8
- Mefford HC, Yendle SC, Hsu C, Cook J, Geraghty E, et al. 2011. Rare copy number variants are an important cause of epileptic encephalopathies. *Ann Neurol* 70: 974-85
- Milligan CJ, Li M, Gazina EV, Heron SE, Nair U, et al. 2014. KCNT1 gain of function in 2 epilepsy phenotypes is reversed by quinidine. *Ann Neurol* 75: 581-90
- Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, et al. 2012. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485: 242-5
- Nevalainen O, Raitanen J, Ansakorpi H, Artama M, Isojarvi J, Auvinen A. 2013. Long-term mortality risk by cause of death in newly diagnosed patients with epilepsy in Finland: a nationwide register-based study. *Eur J Epidemiol* 28: 981-90
- O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, et al. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485: 246-50
- Olson H, Shen Y, Avallone J, Sheidley BR, Pinsky R, et al. 2014. Copy number variation plays an important role in clinical epilepsy. *Ann Neurol* 75: 943-58
- Ottman R, Annegers JF, Risch N, Hauser WA, Susser M. 1996a. Relations of genetic and environmental factors in the etiology of epilepsy. *Ann Neurol* 39: 442-9
- Ottman R, Lee JH, Risch N, Hauser WA, Susser M. 1996b. Clinical indicators of genetic susceptibility to epilepsy. *Epilepsia* 37: 353-61
- Ottman R, Susser M. 1992. Data collection strategies in genetic epidemiology: The Epilepsy Family Study of Columbia University. *J Clin Epidemiol* 45: 721-7
- Ottman R, Winawer MR, Kalachikov S, Barker-Cummings C, Gilliam TC, et al. 2004. LGI1 mutations in autosomal dominant partial epilepsy with auditory features. *Neurology* 62: 1120-6
- Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, et al. 2014. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* 506: 185-90
- Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, et al. 2012. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet*
- Samocha KE, Robinson EB, Sanders SJ, Stevens C, Sabo A, et al. 2014. A framework for the interpretation of de novo mutation in human disease. *Nat Genet* 46: 944-50

- Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, et al. 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* 485: 237-41
- Scheffer IE, Berkovic SF. 1997. Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes. *Brain* 120 (Pt 3): 479-90
- Scheffer IE, Heron SE, Regan BM, Mandelstam S, Crompton DE, et al. 2014. Mutations in mTOR regulator DEPDC5 cause focal epilepsy with brain malformations. *Annals of neurology*
- Shain C, Ramgopal S, Fallil Z, Parulkar I, Alongi R, et al. 2013. Polymicrogyria-associated epilepsy: a multicenter phenotypic study from the Epilepsy Phenome/Genome Project. *Epilepsia* 54: 1368-75
- Shinawi M, Schaaf CP, Bhatt SS, Xia Z, Patel A, et al. 2009. A small recurrent deletion within 15q13.3 is associated with a range of neurodevelopmental phenotypes. *Nat Genet* 41: 1269-71
- Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, et al. 2008. Large recurrent microdeletions associated with schizophrenia. *Nature* 455: 232-6
- Thurman DJ, Hesdorffer DC, French JA. 2014. Sudden unexpected death in epilepsy: Assessing the public health burden. *Epilepsia* 55: 1479-85
- Tomson T, Walczak T, Sillanpaa M, Sander JW. 2005. Sudden unexpected death in epilepsy: a review of incidence and risk factors. *Epilepsia* 46 Suppl 11: 54-61
- Widdess-Walsh P, Dlugos D, Fahlstrom R, Joshi S, Shellhaas R, et al. 2013. Lennox-Gastaut syndrome of unknown cause: phenotypic characteristics of patients in the Epilepsy Phenome/Genome Project. *Epilepsia* 54: 1898-904
- World Health Organization. 2012.
- Yuan H, Hansen KB, Zhang J, Pierson TM, Markello TC, et al. 2014. Functional analysis of a de novo GRIN2A missense mutation associated with early-onset epileptic encephalopathy. *Nature communications* 5: 3251