Electronic RESEARCH INTEGRITY AND COMPLIANCE THE UNIVERSITY OF UTAH

Date: Friday, June 20, 2025 1:27:54 PM Print Close

IRB_00044244 **Created:** 8/17/2010 1:10 PM

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

IRB_00044244

1. Contacts and Title

1. Study Introduction

1. **Responsible Investigator:**
   Hilary Coon

Email	Training	CoI Date
hilary.coon@utah.edu	12/3/2024 SMG	6/17/2025

a. **Position of the Investigator:**

- (●) **Faculty or** Non-Academic Equivalent
- ( ) Student
- ( ) Staff
- ( ) Resident/Fellow
- ( ) Other

2. **Contact Persons for the Responsible Investigator:**

Name	Email	Training
Amanda Bakian	amanda.bakian@hsc.utah.edu	4/25/2024 MG
Hilary Coon	hilary.coon@utah.edu	12/3/2024 SMG
Marie Gibson	marie.gibson@hci.utah.edu	11/16/2022 MCG
Emily Sullivan	Emily.Sullivan@hsc.utah.edu	2/26/2024 MG
Jennifer West	jennifer.a.west@utah.edu	4/7/2025 SMCG

3. **Guests of the Responsible Investigator:**

Last Name	First Name	E-Mail

There are no items to display

4. **What type of application is being submitted?**
   New Study Application (or Amendment/Continuing Review)

5. **Title Of Study:**
   Genetic risk factors in suicide and depression

6. **Study Purposes and Objectives:**

OBJECTIVES

A substantial part of this study is EXPLORATORY. We have an excellent opportunity through these resources to make significant contributions to a highly important area of discovery. Our resource, with high-risk extended

pedigrees and DNA, will help anchor results to meaningful outcomes that we hope will have future clinical relevance.

Broadly, our aims focus on discovery of genetic variants leading to suicide risk. We begin with high-risk extended families; we hypothesize that cases related to common founding couples may have more genetic homogeneity in risk factors. This is the unifying hypothesis of our project. Much of the additional work to determine co-occurring conditions associated with specific genetic risk is necessarily exploratory, as this field is still in its infancy. Some aspects of the study are indeed driven by hypotheses generated from known associations. For example, we hypothesize that genetic variants may be shared among familial suicides with co-occurring conditions already documented in the literature as co-occurring with suicide. These include psychiatric disorders and selected medical disorders. However, it is still unknown if risk variants will show such specificity, or if genetic risk may cut across co-occurring conditions. It is unknown if there are specific patterns of co-occurring conditions that may be more likely to be present with particular mutations.

As we make progress and begin to describe cases in high risk families sharing possible genetic risk, it is critical that we be able to describe these cases as fully as possible. In follow-up work looking for mutations in the gene(s) indicated by these family studies, we will also be able to document the presence/absence of specific characteristics. Much of what we do will necessarily be descriptive, and will require comparison to characteristics of cases in future work with other investigators, or future in vitro or animal studies to investigate mechanisms of genetic mutations.

In addition to our descriptions of cases in high risk families who share genetic risk variants, we will also take full advantage of co-occurring conditions to try to actually select and prioritize unusual families for study. We will use cluster analysis and known clinical groupings of diagnoses to aggregate the diagnostic data. We will then use the familial standardized incidence ratio (FSIR) of each aggregated diagnostic cluster to determine if some high-risk suicide families are also at higher risk for co-occurring conditions. Knowledge of conditions within families can direct genetic results through suggesting gene pathways that may lead to familial co-occurring conditions. Follow-up studies will require replication using data from our many colleagues in consortium groups (e.g., Psychiatric Genomics Consortium, International Suicide Genetics Consortium), and we will again turn to our bench science colleagues for future functional/animal studies.

Finally, we will also use epidemiological methods to determine increased likelihood of co-occurring conditions, with familiality of suicide and the co-occurring conditions as a critical clustering factor. The only previous well powered studies to investigate these questions have only surveyed a small number of potential co-occurring conditions, have not had deep familial data, and/or have not had DNA to allow follow-up of phenotypic associations. Because our data set is so large (>20,000 suicides >12,000 of which are recent enough to have diagnostic ICD9/ICD10 data, and over 2000 families with significant FSIR for suicide), we can employ internal replication techniques, splitting our sample into discovery and replication cohorts. We are pursuing this line of work in addition to prioritizing the highest risk suicide pedigrees, independent of co-occurring conditions. We plan to allow results from both lines of work to inform each other.

This study will involve RECORDS ONLY. Our current work includes:

Genome-wide association using genotyping on suicide deaths in comparison to available Utah controls (e.g., CEPH and the Cache Valley longevity study), and large publicly-available control data sets. Develop an additional resource of genetic and phenotypic data from non-suicide deaths with evidence of prior attempt to elucidate similarities and differences of risks of attempt vs. suicide death. Prioritize suicide deaths for more extensive genetic sequencing using unique high-risk pedigrees available through the Utah Population DataBase (UPDB). Refine the estimate of familial suicide risk by determining the occurrence of suicide deaths using death record data. Enrich the phenotype information in the families by obtaining hospital discharge codes and ICD-9 codes from the UUHSC and IH Electronic Data Warehouse (EDW) for psychiatric and medical conditions (records-only analysis).

Prioritize cases for genetic study, as follows. We will identify characteristics that discriminate high-risk suicide pedigrees (decedents and their relatives) from “non-familial” decedents and relatives (i.e., genealogical records show no increased familial risk of suicide).

Familial characteristics to be investigated will include method of suicide, demographics, and familial co-morbid conditions. We have permission to study basic demographics and method of suicide through death certificate and OME data. We will also study number of children in the family, SES as indexed by the maximum achieved occupation code, and marital history. We will study birth weight, gestational age, and other demographic early life risk factors in analyses led by our co-investigator Amanda Bakian. We will study co-morbid affective disorders, psychosis, suicidal ideation, alcohol/drug-related disorders, and post-traumatic stress disorder (PTSD), including traumatic brain injury. We will study average age at death from all causes (excluding suicide cases) within each family, as an indicator of serious chronic disease. We will include the following specific conditions that are also known to be strongly associated with suicide in the absence of psychiatric diagnoses: chronic pain (Fegg et al., 2016; Calati et al., 2015; Novic et al., 2015), asthma (Goodwin & Eaton, 2005; Clarke et al., 2008) and other lung conditions (Young, 2013; Crump et al., 2013; Goodwin, 2011), epilepsy (Christensen et al., 2007), severe obesity (Zeller et al., 2013; Wagner et al., 2013; Mitchell et al., 2013; Adams et al., 2007), inflammatory bowel disease (Gradus et al., 2010), cardiovascular disease (Crump et al., 2013; Jee et al., 2011; Larsen et al., 2010), immune/autoimmune disorders (Sekar et al., 2016), dementia/neurodegenerative disorders (Takacs et al., 2015). Additional diagnoses will allow us to determine relationships between co-occurring disease and suicide risk that are not already documented. Prescription medication data records will provide additional data to corroborate electronic diagnoses, and will allow us to study characteristics of suicides with prescription opiate use. Suicide subsets with these conditions will be described in comparison to aggregated descriptive statistics from matched non-suicide controls from the Utah Population Database.

High-risk pedigrees were determined as described in the statistical section. Pedigrees that do not show increased risk for suicide will be determined similarly, but for these families, familial risk will not exceed 1 (normal population risk; no aggregation). These comparison families will be screened for accidental and undetermined deaths from death certificate and OME data, as these deaths may be uncoded suicides.

Cases within pedigrees with the highest relative risk will be sequenced using the existing DNA samples on the suicide victims; no new DNA or additional phenotypes will be collected from high-risk families unless additional permission is obtained to do so in future IRB amendments or new studies. We will prioritize top sequence variants for validation and replication through analyses using SGS, VAAST, Condel, and other analysis tools being developed. We will compare sequence data to reference data; use extensive knowledge of exome function from public sources, evolving gene regulation data resources (e.g., RegulomeDB, Boyle et al., 2012), and evolving knowledge of psychiatric gene pathways (e.g., Sullivan et al., 2012). We will test if findings from primary high-risk pedigrees are implicated in the remaining pedigrees, in other suicide deaths, and in collected cases with suicide attempt and/or psychopathology from international consortia.

We will confirm top gene findings in original subjects using Sanger sequencing. We will extend gene results for 2-3 genes per year through targeted re-sequencing in up to 1000 additional selected suicides per gene. We will select cases by linking new Utah suicides with DNA collected during the project to the UPDB to 1) extend the known high-risk pedigrees, and 2) determine new high-risk pedigrees. For each prioritized gene, we will re-sequence new familial suicides, and suicides in the large UT cohort who share highly similar phenotypic profiles to the original cases with the confirmed variants. This re-sequencing is done in these selected cases to identify more complete genetic variation within the prioritized genes of interest.

We will use studies in tissue (brain tissue, hair and derived neuronal cells from fibroblast tissue) to study gene/protein expression and cellular consequences of DNA variants found in these aims. Hair will be analyzed to look for environmental exposures prior to death.

We have concluded studies of suicide risk in Bipolar Disorder (BD) in collaboration with the University of Iowa. This center has detailed phenotyping on BD cohorts with and without suicide attempts. We compared basic demographics, age at first diagnosis, polygenic scores, and frequencies of more rare, putatively functional gene variants across these two cohorts and Utah BD suicide deaths.

Note: For this study, data are analyzed at Utah on CHPC, and a copy of **HIPAA Safe Harbor de-identified data** will be stored on UBox and access to UBox will be managed by the UPDB Data Security Analyst.

7. **Is this a multi-site study, where more than one site needs IRB approval?**
   ○ Yes ● **No**

8. **Background and Introduction:**

Over 48,000 individuals die by suicide in the US every year (https://afsp.org/suicide-statistics/; Hedegaard et al., 2018). Suicide rates are particularly high in Utah, which has the sixth highest in the nation (CDC, 2019). Prediction and prevention has become a high priority for public heath, but suicide rates continue to rise, and identification of risk factors remains difficult. A 50-year meta-analysis of the demographic, diagnostic, and sociological risk factors for suicide shows that prediction is only slightly better than chance (Franklin et al., 2017). Knowledge of biological risk could greatly enhance our ability not only to predict those at risk of death, but also to begin to develop more personalized interventions. While complex environmental variables have undeniable impact, evidence also suggests that genetic factors play a substantial role in *suicide death*, with estimates of heritability from aggregated studies at close to 50% (Pederson & Fiske, 2010; McGuffin et al., 2001). However, discovery of specific genetic risk variants still represents a significant hurdle in our understanding of important biological mechanisms of risk. To date, genetic studies have primarily focused on *suicidal behaviors* in order to maximize statistical power. The rate of suicidal behaviors is relatively common (4.3% per year) compared to suicide death (0.01-0.02% per year; Hedegaard et al., 2018). While genetic studies of suicidal behaviors have become more common in recent years and have resulted in some promising findings (Willour et al., 2012; Mirkovic et al., 2016; Galfalvy et al., 2015; Stein et al., 2017; Erlangsen et al., 2018; Strawbridge et al., 2019), replication has proven challenging, likely due in part to variation in phenotype definition and clinical ascertainment. Importantly, only a small fraction of those who engage in suicidal behavior will go on to die by suicide, and studies of actual suicide death have, to date, been small and underpowered because of the difficulty in obtaining data for this rare, severe phenotype. *To truly understand risk of suicide death and to implement highly effective interventions that provide appropriate, targeted services to those most likely to die, we must understand the risks specifically associated with suicide deaths. While other large data resources exist for the study of suicide behaviors, the Utah Suicide Mortality Risk Study (USMRS) has the world's only large collection of biosamples from suicide deaths.*

Mood disorders and/or substance use are the most common diagnoses among suicide decedents (Beghi & Rosenbaum, 2010; Ernst et al., 2009). Importantly, specific genetic risk for suicide exists above and beyond risk of psychiatric conditions (McGirr et al., 2009; Ernst et al., 2009; Kim et al., 2005; Brent et al., 2002; Egeland & Sussex, 1985). Genetic risk for suicide puts some families already at high risk for these psychiatric diagnoses at additional

high risk for suicide. Suicidologists have postulated that impulsivity and aggression may be intermediate phenotypes for this increased risk (McGirr et al., 2009; Brent & Mann, 2005).

The study will rely upon our large collection, detailed phenotypes, and powerful high-risk pedigrees, ascertained through the Utah Population Data Base (UPDB), a rich resource of health data and family history information for over 9 million individuals (www.hci.utah.edu/groups/ppr/). Shared genetic mutations within these high risk families can reveal potential causes that would be impossible to find in the otherwise vast complexity of suicide risk. The study will also benefit from the availability to follow up findings in an additional large sample of DNA from thousands of Utah suicides. Because of ongoing collection, this sample will grow to > 10,000 by 2025.

Evidence supporting genetic risk for suicide comes from several sources. Risk of suicidal behavior to relatives is increased approximately five fold (Kim et al., 2005; Baldessarini & Hennen, 2004; Brent et al., 2002). Adoption studies show increased risk of suicide for biological relatives of suicidal probands, but no increased risk caused for adopting relatives (Wender et al., 1986; Brent & Mann, 2005). Twin studies provide an estimate of the variation in suicide risk due to genetic factors (heritability). Suicide and suicidal behavior is more concordant among MZ as compared to DZ twin pairs (Pederson & Fiske, 2010; Althoff et al., 2009; Roy & Segal, 2001). Recent estimates aggregating available studies from large data resources indicate that heritability of completed suicide is approximately 45% (Pederson & Fiske, 2010; Althoff et al., 2009; McGirr et al., 2009; McGuffin et al., 2001; Fu et al., 2002), even after accounting for other risk factors, such as psychiatric history, traumatic life events, and socio-economic status (SES) factors.

A hypothesis of recent interest implicates mitochondrial dysfunction. Living at increased elevation may create metabolic stress, and individuals susceptible to suicide may lack compensatory mechanisms to overcome this stress (Haws et al., 2009). Mutations in the mitochondrial genome are therefore also potential candidates. There is overlap in symptomology between psychiatric and mitochondrial disorders, and mitochondrial dysfunction has been suggested for bipolar disorder (Stork & Renshaw, 2005; Hamakawa et al., 2004). Variants in the mitochondrial genome have been associated with altered brain pH and mood disorders (e.g., Kato & Kato, 2000; Kazuno et al., 2006, 2009). In addition, increased brain pH has been directly associated with bipolar disorder (Stork & Renshaw, 2005).

Our large sample also allows us to study demographic and co-morbid characteristics of “non-familial” suicide; that is, suicide probands and their relatives where the genealogical data indicate that there is no increased familial risk of suicide. Comparisons to these “non-familial” suicide probands and relatives will reveal key, quantifiable differences that will allow us to prioritize the high-risk families for genetic analyses (i.e., those most different from “non-familial” suicide). Finally, the large, growing resource of Utah suicide DNA will be critical for follow up and for future studies. Our study is timely, benefiting from: an explosion of new techniques and resources for sequence analysis; increasingly large comparison sets of control subjects and public data sets from studies of other psychiatric disorders; increasing knowledge of gene pathways in psychiatry (e.g., Sullivan et al., 2012); and decreasing molecular costs. Gene variants will provide information about biological etiology and pathways that will become a springboard for additional studies and perhaps targeted interventions.

Importantly, we are requesting permission to compare genetic, phenotypic, and familial risk factors between suicide deaths and individuals who attempt suicide, but did NOT die by suicide. Individuals who die by suicide and individuals who attempt suicide have fundamental epidemiological differences. For example, there is a highly significant difference in sex ratio for suicide death (3.8:1 male to female) vs. attempt (approximately 2:1 female to male) (Hedegaard et al., 2018). In addition, studies show that from population-ascertained data, ~60% of suicide deaths had no prior attempts (Owens et al., 2002; Cavanagh et al., 2003; Levey et al., 2019). Also, though suicide attempt is currently one of the best predictors of suicide death, fewer than 10% of those who attempt go on to die by suicide (Beisher et al., 2019). These statistics suggest that while identifying risks for suicide attempt is important, this knowledge will not necessarily apply to suicide deaths. Understanding risks of this most vulnerable group will be essential to guide limited resources to the task of saving lives. We also now have the opportunity to study limited use data from a genetically informative comparison group of deaths from the Office of the Medical through stored blood spots at the OME dating back to 2005. These deaths will provide an important Utah-population-based non-suicidal control group.

In the past, we collected a limited cohort of human brain tissue; this aspect of the study is now closed.

We continue to collect hair and skin biopsies which can be used to create fibroblast cultures. Analysis of these tissues will complement the pace of discovery of genetic variants associated with suicide risk and will potentially dramatically increase the impact of the parent project. We will conduct pilot studies using this tissue, which will enable studies investigating the molecular, cellular, and physiological impacts of suicide risk genes identified in DNA. Studies in derived neuronal cells from fibroblast cultures can include: RNA and protein expression, methylation, studies of synapse morphology, density, and specificity. These results will help prioritize genes and gene pathways for DNA mutation analyses. Conversely, expression and cellular analyses will also allow us to test whether genetic variants identified using DNA sequencing lead to specific expression and cellular changes in the brain. Studies of hair will yield insight into environmental exposures prior to death.

Viably frozen skin biopsies and established fibroblast cultures will be banked for the future. The resource will allow us to generate human induced pluripotent stem cells (iPSCs) that can be differentiated into neurons from suicide victims

(and corresponding controls) where the suicide cases have been identified as having clear genetic or cellular defects. This will allow us to examine neuronal form and function when comparable underlying genetic mutations are observed in DNA. iPSC-derived neurons will be used in future experiments for 1) electrophysiology and calcium imaging to identify precise functional defects, 2) rescue experiments in which mutated genes are re-expressed in neurons to determine if functional defects can be rescued, and 3) drug screening to identify potential new therapeutics. All will increase our ability to validate genetic risk factors. Because of the ease of collection, we anticipate being able to collect and successfully culture fibroblasts for most suicides. We will create fibroblast cultures as needed depending on the hypothesis of each proposed iPSC experiment. Because this selection hinges upon the outcome of ongoing studies, these experiments will primarily be left for future follow-up studies. However, we must collect and bank the tissue now to enable this future work.

