



Family Investigation of Nephropathy and Diabetes (FIND)

Revised April 1, 2004

List of Acronyms

AA	African American
ACD	Acid Citrate Dextrose
ACE	Angiotensin Converting Enzyme
ARP	Affected Relative Pair
AT1RA	Angiotensin 1 Receptor Antagonist (or blocker)
ASP	Affected Sibling Pair
BPBL	Peripheral Blood Monocytes from Fresh Blood
BUN	Blood Urea Nitrogen
Cols	Co investigators
CORBA	Common Object Request Broker Architecture
CWRU	Case Western Reserve University
DHHS	Department of Health and Human Services
DLT	Digital Linear Tape
DM	Diabetes Mellitus
DMSO	Dimethylsulfoxoide
DN	Diabetic Nephropathy
DRP	Discordant Relative Pair
DSP	Discordant Sib Pair
EA	European American
EAC	External Advisory Committee
EBV	Epstein-Barr Virus
ECDSA	Elliptic Curve Digital Signature Algorithm
EDTA	Ethylenediaminetetraetic Acid
ESRD	End Stage Renal Disease
FBS	Fasting Blood Sugar
FIND	Family Investigation of Nephropathy and Diabetes
FPG	Fasting Plasma Glucose
FSG	Fasting Sugar Glucose
FSP	Fasting Plasma Glucose
GADCC	Genetic Analysis and Data Coordinating Center
GB	Gigabyte
H-E	Haseman and Elston
HGAL	Human Genetic Analysis Lab
IBD	Identity By Descent
ID	Identification
IRB	Institutional Review Board
IS	Information Services

JHU	Johns Hopkins University
MA	Mexican American
MALD	Mapping Admixture Linkage Disequilibrium
MLS	Maximum Lod Score
MOO	Manual of Operations
MRC	Resource Center
NCI	National Cancer Institute
NIDDK	National Institute of Diabetes & Digestive & Kidney Diseases
NIH	National Institutes of Health
NIMH	National Institute of Mental Health
OGTT	Oral Glucose Tolerance Test
PBL	Peripheral Blood Lymphocytes
PBMC	Peripheral Blood Monocytes
PI	Principal Investigator
PIC	Participating Investigator Center
PID	Patient Identifier Number
QC	Quality Control
QTL	Quantitative Trait Locus
RBC	Red Blood Cells
RR	Relative Risk
S.A.G.E.	Statistical Analysis for Genetic Epidemiology
S/MIME	Secure Multipurpose Internet Mail Extensions
SSL	Secure Socket Layer
TAC	Total Allele Content
TDT	Transmission/Disequilibrium Test
UCI	University of California at Irvine
UNM	University of New Mexico
URN	Urine
UTHSCSA	University of Texas Health Sciences Center San Antonio

Table of Contents

1. INTRODUCTION.....	8
1.1 OVERVIEW OF THE MANUAL OF OPERATIONS	8
1.2 STUDY OVERVIEW	8
1.3 STUDY OBJECTIVES	8
1.4 COMMITTEE STRUCTURE	9
1.5 CONSTITUTIONAL POLICY	10
2 RECRUITMENT	11
2.1 RECRUITMENT OVERVIEW	11
2.2 STRATEGIES	11
3 SCREENING, CONSENT, AND ELIGIBILITY	13
3.1 SCREENING LOG	13
3.2 PARTICIPANT QUESTIONNAIRE	13
3.3 INELIGIBILITY OF PROSPECTIVE STUDY ENROLLEES	13
3.4 INFORMED CONSENT	14
3.4.1 Essential Elements of Consent	14
3.4.2 Description of Risks	14
3.4.3 Filing of Consent Form	15
3.4.4 Sample Informed Consent Wording	15
3.5 CERTIFICATE OF CONFIDENTIALITY	18
3.6 ELIGIBILITY INSTRUCTIONS	18
3.7 FIND STUDY PROTOCOLS	18
3.7.1 FIND FAMILY Study Proband Screening Protocol	18
3.7.2 FIND Family Study Eligibility Criteria for Genotyping	22
3.7.3 FIND MALD Protocol	24
4 ENROLLMENT PROJECTIONS	26
4.1 FAMILY STUDY	26
4.2 MALD STUDY	26
5 COLLECTION OF QUESTIONNAIRE AND BIOLOGICAL SAMPLES	27
5.1 DATA FLOW THROUGH THE FIND CONSORTIUM	27
5.2 COLLECTION OF QUESTIONNAIRES, MEDICAL CHART-REVIEW FORM AND FAMILY HISTORY DATA	28
5.3 COLLECTION OF SAMPLES FOR CELL LINE TRANSFORMATION, DNA EXTRACTION, CLINICAL PHENOTYPING AND ARCHIVING	28
5.3.1 Blood and Urine Sample Collection	29
Blood and Urine Sample Collection	29
5.4 DESCRIPTION OF BLOOD COLLECTION AND PROCESSING	30
5.5 DESCRIPTION OF URINE COLLECTION AND PROCESSING	30
5.6 THE SAMPLE SHIPPING INVOICE AND VENIPUNCTURE FORM	30
5.7 BAR-CODING	30
5.8 PROCESSING & SHIPPING OF SAMPLES FROM PIC'S TO NCI	31
5.8.1 Protocols for processing and shipping of blood and urine samples	31
5.8.2 Processing of samples at the NCI Central laboratory	35
5.8.3 Shipment of serum, plasma, buffy coat, whole blood and urine samples from NCI to the GADCC	37
5.8.4 Shipping of whole blood, serum, plasma, and urine to the Central Phenotyping Laboratory	37
5.8.5 Shipping of cell pellets from the NCI to the GADCC molecular laboratory	38
6. FIND DATABASE	39
6.1 DATA INTEGRITY	39
6.2 ERROR CHECKING	39
6.3 PHYSICAL INTEGRITY AND BACKUPS	39
6.4 QUALITY CONTROL OF DATA ENTRY	40
6.5 DATABASE SECURITY MODEL	40



6.6 LABELING AND BARCODES	41
6.7 IDENTIFIER AND BARCODE FORMATS	41
6.8 LOCAL DEMOGRAPHIC DATABASE	43
6.9 CENTRALIZED DATABASE.....	46
6.10 SCHEDULING AND INVENTORY DATABASE.....	51
6.11 DATA ENTRY FOR PIC	52
6.12 DATA ENTRY FOR NCI	53
6.13 GRAPHIC SHOWING INTERACTION BETWEEN PIC AND GADCC THROUGH THE CENTRAL DATABASE	56
7. PROFESSIONAL, SCIENTIFIC, AND TECHNICAL SERVICE	57
7.1 OVERVIEW OF SERVICES	57
7.1.1 Administrative Support.....	57
7.2 PERIODIC PROFICIENCY TESTING.....	57
8 CELL LINE IMMORTALIZATION	58
8.1 OBJECTIVE:.....	58
8.2 MATERIALS:	58
8.3 FROM BLOOD TO DNA--AN OVERVIEW	59
8.4 PROTOCOLS AND PROCEDURES.....	59
8.4.1 Maintaining B95-8 Marmoset Cells.....	59
8.4.2 Preparation of EBV Cocktail	60
8.4.3 Recipe for EBV Cocktail	60
8.4.4 Recipe for Cyclosporin -A Suspension (CSA).....	60
8.4.5 Setting up Feeder Layer Flasks	61
8.4.6 Counting Cells With a Hemacytometer and Trypan Blue Stain.....	61
8.4.7 ACD Blood Separation by Histopaque Gradient Method	62
8.4.8 Viably freezing PBMCs from Blood:.....	63
8.4.9 EDTA Blood Processing Protocol	63
8.4.10 Urine Processing Protocol.....	64
8.4.11 Serum Protocol	64
8.4.12 Fluoridated Plasma Protocol	64
8.4.13 Culturing from Frozen PBMCs	64
8.4.14 Maintaining Cells in Culture	65
8.4.15 Viable Freeze Protocol.....	66
8.4.16 N. Slow Rate Freezing (pre-programmed).....	66
8.4.17 Pelleting Cells from a Roller Bottle	67
8.4.18 DNA Extraction	67
9 DATA ANALYSIS.....	68
9.1B OBJECTIVES OF DATA ANALYSIS SUBCOMMITTEE	68
9.2 DEFINITIONS.....	68
9.3 TWO GUIDING PRINCIPLES	68
9.4 DATA ANALYSIS TECHNIQUES.....	68
10 CLINICAL QUALITY CONTROL.....	69
10.1 INTRODUCTION	69
10.2 MEDICAL CHART REVIEW	69
10.3 BLOOD AND URINE COLLECTION PROCESSING	69
10.3.1 Staff Certification Requirements.....	69
10.3.2 Participant Compliance with Protocol.....	69
10.3.3 Maintaining Proficiency.....	69
10.3.4 Packing Samples for Shipment from PIC to NCI.....	70
10.3.5 Replicate (phantom) Samples	70
10.4 SAMPLE RECEIPT AT THE CENTRAL PROCESSING LABORATORY (NCI)	71
10.4.1 Specimen Identification	71
10.4.2 Maintenance Procedures for Laboratory Equipment.....	71
10.5 CENTRAL PHENOTYPE LABORATORY QUALITY CONTROL (MEDSTAR).....	71
10.5.1 Specimen Identification	71
10.6 MATERIAL ARCHIVING	72

10.7 LONG TERM CONSISTENCY OF METHODS	72
11 MOLECULAR QUALITY CONTROL	73
11.1 OBJECTIVES AND GOALS OF THE MOLECULAR QUALITY CONTROL SUBCOMMITTEE	73
11.2 TO DEVELOP PROTOCOLS AND STANDARD OPERATING PROCEDURES FOR COLLECTION OF MOLECULAR GENETIC DATA FROM SAMPLES ASCERTAINED IN THE FIND STUDY WITH ATTENTION TO QUALITY CONTROL AND ASSURANCE. 73	
11.2.1 DNA extraction from lymphoblastoid cell lines.....	73
11.2.2 DNA extraction from whole blood or buffy coats.....	73
11.2.3 Assignment of DNA sample to 96-well plates for genotyping.....	74
11.2.4 Quality control for DNA concentration.....	74
11.2.5 Shipping of DNA aliquots to the PICs	74
11.3 TO PLAN A GENOME SCAN FOR FIND DATA WITH SELECTION OF APPROPRIATE CONTROLS AND REPLICATE SAMPLES.....	75
11.3.1 Plan I for the genome scan.....	75
11.3.2 Plan II for the genome scan.....	75
11.4 TO EVALUATE THE MOLECULAR RESULTS FROM THE GENOME SCAN AND THE MALD STUDIES.	76
11.5 TO IDENTIFY CHROMOSOMAL REGIONS FOR FOLLOW-UP AND CONFIRMATION OF LINKAGE RESULTS AND CANDIDATE GENES, AS PER THE RECOMMENDATION OF THE DATA ANALYSIS COMMITTEE.	76
11.6 TO PLAN POSITIONAL CLONING PROJECTS FOR THOSE REGIONS WHERE INITIAL AND CONFIRMATORY EVIDENCE FOR LINKAGE HAS BEEN OBTAINED.....	77
11.7 GENOTYPING PROTOCOL FOR MALD ANALYSIS OF MA DN PROBANDS AND CONTROLS.....	77
11.7.1 Quality Control:	78
11.8 GENOTYPING PROTOCOL FOR MALD ANALYSIS OF AFRICAN AMERICANS.	79
11.8.1 Quality Control:	79
12 INTELLECTUAL PROPERTY AND ARCHIVE POLICIES	80
12.1 FUNCTION AND PURPOSE	80
12.2 DUTIES OF THE INTELLECTUAL PROPERTY AND ARCHIVE SUBCOMMITTEE	80
12.3 CATEGORIES OF MATERIALS	80
12.4 BIOLOGICAL SPECIMENS.....	81
12.5 OWNERSHIP	81
12.6 BIOLOGICAL MATERIALS	82
12.6.1 Finite Supply	82
12.6.2 Infinite Supply	83
12.7 FUNDING	84
13 PUBLICATIONS	85
13.1 OVERVIEW AND OBJECTIVES OF THE PUBLICATIONS SUBCOMMITTEE.....	85
13.2 DUTIES OF THE PUBLICATIONS SUBCOMMITTEE	85
13.3 CATEGORIES OF PAPERS	86
13.4 IMPLEMENTATION OF POLICY AND THE REVIEW PROCESS.....	87
13.5 USE OF DATA.....	88
13.6 PRESENTATIONS TO VOLUNTEER PARTICIPANTS	88
14 SUBPROJECTS AND ANCILLARY STUDIES.....	89
14.1 OBJECTIVES OF THE SUBPROJECTS AND ANCILLARY STUDIES SUBCOMMITTEE.....	89
14.2 DEFINITION.....	89
14.3 CRITERIA	90
14.4 PROCEDURES	90
14.5 PUBLICATION.....	91
APPENDIX A: INVESTIGATORS	92
APPENDIX B: SCREENING LOG	98
APPENDIX C: DATA ANALYSIS	99
APPENDIX D: FIND MEDICAL QUESTIONNAIRE.....	116

APPENDIX E: FIND MEDICAL RECORD REVIEW 128
APPENDIX F: SUBCOMMITTEE MEMBERSHIP..... 133
APPENDIX G: STEERING COMMITTEE 138
APPENDIX H: CENTER NUMBER DISTRIBUTION ERROR! BOOKMARK NOT DEFINED.
APPENDIX I: LABORATORY FORMS..... 140
APPENDIX J: CERTIFICATE OF CONFIDENTIALITY 154
APPENDIX K: MEDSTAR PROTOCOL.....ERROR! BOOKMARK NOT DEFINED.

1. INTRODUCTION

1.1 Overview of the Manual of Operations

The Manual of Operations (MOO) provides detailed instruction of study procedures, details of each subcommittee, and instructions regarding all aspects of the study operations, inclusive of human specimen sample shipping and handling.

1.2 Study Overview

The Family Investigation of Nephropathy and Diabetes (FIND) study will have prospective subject entry. The National Institute of Diabetes and Digestive and Kidney Diseases at the National Institutes of Health (NIDDK/NIH) funds FIND.

FIND is a multi-center consortium, comprising eight (8) Participating Investigator Centers (PICs) and a Genetic Analysis and Data Coordinating Center (GADCC), established to study the genetics of Diabetic Nephropathy. The charge of the consortium is to acquire sets of families with well-characterized diabetic nephropathy, establish a secure master FIND database, and perform a genome scan to identify chromosomal regions linked with diabetic nephropathy. Analytic methods will include a) affected sibling pair (ASP), discordant sib pair (DSP), affected relative pair (ARP), and discordant relative pair (DRP) linkage analyses and b) Mapping by Admixture Linkage Disequilibrium (MALD) analyses. To expedite the mapping, and eventual cloning, of nephropathy genes, markers at candidate genes or human chromosomal regions containing putative renal failure susceptibility genes identified in previous genome scans, or syntenic to regions of renal susceptibility in model organisms, will also be examined. For candidate gene analysis both linkage and association (Transmission Disequilibrium Test (TDT)) methodologies will be used.

There are eight (8) Participating Investigator Centers (PICs). The GADCC and PICs, inclusive of each personnel directory, together with the NIDDK/NIH support staff and the External Advisory Committee, are given in APPENDIX A. This list, updated monthly, can be found on the FIND website (<http://darwin.cwru.edu/find>).

1.3 Study Objectives

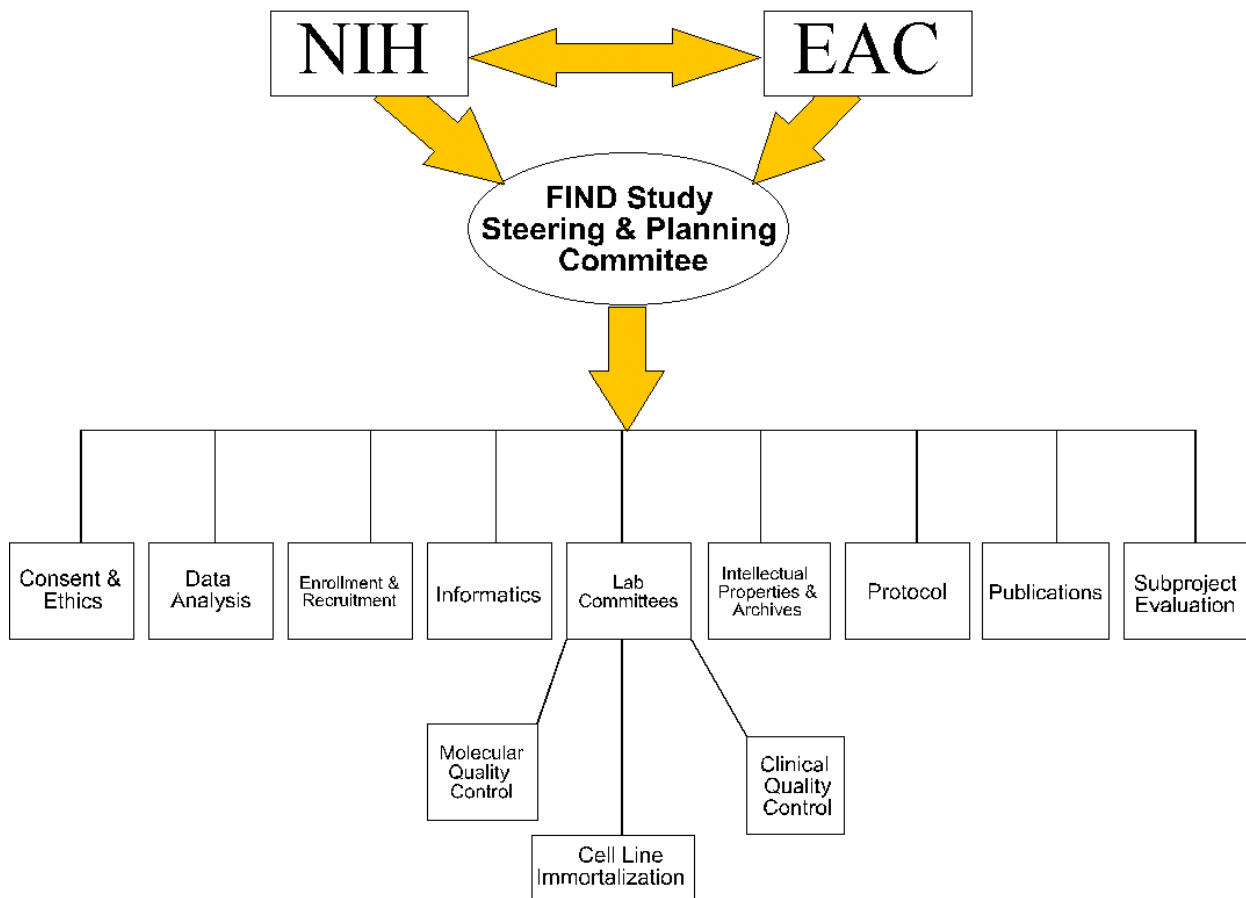
The primary aim of the FIND is to identify genes responsible for diabetic nephropathy and their linkage relationships, if any, to nephropathy.

Specific aims are to:

1. Organize a multi-centered consortium to undertake the recruitment of participants of several ethnicities.
2. Use Affected Relative Pair (ARP), Discordant Relative Pair (DRP), and Mapping by Admixture Linkage Disequilibrium (MALD) linkage analyses to scan the genome for linked chromosomal regions, and the transmission disequilibrium test (TDT) to evaluate candidate genes in those regions with evidence for linkage.
3. Use environmental data as a guide for possible risk factor effects on genetic susceptibility.

This is a clinical research cross-sectional study and longitudinal studies will not be addressed at this time, although may be at some later date.

1.4 Committee Structure



1.5 Constitutional Policy

The Steering committee is the governing body that has the authority to approve and amend the protocol, subcommittee membership, and manual of operations with a six out of ten majority vote. The Steering Committee is made up of the PI's of each center (the eight PICs and the GADCC), and the NIDDK representative (Appendix G). It will meet at least twice a year to establish any changes to the manual of operations, including subcommittee memberships.

2 RECRUITMENT

2.1 Recruitment Overview

The Enrollment and Recruitment committee is responsible for determining FIND enrollment and recruitment policy at the PIC level. Suggestions for changes to the enrollment policy and recruitment standards must be submitted to this committee. Recruitment strategies and monitoring will be done locally at each PIC. Each PIC will keep a record of potential participants interviewed and those who were entered into the study as enrollees. This will be monitored on the GADCC database. The steps are:

1. Contact potential index case. If an index case is not eligible, then enter tracking information (ethnicity) in central database and stop.
2. Obtain study materials from participants.
3. Enter participant into Global Participant Registry. This activates the ID and updates the tracking database.
4. Enter Identifying demographic data into local database.
i.e. name, address, phone number, social security number, birth date, physicians, and information about relatives.
5. Enter remaining data into central database. Double or multiple entry may be added

2.2 Strategies

The following recommendations are made to facilitate the recruitment of study participants:

1. Locate clinics, dialysis, or transplant units that have high numbers of diabetic nephropathy and/or diabetic participants.
2. Include a community spokesperson to act as a liaison for minority or demographically isolated ethnic groups.
3. Offer language translations on site at individual PICs.
4. Troubleshoot transportation barriers.

Sources of potential study participants from each PIC will be site specific. Each site will need to evaluate its own recruitment strategies that best suit the population to be studied.

1. Potential study participants will be recruited at dialysis and transplant units as well as through contacts from primary care physicians. Relatives of probands (index participants) will be ascertained through proband interview and questionnaire. Potential study participants will also be recruited via family gatherings such as picnics and other social gatherings.
2. Recommendations of primary care giver:
When approaching a potential candidate, the endorsement of study participation by her/his primary care giver, or other trusted source, can be

- persuasive. When necessary, a letter with a description of the study can be sent before any study personnel make personal contact.
3. Endorsement of study staff by community leaders, advocacy organizations, dialysis or transplant centers, can serve to increase sources of potential study participants and their families. Further endorsement of confidentiality for all participating enrollees can be stressed as an easement toward study participation.
 4. Study participants may receive compensation for their participation in the study. The more flexible and supportive that each PIC can be, the greater the likelihood of obtaining cooperation from all possible family members.
 5. People with nephropathy may find comfort in being connected to others with the same problem or by contact from support groups. The study coordinators can offer to assist these people by helping to locate any groups that may offer support to them and their relatives.
 6. Booths and tables at well-attended community events can offer exposure to the study, which can be a source of recruitment. Study staff may want to consider attending such gatherings to assist with the recruitment of entire families.
 7. A PIC brochure will provide information about the study, including eligibility and potential benefits.

3 SCREENING, CONSENT, AND ELIGIBILITY

The Consent and Ethics committee is responsible for determining and overseeing FIND consent policy at the PIC level. Suggestions for changes to the FIND questionnaires, or ethical policy must be submitted to this committee. Screening involves identifying likely study participants, describing study objectives and what participation involves, obtaining written informed consent, and determining eligibility. Determining eligibility for the FIND involves confirming various inclusion and exclusion criteria, ascertaining probands, ARPs, DSPs, and controls, etc.

3.1 Screening Log

Screening processes are tracked on the secure website and at each individual PIC. This enables each PIC to ascertain the level of effort, enrollment, and reasons for failure to enroll.

All log entries are completed within 24 hours of completion of eligibility determination.

Reason codes for ineligibility are found on the Screening Log sheet. They are:

- A. Did not sign informed consent: refusal to participate.
- B. Diagnosis not confirmed.
- C. Appropriate siblings not available.
- D. Judged unlikely or unable to follow study protocol.
- E. Ethnicity of parents or grandparent not suitable
- F. Spouse not available.

3.2 Participant Questionnaire

Once an initial verbal dialogue assesses proband eligibility, a questionnaire is filled out by the proband or Study Coordinator . If necessary, the study staff can read this questionnaire to the proband and write the answers for her/him. An interpreter will be on-site at the PIC for language interpretation as needed.

For family members, a questionnaire will also be completed and used as a tool to classify study enrollees based on a variety of demographic, and ethnic criteria.

3.3 Ineligibility of Prospective Study Enrollees

Refusal to participate in the study (give blood and urine samples) after determination of entry criteria renders a prospective enrollee ineligible. After the questionnaire is completed, the study staff determines a level of ineligibility based on the questionnaire information. Further ineligibility can arise if, after completion of the questionnaire, the study staff determines that the participant is not an appropriate candidate for the study.

3.4 Informed Consent

There is an informed consent that each study participant must sign. Each PIC will have its own informed consent form. In all cases, the investigating sites must follow the guidelines of the governing Institutional Review Board (IRB). This includes probands as well as all family participants. An example of the consortium's consent form is shown in section 3.4.4. Specimens, such as urine and blood, will be used in the genetic study to determine the genetic basis of nephropathy, diabetes, or related traits. Also, immortalized cell lines will be established from blood cells and will be stored at the NCI in addition to blood and urine samples. These stored biological samples will be available for all FIND PICs and third parties who have permission to perform research based on the FIND population. However, demographic information, such as name, social security number, and address/phone numbers will stay with the local PICs where the specimens were collected, and will not be available for (future) investigators obtaining these specimens from the NCI FIND registry (i.e., all specimen stored at the NCI will be anonymized).

Because we are obtaining written consent while eligibility is still being evaluated, it should be made clear to the subjects that signing the informed consent does not necessarily mean that s/he will be an eligible participant in this study.

3.4.1 Essential Elements of Consent

Each consent form must include the following:

A statement of the objectives of the study.

1. A fair explanation of the procedures to be followed and their purposes.
2. A description of any attendant discomforts and risks to be reasonably expected.
3. A description of any benefits to be reasonably expected.
4. A disclosure of any appropriate procedures or courses of treatment that might be advantageous for the subject.
5. A statement of confidentiality.
6. A statement of available information, an offering to answer any inquiries concerning the procedures.
7. A compensation statement.
8. An instruction that the person is free to withdraw consent and to discontinue participation in the study at any time without prejudice to the subject.
9. Explanation of whom to contact for answers to any questions concerning the study.
10. Signatures of the subject and additional signatures as determined by the local IRB.

3.4.2 Description of Risks

In addition to administering questionnaires, we are performing various clinical procedures on the subject, including blood drawing and urine collection. Any risks need to be described.

3.4.3 Filing of Consent Form

The subject will be given a copy of the signed consent form and the study coordinator/investigator will keep a copy in a locked secure place at the recruitment site or PIC. No other copy will be issued.

3.4.4 Sample Informed Consent Wording

1. You are being asked to be part of a research study on diabetes and kidney disease. This study is called FIND (Family Investigation of Nephropathy and Diabetes). This study is part of a larger study that is happening in several different cities. The National Institutes of Health (NIH) is heading this study and paying for it. Here at YYY the research is being done by Dr.XXXX (Principal Investigator).
2. The purpose of this study is to find inherited factors (DNA) that change a person's chances of getting kidney disease and other similar diseases. You understand that several of your family members may also be invited to be in this study. If your family members are asked to be in the study, they will be contacted (contact method will be PIC specific), told about the study, and asked if they will be willing to participate in this study.
3. You are being asked to:
 - a. participate in a 20 to 30 minute interview to review your medical history and that of your family members,
 - b. let researchers see your medical records and take information from them
 - c. have a blood pressure measurement
 - d. give blood for testing (a needle will be used to take about 3 tablespoons of blood from your arm)
 - e. have your urine tested.

If the blood or urine that you donate is not enough for the purposes of the study, you may be asked to give more. Blood and urine tests will be used to look for diabetes, to see how your kidneys are working, and for research on inherited factors (DNA research). If the researchers find that you have diabetes, kidney disease, high blood pressure or other medical problems, they will tell you. They will also tell your doctor if you sign a form giving them permission to do so.

4. DNA will be taken from some of the blood cells that you give. The leftover blood cells will be turned into cell lines that could live forever. This allows the researchers to make more DNA if they need it later. Your DNA will be studied to find out if you have inherited factors that change your chances of getting kidney disease and/or other related medical problems (e.g., eye disease, heart disease, and diabetes).

5. At some time the current and/or additional investigators will use your biological specimens to study the genetics of kidney and/or related disease. However, you understand that your identity will always remain strictly confidential.
6. There are no direct benefits to you for being involved in these studies. However, kidney disease runs in families and if DNA factors for kidney disease are identified, it may be of future benefit to your family to know this.
7. There are risks to your health and well-being from this study. You could find out about a serious disease that you did not know about. This may shock or frighten you, but it could also allow you to get treatment earlier than you might have otherwise.
8. Physical risks from this study include:
 - a. bleeding and bruising at the place on your arm where blood is taken,
 - b. temporary, local discomfort that accompanies all needle sticks, and
 - c. infection at the place on your arm where blood is taken (this happens to fewer than 1 out of 1000 participants).
9. Samples of your blood, urine, and DNA will be stored in a secure place. It is possible that a sample may be lost or broken (this happens for fewer than 1 out of 100 participants). If this happens, you may be asked to give another blood and/or urine sample. Mistakes may be made in labeling blood and urine samples (this happens for fewer than 1 in 100 participants). Because mistakes could be made, the researchers will need to double-check before they tell you that you have a disease or medical problem. This might mean that they would have to take blood or urine from you a second time.
10. If needed, a translator will be provided during the study to help you communicate with the researchers.
11. You will be informed of any medical problems found by tests used in this research. If you wish, the study staff will notify the physician of your choice of the study results. If the researchers learn something that endangers you, your child (children), or others, they may discuss it with you, or seek help for the problem.
12. You will be provided with a one time payment of \$XX.XX as compensation for your participation in this research. (Amount varies by PIC).
13. All information about your participation in this research will be kept confidential. Your research records will be kept in a secure locked place. You will not be identified by name in the actual record. All results from the DNA studies are for research purposes only and will be kept strictly confidential.
14. A Certificate of Confidentiality has been obtained from the Federal Government for this study to help ensure your privacy. The certification means that researchers cannot be forced to tell people who are not connected to this study,

including courts, about your participation unless you give written consent. This certificate will be kept on file to help protect your privacy.

15. You may ask questions regarding this research project at any time. Drs. X, Y, and Z will be available by phone (XXXXXXX) or beeper (XXXX) between the hours of 8:00 a.m. and 5:00 p.m., Monday through Friday. For questions or emergencies that arise after-hours or on weekends, you can call Drs. X, Y, and Z at (XXXXXXX) and ask for the faculty nephrologist on call or leave a message for the study staff. You may also contact XXX Institutional Review Board at (XXXXXXX) or Risk Management at (XXXXXXX).
16. You are taking part in this research voluntarily. You will not lose your medical care if you decide not to be in this project. You may quit this research project at any time, without any penalty or loss of services to which you are otherwise entitled. If you wish, your blood, urine, and/or DNA samples will be destroyed or returned to you. You understand that after one year it may not be possible to return these samples.
17. You will be given a signed copy of this informed consent form.
18. If your body is injured as a direct result of participation in this study, emergency medical care will be given to you by the XXXX.
19. By signing this document, you are not giving up your legal rights. State of XXXX laws exist that may help people who may have been treated carelessly, but that requires prompt notice of any claims. For further information, you may write or call either XXXXXX or (names of contact). By signing this document you indicate that you have fully understood, and agreed to participate in this research study.

Subject's name (type or print) _____ Date _____

I agree to participate in this research.

Subject's Signature _____ Date _____

I, as the parent or the guardian of _____, have read this consent form and agree that s/he may participate in this research.

Parent or Guardian's Name /Relationship _____ Date _____

Parent or Guardian's Signature _____ Date _____

Wit
ness to Signature _____ Date _____

Interpreter's Signature

Date

3.5 Certificate of Confidentiality

A Certificate of Confidentiality has been obtained from the National Institute of Mental Health/Department of Health and Human Services (NIMH/DHHS See Appendix K). This certificate protects the participant from data release and means that researchers cannot be forced to tell people who are not connected to this study, including courts, about participation without the subject's written consent.

The Genetic Analysis and Data Coordinating Center (GADCC) formally applied for this certification on behalf of the consortium. This certificate will be kept on file at each PIC as well as at the GADCC. This certification is valid until September 2004 but can be extended if necessary. The informed consent form of each PIC should contain paragraph 14 of the Sample Informed Consent Wording verbatim.

3.6 Eligibility Instructions

There are identical subject questionnaires for both probands and family members. Subject entry is prospective. Subjects who have been previously recruited for other genetic studies may be eligible for the study (subject to approval of the Steering Committee). Since these individuals may not have complete phenotyping data available, they will be first analyzed separately and only pooled with the rest of the data if the groups are found to be homogeneous.

3.7 FIND Study Protocols

The protocol committee is responsible for creating and implementing proband screening standards, ethnic recruiting standards, and enrollment standards for FIND and MALD. Suggestions for changes to the FIND or MALD protocol must be submitted to this committee.

3.7.1 FIND FAMILY Study Proband Screening Protocol

I. Definitions

A. Diabetes Mellitus (DM) Diagnosis:

In FIND, clinical diabetes diagnoses, obtained by medical history, will be accepted. If the history does not include treatment with diabetes drugs (i.e. control achieved with dietary or lifestyle modification), currently or in the past, the diagnosis must be confirmed by medical record review. For all categories of participants, the diabetes diagnostic procedures are as follows:

1. If, at the initial FIND blood draw, the participant is not fasting, measure estimated HbA_{1c}. If the participant has fasted overnight (≥ 8 hours with no caloric intake), both estimated HbA_{1c} and fasting plasma glucose (FPG)

should be drawn. This applies to all participants, whether or not previously diagnosed with diabetes, except those with ESRD, in whom estimated HbA_{1c}, but not FPG, will be measured.

2. ESRD patients will not be tested for diabetes, because their condition and its treatment may make diagnostic tests uninterpretable, although estimated HbA_{1c} will be measured as an analytic variable. Subjects without ESRD and without a previous diabetes diagnosis should be retested with FPG and an oral glucose tolerance test (OGTT) if the estimated HbA_{1c} or FPG is suggestive or provisionally diagnostic of diabetes, according to the table below. In recalling the participant for a retest, we should inform him/her that the estimated HbA_{1c} (and FPG if done) is suspicious for diabetes. Ideally, we will be able to retest in a fasting state, but if not, this needs to be arranged elsewhere. Follow-up testing is clinically indicated and necessary for FIND classification.
3. If the initial classification is “provisionally diabetic” (i.e. no previous diagnosis and FPG is ≥ 126 mg/dl or 2-hour post-load plasma glucose is ≥ 200 mg/dl) and the clinical condition allows it, a repeat FPG or OGTT should be done for confirmation. The criteria for diabetes, in persons without a previous clinical diagnosis, are:
 - Fasting plasma glucose ≥ 126 mg/dl (or venous whole blood glucose ≥ 110 mg/dl or capillary whole blood glucose ≥ 110 mg/dl) *
 - Random plasma glucose ≥ 200 mg/dl (or venous whole blood glucose ≥ 180 mg/dl or capillary whole blood glucose ≥ 200 mg/dl) * and symptoms (polyuria, polydipsia, polyphagia).
 - Two-hour post-load plasma glucose ≥ 200 mg/dl (or venous whole blood glucose ≥ 180 mg/dl or capillary whole blood glucose ≥ 200 mg/dl) * (OGTT).

* The measurement of glucose in serum is discouraged by the WHO. Unless the red cells are immediately removed to prevent glycolysis, serum samples should not be used for diagnosing diabetes. If only serum glucose values are found in the record, however, they should be interpreted as if they were plasma values.

Diabetes Testing in Participants Without ESRD

Clinical diabetes diagnosis	Initial Estimated HbA _{1c} (%) ¹	Initial FPG (mg/dl) ¹	Follow-up with FPG required	Follow-up with OGTT required ²
Yes	--	--	No	No
No	< 6.0	n.d.	No	No
No	6.0-6.9	n.d.	Yes	Optional ³
No	≥ 7.0	n.d.	No ⁴	No ⁴
No	< 6.0	< 110	No	No
No	6.0-6.9	110-125	Optional ³	Optional ³
No	≥ 7.0	≥ 126	No ⁴	No ⁴

Abbreviations:

HbA_{1c} = the A1c fraction of glycosylated hemoglobin. If total glycosylated hemoglobin is measured instead of HbA_{1c}, an estimate of HbA_{1c} should be used for diagnostic purposes.

FPG = fasting plasma glucose (mg/dl).

OGTT = oral glucose tolerance test performed according to World Health Organization recommendations (8-14 hour overnight fast, 75 g anhydrous glucose dissolved in water given orally, plasma glucose drawn 2 hours later).

