Undiagnosed Diseases Network Manual of Operations

June 27, 2018

Table of Contents

Table of Acronyms....2
I. Network Overview and Operating Procedures....5
II. Milestones and Metrics....10
III. Institutional Review Board Communications....13
IV. Publications and Research....17
V. Communications....21
VI. Technology....24
VII. Data Sharing....26
VIII. Clinical Protocols....30
IX. Surveys....46
X. Environmental Exposures....50
XI. Sequencing....54
XII. Biospecimens....69
XIII. Central Biorepository....79
XIV. Metabolomics....83
XV. Model Organisms Screening....95
XVI. Appendices....101
APPENDIX 1: NIH UDP Protocols....101
APPENDIX 2: Example Study Recommendation Letters....106
APPENDIX 3: Applicant Review Form (completed by Clinical Sites)....114
APPENDIX 4: Template Letters for Patients and Health Care Providers....117
APPENDIX 5: Wrap-up Template....134
APPENDIX 6: Participant Follow-up Surveys....137
APPENDIX 7: Feature Request Form....141
APPENDIX 8: Research Concept Sheet....142
APPENDIX 9: Annual Survey Review Worksheet....145
APPENDIX 10: PEER Charter and Application....146
APPENDIX 11: Standard Operating Procedures for Not Accepted Applicants....151
APPENDIX 12: Diagnosis Coding Tool v1.0....154

Table of Acronyms

Acronym	Definition
ACMG	American College of Medical Genetics and Genomics
API	Application programming interface
BCM	Baylor College of Medicine
bp	Base pairs
BWA	Burrows Wheeler Aligner
CC	Undiagnosed Diseases Network Coordinating Center
CDC	Centers for Disease Control and Prevention
CIRB	Central Institutional Review Board
CLIA	Clinical Laboratory Improvement Amendments
COI	Conflict of Interest
CPT	Cell Preparation Tube
CR	Continuing Review
CRC	Clinical Research Center
CS	Undiagnosed Diseases Network Clinical Site
CSF	Cerebrospinal fluid
CSL	Clinical Services Laboratory
dbGaP	Database of Genotypes and Phenotypes
DOB	Date of birth
EEG	Electroencephalogram
EMG	Electromyography
FFQ	Food Frequency Questionnaire
FIPS	Federal Information Processing Standards
FISMA	Federal Information Security Management Act
FWA	Federalwide Assurance
gDNA	Genomic DNA
GSL	Genomic Services Laboratory
HA	HudsonAlpha Institute for Biotechnology
HIPAA	Health Insurance Portability and Accountability Act
HPO	Human Phenotype Ontology
HRPP	Human Research Protections Program
ICD	International Classification of Diseases
ICF	Informed Consent and Assent Forms
IRB	Institutional Review Board
LCF	Lipidomics Core Facility
LIMS	Laboratory Information Management System
MC	Undiagnosed Diseases Network Metabolomics Core
MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
NCV	Nerve Conduction Velocity

NHANES	National Health and Nutrition Examination Survey
NHGRI	National Human Genome Research Institute
NHGRI-IRP	National Human Genome Research Institute-Intramural Research Program
NIH	National Institutes of Health
NIST	National Institute of Standards and Technology
NORD	National Organization for Rare Disorders
OMIM	Online Mendelian Inheritance in Man
ORDR	Office of Rare Diseases Research
OSC	Office of Strategic Coordination
PBMC	Peripheral Blood Mononuclear Cell
PCC	Participant Care Coordinator
PCR	Polymerase chain reaction
PE	Paired-end
PEER	Participant Engagement and Empowerment Resource
PhenX	Phenotypes and eXposures consensus measurement
PHI	Personal Health Information
PI	Principal Investigator
PII	Personally Identifiable Information
QC	Quality Control
QWES	Quick Whole Exome Sequencing
SC	Undiagnosed Diseases Network Sequencing Core
SNP	Single Nucleotide Polymorphism
SRC	Scientific Review Committee
SST	Serum separator tube
TAT	Turnaround time
UDN	Undiagnosed Diseases Network
UDNCB	Undiagnosed Diseases Network Central Biorepository
UDN NIH PO	Undiagnosed Diseases Network National Institutes of Health Program
UDP	Undiagnosed Diseases Program
UDPICS	Undiagnosed Diseases Program Integrated Collaboration System
UUID	Universally Unique Identifier
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
WGL	Whole Genome Laboratory

I. Network Overview and Operating Procedures

A. Network Overview

The Undiagnosed Diseases Network (UDN) consists of 7 Clinical Sites (CSs), a Coordinating Center (CC), 2 DNA Sequencing Cores (SCs), a Model Organisms Screening Center, a Metabolomics Core (MC), and a Central Biorepository.

The CC is located at the following institution, with the following PIs:

- Harvard Medical School, Boston, MA - Isaac Kohane, MD, PhD and Alexa McCray, PhD

The CSs are located at the following institutions, with the following PIs:

- Baylor College of Medicine, Houston, TX - Brendan Lee, MD, PhD
- Duke University Medical Center (with Columbia University Medical Center), Durham, NC - David Goldstein, PhD and Vandana Shashi, MBBS, MD
- Harvard Teaching Hospitals (including Boston Children's Hospital, Brigham and Women's Hospital, and Massachusetts General Hospital), Boston, MA - Joseph Loscalzo, MD, PhD
- National Institutes of Health (NIH), Bethesda, MD - David Adams, MD, PhD; William Gahl, MD, PhD; Cynthia Tifft, MD, PhD; and Camilo Toro, MD
- Stanford Medicine, Palo Alto, CA - Euan Ashley, MD; Jonathan Bernstein, MD; and Paul Fisher, MD
- University of California Los Angeles (with Children's National Medical Center), Los Angeles, CA - Katrina Dipple MD, PhD; Stanley Nelson, MD; Christina Palmer, PhD; and Eric Vilain, MD, PhD
- Vanderbilt University Medical Center, Nashville, TN - Rizwan Hamid, MD, PhD; John Newman, MD, PhD; and John Phillips III, MD

The SCs are located at the following institutions, with the following PIs:

- Baylor College of Medicine, Houston, TX - Christine Eng, MD
- HudsonAlpha (with Illumina), Huntsville, AL - Liz Worthey, PhD

The Model Organisms Screening Center is located at the following institutions, with the following PI:

- Baylor College of Medicine (with University of Oregon), Houston, TX – Hugo Bellen, DVM, PhD

The MC is located at the following institutions, with the following PIs:

- Pacific Northwest National Laboratory (with Oregon Health & Science University), Richland, WA – Thomas Metz, PhD and David Koeller, MD

The Central Biorepository is located at the following institution, with the following PIs:

- Vanderbilt University Medical Center, Nashville, TN – Joy Cogan, PhD and John Phillips III, MD

The purpose of this cooperative research Network is to establish a national network added to and building upon the NIH Undiagnosed Diseases Program (NIH UDP). The objectives of this program are to: 1) improve the level of diagnosis and care for patients with undiagnosed diseases through the development of common protocols designed by a community of

investigators; 2) facilitate research into the etiology of undiagnosed diseases, by collecting and sharing standardized, high-quality clinical and laboratory data including genotyping, phenotyping, and documentation of environmental exposures; and 3) create an integrated and collaborative research community across multiple CSs and among laboratory and clinical investigators prepared to investigate the pathophysiology of these new and rare diseases and share this understanding to identify improved options for optimal patient management.

The major funder of the UDN is the NIH Common Fund, which is managed by the Office of the Director/Office of Strategic Coordination (OSC).

B. Cooperative Agreement Responsibilities

The administrative and funding instrument used for the UDN is the cooperative agreement, an "assistance" mechanism (rather than an "acquisition" mechanism), in which substantial NIH programmatic involvement with the awardees is anticipated during the performance of the activities. Under the cooperative agreement, the NIH purpose is to support and stimulate the recipients' activities by involvement in and otherwise working jointly with the award recipients in a partnership role; it is not to assume direction, prime responsibility, or a dominant role in the activities. Consistent with this concept, the dominant role and prime responsibility resides with the awardees for the project as a whole, although specific tasks and activities may be shared among the awardees and the NIH as defined below.

NIH staff have substantial programmatic involvement that is above and beyond the normal stewardship role as described below:

The NIH Project Scientist(s) have substantial scientific and programmatic involvement during the conduct of this activity through technical assistance, advice, and coordination. However, the role of NIH staff is to facilitate and not to direct the activities. It is anticipated that decisions in all activities are reached by consensus of the UDN and that NIH staff are given the opportunity to offer input to this process. The Project Scientist(s) will participate as members of the Steering Committee and will have one vote. The Project Scientist(s) have the following substantial involvement:

1. Participating with the other Steering Committee members in addressing issues that arise with UDN planning, operation, and analysis. The Project Scientist(s) assist and facilitate the group process and do not direct it.
2. Serving as a liaison, helping to coordinate activities, including acting as a liaison to other NIH Institutes/Centers, and as an information resource for the awardees. The Project Scientist(s) also help coordinate the efforts of the UDN with other groups conducting similar efforts.
3. Attending all Steering Committee meetings as a voting member and all working group meetings, assisting in developing operating guidelines, quality control procedures, and consistent policies for dealing with situations that require coordinated action. The Project Scientist(s) are responsible for working with the grantee(s) as needed to manage the logistic aspects of the resource.
4. Reporting periodically on Network progress to the NIH UDN Working Group (a trans-NIH Common Fund working group made up of staff from multiple NIH Institutes and Centers) and through it to the NIH Common Fund and to the National Advisory Council of Human Genome Research Institute.
5. Serving on subcommittees of the Steering Committee, and Working Groups as appropriate.

6. Assisting awardees in the development, if needed, of policies for dealing with situations that require coordinated action.
7. Providing advice in the management and technical performance of the award.
8. Assisting in promoting the availability of the data and related resources developed in the course of this program to the scientific community at large.
9. Participating in data analyses, interpretations, and, where warranted, co-authorship of the publication of results of studies conducted through the program.
10. Other NIH UDN Working Group staff may assist the awardee as designated by the UDN NIH Program Official (The NIH official responsible for the programmatic, scientific, and/or technical aspects of the grant).

Collaborative Responsibilities:

Close interaction among the participating investigators is required, as well as significant involvement from the NIH, to develop and operate the UDN. Principal investigators participate in in-person Steering Committee meetings on a quarterly basis during the first year of funding cycles and subsequently three times per year; during months in which there are not in-person meetings, there are monthly conference calls as needed to share information on data resources, methodologies, analytical tools, as well as data and preliminary results. Key co-investigators and pre- and post-doctoral trainees, especially those who are members of under-represented minority groups or those from different but related disciplines, are also eligible to attend these meetings.

All Awardees agree to work collaboratively to:

1. Assist in refining a common approach to patients with undiagnosed diseases.
2. Work collaboratively with other UDN investigators to provide for secure, accurate and timely data submission.
3. Participate in presenting and publishing new processes and substantive findings.
4. Participate in the governance of the UDN as a member of the Steering Committee.
5. Interact with other relevant National Human Genome Research Institute (NHGRI) and NIH activities, as needed, to promote synergy and consistency among similar projects.

Additionally, the CS Awardees agree to work collaboratively to:

1. Participate in network-wide processes for patient selection and assignment to a specific CS for evaluation.
2. Identify 10 previously unidentified diseases network-wide per year in FY16 and FY17.
3. Fulfill all principal investigator (PI) primary responsibilities laid out in RFA-RM-13-004.

Additionally, the CC Awardee agrees to work collaboratively to:

1. Share statistical experience and expertise across the UDN and provide advice on statistical methods design.
2. Participate with the current NIH UDP investigators to refine and adapt current single center activities to the requirements of the Network.
3. Fulfill all PI primary responsibilities laid out in RFA-RM-12-020.

Additionally, the DNA SC Awardees agree to work collaboratively to:

1. Fulfill all PI primary responsibilities laid out in RFA-RM-13-018.

Additionally, the Model Organisms Screening Center Awardee agrees to work collaboratively to:

1. Fulfill all PI primary responsibilities laid out in RFA-RM-14-016.

Additionally, the Metabolomics Core Awardee agrees to work collaboratively to:

1. Fulfill all PI primary responsibilities laid out in RFA-RM-15-001.

C. Steering Committee Policies

Guideline: A Steering Committee composed of PIs from all sites (including the CC, CSs (including the NIH UDP), and the other Core Laboratories (including the DNA SCs, Model Organisms Screening Center, MC, and Central Biorepository)) and the NIH Project Scientist(s) will be responsible for the scientific direction of the Network, as set forth in the FOAs RFA-RM-12-020, RFA-RM-13-004, RFA-RM-13-18, RFA-RM-14-016, and RFA-RM-15-001. The Steering Committee is responsible for the scientific direction of the Network.

Policies:

1. The Steering Committee is responsible for policy decisions regarding the Network, and for the discussion and resolution of procedural issues that affect the operation and status of the network as a whole.
2. The UDN Steering Committee will be the operational group through which the NIH UDN Working Group interacts with the UDN.
3. The Steering Committee will have at least monthly conference calls.
4. The Steering Committee will meet in person quarterly during the first year of funding cycles and three times per year or as needed subsequently.
5. The minutes for all Steering Committee discussion will be documented and posted on a CC website (viewable to Steering Committee members).
6. The voting members of the UDN Steering Committee include the Principal Investigator(s) of each CS, the PI(s) of the CC, the PI(s) of each Core Laboratory (including the DNA SCs, Model Organisms Screening Center, Metabolomics Core, and Central Biorepository), and the collective NIH IC Project Scientists. Each site has one vote (multiple PIs may all be members of the Steering Committee, but collectively have one vote for their site) and the NIH Project Scientists group collectively has one vote.
7. The Steering Committee may add additional members, and other NIH Program staff may attend the Steering Committee meetings as desired.

STEERING COMMITTEE

Co-Chairs: Cynthia Tifft (UDP) and Vandana Shashi (Duke)

Members:

CS PIs (1 vote for each CS):

1. Baylor College of Medicine - Brendan Lee
2. Duke University Medical Center (w/ Columbia University Medical Center) - David Goldstein and Vandana Shashi (contact)
3. Harvard Teaching Hospitals - Joseph Loscalzo
4. NIH UDP - David Adams, William Gahl (contact), Cynthia Tifft, and Camilo Toro
5. Stanford Medicine - Euan Ashley (contact), Jon Bernstein, and Paul Fisher
6. University of California, Los Angeles (w/ Children's National Medical Center) - Katrina Dipple, Stanley Nelson (contact), Christina Palmer, and Eric Vilain
7. Vanderbilt University Medical Center – Rizwan Hamid, John Newman, and John Phillips III (contact)

CC PIs (1 vote):

1. Harvard Medical School - Isaac Kohane (contact) and Alexa McCray

Core PIs (1 vote for each core):

1. Baylor College of Medicine - Christine Eng
2. HudsonAlpha (w/ Illumina) - Liz Worthey
3. Baylor College of Medicine (w/ University of Oregon) - Hugo Bellen
4. Pacific Northwest National Laboratory (w/ Oregon Health & Science University) - David Koeller and Thomas Metz (contact)
5. Vanderbilt University Medical Center - Joy Cogan (contact) and John Phillips III

NIH IC Project Scientists (1 collective vote):

1. Anastasia Wise

D. Election of Steering Committee Co-Chairs

Guideline: The position of Chairperson of the Steering Committee of the UDN is filled by co-chairs who serve overlapping terms. The first 2 co-chairs are selected by the NIH UDN Working Group. Subsequent co-chairs are selected by a vote of the UDN Steering Committee.

Principles:

1. The term of the position of Chair is 1-2 years in duration unless an extension is agreed to by the Steering Committee.
2. The individual holding the position of Chair must be a current member of the Steering Committee.
3. The Chair must be either the Principal Investigator of one of the CS or a Core Laboratory or the CC.

E. Executive Committee

The Executive Committee consists of the 2 co-chairs of the Steering Committee, the PIs of the CC, and the NIH Project Scientists. The Executive Committee meets weekly to review and monitor UDN progress.

F. Other Network Committees

Guideline: The Steering Committee may establish working groups as needed to address particular issues, which will include representatives from the program, the NIH, and possibly other experts. The Steering Committee has the overall responsibility of assessing and prioritizing the progress of the various working groups and other needed subcommittees of the working groups.

Working Group Governance:

1. Any individual or group proposing a new working group will present their idea to the Steering Committee. A formal vote of the Steering Committee is needed to create a new working group.
2. Volunteers for chair or co-chairs of the new working group will be solicited when the new working group is proposed. A formal vote of the Steering Committee is needed to confirm the chair or co-chairs.

3. Co-chairs are not required for all working groups, but may be recommended by the Steering Committee.
4. Working group co-chairs may come from the same site.
5. If there are no volunteers, or only one, the Steering Committee may recommend a site or type of site that may be a good fit for the working group and one of the Steering Committee co-chairs will solicit the site(s) for a recommended chair.
6. Any working group proposing to close will present their idea to the Steering Committee. A formal vote of the Steering Committee is needed to close a working group.

Active Committees and Working Groups:

1. Billing Working Group
2. Biosamples and Biorepository Working Group
3. Case Review Committee
4. Clinical Protocols Working Group
    a. Subcommittees: Site Operations, Utility & Utilization (U3)
5. Environmental Exposures Working Group
6. Genetic Counseling & Testing Working Group
7. Metabolomics Working Group
8. Model Organisms Working Group
9. Publications and Research Committee
    a. Subcommittee: Survey Committee
10. Sequencing Working Group

G. Implementing and Revising the Manual of Operations

1. Working groups are responsible for development and maintenance of related chapters for the Manual of Operations.
2. Chapters of the Manual of Operations are ratified by the Steering Committee.
3. Working groups have the authority to make decisions regarding implementation of ratified chapters of the Manual of Operations that are assigned to the working group for implementation.
4. If a working group cannot resolve an implementation decision internally, the Steering Committee is consulted.
5. Working groups consult with other relevant working groups on implementation decisions that involve multiple areas of expertise. A cross-working group liaison may be assigned to facilitate these interactions.
6. All working groups make their agendas and minutes available to other working groups.
7. Working groups that would like to recommend: 1) a change to a ratified Manual of Operations chapter that affects network-wide operations, or 2) addition of a new chapter, should recommend the change to the Steering Committee for ratification.
8. Groups that would like to recommend a change to the network-wide Institutional Review Board (IRB) protocol or consents should recommend the change to the Steering Committee.

II. Milestones and Metrics

One of the core functions of the NIH Program and the CC is to monitor each component of the network (CC, CSs, SCs, other Cores) as well as the network as a whole. The rationale for this monitoring is two-fold: (1) to encourage understanding and continuous improvement of

components of the network and (2) to codify the expectations of our funders in order to ensure that our efforts are aligned with expectations.

NIH milestones are listed in the table below. The NIH Program may update these measures over the course of the project. It is anticipated that these metrics will be compiled and reported quarterly to the Steering Committee during the first two years of the UDN, after which the frequency of evaluation should be revisited by the Steering Committee.

NIH Performance Milestones	
1	By Oct 1, 2014 – UDN to establish working groups to develop the network-wide protocols for the UDN, including design, maintenance, and dissemination of the UDN Manual of Operations and central IRB protocol.
2	By April 1, 2015 – Clinical Sites work together with the UDN Coordinating Center to establish a network-wide application process and data standards for above-the-line data.
3	By Oct 1, 2015 – Establish collaboration amongst laboratory and clinical researchers, and share resulting data and approaches throughout the scientific and clinical communities.
4	By Oct 1, 2015 – UDN Coordinating Center to work together with the Clinical Sites to establish Gateway infrastructure for the UDN application process and above-the-line data.
5	By Oct 1, 2015 - Define the mechanism of at least 1 candidate gene in the pathophysiology of a rare or yet-to-be-described disease.
6	By Oct 1, 2015 - the NIH UDP will identify at least 5 candidate genes annually by working jointly with the UDN through FY2017.
7	By Sept 1, 2016 – 6 UDN extramural Clinical Sites to accept 35 patients, per site, in order to initiate evaluations.
8	By Sept 1, 2016 – UDN Sequencing Cores produce exomes/genomes for analysis in collaboration with the UDN Clinical Sites to identify candidate genes by providing raw data (FASTQ and BAM files) to the UDN Clinical sites within 2 weeks of sample receipt.
9	By Sept 1, 2016 - UDN Model Organisms Screening Center to have established the capability to screen 200 variants per year in collaboration with the UDN Clinical Sites.
10	By Oct 1, 2016 – The NIH UDP to work with the UDN Sequencing Cores to analyze the sequencing results of 300 individuals annually (new and returning).
11	By Dec 1, 2016 – UDN Coordinating Center to promote collaboration and data sharing by posting 50 new cases in dbGaP and other relevant databases; 100 new cases by Oct 1, 2017; and 300 new cases by Oct 1, 2018.
12	By Dec 1, 2016 – create a process for sharing gene variants identified through the UDN with the basic science community and sharing results of these studies back with the UDN.
13	By Jan 1, 2017 – Identify 10 unidentified diseases (previously unknown diseases, novel gene associations with known diseases, or novel phenotype associations with known diseases); by Jan 1, 2018, identify a cumulative total of 20 unidentified diseases.

14	By Aug 1, 2017 - UDN Metabolomics Core to establish background controls for various UDN sample types; and by Aug 1, 2017 to have established the capacity to analyze 150 samples per year in collaboration with the UDN Clinical Sites.
15	By Sept 1, 2017 – All 6 extramural UDN Clinical Sites to accept patients at a rate of 50 patients per year, per site, through FY 2017.
16	By April 1, 2018 – All 6 extramural UDN Clinical Sites to see patients at a rate of 50 patients per year, per site, to continue evaluations.

The CC calculates a range of measures to assess the performance of the network. Select metrics are available on the public website at https://undiagnosed.hms.harvard.edu/about-us/facts-and-figures/.

CS metrics calculated by the CC are described in the table below. It is anticipated that the CC will update these metrics over the course of the project.