IRB_00044244 **Created:** 8/17/2010 1:10 PM

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

IRB_00044244
2. Study Location and Sponsors

2. Study Location and Sponsors

1. **Add all locations applying for approval of research via the University of Utah IRB or Human Research Protection Program (HRPP).**

   Click the appropriate button(s) below to add locations:

	Site Name	Investigators Name	Covered Entity	Sub Sites
View	University of Utah	Hilary Coon	No Yes	

2. **Will a Central IRB (CIRB) or Single IRB (SIRB) model be used for review of this study for the sites listed in this application?**
   ◯ Yes ● **No**

3. **Indicate the source(s) of funding obtained or applied for to support this study.**

	Sponsor	Sponsor Type	Sponsor Contact Information	Prime Sponsor	Prime Sponsor Type	OrgID
View	NIH NATIONAL INSTITUTE OF MENTAL HEALTH	Federal Government				10172

4. **Does this study have functions assigned to a Contract Research Organization (CRO)?**
   ◯ Yes ● **No**

5. **Does this study involve use of the Utah Resource for Genetic and Epidemiologic Research (RGE)?**
   Examples: Utah Population Database (UPDB), Utah Cancer Registry (UCR), All Payers Claims Database (APCD), etc.
   ● **Yes** ◯ No

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244** View/Edit

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Addition of a Site

1. **Site Name:**
   University of Utah

2. **Site Principal Investigator**
   ☑ **Mark if Same as Responsible Investigator (syncs with investigator on the first page)**
   Hilary Coon

Email	Training	CoI Date
hilary.coon@utah.edu	12/3/2024 SMG	6/17/2025

   a. **Position of the Site Principal Investigator**
   Faculty or Non-Academic Equivalent

   b. **Will the Site PI consent participants?** ○ Yes ● **No**

3. **Site Contact Persons, if different from the Site PI:**
   ☑ **Mark if Same as Contacts for Responsible Investigator (syncs with contacts on the first page)**

Name	Email	Training
Amanda Bakian	amanda.bakian@hsc.utah.edu	4/25/2024 MG
Hilary Coon	hilary.coon@utah.edu	12/3/2024 SMG
Marie Gibson	marie.gibson@hci.utah.edu	11/16/2022 MCG
Emily Sullivan	Emily.Sullivan@hsc.utah.edu	2/26/2024 MG
Jennifer West	jennifer.a.west@utah.edu	4/7/2025 SMCG

4. **Site Staff and Sub-Investigators**

Name	Email	Training	Obtaining Consent	CoI Date
Lisa Baird	lbaird@genetics.utah.edu	1/6/2025 MG	☐	9/11/2024
Amanda Bakian	amanda.bakian@hsc.utah.edu	4/25/2024 MG	☐	2/6/2025
Deborah Bilder	deborah.bilder@hsc.utah.edu	2/12/2025 SMG	☐	1/27/2025
Nicola Camp	nicola.camp@hci.utah.edu	2/3/2023 MG	☐	6/6/2025
Danli Chen	u6010503@utah.edu	10/20/2022 MCG	☐	2/6/2025
EMILY DIBLASI	emily.diblasi@hsc.utah.edu	1/16/2024 MG	☐	6/11/2025

Name	Email	Training	Obtaining Consent	CoI Date
Anna Docherty	docherty.anna@gmail.com	7/27/2022 M	☐	2/12/2025
Karen Eilbeck	keilbeck@genetics.utah.edu	10/12/2024 MG	☐	4/23/2025
Elliott Ferris	elliott@neuro.utah.edu	1/2/2024 MG	☐	5/13/2025
Seonggyun Han	seonggyun.han@utah.edu	3/6/2024 M	☐	2/17/2025
Brooks Keeshin	brooks.keeshin@hsc.utah.edu	3/20/2023 MCG	☐	1/6/2025
Anne Kirby	anne.kirby@hsc.utah.edu	11/30/2024 MG	☐	6/10/2025
Vincent Koppelmans	vincent.koppelmans@utah.edu	6/5/2023 MCG	☐	5/29/2025
Gabor Marth	u0950806@utah.edu	5/25/2023 MCG	☐	2/6/2025
Brian Mickey	brian.mickey@utah.edu	11/19/2024 MCG	☐	3/20/2025
Eric Monson	eric.monson@hsc.utah.edu	11/4/2024 MG	☐	4/22/2025
Michael Newman	u0079868@umail.utah.edu	3/10/2025 M	☐	10/14/2024
Aaron Quinlan	aquinlan@genetics.utah.edu	7/10/2024 MCG	☐	2/18/2025
Perry Renshaw	perry_renshaw@yahoo.com	4/22/2025 MG	☐	1/27/2025
Mary Beth Scholand	scholand@genetics.utah.edu	5/18/2024 MCG	☐	6/19/2025
Andrey Shabalin	andrey.shabalin@utah.edu	11/23/2022 MG	☐	9/17/2024
Oleksandr Shcheglovitov	alex8yk@gmail.com	11/10/2023 M	☐	9/24/2024
Michael Staley	mstaley@utah.gov	2/6/2024 MG	☐	2/19/2025
Emily Sullivan	Emily.Sullivan@hsc.utah.edu	2/26/2024 MG	☐	6/4/2025
Douglas Tharp	doug.tharp@hsc.utah.edu	3/5/2025 SMG	☐	9/25/2024
James VanDerslice	jim.vanderslice@utah.edu	9/24/2022 MG	☐	3/17/2025
Rachel Weir	Rachel.Weir@hsc.utah.edu	4/14/2025 M	☐	4/8/2025
Mark Yandell	myandell@genetics.utah.edu	4/3/2023 MCG	☐	6/11/2025
Zhe Yu	zhe.yu@hci.utah.edu	5/17/2024 MG	☐	9/24/2024

Name	Email	Training	Obtaining Consent	CoI Date
Deborah Yurgelun-Todd	deborah.yurgelun-todd@hsc.utah.edu	7/24/2024 MG	☐	6/17/2025
Yue Zhang	zhang.yue@hsc.utah.edu	5/13/2025 MCG	☐	6/7/2025

5. **Site Guests:**

Name	Email	Training

There are no items to display

6. **Select HIPAA coverage for this study:**
Study procedures will be conducted within a HIPAA Covered Entity at this site (HIPAA Privacy Rule applies)
Study procedures will be conducted outside a HIPAA Covered Entity at this site (HIPAA Privacy Rule does not apply)

7. **Select the study procedures that will be conducted at this site:**
Data analysis

8. **Select the University of Utah department responsible for this research:**
PSYCHIATRY

9. **Add any additional sites that are part of this performance group**
There are no items to display

IRB_00044244 **Created:** 8/17/2010 1:10 PM

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

IRB_00044244 IRB Smart Form

Sponsor Information

a. Are you receiving award or contract management for the sponsored funds through the University of Utah Office of Sponsored Projects?

● **Yes** ○ No

If yes, select the associated OSP Proposal ID/DSS through eAward to link it to the ERICA system.

You must have a fully approved Proposal ID/DSS number through eProposal which will show up in eAward after OSP has integrated the ID. To access the eAward application, use the instructions on the OSP website.

Link to a Proposal ID/DSS through eAward

Proposal ID/DSS: 10053740
PI: COON,HILARY H
Sponsor: NIH NATIONAL INSTITUTE OF MENTAL HEALTH
Prime Sponsor:
Department:
Short Title: SUICIDE WGS RISK DISCOVERY
Sponsor Award Number: 5R01MH122412-05
Type: Federal Government
Award Start Date: 4/1/2020
Award End Date: 1/31/2026
Prime Sponsor Type:

IRB_00044244 **Created:** 8/17/2010 1:10 PM

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

IRB_00044244 3. Participants

3. Participants

1. **Ages of Participants:**

7 to 17 years old	(Parental permission and assent form needed)
18 and older	(Consent form needed)

2. **Specific age range of participants (e.g., 7-12 years old, 60+, etc.):**
all DNA on suicide cases collected through Medical Examiner, all ages. Deaths < 7 will not be considered suicides.

3. **Indicate any vulnerable participant groups (other than children) included:**

None

If "Other", please specify:

If "None" and no children are involved, answer the following question.
Has the participant selection process overprotected potential subjects who are considered vulnerable so that they are denied opportunities to participate in research?
○ Yes ● **No**

4. **Number of participants to be included and/or enrolled in this entire study, across all study locations:** ~20,000 cases, 1.5 million controls with ICD9/ICD10 data, and 9 million persons overall to accommodate selection for each of the risk factors that will be analyzed for the study

At Utah prior to October 2019: ~20,000 cases, 1.5 million controls with ICD9/ICD10 data, and 9 million persons overall to accommodate selection for each of the risk factors that will be analyzed for the study

5. **Characteristics of Participants/Inclusion Criteria:**

NOTE: There are >20,000 suicides identified through death certificate data available for epidemiological study. Of these, about 9000 have already been identified and DNA collected through our long-term collaboration with the OME. This number of cases grows by 700 per year. There are ~1,500,000 cases with ICD9/ICD10 codes from which we can select matched comparisons for suicide cases. The ~9,000,000 records in the UPDB will be used to screen for broader de-identified matched controls within the UPDB for epidemiological studies of risk of demographic characteristics/geocoded regions. Note that for subgroups of suicide risk (e.g., obesity/bariatric; PTSD/trauma; developmental delays; peri-/post-partum depression; air pollution exposure; COPD), matching may include appropriate characteristics above and beyond age at death, gender, and year of birth. For example, controls for female suicide cases with peri-/post-partum depression will need to include month/year of delivery of the child in addition to year of birth of the female suicide case.

All proposed studies will be done using limited use or fully de-identified data, and no new subject contact is requested. We anticipate that records for additional family members related to suicide victims will be studied by the UPDB to determine pedigree risk, but only suicide victims with DNA (and in some cases, skin tissue or hair samples) will be part of genetic studies in this protocol.

We continually add samples from decedents to the study. DNA will come from the Utah Office of the Medical Examiner (OME). We anticipate that new samples may add to existing high-risk pedigrees, or may perhaps identify new high-risk pedigrees. Any records that do not link to high-risk pedigrees will add to the follow-up study sample. Samples will continue to be collected and records will be periodically linked to the UPDB during the course of the study.

This is a family genetic study, using linking through the UPDB of existing records on suicide victims. The study will also determine if particular related phenotypes, identified through death records and hospital discharge codes, significantly co-occur with suicide in some families. This study will involve no direct contact with subjects. Identifying information will only be transferred from Vital Records (OME-generated data) only to the UPDB for record linking; no identifying information not already available will be disoclosed or used by University of Utah researchers.

There will be additional sources of de-identified subjects to be studied in the context of suicide risk. These are individuals with early-onset PTSD identified in a study by our co-investigtor Brooks Keeshin (IRB_00082991, Genetic and salivary neuroendocrine markers in youth

treated for traumatic experiences). These individuals, consented and studied in IRB_00082991, will be linked to the UPDB. Their records will be linked to records from suicide cases in the present study (Genetics of Suicide). PTSD cases will be linked to risk factor data for which we have already obtained permission to study in the Genetics of Suicide protocol. All data will be de-identified within the UPDB, as is done currently for the suicide cases, and the resulting de-identified data will be provided to the Genetics of Suicide analysis team. These cases may then be studied as part of high-risk pedigrees, or as follow up cases where phenotypic/demographic profiles match cases in high-risk pedigrees where there are shared genetic risk variants of interest. No PHI will be given to the research team, and, as with our current protocol, no attempt will be made to discover the identities of participants or members of high-risk family clusters. At least 200 de-identified PTSD cases from IRB_00082991 will be studied. The participants from IRB_00082991 will be linked via the Master Subject Index to the UUHSC EDW and the Master Linkage File to the IH EDW. Similarly, cases with gastric bypass, or eligible for gastric bypass (IRB_000006824), Chronic Obstructive Pulmonary Disease (COPD; IRB_00049521), and developmental disorders (IRB_00057455) will be linked to the suicide cohort within the UPDB. Data flow and analysis of de-identified data and samples will be as described above with the PTSD study.

We are additionally requesting to obtain blood spots from the Office of the Medical Examiner or tissue from the IH Biorepository to build a genetically-informative comparison cohort of **non-suicide** deaths with and also without evidence of prior suicide attempts. These comparison cases will provide an essential group for investigating differences in risk between individuals who die by suicide compared to individuals who attempt suicide but do not die by suicide, and finally to control individuals who did not die by suicide and have no suicidal behaviors.

To do this, 1) UPDBD PPR staff will identify **non-suicide** deaths back to 1995 using death certificate information. 2) Using this list of non-suicide deaths, the UUHSC Data Science Service Core and the IH honest data broker will identify the subset of deaths with UUHSC or IH EHR evidence of prior suicide attempt (ICD-9: 304.0*, 304.7*, 305.5*, 965.*, E850.0, E935.0, and ICD10 equivalents) to split the non-suicide deaths into those with vs. without prior suicidal behaviors. 3) For OME blood spots: blood spots are stored for all Utah deaths investigated by the OME back to 2005, so PPR staff will send identifiers of deceased non-suicide deaths with and without prior suicidal behaviors from 2005-present to the OME; OME staff will obtain 2 of the stored spots for DNA; these will be associated with numeric ids and will be picked up and logged using the long-standing system as samples for suicide deaths; DNA will be extracted by the CCTS DNA core and samples will be linked with other study data by PPR staff to create a limited use data set with the same protocols we have been using for suicides.

All samples for this study are from either 1) suicide deaths, 2) non-suicide deaths with prior attempts, or 3) non-suicidal deaths with no prior history of suicidal behaviors. The research team will only receive the limited use data that matches the data elements for the suicides. Given estimated rates, we anticipate the total number of attempts with useable samples to be several hundred per year, and non-suicide control deaths with useable samples to be >1000 per year. We have funding in hand to genotype up to 5000 of these comparison samples, and will seek additional funding for future further comparison studies.

6. Participant Exclusion Criteria:

Only these records will be included:

Records with DNA (and also possibly fibroblast/brain tissue/hair, if collected) are included 1) based on the determination of suicide made by the Office of the Medical Examiner (for suicide deaths); 2) non-suicide deaths both with and without evidence of prior suicidal behavior from the EHR as deceased comparison and deceased control cases, where suicidal behavior is defined using codes specified above.

Records only (no associated samples) for: 1) additional electronic records of suicides as listed on the death certificate were samples could not be collected are added for an accurate estimate of suicide risk in pedigrees; 2) electronic records of population controls where samples could not be collected matched by the UPDB PPR staff.

We request that DNA, fibroblast tissue, hair, and small samples of brain tissue are collected without next of kin consent because complete population ascertainment is critical to the success of discovery, and consent would severely limit inclusion. This is not primarily because next of kin would refuse, but because of logistics of family contact at this difficult time. However, we believe that whole brain donation is much more sensitive, and will only include collection of whole brains with next of kin consent (please see procedures for details). We believe that in the future, we will be able to quantify systematic differences between brain tissue and derived neuronal cells from fibroblasts, allowing us to extrapolate accurate results from derived cells in cases where brain tissue cannot be collected.

7. Is a substantial percentage of the participant population anticipated to be non-English speaking?

○ Yes ● **No**

IRB_00044244 **Created:** 8/17/2010 1:10 PM

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

IRB_00044244 - Vulnerable Populations

Vulnerable Populations

Justification Requirements for the Inclusion of Vulnerable Populations

1. * **How does the nature of the research require or justify using the proposed subject population:**
All records in this study are not considered human subjects for research under regulation 45CFR46.102(f). DNA from all suicide deaths or deceased non-suicide comparison cases collected through our collaboration with the Office of the Medical Examiner (OME) will be used, regardless of age at death.
To these biological samples, we propose to add: 1) pedigree structure for those cases in pedigrees as found by the UPDB (identifying information required for this linking will ONLY be transferred from Vital Records (OME-generated data) to the UPDB, and will not be used for the research), 2) demographic data and veteran status, 3) method of suicide (or method of death for non-suicide comparison deaths), 4) more detailed data for a 1-year cohort of suicide victims already collected by the Medical Examiner, 5) records information on subjects and family members obtained from the UPDB to indicate other suicides and/or hospital discharge and UUHSC EDW codes indicating approved co-morbid conditions. De-identified data from subjects in IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, IRB_00144804, and IRB_00133374 (linked to the UPDB through that protocol) will be linked to the present study within the UPDB. No identifying data used in the UPDB linking will be used in the present study. These subjects will have a score indicating response/non-response to PTSD treatment, or clinical course of COPD, or gastric bypass, or developmental delay, from these studies: IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, IRB_00144804. Limited-use data, linked by UPDB staff from our companion study focusing on diagnostic and demographic, data will also be available for analyses in this study.

NO CONTACT will be made with ANY family members of these deceased subjects for any additional direct information.

2. * **Would it be possible to conduct the study with other, less vulnerable subjects?**
○ Yes ● **No**

If yes, justify the inclusion of vulnerable subjects:

3. * **Is this population being included primarily for the convenience of the researcher?**
○ Yes ● **No**

If yes, explain:

IRB_00044244 **Created:** 8/17/2010 1:10 PM

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

IRB_00044244

4. Study Information

4. Study Information

1. **Design of Study (select all that apply):**

   ☑ **Non-Experimental and/or Descriptive Research Design:**

   Secondary/Archival Data Analysis or Retrospective Chart Review
   Survey/Questionnaire Research

   ☐ **Experimental and/or Interventional Research Design:**

   There are no items to display

   ☐ **Development of a research resource (repositories, databases, etc.)**

   There are no items to display

   ☑ **Other**

   If Other, describe:
   Genetic pedigree study; genetic case-control study

2. **Does your study involve the use of any placebo?**
   ○ Yes ● **No**

3. **Length of entire study, from initiation through closeout:**
   until 6/30/2026

4. **How will participants be recruited or identified for inclusion in the study?**

   a. **Select all methods that will be used:**

   Other

   Subjects are identified through ongoing collection of DNA and tissue on suicides or non-suicide deaths with or without evidence of prior attempt by the Utah Office of the Medical Examiner. Subjects are also linked through approved protocol IRB_00133374, IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, and IRB_00144804; de-identified records on these individuals are then studied by the research team.

   *note: IRB_00133374 is a sub-study of IRB_00044244 but Pedigree and Population Resource staff at UPDB may link new subjects under IRB_00133374. When this occurs, the IDs, minimal demographics, and the fact that a blood spot exists will be shared back to IRB_00044244.

   b. **Describe the recruitment/participant identification process in detail (e.g. who will review charts or records, who can refer participants to the study, where will flyers be posted, how often will recruitment letters be sent, when will follow-up phone calls be made, etc.):**

   Suicide deaths: Vital Records, which stores data from the Office of the Medical Examiner (OME) will transfer identifying information on unnatural deaths (includes the suicide cases to the Utah Population Data Base (UPDB). From the OME, some cases thought to be suicides are occasionally collected that were eventually designated as "undetermined" but at the time of collection the OME determined was highly likely to be a suicide, warranting sample collection.