-- " = this item is not taken into account under these conditions.

n.d. = not determined

Footnotes to table:

1. Go to the row with the greatest abnormality based on estimated HbA_{1c} or FPG, i.e. if estimated HbA_{1c} < 6.0 but FPG = 110-125, use the row for FPG = 110-125.
2. Perform FPG and 2-hour post-load glucose. If FPG can be measured immediately on site, it is permissible to stop the OGTT (i.e. not give the oral glucose) if the FPG ≥ 140 mg/dl (or fasting venous whole blood glucose ≥ 120 mg/dl or fasting capillary whole blood glucose ≥ 120 mg/dl).
3. For the conditions marked "optional", further testing by FIND investigators may be worthwhile for study purposes, as it may result in the participant being eligible as a diabetic sibling. If the investigators do not wish to pursue this testing, these levels of abnormality warrant referral to primary care for clinical purposes, but unless results of followup testing can be documented, such persons will be considered of unknown diabetes status for FIND analytic purposes.
4. In these instances, the participant will be considered diabetic for FIND purposes. Clinical practice, however, dictates follow-up by the investigators or by referral to confirm the diagnosis by glucose testing on a subsequent day.

B. Overt Proteinuria (historical data acceptable):

≥ 500 mg proteinuria/24 hours

or ≥ 300 mg albuminuria/24 hours

or urine protein (mg)/creatinine (mg) ≥ 0.5

or urine albumin (mg)/creatinine (mg) ≥ 0.3

II. Proband screening protocol

A. Proband must have: **1** through **3** below:

1. DM diagnosis

2. Diabetic Nephropathy (DN). Because the goal of this protocol is to identify probands with severe phenotypes, i.e. patients with DN who would be expected to progress to End Stage Renal Disease (ESRD) with relative rapidity, probands must meet one of the following three sets of criteria (2a-2c). Patients with renal biopsies that do not meet criteria described in 2a below, can be entered if they meet criteria defined in 2b or 2c below. Patients with DM and non-diabetic kidney disease caused by known monogenic mutation(s) are excluded.

2a. Biopsy proven DN:

1. Nodular and/or diffuse increases in the mesangial matrix accumulation *and*
2. Thickened glomerular basement membranes and/or arteriolar hyalinization *and*
3. Absence of mesangial immunoglobulin or paraprotein deposits by immunofluorescence microscopy, absence of amyloid deposits by Congo Red staining or electron microscopy, absence of electron dense deposits within the glomerular basement membrane or glomerular capillary subendothelial space

and overt proteinuria

2b. Patients with ESRD (including transplant) from presumed DN:

DM present for at least 5 years prior to the initiation of replacement therapy and retinopathy at any time

- or** DM present for at least 5 years prior to the initiation of replacement therapy and either ≥ 3 gm protein/24 hours, or a urine protein (mg)/creatinine (mg) ≥ 3.0 (historical data acceptable)
- or** retinopathy and either ≥ 3 gm protein/24 hours, or a urine protein (mg)/creatinine (mg) ≥ 3.0 (historical data acceptable)

2c. Patient with presumed DN but not ESRD:

Patient has diabetic retinopathy and either ≥ 1 gram proteinuria/24 hours or a urine protein (mg)/creatinine (mg) ≥ 1.0 (historical data acceptable)

- or** first detection of either ≥ 3 gram protein/24 hours or a urine protein (mg)/creatinine (mg) ≥ 3.0 gram (historical data acceptable) at DM duration ≥ 10 years

3. Two parents (regardless of phenotype) or at least one full sib who is potentially eligible to participate in the Family phenotyping/genotyping protocol below. If possible, affection status of the parents for diabetes and nephropathy will be determined. A medical questionnaire will be completed and entered into the database on these parents.

3.7.2 FIND Family Study Eligibility Criteria for Genotyping

I. Definitions (see FIND Family Study Proband Screening Protocol)

II. Family genotyping protocol.

A. For entry into FIND Genotyping Study, the proband's family must have the following relatives who are willing to participate:

1. Two parents (regardless of phenotype) from whom DNA is available **or** If possible, affection status of the parents for diabetes and nephropathy will be determined. A medical questionnaire will be completed and entered into the database on these parents.
2. At least one diabetic full sib who meets at least one of the following criteria:

Elevated albumin or protein excretion, defined by urine albumin(mg)/creatinine(mg) ≥ 0.03 or albumin excretion ≥ 30 mg/day or protein(mg)/creatinine(mg) ≥ 0.15 or protein excretion ≥ 150 mg/day. Note, this includes overt proteinuria defined in section 3.7.1 I.B. and lesser degrees of abnormality, commonly called microalbuminuria. If possible, the greatest urine protein excretion, urine albumin excretion, urine protein/creatinine ratio or urine

albumin/ creatinine ratio should be ascertained from the medical record (historical data).

- or** azotemia, defined as a creatinine ≥ 1.6 mg/dl for men, ≥ 1.4 mg/dl for women,
- or** unaffected by nephropathy, defined by duration of diabetes for ≥ 10 years AND normal albumin excretion (< 30 mg albumin/24 hours, or a urine albumin (mg)/creatinine (mg) < 0.03) AND without azotemia (defined as above) AND no historical evidence of kidney disease. Essential hypertension is NOT considered an exclusion criterion. If possible, the greatest urine protein excretion, urine albumin excretion, urine protein/ creatinine ratio or urine albumin/ creatinine ratio should be ascertained from the medical record (historical data). The subject's antihypertensive medications at the time of this determination should also be noted. Discordant siblings will have serum creatinine of < 1.6 mg/dl for men or < 1.4 mg/dl for women. IF the diabetic sibling has ESRD, a completed Medical Record Review form should be submitted documenting eligibility using criteria described in 3.7.1, section II 2b. Other affected relative pairs are not acceptable if the proband does not have a sibling that meets inclusion criteria.

B. If one of the criteria defined in IIA. is satisfied, the following, additional family members should be phenotyped (complete questionnaire) and blood and urine collected for laboratory determinations.

1. Parents: Recruit both, if available, or one, if only one available.
2. Additional sibs (if available): First, recruit all other sibs with DM, the oldest sib without DM and all sibs without DM but with nephropathy. Nephropathy is defined as either overt proteinuria or creatinine ≥ 1.6 mg/dl for men, ≥ 1.4 mg/dl for women.

Then, if only one parent is available and 4 sibs are not yet entered, recruit other nondiabetic sibs, giving preference to the older sibs, up to a total of no more than 4 sibs in the sibship.

If no parent is available and 5 sibs are not yet entered, recruit other nondiabetic siblings, giving preference to the older sibs, up to a total of no more than 5 sibs in the sibship.

3. Individual PICs may include half sibs and/or first cousins who meet the criteria described in A2, to develop extended kindreds. The relationship of the subjects to the proband will be provided by each PIC, and blood should be obtained from at least one of the connecting parents.

3.7.3 FIND MALD Protocol

Where possible, PICs that have proposed a MALD design are expected to enter families meeting the criteria described above, and PICs with a family design are expected to also obtain subjects appropriate for the MALD study.

In contrast to the family-based design, the MALD study will use a case-control approach and enroll only African-Americans or Mexican-Americans. The study design will be different in these two ethnic groups as outlined below:

3.7.3.1 African-Americans

African-American cases, spouses and controls must report only African-American ancestry. Their parents and grandparents may not be of other reported ancestry, including European-American, European-African, Asian, or American Indian. Triads will be preferentially recruited, when available. A triad is the index case, a child, age 18 or older, of the index case and the other parent of the child recruited. If it is not possible to recruit a triad, a case and spouse control will be recruited.

A. Diabetic Renal Disease Triads:

1. The index case will have diabetic nephropathy or ESRD (criteria as described above for the family study). If available, the child age 18 or older, recruited should have diabetic nephropathy or diabetes (to allow TdT analysis). The spouse control cannot have renal disease, as defined in the Family study for unaffected sibs (section 3.7.2), but may have diabetes.

B. Nondiabetic Renal Disease Triads:

1. The index case will have nephropathy or ESRD not due to diabetes or monogenic renal disease. Nephropathy is defined as a serum creatinine ≥ 2.0 mg/dl. If available, the child, age 18 or older, recruited should have nephropathy. The other parent of the recruited child cannot have renal disease, as specified in the Family study.

C. Case control study:

For the African-American MALD study, two sets of controls will be recruited: spouse controls and population-based hypernormal controls. Spouse controls are defined as individuals without renal disease, regardless of diabetes diagnosis, who have normoalbuminuria as per the Family study. Population-based "hypernormal" controls will also be defined as per discordant siblings in the Family study. If family history of renal disease is available, persons with a known history of a first degree relative with renal disease will be excluded from the study. Persons with a negative family history of renal disease and those whose family history of renal disease is

unknown will be entered into the study. Multiple hypernormal controls should not be related to each other or only distantly related to each other, i.e., second cousins or more distant. Hypernormal controls cannot be related to renal disease probands.

3.7.3.2 Mexican-Americans

Mexican-American cases and controls must report only Mexican-American ancestry. Their parents and grandparents may not be of other reported ancestry including European-American, European-African, Asian, or American Indian.

A. Case control study:

1. The MALD Study of Mexican-Americans will employ a case-control approach. Triads will not be recruited. Diabetic nephropathy cases will be defined using only the criteria for the family study proband described in 3.7.2 above. Hypernormal Controls will be persons with DM but not DN (as defined above for the AA MALD study).

4 ENROLLMENT PROJECTIONS

4.1 FAMILY Study

Institution	Race	Probands	ARP ¹	DSP ²	Total Number of Persons ³					Number coming from MALD Study
					Year 1	Year 2	Year 3	Year 4	Total	
CWRU	AA	171	133	92	226	226	226	226	904	8
	EA	72	38	21						
JHU	AA	100	100	0	37	38	38	37	150	50
Phoenix	American Indian ⁴	200	214	107	110	220	210	160	700	0
UNM	American Indian ⁵	140	100	40	77	77	77	77	308	0
Harbor-Davis	AA	25	30	20	126	126	126	126	504	100
	MA	50	50	50						
	EA	25	25	25						
UCLA	MA	300	100	198	350	350	350	350	1400	100
UTHSCSA	MA	200	100	100	220	220	220	220	880	100
Wake Forrest	AA	158	150	50	300	300	300	300	1200	15
	EA	152	140	40						
Totals		1559	1180	743	1446	1557	1547	1496	6046	373

¹ Number of distinct ARPs, e.g. 3 for 3 affected sibs.

² Number of distinct DSPs, e.g. 2 for a sibship with 1 ARP and 1 unaffected sib.

³ Total number for whom cells will be initially immortalized (all parents and all sibs, including any sibs who are not classified as either affected or unaffected), estimated to be 10% larger than the number who will be genotyped.

^{4,5} These two groups of American Indians are ethnically different.

4.2 MALD Study

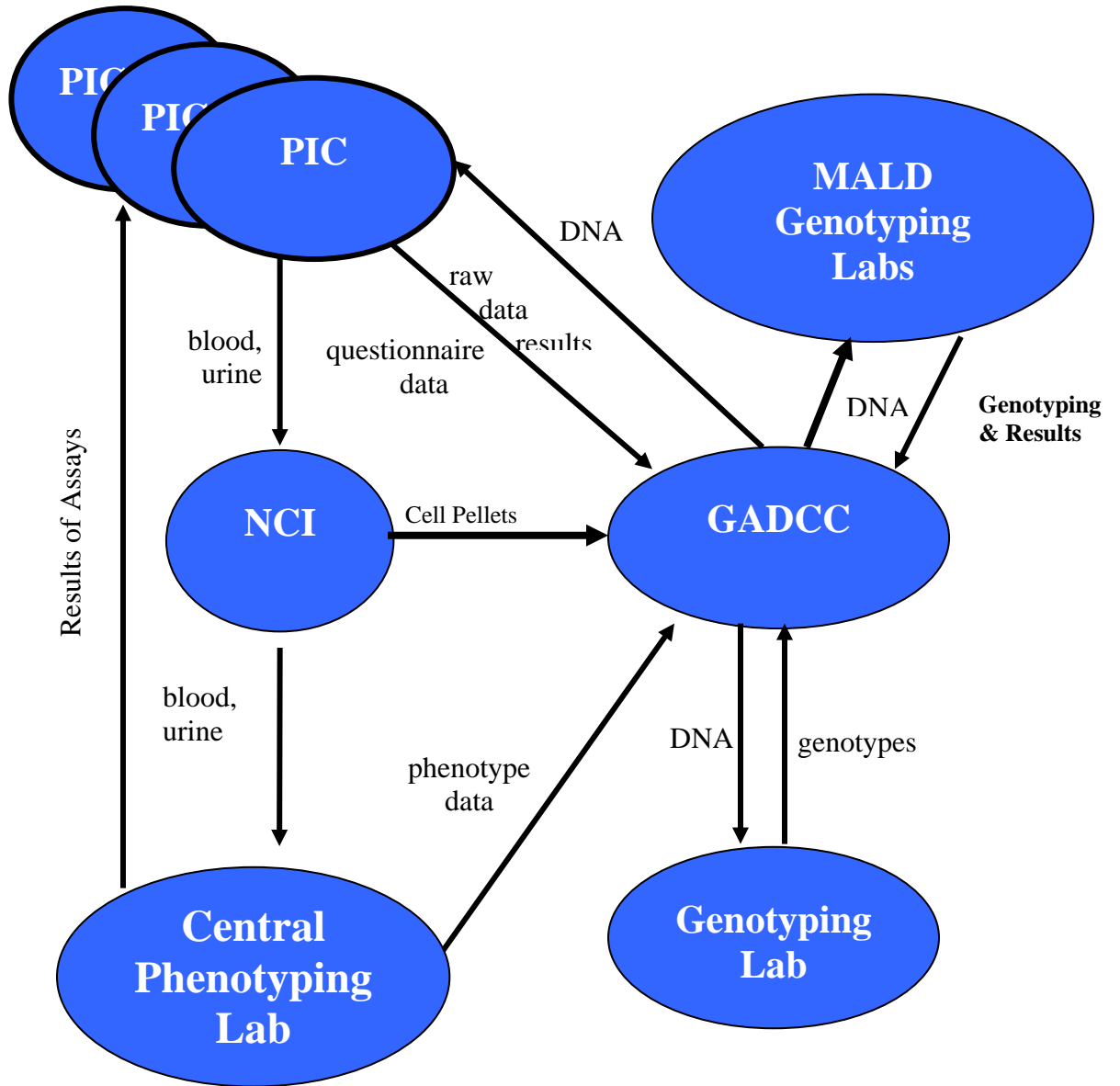
Institution	Race	Probands	Triads	Hyper-Normal Controls	Total Number of Persons ²					Number coming from Family Study
					Year 1	Year 2	Year 3	Year 4	Total	
CWRU	AA	15	15	5	12	13	13	12	50	8
JHU	AA	600	600	400	500	500	500	500	2000	50
Harbor-Davis	AA	0	200	60	388	388	388	387	1551	100
	MA	475	0	475						
UCLA	MA	25	0	25	6	6	6	7	25	25
UTHSCSA	MA	100	0	100	55	55	55	55	220	100
Wake Forrest	AA	15	15	0	11	11	11	12	45	15
Totals		1230	830	1065	972	973	973	973	3891	298

¹ May be a proband for a triad or a case-control pair.

² Total number for whom cells will be initially immortalized (all parents and all sibs, including any sibs who are not classified as either affected or unaffected), estimated to be 10% larger than the number who will be genotyped.

5 COLLECTION OF QUESTIONNAIRE AND BIOLOGICAL SAMPLES

5.1 Data flow through the FIND consortium



5.2 Collection of questionnaires, medical chart-review form and family history data

1. The following information will be collected on all eligible participants
 - a. questionnaire (see appendix D)
 - b. medical chart review (see appendix E)
 - c. family history (see appendix D)
 - d. venipuncture form (see appendix I)
2. If an eligible participant meets the study protocol criteria (see section 3.7) the participant will be invited to enroll in the study. To be enrolled in the study the participant must also provide biological samples as described in section 5.3
3. Once questionnaire, medical history and family history data and consent for procurement of biological samples are obtained, the individual should be logged in the FIND database (see informatics section 6.5). Non-identifying demographic information, questionnaire, medical history and family history data are also entered in the database as per instructions in the informatics section.

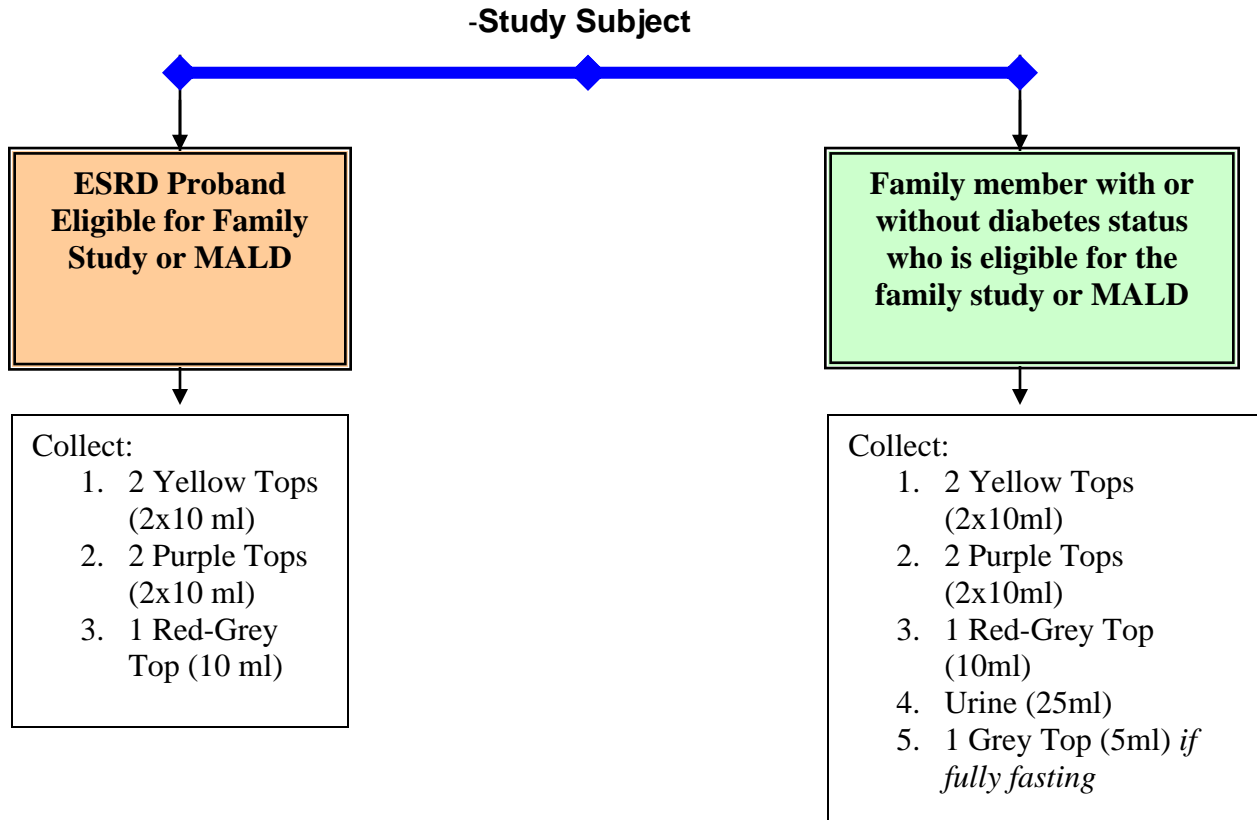
5.3 Collection of Samples for Cell Line Transformation, DNA extraction, clinical phenotyping and archiving

Each clinical site will collect the following types of samples from probands and family members:

1. Two 10ml ACD yellow top tubes (sodium citrate dextrose preservative Fisher # 02-684-26) for establishment of lymphoblastoid cell-lines (B-cell lines)
 2. Two 10ml purple top tubes (EDTA preservative Fisher #02-683-84) for
 - a. Glycosylated hemoglobin assay (estimated HbA_{1c})
 - b. Archiving of white blood cells (or buffy coat), red blood cells, plasma, and whole blood.
 3. One 10ml Red-Grey top tube (serum separator tube, no preservative, Fisher #02-657-11) collection of serum* (for BUN and creatine assays).
 4. One 5ml Grey top tube (Fisher #02-687-70) for fasting glucose assay (Potassium oxalate and sodium fluoride preservatives)*
 5. One 60ml sterile cup screw cap urine container (Fisher # 13-711-64).
- * Items 3, 4 & 5 are only collected on non-ESRD subjects as per the chart below.

CargoPak shipping containers will be used when sending samples from a PIC to NCI. There are two types of shipping packs available that are FIND certified shipping containers (Catalog Number LB06 CP-Bio-combination pack). Orders may be placed by calling: 919-554-3844 or 1-800-266-0652. The shipping containers will be returned to the PIC for re-use. Shipping materials for use with the CargoPaks are outline in Appendix I FIND Approved Lab Materials.

5.3.1 Blood and Urine Sample Collection



Blood and Urine Sample Collection

Subject	Glucose (Ped. gray top) <i>Only if fasting</i>	BUN/ Creatinine (Red-grey top)	HgbA1C (purple top)	Cell Transformation (yellow top)	Urine (25 ml)
PROBANDS					
ESRD					
DM	0	1*	2	2	None
Not ESRD					
DM	1	1	2	2	Yes
FAMILIES/ MALD TRIADS & CONTROLS					
DM	1	1	2	2	Yes
Not DM or unknown	1	1	2	2	Yes

Definitions: Grey top, contains sodium fluoride; purple top, contains EDTA; red-grey top, no additives; yellow top, contains sodium citrate dextrose (ACD tube).

* ESRD Probands – no BUN/Creatinine assay performed. Collection for archival purposes only.

5.4 Description of Blood Collection and Processing

On the day of the clinical evaluation, blood samples will be drawn by placing a needle non-traumatically in a large arm vein distended by a tourniquet filled to just above the diastolic pressure. The tip of the needle will be angled toward the center of the vein with a small cotton wad. The first tube collected must be the red-grey top serum separator (or clot) tube. Next will be two purple top EDTA plasma tubes, followed by two yellow-top ACD tubes, and one 5ml grey-top tube, only if the participant is fasting for ≥ 8 hours with no caloric intake. The purple top, yellow top and red-grey tubes will be gently inverted eight times to mix contents. The grey top tube and the two purple top EDTA tubes must be put on refrigerated until the assay can be performed. Blood drawn from an ESRD patient must be taken pre-dialysis, before heparinization.

The separation process must begin within 3 hours following the blood draw. The separation process (5.8.1) should take about 30 minutes and 8 to 10 samples can be processed in the same period of time.

5.5 Description of Urine Collection and Processing

When the participant is ready to void, a specimen cup (labeled with the participant's ID) is provided, and the participant is instructed to fill the cup if possible. If the sample is insufficient for processing, the participant is requested to void again in a clean container prior to leaving the PIC. Prior to processing, the technician records on the participant's Laboratory form whether a urine sample was obtained, the collection time of the initial (if more than one) urine sample.

5.6 The Sample Shipping Invoice and Venipuncture Form

To avoid ID errors in which information regarding a given participant's samples are written down on the wrong form, the technician should begin filling out the FIND sample shipping invoice and Venipuncture Form before the blood is drawn, verifying the ID of the participant. All sample containers are labeled with bar-codes and the participants barcode ID is affixed to the shipping invoice and Venipuncture Form. A copy of the shipping invoice will be kept at the PIC along with the Venipuncture Form. The same information will be entered into the Scheduling and Inventory database when the samples are sent to NCI. A copy of the invoice and venipuncture form should be kept at the PIC. Paper copies of invoices will not be sent to the GADCC.

5.7 Bar-coding

All participant samples will be labeled with a barcode that includes the following fields:

1. Study ID (2 digits)
2. Sample type (2 digits)

3. Sample sequence/draw number (2 digits)
4. Location/PIC code (2 digits)
5. Individual ID (6 digits)
6. Write ESRD on label for patients that have ESRD to “alert” NCI to process sample immediately

These barcodes will use the Code 128C encoding standard and will be readable by any of the supplied barcode wands and readers. The bar codes will be affixed to all biological samples, the patient questionnaire, the patient medical history form, and the patient pedigree form. Refer to Section 6.8 for barcode printing instructions.

5.8 Processing & Shipping of samples from PIC’s to NCI

5.8.1 Protocols for processing and shipping of blood and urine samples

I. Serum (Red-Grey Top Tube)

Allow the blood in the serum separator tube (red-grey tube) to clot for 30 minutes – 1 hour before centrifuging in a clinical (table top) centrifuge at 800 x g (~ 1500-2500 rpm) for 15 minutes to separate the serum from the clotted red cells. Centrifugation must be performed at room temperature. If centrifugation cannot be performed during this time the serum separator tube must be kept on ice and centrifuged no later than 3 hours after blood draw. The serum separator tube must be centrifuged so that the gel in the tube forms a seal between the clotted red cells and the serum. Remove the serum into one or two 4 - 4.5 ml cryovials (Nunc cryovials Fisher catalog # 12-565-173N or Corning cryovials Fisher catalog # 03-374-24 4ml) labeled with the barcoded ID number that indicate “serum” (SER) on the tube. Additionally, affix a red dot (Avery label AVE05790) to the tube to distinguish it as serum. Place the cryotube in a pocket of the 5 slot bubble wrap pouch and on the ice pack. This sample must be refrigerated (ice-pack) while being shipped. The bubble wrap pouch will also hold the Plasma cryovial and 2 EDTA (purple top) tubes.

II. Plasma for glucose assays (Grey Top Tube)

The grey-top fluoridated plasma tube is collected only on participants that have fasted for ≥ 8 hours with no caloric intake. Participants that are not fully fasted for ≥ 8 hours do not collect this tube. If a fluoridated plasma (grey-top) tube is collected for fasting glucose assays, it should be centrifuged at 800 x g (~ 1500 rpm) for 15 minutes within 30 minutes to 1 hour of the blood draw. Centrifugation must be performed at room temperature. If centrifugation cannot be performed during this time the fluoridated plasma tube must be kept on ice and centrifuged no later than 3 hours after blood draw. After, centrifugation, three layers will be observed (i) the top layer of clear/cloudy plasma, (ii) the ring of white cells at the meniscus of the layer of plasma and the red cells (also called buffy coat), and (iii)

the lower layer of red blood cells. Use plastic disposable pipettes to siphon the clear/cloudy plasma layer into a 4 - 4.5 ml cryovial (Nunc cryovials Fisher catalog # 12-565-173N or Corning cryovials Fisher catalog # 03-374-24 4ml). Do not siphon (suck up) any of the white blood cells in the buffy coat. Label the cryovial with the barcode ID of the patient that identifies it as fluoridated plasma (FDP). Additionally, affix a green dot (Avery label AVE05791) on the tube to distinguish it as plasma. Place the cryovial in the pocket of the 5 slot bubble wrap pouch, and place on the ice pack. The 5 slot bubble wrap pouch will also hold the glucose cryovial and 2 EDTA (purple top) tubes.

III. Other Samples (Purple Top, Yellow Top, Urine)

All other types of samples (ACD and EDTA tubes and urine samples) do not need additional processing at the PICs, but all vacutainers must be labeled with appropriate bar-code IDs that identify the sample as "ACD", "EDTA" or "URN". The remainder of the processing will occur at the NCI central processing and cell-line laboratory. All samples must be placed in the FIND certified container and shipped to the NCI central processing laboratory. Below are packing procedures for shipping samples from the PIC to NCI.

1. Place a refrigerated ice pack in the Styrofoam container of the CargoPak shipping container.
2. The 2 EDTA tubes should be placed in the 5 slot bubble wrap pouch with the glucose cryovial and any plasma cryovial.
3. Wrap the tubes into a roll and place inside of a plastic biohazard ziplock bag. Insert an absorbent strip in the bag.
4. The urine sample should be secured in the adhesive bubble wrap and placed inside of the biohazard ziplock bag with the glucose, plasma and the EDTA tubes. Seal the bag.
5. Place the biohazard bag in a Tyvek bag and seal.
6. Place the Tyvek bag in the styrofoam container with the cryovials and the EDTA tubes closest to the frozen ice pack. These samples (EDTA tubes, glucose, plasma, and urine) need refrigeration during shipment, but should not be frozen on reaching NCI. Therefore, they cannot be in direct contact with the ice-pack.
7. Insert the two ACD tubes into the pockets a 5 slot bubble wrap pouch. Place the tubes inside of Tyvek bag with an absorbent strip and place in the un-refrigerated side of the cargo pack (catalog #LB04 CP Bio-combination pack or LB06 CP Bio-combination pack). ACD tubes should be kept at ambient temperature (not > 90° or exposed to higher temperatures > 1 hour). If ambient storage is not possible, the ACD tubes should be securely packaged and placed in a container with an ice pack such that the ice pack is not in direct contact with the tubes.
8. Complete a paper copy of the shipping invoice. Pack a copy of the shipping invoice with the samples (see 5.8.1. V). Double check the participants barcode against the samples to ensure accuracy.



9. Ship the samples on the same day by Federal Express to the NCI Central laboratories, and alert the appropriate personnel to the arrival of the samples by entering a shipping note in the database (See Section 6.11).

IV. Scheduling shipment of samples using the database

Enter onto the web page log which day the blood will be arriving at NCI. Scheduling for shipment of blood samples can be done two weeks in advance. However, an electronic shipping invoice must be completed the day you are shipping specimens to the NCI. This will insure that the NCI laboratory is prepared to receive the shipment of blood and will also allow the NCI to notify a PIC if the shipment is lost or delayed. For further information please see section 6.11.

V. Labeling samples

All vials of blood samples are labeled with the participant's bar-coded ID. A shipping invoice is enclosed with each shipment giving the IDs for all sets of samples that are enclosed. Complete the top portion of the shipping invoice with the appropriate dates, shipping methods, and tracking number. Affix one participant identifier (barcode) to the shipping list for each participant in the box. Be sure to use a barcode that is for internal use and not one that is labeled for use on the specimens. In the column next to the barcode indicate whether the patient is a proband with ESRD, a proband without ESRD, a family member with diabetes, or a family member without diabetes. In the columns that follow indicate the draw number, the number of tubes that are being sent (mark the appropriate box), and the total number of tubes for each participant. Use the chart in section 5.3.1 to verify that the appropriate tubes have been collected based on the participants classification. This form should be completed by hand and not printed out of the database. Maintain a copy of the shipping invoice at the PIC. Pack the original copy of the invoice with the samples. The person unpacking these samples at the NCI laboratories verifies that the IDs on the vials match the ID on the shipping list. If any discrepancies are detected, the NCI laboratories will contact the PIC to resolve the problem.

VI. Policies on shipping samples to NCI

Blood vials shipped to the NCI Laboratories must be packed securely to avoid breakage.

Federal Express to: Russ Hanson or Cheryl Winkler
NCI/Frederick Cancer Research and Development Center
Building 560, Room 21-6A
Frederick, Md. 21702
301-846-5948



Federal Express biohazard shipping standards are followed if a PIC has prior knowledge that a shipment of samples contains infected blood (such as, HIV, HEP B or C). Federal Express procedures for packaging and shipping blood is available on their website: <http://www.fedex.com/us/services/packaging>.

To avoid delays in transit of samples that may impair the success rates for immortalization of lymphocytes, all samples must be shipped the day they are collected and must be received by the NCI laboratory by 10:00am the following day. Use priority service for a 10:00am delivery. The PICs will use government rates for shipping samples. (i) Complete the FedEx shipping label and mark as "bill recipient", (ii) use the government account number 0200-0407-0 or other government account numbers as applicable (NIDDK-Phoenix), (iii) Ship the package to NCI, (iv) NCI removes the shipping label to maintain weekly record of shipments received, (v) NCI forwards shipping labels to the GADCC for reimbursement, (vi) the GADCC maintains FedEx shipping record and collects the funds from each PIC for samples shipped to NCI, and (vii) the GADCC reimburses NCI for sample shipping.

FIND blood samples are mailed promptly to the NCI Laboratories Monday through Thursday. At present NCI can accept a maximum of 40 participants per week (~ 5/PIC). The current agreement is to ship no more than 8 participants per day, unless previous approval is given by NCI. All shipments are scheduled and recorded in the database. If a PIC schedules a shipment that exceeds the limit of 8 participants per day, the PIC must contact NCI for prior approval. The NOTES section of the REQUEST SHIPMENT form must contain a reason why the PIC exceeded the daily limit for shipping and the name of the person contacted at NCI. These pages will be monitored by the GADCC to ensure standard scheduling and shipping practices.

To avoid delays over weekends or holidays in delivering samples or in moving them to the Central Laboratory freezer once they are delivered to the NCI, all samples are shipped *only* Monday through Thursday (unless a holiday occurs during the week). NCI will post a list of unacceptable shipping dates in the database under the heading ANNOUNCEMENTS if they are scheduled to be closed for a holiday. NCI is unable to receive shipments on Saturday; therefore, **no shipments may be scheduled on Friday**. The NCI laboratory will notify the PIC if sets of samples are received late. If a pattern of delays is encountered with the delivery service a PIC is using, the PIC will change to an alternate delivery service. NCI will not be open for Sunday receipts in order to accommodate Saturday shipping dates. Shipments sent on Saturday will be received and processed Monday morning. **Saturday shipments should only be scheduled when normal shipping dates cannot be accommodated** (family reunions, holiday visits, etc). The NCI laboratories monitor the dates of blood drawing on samples they receive and notify the PIC and the GADCC if they receive samples that were shipped at a later date than that called for under this schedule.



VII. Receipt of samples at NCI

The NCI laboratories monitor the arrival condition of the samples sent from each PIC. If problems are encountered, the laboratories notify the PICs involved. If a pattern of sample damage becomes apparent that suggests a need to modify the materials used to ship samples (e.g. excessive leakage of a certain type of vial) or how samples are packed, the Clinical and Molecular Quality Control Committees take appropriate action.

Important Numbers and email addresses:

To talk with Mr. Russ Hanson: 301/846-5948

r_hanson@ncifcrf.gov

To leave a message on voice mail: 301/846-5363

Dr. C. Winkler 301/846-5747; email: winkler@ncifcrf.gov

To send a fax: 301/846-1909

Always include the following information in your message:

Number of samples arriving

PIDs (Patient Identifier#) of the samples

Date of shipment and expected arrival

Study Coordinator

Your name, phone and institution

ESRD Case Status

Enclose a hard copy with the shipment listing the PIDs, date of sample collection, institution.

5.8.2 Processing of samples at the NCI Central laboratory

1. Upon receipt of the samples at the Central Processing laboratories the sample inventory must be verified and the samples logged in the database as per instructions in the informatics section (section 6.5).
2. After logging the samples the lymphocytes from one of the two ACD (yellow-top) tubes will be transformed as per directions in the cell line transformation procedures given in section 8. The white blood cells from the other ACD tube will be frozen for future use. If the sample is successfully transformed, the white cells from the second ACD tube will be deposited into the NCI archive for the FIND study. If the sample does not transform successfully, the second ACD tube will also be used for transformation. The plasma will be archived at the NCI and the red cells from the ACD tubes will be discarded.
3. After receipt, the contents of one 10ml EDTA (purple top) tube should be inverted several times to mix the contents, and ten 1ml aliquots of whole blood will be obtained. One aliquot of whole blood (1.0ml) will be frozen at



-70°C for shipment to the Central Phenotyping Laboratory[†] for the glycosylated hemoglobin assay. Nine aliquots (1.0ml) will be obtained of which 6 x 1.0ml should be frozen at -70°C and stored at the NCI archive for future use and 3 x 1ml aliquots will be shipped to the GADCC monthly. If the tube contains less than 10ml of blood NCI will use the chart in Appendix I for aliquoting procedures.

4. The second 10 ml EDTA (purple top) tube should be centrifuged at 800 x g (~ 1500-2500 rpm) in a clinical centrifuge to separate the plasma, buffy coat and red blood cells. There should be at least 5 ml of plasma. The clear plasma should be divided into ten 0.5ml aliquots. 7 x 0.5ml aliquots will be frozen at -70°C and archived at NCI for future use. The remaining 3 x 0.5 ml aliquots will be shipped to the GADCC monthly. The buffy coat should be divided into 3ml of WBC and washed in PBS (phosphate buffered saline). Three equal size pellets should then be created. One pellet will be frozen to -70°C and archived at NCI and the remaining two pellets will be sent to the GADCC. Approximately 3mls of red blood cells will be obtained and aliquoted into 3 1.0ml aliquots. 2 x 1.0ml aliquots of red blood cells will be stored at the NCI archive at -70°C. The remaining 1 x 1ml aliquots will be sent to the GADCC. If the tube contains less than 10ml of blood NCI will use the chart in Appendix I for aliquoting procedures.
5. Serum should be divided into six 0.5ml aliquots and one 1ml aliquot. 3 x 0.5ml aliquots will be archived at NCI and 3 x 0.5ml aliquots will be sent to the GADCC. The remaining 1 x 1.0ml aliquot will be sent to the Central Phenotyping Laboratory[†] for the serum creatinine assay. If the tube contains less than 4 ml of serum NCI will use the chart in Appendix I for aliquoting procedures.
6. Plasma should be divided into one 2.0ml aliquot. The 2.0ml aliquot will be sent to the Central Phenotyping Laboratory[†] for the fasting glucose assay.
7. Urine should be divided into 2 x 1.5ml aliquots for shipment to the Central Phenotyping Laboratory[†] for urine creatinine, BUN, urine albumin/creatinine assay. The remaining urine sample will be split into two 10ml conicals and centrifuged at 1600 x g (2500 rpm) to precipitate solids. Five 3ml aliquots will be separated from the centrifuged urine. 3 x 3.0ml aliquots will be stored at the NCI archive. The remaining 2 x 3.0ml aliquots will be shipped to the GADCC laboratory.