CS metrics	
1	Length of Application Review: percentage of applications with a decision within 6 months of submission (target: 100%)
2	External Acceptances: percentage of accepted applicants coming from external sites (target: at least 50% external)
3	Adult Acceptances: percentage of accepted applicants >=18 years old (target: 40%)
4	Time between Acceptance and In-Person Evaluation: percentage of participants with an evaluation start date within 6 months of acceptance (target: not defined)
5	In-Person Evaluations Complete: percentage of evaluations complete relative to relevant milestone
6	Participant records approved for data sharing: percentage of evaluated participants with data sharing approved (target: 100%)
7	Samples Submitted to Biorepository: percentage of evaluated participants with samples in the biorepository (target: not defined)
8	Wrap Up Documents: number of evaluated participants with wrap up documents in the Gateway (target: not defined)
9	Concept Sheets Submitted: number of concept sheets submitted (target: not defined)

III. Institutional Review Board Communications

I. Consent Form Development

1. The UDN PI, Central IRB (CIRB), and CC CIRB Liaison will develop model informed consent and assent form (ICF) templates for the UDN, noting sections of the template that must be customized by each CS.
2. The CC CIRB Liaison will make the template ICFs available to the CS Site Coordinators.
3. The CS Site Coordinators will customize only the areas of the ICFs specified in the template, including:
    a. Placing the consent form on the institutional letterhead
    b. Adding standardized language as required by the CS (due to local policy requirements)
    c. Incorporating HIPAA authorization for use and disclosure of PII if HIPAA is not available as a separate document, as per with the CS institution's standard approach. If HIPAA is provided as a separate document, it does not need to be submitted.
4. The CS Site Coordinators will send the completed site specific ICFs to the CC CIRB Liaison.
5. The CC CIRB Liaison will review the ICFs and send them to the UDN PI to submit to the CIRB.
6. The CIRB will review the site specific ICFs with all of the other submitted site materials provided for site approval.
7. The CIRB will communicate the results of the review to the UDN PI, the CC CIRB Liaison, the local PIs, and the Institutional Designees.
8. The CC CIRB Liaison will communicate the results of the review to the CS Site Coordinators.
9. The CIRB will provide to the CC CIRB Liaison the approved ICFs for each CS. ICFs will have a version control reference in the header and on the last page of the form.
10. The CC CIRB Liaison will make the CIRB-approved ICFs available to the CS Site Coordinators and will store centrally for all CSs to access.

II. Site documentation

1. Site documentation includes:
    a. The principal investigators who will be included on the protocol, a roster of associate investigators, the site name, a description of the site, its location and Federalwide Assurance (FWA) number, documentation that UDN Site Human Research Protections Program (HRPP) training requirements have been met, and the name and contact information of responsible institutional officials.
    b. Documentation of the local conflict of interest (COI) review for all investigators on the protocol indicating whether there are any unmitigated or existing conflicts.
    c. Information about the UDN site's local research context as relevant to the site's role in the protocol.
2. The CC CIRB Liaison will send requests for documentation to the CS and Core Site Coordinators.
3. The CS and Core Site Coordinators will send completed documentation to the CC CIRB Liaison.
4. The CC CIRB Liaison will review and send the documentation to the UDN PI to submit to the CIRB.

III. Reportable Events

1. Unanticipated problems involving risks to subjects or others (including adverse events and protocol violations) and/or serious or continuing noncompliance will be reported by the CS and Core PIs directly to the UDN PI, who will report them to the CIRB. CC will request problem reports from the CS and Core PIs on a regular basis.
2. The CC CIRB Liaison will make a form, generated by the CIRB, available to the CSs and Cores to use for reporting the unanticipated problems to the UDN PI.
3. The CS and Core PIs will report all unanticipated problems or serious and/or continuing noncompliance within 7 or 14 days (depending on the nature of the unanticipated problem) to the UDN PI and copy the CC CIRB Liaison on all correspondences.
4. The UDN PI will report the unanticipated problem or serious and/or continuing noncompliance to the CIRB within a timeframe that does not exceed the timing allowed for the PI.
5. The CIRB will communicate the results of the review to the UDN PI, the CC CIRB Liaison, the local PI, the Institutional Designee, and in some cases, the Institutional Official related to unanticipated problems or serious and/or continuing noncompliance.
6. The CC CIRB Liaison will be copied on all correspondences between the CC, UDN PI, and the CSs and Cores.

IV. Continuing Review

1. Three months prior to the continuing renewal deadline, the CIRB will notify the CC CIRB Liaison regarding information required for Continuing Review (CR) and provide the forms that all CSs and Cores, including the intramural site, must complete.
2. The CC CIRB Liaison will notify each CS and Core Site Coordinator regarding information required for CR, and any other forms required by the CIRB.
3. The CS and Core Site Coordinators will submit their responsive information for CR to the CC CIRB Liaison within 2 months of the CR deadline.
4. The CC CIRB Liaison will review the UDN site forms for accuracy and completeness.
5. The CC CIRB Liaison will provide the individual site-specific CR forms as well as submit a single CR Application to the UDN PI.
6. The UDN PI will review the applications and submit all documents to the CIRB.
7. The CIRB will conduct CR of all submitted materials.
8. The CIRB will communicate the results of the review to the UDN PI, the CC CIRB Liaison, the local PIs, and the Institutional Designees.
9. The CC CIRB Liaison will communicate the results of the review to the CS and Core Site Coordinators.
10. **NOTE: All NHGRI protocols undergo review by the Scientific Review Committee (SRC) every three years. The SRC provides the UDN PI with a written review and a summary of outstanding comments and concerns. The UDN PI will provide the required materials to the SRC in the form of an amendment and no less than two months prior to submission to the CIRB for CR in order to permit sufficient time for SRC review.**

V. Amendments

1. Study-wide amendments
    a. Study-wide protocol amendments will be reviewed with the Steering Committee

before submission to the CIRB.

b. A completed amendment form with all supporting documentation, including tracked and clean copies of any modified documents, will be submitted to the CC CIRB Liaison. Study-wide amendments will be submitted to the CC CIRB Liaison the first Monday of the month.
c. The CC CIRB Liaison will finalize study-wide amendments and submit to the UDN PI.
d. The UDN PI will submit study-wide amendments to the CIRB, including tracked and clean copies of all modified documents with updates to the version control of each document.
e. If the study-wide protocol requires changes to the ICF:
    i. The CC CIRB Liaison will modify the ICF model template.
    ii. The CC CIRB will send tracked and clean copies of the modified ICF model template to the UDN PI to submit for CIRB review.
    iii. Once the model template ICFs are approved by the CIRB, the CC CIRB Liaison will work with the CS Site Coordinators to modify each site-specific document.
f. The CIRB will communicate the results of the review to the UDN PI, CC IRB Liaison, the local PIs, and the Institutional Designees.
g. The CC CIRB Liaison will communicate the results of the review to the CS and Core Site Coordinators.
h. The CC CIRB Liaison will make available to the CSs and Cores the approved amendment documents.
i. If there are updated CS ICFs, the CC CIRB Liaison will make the ICFs available to the CS Site Coordinators and will store centrally for all CSs to access.

2. Study staff changes
    a. The CC CIRB Liaison will provide an amendment template for the CSs and Cores to complete that fulfills the CIRB requirements.
    b. The CS and Core Site Coordinators will submit completed site-specific amendment forms with all supporting documentation, with tracked and clean copies of any modified documents, to the CC CIRB Liaison. Site-specific amendments will be submitted to the CC CIRB Liaison on the first and third Mondays of the month.
        iv. For study staff changes, the local PI must sign the completed amendment template indicating conflict of interest review and completed training requirements.
    c. The CC CIRB Liaison will submit the amendments to the UDN PI.
    d. The UDN PI will submit the amendments to the CIRB.
    e. The CIRB will communicate the results of the review to the UDN PI, CC CIRB Liaison, the local PI, and Institutional Designee.
    f. The CC CIRB Liaison will communicate the results of the review to the CS and Core Site Coordinators.
    g. The CC IRB Liaison will make available to the CS or Core the approved amendment documents.

3. Affiliated studies
    a. CSs and Cores may propose studies affiliated with the UDN that are not network-wide.
    b. These studies will require permission from the Publications and Research Committee and an amendment submitted to the CIRB detailing the site-specific

research.

c. The CC CIRB Liaison will keep track of these studies.
d. The UDN site carrying out an affiliated study that has an IRB-approved protocol will provide the CC CIRB Liaison with information about the study, which may include a brief summary of the study, sites involved, nature and characteristics of proband involvement and consent required, and lead UDN investigator.
e. The CC CIRB Liaison will inform the UDN PI and the CIRB of affiliated studies.

DUE DATES

Site-specific amendments are due to CC CIRB Liaison on the 1st and 3rd Monday of the month.

Study-wide amendments are due to CC CIRB Liaison on the 1st Monday of the month.

IV. Publications and Research

One parameter of UDN success is the number and quality of its publications and presentations. The purpose of this document is to establish a framework that facilitates and streamlines collaborative manuscript submission, as well as antecedent work, including meeting abstracts and presentations. The Publications and Research Committee (PRC) oversees the activities set out herein on behalf of the UDN Steering Committee, and report to it. Changes to the policy described herein, which are expected from time to time, must be approved by the Steering Committee. The publication policy applies to a proposed publication if the results are the product of research that the NIH UDN prime or sub-award funded.

A. Scope

1. To facilitate manuscript submission.
2. To provide input on abstract submission and scientific presentation, when requested.
3. To help the CC with content for the website and, if required, social media.
4. To maintain an up-to-date list of all presentations, abstracts, publications and proposals. The CC assists in tracking and coordinating projects.
5. Notwithstanding anything to the contrary in this document, the scope of PRC does not include evaluation of the scientific merit of any publication produced as a result of UDN participation.

B. Membership

The PRC is made up of one representative from each UDN site, core, and the NIH program.

C. Concept Sheets

Concept sheets are required for research projects, data analysis projects, grant applications, and UDN-led manuscripts. This includes but is not limited to: single-site projects, case reports, analysis of Gateway data from participants evaluated at other sites, and biorepository specimen requests from external groups. Concept sheets are encouraged but not required for manuscripts led by external groups. Manuscripts led by external groups should follow all guidelines outlined in section VIII.C. Manuscript Review.

Concept sheets may be submitted online at https://hms.az1.qualtrics.com/jfe/form/SV_3CyZOKOuiEvnCx7 or by email to the CC (see Appendix 8: Research Concept Sheet). Once a concept sheet is submitted, the CC forwards the concept sheet to the PRC for review. The PRC reviews all concept sheets within 2 weeks of submission. The CC notifies submitters of approval at the end of the review period, including any feedback received from the PRC.

Concept sheets must be reviewed and approved by the PRC before substantial work begins. Investigators are encouraged to review existing concept sheets and manuscripts in the Publications and Research Log (found at: https://hmsharvard.box.com/v/prc) to assess redundancy. In addition to PRC review, investigators should note the following requirements for research projects and grant applications:

1. Research projects involving surveys require review by the Survey Committee prior to the PRC.
2. Research projects involving human subjects research require IRB review and approval.
3. Grant applications and research projects led by external investigators require a Steering Committee presentation and vote.
4. Grant applications and research projects involving external investigators and network-wide data and/or samples require a Steering Committee presentation and vote.

Note that UDN investigators do not need to contact each site individually to obtain approval for analysis of Gateway data from participants evaluated at other sites. Only concept sheet submission is required.

Concept sheets are not required when UDN investigators request biorepository specimens from participants evaluated at their own site.

D. Manuscript Review

Review process

All manuscripts must be reviewed and approved by the PRC prior to publication. Final manuscripts should be submitted to the PRC and journal of interest simultaneously. The PRC reviews all manuscripts within 2 weeks of submission. The CC notifies submitters of approval at the end of the review period, including any feedback received by PRC members.

Corporate authorship

The UDN should be acknowledged as a corporate author in the author list as “Undiagnosed Diseases Network”.

Member list

The UDN member list should be submitted to the journal of interest in accordance with the journal’s requirements for consortia. The UDN member list is updated on a quarterly basis and includes members from the sites as determined by the site PIs. When preparing manuscripts, authors should go to https://hmsharvard.box.com/udnmemberlist to download the current version of the member list.

Author indexing

When possible, authors included in the UDN member list should be indexed in PubMed.

Funding acknowledgement

All UDN papers (network-wide and local) must include the following funding acknowledgement:

> Research reported in this manuscript was supported by the NIH Common Fund, through the Office of Strategic Coordination/Office of the NIH Director under Award Number(s) [INSERT AWARD NUMBERS]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The funding acknowledgement statement and award numbers are available at https://hmsharvard.box.com/udnfa.

Conflict of interest

The CC maintains a list of Conflicts of Interest (COI) for all UDN members at https://hmsharvard.box.com/udncoi. COI updates should be sent to the CC.

Notification of acceptance

Investigators should notify the CC when manuscripts are accepted and provide the expected publication date.

Notification of publication

Investigators should notify the CC when manuscripts are published.

Authorship guidelines

Authors (first, middle, and senior) are determined by the type, scope and site of project. First author takes primary responsibility for the manuscript. Given the nature of the UDN's work, shared first or last authors should be remembered as an option. Generally, it is expected that authors would make contributions to any or all of the following including but not limited to the concept, design, acquisition and analyses of data, drafting of manuscript, editing, and revision of manuscript. The PRC resolves all authorship disagreements.

Scope of manuscript review

PRC approval shall be for purposes of satisfaction of the points in this section alone and, for purposes of clarity, this means that the PRC shall not withhold approval of a manuscript on the basis of its scientific merit.

E. Abstracts and Presentations

Abstracts and presentations do not require PRC review. Abstracts and presentations should reference UDN grant number(s) and include "Undiagnosed Diseases Network" in the author list. Accepted abstracts should be sent to the CC so that they may be posted to the UDN CC website and added to the Publications and Research Log.

If there is an NHGRI or Common Fund co-author, final versions of the abstract must be submitted to the PO for review and approval.

F. UDN Publications and Research Log

The PRC advises the CC in management and population of the log of concepts, presentations, abstracts, and publications submitted, finished, and in process. This log can be found in Box.com (https://hmsharvard.app.box.com/v/prc) and the link is sent to the PRC each time a concept sheet or manuscript is circulated. The publications update sent to the PIs quarterly includes all current accepted and published manuscripts and concept sheets.

G. Start of UDN Publication

The PRC developed a manuscript, which described, defined, and introduced the network to the medical and scientific community: Ramoni, Rachel B., et al. "The undiagnosed diseases network: accelerating discovery about health and disease." *The American Journal of Human Genetics* 100.2 (2017): 185-192.

H. End of Funding Cycle Publications

The PRC is responsible for stimulating the preparation of manuscripts that describe the UDN experience towards the end of each funding cycle.

V. Communications

The purpose of this section is to describe the public-facing UDN websites and outline how the UDN engages with the public. It addresses both official and unofficial UDN social media communications. It provides guidance for UDN investigators interested in sharing information about the UDN on personal social media accounts and profiles. Guidelines for responding to media inquiries are also provided.

A. Website

Information about the UDN is made available on public-facing websites maintained by the CC, NHGRI, NIH Common Fund, and CSs.

- CC: https://undiagnosed.hms.harvard.edu/ (alias: udnconnect.org)
- NHGRI: https://www.genome.gov/27550959/
- NIH Common Fund: https://commonfund.nih.gov/Diseases/index
- CSs
  - Duke: http://www.dumcudn.org/
  - Harvard: http://www.bostonudn.org/
  - Stanford: https://undiagnosed.stanford.edu/
  - UCLA: https://www.genetics.ucla.edu/udn/
  - UDP: https://www.genome.gov/27544402/
  - Vanderbilt: https://mc.vanderbilt.edu/undiagnosed/

The CC website is considered the primary source of up-to-date information about the UDN, as it contains the link to the Gateway, content related to the application process, CSs and core laboratories, current research, and press coverage. The CC website also includes contact information for the Helpdesk located at the CC. The NHGRI, NIH Common Fund, and CS websites are primarily for informational purposes.

B. Social Media

Official UDN social media communications

Since the launch of the Network, the CC has worked to cultivate a social media presence to 1) engage with potential applicants, referring providers, and collaborators, and (2) share the work of the UDN with interested stakeholders (e.g., general public, advocacy organizations, scientific community). A secondary goal is to share the UDN participant pages to identify similar patients.

When representing the UDN on social media, the CC:

- Carefully reviews all content for accuracy and relevance prior to posting on any social media platform
- Deletes spam and comments that are unrelated to the topic in question
- Replies to questions and comments when appropriate
- Avoids making political statements
- Notifies the Executive Committee and CS, if applicable, of comments that rise above the level of general questions or concerns

Official social media accounts managed by the CC include Twitter, Facebook, Reddit, and Instagram, which can be accessed searching for UDNconnect or visiting the following links:

- Twitter: https://www.twitter.com/udnconnect
- Facebook: https://www.facebook.com/udnconnect
- Reddit: https://www.reddit.com/user/udnconnect/
- Instagram: https://www.instagram.com/udnconnect

Unofficial UDN social media communications

Unofficial UDN social media communications are defined as communications about the UDN that have not been reviewed by a UDN investigator or NIH program staff member. Individuals and groups engaging in unofficial UDN social media communications are advised to include a disclaimer such as "Views are my own and do not necessarily reflect the views of the UDN or the NIH" on their social media profiles. They are also asked to route inquiries that may warrant or benefit from an official response to the CC.

Institutional social media communications

The CC engages with institutional social media accounts belonging to the NIH and UDN sites as appropriate. A list of these instituional social media accounts is maintained by the CC and is available to UDN members (https://hmsharvard.box.com/v/communications).

Best practices for personal social media accounts and profiles

UDN investigators are encouraged to share the work of the UDN on their personal social media profiles and engage with the official UDN social media accounts with the following guidelines in mind:

- Do not share sensitive or confidential information about UDN research activities or participants
- Avoid expressions that could be misinterpreted and should not engage in discourse that may reflect poorly on the UDN or the NIH
  - Example: politically charged tweets related to UDN funding
- Avoid making statements about unpublished UDN metrics and outcomes without appropriate context
  - Example: public Facebook posts about the high number of diagnoses made at your site

If the CC observes public social media posts made by a UDN investigator violate any of the above guidelines, the CC contacts the individual to suggest that the post be removed or made private. The CC notifies the UDN Executive Committee of such posts when appropriate, as determined by internal review at the CC.

Social media content suggestions

All members of the UDN are invited to send content suggestions for the UDN websites and social media accounts to UDN@hms.harvard.edu.

C. Media Inquiries

General guidelines

In general, UDN sites are welcome to respond to media inquiries directed at their site. If a site determines that the inquiry may benefit from additional perspectives, the site may opt to contact other UDN sites or the NIH. Comment requests related to NIH funding or other administrative issues should be directed to the CC for routing to NIH, NHGRI, and NIH Common Fund staff, as appropriate. Note that any comments from NIH staff require approval by one or more NIH communications offices. In some circumstances, the CC may engage experts at the UDN sites (e.g., individual with experience in broadcast television) when responding to media inquiries.

These guidelines do not apply to press releases and content developed by the UDN sites. In addition, they do not override institutional policies at the individual UDN sites.

Requests for interviews with participants

The CSs coordinate requests for interviews with participants. Cores and other sites that do not interact with participants should direct interview requests to the CC and include the following in writing:

1. An outline of the proposed interview
2. Criteria for selecting interview participants
3. Expected publication date

Once this information is received, the CC works with the CSs to identify participants or families appropriate for the interview.

Unpublished metrics

Each UDN site is welcome to comment on unpublished metrics for their site. However, the only network-wide metrics that should be shared with the media are those displayed on the UDN website at https://undiagnosed.hms.harvard.edu/about-us/facts-and-figures/.

VI. Technology

The UDN Gateway (https://undiagnosed.hms.harvard.edu/) is the centralized web-based patient and investigator portal for the UDN. The Gateway was built by developers at the CC to meet the unique needs of the network. The Gateway and integrated applications are administered by the CC in a manner that is compliant with Health Insurance Portability and Accountability Act (HIPAA), National Institute of Standards Technology (NIST) guidelines, and the Federal Information Security Management Act (FISMA) Moderate regulations.

Of note, each CS and Core uses local systems to track study data that falls outside the scope of the Gateway, such as medical records. The CC does not dictate which systems a given CS or Core chooses to use as long as that decision does not impair or threaten the overall security and compliance posture of the UDN data and technology network. Examples of local tracking systems include electronic medical record (EMR) systems, laboratory information management systems (LIMS), and REDCap. Data entered into these systems are stored securely and indefinitely under the oversight and policies of the individual CS and Cores.

A. UDN Gateway

Capabilities

The Gateway has two interfaces: patient-facing and investigator-facing. In the patient-facing interface, patients, parents/guardians, and referring providers are able to submit applications and study recommendation letters for consideration by the UDN. When an application is created, it is assigned a universally unique identifier (UUID) and a 6-digit UDN ID (e.g., UDN123456).

The investigator-facing interface is organized into CS- and Core-specific dashboards. This structure allows for sharing of personally identifiable information (PII) and personal health information (PHI) between UDN investigators.

Integrated applications

1. PhenoTips: The PhenoTips tool (https://phenotips.org/) is fully integrated into the Gateway to enable standardized phenotyping using the Human Phenotype Ontology (HPO). Features of PhenoTips include a dynamic pedigree editor, measurement and growth curves, and predictive ontology searches.
2. FileService: FileService is an application programming interface (API) that handles permissions, authentication, upload/download signing to cloud-based storage, metadata handling, searching and archiving of files. It is used as a service to the Gateway for managing secure authorized large file uploads and downloads.

Feature requests

The Gateway continues to evolve to meet the needs of the network. This evolution is primarily driven by requests for additional features. The feature request process described here applies to all Gateway feature requests, both large and small. Individuals or groups within the network are able to make feature requests. If a feature requester is able to provide funding for additional programming and support resources, the requested feature may be able to be addressed more quickly. To determine the resources required, the feature requester should speak with the CC

technology team. “Showstopper”/Critical bugs always jump the queue and take priority. These issues are sent to the CC technology team as they arise for immediate resolution.
Feature request process:

1. Feature requester completes the Feature Request Form (see Appendix 14: Feature Request Form) and sends the request to the CC.
2. The CC logs the feature request in the CC queue.
3. The CC technology team assesses the feasibility of the request from a technical and compliance standpoint. This may necessitate asking for additional information from the requester.
    a. If the request is infeasible on a technical or compliance basis, the CC conveys this information to the feature requester and removes the feature request from the queue.
4. The CC technology team assigns an approximate time to complete the feature request.
5. The feature request is preliminarily prioritized by the CC, after which the Executive Committee vets the prioritization for presentation to and approval by the Steering Committee.
6. The CC technology team executes the feature requests in order of priority.

B. Data Security

Physical security

For physical records (paper, photographs, etc.), appropriate physical access controls (e.g., locked cabinet in a locked room) are used at all UDN sites.