   The participants from IRB_00133374, IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, and IRB_00144804 will be linked via the Master Subject Index to the UUHSC EDW and the Master Linkage File to

the IH EDW by the IH honest data broker.

Non-suicide deaths with and without prior attempts (comparison and control samples) will be studied as follows.

1) UPDBD PPR staff will identify **non-suicide** deaths back to 1995 using death certificate information. 2) Using this list of non-suicide deaths, the UUHSC Data Science Service Core and the IH honest data broker will identify the subset of deaths with UUHSC or IH EHR evidence of prior suicide attempt (ICD-9: 304.0*, 304.7*, 305.5*, 965.*, E850.0, E935.0, and ICD10 equivalents) to divide them into non-suicide deaths with vs. without prior suicidal behaviors. 3) For OME blood spots: blood spots are stored for all Utah deaths investigated by the OME back to 2005, so PPR staff will send identifiers of deceased non-suicides with and without attempts from 2005-present to the OME; OME staff will obtain 2 of the stored spots for DNA; these will be associated with numeric ids and will be picked up and logged using the long-standing system as samples for suicide deaths; DNA will be extracted by the CCTS DNA core and samples will be linked with other study data by PPR staff to create a limited use data set with the same protocols we have been using for suicides.

Limited use data will be studied:

All samples from this study are from either 1) suicide deaths, 2) non-suicide deaths with prior attempts (deceased comparison sample), or 3) non-suicide deaths without evidence of prior attempts (deceased controls). The research team will only receive the limited use data that matches the data elements for the suicides. Given estimated rates, we anticipate the total number of non-suicide deaths with prior attempts with useable samples to be several hundred per year; we anticipate the number of deceased controls to be >1000 per year. We have funding in hand to genotype up to 5000 of these comparison samples.

Identifying information from all participants will only exist at the UPDB, and within IRB_00133374, IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, and IRB_00144804. The study team will not receive this information.

Vital records sends all unnatural deaths to the UPDB; this includes a large number of auto accident deaths that are not linked because they are not suicides, so the Vital Records list is not identifiable. Even so, precautions will still be taken; the Vital Records list will go straight to the UPDB for linking via encrypted FTP transfer.

5. **How will consent be obtained?**

   Waiver or Alteration of Informed Consent

6. **Describe all the procedures chronologically, from screening/enrollment through study closeout, which will be completed in the research project.**

   We have already linked and identified high-risk pedigrees within the UPDB. We note that ongoing collection of suicide DNA and tissue will allow new UPDB linking annually.

   Death certificate information obtained from the OME from their master electronic database for all unnatural deaths is sent to the UPDB via secure FTP transfer to the highly secure servers of the Utah Population DataBase (UPDB) for record linking. The UPDB links information to all cases with suicide determination, and cases with undetermined or accidental designation, but which the OME determined were likely enough to be suicides that samples were collected. Linking also occurs to non-suicide deaths with evidence of prior attempt as described above.

   Records of non-suicide deaths with archived blood spotss will be similarly securely sent from OME to the UPDB using the Master Linkage File; details of the selection and linking are provided above. This deceased comparison group will be linked with the UPDB by PPR staff to the same study data as the suicide deaths.

   The UPDB has linked 99% of the suicide deaths to the UPDB master database and identified high-risk extended pedigrees containing a subset of these cases. Phenotypic data from Vital Records includes: age at death, year of death, race, gender, manner of death, diagnosis at death reported to the OME (if any), medications/drugs reported at death (if any), and OME-generated toxicology data. Year of death will allow us to investigate the influence of temporal effects on suicide, and geographical location of residence prior to death (available only to Vital Records and to investigators within the UPDB) will be useful for future studies of environmental risk factors. At the request of Vital Records, we now attach a list of variables that are transferred.

   We request to receive month and year of approved diagnoses, procedures, and medication data from the EM in the 2 years preceding the suicide death for the suicide and in that same time window for controls matched to the suicide. This detail will still not give exact encounter dates, so will preserve anonymity of cases, but will allow us to test risk prediction models that require knowledge of longitudinal events leading up to the suicide death.

   We additionally request a yes/no flag to be set by the UPDB for codes that occur within one week of the suicide death. Please note: this does not require us to know the actual suicide date, only that certain codes are a result of the

attempt (e.g., cardiac arrest, respiratory failure) and not a pre-existing condition. We will then have the option of deleting these codes in risk modeling.

Significance of demographic characteristics/co-occurring conditions/geocoded area will be determined by querying the ~9,000,000 records in the UPDB for matched Utah controls. The study team will analyze de-identified records data, and no living subjects will be contacted. Data from the OME regarding address of death will be linked to the UPDB; resulting geocoded data will be used by Drs. Bakian, Vanderslice, and Zhang to study spatial clustering of suicide deaths and attempts, and effects of pollution exposure linked to geocodes within the UPDB.

We continually link additional records of DNAs from suicides and non-suicide deaths with and without suicidal behaviors collected since this original linking. These records will be included and definition of high-risk pedigrees will be re-run, as described for the original sample in the methods below. We anticipate these records may add to existing high-risk pedigrees, may define new pedigrees, or will be added to the follow-up sample of cases not in high-risk pedigrees. We will also use de-identified cases from IRB_00133374, IRB_00082991, IRB00049521, IRB_00006824, IRB_00057455, and IRB_00144804 which will be linked within the UPDB, as described above.

In addition to the subjects described above, there are >20,000 suicides recorded in Utah death certificates going back to 1904 that were included to refine the familial risk estimates. We identified families with excess suicides (death certificate and OME), then filtered results to only contain families that were more clustered than expected by chance at a p value ≤ 0.01. We used all the data combined to compute Familial Standardized Incidence Ratios (FSIR; Kerber 1995; Boucher & Kerber, 2001), calculated by taking a cohort of each subject's relatives and checking the occurrence of suicide against population-based controls selected from the UPDB. We obtained 773 families that met the criteria above. We then eliminated families that did not contain 5 or more OME suicides (cases with DNA) which left a set of 184 families. These 184 families were sorted by their FSIR and the top 30 founders (highest FSIR) were selected for cluster analysis. Cluster analysis is a manual process that eliminates the multiple founders claiming the same descendants through marriage. We were left with 22 unique clusters with 5 cases with DNA, and 5 clusters with 3 cases with DNA. By sorting the population-based families on FSIR we obtained the highest clustering of suicide in families when compared to a random distribution. The incidence ratio is also adjusted for genetic (meiotic) distance between suicide victims assuming a simple single allele inheritance. Joining these high incidence families with the OME data insured that the maximum amount of DNA was available per family. The number of suicides per cluster in the 22 suicide families with at least 5 DNAs ranged from 10 to 51; numbers of suicides with DNA ranged from 5 to 11. There were a total of 468 suicides within these families, 138 with DNA. Their FSIR values ranged from 2.26 to 2.91, which implies the clustering of suicides in these high-risk suicide families is 2-3 times greater than expected by chance. For the additional 5 high-risk families with at least 3 DNAs, familial risk is 4-5 times increased.

Planned comparisons with relatives where genealogical records show no increased risk of suicide will be helpful in determining what patterns and degree of co-morbidity may indicate unique risk (see statistical analysis section below for a description of phenotypes). Some families do not show significant co-morbid risk for the categories of co-morbid diagnoses studied so far, but differences may be revealed as we investigate other demographic and co-morbid conditions comparing to pedigrees not at increased risk for suicide.

We also propose to enrich the pedigree phenotype information by obtaining hospital discharge codes and ICD-9/ICD10 codes from the UUHSC and IH EDW for all conditions on cases and family members. Aggregation of conditions will be done through cluster analyses, informed by our expert clinical colleagues' knowledge of symptom clusters. Matching of ICD9/ICD10 data will be done as follows.

a. Cases, comparison deaths, and control deaths are identified by the Office of the Medical Examiner (OME), samples are collected and anonymous numeric id assigned to the sample, and there is an OME case number ID.

b. Once a year, samples are linked to the UPDB; the OME transfers identifying data (this does NOT go to the research team) allowing the UPDB to link cases.

c. A master ID list of samples, comparison cases, and controls exists at the UPDB; this includes cases with samples, but also cases identified as suicides from the death certificate where sample collection was missed; these additional cases allow a more complete data set for epidemiological studies, and to fully characterize pedigree risk.

d. The UPDB determines high-risk pedigrees, and determines age/sex/birth-year matched population controls for the cases; these IDs are also known to the UPDB.

e. The UPDB will give the list of IDs of suicide cases, their family members, and age/gender/birth-year matched population controls to the EDW honest data brokers. NOTE: because the list will include ~20,000 suicides but also ~60,000 family members and up to ~120,000 matched control cases and their family members for a total of ~200,000, identity of the suicides will be protected. Similarly, comparison non-suicide deaths and controls will be matched to EDW data.

f. Only records within this group of cases/family members/comparisons/controls that have ICD9/ICD10 codes will be returned to the UPDB for linking.

g. For the suicide cases, comparisons, and controls, researchers are given month and year of first diagnosis, and number of times that diagnosis was given. Additional detail of month/year of codes for the 2 years prior to death for the suicides and matched population controls will be needed to accurately model risk from these data, as longitudinal sequence of events prior to death is required. Finally a yes/no flag is required for codes assigned within 1 week of death. This does not require us to know the actual suicide date, only that certain codes are a result of the attempt (e.g., cardiac arrest, respiratory failure) and not a pre-existing condition. We will then have the option of deleting these codes in risk modeling. All PHI and codes that allow linking within the UPDB are stripped before data are given to the analysis team.

h. No individual-level family data is given; limited-use control data are used for epidemiological comparison studies.

Our analyses currently focus on affective disorders, psychosis, self-injury, impulse control, substance abuse, alcoholism and PTSD (records-only analysis). Note that peri-/post-partum depression is also included. Women who commit suicide within 1-3 years of delivery (or fetal death) will also be added to those with peri-/post-partum depression codes as additional potential cases of pregnancy-related depression, allowing more powerful studies of this clinically important subgroup. We will be using the Master Patient Index that links the demographic records from the UUHSC and IH EDW to the UPDB, under IRB_00045234, “Master subject index between the Utah Population Database and the University of Utah Health Sciences Center”. Comorbid conditions indicated by these information sources will be used to stratify high risk families, and to identify specific familial risk of other conditions related to suicide. These risk analyses will be performed by the UPDB personnel. Resulting information will be used to prioritize and direct genetic analyses in the families. Again, no attempt will be made to identify family members. Our new pipeline precludes the list of specific codes (see the numbered outline of the pipeline above). For PTSD, we will additionally study ICD-9/ICD10 child maltreatment codes (995.x), suspected abuse codes (V71.x), family problems codes (V61.x), injury codes (E800-E999.1), and codes related to pain, which indicate trauma exposure. The procedure code for electroconvulsive therapy (ECT; 94.27) will be useful to identify cases who had treatment resistant depression. We also request BMI from driver's licenses to more accurately characterize longitudinal BMI as a risk factor.

As we have worked with our data, we have become aware that because suicide is so under-studied, there is really only a very narrow understanding of what conditions might be associated with high risk subsets. Even since our study began, we have noted important work implicating not just psychiatric risk, but also inflammatory conditions, autoimmunity, and other interesting health conditions as possible risks of suicide. While we have tried to tap these categories to explore this in our own data, we believe that the cases themselves could much more accurately tell us which conditions and codes should be the focus of our studies. For this reason, we now request to study all conditions defined by the cases. Requested use of selected prescription medication data also available in the UUHSC and IH EDW will allow a second piece of data to study psychiatric, pain, and inflammatory diagnoses from the electronic records. These prescription data will allow exploration of association of medications with specific mechanism of action with suicide risk, controlling for demographics, co-occurring conditions, and/or exposures. Studies including psychiatric medications will initially focus on selective serotonin reuptake inhibitors (SSRIs). Studies including pain medications will initially focus on opiates. Studies including inflammatory medications will initially focus on mediation/moderation effects of medications on short-term exposure to air pollutants. In each of these studies, comparisons will be cases without medication and/or cases with other related classes of medications with other mechanisms of actions.

We have methods to define significantly associated co-occurring conditions, including: using pedigree analysis methods to identify clusters of familial cases for gene discovery; using classification analysis and cluster analysis to determine significantly over-represented diagnoses (these methods include hierarchical clustering, centroid models such as k-means, distribution models, graphical models, recursive partitioning). Medication data can serve as secondary data points in classification and/or an independent risk variables and/or mediators/moderators of risk. We are collaborating with experts at the Scientific Computing Institute (SCI; Dr. Lex) to develop data visualization tools to better understand the complex patterns in our data.

We will compare high-risk pedigrees to pedigrees of suicides where no familial aggregation was found in the UPBD. In this protocol, these analyses will focus on identifying the most compelling pedigrees for genetic studies. We will also compare to anonymous matched controls taken from the UPDB by the UPDB staff scientists working on this project to obtain aggregated frequencies of conditions/procedures/demographics. Matching may include BMI from drivers license.

Our studies include analyses of subsets of cases defined by our collaborators (e.g., obesity/bariatric cases; PTSD/trauma; early life risk factors; air pollution exposure; COPD; developmental delays; peri-/post-natal depression). Data elements to define these subsets include cross-matching then de-identified analyses of cases in collaborators' studies and/or sets of ICD9/ICD10 codes, other data elements such as air pollution data now linked to the UPDB, early life risk factor scores, birth certificate data, fetal death data, infant death data needed to determine peri-/post-natal depression risk. Data elements are included in the table below.

Aim 2. Genotyping and genetic analysis. Ongoing DNA collection from decedents continues to add to this sample; ~700 new DNAs are collected every year. For these deaths, we will also have access to the same demographic data, diagnostic co-morbidity data through the UPDB, and the OME phenotypic database. From our non-familial case resource, we will prioritize samples with the greatest volume of DNA and the most available phenotype data for genotyping. All samples will be subjected to initial testing for quality and volume. Any sample below a total volume of 10ug of DNA will be subjected to whole genome amplification (WGA) to increase stocks prior to beginning our experiments. This threshold will leave us with a quantity of un-amplified DNA with which we can re-check results obtained on amplified DNA. When WGA is needed, we will use Multiple Displacement Amplification (MDA), a non-PCR based DNA amplification technique (Dean et al., 2002). This method can quickly amplify very small amounts of DNA to larger workable quantities for genomic analysis. Briefly, the MDA reaction starts by annealing random hexamer primers to denatured DNA: DNA synthesis is then carried out at a constant temperature by a high fidelity enzyme (Phi29 DNA polymerase) to produce a DNA product of 7kb to 10 kb long. Phi29's high fidelity and 3'-5' proofreading activity reduces the amplification error rate to 1 in 10^6-10^7 bases, superior to other PCR based WGA that use Taq polymerase (Telenius et al., 1992; Zhang et al., 1992). Phi29 also outperforms Taq in resulting lengths of fragments, accuracy of DNA amplification, and completeness of loci coverage (Telenius et al., 1992; Zhang et al., 1992). The MDA-generated DNA fragments can be directly used for genotyping, real time PCR, sequencing and haplotyping (Shoaib et al., 2008). We now have preliminary experience with these techniques on 12 control DNA samples, and have achieved excellent amplification results, with an average 10-fold increase.

Once samples meet quality/quantity standards, they will be aliquoted for the University of Utah Sequencing core for whole genome and whole exome sequencing. Mitochondrial genotyping will be done at the University of Utah Core Genotyping Lab. We have completed a pilot sample of 60 subjects, and are exploring variants associated with altered brain pH and with psychiatric conditions.

In the past, we participated in the University of Utah Heritage 1K project, sending samples in high risk pedigrees for whole genome sequencing. The study was funded by the Chan Soon-Shiong Family Foundation, the Chan Soon-Shiong NantHealth Foundation, and the Chan Soon-Shiong Institute of Molecular Medicine. Details of the sample transfer appear in the attached Memorandum of Understanding. Briefly, samples were sent along with the study title and the degree of relatedness of individuals (first, second, third, fourth, or fifth or greater). No pedigree drawings or other files that indicate exact pedigree structure were sent. Scientists affiliated with the donor foundation destroyed all data; it is retained for our use on the CHPC Protected servers. Publications resulting from data through collaboration will include investigators from this study, and will be approved first by the RGE.

We had in the past a similar collaboration with Janssen Research (investigator Qingqin Li). These investigators assisted us with the following tasks: study risk factors using epidemiological analyses and/or genetic analyses. Data used by Janssen investigators included: de-identified ids of cases, and associated phenotype data (age at death, method of death, gender, race, diagnostic data). Note that ICD9/ICD10 data had month/year of first diagnosis, and number of times that diagnosis was coded, rather than explicit dates of all diagnoses. Similar limited use data was used on matched controls. For family of cases/controls, ONLY aggregated familial risks of co-occurring conditions was used. When cases were related to other cases, investigators were made aware of familial relationships using the model approved for data sharing for the Utah Genome Project H1K initiative, as follows. Cases who are in an extended family were identified by an anonymous numeric id. One case was randomly selected as proband, and degree of relatedness between this proband and other cases in the family will be given. In this way, specific family structures will be not be disclosed, per RGE regulations (http://rge.utah.edu/policy_updb.php#H5). All publications resulting from this collaboration were first approved by the RGE.

Regarding the U of U sequencing, our pipeline will use locally developed tools (bioserver.hci.utah.edu) to address variant calling and quality analysis (QC). QC is essential for determining which samples and genes are reliably assayed and which variants are true segregating sites; the statistical analysis strategy must be well powered for different possible models by which variation can influence risk. The interpretation of the results will need to span beyond individual variants to evaluation of groups of variants across individual and evidence-based sets of genes. Our software supports tasks from initial alignment of sequence to reference to final annotation and output in ANNOVAR or VAAST formats.

Genetic analysis to prioritize variants will be done as described in the analysis section below, and will also benefit from studies using tissue samples (RNA gene expression; protein expression; methylation; synapse structure, density, specificity; synapse function).

Aim 3. Confirm top gene findings in original subjects using Sanger sequencing. Extend gene results for 2-3 genes per year through targeted re-sequencing in up to 1000 additional selected suicides per gene. Select cases by linking new Utah suicides with DNA collected during the project to the UPDB to 1) extend the known high-risk pedigrees, and 2) determine new high-risk pedigrees. For each prioritized gene, re-sequence new familial suicides, and suicides in the large UT cohort who share highly similar phenotypic profiles to the original cases identified in Aim 2 who have confirmed variants. This re-sequencing is done to identify more complete genetic variation within the prioritized genes of interest.