[†]Samples will be accumulated on a weekly basis, batched by participant into a bio-hazard bag, and shipped in bulk from NCI to MedStar (Central Phenotyping Laboratory) by FedEx every Monday. The shipment will include a requisition of all items that are sent to Medstar for processing.

5.8.3 Shipment of serum, plasma, buffy coat, whole blood and urine samples from NCI to the GADCC.

1. The following samples will be shipped frozen to the GADCC on a monthly basis and an NCI generated inventory shipped with the samples.
 - i) 3 x 1.0ml whole blood EDTA
 - ii) 3 x 0.5ml plasma (EDTA)
 - iii) 2 pellets buffy coat (EDTA)
 - iv) 1 x 1.0ml red blood cells (EDTA)
 - v) 3 x 0.5ml serum (no preservative)
 - vi) 2 x 3ml urine (no preservative)*

*unavailable on ESRD patients

2. The GADCC staff shall check the samples with the NCI inventory, The samples will be separated by PIC and requested samples shipped back to each respective PIC monthly. Each transaction with the sample will be logged in the FIND database.

5.8.4 Shipping of whole blood, serum, plasma, and urine to the Central Phenotyping Laboratory

The NCI will send weekly shipments of samples to the Central Phenotyping Laboratory for analysis. Samples will be sent on Monday for receipt on Tuesday. If a federal holiday falls on Monday, samples will be sent out Tuesday for receipt on Wednesday. NCI will ship the following samples to the Central Phenotyping Laboratory for processing:

- 1 x 1 ml whole blood, for HbA_{1c}
- 1 x 1ml serum, for BUN, creatinine assay
- 1 x 2ml plasma, fasting glucose assay (where diagnosis is non-diabetic or unknown)
- 2 x 1.5 ml urine, for microalbumin, creatinine, and protein assays (not available on ESRD patients)

5.8.4.1 Packaging Procedures

NCI will package all the materials for the same participant in a biohazard zip lock bag. A Blood Sample Collection List (see appendix I) will be enclosed in the front pocket of the bag. All samples are then packed on dry ice (having previously been stored at -70°C and shipped through Federal Express to the Central Phenotyping Laboratory).

NCI will complete an electronic version of the Blood Sample Collection List in the FIND database for tracking purposes.

5.8.4.2 Sample Receipt at the Central Phenotyping Laboratory

Samples received at MedStar will be scanned into Crystal Reports using the FIND barcode ID. A new barcode will be assigned by MedStar. This barcode and the FIND Barcode ID will be linked together in a report generated through an Excel spreadsheet. Information regarding the condition of the samples upon arrival will also be logged into the spreadsheet. This spreadsheet will be sent electronically to the NCI and GADCC weekly.

MedStar will perform the assays upon sample receipt and results from the assays will be completed on the Friday after shipment. The test results will be printed out by PIC and MedStar will Federal Express hard copies of the test results to each site the following Monday. A Principal Investigator will be contacted directly if a test result appears within the alert level of glucose < 50 and > 450 for the fasting glucose assay. If the Principal Investigator is not available, the Study Coordinator will be used as a second point of contact. MedStar will use a contact list provided by the GADCC for shipping test results to the PICs. MedStar will send an electronic report of all test results to the GADCC at the end of the month.

MedStar will be subject to replicate sample shipments for clinical quality control purposes (See section 10.3.5). Results from the replicate sample assays will be compared to previous test results from the same participant. The GADCC will monitor the results of the replicate samples for discrepancies. A report of the findings will be sent to MedStar on a quarterly basis for their records.

MedStar will bill the PICs for assays performed on a monthly basis. The GADCC will provide a list of investigating centers with address and contact information.

5.8.5 Shipping of cell pellets from the NCI to the GADCC molecular laboratory

For extraction of DNA, cell pellets of transformed B-cells will be shipped monthly to the molecular laboratory at the GADCC after approximately 3 months of the start of the study. The GADCC will be responsible for the DNA extraction for the genome scan and candidate gene typing and for distribution of DNA to the PICs and MALD genotyping centers. The number of aliquots, freezer location and usage for each sample will be tracked on the FIND database. The samples should be shipped to:

Dr. Sudha Iyengar
Dept of Epidemiology and Biostatistics
MetroHealth Campus of Case Western Reserve University
Rammelkamp, R445
2500 MetroHealth Drive
Cleveland, OH 44109-1998.
Phone: 216-778-1822
Fax: 216-778-3280

E-mail: ski@po.cwru.edu

Laboratory contact: Karlie Reading
Lab Phone: 216-778-7916
E-mail: kdr@po.cwru.edu



6. FIND DATABASE

The Informatics sub-committee determines policy to monitor data integrity, error checking and security of the FIND database. This includes participant confidentiality, data maintenance, security, and details of the user interface for FIND and MALD. Suggestions for changes to the FIND database user interface must be submitted in writing for approval to the committee.

6.1 Data Integrity

All study data except identifying participant demographic information will be entered into the central database hosted by the GADCC. Each PIC study coordinator, recruiters and sample processing technicians will be responsible for entering recruiting tracking information, participant questionnaire forms and sample tracking and inventory information. These data will be immediately accessible by authorized FIND personnel from the FIND database and secure study website. Furthermore, access to much of the data will be restricted to the PIC that entered it. The GADCC will be able to monitor all website and central database transactions to verify their authenticity and review them for accuracy.

6.2 Error Checking

1. All study data entered into the database can be reviewed by the GADCC.
2. Regular and automatic data integrity checks will be run on the central database to identify data entry and coding errors.

6.3 Physical Integrity and Backups

All study computer servers and equipment will be stored in a secure, dedicated computing facility as a first line of defense against data tampering, accidental or intentional data corruption, and theft. To ensure the reliable storage of all data collected by the database system, a comprehensive system of data duplication will be instituted. Traditional tape backups will be performed using the Legato Networker Backup system, which allows backups of the database to be completed reliably without taking it off-line. Data will be stored to a high capacity backup system such as the Overland Data Model 410 DLT drive, which accepts ten 20GB cartridges and can continue to run unattended backups without human intervention for several months. Digital Linear Tape (DLT) media are among the most reliable available, with a rated shelf life of more than 30 years, 500,000 read/write cycles and a bit error rate less than 10^{-17} ("Tape Drive Reliability"; <http://www.quantum.com/src/whitepapers/reliability.html>). Full backups will be run weekly and supplemented with nightly incremental backups. Once tapes are written they will be stored off-site in a fireproof vault by the GADCC.

The Domino R5 Server will also be used to perform continuous database synchronization with the database server cluster located at the GADCC and

replicating PICs. Should the GADCC or a PIC replica database become unsynchronized because of hardware, software or network failure, the replication manager will automatically re-synchronize. This process is transaction based, so that each PIC can modify and add data to the database while any of the servers in the cluster are down. With this system it is possible to have data "collisions" when two database replicas have incompatible changes made while unsynchronized. Although rare, these replication inconsistencies will be detected and resolved with minimal human intervention.

6.4 Quality Control of Data Entry

Since transcription and data entry errors are frequently easiest to correct during initial entry, more data entry screens will perform validation checks against already entered data. Double entry will be added by a study coordinator on two (2) separate days for 10% of a PICs population. This information will be manually entered data once the basic design has been finalized. Most simple statistics will be available from the Oracle 8 database for each field without having to first export the data into another program. These statistics can be used in any query to generate online reports such as sophisticated data summaries, lists of suspicious values, and to define outliers. Creating a direct interface using the CORBA distributed object technology with existing pedigree and statistical analysis applications will perform more sophisticated data analysis. Thus data are examined "live" and can be updated continuously, instead of working on out-of-date snapshots of the data that would then require manual intervention to incorporate them back into the database. The same 10% of participants that are re-sent to MedStar for replicate sample quality control will be used for double entry.

6.5 Database security model

The FIND database will allow a flexible access model that allows each PIC to have secure access to its data, from any networked location. User level authentication will be required to gain access to any database components and only the appropriate privileges for that user will be granted. For example, only authorized database managers at a given PIC will have full control of the data generated by that PIC. Other PICs will not be able to view, edit or query others' data unless specifically granted access by the GADCC and the authoritative PIC. Such access may be granted for whole tables, subsets of tables, for certain variables, or even at the level of a single field at a time. This allows each PIC to be in total control of how its data are released while being collected. Authorized staff at the GADCC will have full access to all data collected and will be able to generate reports, perform data validation, and quality control procedures. The Domino R5 Server will manage these security features. It delivers proven, flexible security where it is simple to grant different levels of access to intranet and Internet applications, based on both individual and group roles. Fully integrated Public Key Infrastructure with support for X.509 certificates, Secure Multipurpose Internet Mail Extensions (S/MIME), Elliptic



Curve Digital Signature Algorithm (ECDSA), and Secure Socket Layer (SSL) ensure authentication, secured access and interoperability for Web applications.

6.6 Labeling and Barcodes

In order to accurately identify and track study materials, a systematic labeling system has been devised. All participant specific paper forms, biological samples and tracking material will have labels affixed to them as soon as they are collected or created. These labels will identify the study participant, and possibly additional information about the item labeled, with a human-readable identifier as well as a corresponding computer-readable barcode. User-friendly software is provided to print all study labels.

6.7 Identifier and Barcode Formats

There are currently only two identifier formats used to label study materials. These formats include several informative fields that are used to convey information on the identity of the participant, and the PIC where enrolled. Biological samples will include additional information, including an additional sample sequence number and type of sample. Both of these formats can be encoded, or represented, in two ways: a human-readable version that spells out some portions of the identifier, and a purely numeric, computer-readable barcode. This is done to minimize barcode size, because they must fit on extremely small objects, such as 1" cryotubes.

The first and simplest label format is the Participant ID, which consists of the following fields:

	Encodings	
	<u>Human-Readable</u>	<u>Barcode</u>
Study ID	Text	2 digits
Location Code	2 digits	2 digits
Participant ID	6 digits	6 digits
Length	variable	10 digits

The Study ID will always be 'FIND' or '01', in human-readable or barcode format, respectively. The Location Code indicates the center where a participant was recruited. Note that there may be several locations defined for each PIC, since many PICs are collaborations between numerous institutions and consequently have several distinct groups of recruiters. The table in appendix H lists the assigned location codes.

The six digit Participant ID field is specified by each PIC and no assumptions are made about its structure. PICs may, if they wish, choose one of the digits to represent a family code and the remaining digits as an individual number with a family. This option is implemented in the FIND Local Demographic Database as a convenience. See section 6.8 for more information.

The second label format is the Sample Identifier. It consists of all the fields in the Participant ID, as well as a sample type and a sample sequence number:

	Encodings	
	<u>Human-Readable</u>	<u>Barcode</u>
Study	IDText	2 digits
Sample Type	Text	2 digits
Sequence Number		
Draw number	1 digit	1 digit
Sample number	1 digit	1 digit
Location Code	2 digits	2 digits
Participant ID	6 digits	6 digits
Length	variable	14 digits

The StudyID, Location Code and Participant ID fields are as described above. The following table lists the currently assigned sample types:

<u>Sample Type</u>	Encodings	
	<u>Human-Readable</u>	<u>Barcode</u>
EDTA Blood (purple top)	EDTA	01
ACD Blood (yellow top)	ACD	02
Serum separator tube	SERSEP	03
Urine sample	URINE	04
Fluoridated plasma	FDP	05*
Fluoridated plasma separator tube	FDPSEP	06*
Buffy coat	BUFF	07
Plasma (EDTA)	EPLAS	08
Plasma (ACD)	APLAS	09
Red blood cells	RBC	10
Serum	SERUM	11
Viable monocytes	BPBL	13
EDTA Blood (purple top) [overflow]	EDTA2	14

*The FDP Tube coded 05 indicates a tube before centrifuge. The FDPSEP Tube coded 06 indicates post-centrifuge.

The Sequence Number field is composed of two sub-fields. The first is a single digit draw number, which is a sequential count used to differentiate samples collected at distinct times. The initial draw is always 1, and any subsequent draws should use the next sequential value up to 9. The second sub-field is the Sample number,

which distinguishes samples collected within a specific draw. Like the Draw number, it is also a single digit sequential value starting at 1. For example, if 4 tubes of EDTA blood are collected for a second time on a participant, the Sequence numbers of those four tubes will be '21', '22', '23', and '24'.

Examples:

A participant from the CWRU PIC with Individual ID '000401' would be labeled as:

Human-readable: FIND-01-000401
Barcode: 0101000401

Two tubes of EDTA blood, two tubes of ACD blood and a urine sample are collected from the same individual for the first time. Those tubes would be labeled:

EDTA Tube 1:
Human-readable: FIND-EDTA-11-01-000401
Barcode: 01011101000401

EDTA Tube 2:
Human-readable: FIND-EDTA-12-01-000401
Barcode: 01011201000401

ACD Tube 1:
Human-readable: FIND-ACD-11-01-000401
Barcode: 01021101000401

ACD Tube 2:
Human-readable: FIND-ACD-12-01-000401
Barcode: 01021201000401

Urine sample:
Human-readable: FIND-URINE-11-01-000401
Barcode: 01041101000401

6.8 Local Demographic Database

This database is a Microsoft Access database (97 or 2000). After opening the database, a switchboard immediately opens up and displays the choices for the user. Along with the data entry and look up forms, several report queries are available. The first two queries return the participants entered within the last week or month. The other query checks for duplicate records (based on birthdate, first and last names). There are a few notes on the data entry portion. The Family ID and Individual ID are required fields and must be filled in first. These two IDs have 6 character spaces between them and are configured by the Information Services (IS) Staff when the database is set up. The default is 4 for family and 2 for individual. The values of these fields are determined by each PIC. These two fields are combined with the study and

site IDs (previously assigned and specific to a PIC), to form the 10 digit participant ID. First Name, Last Name, and birthdate are also required fields and are used, in conjunction with Middle Name and Maiden Name, to verify that each individual is unique within the database. Therefore, Carol Lorimor, Carol A Lorimor, Carol Ann Lorimor, etc, are all perceived as different individuals. This feature enables the PICs to identify between family members or individuals with the same or similar names. The participant birthdate may be used to verify individuality. Finally, on the relatives tab, to maintain a less cluttered appearance, space for only one relative appears. After the name and relationship is filled in, space for the next relative appears. Currently, the number of relatives that can be entered is restricted to 20.

Start the MS Access database (Find Local Demographics)

A. To Add a Participant

1. Click the 'Add Participant' option on the menu.
2. Fill in the family ID (first 4 digits of the last 6 on the barcodes).
3. Fill in the individual ID (last 2 digits of the last 6 on the barcodes).
4. Fill in the First Name, Last Name, and the birthdate. These fields are required for checking uniqueness. Middle and Maiden Names are optional, but will also be used for uniqueness, if they are completed.
5. Fill in other fields as needed. The remaining fields are optional. To switch to other pages, click on the tab icon (to complete address, doctors, etc.).
6. If adding relatives, Name and Relationship are required. Space for an additional relative will appear after those two fields are completed.
7. When finished adding information for this participant, click the 'Save and Close' icon on the toolbar. If you wish to not save the record, click the 'Cancel and Close' icon.

B. To Review or Edit a Participant

1. Click the 'Participant Lookup' option on the menu.
2. Scan in the Participant ID (barcode) and enter.
3. Review or change any information EXCEPT the Participant ID. If changing the date, re-enter the entire date; otherwise a formatting error will appear. If an error does occur, re-enter the date.
4. When finished, click the 'Save and Close' icon on the toolbar. If you wish to delete the record, click the 'Delete and Close' icon on the toolbar: it will verify that you really want to delete a record. NOTE: You cannot cancel changes and start over. If corrections are needed, save and close the form. Follow the steps in "To Review or Edit a Participant" to make changes.

C. PIC Barcode Printing

Adobe Acrobat Reader is required to the barcode software to run properly. This software is available free on the internet via download. A link to the Adobe website is provided on the barcode print page. Also, a laser printer is required to print the labels so they will always be readable by the barcode scanner.

1. Click on Print Barcode Labels. FIND Barcode Generation Screen will appear.
2. Scan or type in the 10-digit FIND barcode. The barcode follows the following format:

01-(Site ID)-XXXXXX

and XXXXXX is determined by the site.

3. Type in the draw number. 1 for the first time samples are to be collected from a participant, 2 for the second, 3 for the third, etc.
4. Click on Generate Barcodes. Adobe Acrobat Reader will start and the barcodes will be visible in the screen.
5. Make sure the appropriate label stock is loaded into the printer. Click on File – Print. A print window will appear. Click on Ok to print.
6. Close the label printing by hitting the Escape key (Esc).

D. Errors and Enhancements

Errors and Enhancements provides a place for submitting possible errors, questions, and suggestions to the database so they can be prioritized and handled as quickly as possible.

1. Add new request to Errors and Enhancements
 - a. Click on 'Errors and Enhancements'
 - b. Click on 'Add New Request'
 - c. Select Error or Enhancement
 - d. Fill in screen where the error is occurring or enhancement is suggested
 - e. Complete the description of the error or enhancement. This field permits graphics as well as text.
 - f. Click on 'Save and Close' to save the registry document.
 - g. Click on 'Close Window' to return to the main menu.
2. Review request in Errors and Enhancements
 - a. Click on 'Errors and Enhancements'
 - b. Double click on the existing request.
 - c. One can review the document once it is opened.
 - d. Click on 'Save and Close' to close the document (from review)
Click on 'Cancel and Close' to discard changes.
 - e. Click on 'Close Window' to return to the main menu.

E. To Generate a Report

1. Click any of the following options on the menu
'Participants added within the last 7 days'
'Participants added within the last 30 days'



- 2. 'Verify no duplicate records in Participant Entries'
- 2. A table will open displaying the results of the selected reports. If a blank table is opened, then there are no data to display. Click the 'X' icon in the top right corner to close the table.
- F. Exit the local database
 - Click the 'Exit Database' option

6.9 Centralized Database

This database is in Lotus Notes. It contains the information available on the FIND website (staff, discussions) plus the non-identifying participant information. Within the centralized database there are several forms for each individual. They do not all need to be entered at the same time and there is flexibility of the order in which they are entered. Eventually, all forms need to be completed for an individual to be considered "entered" in the study. Navigational icons are included at the top of each screen (ie. Save and Close, Cancel and Close, Close Window, etc.).

- A. Open Centralized FIND Database
 - 1. Start Lotus Notes
 - 2. After entering your password, you will be brought into the FIND Main Menu
- B. Website Information
 - 1. Under the heading Reference Databases are the options: Discussion, Personnel List (by Name, Committee, Location)
 - 2. Click the Discussion option to see the minutes from various meetings.
 - 3. Click any of the three choices under Personnel List to see the staff that is associated with the FIND Study and pertinent information to contact them.

Global Participant Registry

This form is where the social security number and birthdate are converted to a 1-way hash and then compared to other hashes to verify uniqueness throughout all the centers in the study. Duplicate information in the recruitment tracking is entered by clicking the "Enter Recruitment Tracking" option. The only required field is the participant ID. This is how the record will be located at a later time.

- 3. To add participant to the Global Participant Registry
 - h. Click the 'Participant Registry' option on the menu
 - i. Click the 'Add New Participant' option
 - j. Scan Participant ID (barcode)
 - k. Fill in the fields (including Social Security Number and Birthdate)
 - l. Create the recruitment tracking record by clicking the 'Enter Recruitment Tracking' option at the top of the screen.
 - m. Click the 'Save and Close' icon to save the registry document.
 - n. Click the 'Close Window' icon to return to the main menu.

4. To review or edit a participant in the Global Participant Registry
 - f. Click the 'Participant Registry' option on the menu
 - g. Scan Participant ID (barcode) and click the 'Search' option.
 - h. Review document once opened, or click the 'Edit Document' option at the top to make changes.
 - i. Click the 'Save and Close' icon to close the document (after reviewing or editing.) Click the 'Cancel and Close' icon to discard changes.
 - j. Click the 'Close Window' option to return to the main menu.

Medical Questionnaire

The only required field on the Medical Questionnaire is the participant ID. This allows the form to be accessed at a later time. The fields may be entered in any order. When appropriate, the answer for one question will cause a second question to appear. Information can be entered or changed at a later time.

- A. To add a participant's Medical Questionnaire
 1. Click the 'Medical Questionnaire' option on the menu
 2. Click the 'Add New Medical Questionnaire' option
 3. Scan the Participant ID (barcode). This is the only required field on the document.
 4. Fill out the medical record questionnaire screen. As certain options are selected, additional questions will appear where appropriate.
 5. Click the 'Save and Close' icon to save the document. Or click the 'Cancel and Close' icon to delete the document.
 6. Click the 'Close Window' icon to return to the main menu.
- B. To review or edit a participant's Medical Questionnaire
 1. Click the 'Medical Questionnaire' option on the menu
 2. Scan Participant ID (barcode) and click the 'Search' option
 3. Review the document once it is opened or click the 'Edit Document' option at the top to make changes.
 4. Click the 'Save and Close' icon to close the document (after reviewing or editing.) Click the 'Cancel and Close' icon to discard changes.
 5. Click the 'Close Window' option to return to the main menu.

Medical Record Review

The Medical Record Review is similar to the Medical Questionnaire. The Participant ID is the only required field. Once again, as certain answers are selected, additional questions will appear.

- A. To add a participant's Medical Record Review
 1. Click the 'Medical Record Review' option on the menu
 2. Click the 'Add New Medical Record Review' option
 3. Scan the Participant ID (barcode). This is the only required field on the document.

4. Fill out the medical record review. As certain options are selected, additional questions will appear where appropriate.
 5. Click the 'Save and Close' icon to save the document. Or click the 'Cancel and Close' icon to delete the document.
 6. Click the 'Close Window' icon to return to the main menu.
- B. To review or edit a participant's Medical Record Review
1. Click the 'Medical Record Review' option on the menu
 2. Scan Participant ID (barcode) and click the 'Search' option
 3. Review the document once it is opened or click the 'Edit Document' option at the top to make changes.
 4. Click the 'Save and Close' icon to close the document (after reviewing or editing.) Click the 'Cancel and Close' icon to discard changes.
 5. Click the 'Close Window' option to return to the main menu.

Recruitment Tracking

If an individual is entered into the Global Participant Registry and the "Enter Recruitment Tracking" icon has been selected, an entry into the recruitment tracking has been created. Otherwise, an entry can be entered directly. If a group of people are contacted who do not qualify for entry into the study, they can be entered en mass by selecting 'group' and filling in the number of people in the group. However, certain questions would no longer apply (ie. age).

- A. To add a participant to Recruitment Tracking
1. Click the 'Recruitment Tracking' option on the menu
 2. Click the 'Add New Screening' option
 3. Scan Participant ID (barcode)
 4. Select individual or group record and state the number of people in the group if appropriate.
 5. Click the 'Save and Close' icon to save the registry document.
 6. Click the 'Close Window' icon to return to the main menu.
- B. To review or edit a participant in Recruitment Tracking
1. Click the 'Recruitment Tracking' option on the menu
 2. Scan Participant ID (barcode) and click the 'Search' option
 3. Review the document once it is opened or click the 'Edit Document' option at the top to make changes.
 4. Click the 'Save and Close' icon to close the document (after reviewing or editing.) Click the 'Cancel and Close' icon to discard changes.
 5. Click the 'Close Window' icon to return to the main menu.

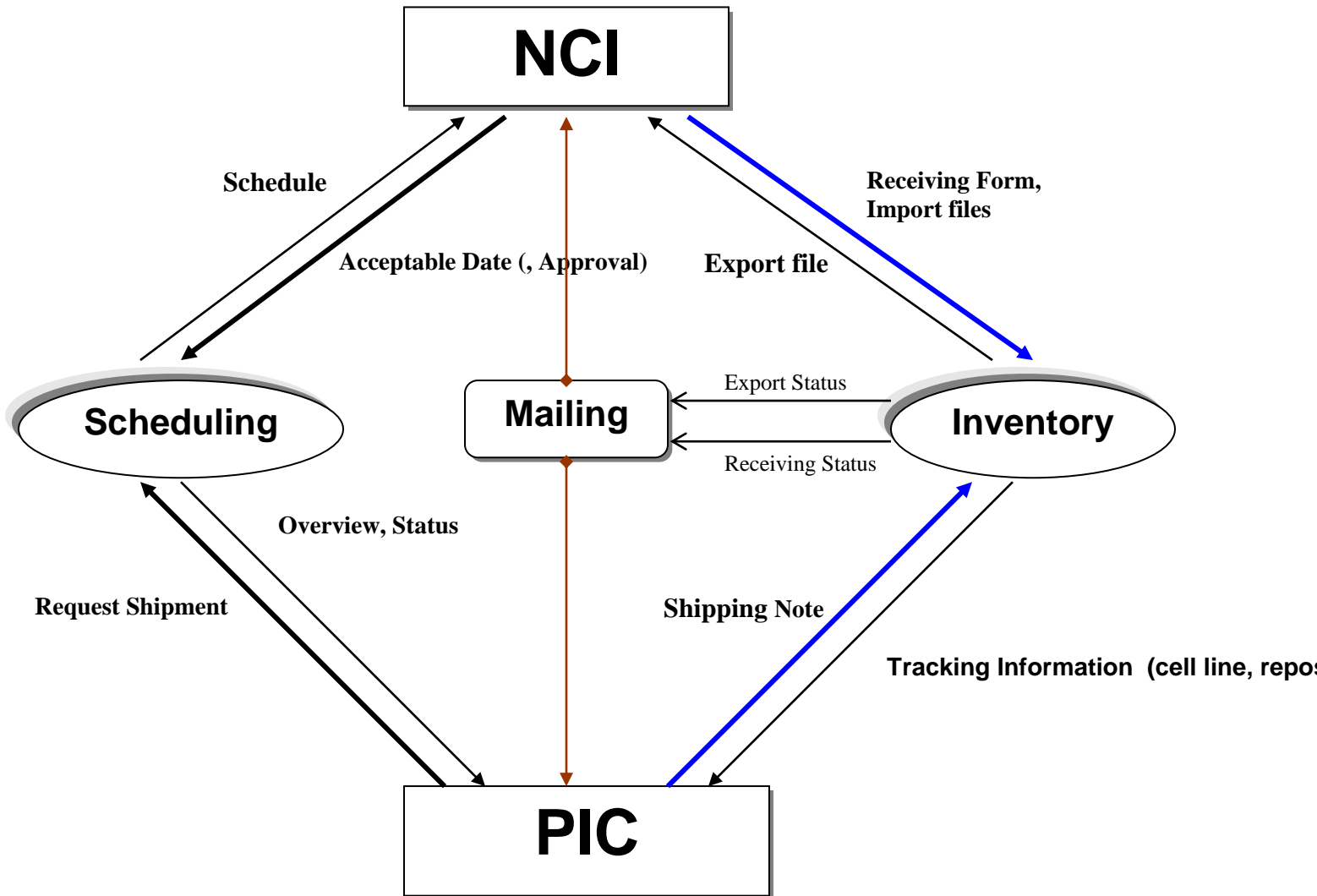
Pedigree Form

All individuals enrolled in the Family Study and AA MALD Triads or Diads must have a pedigree form entered. The system will automatically close MA Cases and all hypernormal controls.

- A. To add a participant or family to the Pedigree
1. Click the 'Pedigree' option on the menu
 2. Click the 'Add New Pedigree' option
 3. Scan the Proband's Participant ID (barcode) in 'Proband ID' section
 4. Select the appropriate designation from the pedigree chart (children, mother, father) and scan in the Proband ID and the remaining Participant IDs. Remember to separate children using a colon before adding more siblings.
 5. Click the 'Save and Close' Option when complete
- B. To add additional nuclear family participants to an existing Pedigree
1. Click the 'Pedigree' option on the menu
 2. Scan the Proband's Participant ID and click the 'Search' option
 3. Double Click on the Participant ID to open the Pedigree
 4. Select the appropriate designation from the pedigree chart (children, mother, father) and scan in the participant ID. Remember to separate children using a colon before adding more siblings.
 5. Click the 'Save and Close' option when complete
- C. To add extended Family participants to an existing Pedigree
1. Click the 'Pedigree' option on the menu
 2. Click the 'Add New Pedigree' option
 3. Scan in the Probands Participant ID in the Proband ID section
 4. Select the appropriate designation from the pedigree chart (children, mother, father) and scan in the participant ID. The Proband ID SHOULD NOT appear within the chart. It should ONLY appear at the top of the page under Proband ID. If collecting a 1st cousin, the connecting parent (aunt or uncle to the proband) must also be included on the pedigree. Remember to separate children using a colon before adding more siblings.
 5. Click the 'Save and Close' option when complete.
- D. To close a pedigree for genotyping
1. Click the 'Pedigree' option on the menu
 2. Scan the Proband's Participant ID and click the 'Search' option
 3. Double Click on the Participant ID to open the Pedigree
 4. Review the pedigree and verify its accuracy
 5. Click the 'Close Family' option when complete
 6. A box will appear that states 'Done. Press the 'Save and Close' button to exit. Click 'OK' option
 7. Click the 'Save and Close' option when complete.
- E. Entering a Venipuncture Form
1. Click the 'Link to Scheduling and Inventory' option on the menu
 2. Click the 'Venipuncture Form' option on the menu on the left
 3. Click 'Add a New Form' from the grey menu bar at the top
 4. Enter the Participant ID and information from the paper form

5. Click 'Save and Close' from the grey menu bar at the top

6.10 Scheduling and Inventory database
Data Flow Diagram



6.11 Data Entry for PIC

PIC Scheduling Database

The Scheduling Database is used to request sending a shipment and to view other PICs' request statuses. Each PIC can schedule its shipment one to four weeks in advance. All scheduled shipments should be confirmed at least one week before the scheduled date.

1. Menu Description of Schedule Shipment

The *Schedule Shipment* screen has three sub-menu options: Calendar, Request Log, and Summary. The Calendar and Request Log view are automatically monitored by the GADCC and removed when the scheduled date has expired. The expiration date of a scheduled date screen is one day after the sample was scheduled for shipment.

- A. Calendar: Each PIC can schedule its shipment and view the shipping statuses of other PICs. After an actual shipment is complete a PIC can check its shipment status, as well as the status of other PICs.
- B. Request Log: This is a text view of the shipping calendar and an alternative way of scheduling shipment.
- C. Summary: This menu provides a summary of information for scheduling shipments. A PIC can see the summary report categorized by Month/Week, PIC, and Date. This includes the total number of shipment requests by each category.

NOTE: The summary view holds all scheduled information for the current month.

2. Creating A Request Shipment Form

- A. Click the *Calendar* or *Request Log* option on the menu
- B. Click the "Request Shipment" icon in the upper-left corner
- C. The Request Shipment form will appear
- D. Enter the number of samples being shipped in the *# of Samples* field.
- E. Select *Scheduled Date* from a pre-assigned date list in the dialog box
- F. Leave priority as default, unless sample's priority is more important
- G. Click the Save & Close icon
- H. Check your result in the *Calendar* or *Request Log* option

PIC Inventory Database

Each PIC must create a *Shipping Note* through the central database before actual shipping occurs. A Shipping Note consists of a FedEx tracking number, shipping date, and all Participant IDs in the shipping box.

1. Description of Process Sample

The inventory system has three menu options: Shipping/Receiving, Sample Information, and Sample Process. Each PIC creates a Shipping Note from the Shipping/Receiving menu. This is the only menu option where a PIC needs to enter sample data into the central database. The next day, the PIC can check the receiving status. The system will generate an e-mail to each PIC that sent a shipment the previous shipping day to notify them of the status of the shipment. If a PIC receives an e-mail notice that the shipment is missing the tracking process should begin immediately.

- A. Call Mary Thompson at NCI to verify that the shipment is missing
- B. Use the Federal Express tracking number and contact FedEx regarding the status of the shipment.
- C. If sample is en route, contact Mary Thompson at NCI and inform her of the status
- D. If sample is lost, begin steps to re-collect the patients specimens.

2. Creating Shipping Note

- A. Click the *Process Sample* option on the menu
- B. Click the *Shipping/Receiving* option
- C. Click the *-** icon from the upper-left menu bar
- D. The *Shipping Note* form will appear
- E. Type in Shipping Tracking Number from FedEx Airbill label.
- F. Scan all Participant IDs (Barcode) through scanner
- G. Leave Shipping Method and Shipping Date as default, unless it is necessary to change: Shipping Method default is FedEx and Shipping Date default is the current date.
- H. Check the ESRD box if the participant is an ESRD patient.
- I. Enter Draw Number: First draw is 1, second is 2, etc.
- J. Repeat steps e. – i. As needed
- K. Click the 'Save & Close' icon
- L. Check your results on the *Shipping/Receiving Log* view or *Calendar Shipped* view

6.12 Data Entry for NCI

NCI Scheduling Database

The NCI will enter acceptable dates for monthly shipping. Also, if there is a request overflow NCI may approve some of the requests through the *Request Log* menu option or by phone. NCI will frequently check the status of requests and shipments through the *Calendar* options.

1. Entering Acceptable Date

- A. Click the *Schedule Shipment* option on the menu and click the *Calendar* option
- B. Click the *Calendar-Overview*
- C. Click the *Enter Acceptable Date* option
- D. Schedule Shipping Date for PIC
(Note) Shipping Date is one day before receiving date
- E. Select an Acceptable Date by clicking the field on the calendar
- F. Click the week option
- G. Click the *Save and Close* option

3. Entering Approve Shipment

- A. Click the *Schedule Shipment* option on the menu and click *Calendar* option
- B. Click the *Calendar-Overview* option
Or *Calendar-Requested*
Or *Request Log*
(Note) all three menu options have Entry Point for *Acceptable Date*
- C. Click the *Approve Shipment* option
- D. Enter the # of *Approved Samples* that NCI needs to approve
- E. Click save and close

NCI Inventory Database

NCI will use this system to confirm sample receipt and provide detailed information on received tubes. This sample information includes each tube's barcode and the date that it was received. The receipt of import files from NCI updates the sample information. The data is always on step behind because the sample information entry is coupled with the exporting of files to the NCI for the shipping note. Primary barcode errors are detected by the NCI, however, each PIC must check for labeling errors before shipping occurs.

1. Description of Process Sample

The *Process Sample* section has three options: Shipping/Receiving, Sample Information, and Sample Processing.

- A. Shipping/Receiving: After receiving the shipped samples, NCI will confirm the arrival by creating the Receiving Form.



- B. Sample Information: In addition to confirmation upon receipt, the NCI will provide each tube's barcode received on a participant basis by providing import files to the central database. Labeling and shipping errors will be checked through this step.
- C. Sample Process: NCI will process each tube's sample, and return a summary process log to the central database.

2. Creating Receiving Form

- A. Click the *Process Sample* option on the menu
- B. Click the Shipping/Receiving View option
- C. Highlight a Shipping Note that the NCI received by single clicking on it
- D. Click the *New Receiving Form* option
- E. The Receiving Form will appear
- F. All necessary information is automatically inherited from the selected Shipping Note

3. Exporting Data

From Scheduling System:

- A. Under Schedule Shipment section, Click the 'Calendar-Shipped; option
- B. Click the 'Export File' icon on the menu bar
- C. System will save shipping information file into directory, c:\find
- D. Each export file naming convention: EF_ReceivingDate_PIC
Example) If CWRU and Wake Forest sent samples on August 20, the system will create two files, EF_821_CWRU and EF_821_WF, in c:\find directory.