Electronic security

For computer systems containing PII and PHI, security controls that are compliant with HIPAA, NIST, and FISMA are utilized and a proper Federal Information Processing Standard (FIPS) Publication 199 assessment will be performed prior to commissioning of these systems.

For computer systems that contain PII but not PHI, appropriate roles-based access and security controls are used and a third-party security assessment will be performed and documented.

For computer systems that contain neither PII nor PHI, roles-based controls are used and conform to the information security and compliance standards of affiliated institutions.

Auditing

The Gateway is assessed periodically via a risk-based approach to determine security, privacy, and compliance requirements. Local systems should have up-to-date virus scanners, disk encryption, workstation-level authentication, and lockouts on all machines accessing the Gateway.

VII. Data Sharing

A. UDN Data Sharing Principles

Data are expected to be shared so as to maximize the value of the significant public investment in the creation and operation of the network. In addition, given the special nature of the project, every effort should be made to share data in a timely and comprehensive manner. Individuals with undiagnosed conditions have often gone for extended periods of time without a diagnosis. Sharing results as rapidly and as broadly as possible holds the promise of accelerating diagnoses and identifying potential treatments for those with these often very serious conditions.

NIH policies

The UDN data sharing principles are consistent with the goals of the NIH Data Sharing Policy[1] which states that "Data should be made as widely and freely available as possible while safeguarding the privacy of participants, and protecting confidential and proprietary data."[2] The NIH Genomic Data Sharing Policy[3] was recently issued, and while it does not yet apply to the UDN, alignment with its goals and requirements is desirable. The Policy expects genomic data generated by NIH-funded research studies to be submitted to a relevant database no later than 3 months after data quality and control have been completed, and expects data to be released up to 6 months after data submission is initiated or at the time of acceptance of initial publication, whichever occurs first.

Expectations of research participants

Participants expect that their data will be shared as early and as broadly as possible in order to maximize the chances of reaching a diagnosis. Several participants have made their data available on the UDN website in the form of participant pages (https://undiagnosed.hms.harvard.edu/participants/).

Commitment of investigators

Investigators agree to further the mission of the UDN by creating new knowledge regarding the biochemistry, physiology, and mechanisms of undiagnosed diseases and improve diagnostic and management options for patients afflicted with them.

Investigators further agree to acknowledge that in pursuit of this mission, common UDN goals may supersede individual goals. Specifically, in the interest of rapid progress, investigators commit to:

1. Model a collaborative, open, interdisciplinary spirit, characterized by mutual trust and respect across disciplines, individuals, areas of expertise, institutions, and by demonstrating interest and engagement beyond their own specific domains.
2. Ensure that data generated at individual sites are comparable and additive by adhering to UDN data standards.
3. Make data available to the Gateway and external databases in a timely and effective manner.

The following outlines the minimum requirements for sharing the data that are collected in the course of participation in the UDN.

B. UDN Data Sharing Procedures

Data Sharing and Use Agreement

Policies and procedures for data sharing within and outside of the UDN are available (https://hmsharvard.box.com/v/dsua).

Informed consent and data security

At the time of application, individuals consent to have their data shared within the UDN. Each applicant is assigned a UUID and UDN ID by the Gateway. These identifiers are used for all data associated with that applicant. Role-based access and physical security controls that are fully aligned with the sensitivity of the data at each point of use and access are employed by the Gateway. De-identified data shared outside of the UDN does not reveal individual identifying information, consistent with the HIPAA Privacy Rule[4]. Individuals accepted into the Network give consent to have their data shared more broadly as part of a network-wide informed consent process.

Sharing identified data

With UDN investigators: All clinical, biospecimen, and sequencing data that are generated by the UDN effort are shared in a secure and compliant manner within the UDN. UDN data include data generated both in human subjects and laboratory research. All UDN investigators have access to the fully identified UDN data. UDN investigators who acquire UDN data must use the data responsibly and must monitor the use of the data by study personnel at their site.

With disease experts: In aid of a possible diagnosis, UDN data may be shared with disease experts who are not currently part of the UDN. In this case, the UDN CS that evaluated the participant must be notified that an outside expert has been consulted.

Sharing de-identified data

The UDN shares de-identified data in the following publicly accessible databases:

dbGaP: dbGaP[5] (Database of Genotypes and Phenotypes) is an NIH-sponsored controlled access repository that was developed to archive and distribute the data and results from studies that have investigated the interaction of genotype and phenotype in humans. For each participant, phenotypic data entered into the Gateway and genotypic data, in the form of BAM and VCF files, are deposited in dbGaP.

PhenomeCentral: PhenomeCentral[6] is a global repository and collaboration platform for secure data sharing in the rare and undiagnosed diseases community. PhenoTips[7] is a web-based tool for collecting phenotypic data that is fully integrated with the Gateway. For each participant, phenotypic data entered through the Gateway and genotypic data, in the form of candidate genes, are deposited in PhenomeCentral.

ClinVar: ClinVar[8] is a freely accessible, public archive of the relationships between specific human variations and phenotype. For each participant diagnosed with a genetic condition, variant level data will be shared in ClinVar. All data are submitted in accordance with ClinVar

requirements. Currently, this includes: gene symbol, reference sequence, descriptions of DNA and protein level sequence variants using Human Genome Variation Society (HGVS) nomenclature, variant type, transcript, cytogenetic and genomic positions, clinical significance and supporting evidence, demographic information, and phenotype information. This information is retrieved from the CS and the Gateway.

Additional databases: As data are generated by additional research cores, relevant databases for deposit of UDN data will be identified. These might include model organism databases, such as Flybase, Zebrafish Information Network, and the Mouse Genome Database, and metabolomics databases, such as the Human Metabolome Database and MetaboLights. In addition, the network will share data publicly on the UDN website. The UDN website[9] is a resource that provides extensive information about the goals, activities, and outcomes of the network. Phenotypic data and candidate gene names are displayed. Sharing data on the UDN website could lead to the identification of additional patients with similar phenotypic and genotypic profiles, and could be useful for advancing scientific investigation and discoveries about these rare conditions.

Data sharing timelines

De-identified participant data must be ready for deposit in dbGaP and PhenomeCentral no later than 3 months after the in-person evaluation is complete and the initial sequencing report is uploaded to the Gateway. This means that phenotyping needs to be complete and, if relevant, candidate genes need to have been identified and entered into the Gateway.

De-identified participant data must be ready for deposit in ClinVar and for display on the UDN website no later than 6 months after the in-person evaluation is complete and the initial sequencing report is uploaded to the Gateway.

Although broad and timely data sharing are expected, UDN investigators have the option of delaying data sharing in certain unusual cases. A rationale for the delay and an estimated data release date must be submitted to the CC. The CC contacts the investigator shortly after the estimated data release date to assess progress and to determine if the data are ready to be shared.

Data release at publication

There is a separate publications section within this document that describes these policies. If broad data release is required as a condition of publication by the authors or the publisher, the CC should be contacted as soon as possible prior to making any commitments to ensure that the data release is feasible.

Role of the CC

The CC facilitates, monitors, and reports on the effective and timely sharing of data within the UDN and beyond.

C. References

1. NIH Data Sharing Policy. http://grants.nih.gov/grants/policy/data_sharing/

2. Final NIH Statement on Sharing Research Data, February 26, 2003. (http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html)
3. NIH Genomic Data Sharing Policy. https://gds.nih.gov/03policy2.html
4. HHS - Office for Civil Rights - HIPAA. http://www.hhs.gov/ocr/hipaa/
5. Tryka KA, Hao L, Sturcke A, Jin Y, Wang ZY, Ziyabari L, Lee M, Popova N, Sharopova N, Kimura M, Feolo M. NCBI's Database of Genotypes and Phenotypes: dbGaP. Nucleic Acids Res. 2014 Jan;42(Database issue):D975-9.
6. Buske OJ, Girdea M, Dumitriu S, Gallinger B, Hartley T, Trang H, Misyura A, Friedman T, Beaulieu C, Bone WP, Links AE, Washington NL, Haendel MA, Robinson PN, Boerkoel CF, Adams D, Gahl WA, Boycott KM, Brudno M. PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. *Hum Mutat.* 2015 Oct;36(10):931-40.
7. Girdea M, Dumitriu S, Fiume M, Bowdin S, Boycott KM, Chénier S, Chitayat D, Faghfoury H, Meyn MS, Ray PN, So J, Stavropoulos DJ, Brudno M. PhenoTips: patient phenotyping software for clinical and research use. Hum Mutat. 2013 Aug;34(8):1057-65.
8. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, Jang W, Katz K, Ovetsky M, Riley G, Sethi A, Tully R, Villamarin-Salomon R, Rubinstein W, Maglott DR. ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016 Jan 4;44(D1):D862-8.
9. The Undiagnosed Diseases Network. https://undiagnosed.hms.harvard.edu/

VIII. Clinical Protocols

A. Introduction

The clinical protocols component of the Manual of Operations "provides preliminary protocols and operating guidelines that will define an initial framework for common approaches to participant selection, data collection, laboratory investigation, and diagnosis, and serve as a base for further refinement by UDN investigators." (From RFA-RM-12-020).

Background of the UDP (see Appendix 1: NIH UDP Protocol for additional information)

Delivery of medical care to patients with rare and yet-to-be described diseases can be fraught with repetitive, inconclusive efforts at diagnosis as patients and their families go from physician to physician in hopes of finding answers. The Office of Rare Diseases Research (ORDR) notes that 6% of individuals seeking their assistance have an undiagnosed disorder and as many as 15% remain in the undiagnosed category for at least 5 years as physicians labor to define cause and pathophysiology. To address these issues, the NIH UDP was established in May 2008, as a joint venture of the NIH ORDR, the National Human Genome Research Institute Intramural Research Program (NHGRI-IRP), and the NIH Clinical Research Center (CRC). The goals of the UDP were to:

1. Provide answers for patients with undiagnosed diseases;
2. Generate new knowledge about disease mechanisms;
3. Assess the application of new approaches to phenotyping and the use of genomic technologies;
4. Identify potential therapeutic targets, if possible.

UDN Clinical Protocols Working Group

The Clinical Protocols Working Group developed these clinical protocols as part of the Manual of Operations and, with input from the Steering Committee, continues to refine it. Should there be a need to vote on matters within the working group, each CS, the NIH Program, and the CC cast a single vote, for a total of 9 votes.

B. Detailed UDN Clinical Protocol

Study design

In this study, individuals with undiagnosed diseases, and their family members when applicable, are investigated. Applicants apply to the UDN through a secure website managed by the CC, called the Gateway, and are assigned to a CS based on an assignment algorithm. The CS collects and reviews the applicant's medical records and accepts or does not accept the applicant. Accepted applicants typically are evaluated at the CS to which they were assigned; however, applicants may be reassigned to a different site based on presenting problems and the expertise of the site. Enrolled individuals undergo a comprehensive medical and family history, physical examination, laboratory testing, imaging studies, consultations, and biological specimen collection, typically over the course of up to a five-day evaluation. Follow-up visits may occur if indicated. See ClinicalTrials.gov record for a publicly available summary of the protocol.

Triaging and accepting applicants

Applicants (or their legal guardians) may learn about the UDN from a variety of sources, including the UDN website, publicity, or from another participant. Healthcare providers not associated with the UDN may learn of the UDN from sources including the UDN website, publicity, colleagues, or medical conferences or publications. Healthcare providers from CSs may refer their own patients for evaluation.

Individuals (or their legal guardians) register and apply to the UDN through the Gateway managed by the CC. In order to be considered for participation, individuals (or their legal guardians) are required to either electronically sign or verbally agree to a disclosure statement allowing the UDN to store the applicant data that is used to: a) assign the applicant to a CS for review, and b) collect characteristics of people who apply to the UDN. If an applicant (or legal guardian) does not speak English, a translator is used to facilitate the verbal consent process.

Once the applicant (or legal guardian) provides consent, either the applicant (or legal guardian), their referring provider, or a CC representative enters the following data into the Gateway:

1. Applicant's name, date of birth, gender, self-described race and ethnicity (for children <18 years the names of their parents are required).
2. Applicant's mailing address, contact information (email address, phone number). For children, the addresses and contact information for both parents are required. If parents are divorced or separated, they must provide information regarding who is legally permitted to sign a consent for medical research on behalf of the child. If parents are separated or divorced they must also both be willing to: (1) provide family history information, and (2) submit DNA samples for genomic analysis. If an adult applicant is unable to consent, the name and contact information of the individual with legal power of attorney who is able to consent on the applicant's behalf must be provided.
3. Evaluation history
4. UDN site preference
5. Travel limitations
6. Referring provider's name and contact information (mailing address, email address, phone number, fax number).
7. Applicant's chief complaint, identification of the system most involved (i.e., cardiac, gastrointestinal), and symptom onset.
8. Environmental exposures. Probands are sent an email asking them to complete an environmental exposures questionnaire. The survey is administered through Qualtrics and data is stored in Qualtrics and in the Gateway. If participants are unable to complete the survey online, paper surveys can be requested. Completed paper surveys are sent to the CSs for data entry. Survey responses are used by the CSs to better inform the evaluation.

Either the applicant (or legal guardian), their referring provider, or a CC representative also uploads to the Gateway a study recommendation letter (see Appendix 2: Example Study Recommendation Letters) summarizing the following: pertinent medical problems, prior diagnoses, history of evaluations and tests, medications, family history, review of systems, and provider's diagnostic impressions.

If an applicant (or legal guardian) does not speak English, a translator is used to facilitate the application process. If an applicant does not have access to the Internet, a paper application can be requested through the CC. Completed paper applications are mailed to the CC for data entry and CS assignment. Applicants (or their legal guardians) are instructed to refrain from

sending additional information, including any records, until assigned to a CS. If an applicant (or legal guardian) or their referring provider sends materials to the CC, the CC holds them until the applicant is assigned a CS, at which time the materials are sent to the CS.

The application is assigned to a CS for triage to determine if the applicant is appropriate for acceptance into the UDN at the assigned site. The assignment to a site is based on an assignment algorithm that takes into account the individual's location, pertinent medical problem(s), site workload and site preference. If a given site is the closest and/or best suited to see a given subject but has already reached enrollment quota, the case would likely be assigned to a different site.

If an application is deemed inaccurate or incomplete by a CS (e.g., study recommendation letter not written by a healthcare provider), the CS contacts the CC. The CC reaches out to applicants accordingly.

The CS makes initial contact with the applicant within 30 days of application submission to request that he/she send all information they have related to the reason for their application to the UDN. This may include: medical records, reports, laboratory studies, radiographic studies, photographs or videos, and pathology slides and reports. If the applicant (or legal guardian) has difficulty obtaining the medical records (including imaging and pathology materials), the site may also contact the referring provider for more information with appropriate release of information from the applicant. Records are reviewed by the CS for completeness. The CS requests any missing components (e.g., images, biopsy slides). The site director or his/her designee assigns the records for review to consultants based upon the specialty involved.

The CS reviews the applicant's records and study recommendation letter and makes a recommendation regarding acceptance into the UDN. (See Appendix 3: Applicant Review Form.) Since few individuals can be accepted into the UDN each year due to limited resources, preference is given to applicants for whom there is the greatest potential to provide a diagnosis or generate new knowledge about disease mechanisms. Inclusion and exclusion criteria are outlined below.

Inclusion criteria:

1. The applicant has a condition that remains undiagnosed despite thorough evaluation by a healthcare provide and has at least one objective finding.
2. The applicant (or legal guardian) agrees to the storage and sharing of information and biomaterials in an identified fashion amongst the UDN sites, and in a de-identified fashion to research sites beyond the network.

Exclusion criteria:

1. The applicant has a diagnosis that explains the objective findings.
2. Review of the records suggests a diagnosis and further evaluation by the UDN is deemed unnecessary.
3. The UDN is unlikely to improve on the comprehensive workup the applicant already received.
4. The applicant is too seriously ill to travel safely to the CS and telemedicine/distance consult is not possible or appropriate.

The Case Review Committee meets on a regular basis (weekly or biweekly) to review (at least initially) all cases at the CSs that have been recommended for acceptance, assign an applicant to a different CS if it feels that the applicant may be more appropriate for another CS based on expertise, and discuss challenging cases. Each CS and SC designate two Case Review Committee members and at least one alternate. Each CS has at least one clinician present at any Case Review Committee meeting. Ordinarily, a pediatrician and an adult internist from each CS are involved in each meeting. The co-chairs of the Committee rotate among the CSs every 6 months. One of them attends each meeting. If both are conflicted for a meeting, then a member of the Committee is asked to chair.

The Case Review Committee meeting is a forum for the CSs to concisely present participants to the UDN clinicians for review and input. The participants fall into two general categories, 1) those that the CS has accepted for evaluation and 2) those for whom the CS is uncertain, has questions about, or thinks may be better served at another CS. Sites are encouraged to present questionable cases for discussion rather than making acceptance decisions at the site level. Each CS presents a one-page summary of the case, and any imaging or pictorials that aid in the decision to accept for evaluation. Presentation and discussion of each case lasts no more than five minutes. If two CSs need prolonged discussion about an applicant, this is taken off line after the meeting. Cases are presented in the order they are submitted to the Case Review queue in the Gateway. The meetings terminate after an hour or 90 minutes. The Case Review Committee co-chairs move the meeting forward and take discussions off line if necessary.
The decision to invite an applicant for evaluation is made primarily by individual CSs, except those for whom the CS is uncertain, has questions about, or thinks may be better served at another CS. The meeting exists to add value to the intended evaluation and to allow the UDN to be informed of the composition of the participant study population,

It is expected that about half of cases are pediatric, half adults. No more than half should be patients known to a CS, half completely new to any CS. “Patients known to a CS” are defined as any patient that is recommended by a healthcare provider from any of the institutions that are on that CS’s award. A patient recommended by an outside provider is not considered “known to a CS” even if s/he has been seen previously at an institution on that CS’s award.

A member of the Case Review Committee may also identify when s/he believes that an accepted or pending case falls outside of the inclusion and exclusion criteria specified in the protocol. In this case, a vote is called. Each eligible CS submits a vote to the Committee co-chairs. The CC, NIH program, and Cores do not have votes. If a majority vote of the Case Review Committee identifies a deviation, the CS may request the opportunity to do a more in-depth review of the case at the following meeting, which would be followed by a final majority vote. If a deviation is finally determined, a pending case would not be accepted, while an accepted case would be addressed in consultation with the IRB.

Application outcomes include:

1. Applicant and site appropriate for acceptance
2. Applicant appropriate for acceptance but reassigned to a different site
3. Applicant requires further testing or evaluation and may be reconsidered following receipt of the results
4. Applicant not appropriate for UDN
    i. A diagnosis was identified based on the review.
    ii. A potentially beneficial referral was identified based on review. An evaluation at the UDN may not be necessary to make a diagnosis.

iii. The applicant is more appropriate for an expert site outside of UDN such as a research protocol at the NIH or elsewhere, the applicant is more appropriate for an expert at an academic medical center or elsewhere, or the applicant is not appropriate for the UDN and no alternative can be identified.

In all cases the referring provider and the applicant is informed of the decision, generally within 60 days after receipt of all the medical records. The CSs send out the disposition letters whether they accept or not accept. (See Appendix 4: Template Letters for Patients and Healthcare Providers). When individuals are not accepted into the UDN, their application information is stored securely and indefinitely in the Gateway. More information related to applicants that are not accepted can be found in Appendix 11: UDN Standard Operating Procedures for Not Accepted Applicants.

Genetic testing

In some cases, it may be useful to have the results of genetic testing (whole-exome/whole-genome sequencing (WES/WGS)) from the enrolled proband and relevant family members prior to the clinical evaluation. Genetic testing is only performed on individuals accepted into the UDN. The sequencing is done at the Baylor College of Medicine and/or HudsonAlpha. The decisions regarding timing of sequencing and WES versus WGS are made on a case-by-case basis as clinically indicated, and left to the discretion of the CS responsible for the evaluation.

If review of the proband's presenting medical problem(s) and medical records suggest that performing sequencing prior to the evaluation would be beneficial and aid in diagnosis, informed consent is obtained over the phone or by videoconferencing (or in person if reasonable) from the enrolled proband, parent, or guardian. The consent form is sent (by mail or email) to the proband (or legal guardian) prior to the remote or in-person consent. The PI, an associate investigator, a genetic counselor, or a project coordinator trained in consenting is available to answer questions and obtain consent. Consent is obtained at this time for the entire study, including for 1) obtaining blood for DNA extraction and sequencing, 2) any research studies performed as part of the evaluation, 3) obtaining other samples (blood, urine, etc.) during the evaluation for research, and 4) the collection of all of the clinical and research data by the UDN for research use. The signed consent forms are sent to the CS responsible for the evaluation and uploaded to the Gateway managed by the CC. The CS also records the consent version signed by the proband and proband's preferences in the Gateway managed by the CC. Genetic counseling is also provided by a physician or genetic counselor to all probands during the consent process. This genetic counseling includes a discussion of the types of results an individual may or may not receive (including primary and medically actionable findings), the likelihood of receiving these types of results, individual preferences for types of results returned, and limitations of the genetic testing including false negative and false positive results. Participants are given the choice of receiving medically actionable findings related to conditions with treatment or management options. Genetic counseling aids were developed to supplement this session.

Assent is obtained over the phone (or in person if reasonable) in the presence of a parent or guardian for all children ages 7-17 years old who are not decisionally impaired. Assent is given for the entire study, including assent for 1) obtaining blood for DNA extraction and sequencing, 2) any research studies performed as part of the evaluation, 3) obtaining other samples (blood, urine, etc.) during the evaluation for research, and 4) the collection of all of the clinical and research data by the UDN for research use. The signed assent form is sent to the CS

responsible for the evaluation and uploaded to the Gateway. The CS also records the assent version signed by the proband in the Gateway.

Once a signed consent form has been received from the proband (or legal guardian), a family history is obtained over the phone (or in person if reasonable) to identify other family members of interest related to the proband's phenotype. Family members are recruited through the proband, i.e., the researchers ask the proband (and/or legal guardian) for permission to contact the family members. Priority is given to those family members who would be most informative for sequencing analysis. Informed consent is obtained over the phone or by videoconferencing (or in person if reasonable) from interested family members. The consent form is sent (by mail or email) to the family member prior to the remote or in-person consent. The principal investigator, an associate investigator, a genetic counselor, or a project coordinator trained in consenting obtains consent and is available to answer questions. Consent is obtained for the collection of blood for DNA extraction and sequencing and for the collection of all of the clinical and research data and pertinent lab specimens for research use. The signed consent forms are sent to the CS responsible for the evaluation and uploaded to the Gateway. The CS also records the consent version signed by the family member and family member's preferences in the Gateway. Genetic counseling is also provided by a physician or genetic counselor to all family members. This genetic counseling includes a discussion of the types of results an individual may or may not receive, the likelihood of receiving these types of results, individual preferences for results returned, and limitations of the genetic testing including false negative and false positive results. During the consent, family members are given the choice of receiving medically actionable findings if they are identified. Genetic counseling aids were developed to supplement this session.