Methods for this aim are primarily detailed in the statistical section below. Sanger sequencing for validation will follow current standards set in the research community; advances in technologies will be incorporated in our studies as they are developed during the funding period. Sanger sequencing (Chou et al., 1996) will be used for validation of

prioritized variants in individuals identified from Aim 2 sequence studies (3-10 suicides for each of ~50 top variants). PCR primers will be designed using the Primer 3 algorithm in order to amplify the DNA surrounding each variant. PCR products will be bi-directionally sequenced using an ABI 3730 fluorescent DNA sequencer. DNA sequence analysis will be performed using Sequencher 4.10.1. Validation will help with further selection of genes in which to detect additional variation. Once 500-1000 decedents are selected for each gene, sequencing across the genes will be done. We assume an average of 10 amplicons per gene, and 2 amplicons for promoters. Primers will be designed, and PCR products will be bi-directionally sequenced as described above. With current methods, we propose to test for additional variation for 2-3 prioritized genes per year, though improvements in efficiency and lowering costs may allow additional genes and/or additional samples to be tested. University of Utah/ARUP labs have experience with sequencing approaches (Margraf et al., 2011; Dames et al., 2010).

In addition to local storage at the University of Utah, samples and data will be shared with the NIMH repository as mandated by our funding. Aliquots of DNA samples will be sent to the Rutgers University Cell and DNA Repository, managed by the NIMH. Samples will be sent annually during the NIH grant award (year 1, current sample; end of year 5, additional samples to complete the full sample ascertained through the course of the funded award). Samples will ONLY be tracked with an anonymous numeric ID. Samples will be sent via overnight FedEx shipping.

Aggregeated, fully de-identified phenotype data associated with these samples will also be sent and will include: suicide designation from the Utah Office of the Medical Examiner autopsy, gender, age at death, race, method of death, yes/no status of diagnoses derived from ICD-9/ICD-10 codes approved for this project indicating co-morbid conditions (currently: affective disorders, drug disorders, alcohol disorders, psychoses, self-injury, PTSD); pedigree cluster anonymous numeric ID indicating which pedigree a sample is part of, if any. Data will be sent using a fully encrypted external hard drive. No detailed pedigree data will be shared. Data and sample use are controlled by application/approval to the NIH. Further use of detailed pedigree data could only occur with explicit approval of the RGE Committee.

All sequence and genotype data generated using NIH funding during the award will be deposited in dbGAP, in accordance with NIH policy. These data will again contain no phi. Data will be sent using fully encrypted external drives.

Additional studies complementary to the ongoing studies using the extended family design are now funded. These include genome-wide association studies using case-control comparisons, and risk prediction modeling using electronic health records and genetic data. We have added an important comparison group of non-suicide deaths with evidence of suicide attempts to greatly enhance our ability to compare the risks of attempt vs. death in the context of family studies, case-control genome-wide association studies, and risk prediction.

Tissue collection procedures: we will collect and bank skin biopsies from each case for which blood for DNA is obtained. A site under the arm will first be disinfected with an alcohol swab, then a sample will be collected using a standard skin punch. Samples will be frozen viably. We have grown cell cultures successfully on 12 fresh tissue samples, and have verified that frozen tissue can be thawed and successfully grown in culture. All tissue samples are stored in secure locked freezers with alarm systems in the University of Utah Psychiatry lab. The resource will allow us to generate human induced pluripotent stem cells (iPSCs) that can be differentiated into neurons from suicide victims (and corresponding controls) where the suicide cases have been identified as having clear genetic or cellular defects. This will allow us to examine neuronal form and function when comparable underlying genetic mutations are observed in DNA. iPSC-derived neurons will be used in future experiments for 1) electrophysiology and calcium imaging to identify precise functional defects, 2) rescue experiments in which mutated genes are re-expressed in neurons to determine if functional defects can be rescued, and 3) drug screening to identify potential new therapeutics. All will increase our ability to validate genetic risk factors. We will bank viably frozen skin biopsies, and create fibroblast cultures from selected deaths as needed for specific hypothesis testing.

Brain tissue collection is now closed. Remaining brain tissue from this small collection is stored in our lab with anonymous numeric identifiers. Because collection of whole brain with next of kin consent was difficult, we collected and studied small samples of brain tissue that will be available when autopsy already dictated the examination of brain tissue. These small samples were taken after the autopsy studies are complete.

Hair Collection: The OME collects hair samples from the posterior vertex scalp region, and hair is stored with the root end clearly marked wrapped in foil inside an envelope. Hair samples will be de-identified and sent to the Mount Sinai HHEAR Targeted Analysis Laboratory for the analysis of environmental analytes.

Methods: RNA expression. High quality RNA will be aliquoted for the University of Utah Microarray and Genomics core (www.cores.utah.edu) for quantitative RNA-Sequencing utilizing the Illumina HiSeq 2000 platform. The core facility generates the cDNA library from total RNA using the Illumina ScriptSeq V2 library preparation kit to generate barcoded 50nt single end reads. Barcoding samples will allow us to run 12 samples per lane, generating an average of 16 million reads per sample. Fastq files will be assessed for quality using FastQC, then aligned to the reference genome (hg19, www.ensembl.org) using TopHat2, which allows for the alignment of reads spanning exons (Trapnell et al., 2009).

Methods: development of protocols to study neuronal morphology, specificity, and density. Individual neurons in fixed tissue will be microinjected with fluorescent dye using iontophoresis, which labels the entire dendritic tree including

spines. Neurons will be imaged on a confocal microscope (this equipment is already available to us). Dendritic length/branching and spine density/shape will be analyzed using NeuroLucida360 software. Methods to detect density of specific classes of synapses will be piloted using immunostaining with antibodies that label highly specific classes of synapses. Our collaborator Dr. Megan Williams pioneered methods to analyze 6 different types of hippocampal synapses at a time using a combination of presynaptic markers (Williams et al., 2011). Immunostaining results can be combined with morphological experiments to determine exact type of synapse formation on particular neurons.

IPSC studies using fibroblast tissue. Human induced pluripotent stem cells (iPSCs) can be prepared from skin fibroblasts using well-established protocols (Nakagawa et al., 2008; Junying et al., 2009). The banking of fibroblasts will allow us to grow neurons from suicide victims (and corresponding controls) that have clear genetic or cellular defects. It will also allow us to examine neuronal form and function from a larger sample for which brain tissue is unavailable, but where comparable underlying genetic mutations are observed in DNA. iPSC-derived neurons can be used in future experiments for 1) electrophysiology and calcium imaging to identify precise functional defects, 2) rescue experiments in which mutated genes are re-expressed in neurons to determine if functional defects can be rescued, and 3) drug screening to identify potential new therapeutics.

DATA TABLE:

UPDB Data Source	**Data Elements to be Accessed**	**Scientific Justification for Access**
Demographic Data	Gender, Race, Ethnicity, Marital Status, Education Level, Industry/Occupation Codes, Year of Birth, Year of Death, Age at Death, Birth Month/Year, Death Month/Year	Demographic data is necessary in comparing subjects to matched controls and in determining familial risk patterns. Average age at death from all causes (excluding suicide cases) within each family will be used as indicator of serious chronic disease.
Geographic Groupings	UDHHS Small Health Statistical Area, County (>20,000), US Census Block, Zip Code	Geographic location will be used as a basis for environmental exposure.
Family Relationship Information	High risk extended pedigrees	We are looking for families with excess suicides.
Birth Certificate Details	Age of parents, Geocodes, details of birth of child of female suicides (for studies of peri-/post-partum depression), including risks and complications with the pregnancy, congenital malformations or anomalies of the child as they appear on the birth certificate either in text, ICD9 codes or check boxes	We are studying early life stressors and later life health outcomes; exposures associated with place of birth could be important to this work. Stressors/exposures will be compared to de-identified matched controls. For the parent study, control cases and their families are chosen from the 9 million records in the UPDB, matched for age at death, gender, and birth year. For the peri-/post-natal risk study, control selection will be similar, but matching will instead be on gender, birth year, and age at delivery of the child.

Death Certificate Details	Primary and secondary causes of death, Geocodes	Identifying suicides from death certificates is key to this study. Causes of death will help define familial risks of co-occurring conditions. Also, Drs. Bakian, VanDerslice, and Zhang are studying acute environmental exposure as a risk mechanism in suicide. Exposures will be compared between de-identified cases and matched controls. In addition to death. Finally, death of a child of a female suicide case within 1 year of delivery is needed as it is an important risk factor for peri-/post-partum suicide.
Fetal Death Certificate	Primary and secondary causes of death; month and year of death	We are investigating risks of peri- /post-partum depression in the female cases in our study. To make this analysis more complete, we will request fetal death records, as miscarriage may be a significant trigger for post-partum depression. Data requested includes: risks and complications with the pregnancy, congenital malformations or anomalies of the child as they appear on the birth certificate either in text, ICD9 codes or check boxes
Driver License Division	Height/Weight (BMI)	We request BMI from driver's licenses to more accurately characterize longitudinal BMI as a risk factor
Health Facility Data (Inpatient, Ambulatory surgery, **Emergency Department)**	Conditions associated with affective disorders (including peri-/post-partum depression), psychosis, alcohol abuse, drug abuse, PTSD/trauma including traumatic brain injury, other psychiatric conditions, asthma and other lung conditions, severe obesity, epilepsy, inflammatory bowel syndrome, cardiovascular diseases, conditions associated with chronic pain, immune/autoimmune conditions, dementia/neurodegenerative conditions. (Codes included in documents). Diagnosis/Procedure Age, Diagnosis/Procedure	These comorbid conditions will be used to determine familial risk patterns. Occasionally diagnoses/procedures/trauma happen within the same year; month will allow us to determine order of these important risks. 12/2016 Amendment allowing all codes on cases and family members will provide opportunity to potentially identify novel co-occurring risk factors. Cluster analysis techniques will be used to identify and aggregate significant co-occurring conditions.

	Month/Year Amendment with 12/2016 renewal: For selected ids representing suicides, relatives of suicides, controls, and relatives of controls, all codes will be provided. Diagnosis/Procedure age, Diagnosis/Procedure month/year	
UUHSC / IHC	ICD-9 Codes for affective disorders, psychosis, self-injury, impulse control, substance abuse, alcoholism, PTSD, and peri-/post-partum depression. (Codes included in documents). Diagnosis age, Diagnosis month/year Amendment with 12/2016 renewal: For selected ids representing suicides, relatives of suicides, controls, and relatives of controls, all codes will be provided. Diagnosis age, Diagnosis month/year	These comorbid conditions will be used to stratify high risk families and to identify specific familial risk of other conditions related to suicide. Occasionally diagnoses/procedures/trauma happen within the same year; month will allow us to determine order of these important risks. 12/2016 Amendment allowing all codes on cases and family members will provide opportunity to potentially identify novel co-occurring risk factors. Cluster analysis techniques will be used to identify and aggregate significant co-occurring conditions.
US Census Data for Utah	Sibship size, SES (as indexed by the maximum achieved occupation code), number of offspring, marital history, geographical location as basis for environmental exposure	For older cases in the cohort, these factors will be approximated using census data
Prescription medication data	Prescription data will be linked to cases within the UPDB; only limited use data will be released to the study team for analysis	Prescription data will provide: 1) a second data point to validate of ICD9/ICD10 diagnosis; 2) data to study characteristics of individuals who have been prescribed opiates who have gone on to die by suicide.
All Payer Claims Database (APCD)	All ICD9/10 Diagnosis/Procedure age, Diagnosis/Procedure month/year Prescription (medication prescribed and filled) data will be linked to cases within the UPDB; only limited use data will be released to the study team for analysis	APCD data will provide more comprehensive diagnosis/procedure and medication information. Uses will again be for: 1) a second data point to validate of ICD9/ICD10 diagnosis; 2) data to study characteristics of individuals who have been prescribed opiates who have gone on to die by suicide.

Samples and associated summary data (genotypes/sequence data, sex, age at death, suicide determination, aggregated yes/no psychiatric diagnosis from the EMR) will be sent to the NIH repositories (RUCDR for samples, NIMH Data Archive--NDA--for data). An anonymous id (the Global Unique Identifier, or GUID) will be created using NIH-developed technology, as follows.

The NIMH requests that we follow a specific process to generate de-identified unique ids for the repository, called Global Unique Identifiers (GUIDs). Because this process involves local use of PHI input into a NIH GUID tool to generate a de-identified hash for each case, which is then encrypted and sent to the NIH, we request that these ids be created by Pedigree Population Resource PPR staff within the Utah Population Database. We request creation of the GUIDs to follow this protocol:

Approval will be secured from the NIH to download a local copy of the GUID tool, which will be run locally, not connected to the internet. The research team will send a list of suicide case numeric anonymous abids (generated by the CCTS DNA core) to PPR staff corresponding to suicide cases with samples/data at the NIH repository. PPR staff will enter into the GUID tool the labids and a list of PHI for each case from birth certificate data, including first name, middle name, last name, sex, date of birth, and place of birth, as requested by the NIH GUID specifications (https://nda.nih.gov/contribute/using-the-nda-guid.html). The GUID tool will create a de-identified hash of these data elements. The algorithm to create the hash has been developed by NIH data scientists to be non-invertible, meaning it cannot be reverse-engineered back into PHI. Resulting de-identified hashes will then be encrypted by the GUID tool. PPR staff will send the encrypted hashes to the NDA staff at the NIH. NDA staff will assign unique GUIDs to each hash and return them to the PPR staff.

NDA staff will also compare hashes to those already generated for all other samples at the NIH repository across all studies. While unlikely, it is possible that a suicide in our study previously participated in a research study in some other location, and has a sample or data attached to a hash already at the NIH. If this is the case, the NDA staff will send the existing GUID back to PPR staff rather than generating a new GUID. This protocol ensures that samples/data from the same case will not exist at the NIH repository associated with more than one id.

Data to be used for this system will remain within the UPDB and will not be used by the research data analysis team for any research purpose.

If any data elements required by the GUID tool are missing (for example, middle name) the GUID tool allows for entry of a missing data code; this will not affect the generation of a hash. For the small number of suicides where the UPDB cannot link to a birth certificate, case ids will be sent separately to the NDA staff and a pseudo-GUID will be generated using a random process that will not allow comparison to existing repository samples, but will at least result in a unique id in the NIH repository.

We are requesting this change because the NIH requires this process to create unique, de-identified GUIDs for samples in the NIH repository. As stated in the NIH instructions for GUID creation (https://nda.nih.gov/contribute/using-the-nda-guid.html):

"Using such a system, the same participant information will return the same GUID whenever or wherever it is entered. This allows NDA to anonymously link participant data records across time and locations, without ever receiving identifying information. The ability to link subject records and the protection of participant confidentiality are both critical components of data sharing."

7. **Are all procedures for research purposes only (non-standard or non-standard of care procedures)?**
   ● **Yes** ○ No

   If no, list the procedures that are performed for research purposes only (non-standard or non-standard of care procedures):

8. **Is there a safety monitoring plan for this study?**
   ○ Yes ● **No**

9. **Provide a summary of the statistical methods, data analysis, or data interpretation planned for this study. Factors for determining the proposed sample size (e.g., power) should be stated.**

   Identifying high-risk pedigrees. Pedigrees were identified as described above. Pedigrees that do not show increased risk for suicide will be determined similarly, but for these families, FSIR will not exceed 1 (normal population risk; no aggregation). These “non-familial” pedigrees will include at least 2500 total relatives to ensure that the FSIR is estimated with accuracy.

   Phenotypes. Phenotype data are available from two sources. First, data are available through the OME via Vital Records for suicides (and for any additional decedents on whom DNA was collected as being likely suicides), including age at death, race, gender, *method of death*, diagnostic information at death reported to the OME, medications/drugs reported at death, OME-generated toxicology data, and residence at time of death. NOTE: Residence data will not be released outside the Vital Records/OME/UPDB, and will only be analyzed by collaborators

within these entities. Data exist in electronic form in the Vital Records/OME database system, and can be associated with cases with DNA using an anonymous numeric ID. In addition to OME phenotype data, we will extract demographic and electronic diagnostic information from the UPDB to determine familial risk patterns. Diagnostic codes studied to date include conditions associated with affective disorders (including peri-/post-partum depression), psychosis, alcohol abuse, drug abuse, PTSD/trauma including traumatic brain injury, other psychiatric conditions, asthma and other lung conditions, severe obesity, epilepsy, inflammatory bowel syndrome, cardiovascular diseases, conditions associated with chronic pain, immune/autoimmune conditions, dementia/neurodegenerative conditions. Female suicide within 1-2 years of delivery will be studied as a comparison group of more formally diagnostically defined peri-/post-partum depression. BMI from driver's licenses will be used to create a longitudinal description of this potentially important risk. Aggregate risk of these co-morbid conditions to 1st and 2nd degree relatives of all suicide cases in our data resource are significantly elevated over population risk, with odds ratios between 2.13 and 4.45 for 1st degree relatives ($p<1x10^{-16}$) to odds ratios of 1.52-2.46 in 2nd degree relatives ($p<0.001$). Many of the pedigrees also show significant elevation of co-morbid risk within the family. Co-morbid conditions may indicate particular genetic susceptibility factors.

NOTE: in order to determine significance of risk in suicide cases, the >9,000,000 UPDB records must be searched for matched controls. All data are de-identified and no living subjects are contacted.

Phenotypes will include those described above, and also: year of suicide (for possible temporal effects), sibship size, SES (as indexed by the maximum achieved occupation code), number of offspring, marital history, and geographical location as the basis for potential environmental exposure. For older cases in the cohort, these factors will be approximated using census data. Note that geographical location data will not leave the UPDB; spatial analyses will be done with the collaboration the UPDB staff. We will study average age at death from all causes (excluding suicide cases) within each family, as an indicator of serious chronic disease. Beyond the psychiatric conditions already approved, we will study conditions that are also known to be strongly associated with suicide in the absence of psychiatric diagnoses listed elsewhere in this application, and will use machine learning analyses to potentially discover new associations with medical conditions. A measure defining response vs. non-response to PTSD treatment will also be linked to records coming from IRB_00082991; clinical course of COPD will be linked to records coming from IRB_00049521; gastric bypass cases will be linked using records coming from IRB_00006824; developmental delay status will be linked using records from IRB_00057455.