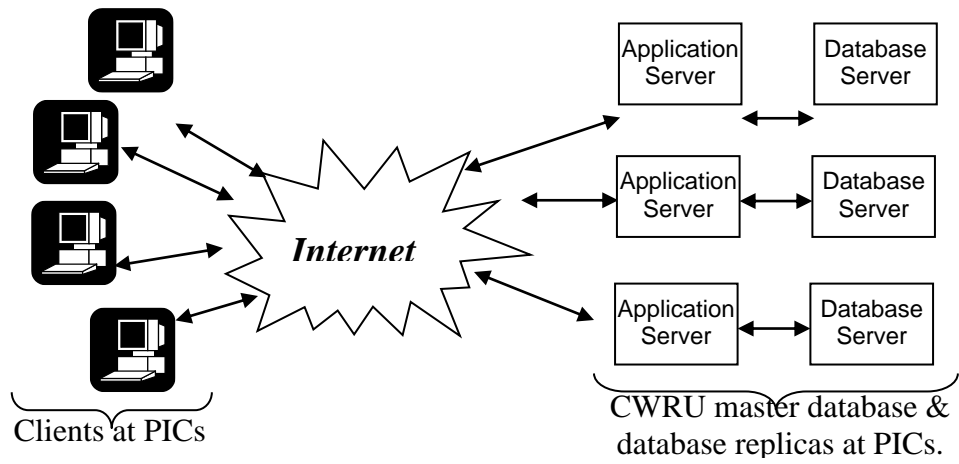
Or from Inventory System:

- A. Click the 'Process Sample' option on the menu
 - B. Click the 'Shipping/Receiving' option on the sub-menu
 - C. Click the 'Export File' icon on the menu bar
- Result is the same as 4C and 4D above

4. Importing Sample Information

- a. Scan tube barcodes by a received date and a PIC into a plain text file
 - b. Save them in directory, /find/SampleInformation
 - c. File Name Convention: SI_PICName.txt
 - d. When you export files from Shipping Note view, the system will automatically upload all files in the specified directory
 - e. After import, the system will delete all files in the directory
- Note: Please keep master log file and send it to GADCC on monthly basis

6.13 Graphic showing interaction between PIC and GADCC through the central database



GADCC & PIC Software and Network Architecture			
		Oracle 8 Database Lotus Domino Server Server	High Speed Internet Connection (> 1Mb/s)
ESRDB2 User Interface (Lotus Notes)	ESRDB2 Application Interface (CORBA)		
High Speed Local Area Network (> 10Mb/s)			

7. PROFESSIONAL, SCIENTIFIC, AND TECHNICAL SERVICE

7.1 Overview of services

These services are established activities for the purpose of this study. These activities require a degree of expertise and training specific to the aim of this study. Activities include advice and counseling regarding ethical questions and validity of participant entry data. Additionally, language translators and technical advisors will be used when necessary.

7.1.1 Administrative Support

The administrative responsibilities for the study functions will be conducted at the GADCC. A designated study coordinator at the GADCC will monitor day to day activities, via the database on the secure website, prepare and organize steering committee minutes, prepare and organize all study meetings for the NIDDK and the External Advisory Committee (EAC). Additionally, this coordinator will perform ongoing daily routine support functions for all business from each PIC that is directly related to the study processes. The coordinator will monitor, and intervene when necessary, on daily data entries coming from all PICs.

7.2 Periodic proficiency testing

Each PIC will have routine monitoring of study procedures to ensure proficiency and quality control. This will be done both with data entry and sample collection from study participants.

8 CELL LINE IMMORTALIZATION

8.1 Objective:

To create immortalized (transformed) cell cultures from primary blood cells in order to have a renewable source of DNA from the subject.

8.2 Materials:

1. B95-8 marmoset cells, which actively produce the Epstein-Barr Virus (EBV) used to transform the human B-cells.
2. MRC-5 human lung cell line (irradiated feeder cells), which provide growth factors and help nourish the B-cells during the transformation phase of growth.
3. Cyclosporin A (CSA) which suppresses the immune response of the T-cells to B-cells infected with the Epstein-Barr Virus.
4. Histopaque, a gradient used to separate the white blood cells from uncoagulated blood.
5. Cell culture medium such as RPMI 1640, supplemented with fetal bovine serum (20%) and extra L-glutamine (final conc. is $2X = 4mM$). Gentamicin sulfate (antibiotic) is used at 50 ug/ml in transformation cocktail. PBS for washes. Trypan Blue Reagent for viable cell counts.
6. Cell culture sterile plastic ware: flasks (T25, T75, T150), roller bottles, centrifuge tubes (15 ml, 50 ml and 500 ml polypropylene), cryovials (NUNC 2 ml), pipettes (1 ml, 5 ml, 10 ml, 25 ml) individually wrapped. Filter units should be purchased from several manufacturers (Corning and Nalgene), all with 0.2u membranes, in varying sizes: 100, 250, 500 ml.
7. A dedicated P-200 micropipettor is recommended for cell counts and for measuring small amounts of reagents for the transformation cocktail. Use sterile tips, with filter barrier if possible.
8. Hemacytometer counting chambers and coverslips, for performing cell counts of fresh and frozen PBMCs, and the growing cultures, if needed.
9. Basic cell culture lab equipment:
 - Biological safety cabinet/laminar flow hood
 - CO₂-incubator
 - Inverted light microscope for viewing cultures
 - Upright light microscope for cell counting
 - Pipette-aid
 - Programmable cell freezing machine (Gordinier or equivalent)
 - Liquid nitrogen storage freezer
 - Refrigerated table-top centrifuge
 - Refrigerated table-top microcentrifuge
 - Lab refrigerator / freezer.

8.3 From Blood to DNA--An Overview

1. Blood to be cultured is received in the lab, usually in ACD vacutainers. Sometimes we receive viable frozen primary blood mononuclear cells (PBMCs).
2. The patient identification information (PID) is logged into the FoxPro database which assigns the HGAL#.
3. Labels for plasma and PBMC freezes are printed and applied to 2 ml NUNC vials, four of each type per sample; PBMC vials also get a Repository number sticker. A log-in record/growing sheet is filled out, including the Repository number.
4. The blood is separated over a histopaque gradient according to the standard lab protocol. PBMCs are thawed by the standard lab protocol.
5. The B-cell transformation culture is set up using feeder-layer flasks and Epstein Bar Virus (EBV) cocktail (contains filtered supernatant of EBV-producing B95-8 marmoset cells, cyclosporin A, and gentamicin sulfate).
6. Plasma is stored at -70 C (4 x 2 ml); excess PBMCs are viably frozen and stored in LN₂ vapor phase.
7. B-cells are fed and observed in T25 flasks during the transformation period, usually 3-4 weeks, until established and growing steadily.
8. Cells are transferred to a larger flask (T75) and maintained in duplicate (split into 2 flasks) getting fed from different bottles of medium to reduce the chance of contamination.
9. After the first viable freezing of the new cell line, termed LCL-1, the cells are transferred into T150 flasks where they are grown for the second freeze (LCL-2) and subsequent expansion into roller bottles.
10. The roller bottles are harvested at confluency (1x10⁶ cells per ml) and the cell pellets are stored frozen at -70°C until DNA extraction is performed.
11. Total cellular DNA is extracted by a standard phenol-chloroform procedure in the case of large cell pellets; a kit method such as Qiagen columns may be used in cases of small amounts of material (buccal swabs, cell dregs, etc.). Quantitation of phenol extracts is determined by spectrophotometer absorption readings; concentration is adjusted to 0.5 mg/ml for the stock solutions.

8.4 Protocols and Procedures

8.4.1 Maintaining B95-8 Marmoset Cells

(Split cells every 7 days (on Friday), collect supernatant from 10-day-old cells (every Monday.)

1. On Fridays, make a 1:20 split of the flasks that are 7 days old:
Put 38 ml of media in each of two clean T75 flasks. Add 2 ml of growing cells from the 7-day-olds. Keep both flasks.
2. On Monday, take the 10-day-olds (the 7-day-olds from Friday), and harvest the supernatant as described below.
3. On Friday, make a 1:20 split of the cultures you set up the last Friday (now the 7-day-olds). Harvest the 10-day-olds on Monday, etc.

4. Make fresh EBV cocktail each Monday from the supernatant of the 10-day-old cultures.
5. NOTE: If B-cell transformation becomes slow or ceases, it may be necessary to thaw a fresh culture of B95-8 cells.

8.4.2 Preparation of EBV Cocktail

1. Transfer B95-8 cells from 10-day-old flasks to 50 ml conical tubes.
2. Centrifuge at 250 x g (1000 rpm) for 10 minutes.
3. Pour supernatant into 100 ml Nalgene filter unit (0.2 μ pore size) and discard the pellet.
4. Filter under vacuum. Pour the filtered supernatant into clean 50 ml tubes.
5. Dispense the proper amount of media and supernatant into a 250 ml corning filter unit (0.2 pore size) according to recipe below. Filter under vacuum.
6. Remove the upper part of filter unit. Add gentamicin sulfate and cyclosporine A (see note #8) to the filtered media/supernatant using a microliter pipettor and sterile tips.
7. Swirl to mix. Cap tightly and label with the date and "EBV cocktail". Store at 4 $^{\circ}$ C. Use within one week.
8. NOTE: Store cyclosporine A (CSA) at 4 $^{\circ}$ C in the dark as it is light sensitive: mix well before using. Stock concentration = 1 mg/ml; Final concentration = 0.02 μ g/ml. (See recipe)

8.4.3 Recipe for EBV Cocktail

	<u>for 50 ml:</u>	<u>for 100 ml:</u>
Media	30 ml	60 ml
Supernatant	20 ml	40 ml
CSA	100 μ l*	200 μ l*
Gentamicin	50 μ l*	100 μ l*

8.4.4 Recipe for Cyclosporin -A Suspension (CSA)

CSA is available from SIGMA Chemical Co. (CAT# C-3662 for a 10 mg vial).

1. Add 1.0 ml of 100% ethanol to 10 mg CSA; Mix until dissolved.
2. Transfer to a 15 ml conical tube and add 0.2 ml Tween-80; mix again.
3. Bring volume up to 10 ml with media RPMI 1640 (no serum).
4. Store at 4 $^{\circ}$ C covered in aluminum foil to protect from light.

Stock concentration = 1.0 mg/ml

MIX THOROUGHLY BEFORE USING.

8.4.5 Setting up Feeder Layer Flasks

One vial of MRC-5 cells makes five T25 flasks of feeder layers, standing upright.

1. Thaw vial in 37° C water bath; swab with alcohol before opening.
2. Transfer contents to a 15 ml conical tube and add about 9 ml PBS to wash out the DMSO.
3. Centrifuge at 250 x g (1000 rpm) for 10 minutes.
4. Resuspend pellet in 10 ml of filtered media (with serum).
5. Dispense 1.0 ml per T25 flask.
6. Add 3 ml of filtered media (with serum).
7. Cap loosely and incubate at 35° C standing upright.

Allow cells to adhere overnight before using.

Use within 10 days.

8.4.6 Counting Cells With a Hemacytometer and Trypan Blue Stain

This method of cell counting is based on the spatial distribution of cells within the space under the coverslip. Clumps of cells are counted as one cell. Dead cells will stain dark blue. Live cells are white, circular, and "glowing".

You will need to know these values:

The Dilution Factor- the ratio of cell suspension to trypan blue (usually 2)

The Conversion Factor- for the area under the coverslip (10^4 for most Neubauer chambers)

The volume of cell suspension from which the aliquot was taken (usually 10 ml)

1. Set up hemacytometer with a coverslip and allow cell suspension to be taken up into the chamber by capillary action.
2. Place the slide on the microscope stage and adjust the focus so you can clearly see the grid lines and the cells. If counting freshly isolated lymphocytes from blood, allow a brief time for the macrophages to "pan" or flatten out; they will appear grey under phase contrast.
3. Each chamber is divided into four quadrants. Count the cells in all four quadrants, keeping track of the number of live cells and the number of dead cells separately using a cell counter; calculate the average for each type and record them on the data sheet.
4. $(\text{average \# of cells per quadrant}) = (\text{total cells}) / (\text{\# quadrants counted})$
5. Calculate the number of live cells per ml.
 $(\text{average \# of live cells per quadrant}) \times (\text{dilution factor}) \times (\text{area conversion factor})$
6. Calculate the total number of live cells
 $(\text{cells per ml}) \times (\text{total volume of cell suspension})$
7. Calculate % viability
 $(\text{\# of live cells counted}) / (\text{live + dead cells counted}) \times 100$

Example:

180 live cells and 20 dead cells in all four quadrants-

Live cells per ml:

$$(180 / 4) \times 2 \times 10^4 = 0.9 \times 10^6$$

Total live cells

$$(0.9 \times 10^6) \times 10 \text{ ml} = 9.0 \times 10^6$$

% viability

$$180 / (180 + 20) \times 100 = 90\%$$

8.4.7 ACD Blood Separation by Histopaque Gradient Method

Have ready in advance:

- Blood prep log sheet for each sample, filled out with HGAL# and patient ID information
 - Labeled tubes for Histopaque separation and washes
 - Pre-labeled vials for plasma and cell freezes (PBMCs)
 - All reagents at room temperature, except 15% DMSO solution (on wet ice)
 - Pre-cleaned hemacytometers and coverslips, and tubes for cell counts
 - EBV cocktail and flasks with feeder layers, for B-cell cultures
1. In addition to standard BL-3 protective gown or jumpsuit, wear Tyvek sleeves and two pairs of gloves when working with blood; shoe covers are recommended as well. Wear a mask or face shield if not working in a hood.
 2. For each 10 ml vacutainer of blood, set up three sterile 15 ml polypropylene tubes with 3 ml of room temperature Histopaque in each.
 3. Using aseptic technique, carefully open the blood tube. Grasp the rubber stopper with a gauze pad or paper towel; twist and pull gently, away from the face, as contents may splatter a little. Discard stopper.
 4. Measure the blood volume with a 10 ml pipette (should be 8-10 ml). Carefully layer 3mL of the blood onto each Histopaque layer, being especially careful not to mix or disturb the interface.
 5. Place the layered tubes in the centrifuge and spin at 400 xg (1350 rpm) for 30 minutes at 22°C. Do not use the BRAKE mechanism; this will sometimes swirl the gradient at the end of the run, causing mixing at the interface.
 6. Carefully remove tubes from centrifuge, keeping them upright so that the layer of PBMCs is not disturbed. The upper layer is plasma, then the Histopaque with a fuzzy white layer of PBMCs near the top, then packed red blood cells (RBCs) in the bottom.
 7. Carefully aspirate off the plasma and dispense about 2 ml into each pre-labeled vial. Store at -20°C, or at -70°C if viral work will be performed later. Do not get too close to the PBMC layer.
 8. With a narrow-tipped pipette, aspirate off the layer of PBMCs (about 2 ml) from each tube and transfer to a new tube. Try not to get too much Histopaque. Fill tube with PBS, recap and mix gently by inversion. (RBCs may be discarded.)
 9. Centrifuge at 250 xg (1000 rpm) for 10 minutes at 22°C. (It is OK to use the brake from now on.)
 10. Pour off the supernatant; this may still be somewhat cloudy from Histopaque, platelets, etc. Combine the pellets in one tube and wash again in 10 ml of

media without serum. After resuspending the cell pellet, remove a small aliquot (50ul) for cell counting.

11. Centrifuge at 250 xg (1000 rpm) for 10 minutes. Count the cells with a hemacytometer after diluting with an equal volume of plain media.
12. Determine from the total cell number how many freeze vials can be made, allowing at least 4×10^6 cells per vial and another 4×10^6 cells for the B-cell culture flask.
13. Resuspend cells in media (with 20% Fetal Bovine Serum) to appropriate concentration. If the cell count is greater than 20×10^6 cells, resuspend the cell pellet in 2.5 ml of media with 20% FBS.
14. Add 4×10^6 cells (0.5 ml of cell suspension) to a labeled feeder flask containing 4 ml freshly prepared EBV cocktail. Place the flask in a CO₂ incubator at 35°C; place tube containing the remaining 2 ml (or less) of cell suspension onto wet ice, for viable freezes.

8.4.8 Viably freezing PBMCs from Blood:

1. To viably freeze the remaining PBMCs, gradually add 2 ml of cold media with 15% DMSO to the 2 ml of cell suspension while stirring gently. (Final concentration = 7.5% DMSO.)
2. Dispense 1.0 ml per labeled freeze vial and keep them on wet ice until you are ready to start the slow-rate freeze (maximum 30 minutes). Prepare a blank vial of 7.5% DMSO for the control probe.
3. At the end of the freeze cycle, immediately transfer vials to LN₂ storage and record the storage location on the log sheet.
4. Enter the cell count information under B-cell Record in the database; enter the PBMC freezes under Freezer Inventory. File the log sheets in the appropriate cohort section of the notebook.

Note: Following processing, all vials should be scanned into the database under "Post-processed Inventory" procedures. If possible, segregate vials going to clinical testing lab, storage (Central Repository) and back to the GADCC for distribution to the PICs, before transferring to the freezer.

8.4.9 EDTA Blood Processing Protocol

EDTA tube 1-

Mix thoroughly and aliquot as whole blood into barcode labeled vials, at 1.0 ml per tube.

EDTA tube 2-

1. Transfer total volume to a 15 ml conical polypropylene tube; record volume of blood.
2. Centrifuge at 1,000 xg (2100 rpm) for 10 minutes at 4° C.
3. Aspirate off plasma and record the volume; dispense into vials in 0.5ml aliquots.
4. Aspirate off buffy coat (WBC) into a clean 15 ml conical tube; fill with ACK Lysing Buffer and incubate at room temperature for 5-10 minutes. Spin at 250

- x g (1000 rpm) for 10 minutes to pellet; wash again with 15 ml PBS and spin as before.
5. Bring the volume in the red blood cell (RBC) tube up to 15 ml with phosphate buffered saline (PBS) and mix gently. Centrifuge at 1,000 xg (2100 rpm) for 10 minutes at 4° C. Aspirate off the PBS and residual white cells. Make 3 x 1-ml aliquots; freeze at -70° C.
 6. Make buffy coat aliquots as follows: resuspend washed pellet in 3 ml of PBS; dispense at 1 ml per vial; spin briefly in a microcentrifuge to pellet. Aspirate off the supernatant and freeze the dry pellets at -70° C.

8.4.10 Urine Processing Protocol

1. Mix thoroughly and remove one 3 ml aliquot for clinical lab testing.
2. Transfer 25 ml of urine to a 50 ml conical tube.
3. Centrifuge at 500 xg (1500 rpm) for 10 minutes at 4° C.
4. Aliquot 5 vials at 3.0 ml each, being careful not to disturb the pellet:
5. Store vials at -70° C.; dispose of excess urine.

8.4.11 Serum Protocol

Serum is aliquoted as follows:

- 1 vial at 2.0 ml for clinical lab testing
- as many vials as possible at 0.5 ml

Store at -70° C.

8.4.12 Fluoridated Plasma Protocol

Place a new barcode label on to the Fluoridated Plasma tube and store at -70°C.

8.4.13 Culturing from Frozen PBMCs

WEAR EYE PROTECTION WHEN THAWING VIALS THAT HAVE BEEN STORED IN LIQUID NITROGEN (LN₂).

(Pre-printed labels from the database are required for correct labeling of tubes and flasks.)

1. Transfer vials from LN₂ freezer to dry ice.
2. Thaw rapidly in a 37°C water bath one at a time.
3. Wipe off vial with 70% ethanol before opening.
4. Mix gently with 1 ml pipette. Transfer to a labeled 15 ml conical tube leaving a small amount of cell suspension in the vial for DNA extraction in case the culture fails.
5. Gradually add PBS to cell suspension while stirring. Bring the volume up to 10 ml. Mix well but gently.
6. Remove 50 µl of cells for viable cell count and add to 50 µl Trypan blue in a 0.5 ml tube.
7. Centrifuge the 15 ml conical at 250 xg (1000 rpm) for 10 minutes. Count cells and calculate viability.

8. Pour off the supernatant and resuspend the cells in EBV cocktail at 1 million cells per ml (1×10^6).
9. Dispense into a labeled flask with a feeder cell layer.

NOTE: If less than 1 million viable cells are recovered, they probably will not survive the transformation process. Transfer all back into the original vial for DNA extraction.

8.4.14 Maintaining Cells in Culture

Have prepared:

MRC-5 feeder flasks
EBV transformation cocktail
Filtered cell culture media with 20% FBS

1. Separate blood on Histopaque to obtain PBMCs or thaw frozen previously-isolated PBMCs. Set up the T25 culture flask with 1×10^6 cells per ml on a feeder layer with EBV cocktail. Incubate at 35°C in a CO_2 incubator.
2. On day 7 the culture should be visually inspected for contamination and media metabolism. If the culture looks healthy feed 2 ml of fresh EBV cocktail and return it to incubator.
3. On day 14 the culture should be visually inspected again for contamination and media metabolism. If the culture looks healthy feed 2 to 4 ml of filtered media with serum.
4. By the third week cells should be starting to show signs of morphological change, under the microscope, indicating successful transformation by the virus. The cells will need to be fed 2-3 times per week depending on their rate of growth. Yellowish media indicates active growth; pink is a bad sign. Continue feeding until the culture volume reaches 15-20 ml and there are large clumps of cells visible.
5. Subculture (split) the cells into two T75 flasks. Put 5-8 ml of cells and 5-8 ml of fresh filtered media with serum in each of the flasks. Leave behind about 2 ml of cells in the T25 for a back-up culture and add 5-10 ml fresh media. Remember to use separate bottles of media to feed each flask for one culture (T75 A, T75 B, T25 backup).
6. Inspect and feed the T75s 2 to 3 times per week, gradually increasing volume until there is 75 to 80 ml in each flask; feed from separate bottles of media to decrease the risk of contamination. Never add more media than half the volume of cells and media already in the flask as they will "drown".
7. When the media is very yellow make the first viable freeze of 4 vials from 50 ml of cells (25 ml from each flask) and transfer the remaining volume from both T75 flasks into a T150 flask. Add 100 ml of fresh filtered media with serum.
8. Continue feeding T150 cultures until the volume reaches 250 ml. When the media is very yellow make another freeze of 4 vials, then transfer the remaining cells to one roller bottle with fresh media to a total volume of 500 ml.
9. When cell growth is heavy and the media is yellow, about 1-2 weeks later, harvest all the cells for DNA extraction (spin down the cells, wash once with PBS and store them as a dry pellet in a 50 ml conical at -70°C until they are extracted).

Total elapsed time from start (EBV culture date) to finish (harvest of roller bottle) is usually 8-10 weeks, but that varies from culture to culture.

8.4.15 Viable Freeze Protocol

1. Begin with 50 ml of cells in a 50 ml conical tube.
2. Centrifuge the tubes at 250 xg (1000 rpm) for 10 minutes.
3. Pour off the supernatant and resuspend the cells in 2 ml of cold media (with serum).
4. Gradually add 2 ml of cold media with 15% DMSO to the 2 ml of cell suspension while stirring gently. (Final concentration = 7.5% DMSO.)
5. Dispense 1.0 ml per labeled freeze vial and keep them on wet ice until you are ready to start the slow-rate freeze (maximum 30 minutes). Prepare a blank vial of 7.5% DMSO for the control probe.
6. At the end of the freeze cycle, immediately transfer the vials to LN₂ storage. Remember to record the location on the log sheet.

8.4.16 N. Slow Rate Freezing (pre-programmed)

1. Open the valve on the tank of LN₂ and turn on the power switch on the back of the machine control unit.
2. Insert the probe into the control vial containing about 1 ml of 7.5% DMSO freeze media with out cells. The tip of the probe should be in the center of the solution not touching the tube walls. Place the probe vial in the chamber near the fan for greater accuracy.
3. Pre-cool the chamber to 4° C by turning the knob to PREP CHAMBER and pushing the black START button.
4. When the chamber is ready, turn the knob to STANDBY (before opening the lid) and transfer the vials of cells to be frozen from the ice bath to a rack in the chamber. Close and secure the lid. Turn the knob back to PREP CHAMBER. Be sure the chamber and the sample temperatures are 4°C before proceeding.
5. Turn the knob to RUN PROGRAM. The pre-programmed run cools the cells approximately one degree per minute with a built-in compensation for the release of heat in the freezing process. After this point, the PLUNGE light goes on. The total time for the run is about 45 minutes.
6. At the end of the cycle, an alarm buzzer will sound. LN₂ will continue to be pumped into the chamber to maintain the temperature until knob is turned. Do not leave samples in the chamber for very long, as the LN₂ tank will run out. The tank can not be changed while the freeze program is still running.
7. Turn the knob to STANDBY before opening the lid. Close the valve to the LN₂ tank. Transfer the sample vials to the LN₂ freezer or cryostorage tank immediately. WEAR HEAVY GLOVES TO PROTECT HANDS AND SAFETY GLASSES OR FACE SHIELD IN CASE OF EXPLODING VIALS.
8. To warm the chamber for another run turn the knob to DRY, close lid and push the START button. This will also prevent the formation of ice clogs in the lines. At the end of the thaw cycle, turn off the power switch and wipe out the condensation in the chamber. Prop the lid open to finish drying out.

8.4.17 Pelleting Cells from a Roller Bottle

1. Pour cells into a 500 ml conical centrifuge bottle. Throw away the roller bottle.
2. Centrifuge at 400 xg (1350 rpm) for 15 - 20 min at 4° C.
3. Pour off the supernatant and resuspend the cells in enough PBS to allow transfer to a 50 ml conical tube; fill tube with PBS and invert the mix. Keep the tubes chilled on wet ice during the procedure.
4. Centrifuge at 250 xg (1000 rpm) for 15 - 20 min at 4°C.
5. Pour off the supernatant and store as dry pellets at -70° C.

8.4.18 DNA Extraction

Stock Solutions:		Lysing solution:	Proteinase K buffer:
NaCl	5 M	0.15M NaCl	10 mm TRIS (pH 8)
TRIS (pH 8)	1 M	0.1M EDTA (pH 8)	5 mm EDTA (pH 8)
EDTA (pH 8)	5 M	0.02M TRIS (pH 8)	0.5% SDS
SDS	20 %	1% SDS	
Rnase	10 mg / ml		
Proteinase K	10 mg / ml		

Procedure:

1. Begin with a -70°C frozen pellet of cells in a 50 ml conical tube. Thaw at room temperature.
2. For a cell pack of 8×10^8 to 1.5×10^9 add 0.5 ml Rnase and 50 ml lysing solution:
3. Place tubes in a 60° C water bath for 10 minutes.
4. Incubate at 45° C for 30 minutes
5. Add proteinase K (100 ug/1 ml of lysing solution)
6. Incubate at 37° C for 2 hours.
7. Add an equal volume of phenol/chloroform/isoamly to lysing solution and mix well by shaking.
8. Centrifuge 10-20 minutes at ___xg (2,000 rpm, IEC centrifuge).
9. Slowly draw off the top layer (DNA) trying not to disturb the interface. Place in a new tube.
10. Add an equal volume of chloroform to DNA and mix by shaking.
11. Centrifuge 10-20 minutes at ___ xg (2,000 rpm, IEC centrifuge).
12. Remove top layer (DNA) do not disturb interface. Place in a new tube.
13. Ethanol precipitate the DNA:
In a plastic disposable container add 2 volumes of 190 or 200 proof alcohol. Pour DNA into alcohol.
- If the amount of DNA is what is expected, the DNA can be spooled out of the EtOH. If the amount is insufficient for spooling, place the DNA / EtOH mixture on dry ice for 15 minutes and centrifuge at 10,000 rpm for 10 minutes.
14. Resuspend the DNA in TE buffer. For 100 ml of buffer, add 1 ml of 1 M TRIS (pH 8) and 0.2 ml of 0.5 M EDTA. If the initial cell pellet was 8×10^8 to 1.5×10^9 use 10 ml of TE buffer, 2 or 3 additional ml can be added if dissolved DNA is too viscous. This step may take 2 to 5 days to dissolve on rotator

9 DATA ANALYSIS

9.1b Objectives of Data Analysis Subcommittee

The function of the Data Analysis Subcommittee is to provide counsel for all data analyses performed by the GADCC. These will comprise all data analyses that use data from more than one PIC, and all data analyses requested of the GADCC by the Subproject/Ancillary Projects Subcommittee.

9.2 Definitions

ARP: DM and overt proteinuria as defined in both relatives in the Protocol

DRP: Duration of diabetes for ≥ 10 years AND normal albumin excretion (< 30 mg albumin/24 hours, or a urine albumin (mg)/gram creatinine < 0.03) AND no historical evidence of kidney disease in one relative and DM and overt proteinuria in the other relative.

Continuous Traits for QTL analysis:

- a. Creatinine concentration
- b. Protein excretion
- c. Estimated HbA_{1c}

9.3 Two guiding Principles

These two principles will be followed in reporting all results

1. Accurate pointwise significance levels will be obtained.
2. Evidence against linkage will be quantified, e.g. by p-values > 0.5 , so that our results will be poolable by others in an unbiased fashion.

9.4 Data Analysis Techniques

Techniques for data analysis are continually being improved. The best and most powerful available when the analyses are performed will be used. Examples of such procedures currently available can be found in APPENDIX C.

10 CLINICAL QUALITY CONTROL

10.1 Introduction

The goal of the Clinical Quality Control Committee is to develop standard procedures for the collection of clinical data. For a multi-center study, the most important step in quality assurance for the collection of clinical data consists of the training and certification process. In FIND, this training and certification will be done at each PIC.

10.2 Medical Chart Review

PIC technicians are trained and certified for medical chart abstraction by PIC coordinators. Five medical charts from each PIC abstractor are xeroxed and sent to the GADCC within the first month of the abstractor beginning data collection. These charts are then reviewed by data abstractors from other PICs, and Kappa statistics calculated for entry criteria by the GADCC to determine the reliability of the data. This process is repeated annually for each data abstractor.

10.3 Blood and Urine Collection Processing

10.3.1 Staff Certification Requirements

The blood drawing and preparation are performed by certified technicians. Each institution or organization is responsible for ensuring that technicians are trained according to institutional guidelines.

10.3.2 Participant Compliance with Protocol

A fundamental function of data collection in FIND is to provide accurate phenotype assessment for diabetes and renal disease. Hemoglobin A1C and history will be used to assess diabetes phenotype. Fasting blood glucose needs to be determined in all participants without ESRD who have an indeterminate Hemoglobin A1C level, per the FIND protocol. In addition, it should be obtained whenever possible to confirm the absence of diabetes. To obtain valid and comparable measurements of fasting blood glucose, FIND participants should fast overnight for at least 8 hours before blood is drawn. Interviewers who administer the medical questionnaire are trained to explain the importance of compliance when recruiting FIND participants and to obtain their agreement to comply. When FIND participants are contacted before their appointment about the scheduled visit, they are to be reminded of the fasting restrictions. Time of the last meal will be recorded. The GADCC analyzes study data for information on length of time fasting and reports the percent of participants at each PIC who do not comply.

10.3.3 Maintaining Proficiency

To maintain their proficiency, technicians are urged to perform blood drawing and processing at least once each week (or 8 times each 2 months). The GADCC analyzes the study data to assure that all technicians collecting and processing blood in the PIC are performing these procedures frequently enough to maintain their proficiency. For PICs using vendors, maintenance of proficiency is dependent on the vendor's protocols. PICs using hospital-based technicians will need to comply with JCHO (Joint Commission Hospital Organization) requirements.

10.3.4 Packing Samples for Shipment from PIC to NCI

All vials of blood samples, as well as the plastic bags that contain all samples for a given participant, are labeled with the participant's barcode ID. A shipping list for all sets of samples is enclosed with each shipment to NCI. The person unpacking these samples at NCI verifies that the IDs on the vials match the ID on the plastic bag and the ID on the shipping list. If any discrepancies are detected, NCI contacts the PIC to resolve the problem.

Blood vials shipped to NCI must be packed securely to avoid both breakage and warming. NCI monitors the arrival condition of the samples sent from each PIC. If problems are encountered, NCI notifies the PIC involved.

Shipment of all samples is coordinated through the GADCC website. To avoid delays in transit to NCI which might cause samples to be warmed or thawed in shipping, it is especially important that the NCI is made aware of samples shipped on Fridays, Saturday, or days prior to a federal holiday.

The GADCC will monitor the Shipping/Receiving Form in the Central Database and notify PICs when a shipment is in violation of FIND approved procedures (See Section 5). The Study Coordinator and the Principal Investigator of the site in question will receive an e-mail from the GADCC citing the following:

- 1) Shipment Date
- 2) Federal Express Tracking Number
- 3) Barcode ID
- 4) Sample Violation or Shipment Violation
- 5) Immediate Corrective Action to be Taken

These violations will be recorded in the monthly reports presented to all members of the Steering Committee, who will also be informed if corrective action is not taken.

10.3.5 Replicate (phantom) Samples

To reduce the burden upon FIND participants, no extra blood will be drawn from any study participants. Instead, an extra set of quality control replicates of whole blood, serum, plasma, and urine from one study participant will be sent out. Replicate samples from approximately 10% of the entire FIND study will be selected for quality control. NCI will generate a list of PIDs that have specimens in the archive and the amount of each specimen. The GADCC will eliminate PIDs that do not have enough specimens in the archive to participate. The remaining specimens will be used to randomly select a 10% sampling for shipment to the Central Phenotyping Laboratory. The NCI laboratory will use the barcode that contains the PID number and the tube number as a quality control ID. The NCI will pull the specimens from their archive, create a shipping form using the quality control ID as described above, and send the samples to the Central Phenotyping Laboratory for processing. The Central Phenotyping Laboratory will treat each sample they receive independent of any previous samples. A new accession number will be assigned and the usual battery of tests performed for that specimen. Results of the tests will be reported to the PICs weekly. The GADCC will receive an electronic report on a monthly

basis from the Central Phenotyping Laboratory. For data analysis by GADCC, results on each laboratory measurement are matched to the appropriate participant results using the PID from the barcode ID. GADCC will generate quality control reports quarterly for review by the Steering Committee. The Central Phenotyping Laboratory will be sent the comparison results on a quarterly basis. These samples will be sent on a regular basis during the entire study duration.

10.4 Sample Receipt at the Central Processing Laboratory (NCI)

10.4.1 Specimen Identification

Upon arrival at NCI, each specimen is verified by NCI technician via the GADCC database. If a discrepancy occurs, the NCI technician calls that specific PIC to resolve the problem.

Procedures for creating and identifying a record for each specimen upon arrival follows current NCI protocol. This record includes a local specimen identification number, as well as the FIND ID number. The local specimen ID number is the linking variable used to update the record for each specimen (either by direct data transfer from the analytic instrument or by entry of results from worksheets) after the specimen is analyzed. It is therefore crucial that care be taken when labeling specimens with local ID's that the local ID recorded for each specimen on the data base matches that by which the specimen is actually identified as it is processed in the lab. Laboratory supervisors periodically review how this process is carried out and instruct laboratory technicians in techniques to use to avoid ID errors.

10.4.2 Maintenance Procedures for Laboratory Equipment

A regular schedule is set up for routine maintenance procedures, with logbooks kept on their performance. The laboratory supervisors review these logs on a regular basis to verify that proper maintenance procedures are being carried out according to the schedule set and that any special maintenance procedures need are carried out.

10.5 Central Phenotype Laboratory Quality Control (MedStar)

10.5.1 Specimen Identification

In FIND, blood samples are collected by each PIC and processed by NCI for shipment to the Central Phenotype Laboratory (MedStar). Procedures for creating and identifying a record for each specimen upon arrival follows current MedStar protocol. This record includes a local specimen identification number, as well as the FIND ID number. The local specimen ID number is the linking variable used to update the record for each specimen (either by direct data transfer from the analytic instrument or by entry of results from worksheets) after the specimen is analyzed. It is therefore crucial that care be taken when labeling specimens with local ID's that the local ID recorded for each specimen on the data base matches that by which the specimen is actually identified as it is processed in the lab. Laboratory supervisors periodically review how this process is carried out and instruct laboratory technicians in techniques to use to avoid ID errors.

Specific guidelines for quality control of each assay performed by MedStar staff can be found in the appendix section of this manual [SEE APPENDIX II].

MedStar routinely archives all samples at -70°C for 30 days after the initial assays are performed. The archived samples are used to repeat assays when necessary. GADCC performs bi-weekly evaluation of laboratory data submitted from MedStar outlying values and systematic patterns. If outlying values or systematic patterns are detected, GADCC generates a list of samples to be repeated and contacts MedStar for repeated assays to be performed.

10.6 Material Archiving

It is important in handling the FIND frozen blood samples (in the case of bulk shipments from the PICs) to avoid any unnecessary exposure to room temperature. Clear procedures for unpacking specimens upon arrival is set out at NCI to minimize such exposure. NCI archiving has provisions for (1) prompt detection of power failure or of failure of freezer to maintain the proper temperature, including both local alarms and alarm signals to a central security office that will notify appropriate laboratory personnel if a problem develops after hours; (2) backup power supplies in the event of power failure; (3) plans for the use of dry ice to maintain the sample temperature until any problems with the freezer can be repaired.

10.7 Long Term Consistency of Methods

The laboratory protocols fully describe the measurement methods and procedures to be used in the FIND. To maintain the long-term comparability of measurement results throughout the FIND, the same measurement methods will be maintained throughout the study. If measurement methods are to be changed, comparison studies must demonstrate to the satisfaction of the FIND Steering Committee that the method to be used has been shown to give results fully comparable to the method used initially. It is important to demonstrate that comparable results can successfully be achieved in the implementation of the method in the FIND laboratory, rather than to rely merely on descriptions of comparisons in the literature, as there are often significant differences in the results achieved by the same analytic method at different laboratories. Use of a back-up method in the event of the temporary failure of the usual FIND method to achieve in-control results should only be done with the permission of the FIND Steering Committee after the back-up has already been implemented in the laboratory and shown to maintain comparable results. In both cases, the experiments to demonstrate method comparability will be designed jointly by the laboratory in question and the GADCC.