Once a signed consent or assent form has been received from the proband and family member(s), a kit is sent to the proband and family member(s) for obtaining DNA. The kit includes: tubes for blood collection, an order form, directions for payment by the CS (direct costs for the collection and shipping of samples are covered by the CS), and an addressed shipping container for the blood to be sent back to the CS. It is expected that the blood collection is completed with the assistance of the proband's local healthcare provider or local laboratory.

For probands selected for pre-evaluation sequencing, DNA is extracted at the CSs and sent to one of the SCs, where the sequencing is performed and analyzed prior to the evaluation in order to have the analysis available by the time the plan for the proband's evaluation is finalized. Significant variants are identified using standard programs for assessing pathogenicity, Mendelian segregation patterns, allele frequencies, and databases of benign variants. Through this process, medically actionable findings may be identified. Findings intended to be reported to probands or participating family members for use in clinical decision making are confirmed by Sanger sequencing in a Clinical Laboratory Improvement Amendments (CLIA) certified laboratory.

If possible, results of genetic testing are disclosed to the proband (or legal guardian) or family member during an in-person genetic counseling session at the CS. However, since some participants may have to travel a significant distance to get to the CS, results of genetic testing may need to be disclosed to the participant (or legal guardian) in a genetic counseling session over the phone. This may occur when blood is drawn for sequencing as part of the evaluation or when the participation of a non-local family member is limited to a blood sample for DNA analysis and a medically actionable finding has been identified. CLIA-certified results of genetic testing are provided to probands (or legal guardians) during a genetic counseling session with a

qualified physician or genetic counselor. Genetic results related to the indication for testing are returned to all probands (or legal guardians) and their referring providers.

If a primary finding is identified, it is listed on the proband's genetic testing report. If other family members undergo genetic testing, the report may include information about inheritance and other family members that carry the primary finding. A parent of a proband may therefore learn information about his/her own genetic status or the status of his/her relatives when receiving his/her child's results. If this report is shared with relatives, they may also learn information relevant to their own genetic results. These results are discussed during a genetic counseling session with a qualified physician or genetic counselor.

The SCs report medically actionable findings in variants in the genes recommended for such reporting by the American College of Medical Genetics and Genomics. In addition, both cores report other medically actionable findings beyond the currently recommended genes, provided these additional findings meet the threshold of having a defined medical treatment or specific management guidelines for disease surveillance. Stringent criteria for interpretation of variants in these medically actionable genes is applied; reported medically actionable variants have either been previously reported as pathogenic or expected to be pathogenic based on the usual molecular mechanism associated with the gene.

EXAMPLES OF MEDICALLY ACTIONABLE CONDITIONS NOT INCLUDED ON THE ACMG LIST	
Syndrome/Disorder	**Gene**
Progressive Familial Heart Block, Type 1B	*TRPM4*
Long QT type 4	*ANK2*
Familial Gastric Cancer	*CDH1*
Von Willebrand disease	*VWF*

Carrier status for autosomal recessive conditions may also be identified. Examples include cystic fibrosis and Tay-Sachs disease. Individuals who are carriers of a specific genetic disorder have one normal, working copy of the gene, and one changed copy of the gene. Because they have one working copy of the gene, carriers do not usually have symptoms of the disorder. Findings without medical treatment or specific management guidelines are not reported by the SCs. Examples of these types of disorders include Alzheimer disease and late onset skeletal dysplasias (e.g., spondyloepiphyseal dysplasia - *TRAPPC2* gene).

It is important to note that the lack of reportable medically actionable findings does not rule out any disease-causing genetic changes in these or other genes. If a medical concern is raised regarding a specific condition, further genetic testing may be warranted.

As a further measure to ensure consistency between the SCs, the SCs communicate and reach consensus on the reporting of each medically actionable variant that they propose to report.

During the pre-test genetic counseling/informed consent session, probands (or legal guardians) are given the option of receiving genetic results that are unrelated to the indication for testing, including results related to: (1) medical conditions with treatment or management options, known as medically actionable findings, and (2) carrier status findings (only applicable for adult probands). The proband's genetic testing results are shared with family members only if the proband (or legal guardian) provides permission. During the pre-test genetic

counseling/informed consent session, family members are given the choice of receiving medically actionable findings if they are identified. If a medically actionable finding is discovered in a proband and a family member is interested in learning if he/she also has this finding, the family member is referred to a clinical genetics program for genetic counseling and testing. As this counseling and testing is not done as part of the UDN, cost associated with this clinical follow-up is billed to insurance.

There are cases where the results of the sequencing performed prior to the evaluation point to a likely diagnosis and in these cases, the CSs are strongly encouraged to continue with the complete evaluation of the individual. The evaluation allows the site to collect phenotypic data about the condition, provide counseling, and make suggestions about management. Exceptions are cases where the diagnosed condition is common enough that established management standards exist and the presenting phenotype is a typical presentation of that disorder. If the clinical presentation varies from the typical clinical presentation of a well-recognized disorder, then phenotyping of the proband is still appropriate.

Planning the evaluation

Once a proband has been enrolled into the UDN and assigned to a CS, the site works with the proband (or legal guardian) and the local team to create a plan to maximize the efficiency of the evaluation. Information gathered by the CS prior to the evaluation includes:

1. Recent medical records (including consultation reports)
2. Previous tests and results
3. Pathology data/slides
4. Imaging/radiography results
5. Review of the medical history and review of systems
6. Family history (including name, age, and contact information of family members)
7. Medication list (including doses, schedule)
8. Contact information for the proband's relevant physicians
9. Environmental assessment
10. Nutritional assessment
11. Proband needs for travel and admission (including ventilators, mobility issues, etc.).
12. Optional: additional surveys and interviews

The CS creates a plan for evaluation that includes:

1. Determining the lead physician.
2. Determining clinical tests, procedures, consults, or research studies to be performed, including a determination of whether additional IRB approvals are required.
3. Scheduling dates for the proband evaluation – typically evaluations occur over several days (expected to be five sequential days in most cases).
4. Determining the sequence and schedule of tests, procedures, and consults.
5. Arranging a "sedation day" if indicated (especially important for children to maximize the efficiency and minimize the number of times a child needs to be sedated for a procedure or test).
6. Arranging travel based on medical needs.

Evaluation schedule

On day 1 of the evaluation, if the proband (or legal guardian) has not already provided consent, consent is obtained for: 1) drawing blood for DNA extraction and sequencing, 2) any research studies performed as part of the evaluation, 3) obtaining other samples (blood, urine, etc.) during the evaluation for research, and 4) the collection of clinical and research data. If not already provided, assent is obtained for all non-decisionally impaired children 7-17 years of age. Following the consent, the CS records the consent version signed by the proband and proband's preferences in the Gateway. Probands may also be consented for other research projects at this time.

The initial visit with the primary care team, including the lead physician, also typically occurs on day 1. This visit often involves review of the medical history, review of the family history, performance of a physical examination, genetic counseling if results of genetic testing are available, surveys and interviews for research, and review of the goals of the visit and schedule. Changes in the schedule based on this initial visit are often made.

On subsequent evaluation days, all tests, procedures, and consultations take place. Clinical investigations during the evaluation may include: laboratory testing, imaging studies, and biological specimen collection. Genetic counseling may occur if results of genetic testing are available. Specialized research studies, such as proteomics, metabolomics, and functional studies, may also be performed to elucidate underlying mechanisms of disease. Surveys and interviews for research may also be administered and consultations/counseling sessions may be recorded (if proband and/or family members give permission). Surveys and interview guides are submitted to the IRB for review prior to their use. For individuals who did not undergo genetic testing prior to the evaluation, if it is determined during the visit that genetic testing is clinically indicated, blood is drawn and DNA is extracted and sent to a SC for sequencing.

On the last day of the evaluation, the team meets with the proband and family to summarize the evaluation and make plans for follow-up.

During the course of the evaluation, clinical diagnostic studies are performed as clinically indicated and within the standards of accepted medical practice. Specialized research studies are performed as deemed relevant. Clinical specimens are collected as medically indicated and at the discretion of the CS where the proband is being evaluated. Recommendations for research specimen collection are presented in the Biospecimens section of this document.

Unanticipated non-genetic medical information

During the course of this study, it is possible that unexpected medical information is discovered that is important to the proband's health care. This information is provided to the proband's health care provider. At the time, the proband (or legal guardian) is given the option of learning this information and referrals are provided as needed.

Change in clinical stability

If, during the course of the evaluation, the proband has a significant change in clinical stability requiring escalation of care or initiation of new treatments not covered by the research protocol, the proband may be offered completion of the UDN protocol at a later date. Probands and their caregivers, as well as referring providers, are apprised of this change of condition necessitating active treatment rather than research-based investigation. If the proband's condition does not allow discharge from the CS at the scheduled completion date, care is assumed either by the

referring provider or appropriate clinical team members at the CS. Payment for further acute care is provided by the participant's insurance company.

Withdrawal procedures

During the UDN study, applicants and participants may withdraw or be withdrawn by UDN investigators. Withdrawal scenarios include, but are not limited to, the following:

1. Applicant or participant elects to withdraw
2. Applicant or participant is not compliant with study procedures or does not follow instructions given by UDN investigators
3. Applicant passes away before full study enrollment
4. Applicant is accepted, but is unable to be evaluated
5. Applicant is accepted, but site is unable to contact to schedule evaluation

If the applicant or participant needs to be withdrawn, the applicant or participant sends a written withdrawal to the PI at the CS (if possible). The site coordinator destroys the applicant or participant's research records and specimens at the CS. The site coordinator then contacts the CC to remove the applicant or participant's Gateway record. The CC changes the status to withdrawn on the Gateway and in written network records. If an individual was a participant, the CC marks the Gateway patient record as inactive. The site coordinator also contacts the biorepository to destroy the participant's research specimens and collaborators to destroy research records or specimens, if possible. If the individual was a participant, the site coordinator saves the signed consent form with the written withdrawal (if available).

Evaluation wrap-up

At the conclusion of the evaluation, and prior to discharge, the lead physician and other members of the care team as appropriate meet with the proband and family to: summarize the results of the clinical evaluation (clinical and research tests performed, procedures performed, consultations provided, results of testing received, and pending test results), provide genetic counseling as indicated, make recommendations for follow-up with the medical home team, provide clear instructions about how to contact UDN team members if additional questions or concerns arise, and answer any questions the proband or family may have. This wrap-up is facilitated using a structured wrap-up form (see Appendix 5: Wrap-up Template).

The wrap-up report and a short letter highlighting the key findings and follow-up recommendations accompanies the transfer of records to the referring healthcare provider and other providers designated by the proband or family. The wrap-up report and a narrative summary of the evaluation are also uploaded to the Gateway.

All consultation and laboratory study reports pending at the time of discharge are included in a revised wrap-up report sent to the proband, referring provider, and any other care providers the proband has designated. If additional revisions occur, updated and revised wrap-up reports are issued.

Return visits to the UDN site

Follow-up visits to the CS are not generally expected but may occur under at least two circumstances: (1) the CS requests additional phenotyping of the proband or family members to

clarify or inform “affected” status or further interrogate candidate genes, and (2) a diagnosis has been made and the family returns for delivery of results.

Billing procedures

The RFA stated that the CSs could bill participants’ health insurance for clinically indicated evaluations, procedures and tests, and use grant funds for underinsured or uninsured participants. The CSs were also required to provide participant transportation and lodging/meals during the one-week stay at the CS. These practices would ensure that participants did not incur out-of-pocket expenses and enable all subjects to have access to the UDN, irrespective of their health insurance status. This would also allow all participants the same experience as at the NIH UDP with no out-of-pocket expenses. However, while establishing billing procedures at the six CSs outside of the NIH UDP, it became evident that there were several challenges to implementing these practices. All the CSs were told by institutional representatives that insurance co-pays and deductibles could not be waived or reimbursed by the grant or institutional funds, due to a federal anti-inducement law that is framed for Medicare and Medicaid but is often applied to other insurance policies {42 U.S.C. § 1320a-7a(i)(6)}. Two sites were told by institutional representatives that they could not both bill insurance and pay for participant travel/lodging due to a federal anti-kickback law {42 U.S.C. § 1320a-7b}.

This led to the formation of a Billing Working Group to resolve the issues so as to not place the CSs and the participants at an undue disadvantage. After considering the legalities and the available choices, two billing options were created. The first option utilizes grant funds solely to cover all the participant evaluations, made feasible by institutional discounts (~80%) for participant care performed as part of NIH-funded studies. The second option bills the participant’s insurance companies for the clinical evaluations and cover underinsured/uninsured participants or tests not reimbursed by insurance with grant or institutional funds. Each CS can choose which option is best based on their institutional policies. In addition, each CS can re-evaluate and change to the other option based on their institutional policies.

To enable payment of co-pays and deductibles at sites that would bill insurance, the UDN is collaborating with the National Organization for Rare Disorders (NORD, https://rarediseases.org/). NORD has established a UDN patient assistance fund with contributions from the Running for Rare Diseases team. Each CS, including the NIH UDP, will receive $5,000 of the NORD fund (total: $35,000) to pay for tests needed for participants before being accepted into the UDN. This amount is utilized at the discretion of the CSs. Each CS will also receive $5000 of the NORD fund (total: $35,000) to pay for record retrieval through PicnicHealth. The remaining NORD fund is utilized by the three CSs that bill insurance to reimburse co-pays and deductibles for financially stressed participants (defined as those with an income below 300% of federal poverty guidelines). This plan would allow for participants at each of these three CSs to be reimbursed $6600 each, the maximum out-of-pocket expenses limit for an individual health insurance plan, as outlined by the Affordable Care Act (www.healthcare.gov).

Similarly, a collaboration with Mercy Medical Angels (http://mercymedical.org/) allows for provision of commercial air travel expenses for participants who are financially stressed. A memo of understanding has been signed between the UDN and Mercy Medical Angels. Each CS will decide if and when they want to use Mercy Medical Angels to arrange travel for the participant and one caretaker meeting the financial criteria (defined as having an income below 300% of the federal poverty guidelines). The CS will provide documentation of financial need and notify Mercy Medical Angels at least two weeks prior to the date of travel to allow sufficient time for them to make the travel arrangements. The CSs will pay Mercy Medical Angels a $200 per ticket administrative fee from their grant funds and Mercy Medical Angels will arrange the travel for these participants and their family member.

The Billing Working Group continues to review issues that arise during the project, and significant changes to the billing structures and/or NORD fund distributions will be presented before the Steering Committee for consideration.

C. Post-evaluation Activities and Follow-up

Transitions of care

Transitions of care programs are designed to promote the safe and timely passage of patients between levels of health care and across care settings. In the context of patient management, suboptimal transitions of care may result in: readmissions, adverse drug events, use of higher-intensity setting of care, decreased functional status, reduced quality of life, unnecessary repetition of tests or procedures, avoidable costs, and/or additional stress on patients, families, and caregivers. Suboptimal transitions of care are a risk for participants undergoing diagnostic evaluations performed by the UDN due to the potential for poor communication between providers; inadequate participant, family, and/or caregiver understanding of findings and follow-up needs of the participant; incomplete diagnostic evaluations; and a lack of clear understanding of which results are returned and which are pending. The purpose of the transitions of care plan within the UDN is to avoid these outcomes.

Best practices at the CSs for transitions of care include providing a written summary of the diagnostic work up to the family upon departure (see section “Clinical evaluation wrap-up”), confirming that probands have made it home safely (via text message, email, or phone call), being available to families and caregivers in case any clinical issues arise, and maintaining open lines of communication.

In order to facilitate communication after the UDN site visit, the CC conducts a series of participant and family surveys.

Obtaining participant feedback

The CC and/or other UDN investigators remain in contact with participants and families after discharge from the CS, which may include contacting participants shortly after the visit to assess satisfaction with the UDN visit and understanding of recommendations, as well as contacting participants periodically to assess clinical and research status. Survey instruments and interview guides were created, vetted by the CC and Genetic Counseling and Testing Working Group, and approved by the CIRB.

Survey summary

Time point	Survey administered
Post-visit (ideally 7-14 days after the in-person evaluation ended)	Care outcomes (Interpersonal Processes of Care (IPC) Survey – modified)
6 months post-visit	Participant-specific clinical and research outcomes (6-month version)
Yearly post-visit	Participant-specific clinical and research outcomes (yearly version)

The immediate post-visit period is a delicate time for participants and families, but is also a unique opportunity to measure strengths and weaknesses of the UDN site visits. A participant and family satisfaction survey after the visit allows the CC to directly assess the participant experience at the CSs.

The CC measures aspects of the participant and family experience with the Interpersonal Processes of Care (IPC) Survey[1-3]. The IPC Survey is a patient and family reported, multidimensional instrument appropriate for individuals from diverse racial and ethnic groups. The survey assesses several sub-domains of communication, patient-centered decision making, and interpersonal style. It is based on data from a sample of over 1,600 adult patients of diverse racial and ethnic background (African Americans, English- and Spanish-speaking Latinos, non-Latino Whites) and was developed to allow reliable, valid, and unbiased comparisons across these groups. The survey was also created to describe disparities in interpersonal care, predict patient outcomes, and examine outcomes of quality improvement efforts to reduce health care disparities. In addition, open-ended questions are asked to assess how the visit went, what went well and what did not, and areas for improvement.

The 6-month and yearly post-visit surveys are participant-specific clinical and research outcomes surveys. The timing of these surveys was chosen for the following reasons: 1) accounts for possible delays in receiving final test results; 2) enough time has passed to assess new treatments; and 3) any problems with the transitions of care to the home team can be identified. At these time points, new topics are introduced, such as whether the investigators discussed research options. These surveys also address the “coordination between health settings” element of the participant experience in more detail.

The CC typically calls the participant or the participant’s representative, administers the IPC Survey, and asks the open-ended questions. Data is collected in real-time by the CC while on the call. At each time point, if the participant or participant’s representative is not reached via phone, an email is sent to collect the data via email. It is important for the person conducting the surveys to check the Gateway to confirm the participant is living. Surveys are not administered to representatives of participants that are deceased. If survey respondents have specific questions or areas that require follow-up, the CC reaches out to the participant’s CS.

Participant Engagement and Empowerment Resource (PEER)

Overview

The purpose of the UDN Participant Engagement and Empowerment Resource (PEER) is to provide the participant and family perspective on UDN research goals and participant experience. The PEER engages with UDN investigators in the development and assessment of participant oriented materials and identifies best practices for receiving participant input in

research. There are many potential benefits of the PEER, including fostering longstanding relationships among participants, families, and researchers; promoting participant empowerment; educating participants and families on the essential dual clinical/research mission of the UDN and similar initiatives; and encouraging future engagement in research studies.

PEER membership

The PEER includes up to 11 participants or legal guardians who are interested in partnering with UDN investigators to assess and contribute to the UDN research process and participant experience. Members of the PEER must be willing to engage in thoughtful conversation about the positive and negative aspects of the research process and respect the perspectives of others. Ideally there is adult, adolescent, and pediatric participant and family member representation.

All UDN participants or legal guardians are offered the opportunity to participate in the PEER following the Transitions of Care survey. The application form is sent to all individuals interested in participating. Applications received are sent to the coordinators of clinical site who know the participant or legal guardian for review. Based on the coordinator review, a small group of individuals are selected by the CC and the CC drafts a proposed member list. The proposed member list is then submitted to the UDN Executive Committee and ratified by the UDN Steering Committee.

Additionally, there are three *ex-officio* members: a CC representative, a Site Operations Working Group representative, and a Genetic Counseling and Testing Working Group representative.

Terms are one year in duration and renewable. Members can reapply for a second term.

Activities

The PEER is responsible for their structure and activities. Activities may include:

1. Providing input regarding various research questions, eligibility criteria, and recruitment and informed consent processes
2. Identifying unmet participant needs
3. Contributing perspectives on risk/benefits of research project
4. Connecting families with one another and to support groups
5. Collecting participant and family experiences with the UDN from participants
6. Providing support for families when they are visiting a site far from home
7. Being a resource for families who have questions or concerns
8. Developing educational materials
9. Organizing participant conferences
10. Leading awareness efforts

Schedule and travel reimbursement

Conference calls are organized by the CC at a frequency determined by the PEER. Annual in-person meetings are held. A PEER co-chair or other PEER members may be invited to attend one in-person Steering Committee meeting annually. Travel expenses are covered for the

annual in-person PEER meeting and for a PEER member to attend the Steering Committee meeting, as needed.

Evaluation

The CC may develop satisfaction surveys to be completed on a regular basis by PEER members. The PEER are asked to develop an annual report on activities to be shared with the Steering Committee. There is an annual Steering Committee review of the PEER.

Transitions to basic science

In most cases, the transition of participant data and/or samples to basic science occurs through the CS to which the participant was assigned, or through cross-site UDN collaborations. There are referrals to the basic science community based on candidate genes and other diagnostic information, and it is important to track these activities in the Gateway.

Participant web pages project

Sequencing has the potential of revealing variants of uncertain significance. The discovery of these types of variants leaves patients in a state of uncertainty, faced with the possibility of never knowing the cause of their symptoms. Finding another patient with the same, or similar, variant can reduce or eliminate this uncertainty and end the diagnostic odyssey. It can take years before enough patients with the same or similar variant are identified to confirm a diagnosis. By using web pages to identify second cases, the UDN hopes to shorten the diagnostic process for participants, generate new knowledge of rare conditions, and connect patients and families to one another for support.

All participants that are undiagnosed or diagnosed with rare conditions are asked if they would like to partake in this project. If a participant is interested in participating, the CS contacts the CC and the CC consents the participant. Consent is obtained over the phone or by videoconferencing (or in person if reasonable) from the enrolled participant, parent, or guardian. Assent is obtained over the phone (or in person if reasonable) in the presence of a parent or guardian for all children ages 7-17 years old who are not decisionally impaired. The consent and assent forms are sent (by mail or email) to the participant (or legal guardian) prior to the remote or in-person consent. The principal investigator, an associate investigator, a genetic counselor, or a project coordinator trained in consenting at the CC is available to answer questions and obtain consent. The signed consent and assent forms are sent to the CC. The CC provides pre-paid, self-addressed envelopes to mail these forms. The CC records the consent and assent versions signed by the participant in the Gateway.