Residential address at time of the decedent's death will be used by the UPDB for geocoding to latitude/longitude. (NOTE: original residential address will not be kept by study investigators). Geocode data will be used by Drs. Bakian, VanDerslice and Zhang for studies of spatial clustering and association of risk with altitude. Methods will include the application of adaptive kernel density functions, with one or more smoothing bandwidths which may vary depending on the spatial extent of the cluster. Application of these methods has been previously used in a de-identified clustering study of autism spectrum disorder in Utah, and results were presented in a sensitive way such that identities of cases within clusters were protected (Bakian et al., 2015). We will adhere to this careful reporting practice, and all manuscripts stemming from these analyses will undergo RGE review. In addition, geocodes will be used within the UPDB as the point location at which ambient air pollution exposure will be assessed. A time-stratified case-crossover design will be used to assess the association between short-term ambient air pollution exposure and risk of suicide. Census block and tract-level information describng SES will also be analyzed as a potential moderator of the relationship between short-term ambient air pollution exposure and suicide risk. The level of geographic data that will be analyzed therefore begins with latitude/longitude geocodes. However, geocode data may be aggregated to larger areal units (census blocks/ tracts or larger).

Because we will also include more broad occurrence of all ICD9/ICD10 data, we must explore significance of occurrence of co-occurring conditions in our data resource. In genetic analyses focused purely on suicide risk in extended families, diagnostic data will be used as a post-hoc descriptive resource. It is unknown if risk variants will show such specificity, or if genetic risk may cut across co-occurring conditions. It is unknown if there are specific patterns of co-occurring conditions that may be more likely to be present with particular mutations. Importance of co-occurring conditions may be obvious based on the nature of risk variants, or may not be clear until future bench studies of gene function have been done.

In addition to our descriptions of cases in high risk families who share genetic risk variants, we will also take full advantage of co-occurring conditions to try to actually select and prioritize unusual families for study. We will use cluster analysis and known clinical groupings of diagnoses to aggregate the diagnostic data. We will then use the familial standardized incidence ratio (FSIR) of each aggregated diagnostic cluster to determine if some high-risk suicide families are also at higher risk for co-occurring conditions. Knowledge of conditions within families can direct genetic results through suggesting gene pathways that may lead to familial co-occurring conditions. Follow-up studies will require replication using data from our many colleagues in consortium groups (Psychiatric Genomics Consortium), and we will again turn to our bench science colleagues for future functional/animal studies.

Finally, we will also use epidemiological methods to determine increased likelihood of co-occurring conditions, with familiality of suicide and the co-occurring conditions as a critical clustering factor. The only previous well powered studies to investigate these questions have only surveyed a small number of potential co-occurring conditions, have not had deep familial data, and/or have not had DNA to allow follow-up of phenotypic associations. We have several methods at our disposal to define significantly associated co-occurring conditions. These methods include hierarchical clustering, graphical models, and recursive partitioning. Because our data set is so large (>20,000 suicides >12,000 of which are recent enough to have diagnostic ICD9/ICD10 data, and over 2000 families with significant FSIR for suicide), we can employ internal replication techniques, splitting our sample into discovery and replication cohorts. We are pursuing this

line of work in addition to prioritizing the highest risk suicide pedigrees, independent of co-occurring conditions. We plan to allow results from both lines of work to inform each other.

We plan to use prescription medication records for the study of medications for psychiatric conditions, pain, and inflammation that are currently part of the UUHSC EDW and IH EDW. These records will allow the following: 1) provide a second data point to validate of ICD9/ICD10 diagnosis, a particularly important issue for psychiatric diagnoses where prescribed medication can provide essential insight into specific psychiatric diagnoses based on drug mechanisms; 2) determine if (as some prior research has indicated) particular psychiatric medications, in combination with specific diagnoses and demographics, may lead to increased suicide risk (e.g., SSRIs); 3) provide data to study characteristics of individuals who have been prescribed opiates, in comparison with use of other pain medications, who have gone on to die by suicide; and 3) provide information about medications associated with inflammatory conditions which will be essential in ongoing studies co-occurring suicide-inflammation risk (e.g., asthma risk, cardiovascular risk, inflammatory metabolic risk)--these studies can use medication data both for confirmation of diagnosis and to determine if medications with particular mechanisms of action (in combination with other diagnoses, demographics, and exposures) may be specifically associated with increased suicide risk.

Analysis methods for comparing high-risk and non-familial pedigrees. Correlations among demographic variables will be addressed by using principal components analysis (PCA) on state-wide data within the UPDB to create aggregate an factor, with confirmation using multivariate mixed models. We anticipate potential strong associations among our chosen demographic measures; previous experience suggests these characteristics may be best captured by one or two factors. Codes for medical conditions will be clustered by broader categories as defined by the International Classification of Diseases (ICD) system (www.cdc.gov). We will test differences between high-risk and no-increased-risk pedigrees using mixed model regression analyses (SAS PROC NLMIXED , www.sas.com), to account for the within-pedigree effects (random) and between-group effects (fixed). We anticipate finding significant differences across groups defined by high-risk and non-familial suicide pedigrees, and will prioritize individual compelling large pedigrees for Aim 2.

We will cross-check our identification of significant risk of co-morbid conditions in the high-risk suicide pedigrees. We will use de-identified electronic diagnostic data, define high-risk pedigrees for the co-morbid condition under study, and determine the extent to which these pedigrees may be at increased risk for suicide. By using this parallel strategy, we may do one more of the following: 1) identify the same set of high-risk pedigrees at significant co-morbid risk that were already identified by starting with the suicide pedigrees; 2) identify other high risk pedigrees that may indicate suicides similarly at high risk for the co-morbid condition. If (2), these new cases will be used in our genetic studies as extensions/ replications for primary studies of suicide risk factors specific to co-morbid conditions.

Power. We approximated power by assuming a logistic response (high-risk vs. non-familial group difference) with response probability of 0.5, a normally distributed quantitative risk predictor, and a normally distributed covariate. Sample sizes were conservatively estimated at 200-300 (individual families range in size from 1000–5000 relatives in high-risk pedigrees, and >2500 in “non-familial” pedigrees). Power to determine group differences in these analyses ranged from 0.75 – 0.90, indicating satisfactory power to detect significant differences even at the level of individual pedigrees.

Pedigree data will be analyzed using several available tools. We will focus on methods where we have local expertise (e.g., SGS and VAAST), but will also test other approaches (e.g., Condel, Gonzalez-Perez & Lopez Bigas, 2011).

Shared Genomic Segment (SGS) analyses. Variant calls from sequence data and chip single nucleotide polymorphisms will be used for SGS. SGS was developed by co-investigator Dr. Nicola Camp and her colleagues to identify genomic segments shared in excess across multiple individuals (Thomas et al., 2008). A pedigree SGS analysis considers sampled, distantly related cases (without genotyping from connecting relatives) and poses the question as to whether the length of consecutively shared loci (identified as identical-by-state, or IBS) is longer than expected by chance. IBS is established simply by determining if allelic types at sequential loci are consistent (phase is ignored). IBS does not infer identity-by-descent (IBD; the same inherited segment from a common ancestor) which is our true interest. However, if the length of SGS shared IBS is significantly longer than by chance (given the known relationships) then IBD is suggested. Theoretically, chance IBD sharing in distant relatives is extremely improbable. Genomewide statistical significance can be found for pedigrees with at least 15 meioses (genome-wide p»0.05). Our high-risk pedigrees have far more than 15 meioses between studied cases.

A brief description of the SGS method is as follows. At any variant position, assume the alleles are labeled 1 and 2. Sharing can be identified based on the possibility (or impossibility) that an allele is shared IBS across all cases. Sharing is impossible at a SNP position if the set of genotypes for the cases contains both a {11} and a {22}. Otherwise sharing is possible. Each SNP position is assigned the value that indicates the length of sharing between all cases that encompasses that SNP. For example, if sharing is possible for a stretch of *N* consecutive SNPs then those *N* SNPs would each be given the value *N*. Due to the possibility of sporadic cases in the analysis, this can be repeated for (n-1) cases. Significance is assessed empirically based on the linkage disequilibrium (LD) structure and allele frequencies for the SNPs studied.

Power for SGS. Using simulated high-risk pedigrees using genotype information only from affected cases, we investigated power for a range of genetic models based on disease prevalence, minor allele frequency (MAF), and heterogeneity to represent disease loci that explain 0.2% to 99.8% of total disease risk (Knight et al., in press). Pedigrees were required to be “high-risk”, based on significant excess ($p<0.001$, $p<0.0001$; see Table 2 for a comparison with our high-risk pedigrees), and contain ≥15 meioses between cases. Results indicated that 3-10

pedigrees was sufficient to gain excellent power to see at least one true positive within any given pedigree (>80%) for the majority of the models. In general, power increased with attributable risk, less heterogeneity, and the excess of disease in the pedigree.

Consideration of sharing in (n-1) cases also increased power by correctly accounting for sporadic disease in the models. For all scenarios considered, genomewide association studies would have negligible power (N = 30,000; power=10%). The table illustrates the number of *pedigrees* necessary to gain ≥80% power *to detect at least one true variant in one pedigree* for a range of models. The characteristics of our ascertained high-risk pedigrees (Table 2 and figure 1) give excellent power for most scenarios considered.

Number pedigrees to detect at least 1 true variant in at least 1 pedigree							
				All cases sharing		n-1 sharing	
Preva-lence	MAF	Hetero-geneity	Attr. Risk (%)	FSIR < 0.0001	FSIR < 0.001	FSIR < 0.0001	FSIR < 0.001
0.5%	0.005	0.5	100	3	3	3	3
		0.8	40	5	6	5	5
	0.0005	0.5	10	2	2	1	2
		0.8	4	3	6	2	4
	0.00005	0.5	1	5	12	4	9
		0.8	0.4	6	21	4	18
1%	0.005	0.5	50	3	3	2	2
		0.8	20	7	9	4	5
	0.0005	0.5	5	5	6	4	4
		0.8	2	10	10	5	6
	0.00005	0.5	0.5	3	12	2	8
		0.8	0.2	23	29	16	26

SGS is based on sharing among all or most cases in high-risk pedigrees and has excellent power using a few pedigrees. However, if the sporadic rate constitutes more than 2 or 3 cases (increased heterogeneity), a new method (weighted mean pairwise SGS, or pSGS) will be used. Estimated power of pSGS. The method has been tested on simulated exome sequence data (GAW17, Cai et al. 2011), and found to successfully identify true variants in 13/18 (72%) of high-risk pedigrees. These pedigrees had p-values for excess disease ranging from 0.002 to 0.007, similar to our ascertained families. Variants successfully discovered by pSGS had population frequencies ranging from 0.0007 to 0.0065, and frequency in cases from 0.044 to 0.91, with ¾ having a frequency in cases of <0.25. Given the characteristics of this dataset, two pedigrees would have more than 90% power using pSGS to detect a genomic region containing at least one true variant. We will select a set of common variants available in both exome and genome sequence data for the SGS analyses. Each of our pedigrees will be analyzed using SGS. A suite of programs to perform SGS will be used (http://medicine.utah.edu/internalmedicine/GeneticEpidemiology/software.htm). The method first establishes the positions and lengths of SGS sharing, then uses graphical modeling to estimate LD structure (from the HapMap CEU founders, who were from Utah). Then empirical assessments of sharing are made. Within each family, we will perform SGS. The number of SGS segments across the genome is estimated for each pedigree across several thousand simulations, and is used to establish a genome-wide significance threshold for that pedigree. If a pedigree yields no significant results, we will use pSGS, with the assumption that heterogeneity may have masked findings using standard SGS. pSGS loops through the standard SGS programs, calculating sharing in all possible pairs and performing a weighted average based on the meiotic distance between pairs. Significance values and genome-wide thresholds will again be determined empirically. Given previous experience with these methods and the simulated power analyses for SGS and pSGS, we anticipate that most of our pedigrees will exhibit genome-wide significant shared regions.

SGS analyses incorporating phenotypic subtyping. Incorporation of phenotypic subtyping may increase genetic homogeneity and identify novel susceptibility variants. As we characterize pedigrees with unique phenotype profiles, we will be able to re-analyze subgroups of decedents within these pedigrees using SGS and pSGS. In this case, we

can restrict pSGS analyses to only cases with the same phenotypic profile, incorporating an additional weight in the statistic based on phenotypic similarity for each pair. This weight will be incorporated in an analogous way to the weighting for number of meioses separating the pair in the original statistic (Cai et al., 2011). In its simplest form, this additional coefficient will be a 0/1 indicator (1 if the subtype matches in the pair, 0 otherwise). More sophisticated quantitative weights of similarity will also be considered.

Statistical follow-up in pedigrees. The prioritization of sequence variants is a challenge and an area of intense current research. Analyses of exome variants (presented below) rely substantially on likely functional importance using knowledge of the consequences of variation in coding regions. Analysis of variants in non-coding regions is truly in its infancy. The prioritization of both exomic and non-coding sequence variants can benefit substantially from knowledge of sharing in pedigrees; this sharing will be used as a weighting factor and also in pathway analyses. In addition to familial evidence (sharing), we will also incorporate the significance of non-coding variation using the recently-developed resource RegulomeDB (Boyle et al., 2012). Finally, once sequence variants have been validated using analyses of direct sequencing (see Aim 3 below), such variants can be used in linkage models as covariate effects. While such analyses traditionally focus on phenotypic covariates, the methods can also be applied to test the effects of particular genetic variation on linkage signals. Such covariate analyses provides effective evaluation of candidate causal variants (Wijsman et al., 2010), and allows comprehensive evaluation of all variants in a linkage region, without first requiring use of possibly imperfect filters to identify plausibly deleterious variants. These analyses will be particularly useful in the analysis variants in non-coding regions.

Condel and other methods. We propose to test coding sequence results using the Consensus Deleteriousness Score (Condel, Gonzalez-Perez & Lopez-Bigas, 2011). This new method produces a weighted average of the normalized scores from five sequence variant analysis tools (Logre, MAPP, Massessor, Polyphen2, and SIFT). The resulting score has been shown to increase accuracy in predicting if sequence variants are deleterious, and their potential impact on protein function. Mutation data (coded in forward strand) can be directly entered on the Condel website (bg.upf.edu/condel/analysis) and scores obtained. We will test mutation candidates from broader regions and lists of genes identified through SGS.

One area of recent development includes analyses of gene pathways (Sullivan et al., 2012). It is possible that while different family-specific variants are identified, co-morbid conditions and demographic characteristics will provide post hoc insights to tie together genes in risk pathways. In addition to genes identified by a growing number of collaborative analyses (e.g., Schizophrenia Psychiatric GWAS Consortium, 2011; Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011), functionally related sets of genes, such as those comprising the “synaptome” and those beginning to be implicated in the neuronal “connectome,” are of interest. Other collaborative efforts have studied >1000 synaptic genes, making use of proteomic analysis of synaptic fractions (Croning et al., 2009; Bayes et al., 2011; genes2cognition.org), which can be treated as a group or can be meaningfully subdivided. In addition to prospective groupings of genes, we can also explore statistical approaches using external published literature (Raychaudhuri et al., 2009) to help interpret test results across many genes simultaneously. For example, we can evaluate whether gene sets are unusually related in literature, protein-protein or expression data sets allowing opportunity to discover groupings of genes that may not represent currently annotated gene sets. Finally, it is possible that variants will be identified in gene pathways that we have not anticipated with our selection of co-morbid phenotypes. If this happens, we will be able to return to the rich databases at our disposal and investigate familial co-morbidity for other medical conditions.

Criteria for prioritizing findings for follow-up studies will focus on the strength of statistical and biological evidence of implicated genes/variants, and evidence across multiple families within our project, suggesting variants of potentially greater clinical impact. We will pursue targeted sequence work for 2-3 significant genes per year to confirm causal variants and characterize other variation across the genes. In addition to validation in the original subjects, we will link records to the UPDB of Utah suicides with DNA collected throughout the project (i.e., those not in our initial UPDB genealogy linking). Some of these will expand existing pedigrees, some may create new pedigrees. From our total set of suicides, we will also identify those who share phenotypic profiles of the decedents with confirmed sequence variants. This phenotypic classification will be done using the SAS K-Nearest Neighbor procedure (www.sas.com), a robust classification technique that does not require assumptions of the distributions of the classifying variables. Our first replication experiments will include these selected decedents (new familial decedents and/or decedents sharing familial phenotypic risk factors). Genes with validated mutations will be sequenced to determine all variation across the entire gene (N: 500-1000 subjects per gene).

We will determine whether confirmed variants are over-represented in the remaining UT suicide decedents by genotyping the remaining unselected UT suicides and comparing to reference data.

Genotyping of variants will be done with standard methods using the University of Utah Health Sciences Center genomics core facility. Note that the most efficient, state-of-the-art methods will be incorporated at the time in the project when this genotyping is needed.

We will use the PLINK software (Purcell et al., 2007) to test for the presence and frequencies of confirmed variants in our large sample of all Utah suicide decedents. PLINK incorporates multiple statistical tests as appropriate given cell sizes of the observed frequencies. It handles other characteristics of the sample; for example, stratification, quantitative variables, single variant and haplotype tests, and single locus and multi-locus tests. Covariates can be included directly in the analyses, allowing us to test specific hypotheses regarding phenotypic subsets in the Utah suicide resource. PLINK offers a distinct analysis advantage with a feature that allows the generation of empirical significance levels. We will compare the presence of confirmed variants in the full suicide sample with presence in publicly available control samples.

Table 5. Power for case-control replication assuming 2000 cases and 2000 controls.

Risk allele frequency	Relative risk	power
0.005	1.5	0.29
0.005	2.0	0.72
0.005	2.5	0.94
0.01	1.5	0.50
0.01	2.0	0.95
0.01	2.5	1.0

Publicly-available, de-identified control samples will be obtained from the best characterized sources available. These will be samples from individuals from other studies who previously consented that their de-identified data could be used for research. Because controls and other external data sets will not be ascertained in Utah we will test for population substructure. An advantage of PLINK is that we can perform permutations only within a cluster if we identify population substructure. For case-control data, PLINK uses a label-swapping procedure for permutations, under the assumption that individuals are independent and therefore exchangeable under the null hypothesis. Table 5 gives an outline of power to replicate a variant given an estimate of 2000 Utah cases and 2000 controls (assuming the potential need for within-cluster comparisons), and assuming a range of characteristics of risk alleles and relative risk. Power was estimated using Genetic Power Calculator (Purcell et al., 2003).

We will use phenotypes in decedents with confirmed variants to direct selection of external datasets from the psychiatric genetics research community. Determine presence of confirmed variation in these external datasets; compare to control data generated in Aim 3b.

Data sets. There are several external data sets that will be available for follow up of our findings of variants. These data sets already have or will soon have sequence data. Through the NIMH Center for Collaborative Genomic Studies on Mental Disorders, WGS is being generated on 1000 bipolar cases and 1000 controls, and 100 schizophrenia cases. Other relevant sequencing data sets include: Sweden Schizophrenia Study (525 subjects, Purcell, Sklar); I2B2-Major Depression (50 subjects, Perlis, Smoller); Autism (1778 subjects, Daly, Gibbs). Dr. Coon is also a member of the PGC. This group has recently published large mega-analyses of bipolar disorder and schizophrenia. The PGC has been a highly collaborative group; we anticipate being able to pursue collaborations with multiple data and DNA resources.