11 MOLECULAR QUALITY CONTROL

11.1 Objectives and Goals of the Molecular Quality Control Subcommittee

The goals of the Molecular Quality Control and Assurance Subcommittee are:

1. To develop protocols and standard operating procedures for collection of molecular genetic data from samples ascertained in the FIND study with attention to quality control and assurance.
2. To plan a genome scan for FIND data with selection of appropriate controls and replicate samples.
3. To evaluate the molecular results from the genome scan and the MALD studies.
4. To identify chromosomal regions for follow-up and confirmation of linkage results and candidate genes, as per the recommendation of the data analysis committee.
5. To plan positional cloning projects for those regions where initial and confirmatory evidence for linkage has been obtained

To meet the goals of the molecular quality control and assurance committee a series of standard operating procedures have been developed. These are listed below.

11.2 To develop protocols and standard operating procedures for collection of molecular genetic data from samples ascertained in the FIND study with attention to quality control and assurance.

11.2.1 DNA extraction from lymphoblastoid cell lines

As described in section 8, cell pellets from lymphoblastoid cell lines will be received at the GADCC laboratory from NCI on a monthly basis, after the first set of samples have been transformed into cell lines at the NCI. The receipt of the samples will be scanned into the FIND database using the FIND barcodes. DNA will be extracted from cell pellets as per protocol on page 65. Stock DNA will initially be diluted to 1 microgram/microliter and vials labeled with barcode labels that are resistant to freezing, moisture and solvent removal. All purified DNA samples will be diluted to a 200 nanogram/microliter concentration in 10 mM Tris, 5 mM EDTA buffer and stored frozen at -80°C in cryoboxes that have been barcoded.

11.2.2 DNA extraction from whole blood or buffy coats

Transformed lymphoblastoid cell lines may be unavailable for DNA extraction on some patients, although, PICs will be informed of lack of success in cell line transformation and will be asked to re-draw crucial samples, if feasible. NCI will inform the GADCC via the

reporting fields in the FIND database that the samples failed to transform, or that yellow top tubes were unavailable for cell line transformation. In these cases the archived buffy coat or whole blood sample will be used for DNA extraction. DNA extraction will be performed using the PureGene Kit [Gentra Systems]. Stock solutions will be diluted to 1 microgram/microliter and vials labeled with barcode labels that are resistant to freezing, moisture and solvent removal. All purified DNA samples will be diluted to a 200 nanogram/microliter concentration in 10 mM Tris, 5 mM EDTA buffer and stored frozen at – 80°C.

11.2.3 Assignment of DNA sample to 96-well plates for genotyping

The assignment of a particular DNA sample (10ng/ul) to a particular address on a specific 96-well Biostar vial will be dependent on whether the individual participated the MALD study or the individual participated in the Family Study. Although, samples will generally be assigned to a plate as collected, the actual assignment will depend on whether a family qualified for the study and other available family members. The assignment of a sample will be determined by the FIND database. For the Family Study, family structures will be retained inasmuch as allowed by concurrent data collection of family members. Each DNA plate will also contain the appropriate number of replicate (N=4) and control CEPH samples (N=4). Using all available samples replicates will be randomly sampled for each plate while ensuring that all 5 ethnicities (Caucasian, African American, Mexican American, Pima Indian and Zuni Indian) are adequately represented. Once a sample is selected as a replicate it will be removed from further consideration.

11.2.4 Quality control for DNA concentration

The purity and concentration of the DNA samples will be ascertained using the spectrophotometer [Spectromax Plus, Molecular Devices]. After samples have been diluted to a concentration of 10 ng/ul, a test PCR reaction will be used to determine if the samples can be amplified using standard primers. If a particular sample has a DNA concentration of less than 200 micrograms, the GADCC will ask for additional blood to be drawn from the individual for genotyping. Samples with DNA concentrations less than 200 micrograms will not be considered to be recruited, since a sufficient quantity of DNA is unavailable for genotyping.

11.2.5 Shipping of DNA aliquots to the PICs

Each PIC may request that ½ of the archival sample, viable cell-lines or extracted DNA be returned to them for candidate gene studies. DNA samples will be shipped to PICs (See Appendix I for addresses) on a monthly basis. Shipment of samples will be pre-arranged between the laboratory personnel at the GADCC and the PIC. Samples will be shipped frozen from Monday through Wednesday and the appropriate personnel alerted to the arrival of the samples at the PICs. In the event that a large quantity of DNA is unavailable for a particular sample, a sufficient amount of DNA from each sample will be retained at the GADCC for the planned genome scan.

11.3 To plan a genome scan for FIND data with selection of appropriate controls and replicate samples.

11.3.1 Plan I for the genome scan

1. In years 2-3.5, an application will be submitted to the Center for Inherited Disease Research (CIDR) for a 10cM genome scan for the entire FIND study. This application will utilize a two-stage approach for the entire sample. First, the DNA sample collected in year 2-3.5 (1/4 of the whole sample) will be sent to CIDR for an initial genome scan. Next, the sample collected in years 3.5-5 (3/4 of the whole sample) will be used to replicate the results of the initial scan. Thus, samples collected in years 3.5-5 will be shipped on an annual basis to CIDR for genome scanning.
2. CIDR will provide both raw and analyzed data for the study. CIDR currently employs 4 CEPH individuals, 4 blind replicates and a moving blank on their gels for quality control but is amenable to modifying plans to suit the consortiums' needs.
3. The rationale for the performing an initial genome scan on 1/4 of the sample is that fine mapping for positive results can be pursued very quickly. Four to eight additional markers will be typed in the whole sample in regions where significant lod scores are obtained from the initial scan (e.g. p-value <0.05 was observed). The size of the interval will determine how to proceed with fine mapping. For an interval larger than 3 cMs, additional markers will be typed to narrow the interval. For a small interval, candidate genes may be tested for mutations and polymorphisms.
4. The results of the genome scan will be utilized in relationship testing and reassigning relationships after assessment of genotyping error (see data analysis section).

11.3.2 Plan II for the genome scan

1. If the genome scan application is rejected by the CIDR access committee because of power considerations or inadequate recruitment, a re-submission will be planned within 6 months of the initial rejection. Currently, the rate of genotyping error at CIDR is less than 1% at CIDR, with a 5% missing data. As a back-up, the molecular laboratory at the GADCC will perform the genome scan. In this case, only 1/4 of the sample will undergo the full genome scan. The whole sample will be typed for additional markers where positive lod scores are observed. For the first stage scan, 405 equidistant markers spanning all human autosomes and the pseudoautosomal region of the X chromosome will be used to type each sample. Markers spaced at an average distance of 10 cM designed from Co-operative Human Linkage Center panels 10, which are composed almost entirely of tri- and tetranucleotide repeats with an average heterozygosity of 0.8, will be purchased from Research Genetics (Huntsville, AL). The current rate of genotyping error at the GADCC laboratory is 5%, with 10% missing data using the ABI 3700.
2. Controls and replicates selected for quality control will remain identical for both plan I and II.

3. Markers selected for fine mapping will depend on results obtained from the first stage. We anticipate that five to six marker regions will “suggest linkage” by chance alone and warrant further investigation in order to maintain our desired level of statistical power. For each positive region, six to eight markers spaced two cM apart on either side of the positive marker will be typed. Further, if putative candidate genes lie within 20 cM of the region with suggestive linkage, then markers close to or within the gene or the candidate gene itself will be typed.
4. Genotype data will be obtained on an ABI 3700 Sequencer. Allele sizes will be assigned using the Genotyper version 2.1 software. Data from the Genotyper program will be dispatched to the GADCC statistical core for entry into the FIND database and for analysis of Mendelian segregation. Data will be analyzed within one week of being run. The average size of raw and processed data files obtained from this program is about 25-30 megabytes. The data will be temporarily stored on the MetroHealth Medical Center server until all genotypic information is extracted. The data from each run will be compressed and moved for remote storage to a zip disk or CD. A version of the raw data files stripped of the large image file and compressed to less than half the size will also be stored on the Unix Servers at the GADCC. The data in these files will be burned onto CDs, media suitable for long-term storage of large amounts of data. Each CD will be carefully labeled with dates of the runs, markers, and study name, and indexed in the FIND database.

11.4 To evaluate the molecular results from the genome scan and the MALD studies.

1. Because data from the genome scan will be accumulated over several years, the Molecular Quality control committee along with the Data Analysis Committee will analyze the data for measurement drift using replicates of samples genotyped in previous years. Population-specific allele frequencies will be estimated from the data and compared using extant data from public databases such as GDB (www.gdb.org). Raw data and reports for each PIC will be made available via web access to the FIND database. The progress of the genome scan and the positive evidence for linkage will be evaluated in terms of potential candidate genes.
2. Raw data and results from the MALD studies will also be deposited in the FIND database. Similar to methods utilized in the Family Based Study, the Molecular Quality control committee along with the data analysis committee will record and analyze the MALD genotyping data for measurement drift using the quality control plan suggested by the MALD centers. The progress of the MALD genotyping and analysis will be periodically evaluated in terms of potential candidate genes.

11.5 To identify chromosomal regions for follow-up and confirmation of linkage results and candidate genes, as per the recommendation of the data analysis committee.

1. After the initial genome scan in ¼ of the samples is completed (Plan I) and regions that demonstrate positive evidence for linkage identified, additional markers in these select regions will be identified. These markers will be genotyped in the original sample that demonstrated evidence for linkage and in a confirmatory sample, if

available. Priority of regions for genotyping will be assigned based on significance of evidence for linkage. Markers at candidate genes selected by the Molecular Access committee and the approved by the Steering Committee will also be genotyped in selected samples.

11.6 To plan positional cloning projects for those regions where initial and confirmatory evidence for linkage has been obtained.

1. If second stage markers confirm evidence for linkage in a particular chromosomal region, the evidence in this region will be evaluated for an eventual positional cloning project. Additional studies that involve reducing the region to a few centimorgans (<2cMs) will first be undertaken. Evaluation of mutations or single nucleotide polymorphisms in specific candidate genes are also under the purview of this aim.

11.7 Genotyping Protocol for MALD analysis of MA DN Probands and Controls.

DNA samples will be aliquoted in a 96 well format. Working stock 96 well template containing 1.0 ng DNA/ μ l TE (10 mM Tris/ 1mM EDTA) volume will be generated using a BIOMEK 2000 robot (Beckman). The working template trays will be used for robotic aliquoting of 1 μ L into two quadrants of a 384 well microtiter trays that will be used for PCR reactions. A Multimek (Sagean/Beckman) robotic unit is used for this procedure and typically at least 4 trays are prepared at the same time at this station. This procedure is repeated with a second 96 well "working stock" template so that each quadrant of the 384 well plate contains 96 wells corresponding to one or the other working stock tray. To minimize any possible run-to-run bias each 384 well plate will contain two quadrants from a proband 96 well working stock and two quadrants from a control 96 well working stock.

All PCR reactions will be performed in 384 well PCR trays. Typically 3 markers will be multiplexed for each PCR reaction and PCR reaction mixes are robotically dispensed into two quadrants of the 384 well PCR tray in which template DNA has been predispensed. This typically repeated so that a total of 8 different PCR reaction mixes that are dispensed at the same time into 4 different 384 well trays. Each of the primer sets are designed to have compatible PCR products (by size or different flours). When additional multiplexing of compatible products is possible a total of upto 12 different PCR reaction mixes can be performed at the same time without exceeded the capacity of the robotic Biomek robotic station or thermocyclers. The final 5 μ l PCR contains 2.65 μ L H₂O, 0.5 μ l PCR buffer, 0.7 μ l of 2.5 mM dNTPs, 0.1 μ L of 10 μ M mix of forward and reverse primers and 0.05 μ l Taq polymerase.

The primers pairs amplify genomic DNA segments of ~100 to 400 bps that contain specific MALD microsatellite or insertion/deletion (I/D) polymorphisms. The forward primer is conjugated with one of three ABI 3700 sequencer compatible fluorescent dyes (FAM, HEX or NED). These are stored in light protected boxes at -20°C prior to PCR setup. For microsatellite polymorphisms most reverse primers will be synthesized to contain tails to minimize the plus A problem that is especially an issue for dinucleotide repeat markers.

To accomplish the PCR setup, a master-mix sufficient for 384 PCRs containing the entire reaction mix except for the template is aliquoted and placed in a Biomek test-tube station. Using one of our standard programs the Biomek 2000 accurately dispenses the required 4 μ L volume into the specified two 96 well quadrants of the appropriate 384 well PCR tray. The PCR tray is then sealed (Micro Amp Clear Adhesive Film), vortexed briefly and then centrifuged.

The completed and mixed 384 well PCR trays are then placed in the thermocyclers. For the typical protocol (as described above), 4 tray PCRs are performed using two of our three dual 384 well thermocyclers (Gene Amp PCR System 9700, PE Biosystems). For most reactions a standard PCR program similar to the Marshfield protocol is used for both microsatellite and I/D polymorphisms (95^oC X 3 min, 95^oC X 45 sec, 58^oC X 1.5 min, 72^oC X 45 sec for 32 cycles). Products are either stored at -20^oC (for upto 2 weeks) at this stage or are prepared for 3700 loading.

After the PCR reaction has been completed and within 12 hours of initiating an ABI 3700 sequencer run, a Multimek robot is used to mix the compatible PCR products and set up loading trays. As indicated above the reaction templates are set up so that a given quadrant is typically matched in four trays by identical templates and PCR products that are size/fluor compatible for multiplexing. The four trays can therefore be combined by simply retaining the same quadrant position in a 384 loading tray. For this tray, 0.5 μ l of each PCR product from each of the four PCR trays is mixed using one of our laboratory's standard Multimek programs. The loading tray is then prepared by adding 4 μ l of loading mixture. The loading mixture contains 90 ml of 3700 BAL ROX standard (Bioventures)/4000 μ L of low conductivity formamide (Amersham). This size standard contains fragments in 20 bp intervals throughout the size range 80 to 400 bp plus 70, 90 and 190 bp fragments.

Two 384 well completed loading trays are then placed in the ABI 3700 (PE Biosystems) for gel capillary size separation. This automated sequencer is run using the standard ABI operating protocol with the POP6 run polymer. After the sequencer run has been completed the run files are transferred from the 3700 dedicated computer to another PC (Dell, Pentium III, 750 mhz) . The Genescan Analysis software is used to size the DNA fragments based on the sizes of the size standard peaks. This "raw data" consisting of DNA fragment sizes for each typed individual is then imported into the Genotyper software program (PE Biosystems) in which templates have been created for each marker, specifying marker name, range of allele sizes and fluorescent dye label and peak height thresholds. This information is used by the software to generate genotypes (to 2 decimal places) for each individual typed and this is displayed in tabular and graphical forms. Each genotype is then manually reviewed to minimize errors in the labeling of alleles by the software. Visual confirmation of each individual peak height is performed by careful review of the electrophoretic profiles. A final table of genotype results is then saved as an Excel file.

Allele binning will then be performed using the analysis tool package in EXCEL (97 or later versions) prior to exporting data for specific statistical analyses.

11.7.1 Quality Control:

To enable checking the quality and reproducibility of the genotyping each of the 96 well templates will contain 4 replicate samples and 4 CEPH controls (as provided by the FIND GADCC). The known positions of the CEPH controls will be used to establish consistency in allele assignments between sets run and genotyped on different days. The replicate samples (with sample positions blinded to technical staff) will be used to determine the concordance rate for individual marker typing. If the marker concordance is less than 95% the marker will be re-genotyped carefully examining the electrophoretograms for artifactual peaks and reviewing the binning criteria. If the marker concordance is still less than 95% the marker will either be discarded or will be retyped starting from the initial PCR step.

11.8 Genotyping Protocol for MALD analysis of African Americans.

DNA samples will be aliquoted into 4.5 ml cryogen type tubes. Working stock templates containing 1.0 ng DNA/ul will be generated using a Packard Multiprobe II robot. The 300 ul of the working template stocks will then be robotically aliquoted into deep-dish 384-well plates. Those daughter plates will be transferred into 384-well PCR granddaughter plates hydra containing 5ng/well with a Robbins 384 well hydra. Plates are then dried overnight in a laminar flow hood, sealed with tape and stored for up to six months at -20 C for subsequent PCR analysis.

The NCI will perform these analyses using state of the art technologies that are constantly being evaluated and implemented to improve the genotyping process. If 1532-well PCR systems are commercially available and reliable at the time the MALD analysis begins those will be used in the PCR analysis using a Robins 384 Tango hydra system for aliquoting and pooling. Subsequent discussions of these plans will assume 384-well density even though the higher 1532-well density will hopefully be available in the next one to two years.

All PCR reactions will be performed in 384 well PCR trays. PCR reaction mixes are robotically dispensed into the 384 well PCR trays in which template DNA has been pre-dispensed and rehydrated. This is typically repeated so that many different PCR reaction mixes that are dispensed at the same time into different 384 well trays. Using the newly released five-dye system from Applied Biosystems, each of the primer sets will be designed to have compatible PCR products (by sizes up to 500 bp and four different flours). Up to twenty loci can be independently amplified, pooled and then analyzed.

Fragments will be size separated on Applied Biosystems 3100 sequencers, or their successors depending on availability. Electropherograms that are produced by the sequencers will be analyzed in Gene Mapper (Applied Biosystems) to bin alleles for subsequent analysis. These data will be deposited into a database at the NCI and analyzed for association between the FIND phenotypes and MALD genome scan markers.

11.8.1 Quality Control:

To enable checking the quality and reproducibility of the genotyping each of the 384 well templates will contain 8 replicate samples used in Dr. Michael Smith's laboratory in all 384 well plate genotyping applications. The internal duplicates will be used to establish the consistency of genotyping in the laboratory. Internal blinded replicate samples will be used to determine error rates and determine which markers must be retyped or redesigned.

12 INTELLECTUAL PROPERTY AND ARCHIVE POLICIES

12.1 Function and purpose

The Intellectual Property and Archive Subcommittee will develop and propose policies relating to ownership and storage for data and biological specimens generated by the F.I.N.D.

12.2 Duties of the Intellectual Property and Archive Subcommittee

Develop policies relating to ownership and storage for data and biological specimens generated by the F.I.N.D., and to the distribution of these to non-F.I.N.D. investigators (in the latter regard, the duties may overlap with the Subproject Evaluation Subcommittee). These policies will require approval by the Steering Committee.

Evaluate and make recommendations for use of data or biologicals by specific outside (i.e. non-F.I.N.D.) investigators. (This duty may overlap with the Subproject/Ancillary Studies Subcommittee.)

Serve as the initial forum for decision and appeal of intellectual property issues. Decisions in this regard are to be recommended to the Steering Committee for final decision.

All policies and recommendations need to be in accordance with the arrangements made with the specific population involved (e.g. Native American groups), and in accordance with NIH and federal guidelines, and when possible in accordance with individual institutional guidelines. Policies and recommendations [internal and external] will be dictated by the nature of the informed consents signed by the subjects.

12.3 Categories of Materials

There are several categories of data, as follows:

Category 1:

Data are project wide, e.g., phenotypes being recorded on all participants, genotyping data for genome scans. In general these are data specified in the MOO (Manual of Operations).

Category 2:

Data that are part of the designed study, but generated by a limited number of centers (e.g. 2 PIC's, or 1 PIC and the GADCC).

Category 3:

Data that are the results of an approved subproject (ancillary) study that is funded by the NIDDK as a supplement to the F.I.N.D. or formally in collaboration with the F.I.N.D. (e.g. the NEI retinopathy study).

Category 4:

Data that are the result of an approved subproject (ancillary study) not funded by the NIDDK.

Category 5:

Data that are generated independently from F.I.N.D., but utilizes in part F.I.N.D. subjects, whether conducted by F.I.N.D. investigators or non-F.I.N.D. investigators.

12.4 Biological Specimens

There are two basic categories:

Category 1: Material in finite supply.

- a. DNA
- b. Serum/plasma
- c. Urine

Category 2: Material of unlimited supply.

This can include but is not limited to:

- a. Immortalized cell lines
- b. DNA from immortalized cell lines

12.5 Ownership

The basic concept here is that ownership of material is vested in the participating investigators (and their institutions), limited by agreements with, and the rights of, participating populations, NIH and federal polices, and policies of the individual funding agencies. Confidentiality of individual participants will be maintained with all data shared between centers and with all releases of data. Furthermore, it is deemed advisable that, after some period of time, certain categories of data and possibly biologic materials be made available to non-FIND investigative teams.

Category 1

These data will be owned jointly by the individual PIC's and the GADCC. They will be kept at the GADCC, and at each PIC that desires them. While each group is free to examine and evaluate the data for informal/internal purposes, the formal use of these data will be dictated by the Steering Committee via the Analysis, Protocol, and the Publication Subcommittees. The GADCC and PIC's will not disclose the data to other parties except as specified by the above committees and the Manual of Operations.

After each F.I.N.D. publication that involves major data, the GADCC will make the data used in that publication available to the public (the exact method to be proposed by the GADCC) as soon as is reasonably possible, but no later than within one year. These data will not include the American Indian ethnic group. These data will be anonymized, i.e. identifiers (e.g. birthdates) will be removed, and ethnic group identified by only one of 3 categories (Caucasian, African-American, Hispanic). This anonymization must be done before any data are released to the public.

When the GADCC (or equivalent successor) ceases to function as the data coordinator and analytic resource to the F.I.N.D., it will release a fully documented copy of all F.I.N.D. data to each PIC (limiting identifiers to the PIC from which the data were derived), and an anonymized data set to the NIDDK.

One to Two years after the termination of the F.I.N.D. and/or this release (whichever comes first), the anonymized data set will be made public (this will probably require an archiving mechanism, to be discussed). Decisions regarding any other disclosure of data to other parties will be brought to the Intellectual Property Subcommittee, which will make recommendations for decision by the Steering Committee.

Note: The responsibilities of any users (e.g. the GADCC) to return these data to the central database will continue, the data being sent to the archives/repository mechanism established after the GADCC ceases to function.

Category 2

For the most part these are to be treated in the same way as category 1 data, but ownership of the data is to be vested in the individual participating centers (e.g. the GADCC and a PIC or 2 PICs). The location of the data should be in the GADCC and participating centers.

Category 3

These are to be treated as category 1 or category 2 data, depending on the number of Centers involved.

Category 4

These data are to be owned by the responsible investigators. Data analysis is the responsibility of those investigators, with whatever involvement of the GADCC the Subproject and Steering Committee will have designated. Since this is an approved project of the F.I.N.D., even if the GADCC is not involved in analysis, the data (e.g. the primary molecular genotyping data) should be immediately available to the GADCC and the Steering Committee upon publication. However any public release of these data will be at the discretion of the responsible investigators, but will be included in the full public data release.

Category 5

The responsible investigators own these data, and their analysis. After publication, the data are to be made available to the GADCC central database (and/or successor) within 3 months. Public distribution of these data is at the discretion of those investigators but will be included in the full public data release.

12.6 Biological materials

The basic principal here is that the first priority is the desire of populations, second the needs of the study, third that of the investigators, and fourth posterity.

12.6.1 Finite Supply

PIC's will have the option of receiving a sample of DNA (quantity to be determined) from all FIND participants recruited at that site during the initial 5 year enrollment period. Within one to two years of the conclusion of the study, any remaining biological

materials in this category (e.g. DNA, serum) residing at the GADDCC or other central laboratory (e.g. assay laboratory) will be treated as follows:

- a. If a given population has requested return of samples, that request will be honored.
- b. Of the remaining material at the GADCC (or other central core), half will be returned to the individual PIC, and half will be archived for the NIDDK (mechanism to be determined).
- c. The ownership of these samples will reside with the possessors (e.g. the NIDDK and the individual PICs, respectively).

Note, excess materials that have been accumulated at a given PIC will be owned by that PIC, subject to restriction (a) above.

12.6.2 Infinite Supply

First, as cell lines are created from any samples from one or more PICs, the individual PICs will be given the opportunity to receive viable aliquots of those lines as soon as feasible by the central laboratory, but no later than 3 months after establishment of the lines.

Within one to two years of the conclusion of the study, any biological materials in this category residing at the GADDCC or other central core laboratory will be treated as follows:

If a given population has requested return of the samples, that request will be honored. A viable aliquot of the material will be provided to each individual PIC that generated it. A viable aliquot will be archived for the NIDDK (mechanism to be determined). Ownership of the samples will reside with the possessors (e.g. the NIDDK and the individual PICs, respectively).

It is currently the intention of the NIDDK to establish a repository of cell lines with accompanying data to be available to outside, i.e. non-F.I.N.D., investigators. The full sample repository (final phase repository) will be established within 1 to 2 years, after the termination of the F.I.N.D. These samples must be anonymized, i.e. all identifiers removed. Ethnic groups will be identified by only one of 3 categories: Caucasian, African-American, Hispanic. The American Indian population will not be included in the final repository available to the public.

There may be a desire to establish an early phase repository based on relevant publications (e.g. the initial genome scan). Such an early phase repository would include the relevant publication data and samples available at that time. They would be anonymized. The desire/need for such a repository will be assessed by a to-be-established committee (such could include representatives from the F.I.N.D. Steering Committee, EAC, NIDDK, and non-F.I.N.D. representatives).

Access to the sample repository(ies) would be decided by a to-be-established committee, likely to consist of F.I.N.D. Steering Committee members, EAC members,

NIDDK, and possibly outside (non-F.I.N.D.) representatives. This will require an application process, and formal commitments by the requested investigators (whether in academia or industry) to commit to utilize the samples only for the F.I.N.D. purposes (i.e. to find the genes for kidney disease, diabetes, and associated diseases, e.g. retinopathy and atherosclerosis), and to deliver to the central database the primary data of any publication within 3 months after publication. Provision of these samples and accompanying data will require a funding mechanism to be developed, ranging from peer reviewed funding to a fee-per-sample mechanism.

12.7 Funding

All activities that are mandated by these policies will receive appropriate funding for the responsible parties to carry them out, otherwise these parties will not be held responsible for the tasks (i.e. there are to be no unfunded mandates).

13 PUBLICATIONS

13.1 Overview and Objectives of the Publications Subcommittee

The Publications Subcommittee will coordinate, monitor, review, and assume responsibility for, arranging the preparation of all study-wide press releases, interviews, presentations, and publications relating to the scientific aspects of FIND. During the planning, conduct, and analysis of FIND, there will be no communications of study plans or results that have not been reviewed and approved following the procedures as described below.

13.2 Duties of the Publications Subcommittee

Recommend policy and procedures for review and approval of all scientific communications regarding FIND to outside groups.

Identify publications to be written and presentations to be made during the course of the study, with target dates for each. The Publications Subcommittee will also review proposals for FIND-related publications or presentations. These will include publications related to the main study and those resulting from projects previously approved by the Subprojects Evaluation Subcommittee.

Propose policy guidelines for authorship of FIND publications and appoint writing groups.

Monitor the writing of each paper to ensure publication in a timely fashion.

Suggest appropriate journals for FIND publications and monitor the process of publication.

Perform other writing, reviewing, or editing tasks assigned by the Steering Committee.

Assist writing groups in publishing papers of the highest quality and clarity. Review and edit all FIND publications and presentations prior to submission (see section 13.4), enlisting the special assistance of other members of FIND Research Group and subcommittees whenever appropriate. The activities of the Publications Subcommittee will be conducted pursuant to the following policy regarding publications and presentations from FIND:

1. ensure accurate, uniform, timely, and high quality reporting of FIND activities and results;
2. preserve the scientific integrity of the study;
3. safeguard the rights and confidentiality of participants;
4. ensure that the timing of publications and presentations serves the right of the public to know the results of the study without jeopardizing the privacy of the participants.

In all of these matters, the Publications Subcommittee will, by majority vote, make recommendations to the Steering Committee. Final decisions will be by majority vote of the Steering Committee.

13.3 Categories of Papers

Definitions

Paper: a paper written for publication in a journal, monograph, or book, whether initiated by the group or invited.

Writing group: the individuals directly responsible for writing a paper, even though many others in the research group may be involved in review and approval before the paper is submitted. A writing group will usually be composed of three to seven people and include representatives from all the disciplines involved in the paper (e.g., clinicians, molecular geneticists, genetic epidemiologists, or statisticians).

Research Group: all investigators at the PICs, GADCC, NIDDK project office, and any subcontracting units, such as central laboratories. Investigators are considered those working at professional levels (and usually trained at the doctoral level), regardless of professional discipline (e.g., clinicians, molecular geneticists, genetic epidemiologists, or statisticians). The program coordinator at each PIC might be included regardless of degree or professional credentials.

Each paper will be categorized before the analysis and writing are undertaken. Categories are determined by the scope and content of each paper, rather than by the particular protocol or protocol(s) under which the data were obtained. Results obtained under the full protocol followed by all centers will probably be designated as category A (defined below), whereas results obtained in ancillary or sub-studies may be of any category, depending on their scope and content.

A. Full group papers

These papers represent the work of the entire consortium, including the PICs, GADCC, NIDDK project office, and any subcontracting units, such as central laboratories. Examples of topics are: overall study design and methods, descriptions of recruited populations, and findings representing all study subjects and PICs.

Authorship will be by the FIND Research Group. Group members will be listed in a footnote or appendix (depending on editorial policy of the journal). The list may designate certain roles of individuals, such as PIs, chairs of the Steering Committee and subcommittees, and members of the writing group. The corresponding author will be the GADCC, with no individual named, unless required by the journal, in which case the PI of the GADCC will be named. The appendix may also acknowledge by name or by group the nonprofessional staff, external advisors and consultants, and others as appropriate.

B. Subgroup papers

These papers represent the work of part of the consortium, consisting of at least one center, which may include the PICs, GADCC, NIDDK project office, and any subcontracting units, such as central laboratories. Examples of topics are: study design, methods, subjects, and findings pertaining to a subset of PICs or a single PIC; or methods (but not study results) used by the GADCC alone.

The involved centers may be a single PIC (including affiliated laboratories or consultants) if the services of the GADCC were not used in the substudy, the single center plus the GADCC, or more than one (but not all) PICs, with or without the GADCC. Depending on the roles of the GADCC in the substudy (i.e. statistical analysis or genotyping), the authorship list may include only some members of the GADCC.

Authorship will be the investigators of the centers involved. A subtitle, footnote, acknowledgment, or appendix will indicate that this represents a subgroup of FIND, and if possible, a FIND paper (category A) will be referenced.

C. Miscellaneous papers

These papers represent work funded by FIND but not central to the FIND consortium. Examples are: development of statistical or genotyping methods not illustrated by FIND results (even if subsequently used in FIND), descriptions of aspects of the study populations not directly related to FIND analyses (such as correlations among study variables). These papers may emanate from a single center (PIC or GADCC) or from a combination of one or more PICs, with or without the GADCC.

Authorship will be by the investigators of the centers involved. A subtitle, footnote, acknowledgment, or appendix will indicate how this paper is related to FIND, and if possible, a FIND paper (category A) will be referenced.

Abstracts and meeting presentations will be classified and approved as above. When required by the organization sponsoring the meeting, a corresponding author can be named, e.g. in a footnote or in the authors line, such as A. B. Member, for the FIND Research Group. If any member of the FIND Research Group receives an invitation to make a FIND presentation, the invitation should be forwarded to the Publications Subcommittee. The Publications Subcommittee will recommend whether the invitation should be accepted, and may suggest that the presentation be given by a member of the Research Group other than the person originally invited (to spread the workload of presentations and facilitate public recognition of members who may not be frequently invited).

13.4 Implementation of Policy and the Review Process

Topics for papers and members of writing groups may be proposed by the Publications Subcommittee or any member of the Research Group. The Publications Subcommittee will either reject the proposal or assign each proposal a category and writing group. A log of

such decisions will be maintained on the FIND website accessible to all members of the Research Group. Any members disagreeing with a Publications Subcommittee decision on categorization of a paper, membership in the writing group, or approval of a paper should express their concerns to the Chair of the Publications Subcommittee, who will relay them to the full Publications Subcommittee, if necessary, for resolution. Unresolved disagreements within the Publications Subcommittee will be referred to the Steering Committee for resolution. A decision of the majority of the Steering Committee will be binding.

The Publications Subcommittee will review a completed paper of any category. The review will include attention to scientific accuracy, clarity, and style. If a majority does not approve it, suggestions for revision will be given to the writing group for revision, after which the Publications Subcommittee again reviews the paper. When the majority of the Publications Subcommittee approves a paper, it will be forwarded to the Steering Committee for further suggested revisions or final approval by majority vote.

13.5 Use of Data

For purposes of publication and presentation policies, study data are defined as all data specified in the MOO pertaining to enrolled participants. Subjects evaluated for eligibility but not enrolled (for whatever reason) can be used by the individual PICs for other studies.

Study data will be owned jointly by the individual PICs and the GADCC, but will be kept at the GADCC. The GADCC and PICs will make no use of study data nor disclose them to any other parties except as specified in the Protocol or MOO, unless such use or disclosure is approved by a majority of the Steering Committee.

Final disposition of the data is covered by the policies developed by the Intellectual Property and Archives Subcommittee.

13.6 Presentations to Volunteer Participants

Presentation of results with clinical significance determined by the steering committee will be offered to the volunteer participants in an appropriate manner before they are presented to the public or the professional communities.

14 SUBPROJECTS AND ANCILLARY STUDIES

14.1 Objectives of the Subprojects and Ancillary Studies Subcommittee

Subprojects and ancillary studies will be evaluated with regard to scientific merit and the potential impact on the objectives of FIND. Projects that complement the objectives of FIND will enhance the value of the study and, therefore, are encouraged. The Subproject Evaluation Subcommittee will evaluate each proposal and make a recommendation to the Steering Committee. The Steering Committee's final approval is necessary before the sub-study can begin. In addition, the customary approval of the relevant Institutional Review Board(s) must be obtained.

14.2 Definition

Subprojects and ancillary studies are defined as investigations that are not set forth in the primary FIND protocol, but which require resources that are shared by all of the PICs (including analytic resources and genotypic data), or which relate to the molecular genetics of nephropathy. These include the following types of proposals:

1. Proposals that the GADCC conduct analyses of data collected by one or more individual PICs which are not part of the analyses proposed by the overall FIND protocol.
2. Proposals that the GADCC type candidate genes, which are not proposed in the FIND protocol, in samples collected by one or more individual PICs.
3. Proposals that individual PICs, or other collaborators, type candidate genes in samples obtained from more than one PIC, which are collected for the primary FIND protocol.
4. Analyses conducted by individual PICs or other collaborators, which employ genotypic data generated by the main FIND.
5. Proposals that all PICs collect data that is not specified in the primary protocol.
6. Proposals that an individual PIC investigate the molecular genetics of nephropathy (*e.g.*, by association or linkage studies utilizing candidate genes or genomic segments), in data collected by the individual PIC under the FIND protocol.

Analyses, conducted by an individual PIC, of clinical phenotypic data collected by that PIC or of molecular genetic data that are unrelated to nephropathy, do not require evaluation by the Subproject Evaluation Subcommittee. (Note that any such studies using data in FIND subjects are limited by the participants' consent.)

Subprojects are defined as proposals, meeting the above definition, that are to be conducted using FIND resources. Ancillary studies are studies that require supplemental funding or resources in addition to FIND resources. (For example, ancillary studies may be conducted with funding obtained from an RO1 grant from NIH. Since approval from FIND will be necessary before proceeding with such studies, it will generally be to the investigators' advantage to obtain approval for such studies from FIND prior to applying for supplemental funding.)

14.3 Criteria

The Subproject Evaluation Committee will evaluate proposals with regard to the following:

1. Impact on the objectives of FIND (no subproject should compromise the scientific integrity of FIND, adversely effect its public image, or cause a significant diversion of Find's resources. In addition, subprojects should not compete with or duplicate the main results of FIND).
2. Scientific merit (proposals should test an important hypothesis and the design and power of the proposal should be adequate to ensure a reasonable test of this hypothesis).
3. Feasibility (both from a practical and ethical standpoint- the input of the Consent and Ethics Subcommittee will be sought when necessary to ensure that the proposed project is within the scope of participants' consent).

14.4 Procedures

The Subproject Evaluation Subcommittee will meet at each FIND Investigators' meeting. Proposals for subprojects or ancillary studies should be submitted at least one month prior to each meeting to ensure that sufficient time is available for review.

A written request for approval of a proposed subproject or ancillary study should be submitted to the Chair of the Subproject Evaluation Subcommittee. The request should be brief (no more than 5 pages), but should contain sufficient detail so that scientific merit can be assessed. It should contain the following elements in order:

- A. description of the hypothesis(es) to be evaluated.
- B. background and significance of the study.
- C. the proposed methods.
- D. analysis strategy.
- E. plans for publication.
- F. source of funding.

If funds are being sought from outside FIND, the proposed source of such funds should be clearly identified. If additional data from subjects are to be collected (e.g., by collection of additional samples or by clinical examination), full details need to be provided. It should

also be specified whether molecular or statistical analyses are to be conducted by the GADCC.