A web page about the participant is created by the CC based on information in the Gateway. The web page is reviewed by the CS and participant or parent/guardian. The web pages may include the following:

1. Genetic variants (including gene name(s) and variant)
2. Symptoms/signs, both past and present, using both Human Phenotype Ontology (HPO) terms and simplified language
3. Past medical history (including suspected diagnoses with dates or on a timeline)
4. Past and current treatments, procedures, and medications and their outcome and effectiveness (treatments, procedures, and medications with dates or on a timeline)
5. Gender, race, and ethnicity

6. Pictures and/or videos, if participant consents (pictures at multiple time points would be ideal)
7. Contact information (phone number, email address) for CC

Google alerts for all gene names of interest, specific variants of interest, and rare symptoms are created. The CC utilizes search engine optimization techniques to ensure that the pages are found.

When searching for symptoms or genetic variants, clinicians and families may find a UDN participant web page. The CC triages and records contacts made regarding these web pages. For possible matches, additional information is collected. The CC facilitates conversations between the potential match or clinician, the volunteer CS, and the participant. Participants are asked if they would like to communicate with each potential match or clinician before they are connected. If true matches are made, the CC updates the web pages accordingly. The CC records the amount of time spent processing each inquiry, how many participants receive diagnoses as a result of the web pages, and the length of time to each diagnosis. The UDN collaborates with the UDN patient advisors and participants to further refine this process and foster the creation of rare disease patient communities.

D. References

1. Napoles AM, Gregorich SE, Santoyo-Olsson J, O'Brien H, Stewart AL. Interpersonal processes of care and patient satisfaction: do associations differ by race, ethnicity, and language? *Health Serv Res.* 2009;44(4):1326-1344. doi: 1310.1111/j.1475-6773.2009.00965.x. Epub 02009 Mar 00931.
2. Stewart AL, Napoles-Springer A, Perez-Stable EJ. Interpersonal processes of care in diverse populations. *Milbank Q.* 1999;1999;77(3):305-339.
3. Stewart AL, Napoles-Springer AM, Gregorich SE, Santoyo-Olsson J. Interpersonal processes of care survey: patient-reported measures for diverse groups. *Health Serv Res.* 2007;42(3 Pt 1):1235-1256.

IX. Surveys

A. Background

Surveys offer an opportunity to learn from the unique participant population with rare and undiagnosed diseases. The purpose of this section is to outline the survey implementation approach for the network.

B. Survey Committee

Purpose

The Survey Committee ensures that surveys are consistent with the UDN mission and that participants are not overburdened. The Survey Committee is under the umbrella of the Publications and Research Committee (PRC).

Members

The Survey Committee includes:

1. One voting member from each CS (7 members)
2. One voting member from the CC
3. At least 1 survey methodologist from a UDN site. The survey methodologist could also be the voting member from their site.
4. One member from the Clinical Protocols Working Group, who acts as a liaison
5. Optional: One voting member from each Core

Ad hoc members:

1. Members (1-2) from any working group impacted by a survey under review are asked to join the Committee as ad hoc members.
2. UDN members with expertise in an area covered by a survey under review may be asked to join the Committee as ad hoc members.

Tasks

The Survey Committee reviews all surveys proposed after the initiation of the Committee. For survey research using the CIRB, the Survey Committee makes a recommendation regarding whether or not the survey should be integrated into the UDN. For survey research not using the CIRB, the Survey Committee reviews the survey in order to have a complete picture of surveys being administered to UDN participants, but does not issue recommendations.

Each survey using the CIRB, including those approved prior to the initiation of the Survey Committee, are reviewed once a year, starting one year after IRB approval was obtained. A subgroup of the Survey Committee (2-3 members), and ad hoc members as appropriate, are tasked with reviewing the survey performance characteristics and make a recommendation to the Survey Committee regarding survey continuation. The Survey Committee decides who the reviewers are for each survey, with a goal of dividing up the surveys among the Committee members. Surveys not using the CIRB are not reviewed yearly, but are requested to inform the Survey Committee annually as to the status of the research (completed; ongoing; other).

C. Survey Development

Creation of survey plan

The group proposing a survey must formulate a plan, which should include:

1. Purpose of survey, hypotheses, and questions answered by the survey
2. Instrument that will be used and validation of that instrument
3. UDN sample population (what sites will be involved; what populations within the UDN sites will be surveyed, e.g., all members or a subset)
4. Plans for administering the survey to non-English speaking individuals
5. Administration timeline
6. Method of survey administration (online, phone, etc.)
7. Text for all correspondence with participants
8. Individuals administering survey
9. Metadata from the Gateway that should be associated with survey (e.g., dates of visit, demographics, etc.)
10. Where data will be kept (Gateway or separate from Gateway) and who should have access to the data
11. Analysis plan, including who will do the analysis
12. Requirement for and availability of additional funds to conduct and analyze the survey

Plan proposal and approval

The group proposing a survey must complete and submit a research concept sheet. Research concept sheets are submitted to the CC and sent directly to the Survey Committee for review.

<u>The rest of this section is ONLY applicable to surveys using the CIRB</u>

After receiving a research concept sheet, the Survey Committee chair informs working groups impacted by the survey, UDN members with expertise in an area covered by the survey, and the Steering Committee of the survey review date. Updates to the Steering Committee are provided during the working group updates segment of the meetings. Applicable working groups and experts are offered the opportunity to join the review as ad hoc members.

At least one week prior to a review, the Survey Committee chair distributes the research concept sheet to Committee members, including ad hoc members. The Survey Committee documents their deliberations and recommendations in meeting minutes.

Working group ad hoc members report back to their groups within 2 weeks regarding the Survey Committee recommendations. The Survey Committee provides the working groups with: 1) the research concept sheet and 2) meeting minutes. The ad hoc members bring questions and feedback from the working groups back to the Survey Committee for discussion. Based on this discussion, the Survey Committee finalizes their recommendation.

The Survey Committee may not come to a final recommendation and recommend that the research concept sheet be revised. In this case, the group proposing the survey provides the revised concept sheet to the Survey Committee. The Survey Committee votes on a final recommendation, with a majority vote needed to decide on an outcome.

The Survey Committee sends their final recommendation to the PRC for review, expected to be within 3 weeks of the review. The Survey Committee provides the PRC with: 1) the research concept sheet and 2) meeting minutes. Following PRC review, the CC sends the outcome of the review to the chair of the Survey Committee, chair(s) of the working groups impacted by the survey, members of the proposing group, and the CC. The CC adds the survey review to an upcoming Steering Committee meeting agenda, expected to be within 1 month of the completion of the PRC review. New surveys are presented to the Steering Committee by the proposing group for final consideration on the assigned date.

The Steering Committee may approve the survey research, raise concerns and send the research concept sheet back to the group for revision, or reject the survey research.

IRB amendment submission

If the survey research is approved by the Steering Committee, the group proposing the survey must complete an amendment form detailing the survey plan and submit the form to the CIRB Liaison (see Section III. Institutional Review Board (IRB) Communications). The amendment must include any required changes to the UDN protocol, the survey plan, a paper version of the survey, and all language for communications with participants.

If the survey is not integrated into the Gateway, the group can start administering the survey once IRB approval has been obtained.

Survey tracking

The Survey Committee keeps a record of all UDN surveys. For those using the CIRB, the table includes the date when the survey was approved by the CIRB and the date of the yearly review.

Gateway integration

If the survey needs to be integrated into the Gateway, the UDN group should contact the CC.

Ongoing review

Surveys are reviewed once a year by a subgroup of the Survey Committee to determine if the survey should continue to be administered. The Survey Committee contacts the group conducting the survey 3 months prior to the annual review date. The Annual Survey Review Worksheet (Appendix 9: Annual Survey Review Worksheet) should be submitted to the chair of the Survey Committee on or before the review date.

After receiving an annual survey review worksheet, the Survey Committee chair informs working groups impacted by the survey, UDN members with expertise in an area covered by the survey, and the Steering Committee of the survey review date. Updates to the Steering Committee are provided on at least a monthly basis. Applicable working groups and experts are offered the opportunity to join the review.

At least one week prior to a review, the Survey Committee chair distributes the Annual Survey Review Worksheet to subgroup members, including ad hoc members. The subgroup documents their deliberations and recommendations in meeting minutes.

Working group ad hoc members report back to their groups within 2 weeks regarding the subgroup recommendations. The subgroup provides the working groups with: 1) the Annual Survey Review Worksheet and 2) meeting minutes. The ad hoc members bring questions and feedback from the working groups back to the subgroup for discussion. Based on this discussion, the subgroup finalizes their recommendation. The subgroup may not come to a final recommendation and may recommend that the Annual Survey Review Worksheet be revised. In this case, the group conducting the survey provides the Annual Survey Review Worksheet to the Survey Committee before a final recommendation is made.

The Survey Committee votes on final recommendations, with a majority vote needed to decide on an outcome. Surveys recommended for continuation or discontinuation by the Survey Committee are sent to the PRC for review. The Survey Committee provides the PRC with: 1) the Annual Survey Review Worksheet and 2) meeting minutes. Following PRC review, the CC sends the outcome of the review to the chair of the Survey Committee, chair(s) of the working groups impacted by the survey, and members of the conducting group. Survey continuations and discontinuations are presented to the Steering Committee during the working group updates segment at the next meeting. The Steering Committee may approve the continuation of the survey research, raise concerns and send the Annual Survey Review Worksheet back to the group for revision, or reject continuation of the survey research.

Significant survey modifications

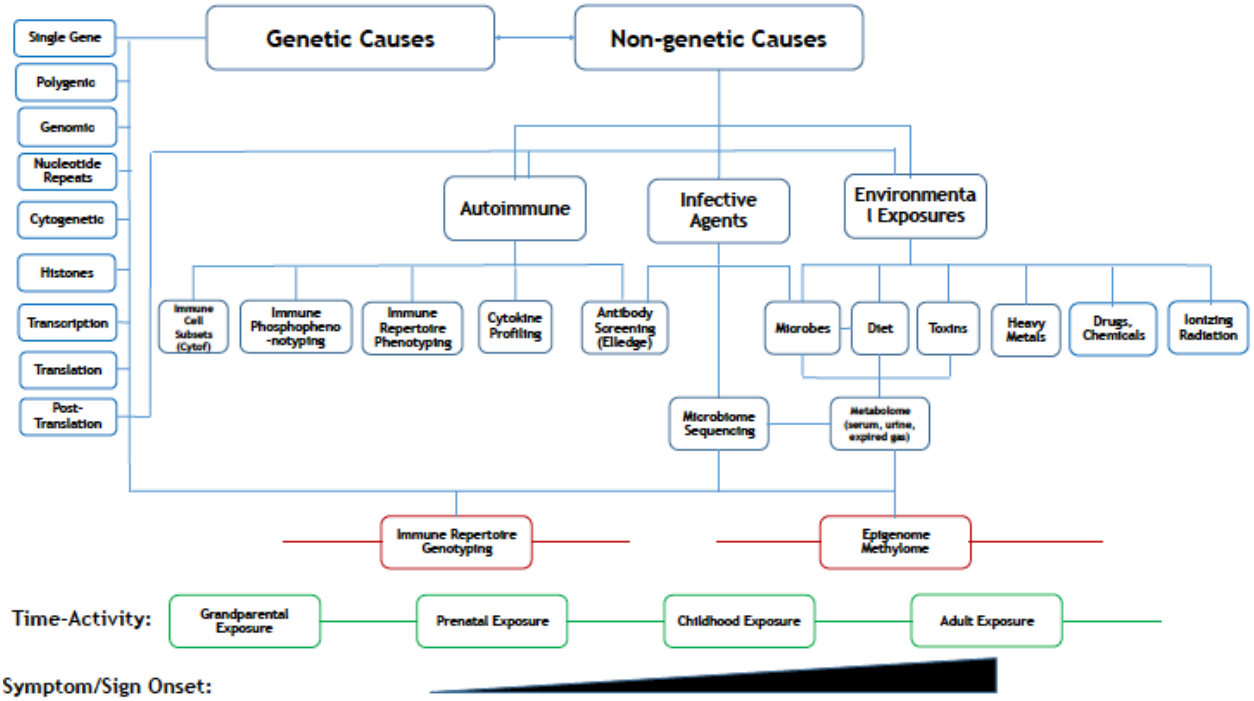
If significant changes are made to a survey (more than just minor edits), the group responsible for the survey submits the revised survey to the Survey Committee for review. The Survey Committee chair reviews the changes to determine if they require review by the Survey Committee. If the Survey Committee chair feels that the changes require Committee review, the chair informs the Committee, working groups impacted by the survey, UDN members with expertise in an area covered by the survey, and the Steering Committee, and schedules a survey review date. If there are no concerns, the review decision is expedited and the decision is reported to the Steering Committee during the working group updates.

X. Environmental Exposures

A. Background

The UDN fully recognizes that all diseases are likely determined by many factors, including "minor modifying" genes and environment, through molecular mechanisms such as by epigenetics, transcriptomics, post-translational modification of the proteome, the microbiome, infections, and immune responses. The UDN protocol for identifying possible environmental factors, outlined in the figure below, is intended to systematically screen all UDN applicants for such factors, expedite the efforts of clinical investigators, and perhaps improve the participants' health by intervening against current environmental hazards.

Figure: A holistic view of the multifactorial, ecogenetic hypothesis of disease etiology.



In practice, this protocol begins with a screening of environmental exposures (exposome) via a self-administered survey completed by the UDN applicant (or their relatives). Environmental Exposures Working Group members with expertise in environmental and occupational medicine and teratology comment on responses that may deserve possible exploration by the UDN CS investigators. Recommendations may include further clinical evaluation by an Environmental and Occupational Medicine physician, a teratologist, or appropriate laboratory assays of drugs and environmental toxins, imprinting, RNA sequencing, immune response characterization, and possible infections.

B. The Environmental Exposures Working Group

Purpose

The Environmental Exposures Working Group facilitates: 1) the screening of UDN applicants for self-reported environmental exposures, 2) the interpretation of the screening questionnaire results, 3) recommendations for further evaluations of putative exposures by targeted clinical and laboratory investigations to corroborate the exposure(s), and 4) the preparation of presentations, abstracts, and manuscripts about the data related to environmental exposures. The Environmental Exposures Working Group follows the procedures of all Working Groups, as detailed elsewhere in this Manual of Operations.

Part of the consultation service of the Environmental Exposures Working Group is to facilitate the use of validated laboratory assays on the best specimen type to aid the UDN clinicians. It may invite laboratory consultations and/or assays, in coordination with the Metabolomics Core. Biospecimen handling follows procedures of this Manual, although special handling of rare sample types (teeth, hair, nails, etc.) and special specimen containers may supersede general recommendations.

The Environmental Exposures Working Group initiates aggregate analyses of the results of environmental exposure assessments as reported on the questionnaire, especially when corroborated by laboratory or other independent data. The procedures of the Publications and Research Committee are followed.

Members

The Environmental Exposures Working Group includes two co-chairs, to be approved by the UDN Steering Committee. Decisions are reached by consensus and are only advisory, so no distinction is made between voting and non-voting members. Other members include:

1. Members from any of the CSs
2. One member from the CC
3. Optional: One member from each Core
4. Optional ad hoc members: Up to three other interested parties, preferably with expertise in occupational and environmental medicine or teratology

C. Environmental Exposures Questionnaire

Current and previous versions of the Environmental Exposures Questionnaire are available (https://hmsharvard.box.com/v/envsurvey).

The initial application had two survey questions about any possible environmental factor related to the illness and, for affected children, any suspected preconceptual and prenatal exposures. Subsequently, a longer environmental exposures questionnaire was developed by an environmental exposures ad hoc group, which included experts in occupational and environmental medicine, teratology, toxicology, epigenetics, and genetic epidemiology. This questionnaire was designed for self-completion by participants or parents/guardians and was implemented with electronic data capture by the CC using Qualtrics. To create the questionnaire, existing questionnaires were leveraged including the American Academy of Clinical Toxicology Exposure History Form,[1] CDC Environmental Exposure History,[2] PhenX project,[3] and NHANES,[4] in addition to questions developed with the assistance of various experts in multiple environmental exposures.

The areas addressed by the questionnaire include:

a) occupational exposures to dust and chemicals,
b) residential exposures to chemicals and pollution,
c) use of alcohol, tobacco, and other drugs (prescribed and recreational),
d) exposures from hobbies and travel,
e) exposures to infectious disease vectors and sexually transmitted diseases,
f) unusual dietary histories, and
g) parental environmental and occupational exposures before and during pregnancy that resulted in the birth of the participant.

This questionnaire is intended to be used as a clinical screening tool to identify applicants who would potentially benefit from a more intensive environmental exposure assessment at the UDN CSs. It was initially administered to accepted participants, but is to be administered to all applicants in the future. It is expected that revisions to this questionnaire will be made over time. After a year's experience, a shortened version was implemented.

Members of the Environmental Exposures Working Group review questionnaires during monthly meetings. The group reaches consensus on either making no recommendation for UDN CS clinicians or making specific recommendations to be implemented at the discretion of the clinicians. The Environmental Exposures Working Group communicates these recommendations to the CSs. Typically these recommendations include evaluation by an Environmental/Occupational physician, a teratologist, additional laboratory testing, or collection of additional clinical history; sample advisory comments are below.

Dear Dr. X,

The UDN Environmental Exposures Working Group has reviewed additional Environmental Exposures Questionnaires that have been completed by UDN patients at the XX Clinical Site.

The following patient accepted by your site had potentially relevant environmental exposures identified from the Environmental Exposures Questionnaire:

UDNXXX: This patient reported exposures to xxx. Therefore, we recommend an Environmental/Occupational Medicine evaluation.

Please note that we did not evaluate the full medical record for this subject, and we are basing this recommendation on the limited information available to us. Please use your judgement to determine whether this recommendation is appropriate. If you have any questions, we would be happy to discuss them.

Thanks very much for your help.
UDN Environmental Exposures Working Group

The Environmental Exposures Working Group may, on infrequent occasions, present its rationale for recommending acceptance of an applicant mostly because of a very compelling environmental history of possible etiologic importance, one that is addressable by UDN clinicians and laboratory investigators. Guidance may be offered in preferring an alternate CS to be used rather than the initially assigned one, on the basis of expertise to address the alleged exposure and access to laboratory approaches to investigate the putative pathogenicity. Any time CS reviewers become aware of an environmental exposure of special concern, they can

request urgent review by the Environmental Exposures Working Group, which strives to provide an interpretation within one week.

D. References

1. American Academy of Clinical Toxicology Exposure History Form http://www.clintox.org/sections/occupationalenvironmental
2. Agency for Toxic Substances and Disease Registry. Case Studies in Environmental Medicine (CSEM): Taking an Exposure History. https://www.atsdr.cdc.gov/csem/exphistory/docs/exposure_history.pdf
3. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, Hammond JA, Huggins W, Jackman D, Pan H, Nettles DS, Beaty TH, Farrer LA, Kraft P, Marazita ML, Ordovas JM, Pato CN, Spitz MR, Wagener D, Williams M, Junkins HA, Harlan WR, Ramos EM, Haines J. The PhenX Toolkit: Get the most from your measures. *Am J Epidemiol.* 2011;174(3):253-60. PMID 21749974..
4. Centers for Disease Control and Prevention, Nagtional Center for Health Statistics, National Health and Nutrition Examination Survey. http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm.

XI. Sequencing

A. Flow of Samples to SC

I. Sample collection and DNA extraction

1. CSs arrange for blood sample collection before or during the clinical evaluation. If collected off-site, blood samples for DNA should be shipped to the CS.
2. Additional sample collection considerations:
    a. If the individual has had a bone marrow transplant or recent blood transfusion, DNA from fibroblasts is preferred.
    b. Samples from deceased individuals should be discussed with the SCs on a case-by-case basis.
3. The CS arranges DNA extraction and QC. DNA samples submitted for sequencing should meet the following conditions:
    a. WES: at least 6ug of 50-200ng/ul DNA
    b. WGS: at least 10ug of 50-200ng/ul DNA
4. Additional DNA is stored at the CS with other biospecimens collected during the clinical evaluation.

II. Sample shipping guidelines

1. CS prepares DNA samples for shipment to the appropriate SC. DNA samples should be sent as complete families (including all family members that will be included in the analysis) excepting clinically urgent samples that warrant prioritized sequencing. Urgency is at the discretion of the CS.
2. CS completes a sequencing request form in the Gateway for each DNA sample being sent for sequencing.
3. CS enters and releases updated phenotype information (i.e., PhenoTips) in the Gateway for use by the SC in their analyses.
4. CS enters shipping information (date DNA sent and tracking number) in the Gateway and ships samples.
    a. Please note that the Gateway provides alerts for shipping of UDN samples, but shipment tracking needs to occur at the CS/SC level.
5. Gateway sends an automated email to alert the appropriate SC of sample shipment and available phenotypic data.
6. SC acknowledges receipt of samples by entering date DNA received in the Gateway.
7. If a submitted DNA sample does not pass QC at the SC or is otherwise deemed unacceptable, the SC contacts the CS site directly via phone or email to request a replacement.
8. Sample labeling discrepancies are addressed on a case-by-case basis at the discretion of each SC.

B. Flow of Clinical Information to SC

1. The CS are responsible for collection of blood specimens and DNA extraction. The CC and SCs will not know about or track the DNA specimens until they are shipped to the SCs. The CSs are encouraged to collect all specimens for a family before sending them, but additional family members may be added at a later date if necessary.

2. Typically, a CS send an aliquot of DNA extracted in a CLIA-certified lab and keep the remainder of the DNA for future procedures.
3. Samples may be sent either before or after the in-person evaluation of the proband. In either case clinical information and a pedigree, including the relationships of all submitted family members to the proband, should be added to the PhenoTips as soon as possible for samples submitted for sequencing. This information is used by the SCs for their analysis.
4. DNA samples submitted for sequencing must be labeled with participant name, date of birth, and the initials "UDN".
    a. Local identifiers may also be included at the discretion of the CS. Local identifiers will appear with participant name on sequencing reports (i.e., John Doe, UDP123).
5. Other required information for sample submission includes:
    a. Gateway consent form
    b. Gateway sequencing form
    c. Affection status of family members

C. Exome Sequencing

This section covers sample intake, library preparation, whole exome capture, and sequencing at Baylor College of Medicine. This section describes the sample flow from DNA sample receipt to production of WES data, including appropriate quality control and assurance procedures.