Gene expression methods. After alignment of expression data, raw counts are quantified for each gene using Rsubread (Liao et al., 2013), and initial quantifications can be examined for confounding variables (brain pH, post-mortem interval, sequencing lane, sequencing date) using combat (Chen et al., 2011). Normalized count data can then be tested for differential expression using the edgeR package (Robinson et al., 2010). Count data will also be used with the Weighted Gene Co-expression Network Analysis (WGCNA) software, which identifies networks of co-expressed genes (Langfelder & Horvath, 2008). Network analysis is an agnostic approach blind to case/control membership. Differentially expressed genes and gene co-expression networks can be tested for functional enrichment using Gene Set Enrichment Analysis (Subramanian et al., 2005; Mootha et al., 2003). Expression will be used to validate DNA results and/or to prioritize DNA variants for additional study. Other studies in tissue samples will similarly serve as directed follow-up experiments.

Additional replication and extension studies will be possible through collaboration with the Psychiatric Genomics Consortium (PGC). The PGC is an international collaborative group focused on genetic analyses of psychiatric disorders. We share de-identified summary-level data (not individual-level data) with this highly productive research group. No family structure data will be shared. Data access for summary-level data is by application only, and is controlled by the PGC data access committee, with additional review of proposed use of data by a Utah OME representative. Applications for access follow those used by the NIH data repositories.

We have studied suicide risk in Bipolar Disorder (BD) in collaboration with the University of Iowa. This center has detailed phenotyping on BD cohorts with and without suicide attempts. We compared basic demographics, age at first diagnosis, polygenic risk scores, and frequencies of more rare, putatively functional gene variants across these two cohorts and Utah BD completed suicides. Analyses included both descriptive characterization and cluster modeling to determine prediction of risk of attempt and completion in individuals with BD. All analyses were done at Utah; no data were sent to the University of Iowa. This collaboration is now closed.

We will conduct genome-wide association studies using case-control comparisons of suicide deaths, non-suicide attempt cases, non-suicide deaths without evidence for prior suicidal behaviors, and publicly-available controls. We will investigate similarities and differences of suicide deaths vs. non-suicide attempts vs. controls, making use of demographics (age, sex, race/ethnicity), co-occurring diagnoses from the electronic health records, and polygenic risk scores. Analysis will use machine-learning cross-validation methods, and will be enhanced by our ability to compare results with other data resources in the International Suicide Genetics Consortium, which now studies data from >20,000 individuals with suicidal behavior and >170,000 controls.

Amendment 3/2021 requests APCD medication data as a more comprehensive source for medications. These data will be used to verify ICD-9/ICD-10 diagnostic category determination, and to allow more accurate study of individuals in particular high-risk subgroups (e.g., chronic pain, transgender, opioid use/misuse).

Amendment 7/2021- The analysis of hair will be descriptive and include measures of the 1) presence or absence of each analyte by decedent, 2) the mean concentration and variation of each analyte among decedents over the entire sample and by month, and 3) the correlation among analytes across the entire sample. The presence/absence of

analytes and mean analyte concentrations will be compared by decedent characteristic including sex, age group, and season of death using chi-square tests and t-tests/ANOVAs.

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244 - Request for Waiver of Consent

PI: Hilary Coon Ph.D.

Submitted: 3/22/2011

Title: Genetic risk factors in suicide and depression

Request for Waiver or Alteration of Consent

* Requested Waivers

	Date Created	Type of Request	Purpose of Waiver Request
View	12/11/2012	Waiver of Informed Consent	Suicide victims and non-suicide deaths with and without evidence for suicide attempts cannot be contacted for consent. While extended family information will be used to determine familial risk, and co-morbid risk, no actual living family members will be studied, and no identifying information will be used from family members. To determine significance of the risk of co-occurring conditions/demographic characteristics/geocoded clustering, matched controls must be found from the ~9,000,000 records in the UPDB. Matching is done within the UPDB. All data are made into limited use datasets prior to analysis.

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244** IRB Smart Form

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Request for Waiver or Alteration of Consent

1. **Purpose of the Waiver Request:**
   Suicide victims and non-suicide deaths with and without evidence for suicide attempts cannot be contacted for consent. While extended family information will be used to determine familial risk, and co-morbid risk, no actual living family members will be studied, and no identifying information will be used from family members.
   To determine significance of the risk of co-occurring conditions/demographic characteristics/geocoded clustering, matched controls must be found from the ~9,000,000 records in the UPDB. Matching is done within the UPDB. All data are made into limited use datasets prior to analysis.

2. **Type of Request:**
   Waiver of Informed Consent

3. **List the identifying information you plan to collect or keep a link to (e.g. names, dates, or identification numbers such as social security numbers or medical record numbers, etc.).**
   Identifying information needed by the UPDB to link suicide cases and non-suicide deaths with and without prior attempts to health data and extended high-risk pedigrees includes: name, birth date, death date, location at time of death, and the State Death Certificate number. These identifying elements will ONLY be transferred directly from the Office of the Medical Examiner or the IH Biorepository to the UPDB. This information will ONLY be used by the UPDB for linking, and will not be transferred to the research personnel on this project. The research personnel will receive: pedigree structure, numeric DNA/fibroblast tissue identifying number, IH Biorepository number, demographic data, and coded UPDB heath information (e.g., ICD-9 and ICD-10 co-morbidity data). For the 1-year cohort where Dr. Doug Gray collected additional risk factor information, these data will be used WITHOUT any identifying information; data will be linked to DNA/fibroblast tissue and pedigree structure only through numeric IDs.

4. **Explain why the research could not be practicably conducted without using identifiable information. Examples of such explanation could include the following:**
   Identifying information is needed by the UPDB only o link suicide cases and non-suicide deaths with and without prior attempts to health data for risk analysis. Without use of identifying information by UPDB staff, linking of cases to their health data would not be possible. This linking is essential for the primary aims of this study, which include characterization of high-risk subgroups and development of risk models by the analysis team. This analysis work depends upon health data, membership in high-risk families, and genetic risk, where clinical, demographic, and pedigree data are linked to cases, then identifiers stripped to form a limited use data set.

5. **Explain why the research could not practicably be conducted without the waiver or alteration. For example, complete the following sentence "If I had to obtain consent, the research could not be conducted because...":**
   Permission cannot be obtained from the suicide victims, or from deceased non-suicide comparison and control deaths. Contacting family members for permission would be extremely intrusive. We will not know the identity of individuals and are not requesting any contact with these families, and will use pedigree structure and diagnostic data with the utmost caution to prevent revealing the identity of any family member.

6. **Explain why the research and privacy risk of the research are no more than *minimal:***
   There will be no subject contact. Only family structure, suicide status, suicide attempt status for non-suicide deaths, manner of death for comparison and control deaths, and UPDB record information will be used. Identifying information will ONLY be transferred from the OME or the IH Biorepository to the UPDB for genealogical linking, and will not be used for analysis.

7. **Describe the measures you will take to ensure the waiver or alteration will not adversely affect the rights and welfare of the *subjects:***
   Any publication involving diagnosis and pedigree structure will be disguised to protect the identity of subjects. No attempt will be made to discover the identity of family members.

8. **Explain how you will, if applicable and appropriate, provide the subjects with additional pertinent information *after* they have participated in the study, or indicate *"Not applicable":***
   Not applicable.

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244**

5. Data Monitoring

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

5. Data Monitoring Plan

1. **Privacy Protections:** Privacy refers to persons and to their interest in controlling access of others to themselves. Privacy can be defined in terms of having control over the extent, timing and circumstances of sharing oneself (physically, behaviorally, or intellectually) with others. **What precautions will be used to ensure subject privacy is protected?**

Select all that apply:

The research intervention is conducted in a private place

The collection of information about participants is limited to the amount necessary to achieve the aims of the research, so that no unneeded information is being collected

Other or additional details (specify):

Research done in this study will not use any identifying information on subjects or family members. For the 1-year cohort of suicide victims where additional risk information was collected by the OME, only de-identified information from this risk data (linked through an anonymous numeric ID) will be used. Similarly, only de-identified data from subjects linked through IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, and IRB_00144804 will be used. Because pedigree structure together with suicide or co-morbid disorder status may still possibly reveal identity, all files revealing such structure will be kept in locked, secured rooms within the University covered entity. All computer files with such structure and suicide status will be kept on the CHPC Protected environment servers with password protected access. Any data analysis will be done only on computers with password access, behind the university firewall, and with full encryption.
De-identified data (phenotype, sequence, and genotype) will be deposited into the NIMH Repositories (DNA, data) and as specified by the NIH. All data will be de-identified. No detailed pedigree data will ever be put into the repository. Rather, a numeric ID will indicate which extended pedigree a sample is part of, if any. Only summary-level (not individual level) data will be available to the PGC consortium. No individual-level data are stored on PGC servers. Access to NIMH repository data is governed by an application/approval process through the NIH and a similar process exists for the PGC summary-level data with additional review of the proposed data use by a OME representative. Use of pedigree data by those accessing the repository will only occur through explicit additional approval by the RGE.
All samples sent for Heritage 1K sequencing were associated only with anonymous numeric ids. No pedigree structures were sent, only degree of relatedness.Only completely de-identified data from deceased individuals were sent.
All data used by PGC investigators is completely de-identified and is summary-level, not individual-level data. All publications resulting from Heritage 1K sequencing or PGC consortium members will include U of U study team members, and will first be approved by the RGE, as appropriate. The Heritage 1K scientists associated with the donor have destroyed all data; the Heritage 1K data is archived for UU study team use. Similar protocols and restrictions have applied to data and analyses generated with our Janssen colleagues; this collaboration is also now closed and data have been deleted from the Janssen site.

For the NIH repositories, in the past, a de-identified Global Unique Identifier (GUID) was previously created by PPR staff as described in the Procedures section of this protocol. However, we have recently obtained permission to only share data through the NIMH repository at Rutgers which uses only an anonymous numberic ID and does not require the GUID.

Locally, data will be stored on the secure CHPC Protected Environment, and on the HIPAA-compliant UBox. UBox was developed for use by University of Utah faculty and staff. It is HIPAA-compliant, and is among the select number of platforms for sensitive data storage. Level of access is controlled by the study PI, who will restrict access to only those approved on the IRB. In addition, study team must access UBox from approved University of Utah devices that have been encrypted using accepted encryption algorithms. Box is encrypted using AES 256bit for data at rest and TLS 1.2 for data in motion. Backups are encrypted using this same methodology and stored in the cloud. The UofU has also negotiated in the Business Associate Agreement with Box that backups are not stored in any datacenter outside the United States. Note: For this study, only HIPAA Safe Harbor de-identified data will be stored on UBox and access to UBox will be managed by the UPDB Data Security Analyst.

2. **Confidentiality Precautions:** Confidentiality is an extension of the concept of privacy; it refers to the subject's understanding of, and agreement to, the ways identifiable information will be stored and shared. Identifiable information can be printed information, electronic information or visual information such as photographs. **What precautions will be used to maintain the confidentiality of identifiable information?**

Select all that apply:

Storing research data on password protected computers or in locked cabinets or offices

Complete de-identification of study data

All data that will be transferred or transported outside of the institution will be encrypted

Other or additional details (specify):

Other or additional details (specify):

Any publication involving printed drawings of pedigree structure and suicide/psychiatric disorder status will be disguised as recommended by the UPDB, and as practiced previously by our research group in our autism pedigree studies. All publications will first be reviewed by the RGE, as appropriate.
No PHI are available to the study team. By extension, no PHI are shared with outside investigators. Detailed pedigree structure data will not be part of the NIMH repository, as described above. Pedigrees were not sent with samples for Heritage 1K sequencing, or to Janssen collaborators, or to the PGC consortium, as noted above. The Heritage 1K and Janssen scientists have destroyed study data; these efforts are now closed.PGC collaborators only request summary-level data.
Locally, data are stored on the secure CHPC Protected Environment, and in a completely de-identified state on the HIPAA-compliant UBox. UBox was developed for use by University of Utah faculty and staff. It is HIPAA-compliant, and is among the select number of platforms for sensitive data storage. Level of access is controlled by the study PI, who will restrict access to only those approved on the IRB. In addition, study team must access UBox from approved University of Utah devices that have been encrypted using accepted encryption algorithms. Box is encrypted using AES 256bit for data at rest and TLS 1.2 for data in motion. Backups are encrypted using this same methodology and stored in the cloud. The UofU has also negotiated in the Business Associate Agreement with Box that backups are not stored in any datacenter outside the United States. Note: For this study, only HIPAA Safe Harbor de-identified data will be stored on UBox and access to UBox will be managed by the UPDB Data Security Analyst.

3. Will photos, audio recordings, or video recordings, or medical images of participants be made during the study?

○ Yes ● **No**

If yes, describe the recording/images and what will become of them after creation (e.g., shown at scientific meetings, stored in the medical/research record, transcribed, erased, etc.):

4. How will study data and documentation be monitored throughout the study?

Select all that apply:

Periodic review and confirmation of participant eligibility

Periodic review of the transfer/transcription of data from the original source to the research record

Other additional details (specify):

5. Who will be the primary monitor of the study data and documentation?

Select all that apply:

Principal Investigator

Other or additional details (specify):

Other or additional details (specify):

Investigators at the OME, the UPDB, and IH (David Crockett) will also monitor data (these individuals will be the only study personnel with access to identifying data).
The NIH will oversee use of de-identified data that are part of the NIMH repository as a result of this study.
The Heritage 1K scientists affiliated with the donor managed the whole genome sequence data generated on de-identified subjects, transferred that data and any findings resulting from that data to U of U investigators, and have now deleted data. Janssen scientists managed generation and transfer of molecular data, and analyses and communication of results to the U of U team, and have now deleted data. The PGC only allows access for analysis of completely de-identified, summary-level data after an approval process similar to that of the NIH. At this point, this access for PGC investigators does not include individual-level data. For studies in collaboration with the University of Iowa, all analyses were done at the University of Utah; no data were transferred.
Data on CHPC PE is monitored continuously by safeguards set up by CHPC management. UBox use will be via encrypted approved UU devices. Access will be controlled by the study PI; only IRB-approved study team members will

be allowed access. Use of UBox will be accompanied by the following additional security aspects:
1) Periodic (at least annual) external audit whereby the study PI provides access to UPDB Data Security Officer to ensure personnel who have been given access to the UBox folders all are IRB-approved for the study.
2) Additional audit by the UPDB Data Security Officer triggered by any IRB amendment that results in addition or removal of study staff.

There is an approved COI management plan for Dr. Mark Yandell. Per this plan, a non-conflicted investigator on the study will review all research results where Dr. Yandell has been involved prior to publication, and this review will also be subject to oversight by the COI committee.

6. How often is study data and documentation monitoring planned (e.g., monthly, twice a year, annually, after N participants are enrolled, etc.)?

Data monitoring will occur each time data are transferred from the OME to the UPDB, and each time de-identified data are obtained from the UPDB (approximately quarterly). Additional oversight by the RGE of representation of pedigree data will be done any time a publication is submitted.
Data will additionally be checked with each deposition into the NIMH repositories (approximately annually throughout the NIH award).
Samples and data were checked for quality prior to sending for Heritage 1K sequencing, or Janssen molecular work. The PI assured that data accompanying those samples did not include explicit pedigree information.
Data on CHPC PE is monitored continuously by safeguards set up by CHPC management. UBox use will be via encrypted approved UU devices. Access will be controlled by the study PI; only IRB-approved study team members will be allowed access. Use of UBox will be accompanied by the following additional security aspects:
1) Periodic (at least annual) external audit whereby the study PI provides access to UPDB Data Security Officer to ensure personnel who have been given access to the UBox folders all are IRB-approved for the study.
2) Additional audit by the UPDB Data Security Officer triggered by any IRB amendment that results in addition or removal of study staff.

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244** 6. Risks and Benefits

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

6. Risks and Benefits

1. Describe the reasonable foreseeable risks or discomforts to the participants:

This is a records-only study. The researchers will have no access to identifying data for suicide victims, non-suicide deaths with and without prior attempts, or family members. There will be no procedures performed on subjects. There will be no direct contact of any family members.

The only risk is potential loss of confidentiality. Data will be encrypted before being transferred from Vital Records/OME. Data firewalls are firmly in place at the OME and at the UPDB to safeguard data.

Data transfers to the NIMH repositories, the Heritage 1K scientists and Janssen scientists will include only de-identified data; however, transfers will still occur only using fully encrypted devices. For all studies done in collaboration with the University of Iowa, analyses will only be done by the UU research team; no data will be transferred. PGC collaborators only use completely de-identified summary-level (not individual-level data).

For the NIH repositories, a de-identified Global Unique Identifier (GUID) was created in the past by PPR staff as described in the Procedures section of this protocol. New arrangements specify sharing only with the NIMH sample repository at Rutgers which uses an anonymous numeric id; no GUID is required.

2. Describe the potential benefits to society AND to participants (do not include compensation):

In this study, we hope that by identifying familial suicide, we may begin to reduce this heterogeneity and thereby identify specific genetic risk factors in these families. While these risk factors may be rare and specific to particular families, genes that are discovered to be involved could shed light on biological pathways that are more general risk factors.

Because the study is now allowed to be part of the NIH repositories and because we will participate in the Heritage 1K initiative, and analyses by world experts in the Psychiatric Genomics Consortium, there is a greater chance that substantive findings will result from the work.

3. Are there any costs to the participants from participation in research?

○ Yes ● **No**

If yes, specify:

4. Is there any compensation to the participants?

○ Yes ● **No**

a. **If yes, answer the following:**
Specify overall amount:

b. **Specify when participants will be paid (e.g. at each visit, at end of study, etc.):**

c. **If applicable, please specify payment by visit or other time interval (e.g. $10 per visit, etc.):**

d. **If applicable, explain plan for prorating payments if participant does not complete the study:**

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244
7. HIPAA & the Covered Entity

PI: Hilary Coon Ph.D.

Submitted: 3/22/2011

Title: Genetic risk factors in suicide and depression

7. HIPAA and the Covered Entity

1. **Does this study involve Protected Health Information (PHI) or de-identified health information?**
   ● **Yes** ○ No

   a. **Select the method(s) of authorization that will be used:**

   Waiver or Alteration of Authorization

   De-identified

   Select the method of de-identification: Safe Harbor De-Identification

   b. **Will PHI be disclosed outside the Covered Entity?**
   ○ Yes ● **No**

Does this study involve any of the following:

2. **The investigational use of a drug?**
   ○ Yes ● **No**
   Mark yes, for an expanded access application.