Proposals may be assigned to have a detailed or abbreviated review. Any proposals that utilize genotypic data, clinical data or analytic resources common to more than one PIC (study types 1-5 above) will receive a detailed review; proposals that relate to the molecular genetics of diabetic nephropathy, but in which all of the work, including genotyping and data analysis, will be done by a single PIC (study type 6) will receive an abbreviated review. If a detailed review is necessary, the Chair of the Subproject Evaluation Subcommittee will assign a primary and secondary reviewer from among the committee members for each proposal. The opinions of outside experts will be sought at the discretion of the committee. A statistical review by the GADCC will also be conducted. The Subproject Evaluation Subcommittee will vote on approval or disapproval; approval is based on a simple majority of committee members. The Chair will then prepare a written summary of the committee's decision and forward this to the investigator(s). Copies of proposals designated for an abbreviated review will be forwarded by the Chair to the Subcommittee, which will vote on whether the proposal is consistent with the objectives of FIND. Since such proposals will not utilize the shared molecular or analytical resources of FIND, they will not require a detailed review by primary, secondary and statistical reviewers.

If the Subproject Evaluation Subcommittee approves the proposal, the proposal will be forwarded to the Steering Committee, along with the summary and recommendation of the Subproject Evaluation Subcommittee. Final approval of the proposal will be based on a majority vote of the Steering Committee. If the Steering Committee approves a subproject, it will proceed. If the proposal is an ancillary study, it may proceed with Steering Committee approval once funding is obtained.

The Subproject Evaluation Subcommittee may not initially approve the proposal, but may request further information from the investigators before rendering a final decision. In such cases, the investigators will be invited to submit a revised proposal.

If the Subproject Evaluation Subcommittee disapproves the proposal, the investigators may appeal to the Steering Committee, which may override the decision of the Subcommittee. If the Steering Committee also disapproves the subproject, it will not be conducted.

14.5 Publication

Publication of data from all subprojects and ancillary studies must comply with the policies of the FIND Publications Subcommittee.

APPENDIX A: INVESTIGATORS

Institutions and addresses	Investigators, Email, Addresses, Phone Numbers:
<p>Genetic Analysis and Data Coordinating Center (GADCC) ¹Case Western Reserve University Department of Epidemiology and Biostatistics 10900 Euclid Ave Wolstein Research Building 1308 Cleveland, OH 44106-7281 Fax: 216-778-3280</p>	<p>PI: Robert C. Elston ¹ rce@hal.cwru.edu Phone: 216-368-4312 Co Is: Jane Olson ¹ olson@darwin.cwru.edu Phone: 216-368-4314 Sudha Iyengar ¹ ski@po.cwru.edu Phone: 216-368-4388 Katrina Goddard ¹ katrina@darwin.cwru.edu Phone: 216-368-4411 FIND Study Coordinator: Sarah Ialacci ¹ sialacci@darwin.cwru.edu Phone: 216-368-5636 Informatics: Gyungah Jun ¹ gyungha@darwin.cwru.edu Phone: 216-368-5638 Carol Lorimor ¹ lorimor@darwin.cwru.edu Phone: 216-368-5628</p>
<p>¹University of California Los Angeles 924 Westwood Blvd., Suite 335 (Mail Box 15) Los Angeles, CA 90024 Fax: 310-794-6206</p> <p>²Medical Genetics SSB-378 Cedars Sinai Medical Center 8700 Beverly Boulevard Los Angeles, CA 90048 Fax: 310-423-0237</p>	<p>PI: Mohammed Saad ¹ Msaad@mednet.ucla.edu Phone: 310-794-6200 Co Is: Jerry Rotter ² Jerome.Rotter@cshs.org Phone: 310-423-6451 Kent Taylor ² kent.taylor@cshs.org Phone: 310-423-6451 Study Coordinator: Maria Budgett mbudgett@mednet.ucla.edu Phone: 626-979-4920</p>
<p>¹Harbor-UCLA Medical Center, Division of Nephrology 1000 West Carson Street Torrance, CA 90502 Fax: 310-782-1837</p> <p>²University of California, Davis Biological Chemistry and Medicine Tupper Hall Room 4303 Davis, CA 95616 Fax: 530-754-6015</p> <p>³University of California-Irvine UCI Medical Center Bldg 53, room 125, route 81 101 City Drive Orange, CA 92868</p>	<p>PI: Sharon Adler¹ sadler@rei.edu Phone: 310-222-3891 Eli Ipp ¹ ipp@gcrc.humc.edu Phone: 310-222-2503 Fax: 310-533-6972 Co Is: Michael F. Seldin² mfseldin@ucdavis.edu Phone: 530-754-6016 Hongzhe Li ² hli@ucdavis.edu Phone: 530-754-9234 Madeline Pahl ³ mpahl@uci.edu Phone: 714-456-5142 Study Coordinator: Janine LaPage ¹ jlapage@rei.edu Phone: 310—222-3891/vm: 310-222-4104 Carolyn Garcia ¹ cgarcia@rei.edu Phone: 310-222-3891 / vm 310-222-4104 Barbara Walker ³ bawalker@uci.edu Phone: 714-456-6554 Fax: 714-456-6236 Joyce Gonzalez ¹ Phone: 310-222-3891 vm: 310-222-4104</p>

<p>¹University of Texas Health Science Center at San Antonio Department of Medicine / Division of Nephrology 7703 Floyd Curl Drive, MSC 7882 San Antonio, Texas 78229-3900</p> <p>Fax: 210- 567-4712 (Dr. Abboud) Fax: 210-949-3348 (FIND Study)</p>	<p>PI: Hanna E. Abboud ¹ abboud@uthscsa.edu Phone: 210-567-4700 Co Is: Nedal Arar ¹ arn@uthscsa.edu Phone: 210-567-0075 Michael Stern ¹ stern@uthscsa.edu Phone: 210-567-4737 Ravi Duggirala ¹ duggirala@uthscsa.edu 210-567-4735 Rosemarie Plaetke ¹ plaetke@uthscsa.edu Phone: 210-567-0016 Fax: 210-949-3348 Pager: 210-751-6401 B.J. (Kashi) Kasinath ¹ kasinath@uthscsa.edu Phone: 210-567-4707 Study Coordinator: Cary Goyes ¹ goyes@uthscsa.edu Phone: 210-567-0015 Fax: 210-949-3348 Pager: 210-235-0680 Valeria Sartorio ¹ sartorio@uthscsa.edu Phone: 210-567-6429 Pager: 210-715-5881</p>
<p>¹Wake Forest University School of Medicine Dept of Internal Medicine/Nephrology Medical Center Blvd. Winston-Salem, N.C. 27157-1053 Fax: 336-716-4318</p>	<p>PI: Barry Freedman ¹ bfreedma@wfubmc.edu Phone: 336-716-6192 Co I: Stephen Rich ¹ srich@wfubmc.edu Phone: 336-716-5778 Fax: 336-716-5425 Don W. Bowden ¹ dbowden@wfubmc.edu Phone: 336-716-3912 Fax: 336-716-7200</p> <p>Study Coordinators: Sharon Warren ¹ shwarren@wfubmc.edu Phone: 336-716-6324 Gloria Brooks ¹ brooks@centernet.net Phone: 336-713-0168 Cell: 919-548-4004</p>
<p>¹Rammelkamp Center for Research and Education CWRU – MHMC G531 MetroHealth Medical Center 2500 MetroHealth Drive Cleveland, OH 44109-1998 Fax: 216-778-8248</p>	<p>PI: John R. Sedor ¹ jrs4@po.cwru.edu Phone: 216-778-4993 Co Is: Jeff Schelling ¹ jrs15@po.cwru.edu Phone: 216-778-3079</p> <p>Study Coordinators: Angel Pickens ¹ ahw@po.cwru.edu Phone: 216-778-1792/Room R614 Lisa Humbert ¹ lxh18@po.cwru.edu Phone: 216 778-7129 Fax: 216 778 – 8248 / R614 Linda Goetz Fradley¹ lfradley@metrohealth.org Phone: 216-778-4833 Ashwini Sehgal¹ axs81@po.cwru.edu Phone: 216-778-4156 Marigel Constantiner ¹ mconstantin@metrohealth.org Phone: 216-778-2986</p>

<p>¹Johns Hopkins Suite 2-600, 2024 Monument St. Baltimore, MD 21234 Fax: 410-955-0476</p>	<p>PI: Michael J. Klag ¹ mklag@welch.jhu.edu Phone: 410-955-0496 Co Is: Josef Coresh ¹ coresh@jhu.edu Phone: 410-955-0495 Lucy Meoni ¹ lmeoni@jhu.edu Phone: (410) 955-3700 Rulan Parekh ¹ rsparekh@jhmi.edu Phone: (410) 955-2467 Linda Kao ¹ wkao@jhsph.edu Phone: (410) 614-3822 Study Coordinators: Nancy Fink ¹ nfink@jhsph.edu Lihan Wei ¹ lwei@welch.jhu.edu Patricia Bayton ¹ pbayton@jhsph.edu Phone: (410) 614-3994/fax (410) 614-9793</p>
---	--

<p>¹Department of Internal Medicine 2211 Lomas Blvd., NE, ACC-5th Floor Albuquerque, NM 87131-5271 Fax: 505-247-1297</p> <p>³Department of Family & Community Medicine UNM School of Medicine 2400 Tucker Avenue, NE Albuquerque, NM 87131-5241 Fax: 505-272-4494</p> <p>⁴Department of Genetics Southwest Foundation P.O. Box 760549 San Antonio, TX 78245-0549 Fax: 210-670-3317</p> <p>⁵Indian Health Service Kidney Program 801 Vassar Drive, NE Albuquerque, NM 87106 Fax: 505-248-7697</p>	<p>PI: Philip Zager ¹ Pzag@unm.edu Phone: 505-272-4750 Meg Lamey/Secretary megl@unm.edu Co Is: Vallabh Shah ¹ Vshah@salud.unm.edu Phone: 505-272-9615 Fax: 505-272-2614 Christine Stidley ³ cstidley@salud.unm.edu Phone: 505-272-8704 Jean W. MacCluer ⁴ jean@darwin.sfbr.org Phone: 210 –258-9490 Andrew Narva ⁵ Phone: 505-843-3309 Study Coordinators: Arlene Bobelu ¹ Abobelu@salud.unm.edu Phone: 505-782-2578 Marina Scavini ¹ mscavini@unm.edu Phone: 505-247-4044</p>
---	---

<p>¹ NIDDK/ Phoenix Diabetes and Arthritis Epidemiology Section 1550 E. Indian School Rd Phoenix, AZ 85014 Fax: 602-200-5225</p> <p>² Phoenix Indian Medical Center 4212 N. 16th Street Phoenix, AZ 85016</p>	<p>PI: William C. Knowler ¹ knowler@nih.gov Phone: 602-200-5206 Co Is: Vincent Berkley ² vincent.berkley@mail.his.gov Phone: 602-263-1568 Robert Hanson ¹ rhanson@phx.niddk.nih.gov Phone: 602-200-5207 Robert Nelson ¹ rnelson@phx.niddk.nih.gov Phone: 602-200-5205 Study Coordinators: Lois Jones ¹ ljones@phx.niddk.nih.gov Phone: 602-200-5213 Camas Luethe ¹ cluethe@phx.niddk.nih.gov</p>
--	--

External Advisory Committee:

Dr. David Warnock-Chair
THT 647
1530 Third Avenue S
Birmingham, AL 35294-0006
Phone: 205-934-3585
Fax: 205-934-1879
dwarnock@nrtcdom.uab.edu

Dr. Georgia M. Dunston
Department of Microbiology
Howard University College of Medicine
520 West Street, N.W.
Washington, D.C. 20059
Phone: 202-806-6484/6288
Fax: 202-806-4508
gdunston@howard.edu

Dr. Stephen O'Brien
Geneticist
Chief, Laboratory of Genomic Diversity
National Cancer Institute
Frederick, MD 21702-1201
obrien@ncifcrf.gov
Laraine Main mainl@ncifcrf.gov

Dr. Ranajit Chakraborty, Ph.D.
Robert A. Kehoe Professor & Director Ctr for Genome Infor.
Department of Environmental Health
University of Cincinnati
3223 Eden Ave., Room 110 Kettering Lab
Cincinnati, Ohio 45267-0056
Phone: 513-558-4925
Fax: 513-558-0071
ranajit.chakraborty@uc.edu or chakrar@ucmail.uc.edu

Richard P. Lifton, M.D., Ph.D.
Yale University School of Medicine
295 Congress Avenue, BCMM 154D
New Haven, CTn06516
Richard_lifton@yale.edu

Professor Richard Spielman
Department of Genetics, 464 Clinical Res Bldg
University of Pennsylvania School of Medicine
415 Curie Boulevard
Philadelphia, PA 19104-6145
Phone: 215-898-5763
Fax: 215-573-5892
Spielman@pobox.upenn.edu
Megan O'Karma Phone: 215-573-7516

NIDDK/NIH:

¹ Director, Diabetic Nephropathy & HIV
Programs, NIDDK/NIH
Room 607
2 Democracy Plaza
Bethesda, MD 20892-5458
Fax: 301-480-3510

² Director, Division Of Kidney, Urology, &
Hematology NIDDK, NIH
Two Democracy Plaza
6707 Democracy Blvd., MSC 5458
Bethesda, MD 20892-5458 (20817 courier)
Fax: 301-402-4874

³ Program Director, Genetics and Genomics
Division of Kidney, Urologic and Hematologic
Diseases NIH-NIDDK
Two Democracy Plaza
6707 Democracy Blvd., MSC 5458
Bethesda, MD 20892-5458 (20817 courier)
Fax: 301-480-3510

Paul Kimmel¹ kimmelp@extra.niddk.nih.gov
Phone: (301)-594-7713
Josephine Briggs² BriggsJ@hq.niddk.nih.gov
Phone: (301)-496-6325
Rebekah Rasooly³ rasoolyr@extra.niddk.nih.gov
Phone: (301) 594-6007



<p>¹ Senior Scientist, SAIC Principal Investigator, NCI Laboratory of Genetic Diversity Frederick Cancer Institute, P. O. Box 8 Frederick, MD 21702 Fax: 301 – 846 - 1909</p> <p>² Director – Central Laboratory & Cell Immortalization Program National Center Institute Frederick Cancer Research & Development Building 560, Room 21-6A Frederick, MD 21702 Fax: 301-846-7139</p>	<p>Michael W. Smith¹ smithm@ncifcrf.gov Phone: 301-846-1913 Cherie Winkler² winkler@ncifcrf.gov Phone: 301-846-5747</p> <p>Mary Thompson² thompsonm@ncifcrf.gov Phone: 301 846–7055</p> <p>Russ Hanson² r_hanson@ncifcrf.gov Phone: 301 846-7055</p>
--	--

<p>¹ Fundus Photograph Reading center Park West One 406 Science Drive, Suite 400 Madison, WI 53711-1068 Fax: 608-263-0525</p>	<p>PI: Matthew Davis¹ davis@rc.ophth.wisc.edu Phone: 608-263-5749</p> <p>Co Is: Barbara Blodi¹ blodi@rc.ophth.wisc.edu Phone: 608-263-5749 Michael S. Ip¹ msip@rc.ophth.wisc.edu Phone: 608-263-5749 Michael M. Altaweel¹ mmltaweel@facstaff.wisc.edu Phone: 608-263-5749 Photography Unit Mike Neider¹ neider@rc.ophth.wisc.edu Phone: 608-263-9858 Hugh Wabers¹ wabers@rc.ophth.wisc.edu Phone: 608-263-0740 Marcia Schiffman¹ schiffman@rc.ophth.wisc.edu Phone: 608-263-6468 Grading Unit Larry Hubbard¹ hubbard@rc.ophth.wisc.edu Phone: 608-263-2245 Jim Reimers¹ reimers@rc.ophth.wisc.edu Phone: 608-263-9370 Coordination Unit Nancy Robinson¹ robinson@rc.ophth.wisc.edu Phone: 608-265-6317 Kathleen Glander¹ glander@rc.ophth.wisc.edu Phone: 608-263-6983 Data Management Linda Kastorff¹ Kastorff@rc.ophth.wisc.edu Phone: 608-263-4611 Finance and Contracts Jim Onofrey¹ onofrey@rc.ophth.wisc.edu Phone: 608-263-8836</p>
--	---

<p>University of Chicago Department of Medicine Section of Nephrology 5841 S. Maryland Ave. Suite S511 MC 5100 Chicago, Illinois 60637 Fax: (773) 702-8518</p>	<p>Co: Orly F. Kohn okohn@medicine.bsd.uchicago.edu Phone: (773) 702-1473</p> <p>Coordinator: Cathy S. Brown cbrown@medicine.bsd.uchicago.edu Phone: (773) 834-7859</p>
<p>Loyola University Chicago Loyola University Medical Center Division of Nephrology 2160 South First Avenue Maywood, Illinois 60153 Fax: (708) 216 – 4060</p>	<p>Co: David J Leehey dleehey@lumc.edu Phone: (708) 216 – 3306</p> <p>Coordinator: Gail Barone</p>
<p>The University of Alabama at Birmingham School of Medicine Division of Nephrology PB 234 1530 3rd Avenue South Birmingham, Alabama 35294-0007 Fax: (205) 975-0215</p>	<p>Denyse Thornley-Brown dtb@nrtc.uab.edu Phone: (205) 934-0107</p> <p>Coordinator: Carthenia Jefferson cwj@uab.edu Phone: (205) 996-2225</p>

APPENDIX B: SCREENING LOG

Recruitment Tracking

Log entry:

Clinical Center ID: MSP
Barcode: 3F2C55A96C8FB4B2852568CF004FF22A
Date of Contact: (mm/dd/yyyy) 03/30/2000
Age: (99=unknown) 45
Ethnicity: Caucasian
Gender: Male Female
Ineligible:
Eligible: (Study ID)

Log report:

Barcode	Clinical Center ▲	Date of Contact ▲	Age	Ethnicity ▲	Gender ▲	Ineligible	▲ Eligible
FIND							
13AB502AC852568	Det	03/10/2000	73	99	1	Not interested	
162AFAB13AB502	Cle	05/01/2000		99	2	Wrong Diagnosis	
3F2C55A96C8FB4	MSP	03/30/2000	45	1	1		
502AC852568C800	Det	03/01/2000		1	2		
62AFB13AB502AC	Cle	05/01/2000		99	1		FIND
BC0DF216502AC8	MSP	07/13/2000	23	5	2		
CB88F46D378E01	Cle	04/01/2000	81	2	1	No siblings	
E34577D5FCBAF8	Det	04/30/2000	52	1	1		

APPENDIX C: DATA ANALYSIS

Overall Approach

We propose to utilize model-free linkage analysis methods in general, and ARP analysis in particular, as the leading analysis tool because Diabetic Nephropathy is a complex disease with an unknown mode of inheritance, and we are uncertain of the homogeneity of collection methods across PIC's. To take full advantage of the information in all contributed data sets, we will follow-up ARP analysis with methods that incorporate both ARP and DSP information. Periodically, the results of data analysis will be presented to the Steering Committee and the EAC.

Genome Scan Using Sib Pairs

The basic genome scan using ARP analysis will be performed in two phases. We conservatively estimate that the first set of ARPs/DSPs will be available in 2002, if only new data are to be collected. If ARPs and DSPs have already been ascertained by some PICs, the timetable to perform the genome scan will be revised and genotyping will begin earlier than presently expected. In the first stage, genetic markers approximately 10 cM apart will be typed and lod scores computed using multipoint ARP or ARP/DSP techniques. For chromosomal regions that achieve a maximum multipoint lod score of 0.8 (approximately $p=0.05$), we will type additional 6-8 markers in the 10-20 cM region surrounding the highest lod score and reanalyze the region. This two-stage design was selected with the computer program DESPAIR, part of the S.A.G.E. suite of programs. DESPAIR chooses the most cost effective two-stage strategy for a genome scan. In the first phase, a two-stage scan will be used on an initial data set of about 400 ARPs and about 250 DSPs.

In the second phase of the genome scan we will include the second set of ARPs and DSPs, which will become available by the fourth year of the study. We anticipate that the combined total from both sets will be approximately 1200 ARPs and 750 DSPs. First we will genotype the second set of ARPs and DSPs with markers at chromosomal regions found in the initial set to have positive lod scores. (This will provide enough information to detect relationships other than full sibs). In those regions where we observe concordance in positive results between sets 1 and 2, we will follow-up with additional markers to reach a final inter-marker density of 2 cM or less. In this way, we will focus immediate attention on regions that appear to be interesting at first pass. If an application to the Center for Inherited Disease Research for initial genotyping is successful, the second set of data will undergo exactly the same genotyping procedure as the first set.

The second set, analyzed separately, will provide a replicate that can be used to informally discriminate between true and false positives. Analyzed together, the 1200 ARPs and 750 DSPs will provide considerable power for detecting susceptibility loci.

Relative Pair Analysis

Conceptually, ARP analysis is straightforward; it determines if the numbers of marker alleles shared identical-by-descent (IBD) in a set of affected relatives deviate significantly from Mendelian expectations. At a marker locus unlinked to disease, the proportions of ARPs that share 0, 1 or 2 alleles under Mendelian expectations are 0.25: 0.5: 0.25, respectively. Significantly increased sharing among affected siblings suggests that the marker is linked to the disease locus (Kruglyak and Lander, 1995; Lander and Kruglyak 1995; Risch, 1990). In practice, the situation is more complicated because one cannot unambiguously determine the number of alleles shared that are IBD at every position along the genome. Maximum lod scores (MLS) are calculated using the estimated probabilities of each ARP sharing 0, 1, or 2 alleles IBD at each location in the genome using multipoint methods. The vector denotes the corresponding allele-sharing parameters in the likelihood $z = (z_0, z_1, z_2)$, where $z_j = \Pr\{\text{ARP shares } j \text{ alleles IBD}\}$, subject to $z_0 + z_1 + z_2 = 1$ (Risch, 1990) At the disease locus, for full sibs $z = [1/4\lambda_s, \lambda_o/2\lambda_s, \lambda_m/4\lambda_s]$, where λ_s , λ_o and λ_m are locus-specific relative risks for siblings, offspring, and monozygotic twins, respectively. For a set of n ASPs, the lod score at a given point in the genome is:

$$\text{lod}(z) = \sum_{i=1}^n \log_{10} \left[\frac{P(I_{m_i} | \text{ASP}; z)}{P(I_{m_i} | \text{ASP}; z = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4}))} \right], \text{ where } I_{m_i} \text{ is the marker data for family } i.$$

The parameter vector z is usually maximized within the “possible triangle” boundaries (Holmans, 1993) that are consistent with genetic models (i.e. $z_0 \geq 0$, $z_1 > 1/2$, $z_2 > 2z_0$). However, maximization will also be performed (i) without the “possible triangle” restriction, because this restriction leads to false positive inferences if the model assumptions do not hold and (ii) with an analogous restriction to be negative, which constrains the variance components (Olson, 1997); if the latter leads to a larger lod score than for the possible triangle; this is evidence against linkage.

For each ARP, the number of marker alleles shared IBD will be either definitively known based on the parental and unaffected sib genotypes, or inferred probabilistically using the allele frequencies at each locus. The latest version of the S.A.G.E. software will be used to calculate lod scores and p-values at each marker locus, and at any point between two markers, using full multipoint information. One limitation of the originally proposed methods is that they only consider IBD sharing at a single locus, which limits their power to detect linkage. To overcome this limitation, several multipoint IBD methods have been proposed. Because multipoint IBD sharing maximizes the available marker and sib information, it provides higher power for detecting linkage than the single-point methods. For most analyses, IBD sharing probabilities will be calculated using multipoint techniques (Lander and Kruglyak, 1995) using the S.A.G.E. program GENIBD. We will also perform single-point analyses to ensure that the multipoint lod score peaks are not a result of a single marker with substantial genotyping error. Population allele frequencies will be estimated using maximum likelihood from all the family data collected, with the underlying assumption is that none of the marker alleles are in linkage disequilibrium with a disease locus. While this assumption may not be valid for all marker alleles, it is robust in that linkage disequilibrium will not lead to a false finding of linkage. The significance of the ARP lod score can be assessed under a large sample assumption of a mixture of chi-squareds, but, due to the uncertainty of the large sample assumption for the ARP approach, simulation will be used to assess significance levels for any potentially interesting results.

To perform a combined analysis, of ARPs and DSPs, we estimate z_0 , z_1 , and $(y_0 = \Pr\{\text{DSP shares 0 alleles IBD}\})$. We will extend the method to include covariates such as the PIC type of predisposing condition (Type I diabetes, Type II diabetes, polycystic kidney disease, etc.) and age, so that we can increase power and better evaluate the features that best predict linkage to a particular chromosomal region. To incorporate covariates, we will reparameterize the likelihood by modeling:

$$\log \left\{ \frac{2 P(\text{ASP} | I_m; x)}{P(\text{DSP} | I_m; x)} \right\} = \log \left\{ \frac{P(\text{affected sib} | \text{affected proband}; I_m; x)}{P(\text{unaffected sib} | \text{affected proband}; I_m; x)} \right\} = \alpha + \gamma_1 \pi + \gamma_2 f_2 + \beta^T x,$$

where x is a vector of (possibly centered) covariates, π is the mean allele-sharing, and f_2 is the probability that the pair shares exactly two alleles IBD. If $x=0$, the model reduces to (a reparameterization of) the usual affected/discordant-sib-pair likelihood described above. Genetic variance components, recurrence and relative risks, and allele-sharing probabilities, conditional on the value of x , are functions of the model parameters and are easily obtained.

Quantitative Trait Linkage Analysis

We will analyze phenotypes such as proteinuria, AC ratio, HbA1C to detect QTLs. If ARPs and DSPs and concordant unaffected sib pairs are all present, we can analyze the combined data set by treating the binary disease indicator as a quantitative variable and applying statistical methods for quantitative trait linkage analysis. One of the most commonly used methods for examining linkage between a complex trait and marker loci is the Haseman-Elston (H-E) method (Haseman and Elston, 1972). The basic idea, underlying the method is that sib pairs who are IBD at a marker locus will be phenotypically similar for traits influenced by a nearby linked gene. Conversely, sib pairs who do not share genes IBD at the linked marker will tend to be phenotypically different. The sib-pair trait difference is squared and regressed on the proportion of marker alleles the pair is estimated to share IBD. All sib pairs are used, whether or not the IBD sharing can be unambiguously determined. Linkage is suggested if there is a significant negative regression coefficient; if there is linkage, similarity in the sibs' trait values should be correlated with similarity in the marker alleles they inherit. Recently, the original H-E method has been modified by changing the dependent variable from the squared difference to the mean-corrected product of the sibs' trait values, thus modeling the sib covariance directly; this includes information available in the sib-pair trait sum that was ignored in the traditional H-E regression. It is simple to study multiple trait loci, and epistatic interactions, since they involve only multiple linear regression. The regression procedure also allows for covariates, and hence can be more powerful, in a computationally efficient manner. Investigations are underway to increase the power of the method further. The most powerful valid extension of this method that is available at the time the data have been collected will be used for the analysis. Simulation will automatically be used to obtain valid p-values in the latest implementation of this method in S.A.G.E.

Preliminary Analysis and Data Checking

Descriptive Statistic Analysis

The PEDINFO program of the S.A.G.E. package will produce descriptive statistics about nuclear family data. These include means, variance, histograms of sibships as well as size and count of each type of relative pairs etc.

Relationship Testing

In a sib pair study, particularly one in which parents are not genotyped, it is crucial that putative sib pairs be confirmed as actual sib pairs to avoid bias in the linkage results. We will subject each putative sib pair in our study to a genome-wide relationship test prior to inclusion of the sib pair in the linkage analysis.

Estimating Age-at-Onset

Since unaffected siblings cannot be considered non-susceptible with a probability of 1, an age of onset (or duration of diabetes at onset) probability distribution function, which will be estimated from the sample, will be included in the calculations that use discordant pairs. Let a_j be the age of sib j , p the power of the age of onset that can be assumed to be approximately normally distributed, μ the mean of the distribution of a^p where a is the age of onset, σ the standard deviation of a^p , and γ the unconditional prior probability that a sib is susceptible. Then the conditional probability that sib j is susceptible, given whether affected or not at age a_j , is given by:

$$x = \begin{cases} 1 & \text{If affected} \\ \frac{\gamma - \gamma \Phi\left(\frac{a_j^p - \mu}{\sigma}\right)}{1 - \gamma \Phi\left(\frac{a_j^p - \mu}{\sigma}\right)} & \text{If not affected by age } a_j \end{cases}$$

where Φ is the standard cumulative normal distribution function. In this study, the program AGEON in the S.A.G.E. program package will be used to estimate μ , σ , p and γ . The program can also estimate γ as a function of parental affection status if sufficient parental data are available. Alternatively, in order to investigate linkage to genes that affect age of onset, rather than susceptibility, we define the survival analysis residual

$$x = \begin{cases} 1 - \gamma \Phi\left(\frac{a_j^p - \mu}{\sigma}\right) & \text{If affected at age } a_j \\ -\gamma \Phi\left(\frac{a_j^p - \mu}{\sigma}\right) & \text{If unaffected by age } a_j \end{cases}$$

Using one of these values of x as a quantitative trait can be more powerful in the usual Haseman-Elston test for linkage than using disease status as a simple binary trait (Zhu, et al, 1997, Hanson and Knowler, 1998). Not only can this quantitative trait determined by

AGEON be used to obtain more power in the new H-E test (Elston et al., 2000), but it can also be added to the DSP likelihood analyses. The DSP likelihood contribution can be written in the form: $x L_{ARP} + (1-x) L_{DSP}$, where L_{ARP} and L_{DSP} are the ARP and DSP likelihood contributions, respectively.

Power To Detect Linkage

Required sample size estimates, based on affected sib pairs alone, were computed for the most economical two-stage affected sib pair design and are based on an initial scan using 380 markers (average distance < 10 cM). Because we cannot know, *a priori*, the effect size (λ) of any single gene, we have calculated the sample size requirements for λ ranging from 1.2 to 2.0. The data on other family members will increase the statistical power of our analysis and allow a margin of error if the sample of affected siblings alone is insufficient to conclusively show linkage. Because the statistical power estimates in the following tables are based on a less than optimal statistic and only on expected affected sibling pair data, they should be considered conservative lower bound estimates

Power Analysis for Two Stage Genome Scan

The following tables give, for each value of the locus-specific recurrence risk λ and initial marker spacing in cM, the sample sizes needed to obtain a lod score of 3 with power .99 when using a two-stage design. In parenthesis, in selected cells in the following table, is the first stage significance level that determines where second stage markers should be typed. In each cell, the following three quantities are given:

Number of Sibpairs / Number of Flanking markers on each side at second stage

Total cost (unit = cost of 1000 marker assays)

I. Assuming markers with a Polymorphic Information Content value of 0.7

Cost of recruiting 1 individual = cost of 1000 marker assays

8	Initial Marker Spacing			
	5	10	20	27
1.2	3608 / 5 12,724	3928 / 3 11,465 (.043)	4948 / 2 14,344	4888 / 3 16,739
1.4	1096 / 5 3,880	1200 / 3 3,546 (.052)	1534 / 3 4,511	1514 / 3 5,234
1.6	574 / 5 2,041	600 / 5 1,963 (.074)	821 / 2 2,445	810 / 3 2,822
1.8	372 / 5 1,330	413 / 3 1,253 (.070)	544 / 2 1,639	536 / 3 1,880
2.0	270 / 5 970	301 / 3 928 (.081)	403 / 2 1,227	397 / 3 1,440

Cost of recruiting 1 individual = cost of 500 marker assays

Initial Marker Spacing				
8	5	10	20	27
1.2	3608 / 5 9,116	3928 / 3 7,357 (.043)	4948 / 2 9,396	5639 / 2 11,223
1.4	1006 / 5 2,782	1267 / 2 2,339 (.041)	1534 / 2 2,977	1761 / 3 3,544
1.6	579 / 4 1,464	670 / 2 1,251 (.048)	821 / 2 1,624	949 / 2 1,928
1.8	376 / 4 9,554	439 / 2 829 (.056)	544 / 2 1,095	633 / 2 1,296
2.0	273 / 4 697	321 / 2 614 (.064)	403 / 2 824	472 / 2 973

II) Assuming completely informative markers

Cost of recruiting individual = cost of 1000 marker assays

Initial Marker Spacing				
8	5	10	20	27
1.2	3520 / 7 12,264	3608 / 5 10,006	3821 / 4 9,613	4107 / 3 10,214
1.4	1069 / 6 3,720	1083 / 7 3,042	1165 / 4 2,963	1259 / 3 3,70
1.6	557 / 7 1,944	570 / 6 1,599	613 / 4 1,577	665 / 3 1,696
1.8	360 / 8 1,261	369 / 6 1,039	399 / 4 1,039	436 / 3 1,125
2.0	260 / 8 912	270 / 5 758	301 / 3 767	319 / 3 834

Cost of recruiting 1 individual = Cost of 500 marker assays

Initial Marker Spacing				
8	5	10	20	27
1.2	3528 / 6 8,745	3608 / 5 6,398	3928 / 3 5,756	4107 / 3 6,107
1.4	1073 / 5 2,654	1096 / 5 1,951	1200 / 3 1,780	1355 / 2 1,904
1.6	557 / 7 1,387	579 / 4 1,026	632 / 3 950	719 / 2 1,023
1.8	363 / 5 899	372 / 5 668	413 / 3 629	416 / 4 720
2.0	260 / 8 652	273 / 4 488	301 / 3 466	348 / 2 507

Conclusion: the project enrollment of ARP (alone) will have ample power ($> .99$) to detect a linked locus with $\lambda \geq 1.4$ segregating in all ethnic groups, or a linked locus with $\lambda \geq 2$ segregating in either AA or MA alone. There will still be considerable power (e.g. $> .8$) to detect smaller values of λ .

Family-Based Association Studies

For fine mapping of the regions of interest obtained from the first stage genome-wide screening and for examining candidate genes, the following methods will be employed. The primary analysis tool for family-based association tests will be the sib TDT. However, we describe both the original TDT and sib TDT methods in case other types of data are also available for analysis.

Transmission/Disequilibrium Tests (TDT)

The TDT was first introduced by Spielman et al. (1993) as a test for linkage between a complex disease and a genetic marker, using nuclear family data. The TDT statistic compares the frequencies of alleles transmitted to an affected child and alleles that are not transmitted. This approach eliminates the concern that the population substructure may be the sole cause of the association. One of the requisites for using this method is that there is an association between the marker and the disease; otherwise, the TDT statistic has no power to detect linkage. Although the main purpose in developing the TDT statistic was to provide a test for linkage in family-based allelic association studies, the TDT statistic has also been recommended as a test for fine localization of disease genes, when linkage is not in question (Risch and Merikangas, 1996). The TDT can be used to test linkage in simplex (data are from single affected sibs), multiplex (data are from two or more affected sibs) or multigenerational families (Spielman and Ewens, 1996). The statistical significance is tested by the standard χ^2 (McNemar's test). The disadvantage of this approach is that the TDT has no power to detect linkage if association is not present.

Sib Transmission/Disequilibrium Test (TDT)

We expect that parental DNA will be unavailable for a substantial proportion of families, and the sib-TDT statistic can be used for analyses in place of the regular TDT. This sib-TDT compares marker allele frequencies in affected and unaffected siblings. However, Knapp (1998) has recently proposed a better method. This method will be included in the S.A.G.E. package.

Subanalyses and Covariates

Subcategorizations include racial group, gender, contributing PIC and clinical subcategories such as level of disease progression and other clinical aspects of diabetic nephropathy. In general, initial phenotypic stratifications will subdivide the sample into two sets, with further stratifications made where statistical power permits. Analyses of different subgroups will allow us to determine whether: 1) positive findings in the whole sample are found within subgroups; 2) subgroups lead to positive findings not present in the whole sample. In addition, overall analysis will be performed that combine results for each stratum in a Mantel-Haenszel fashion.

Environmental factors may either enhance or inhibit gene expression. Familial dependence may limit analyses to the extent that entire families share the same exposure. For example, within an individual family the exposure may be “confounded” with genetic factors in that all subjects from that family, affected and unaffected, those with and without genetic factors, are concordant with respect to the environmental factor. The analyses of environmental factors will focus on the interactions of genes identified in the analysis of genetic data and environmental factors that are associated with Diabetic Nephropathy and show some heterogeneity within families. Interactions between genetic and environmental factors will be tested using matched-pair case/control analyses comparing affected vs. unaffected siblings. Interactions will be represented in logistic models as the product of an indicator variable for presence of the gene and an indicator variable for the presence of the environmental factor. The ARP ascertainment scheme will not invalidate any tests of gene-environment interaction, but the interpretation of such tests, if significant, will not be trivial.