Sample intake

DNA samples are received at the Whole Genome Laboratory (WGL). A visual inspection of the sample tubes is conducted. Sample tube label is compared with information entered in the Gateway to ensure consistency and completeness of the Gateway data, consent forms, proper sample labeling, and sample tube integrity. Samples are be accepted if no discrepancies are found, sample labels match, and no tube damage is observed. If any of the above criteria is not met, Baylor notifies the referring CS.

Once accepted, samples are accessioned into the WGL Laboratory Information Management System (LIMS) system. Sample information in the Gateway is entered into the WGL LIMS database. Each sample is assigned an internal six-digit lab number, as well as a six-digit family number in LIMS. 1D bar code labels with participant specific information (unique identifier) including participant name, date of birth (DOB), lab number and family number are attached to the stock DNA tube. Subsequently samples are aliquoted from the stock tubes into 2D barcode tubes. The samples in 2D bar code tubes are processed for exome sequencing. Before sample transfer, the record for a sample is first opened in LIMS, then the 1D barcode label on the stock tube of the sample is scanned and the LIMS automatically verifies if the sample ID in LIMS matches that on the label. Then, the 2D bar code on the aliquot tube is also scanned to link the two bar codes in LIMS before sample transferring occurs. These steps are to ensure the chain of custody remains intact during sample transfer.

DNA sample QC

DNA samples are screened to quantify DNA as well as determine DNA quality. To determine DNA concentration and purity, the samples are evaluated using the Quant-iT PicoGreen dsDNA assay on the BioTek Synergy 2 microplate reader. Passing criteria include:

1. The R-squared value for the standard curve must be ≥99.9%DNA concentration
2. Sample contains 1ug of DNA

To verify DNA integrity and relative size, the same dilution of sample is loaded on a 0.8% E-Gel. Passing criteria include:

1. Gel image is clear and shows no DNA degradation

If a sample does not meet the criteria above, the CS is notified.

Pre-capture library preparation

In order to meet UDN sequencing objectives, we use our quick whole exome sequencing (QWES) protocol. QWES is an optimized version of the standard Illumina (ILM) library preparation workflow that reduces library construction time to 5-6 hours.

Library construction is a completely automated process on the Span-8 Biomek NXP with an incorporated LIMS tracking system. Before starting library preparation, all primers and adapters lots are validated and the appropriate dilutions are prepared. Negative (H_2O) and reagent blank controls are included Robot operator closely monitors each transfer step. Pre-capture library preparation involves the following steps:

Normalization and shearing

DNA samples are normalized to 750 ng total. Samples are loaded into Covaris microtubes in 50 ul aliquots and sheared to approximately 250-500 bp using the Covaris E220 ultrasonicator. Shearing efficiency is assessed using a 2.2% flash gel. Fragments should range from 100-600bp with average of 250-500bp. If the majority of sheared fragments are larger than 800 bp, the sample is re-sheared.

End repair

Fragmented DNA samples are treated with NEBNext® END REPAIR Module (catalog#: E6050L) at 20°C for 20 minutes to make blunt ended DNA. Then 1.8X Beckman SPRI beads (Agencourt AMPure XP Solid Phase Reversible Immobilization magnetic beads) and 70% Ethanol are used for cleanup. Treated fragmented samples are eluted with 40 ul elution buffer while SPRI beads remaining in the solution.

3’ Adenylation

The treated DNA samples are incubated with NEBNext® dA-Tailing Module (catalog#: E6053L) at 37°C for 20 minutes to incorporate a non-templated dAMP on the 3’ end of a DNA fragment. The binding buffer (BB) made with Polyethylene Glycol (PEG) and 5 M Sodium Chloride (NaCl) (final concentration of PEG and NaCl: 20% and 2.5 M respectively) is applied to help dA-Tailing fragment DNA binding back to SPRI beads while the rest of solution is discard. 70% ethanol then is used for cleaning up the DNA bounded SPRI beads. DNA samples again are eluted with 40 ul elution buffer while SPRI beads remaining in the solution.

Ligation

Post dA-Tailing, DNA Samples are ligated with Illumina multiplexing paired-end (PE) adapters by using Invitrogen Expresslink ligase (catalog#: A13726101) and buffer at room temperature for 20 minutes. The same binding buffer is used to allow ligated DNA binding back to SPRI beads and 70% ethanol is used for cleaning up the SPRI beads. SPRI beads are removed post 40 ul elution buffer added.

Enrichment

Ligated DNA samples are enriched for total 6 cycles with 2X KAPA HiFi HotStart Ready Mix PCR kit (catalog#: KK2612) and Illumina PE PCR primers. AB GeneAmp PCR System 9700/Veriti are used for amplification Enrichment.

Post-enrichment QC

The enrichment PCR efficiency analysis is performed on 2.2% FlashGel by checking the product intensity. The FlashGel analysis is preformed after 6 cycles and is re-run if the band is too weak. Additional PCR cycles can be added for samples with low yield.

1. No more than 9 cycles total can be run for samples
2. If the amount of the post-PCR product is insufficient after a total of 9 cycles, the whole process needs to be repeated.
3. The negative (H_2O) and reagents blank control should give no product

To check size distribution and quantify the final library, the sample is run on an Agilent Bioanalyzer 2100 DNA 7500 Chip.

1. The library sizes should range from 200 – 750 bp (Majority are 250-550bp) with the peak ranging from 250 - 350 bp.
2. The yield of library should be more than 1.5 ug. If pre-capture library yield is lower than 1ug, the library preparation is repeated. No adaptor dimer and free primers are visible.
3. See Figure 1 for an example of a passing pre-capture library.

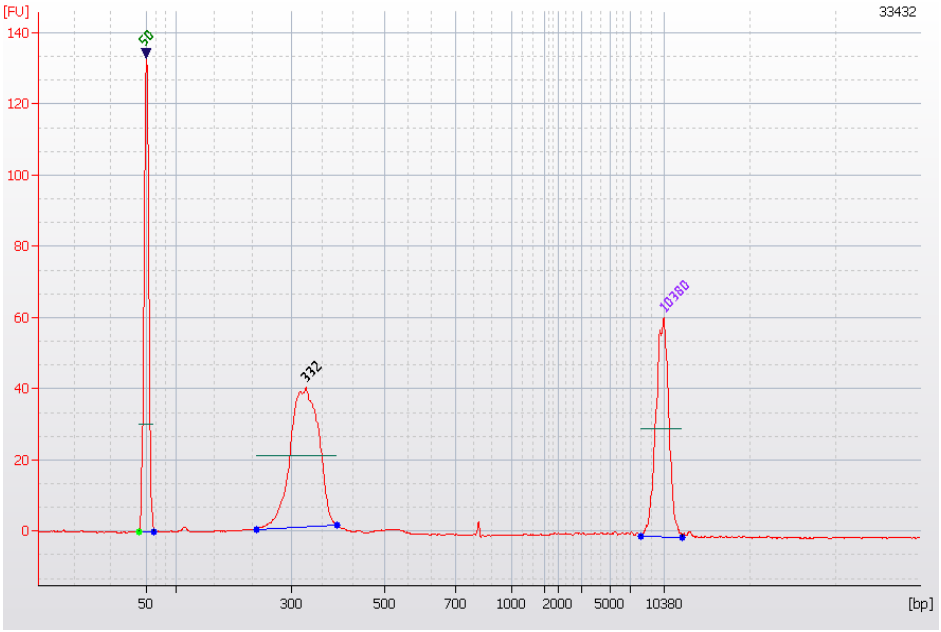


Figure 1: Illumina paired end DNA library on Agilent Bioanalyzer 2100

Whole exome capture

The whole exome capture utilizes the NimbleGen liquid capture on HGSC VCRome 2.1, that targets approximately 34Mbp of genomic DNA including all coding exons of currently known disease genes (OMIM, HGMD, and GeneTests). To enhance the coverage of clinically relevant disease genes, the currently developed spike-in probe set (Exome 3 – PKV2) is used in 1:1.25 equimolar ratio with the VCRome exome capture design in combination with the QWES protocol.

Solution capture is initiated by combining 1.5 ug of the pre-capture library, 40ul of 1mg/ml human Cot1 DNA, and adding 0.65 ul of each 1,000uM Hybridization enhancing (HE) oligos. Full-length hybridization enhancing oligos are used to augment capture efficiency.

This mixture is dried down in a DNA vacuum concentrator on high heat setting and re-suspended in Hybridization buffer and Formamide. The mixture is denatured for 10 minutes and VCRome probe with Panel Killer V2 combined in 1:1.25 ratio are added. The mixture is incubated at 56°C for 16 hours. The following day, captured DNA is washed and recovered. Post-Capture PCR amplification is performed using KAPA HiFi HotStart DNA polymerase with total 12 cycles.

Final library QC

FlashGel

Capture efficiency is checked using a 2.2% FlashGel. If the intensity is too low, the capture process needs to be repeated. Only primer bands should be seen for negative and reagent blank controls.

Bioanalyzer

To assess size distribution and quantify the final post-capture product, libraries are loaded onto DNA 7500 Chip for assessment on the Agilent 2100 Bioanalyzer. The majority size should be around 300-400 bp with the concentration above 20 nmol/L. See Figure 2 for an example of a passing final library.

qPCR

Capture efficiency evaluated by SYBR green-based qPCR with known four loci assays. Successfully enriched capture libraries have an average delta Ct of the four loci >6 with delta CT of the individual assay is >5.

Post-capture libraries must pass all three QC checks to proceed to cluster generation and sequencing. For each sequencing run, three individual barcoded libraries are normalized to 10nM and pooled into one pool.

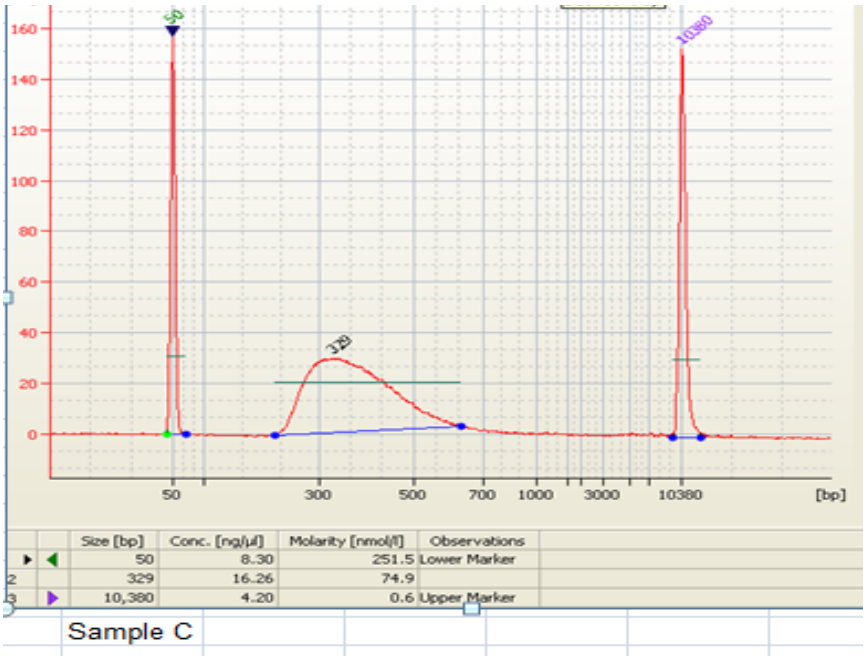


Figure 2: Final post-capture library on Agilent Bioanalyzer 2100

cBot cluster generation

Post-capture libraries are denatured prior to loading on the cBot. Denatured PhiX control is spiked-in into lanes 1 and 2 for HiSeq 2500 Rapid runs. The optimal library concentration for cluster generation is 10nM. Library and flow cell information are entered into WGL LIMS system prior to starting a cBot run.

After the cBot run is completed, each strip is checked to ensure that correct volumes were drawn for each lane. If volume of any reagent tube or DNA strip tube is not equal to the others (most frequently more volume is left), the reagent and/or DNA was not delivered properly for that lane.

Sequencing on HiSeq 2500

The HiSeq 2500 is employed for sequencing in rapid mode (27 hr cycle time) to generate 100bp paired-end reads in a format of 3 samples per lane to generate 10-12 GB per sample. Target coverage for proband and parental samples is >100x. The WGS LIMS system is utilized to track the run set up, status and quality metrics. Pertinent metrics and passing thresholds are provided in the tables that follow.

The run stops after imaging is complete for the first cycle and generate a first base report. The first report is confirmed by manual review and the run resumes if all the metrics correlate with the determined standards.

G intensity	>6000
Cluster density	>400k/mm2

The performance of the run is monitored, and the metrics below are recorded to assess quality at a particular step of the sequencing run, evaluate library quality and concentration, detect any potential sequencing reagents and/or optical issues.

Cluster density at cycle 5	900-1100 k/mm2
Phasing/pre-phasing at 25 cycles	<0.3/0.7%
Passing filter rate	>80%
PhiX error rate	<3%
Q30	>80%

After the run is complete, a comprehensive set of post-sequencing production metrics are continuously monitored and are reviewed at weekly meetings to facilitate timely troubleshooting to maintain overall pipeline performance. Overall run performance is evaluated by metrics from the off-instrument software (Casava) and from mapping results generated by the Mercury analysis pipelines using Burrows Wheeler Aligner (BWA) software to ensure that production standards are met.

Pass filter	>80%
Aligned reads	>80%
Error rate	<4%
Unique reads	>90%

The capture analysis is incorporated in the Mercury analysis pipeline and provides metrics to gauge the overall quality of the capture process. This pipeline reports:

1. Proportion of the aligned reads that map to the targeted region, which is relative to the effective enrichment of the capture
2. Distribution of coverage across the targeted bases; specifically, the fraction of targeted bases covered at 1x; 10x, 20x, 40x

The complexity of the capture library is assessed by calculating the number of alignment reads that occur from PCR duplicates. If needed, these reads can be removed from the analysis.

Key metrics that have been developed and are reviewed in weekly meetings are presented in the table below

Reads Aligned to target	>50%
Target bases covered at >20x	>90%
Target bases covered at >40x	>80%
Mean coverage of target bases	>100x

As an additional quality control measure, samples are also analyzed by single nucleotide polymorphism (SNP) array. SNP data is compared with WES data to ensure correct sample identification and to assess sequencing quality. The data is analyzed using an automated pipeline that produces concordance and contamination scores.

Concordance	>90%
Contamination	<5%

D. Genome Sequencing

This section covers sample intake, library preparation, and whole genome sequencing at HudsonAlpha Institute for Biotechnology. This section describes the sample flow from DNA sample receipt to production of WGS data, including appropriate quality control (QC) and quality assurance procedures.

DNA samples are received by the Genomic Services Laboratory (GSL) at HudsonAlpha Institute for Biotechnology. A visual inspection of the samples and subsequent accessioning is performed by two GSL employees. Accessioning includes entering the samples into a project in the LIMS of the Clinical Services Lab (CSL) and assigning two identifiers to the samples. The first identifier is the CSL identifier and is formatted as a project number, submitter's initials, and a unique sample number incrementing up from sample 0001 (i.e., C1001-SL-0001). The second identifier is a unique identifier created by the LIMS using the date of accessioning and another digit, which indicates the order in which samples were accessioned. These identifiers deidentify the sample and are used to track the samples through all handling performed by the CSL. This accessioning is performed by the HAIB CSL laboratory manager or a designee.

If the submitted sample is DNA from a CLIA-certified laboratory, it proceeds to QC procedures as described below. If the submitted sample is blood, DNA is isolated in conjunction with its accessioning and another unique sample identifier for the DNA is created by adding a '.1" to the GSL ID (i.e., C1001-SL-0001.1). DNA is extracted from 1ml of whole blood on the QIAsymphony instrument using the Blood_1000_V7_DSP protocol. This protocol yields on average of 10-18ug of genomic DNA (gDNA) from 1 ml of whole blood.

Sample information in the Gateway is reviewed by the HAIB UDN Project Manager to ensure consistency and completeness of the Gateway sequencing and consent forms, proper sample labeling, and sample tube integrity. Coordination between the UDN HAIB Project Manager and the CSL laboratory supervisor ensure that sample labels are cross-checked with information in the Gateway. Sample tubes are visually compared to information in the Gateway for accuracy prior to accessioning by the CSL laboratory supervisor or an assignee.

Samples are accepted if no discrepancies are found, sample labels match, and no tube damage is observed. If any of the above criteria is not met, HudsonAlpha notifies the referring CS.

DNA sample QC

All DNA samples are evaluated for concentration by fluorometric assay (Qubit® or Picogreen®) and for integrity by agarose gel. Ideally, there are at least 500ng of intact, high quality DNA available to enter the GSL whole genomic sequencing (WGS) library preparation. For fragmented samples or other non-traditional DNA preparations, such as low input samples, QC may be performed with the Agilent Bioanalyzer or Caliper GX. Samples not meeting the minimum requirements for clinical WGS may still proceed into library preparation, but as research samples, not clinical.

Library preparation

DNA samples are normalized to 1000ng of DNA in 50ul of water. All gDNA samples require fragmentation in a random manner to create the fragments that become the inserts in the final

library. The Covaris L-series and E-series instruments are used to shear the DNA to a final insert size of ~350bp. This longer insert size improves overall library performance and allows the longer sequencing read lengths on the Illumina HiSeq X platform (150bp) to be efficiently used without producing a significant number of over-lapping reads. QC is performed after sonication to ensure that yield and fragment size are within expected ranges. Library preparation is then performed using a proprietary GSL methodology with key QC steps performed throughout. QC is performed during preparation, after ligation, to assess yield. After the library preparation is complete, final library yield, fragment size, and fragment distribution are measured. Finally, real-time PCR quantitation to determine the molar fraction of the sequenceable library is performed. Yields are determined with fluorescent measurements (PicoGreen) and fragment sizes and distribution is determined with either the Agilent Bioanalyzer or the Caliper LabChip GX, depending on batch size. Example traces for sonicated gDNA and final sequencing library are below. The final libraries are diluted to 3 nM stocks for use in clustering and sequencing.

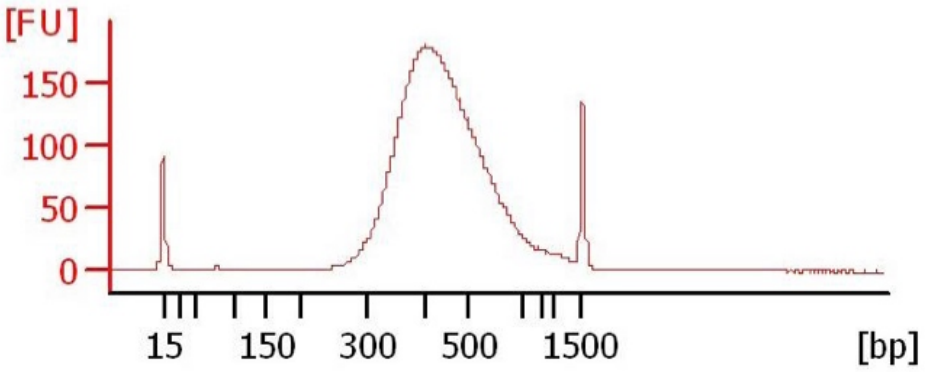


Figure 3: post sonicated gDNA QC

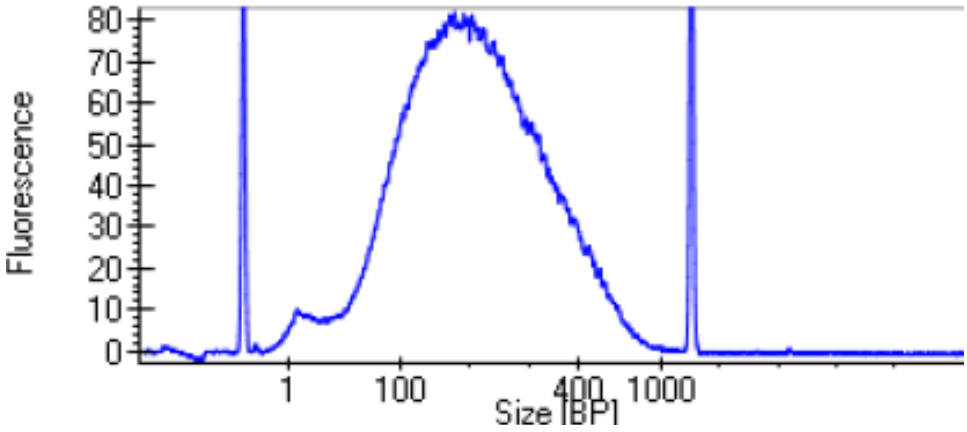


Figure 4: post library prep QC

Sequencing on HiSeq X

Once a library passes QC, the production sequencing on the HiSeq X is performed. Clustering and sequencing are performed as per standard Illumina protocols for HiSeq X sequencing. Each UDN sample is sequenced by itself in one lane, plus also sequenced in a pool of 3 samples across a final lane. One lane of sequencing on the HiSeq X instruments generates approximately 30X coverage of the human genome when duplicates removed from consideration. The UDN decided on a minimum of 40X coverage. Therefore, for a given trio of samples, a pool of the 3 samples is run across a fourth lane to supply an additional ~10X coverage per sample, yielding 40X total. Approximately 360 million paired-end reads, each 150 bp in length, are generated for each sample, with typical flow-cell runs lasting ~3 days each. Over 105 Gb of sequence per sample is generated per lane and a 40X UDN WGS sample receives a minimum of 150 Gb of data.

D. Analysis

Overview

The steps in the analysis of WES or WGS data can generally be divided into 4 phases: primary analysis, secondary analysis, tertiary analysis, and interpretation. Secondary analysis can be further subdivided into read mapping and variant calling phases. Best-of-breed standards in analysis of WES or WGS sequence data are followed (as defined in this manual and agreed to across all SCs, CSs, and the CC).

Annual review of current methodologies is performed by the SCs with an aim of identifying and potentially incorporating advances of note in analytical approaches supporting interpretation of sequencing data. Any alterations that are considered for inclusion are shared with the Sequencing Working Group. Those prioritized are implemented leading to revision of the analysis steps outlined in this section. Each CS may conduct analyses on UDN cases sequence data as they see fit, but the SCs undertake primary, secondary, and tertiary analysis of the sequencing data. The purpose of this is to provide consistency in the format and quality of the data provided and to create maximal utility for the widest range of consumers of these data. For example, this method ensures that sites without existing clinically certified variant annotation and prioritization pipelines have access to richly annotated data.