3. **The investigational use of a medical device or humanitarian use device?**
   ○ Yes ● **No**
   Mark yes, for an expanded access application.

4. **The investigational use of a dietary supplement, food, or cosmetic?**
   ○ Yes ● **No**

5. **Is this an investigator-initiated drug or device trial lead by the Principal Investigator'**
   ○ Yes ● **No**
   All investigator-initiated drug or device trials are required to have a full research protocol attached to the Documents and Attachments page.

6. **Will this study involve the use of an imaging modality from the department of Radiology?**
   ○ Yes ● **No**

7. **Exposure to radioisotopes or ionizing radiation?**
   ○ Yes ● **No**

8. **Genetic testing and/or analysis of genetic data?**
   ● **Yes** ○ No

9. **Creating or sending data and/or samples to a repository to be saved for future research uses?**
   ● **Yes** ○ No

10. **Are you:**

    Collecting samples of blood, organs or tissues from participants for research purposes;

    Introducing Recombinant or Synthetic Nucleic Acids (e.g. viral vectors, oligonucleotides) or cells containing recombinant nucleic acids (e.g. CAR-T) into participants; OR

    Introducing other biological materials (e.g. bacteria, viruses) into participants.

    ○ Yes ● **No**

11. **Does this study involve any of the following?**

    Cancer Patients

    Cancer Hypothesis

    Cancer risk reduction

    Cancer prevention

    ○ Yes ● **No**

12. **Any component of the Clinical and Translational Science Institute (CTSI)?**
    ● **Yes** ○ No

    The Clinical Research Unit (CRU)?
    ○ Yes ● **No**

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244 - Request for Waiver of Authorization

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Request for Waiver or Alteration of Authorization

Request for Waiver of Authorization for Recruitment Only

This option must only be used if you are reviewing PHI in order to identify eligible participants BEFORE approaching them to obtain consent and authorization. All other waiver requests must be entered below.

Other Requests for Waivers of Authorization:

Click "Add" below to add a new waiver request to this application.

Click the waiver name link to edit a waiver that has already been created.

To delete a waiver request, contact the IRB.

	Date Created	Type of Request	Purpose of Waiver Request
View	12/11/2012	Waiver of Authorization	We cannot obtain authorization from suicide victims or comparison non-suicide deaths with and without evidence of prior suicide attempts. Family members are only included in the study to identify familial risk and co-morbid risk, and to define clustering of suicides for DNA/tissue studies. No identifying information is used on any subjects. No attempts will be made to contact living relatives. To determine significance of the risk of co-occurring conditions/demographic characteristics/geocoded clustering, matched controls must be found from the ~9,000,000 records in the UPDB. Matching is done within the UPDB. All data are made into limited use data sets prior to analysis.

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244** IRB Smart Form

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Request for Waiver or Alteration of Authorization

1. **Purpose of the Waiver Request:**
   We cannot obtain authorization from suicide victims or comparison non-suicide deaths with and without evidence of prior suicide attempts. Family members are only included in the study to identify familial risk and co-morbid risk, and to define clustering of suicides for DNA/tissue studies. No identifying information is used on any subjects. No attempts will be made to contact living relatives.
   To determine significance of the risk of co-occurring conditions/demographic characteristics/geocoded clustering, matched controls must be found from the ~9,000,000 records in the UPDB. Matching is done within the UPDB. All data are made into limited use data sets prior to analysis.

2. **Type of Request:**
   Waiver of Authorization

3. **List the identifying information you plan to collect or keep a link to (e.g. names, dates, or identification numbers such as social security numbers or medical record numbers, etc).**
   Identifying information needed by the UPDB to link suicide cases and non-suicide deaths with and without prior attempts into pedigrees and to health data will ONLY be transferred directly from the Office of the Medical Examiner or the IH Biorepository to the UPDB. This information, including State Death Certificate number and names, will ONLY be used by the UPDB for linking, and will not be transferred to the research personnel on this project. The research personnel will receive: pedigree structure, numeric DNA/fibroblast tissue identifying number, IH Biorepository number, demographic data, and UPDB record information (death certificate data, ICD-9 co-morbidity data). For the 1-year cohort where Dr. Doug Gray collected additional risk factor information, these data will be used WITHOUT any identifying information; data will be linked to DNA/fibroblast tissue and pedigree structure only through numeric IDs.

4. **Explain why the *PHI* to be used or disclosed is the minimum necessary to accomplish the research objectives:**
   PHI will only be exchanged between the Office of the Medical Examiner (OME) or the IH Biorepository and the UPDB. This information is necessary to link those suicide victims whose DNA/tissue has been collected with heath data, and with the genealogical records to determine larger pedigree structures of suicide victims who are distantly related.
   PHI will be used as necessary to match control comparison subjects within the UPDB, but will then be stripped and will not be used by study investigators.

5. **Explain why the research could not practicably be conducted without the waiver of authorization.** *For example, complete the following sentence: "If I had to obtain authorization, the research could not be conducted because..."*
   Suicide victims or non-suicide deaths with and without prior suicide attempts cannot be contacted for consent. It is impossible to know ahead of time which families may be linked into larger extended families. Contact of family members of all suicide victims would be prohibitive, and would be needlessly intrusive for these families.

6. **Describe your plan to protect the identifiers from improper use and disclosure, and indicate where the *PHI* will be stored and who will have access:**
   Identifiers will only be accessible to the OME or the IH Biorepository. They will be securely transferred to the UPDB.

7. **The identifiers must be destroyed at the earliest opportunity consistent with conduct of the research, unless there is a health or research justification for retaining the identifiers or such retention is otherwise required by law. Describe how and when you will destroy the identifiers, or justify their retention:**
   The research personnel will never have access to identifiers. The OME and the IH Biorepository have the right to retain access. The UPDB will destroy identifiers at the end of the project.

8. **Describe the measures you will take to ensure the PHI will not be reused or disclosed to any other person or entity, except as required by law, for authorized oversight of the research study, or for other research approved by the IRB:**
   Because the research personnel will never have the identifiers, they will not be able to disclose them to any other entity.

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244 - Safe Harbor De-Identification Agreement

PI: Hilary Coon Ph.D.

Submitted: 3/22/2011

Title: Genetic risk factors in suicide and depression

Safe Harbor De-Identification

1. **This declaration applies to the following part(s) of this study:**
   - A. ☐ All of the information used or disclosed in this study.
   - B. ☐ The information received or collected from these sources:
   - C. ☑ The information shared with or disclosed to these groups: Mt. Sinai, PGC, University of Iowa, University of California Irvine
2. **As the principal investigator for this study, I declare the following:**

   1. To the best of my knowledge, the information could not be used (alone or with other information) to identify an individual who is a subject of the information, and

   2. None of the following types of information, regarding subjects or relatives, employers, or household members of subjects, are used or disclosed in the part of this study indicated above:

      a. Names;

      b. All geographic identifiers except state or the first three digits of a zip code (however, all data from the following 17 3-digit zips are combined together under "000": 036, 059, 063, 102, 203, 556, 692, 790, 821, 823, 830, 831, 878, 879, 884, 890, and 893)

      c. The month and day (the year can be kept) from all dates directly related to an individual, including birth date, admission date, discharge date, date of death. Ages over 89 are combined in a single category of "Age 90 and older."

      d. Telephone numbers;

      e. Fax numbers;

      f. Electronic mail addresses;

      g. Social security numbers;

      h. Medical record numbers;

      i. Health plan beneficiary numbers;

      j. Account numbers;

      k. Certificate/license numbers;

      l. Vehicle identifiers and serial numbers, including license plate numbers;

      m. Device identifiers and serial numbers;

      n. Web Universal Resource Locators (URLs);

      o. Internet Protocol (IP) address numbers;

p. Biometric identifiers, including finger and voice prints;

q. Full face photographic images and any comparable images; and,

r. Any other unique identifying number, characteristic, or code, except as permitted for re-identification.

3. If I assign a code or other means of record identification to allow de-identified information to be re-identified,

1. The code or other means of record identification is not derived from or related to information about the individual and is not otherwise capable of being translated so as to identify the individual, and
2. I will not use or disclose the code or other means of record identification for any purpose other than re-identification, and I will not disclose the mechanism for re-identification.

4. Before I allow a code to be used to re-identify this information,

1. If the purpose of the re-identification is within the scope of the original protocol, I will obtain approval of an amendment from the IRB and comply with the requirements of HIPAA; or
2. If the purpose of the re-identification is outside the scope of the original protocol, I will submit a full New Study Application, obtain IRB approval, and comply with the requirements of HIPAA.

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244** - Genetic Research

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Genetic Research

1. **Describe the risks to participants in regard to genetic testing, including applicable risks to privacy and confidentiality, as well as psychological and social risks.**

   This is a records-only study. The researchers will have no access to identifying data for suicide victims, non-suicide deaths with and without prior attempts, or family members. There will be no procedures performed on subjects. There will be no direct contact of any family members.

   The only risk is potential loss of confidentiality. Data will be encrypted before being transferred from Vital Records/OME. Data firewalls are firmly in place at the OME and at the UPDB to safeguard data.

   Data transfers to the NIMH repositories, the Heritage 1K scientists and Janssen scientists will include only de-identified data; however, transfers will still occur only using fully encrypted devices. For all studies done in collaboration with the University of Iowa, analyses will only be done by the UU research team; no data will be transferred. PGC collaborators only use completely de-identified summary-level (not individual-level data).

   For the NIH repositories, a de-identified Global Unique Identifier (GUID) was created in the past by PPR staff as described in the Procedures section of this protocol. New arrangements specify sharing only with the NIMH sample repository at Rutgers which uses an anonymous numeric id; no GUID is required.

2. **Describe the privacy protections in place for participants in regard to genetic testing. This includes how family member privacy will be protected.**
   Research done in this study will not use any identifying information on subjects or family members. For the 1-year cohort of suicide victims where additional risk information was collected by the OME, only de-identified information from this risk data (linked through an anonymous numeric ID) will be used. Similarly, only de-identified data from subjects linked through IRB_00082991, IRB_00049521, IRB_00006824, IRB_00057455, and IRB_00144804 will be used.
   Because pedigree structure together with suicide or co-morbid disorder status may still possibly reveal identity, all files revealing such structure will be kept in locked, secured rooms within the University covered entity. All computer files with such structure and suicide status will be kept on the CHPC Protected environment servers with password protected access. Any data analysis will be done only on computers with password access, behind the university firewall, and with full encryption.
   De-identified data (phenotype, sequence, and genotype) will be deposited into the NIMH Repositories (DNA, data) and as specified by the NIH. All data will be de-identified. No detailed pedigree data will ever be put into the repository. Rather, a numeric ID will indicate which extended pedigree a sample is part of, if any. Only summary-level (not individual level) data will be available to the PGC consortium. No individual-level data are stored on PGC servers.
   Access to NIMH repository data is governed by an application/approval process through the NIH and a similar process exists for the PGC summary-level data with additional review of the proposed data use by a OME representative. Use of pedigree data by those accessing the repository will only occur through explicit additional approval by the RGE.
   All samples sent for Heritage 1K sequencing were associated only with anonymous numeric ids. No pedigree structures were sent, only degree of relatedness.Only completely de-identified data from deceased individuals were sent.
   All data used by PGC investigators is completely de-identified and is summary-level, not individual-level data. All publications resulting from Heritage 1K sequencing or PGC consortium members will include U of U study team members, and will first be approved by the RGE, as appropriate. The Heritage 1K scientists associated with the donor have destroyed all data; the Heritage 1K data is archived for UU study team use. Similar protocols and restrictions have applied to data and analyses generated with our Janssen colleagues; this collaboration is also now closed and data have been deleted from the Janssen site.

   For the NIH repositories, in the past, a de-identified Global Unique Identifier (GUID) was previously created by PPR staff as described in the Procedures section of this protocol. However, we have recently obtained permission to only share data through the NIMH repository at Rutgers which uses only an anonymous numberic ID and does not require the GUID.

   Locally, data will be stored on the secure CHPC Protected Environment, and on the HIPAA-compliant UBox. UBox was developed for use by University of Utah faculty and staff. It is HIPAA-compliant, and is among the select number

of platforms for sensitive data storage. Level of access is controlled by the study PI, who will restrict access to only those approved on the IRB. In addition, study team must access UBox from approved University of Utah devices that have been encrypted using accepted encryption algorithms. Box is encrypted using AES 256bit for data at rest and TLS 1.2 for data in motion. Backups are encrypted using this same methodology and stored in the cloud. The UofU has also negotiated in the Business Associate Agreement with Box that backups are not stored in any datacenter outside the United States. Note: For this study, only HIPAA Safe Harbor de-identified data will be stored on UBox and access to UBox will be managed by the UPDB Data Security Analyst.

3. **Are you performing whole genome or whole exome sequencing?**
   ● **Yes** ○ No

4. **Describe the confidentiality protections in place for participants' genetic information. Discuss if and how data will be shared and protected outside the local study team.**
   Any publication involving printed drawings of pedigree structure and suicide/psychiatric disorder status will be disguised as recommended by the UPDB, and as practiced previously by our research group in our autism pedigree studies. All publications will first be reviewed by the RGE, as appropriate.
   No PHI are available to the study team. By extension, no PHI are shared with outside investigators. Detailed pedigree structure data will not be part of the NIMH repository, as described above. Pedigrees were not sent with samples for Heritage 1K sequencing, or to Janssen collaborators, or to the PGC consortium, as noted above. The Heritage 1K and Janssen scientists have destroyed study data; these efforts are now closed.PGC collaborators only request summary-level data.
   Locally, data are stored on the secure CHPC Protected Environment, and in a completely de-identified state on the HIPAA-compliant UBox. UBox was developed for use by University of Utah faculty and staff. It is HIPAA-compliant, and is among the select number of platforms for sensitive data storage. Level of access is controlled by the study PI, who will restrict access to only those approved on the IRB. In addition, study team must access UBox from approved University of Utah devices that have been encrypted using accepted encryption algorithms. Box is encrypted using AES 256bit for data at rest and TLS 1.2 for data in motion. Backups are encrypted using this same methodology and stored in the cloud. The UofU has also negotiated in the Business Associate Agreement with Box that backups are not stored in any datacenter outside the United States. Note: For this study, only HIPAA Safe Harbor de-identified data will be stored on UBox and access to UBox will be managed by the UPDB Data Security Analyst.

5. **Will incidental findings relevant to individuals or families be communicated to the participants?**
   ○ Yes ● **No**
   If yes, answer the questions below:

   a. **Describe the process for determining which incidental findings will be returned to the participants. Describe the information and expert consultation that will be used to make this determination.**

   b. **Indicate the process that will be used to return information about incidental finding to participants:**
      There are no items to display
      If Other, describe and justify the process that will be used:

6. **Will genetic information or samples be submitted to a national or international database because of this research?**

   ● **Yes**
   ○ No

   Indicate all databases with which the genetic information will be shared.

Database Short Name	Name
There are no items to display	

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244 - Data & Tissue Banking

PI: Hilary Coon Ph.D.

Submitted: 3/22/2011

Title: Genetic risk factors in suicide and depression

Data & Tissue Banking

1. **Select the items that will be banked:**
   - ☐ Biological samples
   - ☐ Data
2. **What type(s) of future research will be allowed on the data/samples?** Samples will be allowed to be used for any medical research. Approval of this research locally will be through additional specific IRB protocols; approval of this research through the NIMH repositories will be through a review process and oversight by the NIH. Approvals of Psychiatric Genomics Consortium analyses will be similarly approved through a PGC oversight committee, approval from a PGC scientific work group, and review of proposed data use by an OME representative. Data use at the PGC will only include summary level (not individual level) data.
   Hair analysis will be done at the Mount Sinai HHEAR Targeted Analysis Laboratory.
3. **Who manages the repository and where will the data/samples be stored?**
   University of Utah Department of Psychiatry Genetics Laboratory.
   NIMH repository (Rutgers University Cell and DNA Repository, RUCDR).

   Psychiatric Genomics Consortium Data Access Committee; summary-level data can be obtained from the PGC Suicide Working Group via our site through the permission process noted above (no individual level data are stored).
4. **Indicate whether the data/samples in the repository will be identifiable directly or through a code/link.**
   a. **Select one of the following options:**
   OPTION 3: All data/samples will be de-identified to all individuals who have responsibilities to manage or oversee the repository. No link or code will be accessible to or maintained by the repository.

   c. **If you selected OPTION 2 or 3 above, describe the process for de-identifying the data/samples:**

   ### Who will de-identify the data/samples?
   Samples will be de-identified by the UPDB. Samples will only be associated with an anonymous ID generated by the UPDB, an anonymous ID generated by the Office of the Medical Examiner, and by their anonymous CCTS LABID number.
   For the samples/data at NIH repositories, a de-identified numeric ID will be used.

   ### When will the data/sample be de-identified?
   Data will be made into a limited use dataset (month and year will be included for events) after being securely transferred by the Office of the Medical Examiner to honest data brokers within the UPDB, where it is then linked with UPDB genealogical, demographic, and co-morbidity data.
   All samples at the NIH repositories will be sent by research staff in a compltely de-identified state (deleting month data), and will be tracked by an anonymous numeric ID assigned by the NIH repository.
5. **Describe the procedures for participants to withdraw their data/samples from the repository. If participants will not be able to withdraw their samples, please provide an explanation:**
   Subjects are deceased and will therefore not be able to withdraw their samples. Families will not be able to withdraw samples of relatives because they will be stored anonymously.
6. **Will future research results or findings be communicated to the participants?**
   ○ Yes ● **No**
7. **Describe the procedures for other researchers to obtain data/samples from the repository for use in future research.**

Use of de-identified data in the NIMH data repository will occur only through application and approval as described above. Summary-level data will be available to PGC collaborator also through an application process as described above. DNA and tissue from these subjects is a limited resource. If available, collaborators will be able to obtain it for future research with approval of University of Utah investigators still involved in the research (if any), and ONLY with the approval of the RGE committee, the University of Utah IRB, the Utah Health Department IRB, and a fully executed material transfer agreement.

Use of samples in the NIMH repositories will be through an application/approval process governed by the NIH. Use of additional detailed pedigree structure data not in the repository by investigators requesting repository access will only occur with the explicit approval of the RGE. Use of fully de-identified, summary-level data for consortium analysis will first be approved using a process similar to that used by the NIH, with additional review of proposed data use by an OME representative.