Genome-wide Association Study

Meta-Analysis

Meta-analysis will enable us to confirm initial findings using the FIND samples, and will increase our power to detect genetic variants of small effect. We will first identify DN cases and controls from the Genetics of Kidneys in Diabetes (GoKinD) Study (Mueller et al., 2006) and from the Finland-United States Investigation of Non-insulin-dependent Diabetes Mellitus Genetics (FUSION) Study type 2 DM case sample (Scott et al. 2007), families containing at least one DN affected or discordant sib pair from the Framingham Heart Study SHARe Genome-wide Association Study (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v2.p1) and cases and controls from the Type 1 Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications (DCCT/EDIC) study (The Diabetes Control and Complications Trial Research Group. 1993, Epidemiology of Diabetes Interventions and Complications. 1999). DN cases and DM-positive, DN-negative controls will be selected using criteria as similar as possible to the FIND criteria (Knowler et al. 2005), subject to the available phenotype data. In addition to the studies mentioned previously, the FIND consortium plans to request data from the following cohorts in order to augment the power of our meta-analysis and confirm our initial findings: African American Study of Kidney Disease (AASK), Action for Health in Diabetes (Look AHEAD), Diabetes Prevention Program (DPP), The Finnish Diabetic Nephropathy Study (FinnDiane), Chronic Renal Insufficiency Cohort Study (CRIC), The Diabetic Retinopathy Clinical Research Network (DRCR), The Hypertension Genetic Epidemiology Study (HyperGEN), The Diabetes Genome Anatomy Project (DGAP), Women’s Health Study (WHS), European Rational Approach for the Genetics of Diabetic Complications (EURAGEDIC), Genetics of Hypertension Associated Treatment (GENHAT), Accelerated Mortality on Renal Replacement (ArMORR), The Jackson Heart Study, The Golden Years Cohort, The Skaraborg Diabetes Registry, The University College Dublin Diabetes Center, and the Queen’s University in Belfast Diabetes Center. All of these studies contain data on diabetes, diabetic retinopathy, or diabetic nephropathy that will be an invaluable resource for increasing power to detect genetic variants of small effect.

Genotypes from FIND samples will be called from raw fluorescence data using an empirical Bayes approach (Lin et al., 2008). All genotype data will be subject to standard quality control criteria with regard to call rate, deviation from Hardy-Weinberg equilibrium, and, where possible, Mendelian inconsistency (WTCCC, 2007; Zeggini et al., 2007; Zeggini et al., 2008). In addition, we will enhance the coverage of the Affymetrix 6.0 chip by imputing genotypes for SNPs in the HapMap II marker set (International HapMap Consortium, 2007) not in Affymetrix 6.0, via the IMPUTE (Marchini et al., 2007) and/or MaCH (<http://www.sph.umich.edu/csg/abecasis/MaCH/download/>) programs. Based on previous investigations (Scott et al., 2007; WTCCC, 2007; Zeggini et al., 2007; Zeggini et al., 2008), we expect to be able to add about 1.5 million SNPs to the overall scan through imputation.

Population stratification is a common cause of type I (false-positive) error in association studies, even when samples are ethnically uniform. We will take a multilevel approach to correcting for population stratification, with separate means of adjustment across and within ethnic groups. We will examine the individual study samples for genetic substructure by multidimensional scaling based on identity-in-state allele sharing (Purcell et al., 2007), and remove or reassign individuals who appear to be misclassified with respect to study population. For joint analyses incorporating multiple FIND ethnic groups, we will include covariates indicating group membership into the regression model in the S.A.G.E. ASSOC program (see below) (Haiman et al., 2008). We will test for heterogeneity in SNP effect across ethnic groups by conducting likelihood ratio tests comparing a model including terms for SNP \times ethnicity interactions with a model lacking such terms. Next, we will evaluate population substructure within case/control subsamples by the method of principal components (Price et al., 2006). If necessary, we will adjust for significant principal components in regression-based analyses (see below). Finally, we will inspect the distribution raw p values for residual inflation of significance by estimating the genomic-control parameter λ (Devlin and Roeder, 1999).

We will adopt a two-stage design for genome-wide association analysis, in which both Stage 1 and Stage 2 samples will be included in the follow-up analysis (Skol et al., 2006). Quality control measures and criteria may differ among data sets from different studies. For this reason, and because of the likely variation in phenotype definition, we will not attempt to merge samples from different studies. Instead, we will conduct meta-analysis at any point at which the sample includes multiple data sets (Scott et al., 2007; Zeggini et al., 2007; Zeggini et al., 2008). Briefly, SNPs that pass a significance criterion in stage 1 will be tested in the pooled Stage 1 and Stage 2 samples. Genotypes not available in the Stage 2 samples will be imputed (see above). Given the sizes of the available samples that meet our criteria for DN affected/unaffected status from GoKinD, FUSION and Framingham SHARe, we will choose the criteria for promoting SNPs to Stage 2 that optimize the overall power to detect association (Skol et al., 2006). We will genotype imputed SNPs in the FIND families that meet the significance criteria, and will select them for Stage 2 only if the p values from directly measured genotypes also meet the criteria.

We will conduct initial tests of association either by logistic regression (for the binary DN phenotype) or by linear regression (for quantitative measures of kidney function), using an additive model for SNP genotype, and adjusting for relevant covariates. We will analyze family-based data via the ASSOC program in S.A.G.E., and case-control data using PLINK

(Purcell et al., 2007). We will combine results across study samples by one of two approaches. In the case of binary traits, we will construct a weighted average of the odds ratios (ORs) from all study samples, weighted by the variance of the estimate of the log(OR). For quantitative traits, we will calculate a similar weighted average of the SNP-specific estimates of regression coefficients weighted by their variances. Evidence for disease-predisposing variants in a meta-analysis is most compelling when the estimates of effect are consistent among samples. We will assess consistency of ORs and allelic effects through the I^2 statistic of Higgins et al. (2003).

MALD Study

Haplotype Based Method

If a marker locus is linked to a locus where genetic variation underlies a difference in disease risk in two founding populations, the proportion of alleles at the marker locus that have ancestry from the high-risk population will be higher in affected individuals than in unaffected individuals (or, alternatively, disease protecting alleles can be examined). In particular, we will consider the AA population, which is regarded as a recently admixed population between European Americans (EA) and Africans. Assume that we have $m/2$ diabetic AA patients with nephropathy (cases) and $n/2$ diabetic AA patients without nephropathy (controls), and these patients are genotyped at markers along a given chromosome. The goal is to identify a chromosomal region where the alleles of the haplotype in this region are more likely to have a specific ethnicity, eg. EA, by descent in the case population than in the control population.

Let a, b be the two endpoints of the interval of a given region. We define the following statistic:

$$S(a,b) = \frac{\sum_{i,j} \sum_k |p_{i,k} - q_{j,k}|}{mn}, \quad (1)$$

$i = 1, \dots, n$
 $j = 1, \dots, m$
 $k = 1 \dots K(a,b),$

where the first sum is over all possible $m \times n$ comparisons of the case and control haplotypes, and the second sum is over all the $K(a,b)$ markers in the interval $[a,b]$, $p_{i,k}$ is the allele frequency in the (EA) population of allele k in the i^{th} case haplotype, and $q_{j,k}$ is the allele frequency in the (EA) population of the allele k in the j^{th} control haplotype. This statistic measures the average difference of the alleles that are EA by descent between case and control haplotypes in the interval $[a,b]$.

Let $[a_0, b_0]$ be the interval that corresponds to the maximum value of $S(a,b)$, i.e., $S(a_0, b_0) = \max_{a,b} S(a,b)$. This interval identifies the interval that shows the largest segment of EA ethnicity by descent when comparing case and control haplotypes. This interval is most likely to harbor the disease genes that underlie the ethnic differences in disease risk.

If the haplotypes cannot be uniquely determined for some interval $[a,b]$, we can replace $S(a,b)$ by its expectation.

Significance will be determined by permuting case and controls.

Another way of analyzing the haplotype data is similar in spirit to multipoint linkage analysis using the Hidden Markov model, i.e., the Lander-Green algorithm. Consider the region the haplotype covers. For any given point, using the Hidden Markov model, we can calculate the probability of being EA by descent for the i^{th} haplotype of the cases, denoted here p_i , for $i=1, \dots, n$, where n is the number of case haplotypes. Similarly, we can calculate the probability of being EA by descent for the j^{th} haplotype of the controls, denoted here q_j for $j=1, \dots, m$, where m is the number of the control haplotypes. We can then calculate the Wilcoxon rank-sum score statistic for comparing $p = (p_1, \dots, p_n)$ and $q = (q_1, \dots, q_m)$. The site with the largest score is the location that shows the largest difference in the probability of an allele being EA by descent between the cases and the controls. This is also the location that very likely harbors the disease gene.

Power Analysis For MALD

In general the power of this approach is based on: 1) the difference in disease susceptibility genes in the two admixed populations; 2) the number of generations since admixture; 3) other factors that may influence meiotic recombination between DNA of diverse ethnic origin; 4) the frequency of the “disease” genes in the general population of the admixed population conferring increased risk or the frequency of “disease protective” genes in the general population of the lower risk population; 5) the information content (ethnic discrimination) of the markers; 6) the % admixture of the population; and 7) the genotypic or relative risk (RR) of the disease genes. In addition, the DNA pooling strategy that is a component of this protocol is also constrained by the frequency of the relevant genes in the admixed and control populations and the % admixture, because small differences in allele frequencies between probands and controls cannot be discriminated.

The potential power of these studies has been examined by simulation analyses. These suggest that this method can identify critical regions of the genome for “disease modifying” genes of low genotypic relative risk even in generally unfavorable inheritance models (e.g. recessive) (McKeigue, 1998). These recent power estimates assumed we use 2000 diallelic markers that are highly informative for ethnicity and that conditioning on the parental contribution of admixed genes provides a measure of the potential value of MALD. When admixture occurs 2 to 10 generations prior to analysis, admixture ratios are between 20% and 70% and the simulations suggest that 500 to 1000 families can provide greater than 80% power to detect a “disease gene” with relative risk ratios > twofold. A similar level of power should be obtainable using a substantially smaller number of highly informative multiallelic markers and an ethnicity-by-descent haplotype based method of analysis (see analysis section). Modeling is currently being explored to further examine this issue.

References

(1993) The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Engl J Med*. 1993 Sep 30;329(14):977-86.

- (1999) Epidemiology of Diabetes Interventions and Complications (EDIC). Design, implementation, and preliminary results of a long-term follow-up of the Diabetes Control and Complications Trial cohort. [Diabetes Care](#). Jan;22(1):99-111
- Devlin B, Roeder K (1999). Genomic control for association studies. *Biometrics* 55:997-1004.
- Elston RC, Buxbaum S, Jacobs KB, Olson JM (2000). Haseman and Elston revisited. *Genet Epidemiol* 19:369-372.
- Haiman CA et al. (2008). Comprehensive association testing of common genetic variation in DNA repair pathway genes in relationship with breast cancer risk in multiple populations. *Hum Mol Genet* 17:825-834.
- Hanson R, Knowler W (1998). Analytic strategies to detect linkage to a common disorder with genetically determined age of onset: diabetes mellitus in Pima Indians. *Genet Epidemiol* 15:299-315.
- Haseman, JK., Elston RC (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 2:3-19.
- Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. *Brit Med J* 327:557-560.
- Holmans P (1993). Asymptotic properties of affected sib-pair linkage analysis. *Am J Hum Genet* 52:362-374.
- The International HapMap Consortium (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449:851-861.
- Knapp M (1998). The transmission/disequilibrium test and parental genotype reconstruction: The reconstruction-combined transmission/disequilibrium test. *Am J Hum Genet* 64: 861-870.
- Knowler WC et al. (2005). The Family Investigation of Nephropathy and Diabetes (FIND): design and methods. *J Diabetes Complications* 19:1-9.
- Kruglyak L, Lander ES (1995). Complete multipoint sib-pair analysis of qualitative and quantitative traits. *Am J Hum Genet* 57:439-454.
- Lander E, Kruglyak L (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241-247.
- Lin S, Carvalho B, Cutler DJ, Arking DE, Chakravarti A, Irizarry RA (2008). Validation and extension of an empirical Bayes method for SNP calling on Affymetrix microarrays. *Genome Biol* 9:R63.
- McKeigue PM (1998). Mapping genes that underlie ethnic differences in disease risk: methods for detecting linkage in admixed populations, by conditioning on parental admixture. *Am J Hum Genet* 63:241-251.

Mueller PW et al. (2006). Genetics of Kidneys in Diabetes (GoKinD) Study: a genetics collection available for identifying genetic susceptibility factors for diabetic nephropathy in type 1 diabetes. *J Am Soc Nephrol* 17:1782-1790.

Olson JM (1997). Likelihood-based models for genetic linkage analysis using affected sib pairs. *Hum Hered* 47:110-120.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D (2006). Principal components analysis corrects for stratification in genomewide association studies. *Nat Genet* 38:904-909.

Risch, N (1990). Linkage strategies for genetically complex traits. II. The power of affected relative pairs. *Am J Hum Genet* 46:229-241.

Risch N, Merikangas K (1996) The future of genetic studies and complex human diseases. *Science* 273:1516-1517.

Scott LJ et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316:1341-1345.

Skol AD, Scott LJ, Abecasis GR, Boehnke M (2006). Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet*

Spielman RS, McGinnis RE, Ewens WJ (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506-516.

Spielman RS, Ewens WJ (1996). The TDT and other family-based tests for linkage disequilibrium and association [editorial]. *Am J Hum Genet* 59:983-989.

Wellcome Trust Case Control Consortium (WTCCC) (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661-678.

Zeggini E et al. (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1336-1341.

Zeggini E et al. (2008). Meta-analysis of genome-wide association data and large-scale replication identified additional susceptibility loci for type 2 diabetes. *Nat Genet* 40:638-645.

Zhu X, Olson JM, Schnell AH, Elston RC (1997). Model-free age of onset methods applied to the linkage of bipolar disorder. *Genet Epidemiol* 14:711-716.

FIND "FAMILY STUDY" ANALYSIS PLAN (5/02/03)

1. Genome screen* for diabetic nephropathy in all families
2. Genome screen for diabetes mellitus in all families
3. Genome screen for urine ACR and protein to creatinine in all families
4. Genome scan for serum creatinine concentration and calculated GFR in all families
5. Genome screen for hemoglobin A1C in all families
6. Genome screen for diabetic retinopathy
7. Analyses 1-6 will be performed both separately in European American, African American, Native American, and Hispanic American family sets, and with ethnic group as a covariate.
8. Analyses 1-6 will be performed using the MALD approach separately in Mexican Americans and African Americans.
9. Diabetic retinopathy and nephropathy associations (calculated GFR, ACR, urine protein:creatinine ratio) and with level of glycemic control
10. Race and gender differences in familial clustering of diabetic nephropathy
11. The following candidate gene and candidate region analyses will be performed in the entire family set and within each racial group

Name	Symbol	Chromosome	PIC
Angiotensin Converting Enzyme	ACE ACE (I/D)	17q23	CWRU, HUCLA, JHU, UCLA, UNM
Advanced Glycosylation End Product-Specific Receptor	AGER	6p21.3	HUCLA
Angiotensin	AGT	1q42-q43	CWRU, HUCLA, JHU, UCLA
Angiotensin Receptor 1, 2	AGTR1 AGTRII	3q21-q25 Xq22-q23	CWRU, HUCLA, JHU, UCLA
Aldose Keto-Reductase Family 1	ALDR1	7q35	HUCLA UNM
Apolipoprotein A-II Apolipoprotein E	AII APOE	1q21-q23 19q13.2	HUCLA, JHU UNM
Arachidonate 12-Oxidoreductase	ALOX12	17p13.1	HUCLA
Properdin Factor B	Bf	6p21.3	JHU
Complement Component 3, 4	C3 C4	19p13.3-p13.2 6P21.3	JHU
Connective Tissue Growth Factor	CTGF	6q23.1	HUCLA
Endothelin 2	EDN2	1q34	WFU
Epidermal Growth Factor	EGF	4q25	CWRU, JHU
Epidermal Growth Factor Receptor	EGFR	7p12.3-p12.1	CWRU
Fibronectin	FN1	2q34	HUCLA
G-protein $\beta 3$	GNB3	12p13	CWRU
Hepatocyte Nuclear Factor-1 β	HNF1B/ TCF2	17cen-q21.3	HUCLA, JHU
Major Histocompatibility Complex	HLA	6p21.3	JHU

Name	Symbol	Chromosome	PIC
Insulin	INS	11p15.5	JHU
Insulin Like Growth Factor 1	IGF1	12q22-q24.1	HUCLA
Interferon Gamma	IFNG	12q14	CWRU
Interleukin 1 Receptor Antagonist	IL1RN	2q14.2	HUCLA, JHU, WFU
Kallikrein 1,3	KLK1 KLK3	19q13 19q13	CWRU, JHU, WFU
Kinin B1 Receptor	BKR2	14q32.1-q32.2	JHU
Matrix Metalloproteinase 2, 3, 9	MMP2 MMP3 MMP9	16q13 11q23 20q11.2-q13.1	JHU, WFU
Methylenetetrahydrofolate Reductase	MTHFR	1p36.3	UNM
Solute Carrier Family 9, Isoform 5	NHE5 SLC9A5	16q22.1	HUCLA, JHU, WFU
Nitric Oxide Synthase 1, 2, 3	NOS1 NOS2 NOS3	12q24.2-q24.31 17cen-q11.2 7q36	HUCLA, UNM, WFU
Plasminogen Activator Inhibitor 1	PAI1	17q21.3-q22	HUCLA
Platelet Derived Growth Factor	PDGFA PDGFB	7p22 22q12.3-q13.1	HUCLA, JHU
Platelet Derived Growth Factor Receptor	PDGFR	5q31-q32	CWRU
Protein Kinase C β 1	PRKCB1	16p11.2	HUCLA
Renin	REN	1q32	CWRU, HUCLA, JHU
Selectin Cluster Genes	SELE SELL SELP	1Q24-25	UNM
T-Cell Antigen Receptor Beta	TCRBC	7q35	UNM
Transforming Growth Factor β 1, β 2, β 3	TGFB1 TGFB2 TGFB3	19q13.1 1q41 14q24	CWRU, HUCLA, JHU, UNM
Transforming Growth Factor β Receptor 1, 2, 3	TGFBR1 TGFBR2 TGFBR3	9q33-q34 3p22 1p32-p33	CWRU
Tumor Necrosis Factor	TNF	6p21.3	JHU
Tumor Necrosis Factor Receptor 1, 2	TNFR1 TNFR2	12p13.2 1p36.3-p36.2	CWRU
Type IV Collagen α -1 Chain	COL6A1	21q22.3	JHU
Vascular Endothelial Growth Factor	VEGF	6p12	HUCLA, UNM
Krolewski Region		Chromosome 3	CWRU, HUCLA
Pima Genome Scan		Chromosome 7 7q35	CWRU, HUCLA, Phoenix, UNM
Pima Genome Scan		Chromosome 9	HUCLA
Syntetic to Rodent Gene <i>Rf</i> - 1	<i>Rf</i> - 1	Chromosome 10 10q24-q26	CWRU, WFU

Name	Symbol	Chromosome	PIC
Synthetic to Mouse Os		16q13-q22	CWRU
Pima Genome Scan		Chromosome 18	Phoenix
Syntetic to Rodent Gene <i>Rf - 2</i>	<i>Rf - 2</i>	Chromosome 19 19q13.1	
Pima Genome Scan		Chromosome 20	CWRU

*Genome screens will control for various intermediate phenotypes (diabetes duration, age at diabetes onset, medications (particularly ACE inhibitor and angiotensin receptor blocker use), and additional covariates. Where age is involved, analyses will also be performed on "susceptibility" allowing for age, and for time to disease, as implemented in SIBPAL. Analyses of quantitative traits will be done using the most powerful valid method available in SIBPAL. Analysis of dichotomous traits will be performed using SIBPAL and LODPAL.

PIC	PHASE 1 CAND. GENE
CWRU	ITGA8, Rrm2b, PON2
WFU	KLKB1, MMP9
HUCLA/LGD-NCI	ACE, P-Cadherin, ACTN4, Podocin
UNM	HMGCR, CTGF, TGFBI
UTHSCSA	ZO1, SMAD, Gremlin
JHU	VEGF, SLC2A1
UCLA	CRP
Phoenix-NIDDK	NOS3, AKR1B1

APPENDIX D
APPENDIX D: FIND MEDICAL QUESTIONNAIRE

Barcode and
Participant ID

--

10. **Sex:** Male Female

11. **Ethnicity:** African American American Indian; tribal identifier ____
European American Mexican American Other Hispanic Other _____

12. **Country of birth:** _____ **City of Birth:** _____

13. **Name of your primary doctor:**

Telephone: _____ - _____ - _____ **OR if outside the U.S.**
Country code _____ Tel _____

Location: _____
Street *Suite* *City*

State/Province *Postal code* *Country*

14. **Name of your dialysis or transplant unit:**

Telephone: _____ - _____ - _____ **OR if outside the U.S.**
Country code _____ Tel _____

Location: _____
Street *Suite* *City*

State/Province *Postal code* *Country*

15. **Name of your nephrologist:**

Telephone: _____ - _____ - _____ **OR if outside the U.S.**
Country code _____ Tel _____

Location: _____
Street *Suite* *City*

State/Province *Postal code* *Country*

16. **Name of your eye doctor:**

Telephone: _____ - _____ - _____ **OR if outside the U.S.**
Country code _____ Tel _____

Location: _____
Street *Suite* *City*

State/Province *Postal code* *Country*

Barcode and
Participant ID

17. Are you currently participating in another study?

No Yes

If yes, please specify: _____

B. HEIGHT AND WEIGHT

18. Current height and weight: (ft) ___ (in) ___ (cm) ___ (weight lbs) ___ (kilos) ___

19. Maximum weight or dry weight (lbs.): ___ (kilos) ___

20. Age and height at maximum weight: (years) ___ (ft) ___ (in) ___ (cm) ___

21. Have you ever had a major leg amputation (*above the ankle*)?

No Yes

If **Yes**, specify locations (*Mark all that apply*): Left below the knee Right below the knee
Left above the knee Right above the knee

C. KIDNEY DISEASE

22. Do you have high blood pressure?

No Yes Don't know

If **Yes**, when was your high blood pressure diagnosed (year)? _____

23. Have you ever had a kidney biopsy?

No Yes Don't know

If **Yes**, when did you have the kidney biopsy (year)? _____

Where was the kidney biopsy performed? _____

24. Do you have kidney failure requiring dialysis or transplant?

No Yes

If **Yes**, indicate the date of onset of dialysis treatment (MM/DD/YYYY): ___ | ___ | _____

If **No**, skip to section D.

25. Have you ever had an organ transplant?

No Yes Please Specify _____

26. What caused your kidney failure/renal insufficiency (*Mark all that apply*)?

- | | |
|--|--|
| <input type="checkbox"/> Diabetes | <input type="checkbox"/> Polycystic kidney disease |
| <input type="checkbox"/> Hypertension | <input type="checkbox"/> Lupus nephritis |
| <input type="checkbox"/> IgA nephropathy | <input type="checkbox"/> Kidney cancer |
| <input type="checkbox"/> Membranous glomerulonephritis | <input type="checkbox"/> Obstruction |
| <input type="checkbox"/> Focal glomerulosclerosis | <input type="checkbox"/> Don't know |
| <input type="checkbox"/> Other _____ | <input type="checkbox"/> None |

D. HISTORY OF DIABETES

27. Has a doctor or other health care provider told you that you had diabetes?

No Yes Don't know

If **Yes**, at what age were you diagnosed? ___ OR in what year were you diagnosed? ___

Barcode and
Participant ID

28. **Have you ever been treated for diabetic ketoacidosis or coma?**

- No Yes Don't know

29. **Have you ever taken insulin shots to treat your diabetes?**

- No Yes Don't know

If **Yes**, how soon after diagnosis did you first begin taking insulin?

- Within 1 year At least 1 year after diagnosis

Once you started taking insulin shots, did you ever go at least one month without insulin treatment?

- No Yes Don't know

E. DIABETIC EYE DISEASE

30. **Have you ever had laser treatment of your retina(s) done for diabetes?**

- No Yes Don't know

31. **Have you ever been treated for bleeding inside your eye that was not due to trauma?**

- No Yes Don't know

32. **Have you ever seen an ophthalmologist or optometrist (eye doctor)?**

- No Yes Don't know

Continued on Next Page

**Family Investigation of Nephropathy and Diabetes (FIND)
Family Pedigree Form**

Barcode and Participant ID	
-----------------------------------	--

Date: ____/____/____

Screened by: _____

Name: Last, First, MI	
Address:	
Telephone number:	()-
Clinic/Dialysis Unit Contact physician City/State	

Language preference (please circle): **1) English** **(2) Spanish** **(3)Other (specify)**_____

1. Date of Birth: _____

2. Gender: Male
 Female

3. Ethnicity: African-American
Mark all that Apply European-American
 Mexican-American
 Other Hispanic (please, specify)_____
 Native American
 Other (please, specify)_____

4. Do you have any living, blood related relatives with diabetes (siblings/parents)?

No, I am not aware of relatives having diabetes.

I do not know, I am not sure.

Yes, I have relatives with diabetes.

If YES obtain consent to get family information and continue with Table #1

Barcode and
Participant ID

--

**Table 1 – Grandparent Pedigree
Screening for the FIND Study**

	Maternal Grandmother	Maternal Grandfather	Paternal Grandmother	Paternal Grandfather
Last Name				
First Name				
Middle Initial				
Birth Date				
Ethnicity Check all that apply	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____

**Table 2: Family Members Screening for the FIND Study
MALD Screening for Spouse and Child for the FIND Study**

**Barcode and
Participant ID**

	Proband	Mother or Spouse	Father or Child	Sib/1	Sib/2
Ethnicity		<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other
Barcode ID					
Name/Address/Phone # or note that proband intends to contact this person		<input type="checkbox"/> proband will contact	<input type="checkbox"/> proband will contact	<input type="checkbox"/> proband will contact	<input type="checkbox"/> proband will contact
Current Age					
Gender		FEMALE	MALE		
Diabetes (age at onset or yr of diagnosis) X if not diabetic or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Kidney disease (age at onset or yr of diagnosis) X if no KD or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Hypertension (age at onset or yr of diagnosis) X if no HTN or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Eye disease (age at onset or yr of diagnosis) X if no ED or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Does proband think this person will participate ?					
Eligible for FIND (Y or N) if Y give study ID					

Table #2: Extension Page Family Members Screening for the FIND Study

Barcode and Participant ID	
----------------------------	--

	Sib/3	Sib/4	Sib/5	Sib/6
Ethnicity	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____	<input type="checkbox"/> African-American <input type="checkbox"/> European-American <input type="checkbox"/> Mexican-American <input type="checkbox"/> Other Hispanic _____ <input type="checkbox"/> Native American <input type="checkbox"/> Other _____
Barcode ID				
Name/Address/Phone # or note that proband intends to contact this person	<input type="checkbox"/> proband will contact	<input type="checkbox"/> proband will contact	<input type="checkbox"/> proband will contact	<input type="checkbox"/> proband will contact
Current Age				
Gender				
Diabetes (age at onset or yr of diagnosis) X if not diabetic or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Kidney disease (age at onset or yr of diagnosis) X if no KD or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Hypertension (age at onset or yr of diagnosis) X if no HTN or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Eye disease (age at onset or yr of diagnosis) X if no ED or ? if unknown	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____	Age _____ or Year _____
Does proband think this person will participate?				
Eligible for FIND (Y or N) if Y give study ID				

List of Primary Causes of End Stage Renal Disease

FIND Accepted Diabetic Reasons

<u>ICD-9</u>	<u>LTR</u>	<u>Narrative</u>
25000	A	Type II, adult-onset type or unspecified type diabetes
25001	A	Type I, juvenile type, ketosis prone diabetes

FIND Accepted Non-Diabetic Reasons

<u>ICD-9</u>	<u>LTR</u>	<u>Narrative</u>
5829	A	Glomerulonephritis (GN)
5821	A	Focal glomerulosclerosis, focal sclerosing GN
5831	A	Membranous nephropathy
5832	A	Membranoproliferative GN type1, diffuse MPGN
58381	B	IgA nephropathy, Berger's disease (proven by immunofluorescence)
5804	B	Rapidly progressive GN
5834	C	Goodpasture's Syndrome
7100	E	Lupus Erythematosus, (SLE nephritis)
2870	A	Henoch-Schonlein syndrome
7101	B	Scleroderma
4464	B	Wegener's granulomatosis
4462	A	Vasculitis and its datives
4039	D	Renal disease due to hypertension (no primary renal disease) <i>must prove hypertension before diabetes, else PID considered diabetic</i>
5832	C	Dense deposit disease, MPGN type 2
58381	C	IgM nephropathy, (proven by immunofluorescence)
5800	A	Post infectious GN, SBE
5820	A	Other proliferative
2831	A	Hemolytic uremic syndrome
4460	C	Polyarteritis
5839	C	Nephropathy due to heroin abuse and related drugs
5839	B	Secondary GN, other
9659	A	Analgesic abuse
9849	A	Lead nephropathy
5909	A	Nephropathy caused by other agents
27410	A	Gouty nephropathy
58389	B	Chronic interstitial nephritis
58089	A	Acute interstitial nephritis
4401	A	Renal artery stenosis
59381	B	Renal artery occlusion
59381	E	Cholesterol emboli, renal emboli
7598	B	Hereditary/familial nephropathy
28260	A	Sickle cell disease/anemia
28269	A	Sickle cell trait and other sickle cell (Hbs/Hb other)
0429	A	AIDS nephropathy
7999	A	Etiology uncertain

FIND Unacceptable Causes of End Stage Renal Disease

<u>ICD-9</u>	<u>LTR</u>	<u>Narrative</u>
5830	B	Radiation nephritis
5920	C	Nephrolithiasis
5996	A	Acquired obstructive uropathy
5900	C	Chronic pyelonephritis, reflux nephropathy
5929	B	Urolithiasis
2754	A	Nephrocalcinosis
75313	A	Polycystic kidneys, adult type (dominant)
75314	A	Polycystic, infantile (recessive)
75316	A	Medullary cystic disease, including nephronophthisis
7595	A	Tuberous sclerosis
7598	A	Hereditary nephritis, Alport's syndrome
2700	A	Cystinosis
2718	B	Primary oxalosis
2727	A	Fabry's disease
7533	A	Congenital nephritic syndrome
5839	D	Drash syndrome, mesangial sclerosis
7532	A	Congenital obstructive uropathy
7530	B	Renal hypoplasia, dysplasia, oligonephronia
1890	B	Renal tumor (malignant)
1899	A	Urinary tract tumor (malignant)
2230	A	Renal tumor (benign)
2239	A	Urinary tract tumor (benign)
2395	A	Renal tumor (unspecified)
2395	B	Urinary tract tumor (unspecified)
20280	A	Lymphoma of kidneys
2030	A	Multiple myeloma
2030	B	Light chain nephropathy
2773	A	Amyloidosis
99680	A	Complication post bone marrow or other transplant
64620	A	Post partum renal failure
8660	A	traumatic or surgical loss of kidney(s)
5724	A	Hepatorenal syndrome
5836	A	Tubular necrosis (no recovery)
59389	A	Other renal disorders

APPENDIX E: FIND MEDICAL RECORD REVIEW

FIND Medical Record Review

Barcode and
Participant ID

Study Coordinator ID _____

Date of review (MM/DD/YYYY) ____ | ____ | ____ _

Instruction to record abstractor: During review of the medical record(s), please review the Medical Questionnaire (Form 01), since it may be necessary to add or correct some data on the questionnaire (e.g., medicines or date of onset of diabetes or dialysis).

RECORD TYPE

1. Indicate type of record(s) reviewed (*Mark all that apply*).

- Dialysis unit
- Primary care (*Including records from the endocrinologist and/or nephrologist*)
- Hospital
- Other _____

PARTICIPANT INFORMATION

2. **Name:** _____
Last First Middle Maiden

3. **Birth date** (MM/DD/YYYY): ____ | ____ | ____ _

4. **Sex:** Male Female

Barcode and
Participant ID

--

KIDNEY DISEASE

5. Is the level of urinary protein/albumin recorded in the record?

No Yes

If **Yes**, then specify the following (*Mark the maximum value only*):

	<u>Protein excretion rate</u>		<u>P/C ratio</u>		<u>Albumin excretion rate</u>		<u>A/C ratio</u>
a. <input type="checkbox"/>	<50 mg/24h	or	<0.15 mg/mg	or	<30 mg/24h	or	<0.03 mg/mg
b. <input type="checkbox"/>	≥50 mg/24h	or	≥0.15 mg/mg	or	≥30 mg/24h	or	≥0.03 mg/mg
c. <input type="checkbox"/>	≥500 mg/24h	or	≥0.5 mg/mg	or	≥300 mg/24h	or	≥0.3 mg/mg
d. <input type="checkbox"/>	≥1.0 g/24h	or	≥1.0 mg/mg	or	≥1.0 g/24h	or	≥1.0 mg/mg
e. <input type="checkbox"/>	≥3.0 g/24h	or	≥3.0 mg/mg	or	≥3.0 g/24h	or	≥3.0 mg/mg
f. <input type="checkbox"/>	Nephrotic (3.0g – 3.5g)						
g. <input type="checkbox"/>	None						

If **a**, record date of **LAST** value (*MM/DD/YYYY*): ___ | ___ | ___

Also record whether subject is receiving antihypertensive therapy at **LAST** value: No Yes
HIGHEST value: No Yes

If **e**, record date of **FIRST** value at this level (*MM/DD/YYYY*): ___ | ___ | ___

Equivalent Measures of Urinary Albumin Excretion

30 mg/24h	=	20 µg/min
300 mg/24h	=	200 µg/min

If the excretion reported above is a ratio, record the urine protein (albumin) and creatinine concentrations:

Urine protein (*mg/l*) ___ . ___ or Urine albumin (*mg/l*) ___ . ___

Urine creatinine (*g/l*) ___ . ___

6. Is the participant receiving chronic renal replacement therapy?

No Yes If **Yes**, record the date of onset (*MM/DD/YYYY*): ___ | ___ | ___
 Yes AA MALD with Nephropathy but no DM

If **No**, record the **HIGHEST** serum creatinine concentration (*mg/dl*): ___ . ___

Date of **HIGHEST** serum creatinine concentration (*MM/DD/YYYY*): ___ | ___ | ___

7. Primary cause of kidney failure / renal insufficiency (*Mark all that apply*)?

- | | |
|--|--|
| <input type="checkbox"/> Diabetes | <input type="checkbox"/> Polycystic kidney disease |
| <input type="checkbox"/> Hypertension | <input type="checkbox"/> Lupus nephritis |
| <input type="checkbox"/> IgA nephropathy | <input type="checkbox"/> Kidney cancer |
| <input type="checkbox"/> Membranous glomerulonephritis | <input type="checkbox"/> Obstruction |
| <input type="checkbox"/> Focal glomerulosclerosis | <input type="checkbox"/> Don't know |
| <input type="checkbox"/> Other _____ | <input type="checkbox"/> None |

Barcode and Participant ID	
-----------------------------------	--

8. Is a kidney biopsy recorded in the record?

No Yes If **Yes**, record the biopsy date (MM/DD/YYYY): ___ | ___ | ___ _ _ _

If **Yes**, specify methods of evaluation (*Mark all that apply*).

Light microscopy Electron microscopy Immunofluorescence

Also specify the histologic findings (*Mark all that apply*).

- Increased nodular mesangial matrix.
- Increased diffuse mesangial matrix.
- Thickened glomerular basement membrane.
- Arterial hyalinization.
- Arteriolar hyalinization.
- Mesangial immunoglobulin or paraprotein deposits by immunofluorescence.
- Amyloid deposits by Congo red staining or electron microscopy.
- Electron dense deposits within the glomerular basement membrane or glomerular capillary subendothelial space.
- Non-Diabetic Pahtological Diagnosis

HISTORY OF DIABETES

9. Has a diagnosis of diabetes been made?

No Yes

If Yes, record the source of the diagnosis (*Mark all that apply*):

- Fasting plasma glucose \geq 126 mg/dl (or venous whole blood glucose \geq 110 mg/dl or capillary whole blood glucose \geq 110 mg/dl) *
- Random plasma glucose \geq 200 mg/dl (or venous whole blood glucose \geq 180 mg/dl or capillary whole blood glucose \geq 200 mg/dl) * and symptoms (*polyuria, polydipsia, polyphagia*).
- Two-hour post-load plasma glucose \geq 200 mg/dl (or venous whole blood glucose \geq 180 mg/dl or capillary whole blood glucose \geq 200 mg/dl) * (*OGTT*).
- Clinical diagnosis without documented plasma glucose concentration.