Sequencing files in the Gateway

Output files from various stages in the analysis process derived from various applications are uploaded to the Gateway by the SCs. Files may also be stored locally according to clinical data retention policies in place at the SCs. As currently defined these are:

1. Standard compliant format FASTQ files
2. Standard compliant BAM files
3. Standard compliant VCF files
4. Annotated variant files
    a. A tab delimited text file format
    b. SCs work to ensure that the format and data encapsulated in this file is equivalent at both sites
5. A spec sheet listing software versions and patches, analysis tools, and annotation repositories is provided, along with exact parameters used in the analysis.

    a. This allows sites to unambiguously determine the exact steps for reproduction of analysis and, perhaps allow for case based additional optimization of analysis parameters at capable CSs. Format of this file TBD; one suggestion is an XML spec sheet for unambiguous representation and downstream automation.
    b. In addition, where applicable (for example as the header of the generated VCF file), a human verifiable description of the applications, version, and reference datasets used is encapsulated in the output files themselves.
6. An interpreted clinical report is also be provided by the SCs.
    a. The format of this report follows the existing industry standards for clinical sequencing reports. Clinical reports include the following report sections:
        i. Lab contact information and general test information
        ii. Participant name and date of birth
        iii. Indication for testing
        iv. Primary findings (pathogenic, likely pathogenic, and variants of unknown clinical significance) in tabular format
        v. Medically actionable findings in tabular format
        vi. Interpretation of findings – textual discussion of the relevance of the findings given the clinical presentation of the proband
        vii. Specific recommendations
        viii. A description of the methods used
        ix. Limitations for both the sequencing technology and analytical processes
        x. References
7. Each family member receives a report containing one of three possible results:
    a. Medically actionable finding identified
    b. No medically actionable finding identified
    c. Family member opted out of receiving medically actionable findings

Primary analysis – demultiplexing

Primary analysis (demultiplexing) is performed on the HiSeq instrument workstation according to Illumina guidelines. Software used for primary analysis is described in Table 1.

The primary analysis steps at each site are equivalent, although they may have version differences reflecting the software update timetables in place at the SC. The primary analysis software version used is listed in the spec sheet provided by the SCs. Changes and updates are communicated to the CSs.

Secondary analysis – read mapping

It is important to note that the secondary analysis steps performed at the SC are not necessarily be identical across all steps. They are, however, comparable and clinically appropriate as defined by their existing usage in the CLIA and College of American Pathologists (CAP) accredited clinical laboratories at both SCs. The secondary analysis steps at each SC are equivalent although they may have version differences reflecting the software update schedules in place at the SCs. Each SC performs clinically appropriate validation of all datasets and algorithms/software applications in use within its clinically validated analysis pipelines. Significant pipeline component changes undergo re-validation at the discretion of the SCs.

Secondary analysis step 1 (mapping and realignment) is performed at each SC using the methods and tools described Table 2.

GRCh37/hg19 (b37d5) will continue to serve as the alignment template until it is superseded and adopted by leaders in the field. Alignment of reads to GRCh37/hg19 (b37d5) is performed without truncation of the data. More specifically, duplicates are marked but not removed from the dataset.

Secondary analysis Step 2 (variant calling) is performed at each SC using the methods and tools described Table 2.

Tertiary analysis – variant annotation

It is important to note that the variant annotations produced by each SC are not identical but they are equivalent. The reason behind this is that to make them identical would require re-working the clinical pipelines in place at the SCs. This activity was not planned for or funded. The existing clinically validated pipelines at the two sites are used. The SCs liaise closely to ensure equivalency and continue to work towards a unilateral set of variant annotations. The specific variant annotations and the tools used to produce them are outlined in Tables 3 and 4. The SCs aim to ensure that the data is labeled the same in the datasets produced from each CS.

Coverage analysis

In addition to primary, secondary, and tertiary analysis, the SCs also provide a summary of the sequencing coverage for each sample. This report details the coverage from the sample run at the gene model, transcript, and exon levels. This provides the CSs with an indication of regions of likely importance that are not well covered. In cases where a trio or an extended pedigree is submitted for sequencing, this coverage report is provided for both the proband and the other individuals in the pedigree. The CSs and CC is appraised of any updates to the coverage algorithms made by the SCs.

BCM uses ExCID, an in-house developed tool, to assess capture efficiency and coverage at desired cutoffs. Files from this tool can be viewed in any genome browser to see the actual coverage on genes of interest. This also reports a summary of poorly covered regions.

HA assesses coverage with GapMine v 3.0.1, a software package developed the Medical College of Wisconsin for coverage at the gene, transcript, and exon level. Files can be reviewed in any genome browser. The coverage at these levels and for specific gene lists are provided as a report if requested. Coverage gaps at a defined depth threshold are also reported.

Interpretation

The SCs also provide an interpreted clinical report. A systematic process is followed in accordance with published American College of Medical Genetics and Genomics (ACMG) guidelines to determine the clinical significance of each variant considered for reporting.

Analysis output delivery turnaround times

Initial analysis (to end of tertiary analysis phase) is completed within a 2-week turnaround time (TAT) at both SCs.

Preliminary clinical reports are typically available 7-8 weeks after raw data is uploaded to the Gateway. All variants included in clinical reports are confirmed by Sanger sequencing. CSs may

request additional Sanger sequencing of variants identified during their analysis. Sanger sequencing is performed by the SCs. The SCs Sanger confirm up to 8 variants per case. Unused Sanger confirmations may be distributed among each CS's case cohort (e.g., 25 cases = up to 200 Sanger confirmed variants). The turnaround time for final clinical reports depends on the timing of CS data analysis.

Requesting release of sequencing data

Each SC follows its existing institutional policy for fulfilling raw data requests. Raw data is only released after a final report has been returned to the CS and test results have been communicated to the participant.

Baylor College of Medicine can release WES data to physicians and qualified researchers at the participant's request. See https://baylorgenetics.com/consent/ for more information.

HudsonAlpha can release WGS data to participants, physicians, and qualified researchers. Interested participants should contact the Clinical Services Laboratory directly at clinical@hudsonalpha.org or (256) 327-9413.

Table 1. Primary analysis tools used by each SC.

Phase	Step	BCM	HA
Demultiplexing	Bcl2Fastq	bcl2fastq-1.8.3	Dragen executable v01.002.091.0 1.03.17.16114

Table 2. Analysis steps and applications/algorithms/platforms used for secondary analysis at each SC.

Phase	Step	BCM	HA
Read mapping	Alignment	bwa v0.6.2	BWA-mem v0.7.12
	Fixmate, Sort & Index	Picard v1.8.4	SAMbamba v0.5.4
	Mark duplicates	Picard v1.8.4	sambamba_v0.5.4 markdup
	Realignment and recalibration	GATK v2.5.2	GATK v3.3
Variant calling	SNV	Atlas2-SNP v1.4.3	GATK v3.3
	INDEL	Atlas2-Indel v1.4.3	GATK v3.3

Table 3. Tertiary analysis tools used by each SC.

Phase	Step	BCM	HA
Variant annotation	Annotation	Cassandra v15.4.29	CarpeNovo v 6.0.1
Variant prioritization	Trio analysis	Trio Afterburner v2	CarpeNovo v 6.0.1

Table 4. A depiction of the set of variant annotations and tools used by each SC.

Annotation	BCM	HA	Notes
Annovar	yes	no	commercial product
Alamut HT	no	no	commercial product
splice sites	yes	yes	
near splice site	yes	yes	BCM: +/- 5bp , HA +/- 6bp donor & 25bp acceptor
protein coding flag	yes	yes	
syn change flag	yes	yes	
non-syn change flag	yes	yes	
AA change	yes	yes	
Sift prediction	yes	yes	
Polyphen2 HVAR prediction	yes	yes	
Polyphen2 HDIV prediction	yes	yes	
Mutation Taster prediction	yes	yes	stand-alone
Condel prediction	no	no	BCM uses dbNSFP (PMID: 25552646)
MutationAssessor prediction	yes	yes	
AlignGVGD	no	no	web service; not in pipeline
stop gain flag	yes	yes	
stop loss flag	yes	yes	
start loss flag	no	yes	
frameshift flag insertion/deletion/indel	yes	yes	
non-frameshift flag insertion/deletion/indel	yes	yes	
location: intron/exon	yes	yes	
location: 5'UTR/3'UTR/Intergenic/Promotor	yes	yes	
HGNC appropriate Gene Symbol	yes	yes	RefSeq
Transcript ID	yes	yes	
COSMIC	yes	no	BCM: additional HA: scheduled for 2016 inclusion
HGMD ID	yes	yes	
HGMD variant level association	yes	yes	
HGMD gene level association	yes	yes	
OMIM ID	yes	yes	
OMIM variant level association	yes	yes	
OMIM gene level association	yes	yes	
ClinVar ID	yes	yes	

ClinVar metadata (various; to be clarified)	yes	yes	
dbSNP ID	yes	yes	
dbSNP AF	yes	yes	
1000 Genomes AF	yes	yes	
ESP EVS AF	yes	yes	
Mappability score	yes	yes	

XII. Biospecimens

A. Recommendations for Research Specimen Collection and Processing

Overview

The following specimens should be collected on all probands and affected family members unless they are refused by the participant or their collection would compromise participant safety. Clinical sample collection is understood to be the priority and blood draw volume limits are guided by institutional policies. Blood volume limits in pediatric probands are addressed below. Specimens may also be collected from unaffected family members at the discretion of the CS. When blood is drawn for research purposes, DNA collection and plasma for metabolomics should be given the highest priority. It is then up to the individual CS to prioritize the remaining available blood for serum, plasma, and peripheral blood mononuclear cells (PBMCs) based on planned studies.

Serum: 3 ml serum in 0.5 ml aliquots

Plasma: 3 ml plasma in 6 x 0.5 ml aliquots (for general use) and 3 x 0.05 ml (for metabolomics and lipidomics analyses)

DNA: at least 20 micrograms (with goal of 50 micrograms) at a target concentration of 100-200 ng/ul (with a minimum of 50 ng/ul)

PBMCs: stored in 4 aliquots of $2x10^6$ cells followed by multiple aliquots of $5x10^6$ cells each

Urine: 10 ml urine in 10 x 1.0 ml aliquots (for general use) and 3 x 0.1 ml (for metabolomics analyses)

In the instructions below, aliquots for the Metabolomics Core must be placed in 0.5ml Sarstedt Biosphere® SC Micro Tubes (Sarstedt Catalog # 72.730.217), snap frozen, and shipped to the lab within 3-6 months for analysis. Collection of samples for metabolomics analysis is at the discretion of the CS.

General sample collection guidelines

Blood samples for serum and plasma should be obtained in the fasting state, defined as an overnight fast for adults and at least 3 hours of fasting for children. If a subject is unable to fast, samples should still be obtained. The CS should record whether blood samples were collected as fasting (including duration) or non-fasting.

Blood for PBMCs should be collected in CPT Vacutainer® tubes with citrate (one 8 ml tube for pediatric subjects and as total blood draw volume allows, and two 8 ml tubes for adult subjects; if sample volumes are limited by proband body weight use the smaller 4 ml CPT tube). If the CS has expired 8 ml CPT tubes or wants to be cost-efficient and keep only 4 ml CPT tubes in stock, the CS uses as many 4 ml CPT tubes as needed.

Blood for DNA should be collected in one 10 ml purple top EDTA Vacutainer® tube and sent to a local CLIA laboratory for DNA extraction and quantification. If DNA from blood cannot be obtained, an alternative source of DNA such as skin fibroblasts should be considered.

Pediatric probands (and other participants with limited blood collections)

The volume of blood drawn should be consistent with the allowable blood collection based on participant body weight. In cases where the blood volume that can be obtained is the limiting factor, samples should be obtained in the following order:

1. 3 ml blood for DNA (EDTA)
2. 3 ml blood for plasma (EDTA)
3. 3 ml blood for serum (SST)
4. 4 ml blood for PBMCs (Citrate CPT tube)

Sample tracking and storage

Barcode labels of samples to be stored at a local or central biorepository should include the UDN ID, sample type, and date of sample collection.

Samples should be stored as follows:

- Serum: -80°C or colder (e.g., liquid nitrogen)
- Plasma: -80°C or colder (e.g., liquid nitrogen)
- Urine: -80°C or colder (e.g., liquid nitrogen)
- DNA: -20°C or colder
- PBMC: cryopreserved in liquid nitrogen

Note that aliquots for metabolomics analysis should not be subjected to freeze thawing.

Sample collection and processing protocols (I-VII)

I. Blood sample collection and processing

1. Overview
    a. Processing of blood for plasma and serum, as well as urine, should be performed within two hours of sample collection. PBMCs should be processed within 24 hours of sample collection. All steps should be performed as quickly as possible and on wet ice or otherwise at 4°C, as appropriate, to minimize artifacts in metabolomics data.
    b. The standard large serum, plasma, urine, DNA, and PBMC aliquots should be stored in screw-cap cryovials appropriate for ultra-cold storage (e.g., Nalgene NUNC 1.8 ml Cryovials, Fisher Scientific Catalog #: 12-565-170N).
    c. Aliquots for the Metabolomics Core need to be stored in Sarstedt Biosphere® SC Micro 0.5ml Tubes, Sarstedt Catalog # 72.730.217.
2. Serum sample collection and processing
    a. Blood for serum should be collected in one 10 ml red top Vacutainer® Serum Separator Tubes (SST) with clot activator
    a. After obtaining the SST sample, allow sample to clot 30 minutes in a vertical position

b. Centrifuge vacutainer tube at 2500 RPM (1-1.3 x 10^3 g) for 10 minutes either at ambient temperature or at 4°C.
c. Carefully remove rubber stopper from vacutainer tube and remove caps from cryovial tubes
d. Aliquot 0.5 ml of serum into 6 cryovial storage tubes (3 ml total), and store samples in a -80°C or liquid nitrogen freezer with appropriate labels (UDN ID, sample type, and collection date). Pediatric samples generate fewer tubes.

3. Plasma sample collection and processing
    a. Blood for plasma should be collected in one 10 ml purple top EDTA Vacutainer® tube
    b. Centrifuge vacutainer tube at 2500 RPM (1-1.3 x 10^3 g) for 10 minutes at 4°C
    c. Aliquot 0.5 ml of plasma into 6 cryovial storage tubes (3 ml total), and store samples in a - 80°C or liquid nitrogen freezer with appropriate labels (UDN ID, sample type, and collection date). Pediatric samples generate fewer tubes.
    d. Aliquot plasma to storage tubes for metabolomic studies (Sarstedt Biosphere® SC Micro Tube 0.5ml, Sarstedt Catalog # 72.730.217).
        i. For metabolomics and lipidomics analyses: 3 aliquots of plasma (0.05 ml each).
        ii. For optional targeted oxylipid analysis (performed in coordination with the Metabolomics Core): 1 aliquot of plasma (1 ml). Please note that hemolyzed samples CANNOT be used for oxylipid analysis.
    e. Flash freeze all samples in liquid nitrogen or quick freeze in dry ice/ethanol prior to storage in a -80°C freezer, or in liquid nitrogen, with appropriate labels (UDN ID, sample type, and collection date).
4. DNA extraction at CLIA laboratory
    a. Multiple acceptable DNA extraction protocols for the EDTA Vacutainer tube blood samples can be used (Examples of suitable Extraction kits: Qiagen Gentra Puregene Blood Kit Catalog # 158445, or Qiagen DNeasy Blood and Tissue Kit Catalog# 69504). DNA should be stored in TE (Tris-EDTA) buffer).
    b. DNA stored at the CS or UDN Central Biorepository (UDNCB) should be labeled with UDN ID, date of birth, sample type, and collection date.
    c. DNA aliquots sent to the SCs must be labeled with participant name, date of birth, and “UDN” (do NOT include UDN ID).
    d. DNA quantification should be performed with PicoGreen (not NanoDrop), and DNA concentration should be between 100 to 200 ng/ul (with a minimum of 50 ng/ul).

II. PBMC isolation and cryopreservation

1. PBMC isolation
    a. CPT tubes are the recommended cell separating device (refer to specific CPT tube manufacturer instructions for complete details for steps 1-4 and those below are offered as suggestions). Other cell separating devices may be utilized at individual sites for local biorepository.
    b. Centrifuge blood collected in CPT tubes at room temperature at 1500 to 1800 x g in a swing bucket rotor for 20-30 minutes with no brake following specific instructions from the CPT tube documentation. Visually inspect the CPT gel plug in addition to other guidance in the manufacturer’s instructions.
    c. Use an aspirating pipette to remove the PBMC layer located at the gel interface in the CPT tube.
    d. Place the PBMCs in a new 50 ml conical tube.

e. Wash cells by gently resuspending the cell pellet in 10 ml sterile 4°C or ambient temperature PBS (or other physiologic buffer) followed by centrifugation at 250-400 x g. Repeat once.
f. Count the PBMCs on a hemocytometer, cellometer, or other standardized cell counting device.
g. Separate into 4 aliquots of $2x10^6$ cells followed by multiple aliquots of $5x10^6$ cells each in separate polypropylene tubes and centrifuge at 250-400 x g to create the final cell pellet. The maximum number of aliquots should be made to allow for as many separate PBMC samples as possible to be saved from any one donor.

2. Preparation of PBMCs for storage in a cryorepository
    a. To prevent contamination, all processing shall be completed in a sterile biological safety cabinet by wiping all inside surfaces with 70% alcohol and performing UV light treatment for at least 5 minutes.
    b. Remove as much of wash buffer as possible from each aliquoted PBMC pellet. You may gently flick pellet to loosen prior to next step (resuspension).
    c. Resuspend each washed PBMC pellet containing 2 or 5 million cells in 1 ml of cold (2-8°C) CryoStor CS10 Freeze Media (BioLife Solutions).
    d. Mix cells by gently tapping the tube; do not use a pipette.
    e. Incubate resuspended cells at 2-8°C for 10 minutes.
    f. Pipet gently to minimize shear force and transfer into a labeled cryopreservation vial.
    g. For cryopreservation, transfer vials to a Controlled Rate Freezer to decrease the temperature in a controlled fashion (or if that is unavailable into an isopropanol containing cryopreservation system followed by transfer into a -80°C freezer for a minimum of 12 hours).
    h. Once cryopreservation vials have been appropriately cooled and contents frozen, transfer to designated receptacle within the liquid nitrogen storage unit for long-term storage with appropriate labels (UDN ID, sample type, and collection date).

III. Urine sample collection and processing

1. Urine samples should be the first morning void urine collected in a polypropylene container. A 24-hour urine sample is not required, but may be elected by the CS.
2. Centrifuge urine at 1000 x g for 5 minutes at 4°C to remove any cells and particulates.
3. Transfer supernatant to storage tubes
    a. Ten 1.0 ml aliquots in 1.8 ml screw-cap cryovials
    b. Three 0.1 ml aliquots in Sarstedt Biosphere® SC Micro 0.5ml Tubes (for metabolomics analyses)
4. Flash freeze all samples in liquid nitrogen or quick freeze in dry ice/ethanol prior to storage in a -80°C or liquid nitrogen freezer with appropriate labels (UDN ID, sample type, and collection date).

IV. Sample shipment to the Metabolomics Core

1. Prior to shipping samples to the Metabolomics Core, specimen and request details must be entered into the Gateway.
2. A hard copy of the Gateway metabolomics request form should be sent with each sample/family. If metabolomics analyses are performed on more than one family member, all samples should be shipped together.

3. Samples should be shipped to the Metabolomics Core frozen and on dry ice.
    a. Samples should be shipped Monday - Wednesday and not during or adjacent to holidays.
4. If a submitted sample does not pass QC at the Metabolomics Core or is otherwise deemed unacceptable, the Metabolomics Core contacts the CS directly via phone or email to request a replacement.
5. Sample labeling discrepancies are addressed on a case-by-case basis at the discretion of the Metabolomics Core.

B. Recommendations for Optional Research Biospecimen Collection and Processing

I. CSF collection and processing

CSF should be collected from neurological cases for clinical laboratory studies and for research use, including metabolomics, lipidomics, glycomics testing, and luminex inflammome studies. All samples should be collected by lumbar puncture in the L3/L4 or the L4/L5 inter-space. If neurotransmitters are to be analyzed, they should be collected first and will pre-empt the possibility of obtaining an opening pressure. Neurotransmitters are volatile and have a cephalo-caudal concentration.

1. Microtubes
    a. Microtubes 1-5 (3.5 ml total): for shipment to Medical Neurogenetics (Note: Medical Neurogenetics and Baylor both supply 5 pre-marked CSF microtubes (total volume if filled only to the lines is 3-3.5 ml) - the studies offered by each lab varies though the methods are the same.
    b. CSF must be collected from the 1st drop into the designated tubes in the order indicated in the following table. Do not collect the CSF in one large tube and aliquot into the tube set.
    c. Fill each tube to the marked line with the following volumes, indicated in the table below. The total CSF volume required is 3-3.5 ml.

Tube Number	Required Volume
1	0.5 ml
2	0.5 ml
3	1.0 ml
4	1.0 ml
5	0.5 ml

    d. If the samples are not blood contaminated, place the tubes on ice (or dry ice if available), and then transfer to a -80°C freezer. If the samples are blood contaminated, then centrifuge immediately (prior to freezing) and transfer the clear CSF to new similarly labeled tubes then freeze and store at -80°C.
    e. Microtubes are collected in 1 ml aliquots and placed in wet ice until they can be aliquoted: 5 x 1 ml (e.g., for energetics), 3 x 0.2 ml (for metabolomics, glycoproteomics, and lipidomics analyses), and the remainder in a final aliquot (e.g., Sarstedt Polypropylene 2ml tube #72.694.107). The Metabolomics Core

requires that its 3 x 0.2 ml aliquots be in Sarstedt Biosphere® SC Micro Tube 0.5ml, Sarstedt Catalog # 72.730.217.

f. Polystyrene tubes 1-3 (0.5 ml each minimum): for measurement of glucose, cell count, protein, sterile fluid culture (Note: a standard lumbar puncture tray contains 4 polystyrene tubes with caps).
g. Polystyrene tube 4: for IgG index, Oligoclonal bands, etc. (Note: use 4th tube from lumbar puncture tray here).