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244
8. Resources and Responsibilities

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

8. Resources and Responsibilities

1. * **State and justify the qualifications of the study staff:**

Dr. Coon will take the administrative lead to ensure that all study and confidentiality requirements are met, and will manage other aspects of the project, including sample selection, sample management, transfer of samples to the genotyping core, management of genotype data, direction of data flow between Vital Records/OME and the UPDB, genetic and phenotype analysis of resulting de-identified UPDB information, scheduling of regular meetings with collaborators, directing overall data analysis and the interpretation and writing of grants and manuscripts. Dr. Coon will also manage the transfer of data and samples to the NIMH repository, with the assistance of Lisa Baird (senior lab manager), and data staff. Dr. Coon will ensure that annual review and approval occurs for an IRB with the Utah State Department of Health. Dr. Staley and Mr. Callor (external investigator) will create the lists of subjects to be merged with UPDB and receive a copy of the relevant data back from UPDB. They will help with study design and analysis of data issues of concern for the OME. They will also provide checking of subject status. Ms. Emily Sullivan will assist with study coordination, data and sample permissions and transfer, coordination among collaborators, and data management issues. Dr. Camp will provide expertise in the analysis of genetic sequence data in the families, and in follow-up studies. Drs. Shabalin, Monson, Mickey, Bryan, Bakian, Docherty, Kirby, Weir, Tharp, and Bilder will assist with data analysis and interpretation, and with writing of papers and grants. Drs. DiBlasi and Monson will work with genomic data and de-identified, aggregated familial risk data to determine risk factors, and will also assist with writing papers and grants. Drs. Kirby and Bilder will specifically study co-occurring risk of developmental disorders and de-identified analyses of cases linking to Dr. Bilder's study (IRB_00057455). Drs. Yurgelun-Todd, and Renshaw will provide expertise in phenotype issues, genetic pathways, and study design, and will also assist with writing of papers and grants. Drs. Keeshin, Yurgelun-Todd, and Renshaw will provide expertise on studies including psychiatric and pain medication data. Dr. Camp will assist with pedigree assessment, early life risk analysis, and with writing of papers and grants. Dr. Keeshin will direct studies of co-morbid PTSD risk. Dr Koppelmans apply his expertise in the study of neurodegenerative disorders, and how this could be related to suicide. Dr. Alex Shcheglovitov will lead experiments to create iPSCs from fibroblast tissue. Zhe (David) Yu and Michael Newman, who is part of the UPDB team of scientists, work with data linking and high-risk pedigree identification with UPDB data and resources. Drs. Yandell, Marth, Quinlan, and Eilbeck will collaborate with applying sequence variant pipelines and novel software analysis tools for detecting sequence variants associated with risk. All deaths listed as "unnatural" are transferred (including a large number of auto accident deaths, this transfer keeps the list from being identifiable as matching to our suicide cases.

External consultants: Dr David Brent will offer expertise in familial co-morbid risks, and in analyses and interpretation of data.

Dr. David Crockett is the PI of the IH IRB. This IH IRB will duplicate all study amendments in that IRB (a current approval letter is attached). Dr. Crockett will manage our collaboration with samples from the IH Biorepository. He will assist with secure transfer of information to the UPDB for data linking. Data management of the IH data will be facilitated by an IH 'honest broker' supervised by Dr. Crockett. The honest broker pulls records with ICD9/ICD10 codes for transfer to the UPDB with the IH Linkage ID used by the UPDB. The honest broker does not keep a copy of these data. After linking, the UPDB gives the analysis team only de-identified data that does not include any PHI or this IH Linkage ID, or any other id that could link to PHI.

Drs. VanDerslice and Yue Zhang will help us in a project to examine the effects of environmental exposures using de-identified DNA and data from suicide cases. In particular, we will investigate these exposures in individuals with genetic mutations in gene pathways consistent with particular vulnerability to environmental exposures. In order to do this project, Dr. VanDerslice will geocode our existing research records with already collected address information, then match already collected environmental data to these records. Drs. Vanderslice and Yue Zhang will use their analytical expertise to help us look for familial associations, associations with other medical and psychiatric co-morbidities that we are already approved to study, and genetic associations. These analyses will include psychiatric and inflammatory medications as potential mediating/moderating effects.

Drs. Qinggin Li of Janssen Research will assist with the creation of additional molecular data (Illumina PsychArray genotyping and sequencing), and with analysis of de-identified data, as previously specified. Her particular focus will be on co-occurring major depression and bipolar disorder. She will also study psychiatric medication data as contributing directly to risk and/or as a mediating/moderating variable.

Dr. Virginia Willour at the University of Iowa is an internationally known expert in genetic and epidemiological risks of suicide in individuals with bipolar disorder. She leads the American Foundation for Suicide Prevention grant; the

University of Utah will participate in a subaward of that grant for collaboration (now closed). Analyses of data will be done at the UU, and completely de-identified summary results will be worked on for publications. While an MTA was established to transfer de-identified data, analyses were instead only done at the University of Utah. Dr. Willour's research activities will be covered by her own institution.

Mr. Elliott Ferris will assist with the analysis of whole genome sequence data.

Seonggyun Han is now a postdoc in our group. He has training in analysis of genomic data and in machine learning methods. that will be applied to this study. We will assist him in devising projects using our de-identified data, allowing him to gain real-world experience with these methods.

2. * **Describe the training that study staff and investigators will receive in order to be informed about the protocol and understand their research-related duties and functions:**

Communication will be ongoing by phone and email. Meetings will occur regarding specific grants and analyses in conjuction with the study. These meetings are project-dependent, but happen at least bi-weekly with relevant subsets of investigators. Drafts of grants and papers are circulated among all study personnel for review.

3. * **Describe the facilities where the research activities will be performed (e.g. hospitals, clinics, laboratories, classrooms/schools, offices, tissue banks, etc.).**

Resources for bio-specimen collection, storage, and handling: Blood from decedents is collected from Utah suicides when possible by the Office of the Medical Examiner (OME) personnel. Samples are transported to the University of Utah Center for Clinical and Translational Science (CCTS) Translational Technologies and Resources core facility for DNA extraction. This core uses the Qiagen Autopure LS automated DNA extractor, a highly reliable and repeatable DNA extraction method. All samples have been stored in -80 freezers housed in the University of Utah Psychiatry Department lab. Freezers are equipped with automatic alarm systems. This lab has a computerized database indexed by anonymous sample ID number. The database can store freezer location, amount, concentration, and quality of each sample. All samples are bar-coded with an anonymous numeric OME ID and another anonymous sample ID assigned by the CTSI system. Suicide research in Utah builds on over 15 years of collaboration on projects involving University of Utah researchers, the Office of the Medical Examiner, government agencies including the Utah Department of Health, and many community-based organizations. The common goal is to learn as much as we can about the science related to death by suicide. We are fortunate to have an Office of the Medical Examiner that handles autopsies for the entire state, and a Medical Examiner (Dr. Erik Christensen) who is dedicated to suicide research. The OME also houses the Utah Medical Examiner Database (UMED); data from this comprehensive database also exists in the Vital Records database. With IRB approval, this project has obtained specific data that allowed linking of suicide decdents into extended pedigrees. Additional data in Vital Records/UMED that will be available for this project includes age at death, race, gender, manner of death, diagnosis at death reported to the OME, medications/drugs reported at death, OME-generated toxicology data, and zip code location at time of death. Data will be indexed by anonymous numeric IDs, and can be easily linked to DNA sample IDs. Utah Population Data Base (UPDB). The UPDB is one of the premier data bases for familial and genetic research. More than 30 years old, the UPDB is one of the world’s richest sources of in-depth information that supports research on genetics, epidemiology, demography, and public health. The UPDB includes the extensive genealogical records from the LDS (Mormon) Church, Utah Cancer Registry Records, Utah hospital discharge data, and public health data. It also has data linkage crosswalks to clinical records of patients treated within the Intermountain Health Care System and the University of Utah system (well over half the covered lives in the state and the surrounding region), Utah drivers’ license data, and an increasing amount of environmental data. The UPDB is the only database of its kind in the United States and one of few such resources in the world. Because of the privacy issues posed by pedigree research, the UPDB has strict approval protocols in place to govern research proposals. This study has gained approval for the proposed use of UPDB genealogical records and clinical codes to investigate co-morbidity. The UPDB data and analysis team are currently working with data linking and pedigree identification for our study. More than 185,000 Utah families were identified on "Family Group Sheets" from the archives at the Utah Family History Library. These sheets contain demographic and kinship information on three generations. For a family to be selected for the UPDB, at least one member had to have a vital event (birth, marriage, or death) on the Mormon Pioneer Trail or in Utah. These families have been linked across generations. These records are different from those available on the Web through FamilySearch or other publicly available genealogical databases and have been maintained as a resource only for biomedical researchers. The Genealogy File includes migrants to Utah and their Utah descendants, about 1.6 million individuals born from the early 1800s to the mid-1900s, birth and death information that occurs in any state or country, and linked multi-generation pedigrees Information from Utah death certificates is available from 1904 to the present. For records from 1904 to 1956, cause of death is available as a literal field; it has been coded to ICD Revision 10, using the 2000 Medical Data System and supplemented by hand coding. A wealth of medical information also appears in the UPDB. The hospital discharge data system is a database containing statewide, population-based healthcare information associated with 53 hospitals in Utah. Beginning in 2003, a new research tool was created for investigators studying the familiality of disease: a link between the medical information contained in the University of Utah Health Sciences Center (UUHSC) Data Resource Center and the family history information in the UPDB. This resource covers hospitals and out patients clinics associated with the University of Utah. The UUHSC Data Resource Center maintains a data warehouse with medical information on more than 1.8 million patient records. More than 1.3 million (74 percent) demographic records from the UUHSC have been matched

to a “person record” in UPDB. Neither the patient demographic records nor the medical information from the UUHSC are incorporated into the UPDB; each resource maintains its own data. However, the medical data and UPDB records can be accessed at the time that a research-specific project has been approved (this project has been approved already for access to codes for the disorders listed in the proposal). Computer: Each investigator and research assistant has a PC with the appropriate word processing and data analysis programs. The University of Utah Psychiatry Department also has two RAID-5 Linux servers with sufficient space and speed for analysis used by Dr. Coon. These servers are behind the University firewall, which also provides VPN ICSA-certified secure network access, network anti-virus scanning, content filtering, and authentication services. Human Genetics and Bioinformatics researchers have access to a large capacity, high-speed computer systems. Working spaces for Dr. Camp has multiple additional computers with large capacity for data storage as part of the University of Utah’s Genetic Epidemiology and Psychology Department computer network systems (respectively). Psychiatry, Human Genetics, and Genetic Epidemiology computers all have ample analysis software packages for familial linkage and sequence analysis. All three research groups are highly collaborative, and have active teams for analysis using existing tools and development of new software tools. The master database will be housed with the University of Utah CHPC Protected environment servers. Nightly and weekly backups are part of controlled recovery plan which includes offsite storage. The database servers currently support 16 TB of data space. Office: Investigators have offices within in the U of U Research Park or at the main Medical school campus, in close proximity to each other. All investigators have arranged computer access behind the university firewall so that data and analyses can be easily shared. There is ample office space available at the OME. Standard office equipment is accessible to all personnel. Laboratory, DNA extraction, storage, quality control, preparation (includes listing of major available equipment): DNA for this study is made at the University of Utah Clinical and Translational Science Center (CTSI) DNA facility using standard, automated procedures. Also available in nearby equipment rooms are: ultracentrifuges, refrigerated centrifuges, and tissue culture rooms. DNA is housed in the Molecular Genetics Laboratory in the Department of Psychiatry, University of Utah School of Medicine. In addition, DNA quality assurance, Whole Genome Amplification (WGA), and DNA preparation for assays will be done in this laboratory. The laboratory consists of a total of 1400 square feet of modern lab space in two adjoining rooms. The labs are fully stocked with all necessary equipment and supplies to carry out DNA purification, Whole Genome Amplification, DNA quality control and DNA storage. The DNA is stored in three locked -800 C upright freezers, that are monitored 24/7 by an auto alarm system. The lab is also equipped with a tissue culture room containing a Baker SterilGard Biological hood, two Forma C02 incubators and all necessary supplies needed to grow cells for DNA extraction. Other equipment in the lab consists of, four centrifuges, three balances, four microscopes, one 40 refrigerator, one -200 freezer, a chemical hood, and an autoclave. The lab has unlimited access to a NanoDrop ND-1000 Spectrohotometer to quantify and qualify all DNA samples. All equipment is on a yearly preventative maintenance schedule to ensure that is always in good working order and calibrated. The lab has two password-protected computer work stations that are connected to the University of Utah secure servers, so that all sample data, sample inventories and information are securely stored and backed up routinely. Laboratory for whole genome sequence (WGS) and whole exome sequence (WES) (includes listing of major available equipment). Sequencing will be done at the University of Utah Huntsman Cancer Institute Core Laboratory (www.cores.utah.edu/labs/DNASequencing/dnasequencing.html). This lab supports massively parallel sequencing on the Illumina platform. The lab has ample trained staff and instrumentation to support this project. Support equipment includes a NanoDrop spectrophotometer, and Agilent Bioanalyzer, six BioRad thermal cyclers, a GeneAmp PCR system 9700, a Bio-Rad CFX Connect qPCR instrument, an Agilent Bravo Liquid Handling system, an Illumina cBot, an Illumina MiSeq, two Illumina HiSeq 2000 instruments, and and Illumina Genome Analyzer IIx. A state-of-the-art database and LIMS system (GNomeEx) has been designed specifically to interface with the core. This system tracks, stores, visualizes, and analyzes experimental information and data from the sequencing experiments. The NIMH has been a leader in the NIH for data and tissue sample repositories. DNA and cell aliquots from this study will be deposited in the Rutgers University Cell and DNA Repository (RUCDR) managed by the NIMH. This repository houses thousands of samples from multiple NIH-sponsored studies. Researchers apply for access, and are granted through a formal approval process directed by the NIH. De-identified, aggregated data matching these samples will be deposited into the accompanying companion phenotype database. This system is governed by the highest standards of data quality and security. Finally, sequence and genotype data will be deposited into dbGAP. This NIH-sponsored genetic database contains molecular data from studies of many disease states, and allows creative research within and across disorders by approved scientific investigators. Access by researchers will follow the NIH protocols for access to DNA. Dr. Megan Williams' lab is equipped for studies of synapse characteristics. This includes the capacity to microinject individual neurons with fluorescent dye using iontophoresis, and then to image these neurons using a confocal microscope. Dendritic length/branching and spine density/shape will be analyzed using NeuroLucida360 software. Methods to detect density of specific classes of synapses will be piloted using immunostaining with antibodies that label highly specific classes of synapses; the capacity for these experiments also exists in the Williams lab. Dr. Alex Shcheglovitov is currently setting up a lab with the necessary resources for iPSC studies.

Hair Analysis will be conducted HHEAR Targeted Laboratory at Mount Sinai. An MTA will be put in place between the University of Utah and Mount Sinai.

4. * **Describe the medical or psychological resources available at this site (and other participating sites, if applicable) that participants might require as a consequence of the research. If not applicable, please state.**

NA

IRB_00044244 **Created:** 8/17/2010 1:10 PM

IRB_00044244
Documents and Attachments

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Documents and Attachments

If any of your documents (such as investigational brochures, sponsor protocols, advertisements, etc.) are not available in an electronic format, please scan and save them as PDF files or contact our office for assistance.

Naming Documents: Please use the title field to clearly indicate the content of each form. The name you enter will be listed on your approval letter. Use names that will differentiate from earlier versions.

Examples:
Consent Document Control Group 04/14/05
Consent Document Treatment Group 4/14/05
Sponsor Protocol 04/14/05 Version 2
Assent Document(Highlighted Changes)

Apple/Macintosh Users:MS Word documents must have a .doc file extension. See ERICA home page for instructions.

Print View: IRB Draft Protocol Summary

eProtocol Summary:

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

Consent Documents, Consent Cover Letters, Consent Information Sheets, Consent Scripts, etc.:

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

Parental Permission Documents:

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

Assent Documents:

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

VA Consent Documents:

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

Surveys, Questionnaires, Interview Scripts, etc.:

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

Full Protocol (company protocol, sponsor protocol, investigator-initiated protocol, etc.):

Name	Version	Date Created	Date Modified	Date Approved
There are no items to display				

Investigational Brochure (IB) for Investigational Drug or Drug/Device Package Insert:

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

Grant Application:

The Federal Government is a direct or indirect sponsor of your research. You are required to provide a copy of the grant proposal, grant award, or sub-award.

By submitting to the IRB, you are confirming the grant and the study protocol are consistent (Design, Study Population, Study Objectives and Goals, Test Interventions and Procedures, etc.)

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

Literature Cited/References:

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

Principal Investigator's Scholarly Record (CV/Resume):

Name	Version	Date Created	Date Modified	Date Approved
Coon Biosketch_2023.docx(0.01)	0.01	4/11/2023 2:59 PM	4/11/2023 2:59 PM	

Faculty Sponsor's Scholarly Record (CV/Resume):

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

Other Stamped Documents:

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

Recruitment Materials, Advertisements, etc.:

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

Other Documents:

Name	Version	Date Created	Date Modified	Date Approved

There are no items to display

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244**

Ancillary Applications

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Ancillary Application

This page should be used for submitting human research applications to the following ancillary committees:

Resource for Genetic and Epidemiologic Research (RGE) for access to the Utah Population Database (UPDB) and Utah Cancer Registry (UCR)
Phone: 801-581-6351
Website: https://rge.utah.edu/

Radiological Drug Research Committee Human Use Subcommittee (RDRC-HUS)
Phone: 801-581-6141
HUS Website: https://rso.utah.edu/committees/hus-rsc.php
RDRC Website: https://rso.utah.edu/committees/rdrc.php

Institutional Biosafety Committee (IBC)
Phone: 801-581-6590
Website: https://ibc.utah.edu

	ID	Name	Status
●	RGE_00000283	Genetic risk factors in suicide and depression	RGE Approved

IRB_00044244 **Created:** 8/17/2010 1:10 PM **IRB_00044244** Finish

PI: Hilary Coon Ph.D. **Submitted:** 3/22/2011

Title: Genetic risk factors in suicide and depression

Finish Instructions

Finish Instructions

1. **To view errors, select the "Validate" option at the top-left of the page. If you have errors on your application, you won't be able to submit it to the IRB.**
2. **Selecting the Finish button will NOT submit the application to the IRB.**
   You MUST select the "Submit" option on the workspace once you've selected the "Finish" button.
3. **If your study has a faculty sponsor: Once the PI submits the application, it will be sent to the faculty sponsor for final approval. The IRB cannot review the study until the faculty sponsor submits the application to the IRB.**