Record the earliest date of diagnosis (MM/DD/YYYY): ___ | ___ | ___ _ _ _

*The measurement of glucose in serum is discouraged by the WHO (World Health Organization). Unless the red cells are immediately removed to prevent glycolysis, serum samples should not be used for diagnosing diabetes. If only serum glucose values are found in the record, however, they should be interpreted as if they were plasma values.

10. Is a diagnosis of diabetic ketoacidosis or diabetic coma recorded in the record?

No Yes

Barcode and Participant ID	
-------------------------------	--

DIABETIC EYE DISEASE

11. Is a diagnosis of diabetic retinopathy recorded in the record?

No Yes

If **Yes**, specify the severity of the retinopathy (*Mark all that apply*).

Background retinopathy

Pre-proliferative retinopathy

Proliferative retinopathy

Vitreous hemorrhage

Macular edema

Photocoagulation therapy

APPENDIX F: SUBCOMMITTEE MEMBERSHIP

Consent & Ethics Subcommittee

Chair

Phil Zager- University of New Mexico

Nedal Arar – University of Texas Health Sciences Center at San Antonio

Nancy Fink – Johns Hopkins University

Sudha Iyengar – Case Western Reserve University

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Diseases

Madeline Pahl – University of California Irvine

Rosemarie Plaetke – University of Texas Health Sciences Center at San Antonio

Ashwini Sehgal – Case Western Reserve University

Vallabh Shah – University of New Mexico

Data Analysis Subcommittee

Chair

Robert C. Elston – Case Western Reserve University

Joe Coresh – Johns Hopkins University

Robert Hanson – National Institute of Diabetes and Digestive and Kidney Diseases Phoenix

Gyungah Jun – Case Western Reserve University

Paul Kimmel - DEA- National Institute of Diabetes and Digestive and Kidney Diseases

Hongzhe Li – Harbor-Torrance/University of California Los Angeles

Jean W. MacCluer – Southwest Foundation of Biomedical Research at San Antonio

Jane Olson – Case Western Reserve University

Rosemarie Plaetke – University of Texas Health Sciences Center at San Antonio

Steven Rich – Wake Forest University

Jerry Rotter – University of California Los Angeles

Christine Stidley – University of New Mexico

Enrollment & Recruitment Subcommittee

Chair

Jeff Schelling – Case Western Reserve University

Paul Kimmel - DEA- National Institute of Diabetes and Digestive and Kidney Diseases

Cell Line Immortalization and Transformation Subcommittee

Chair

Mike Smith – National Cancer Institute

Paul Kimmel - DEA- National Institute of Diabetes and Digestive and Kidney Diseases

Informatics Subcommittee

Chair

Carol Lorimor – Case Western Reserve University

Joe Coresh – Johns Hopkins University

Sarah Lalacci – Case Western Reserve University

Kevin Jacobs – Consultant

Gyungah Jun – Case Western Reserve University

Paul Kimmel - DEA- National Institute of Diabetes and Digestive and Kidney Diseases

Intellectual Properties & Archives Subcommittee

Chair

Jerry Rotter – University of California Los Angeles

Arlene Bobelu – University of New Mexico

Sudha Iyengar – Case Western Reserve University

Paul Kimmel - DEA- National Institute of Diabetes and Digestive and Kidney Diseases

Mike Seldin – Harbor Torrance / University of California Los Angeles

Vallbh Shah – University of New Mexico

Protocol Subcommittee

Chair

John Sedor – Case Western Reserve University

Hanna Abboud – University of Texas Health Sciences Center at San Antonio

Sharon Adler – Harbor-Torrance University of California Los Angeles

Barry Freedman – Wake Forest University

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Diseases

Bill Knowler – National Institute of Diabetes and Digestive and Kidney Diseases – Phoenix

Jean W. MacCluer – Southwest Foundation of Biomedical Research at San Antonio

Robert Nelson – National Institute of Diabetes and Digestive and Kidney Diseases – Phoenix

Jane Olson – Case Western Reserve University

Rosemarie Plaetke – University of Texas Health Sciences Center at San Antonio

Mohammed Saad – University of California Los Angeles

Vallabh Shah – University of New Mexico

Jeff Schelling – Case Western Reserve University

Phil Zager – University of New Mexico

Publications Subcommittee

Chair

Bill Knowler – National Institute of Diabetes and Digestive and Kidney Diseases - Phoenix

Sharon Adler – Harbor-Torrance University of California Los Angeles

Barry Freedman – Wake Forest University

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Disease

Mike Klag – Johns Hopkins University

Jane Olson – Case Western Reserve University

Mohammed Saad – University of California Los Angeles

Jeff Schelling – Case Western Reserve University

Mike Seldin – Harbor Torrance / University of California Los Angeles

Phil Zager – University of New Mexico

Quality Control and Assurance Subcommittee– Clinical

Chair

Mike Klag – Johns Hopkins University

Eli Ipp – Harbor-Torrance / University of California Los Angeles

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Diseases

Lucy Mead – Johns Hopkins University

Ashwini Sehgal – Case Western Reserve University

Quality Control and Assurance Subcommittee – Molecular Genetics

Chair

Sudha Iyengar – Case Western Reserve University

Don Bowden – Wake Forest University

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Diseases

Mike Seldin – Harbor Torrance / University of California Los Angeles

Vallabh Shah – University of New Mexico

Kent Taylor – University of California Los Angeles

Subproject Evaluation Subcommittee

Chair

Robert Hanson – National Institute of Diabetes and Digestive and Kidney Diseases-Phoenix

Hanna Abboud – University of Texas Health Sciences Center at San Antonio

Sharon Adler – Harbor-Torrance University of California Los Angeles

Katrina Goddard – Case Western Reserve University

Sudha Iyengar – Case Western Reserve University

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Diseases

John Sedor – Case Western Reserve University

Vallabh Shah – University of New Mexico

APPENDIX G: STEERING COMMITTEE

Chair

Barry Freedman – Wake Forest University

Hanna Abboud – University of Texas Health Sciences Center at San Antonio

Sharon Adler – Harbor-Torrance University of California Los Angeles

Robert C. Elston – Case Western Reserve University

Paul Kimmel - DEA-National Institute of Diabetes and Digestive and Kidney Diseases

Mike Klag – Johns Hopkins University

Bill Knowler – National Institute of Diabetes and Digestive and Kidney Diseases – Phoenix

Mohammed Saad – University of California Los Angeles

John Sedor – Case Western Reserve University

Phil Zager – University of New Mexico

APPENDIX H: LABORATORY FORMS

FIND Approved Lab Materials

Fisher Scientific

1-800-766-7000 (Customer Service)

3970 Johns Creek Court

Suite #500

Suwanee, GA 30024

Description	Catalog Number
Yellow Vacutainer Blood Bank Tube 10 ml (ACD Tube)	02-684-26
Purple Vacutainer Hematology Tube 10 ml (EDTA Tube)	02-683-84
Red-Grey Glass Serum Gel Separation Tube 10 ml (Tiger-Top)	02-657-11
Grey-Glass Plasma with Glycolytic Inhibitor 5 ml (Grey-Top)	02-687-70
Nunc Cryovial 4.5 ml	12-565-173N
60 ml Sterile Cup Screw Cap Urine Container	13-711-64

CargoPak

919-878-9933 (Contact: Rob Smith)

3215 Wellington Court – Suite A

Raleigh, NC 27615

Description	Catalog Number
CP-Combination-Diagnostic Shipper	LB-06 Interlocking Styrofoam Container
8 ounce Gelpak	Frigid – 08
Pressure Tested Tyvek Bag and Plastic Pouch*	SAF – T – PAK - 710
Adhesive One-Side Bubble Wrap 12” x 12”	SAF – T – RAP – 610
5 Slot Bubble Pouch	SAF – T – Pouch 600
100 ml Absorbent Strips	SAF – T – PAK 151

*Ziplock biohazard bag

These materials are approved for collecting and shipping specimens from the PIC to the NCI. These materials will ensure safe receipt of the specimens. All recyclable materials (CargoPak shipping containers, gelpak, bubble wrap, absorbent strips, etc) will be returned to the PIC for re-use. Substitutions will not be returned to the PIC for re-use.

FIND Approved Lab Materials (continued)

Miscellaneous FIND Materials (Office Products)

Description	Catalog Number
Avery Color Coding Dots ¼" – Red	AVE05790
Avery Color Coding Dots ¼" – Green	AVE05791

The color coding dots may be purchased through your office supply distributor.

Electronic Imaging Materials

Contact: Mike Blaudschun (603) 357-1459

180 Emerald Street

Keene, New Hampshire 03431

Fax: (603) 357-1542

Description	Catalog Number
1.75" x .5" white latex labels (80 per page) 1000/case	815175050

Barcode labels can be purchased directly from Electronic Imaging. The cost per case depends on how many cases are purchased. (~\$350 for 1 case, ~\$300 for 5 cases).

Family Investigation of Nephropathy and Diabetes Shipping Invoice

PIC: _____ Sender: _____

Date Shipped: _____ Scheduled Arrival Date: _____

Shipping Tracking No.: _____ Shipping Method: _____

Date Collected: _____

Participant ID	Participant Status	Draw #	FDP Grey	Urine	Serum Red-Grey	ACD Yellow	EDTA Purple	Total # Samples
	<input type="checkbox"/> Participant with ESRD <input type="checkbox"/> Participant Fasting <input type="checkbox"/> Participant Not Fasting		<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	
	<input type="checkbox"/> Participant with ESRD <input type="checkbox"/> Participant Fasting <input type="checkbox"/> Participant Not Fasting		<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	
	<input type="checkbox"/> Participant with ESRD <input type="checkbox"/> Participant Fasting <input type="checkbox"/> Participant Not Fasting		<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> 0 <input type="checkbox"/> 1	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	<input type="checkbox"/> 0 <input type="checkbox"/> 1 <input type="checkbox"/> 2	

Comments: _____

Please include a copy of this form with the specimen shipment to NCI. Any Shipments that cannot be entered into the Shipping Calendar must be called in to the LGD-NCI at 301-846-7055. If possible, a fax of this form should be sent also to 301-846-7139. Please call the GADCC to notify them that a shipping note was not entered 216-778-4526 or 216-778-5652. Send Shipments to: Russ Hanson, LGD-NCI Frederick Cancer Research and Development Center, Building 560 Room 21-6A, Frederick MD 21702.

Blood Processing and Aliquoting Scheme for the FIND Study

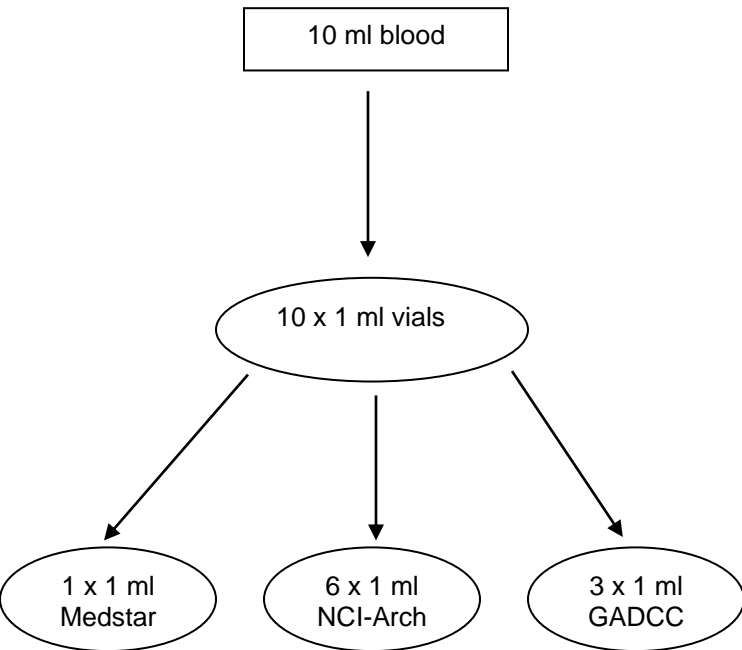
12/20/00

Tube: Purple (EDTA, first of two tubes)
 Extraction: Whole blood
 Purpose: Hemoglobin A1C

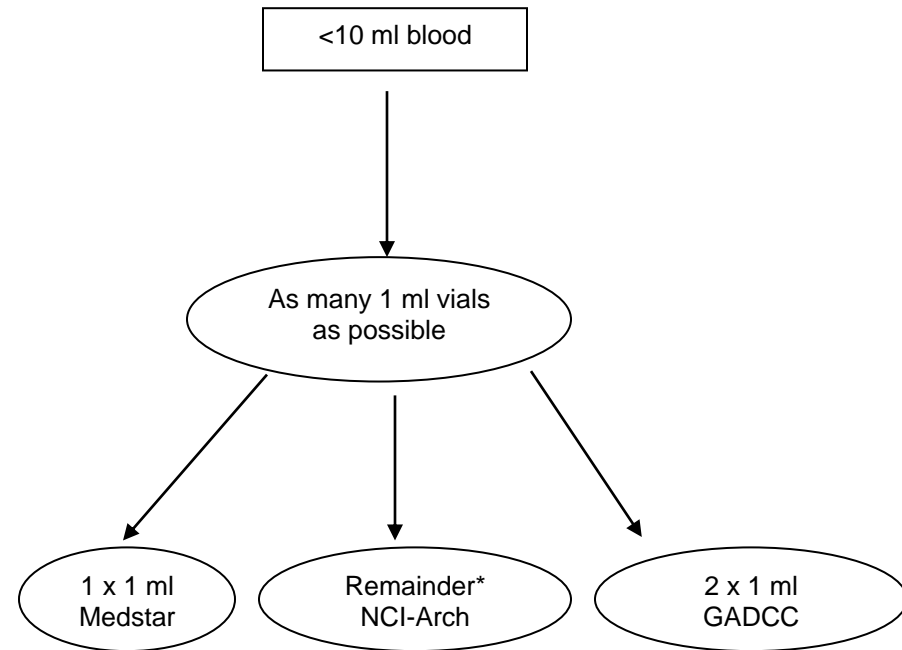
Processed by PIC

Processed by NCI-F
 (NCI-Frederick)

If the tube contains 10 ml of blood



If the tube contains < 10 ml of blood



If there is > 10 ml of blood, make as many 1 ml vials as possible. Freeze ½ of the extra ones at -70°C at NCI-Archive and send the other ½ of the extras to GADCC.

* If two vials remain after sending the other vial to MedStar, then send one vial to NCI-Archives and the other vial to GADCC. If only one vial remains after sending the other vial to MedStar, send the one remaining vial to NCI-Archive.

Blood Processing and Aliquoting Scheme for the FIND Study

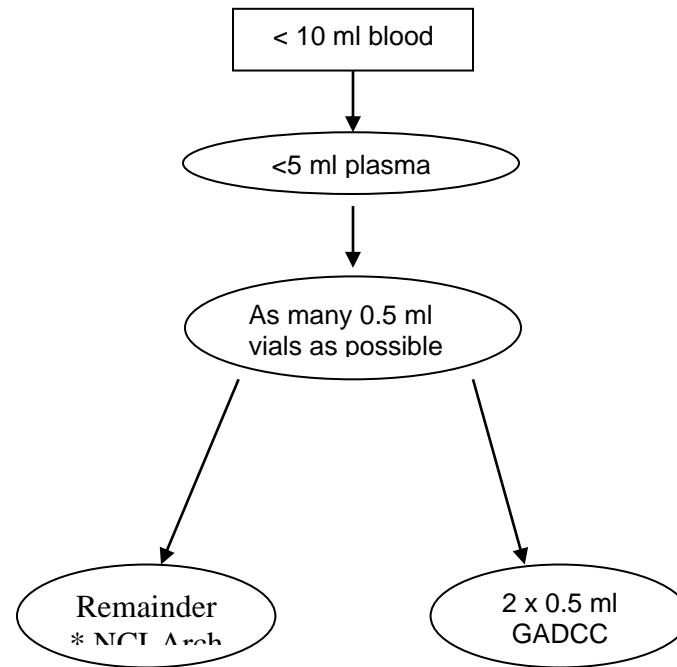
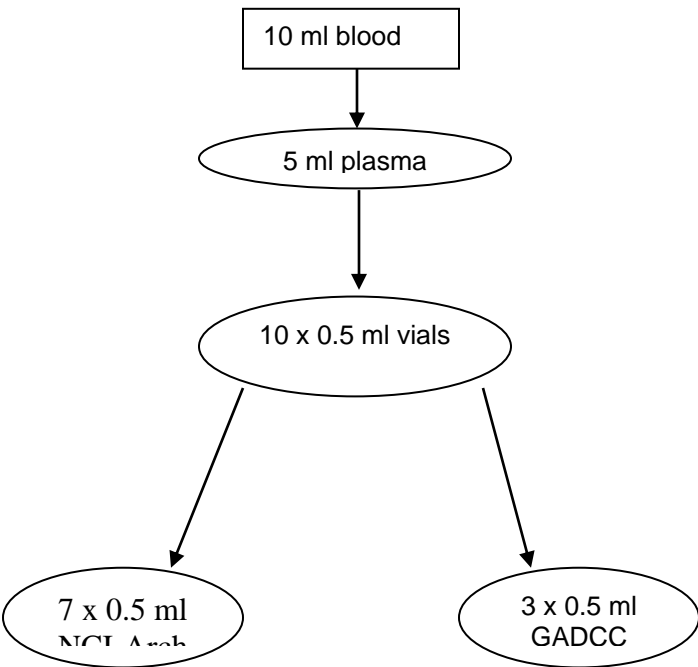
12/20/00

Tube: Purple (EDTA, second of two tubes)
Extraction: Plasma
Purpose: Storage



If the tube contains 5 ml of plasma

If the tube contains < 5 ml of plasma



If there is > 5 ml of plasma, make as many 0.5 ml vials as possible. Freeze ½ of the extra ones at -70°C at NCI Archive and send other remaining ½ of the extras to GADCC.

* If only two vials of plasma are made, then freeze one vial at -70°C at NCI-Archives and send the other vial to GADCC. If only one vial of plasma is made, send it to NCI-Archive.

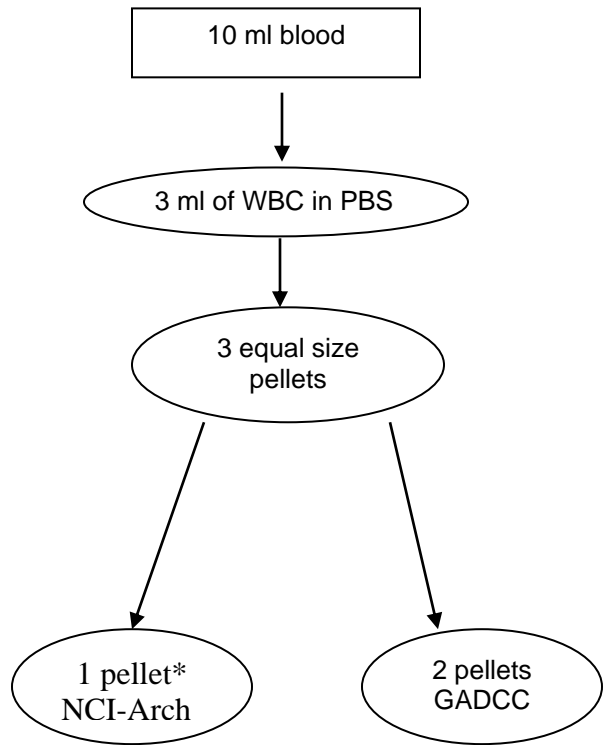
Blood Processing and Aliquoting Scheme for the FIND Study

12/20/00

Tube: Purple (EDTA, second of two tubes)
Extraction: Buffy coat
Purpose: DNA extraction and storage

Processed by PIC

Processed by NCI-F
(NCI-Frederick)



*If only one pellet can be made, send it to NCI-Archive.

Blood Processing and Aliquoting Scheme for the FIND Study

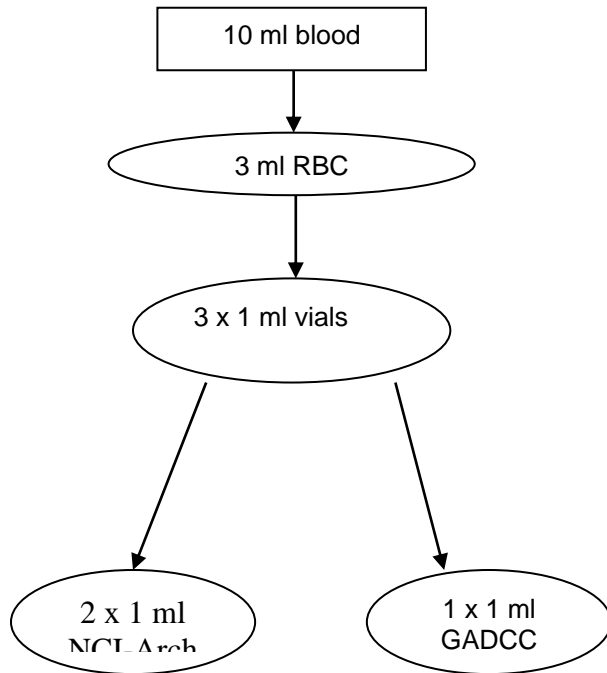
12/20/00

Tube: Purple (EDTA, second of two tubes)
Extraction: Red blood cell
Purpose: Storage

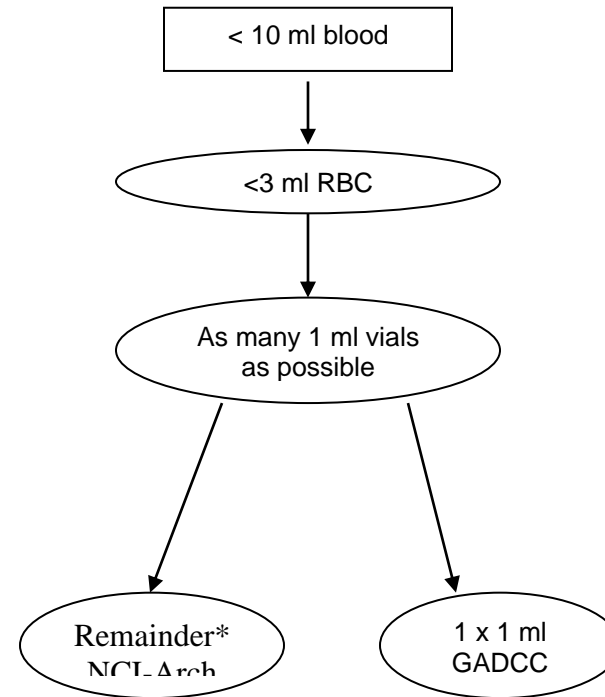
Processed by PIC

Processed by NCI-F
(NCI-Frederick)

If the tube contains 3 ml of red blood cell



If the tube contains < 3 ml of red blood cell



*If there is > 3 ml of RBC, make as many 1 ml vials as possible. Freeze ½ of the extra vials at -70°C at NCI Archive and send the remaining ½ of the extra vials to GADCC.

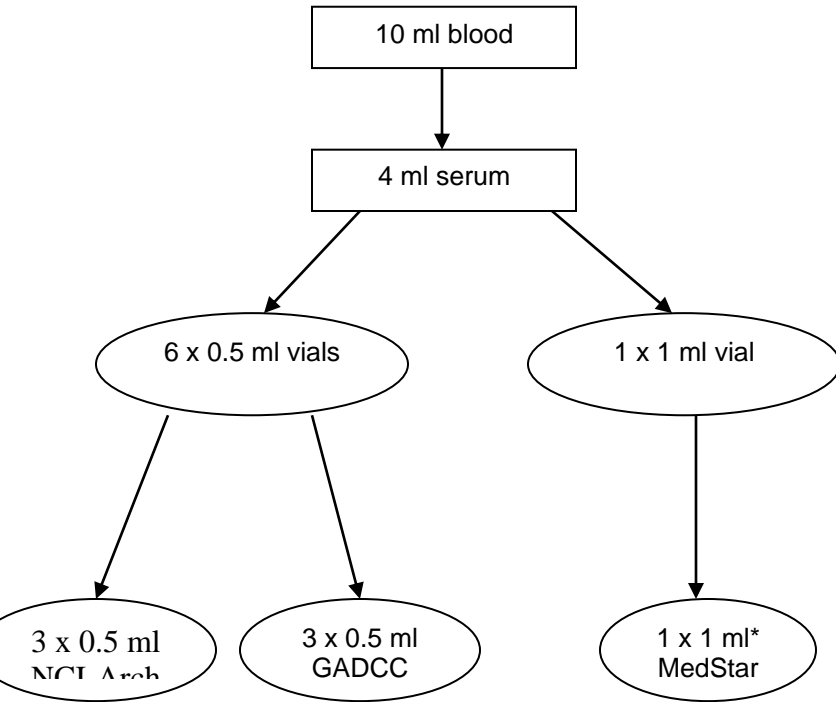
* If only one vial is made, send it to NCI-Archive.

Blood Processing and Aliquoting Scheme for the FIND Study

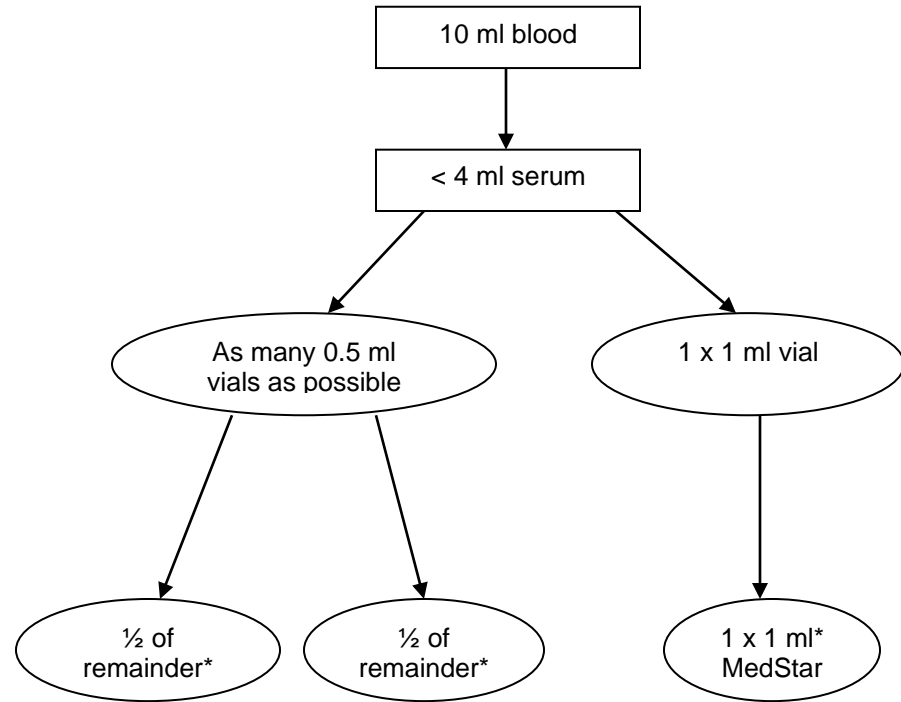
12/20/00

Tube: Red-Gray
 Extraction: Serum
 Purpose: BUN, creatinine

If the tube contains 4 ml of serum



If the tube contains < 4 ml of serum



If there is > 4 ml of serum, make as many 0.5 ml vials as possible. Freeze 1/2 of the extras at -70°C at NCI Archive and send the remaining 1/2 to GADCC.

* If only one vial is made, send it to NCI-Archive.

** BUN and creatinine will not be measured in patients with ESRD; therefore, no serum of patients with ESRD will be sent to MedStar.

Blood Processing and Aliquoting Scheme for the FIND Study

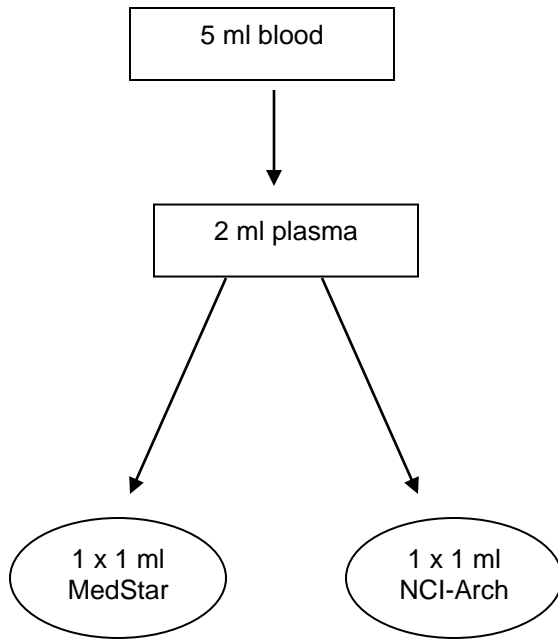
12/20/00

Tube: Gray (FDP)
Extraction: Plasma
Purpose: Fasting glucose

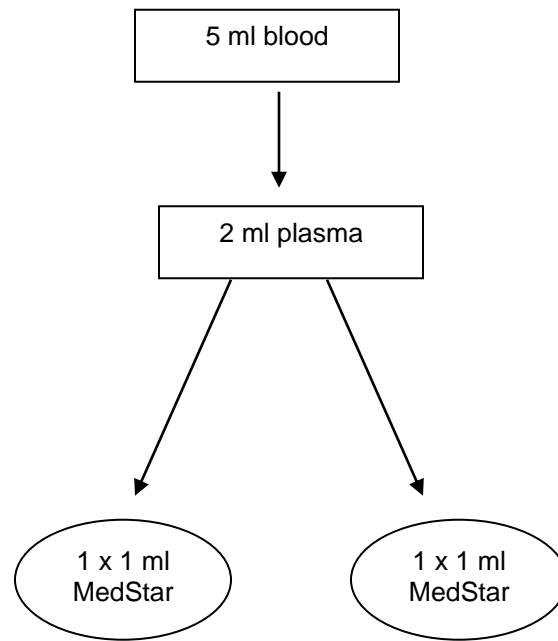
Processed by PIC

Processed by NCI-F
(NCI-Frederick)

Regular Samples



Quality Control Sample

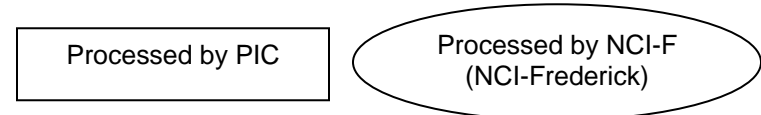
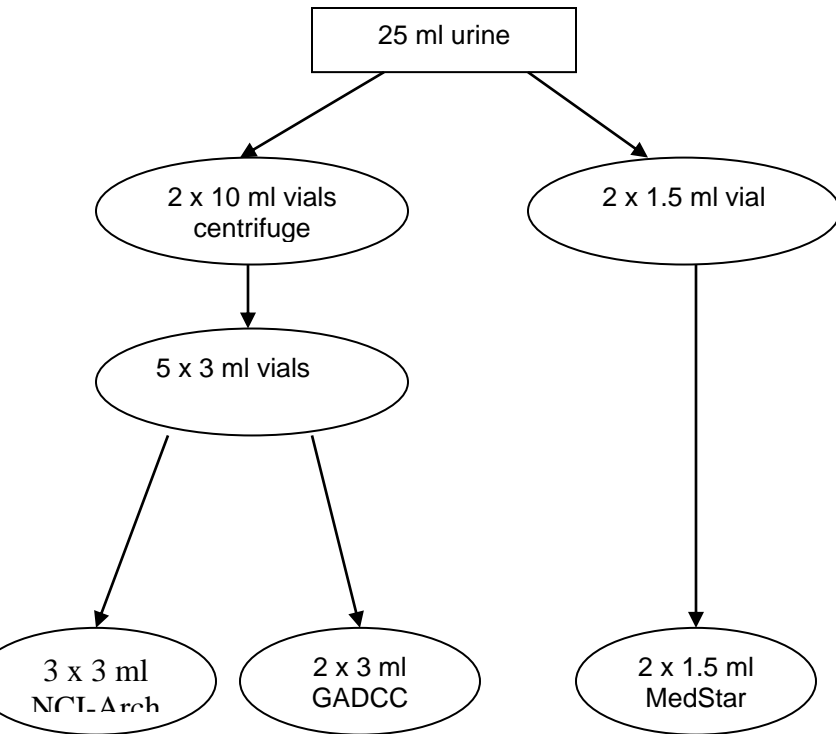


Blood Processing and Aliquoting Scheme for the FIND Study

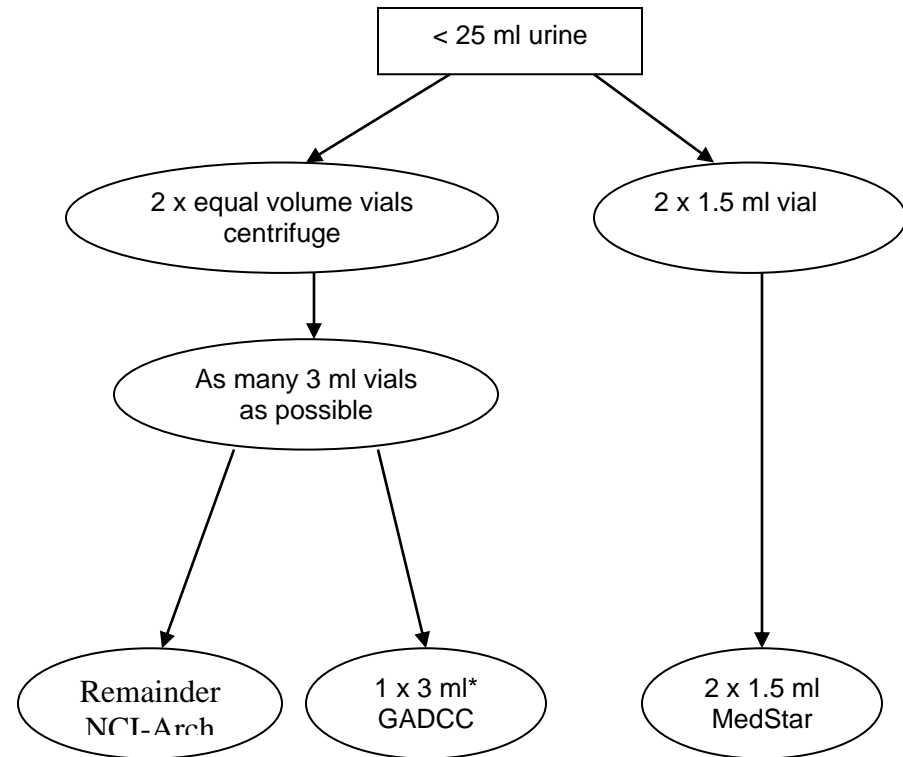
12/20/00

Tube: Urine Cup
Extraction: Urine
Purpose: Microalbumin, protein, creatinine

If the cup contains 25 ml of urine



If the cup contains < 25 ml of urine



If there is > 25 ml of urine, make as many 3 ml vials as possible. Freeze ½ of the extra vials at -70°C at NCI Archive and send the other ½ of the extra vials to GADCC.

* If only one vial of urine remains after sending the other to MedStar, send it to NCI-Archive.

Blood Sample Collection List

FIND Blood Sample Collection List

Shipping Date:

Penn Medical Laboratory
Irving St, NW Annex # 2
Washington, DC 20010

Participant ID No.(FIND):	Site ID (PIC):	Collection Date:	(PML accession labels) here	
Check below to indicate sample was collected	Sample Type and Volume	Test	FIND ID (type-specific)	STORAGE AND SHIPPING REQUIREMENTS
<input type="radio"/> 0 <input type="radio"/> 1	SERUM 1 x 1.0 ml	BUN, Creatinine		Frozen
<input type="radio"/> 0 <input type="radio"/> 1	EDTA Whole Blood 2 x 1.0 ml	HemoglobinA1c		Frozen
<input type="radio"/> 0 <input type="radio"/> 1	Fluoridated Plasma (FDP) 1 x 2.0 ml	Glucose		Frozen
<input type="radio"/> 0 <input type="radio"/> 1 <input type="radio"/> 2	Urine 2 x 1.5 ml	Microalbumin, Protein & Creatinine		Frozen
Site Comments: <input type="radio"/> Participant with ESRD <input type="radio"/> Participant Fasting <input type="radio"/> Participant Not Fasting NCI Contact Person: Mary Thompson (301) 846-5948				Laboratory Comments: Date and Time Samples Received: Processing Tech Initials: Other Comments – provide below:

Find Central Laboratory Test Schedule

FIND CENTRAL LABORATORY TEST SCHEDULE:

A

B

C

D

SUBJECT TYPE:	Proband ESRD Diabetic	Proband NON-ESRD Diabetic	Families/ MALD Triads & Controls Diabetic	Families/M ALD Triads & Controls Non- Diabetic or Unknown
TEST:				
BUN, Creatinine		X	X	X
Glycosylated Hemoglobin	X	X	X	X
Glucose		X	X	X
Urine Protein, Microalbumin, Creatinine		X	X	X

APPENDIX J: CERTIFICATE OF CONFIDENTIALITY