2. CSF handling
    a. CSF is collected by lumbar puncture, place the CSF on wet ice immediately and transport to the laboratory. If dry ice is available and the CSF is not bloody the microtubules for neurotransmitters can be placed in the dry ice immediately.
    b. If CSF is bloody, excessive blood may interfere with neurotransmitter and metabolomics testing
    c. Flash freeze all samples in liquid nitrogen or quick freeze in dry ice/ethanol prior to storage in the -80oC freezer with appropriate labels (UDN ID, sample type, and collection date)
    d. Samples should not be subjected to freeze-thawing

II. Optional blood collection and processing

1. General guidelines
    a. As noted above, blood is collected from all probands for DNA, plasma, serum, and PBMCs. Various other optional samples may be considered as well, including: PaxRNA, buffy coat, platelets, oxylipids, and blood spot cards.
    b. Blood for RNA is collected in PAXgene blood RNA Vacutainer® tubes (VWR 77776-026).
    c. Blood for additional PBMC or other buffy coat collection is collected in additional CPT (citrate) tubes or other site-specific collection tubes.
    d. If platelets are to be collected, draw 31.5 ml blood in 7 tubes of light blue top sodium citrate tubes (BLU).
    e. If oxylipids are to be analyzed, then blood should be collected in a chilled EDTA tube and placed on ice.
    f. Blood spot cards may also be obtained and stored at room temperature.
2. RNA processing
    a. Use PAXgene Blood RNA Kit (Qiagen 762164)
    b. Aliquot in 80ul aliquots and store in 500ul sterile, RNase- DNase-free tubes at -20^{o}C until needed
3. RBC isolation
    a. After removal of plasma (see procedures above), discard the remaining supernatant above the porous barrier using a plastic Pasteur pipette (wide orifice)
    b. Using a glass Pasteur pipette (narrow orifice), transfer the erythrocyte (RBC) pellet to a 50 ml conical tube
    c. Fill 50 ml conical tube to 40 ml with Phosphate Buffered Saline (PBS), pH 7.4 and invert several times to mix
    d. Centrifuge for 5min at 1811 x g at 4^{0}C
    e. Remove saline layer and discard
    f. Repeat wash with PBS pH 7.4 until PBS is clear (minimum of 3 times)

g. Aliquot 1 ml of the erythrocyte (RBC) pellet to clean cryovials. Store in -80°C freezer with appropriate labels (UDN ID, sample type, and collection date). *Note: this fraction also contains granulocytes*

4. Platelet Isolation
    a. Add 1 volume HEP buffer + PGE1
    b. Mix very gently by inverting the tube slowly
    c. Spin at 100 x *g* for 15-20 min at room temperature (with no brake applied) to pellet contaminating red and white blood cells
    d. Transfer the supernatant into new plastic tube using a transfer pipette (wide orifice)
    e. Pellet platelets by centrifugation at 800 x *g* for 15-20 min at room temperature (with no brake applied). Discard the supernatant.
    f. Rinse the platelet pellet two times with platelet wash buffer by gently adding wash buffer and removing it slowly with a pipette. (DO NOT RESUSPEND! to avoid platelet activation)
    g. Store the dry platelet pellet at -80ºC freezer with appropriate labels (UDN ID, sample type, and collection date). *Note: Freezing the pellet disrupts the platelet granules. This pellet is only to be used for determination of glycomics, lipidomics and proteomics that does not include the platelet granules.*
5. Oxylipid sample processing
    a. Centrifuge chilled blood for 10 min at 1200 x g at 4°C
    b. Aliquot 1 mL plasma into a single storage vial
    c. Keep on ice and flash freeze in liquid nitrogen or quick freeze in dry ice/ethanol the samples prior to storage at -80 ºC within 4 hours (optimal is to freeze immediately and store at -80ºC) and do not subject to freeze thaw cycles. If that is not possible then the time should be adjusted to be the same for all samples. Do not store at -20 ºC. Redraw the sample if it is extensively hemolyzed.

III. Skin biopsy collection and processing

Skin biopsies are used for culturing fibroblasts. These cell lines are then be used for various research purposes as well as glycomics testing. For subjects who are unable to provide PBMCs, skin fibroblasts provide an alternate source of living cells for future research.

1. Reagents
    a. DMEM High Glucose (Invitrogen #11965-118)
    b. Fetal Bovine Serum (FBS), Certified, Heat-Inactivated, US Origin (Invitrogen #10082-147)
    c. 100X Antibiotic-Antimycotic (Invitrogen #15240-062)
    d. 0.25% Trypsin-EDTA (Invitrogen #25200-056)
    e. 1 x PBS pH 7.4 w/o Calcium or Magnesium (Invitrogen #10010-023)
    f. Dimethyl sulfoxide (Sigma #D8418-100ML)

2. Collection protocol
    a. 3-5mm punch full thickness skin biopsy obtained according to standard medical procedure
    b. Place biopsy in sterile tissue culture medium (DMEM, 10% FBS, 1% antimycotic, antibiotic) contained in a 15 ml conical tube
    c. Store and transport the biopsy at ambient temperature
    d. Deliver the biopsy to the laboratory within 24 hours (Up to 96 hours is acceptable if shipped)

3. Initiation of skin fibroblast culture
    a. Spray hood and scalpels with Ethanol and wipe with Kimwipe
    b. Clean the biopsy tube by spraying well with Ethanol before placing in the hood
    c. Label 6-well tissue culture (TC) plate and place in TC hood
    d. Aspirate medium from biopsy sample using a 2 ml aspirating pipette
    e. Remove biopsy sample and place in one well of the 6-well dish
    f. Using the scalpels, cut the biopsy sample into 6 pieces (Try to attach the biopsy to the plate with the scratches made by the scalpel)
    g. Add 1 ml of pre-warmed culture medium (DMEM, 10% FBS, 1% antimycotic, antibiotic) to each well, being careful not to dislodge the biopsy
    h. Gently swirl the 6-well dish to coat the wells with culture medium
    i. Place in the 37°C, 5% CO_2 TC incubator for 4-5 days to allow the biopsy to attach to the well
    j. Gently add 2 ml of fresh, pre-warmed culture medium to each well being careful to not dislodge biopsy sample
    k. Allow the sample to remain in the 37°C incubator until a monolayer of cells is present in the wells, feeding cells with fresh culture medium every 3-4 days
    l. Once adequate cells have grown out of the biopsy fragment, remove the cells from each of the 6-wells by washing with 2 ml PBS, and following removal of the PBS, by adding 800ul trypsin (0.25% Trypsin EDTA, Invitrogen). Incubate at 37°C until cells are released from substrate. Then add 2 ml of tissue culture medium and transfer cells to one T75 culture flask.
    m. After 2 days, aspirate all but 1 ml of medium from flask. Scrape the bottom of the flask using a cell scraper, and wash cells with the 1 ml of medium remaining in the flask. Remove the 1 ml of culture, and place in a 1.5 ml eppendorf tube. Perform mycoplasma testing using ATCC Universal Mycoplasma Testing Kit (see section IX.B.III.7 below) on this 1 ml aliquot and record results.
    n. Add 10 ml of fresh DMEM to flask and return to 37°C tissue culture incubator to allow the remaining cells to proliferate.
    o. Allow samples to reach confluency
    p. Remove medium and wash cells with 10 ml PBS
    q. Detach cells as described below (Detaching and Passaging Cells). Add 8 ml of tissue culture medium to collect cell suspension, use 1 ml of the culture to start another T75 culture flask. As described below (Freezing Cells), count the cells remaining in suspension, centrifuge at 1000 RPM for 10 minutes, add 3 ml of freezing medium, and freeze the remaining culture into 3x 1 ml cryovials at -80°C in a cool cell for 3 days with appropriate labels (UDN ID, sample type, collection date, and passage number). Transfer to -150°C for permanent storage (Passage 1 cells).
    r. Allow the second T75 flask to reach confluency, feeding with fresh DMEM every 2-3 days. Once confluent, trypsinize the flask, add 9 ml of DMEM count the cells, centrifuge at 1000 RPM for 10 minutes, aspirate supernatant, add 3 ml of freezing media, and freeze at -80°C in a cool cell for 3 days with appropriate labels (UDN ID, sample type, collection date, and passage number). Transfer to -150°C for permanent storage (Passage 2 cells).

4. Freezing fibroblast cells
    a. Prepare 3-4 ml of fresh freezing medium (10% DMSO, 90% FBS) per T75 flask and warm to 37°C.
    b. Place the 9 ml fibroblast culture into a sterile 15 ml conical tube.

c. Count cells using the cell counter: Add 10ul of culture to BioRad cell counting slide and insert into the BioRad TC20 Automated Cell Counter. Multiply this number of cells by 9 (volume of total culture) and divide by the number of frozen culture aliquots you are making to determine the amount of cells frozen per tube.
d. Centrifuge culture at 1000 RPM for 5 minutes.
e. Aspirate off supernatant.
f. Add prepared freezing medium and mix by pipetting up and down.
g. Aliquot 1 ml into labeled cryovial.
h. Place cryovial into Cool-Cell that is labeled with your name and date.
i. Place Cool-Cell into -80°C freezer for 3 days (36 hours) with appropriate labels (UDN ID, sample type, collection date, amount of cells, and passage number) and then transfer into -150°C storage.

5. Tissue culture for frozen skin fibroblasts
    a. Once the tissue culture medium has warmed, remove cell vial from -150°C and immediately place at 37°C.
    b. Prepare a T75 tissue culture flask or 10cm petri dish by adding 10 ml of pre-warmed tissue culture medium.
    c. Remove cells from vial using a sterile 1 ml pipette tip or 1 ml pipette.
    d. Add cells to flask and gently mix.
    e. Place inoculated flask in a 37°C, 5% CO_2 tissue culture incubator for 24 hours to allow cells to attach to the dish.
    f. After 24 hours, remove medium using a 2 ml aspirating pipette and replace with 10 ml of fresh tissue culture medium. Place at 37°C, 5% CO_2 in tissue culture incubator.
    g. Feed cells every 2-3 days by aspirating off old tissue culture medium and replacing with fresh, pre-warmed, tissue culture medium.
    h. Once cells have reached confluency, cells must be passaged and split into new T75 culture dishes.

6. Detaching and passaging cells
    a. Warm 0.25% Trypsin EDTA to 37°C.
    b. Warm DMEM, 10% FBS, and 1% Anti-Anti to 37°C.
    c. Once reagents have warmed, aspirate medium from flask containing cells using a 2 ml aspirating pipette.
    d. Rinse cells with 10 ml of 1X PBS pH 7.4 without Calcium or Magnesium.
    e. Aspirate 1X PBS using 2 ml aspirating pipette.
    f. Add 1 ml of pre-warmed 0.25% Trypsin EDTA to TC flask and spread across the attachment area by swirling the flask.
    g. Incubate flask at 37°C for ~ 5 minutes (or until cells are rounding), then gently tap the flask to release the cells.
    h. Add 8 ml of pre-warmed tissue culture medium and wash cells to the bottom of the flask.
    i. Add 3 ml of culture to fresh T75 flask containing 7 ml of tissue culture medium. Place at 37°C, 5% CO2 in tissue culture incubator.

7. Mycoplasma testing
    a. After 2-3 days of cell growth, aspirate all but ~1 ml of medium from flask.
    b. Using a cell scraper, scrape the cells from the bottom of the flask only.
    c. Using a 1 ml serological pipette, wash cells using the 1 ml of media in the flask.
    d. Remove the 1 ml aliquot and place in a sterile 1.5 ml eppendorf tube.

e. Add fresh pre-warmed DMEM to flask and place in 37°C, 5% CO_2 tissue culture incubator for future use.
f. Perform mycoplasma testing using the Universal ATCC Mycoplasma Testing Kit (ATCC #30-1012K) according to the protocol on the 1 ml aliquot.
g. Record mycoplasma results and upload gel image into LIMS. If mycoplasma free, you can continue to passage. If mycoplasma positive, cells must be treated with Plasmocin (InvivoGen #ant-mpt) according to the protocol for 2 weeks and retested.

8. Growth and isolation of fibroblasts for metabolomics analysis
    a. Grow passaged fibroblasts (passage 1 or 2) to a density of ≥ 1 million cells, in replicate (n = 3 – 5)
    b. Remove cell culture media, and place 1 mL of spent media into a storage tube. Also place 1 mL of fresh media into a storage tube (Sarstedt Biosphere® SC Micro Tube 0.5 ml, Sarstedt Catalog # 72.730.217)
    c. Quench cell metabolism using a wash of ice-cold PBS
    d. Scrape cells from culture plate into storage tube (Sarstedt Biosphere® SC Micro Tube 0.5 ml, Sarstedt Catalog # 72.730.217). If needed, then add a small amount of ice-cold PBS to facilitate scraping and collecting of cells, followed by centrifugation to pellet cells (remove PBS and discard).
    e. Flash freeze all samples in liquid nitrogen or quick freeze in dry ice/ethanol prior to storage in a -80°C freezer, or in liquid nitrogen, with appropriate labels (UDN ID, sample type, and collection date)

XIII. Central Biorepository

The UDN Central Biorepository (UDNCB) will accept samples from approximately 4500 participants during the total grant period. Probands and affected family members should be prioritized followed by unaffected family members that underwent sequencing. Samples from family members are not required. The maximum number of samples that may be submitted is as follows:

- Baylor, Duke, Harvard, Stanford, UCLA, Vanderbilt CS: 500 participants each
- NIH UDP CS: 1500 participants

A. Flow of samples to UDNCB

Samples are submitted to the UDNCB through the Gateway. CS are able to submit samples, view samples available, request samples, and view contact information for the UDNCB in the Gateway.

Submitting samples

1. CS follows specimen collection, processing, and storage guidelines described in Biospecimen section.
2. CS selects half of the processed samples for shipping to the UDNCB. The following samples are prioritized for submission to the UDNCB when available: DNA, plasma, serum, urine, and PBMCs. Other samples such as fibroblasts, CSF, and RNA may be submitted at the discretion of the CS. Samples may be shipped in batches (recommended frequency: at least every 2 months).
3. CS enters the sample information in the Gateway by navigating to the Biorepository tab and clicking on “Submit Specimens>Create New Submission.”
4. Once the samples are submitted click the Submission ID link to download a CSV file containing the details of the submission. Keep this for your records.
5. CS prepares the samples for shipment on dry ice, FedEx priority overnight (Mon-Weds only), include printed copy of sample submission form (see step 3). No shipping during a holiday week or two weeks before Christmas through New Year’s due to issues with FedEx deliveries during these times.
6. Triple pack your samples, select an appropriate dry ice safe shipping box and enough dry ice to last 48 hours in case of shipping/delivery delays.
    a. Example: Thermosafe EPS foam box w/corrugated carton #448UPS & 10 lbs dry ice
    b. Triple Packing Components:
        i. Primary container – to contain the sample
        ii. Secondary container – to contain leaks, should include absorbent material
        iii. Outer package – to give form to the package and protect inner contents (e.g., a fiberboard box surrounding the styrofoam box)
        iv. Note: The styrofoam box doesn’t count as one of the 3 containers here (it is to keep samples frozen)
7. Ship samples to: Lynette Rives, Vanderbilt Medical Center, DD-2205 MCN, Nashville TN 37232-2578, Phone: 615-875-7198
8. UDNCB notifies CS when samples have safely arrived.
9. Questions about preparing samples to ship, sample submission form, etc.: email UDNCB@vanderbilt.edu or phone 615-875-7098.

Viewing Samples Available

Samples available can be viewed two ways in Gateway:

1. Within the individual participant record, navigate to the Samples tab and click “View Inventory” to view all samples available for that participant.
2. From anywhere in the Gateway, navigate to the Biorepository tab and click “Specimen Inventory” to view all samples currently available from the UDNCB.

Requesting Samples

Samples stored by the UDNCB are available to UDN investigators and their collaborators. An investigator wishing to obtain samples from the UDNCB should check the availability of the samples of interest through the Gateway. If the investigator wishes to obtain samples from the UDNCB that were deposited by their own site, no approval by the Publications and Research Committee (PRC) is required.

Since quantities are limited, sample requests for samples from a site other than the investigator’s require prior approval from the UDN. If the investigator wishes to obtain samples from the UDNCB that were deposited by another site, they must first obtain approval from the UDN Publications and Research Committee (PRC). After confirming the availability of samples, the investigator must submit a research concept sheet through Qualtrics (https://hms.az1.qualtrics.com/jfe/form/SV_3CyZOKOuiEvnCx7). The concept sheet should describe the proposed studies and the samples needed from the UDNCB. The UDN Publications and Research Committee (PRC) reviews the application and the Director of the UDNCB, who is part of the PRC, reconfirms sample availability. If the request is approved, a UDNCB approval number is sent to the investigator.

The investigator then completes the Sample Submission Form in the Biorepository section of the Gateway. Since outside investigators do not have Gateway access, a member of the UDN with Gateway access may submit the request on the investigator’s behalf.

The UDNCB contacts the investigator to arrange a time to ship samples. The investigator is asked to notify the UDNCB when samples arrive. The UDNCB notifies the relevant CS whenever samples are released to a UDN investigator or third party.

B. UDNCB Procedures

Storing and archiving of biological specimens

Mailed samples are opened in a clean “no amplified DNA” laboratory. The frozen samples, pre-aliquoted in screw cap cryotubes and labeled by the sender, are placed on dry ice while the labels are checked against the Sample Submission Form accompanying the package for confirmation. This form is downloaded from the Gateway and must be included with all samples shipped to the biorepository. The Sample Submission Form contains the participant UDN ID (also printed on the sample tubes). All tubes have participant UDN ID, sample type, and date of collection typed on the label. Receipt of the samples is recorded in the laboratory sample intake book and their condition noted (if dry ice is gone, samples are partially thawed, any tubes are cracked, etc.). Samples are not required to have barcodes, but the UDNCB has the ability to

read 1D and 2D barcodes. If barcodes are included on the labels by the CS, the information outlined in the Biospecimen section must still be typed legibly on the sample labels.

Storage of samples: The biological samples are placed into liquid nitrogen cryotanks (PBMCs and possibly fibroblasts) and -80°C freezers (DNA, serum, plasma, urine) for long-term storage in the locations assigned by the Progeny Laboratory Information Management System (LIMS) database. Samples with multiple tubes are divided into 2 separate freezers/cryotanks. All UDNCB equipment is on the Vanderbilt Delta alarm system with temperature and nitrogen fluctuation notification automatically going to the Director's and the senior Research Assistant's cell phone/pager.

Documentation: The sender is notified by email of shipment arrival and any problems that may have occurred with the shipment (late arrival, partially thawed tubes, broken tubes, etc.). Any problems with the shipment are also recorded into the Progeny LIMS database. Documentation into the Gateway Sample Submission Form is completed by the CS submitting the samples, and a printed copy of the CSV file containing the details of the submission is included with the samples when they are shipped. No participant names or identifying information is recorded in the Progeny LIMS, the sample intake notebook, or requested on the Sample Submission Forms.

Retrieval and shipping of biological specimens

The UDNCB retrieves biological samples from liquid nitrogen and/or -80°C freezers, packages samples in dry ice, and ships to UDN investigators and collaborators. Sample information in Progeny LIMS is used to track quantities and distribution of biological samples.

Locating biological specimens in storage: The Progeny LIMS database contains participant UDN ID, date of birth, sample type, date of collection, and sample location. In addition to sample and location information, the Progeny LIMS database keeps track of original and current quantities of the biological samples and record the distribution of samples to investigators.

Sample retrieval and transfer: The Biorepository Working Group and Steering Committee must approve sample requests by reviewing the Sample Submission Form prior to electronic application to the UDNCB. After UDN approval by the applicant's receipt of a UDN approval number and completion of a Sample Submission Form through the Gateway, the UDNCB contacts the CS or CS-designated investigator requesting the samples by email to pre-arrange a date for shipment. The samples are then located using the Progeny LIMS database, retrieved, and placed on dry ice to prepare for transfer to the investigator requesting the sample.

Documentation of retrieval: The type of sample, amount transferred, date of retrieval, and the CS-designated investigator receiving the sample are recorded in the Progeny LIMS database. The Biorepository sample inventory on the Gateway is updated so that all UDN investigators can see which samples and the amount(s) of each remain in the system.

Packaging and shipping: The biological samples, already labeled and in screw cap tubes, are packaged and shipped per International Air Transportation Association (IATA) requirements that apply to all dangerous goods (such as dry ice) by air. Samples must be triple packed which includes a leak proof bag with absorbent material. UDNCB ships frozen samples in EPS foam containers (1.5 inch minimum thickness) with corrugated cartons, 10 lbs dry ice by FedEx priority overnight (Monday–Wednesday). Average dry ice sublimation in a 1-1/2 inch thick wall EPS container with corrugated container is 5 pounds over 24 hours (< 10 lbs in 48 hours).

Delays can arise with FedEx and the extra dry ice is a safeguard to protect the samples in case of delays in delivery. A list of sample content is included with shipment and the CS and CS-designated investigator(s) are notified by email that the sample has shipped, given the FedEx tracking number, and an electronic copy of the sample sheet. The email requests that the UDNCB be notified upon receipt of the shipment and of any problems with the samples (e.g., tubes thawed or damaged, etc.).

Packaging and shipping budget: Shipping samples to the UDNCB is at the CS expense and can be in batches to reduce costs. Shipping samples out from the UDNCB to investigators is at the expense of the UDNCB. The budget accommodates for a total of 150 shipments over 3 years, as directed by the NIH. If sample approvals by the UDN exceed the number of shipments allotted per year, then additional funds will need to be made available for the UDNCB or, alternatively, the CS-designated investigator(s) receiving samples will be required to pay the shipping costs (estimated to be $175/request).

Quality control: The UDNCB keeps records on the number of samples received, their condition, date shipped/date arrived, etc. The UDNCB also keeps records on the number of samples shipped out, date shipped/date received, condition upon arrival, any problems reported by the recipient, etc. This information is compiled into a quality control report, along with total samples received and shipped, and presented to the Steering Committee 3 times per year.

Updating UDNCB inventory in the Gateway

Samples collected by the CS and entered into the UDNCB Sample Submission Form prior to shipping are automatically entered into the sample inventory and can be viewed in the Gateway. When samples are requested and shipped out through the UDNCB, the UDNCB edits/updates the sample inventory on the Gateway.

XIV. Metabolomics

A. Introduction

The Metabolomics Core (MC) at Pacific Northwest National Laboratory (PNNL) and Oregon Health & Science University (OHSU) provide metabolomics and experimental design consultation, laboratory analysis, and bioinformatics/computational expertise to the UDN. More specifically, the MC supports the objectives of the UDN by:

1. Performing comprehensive untargeted and targeted 'omics measurements to identify quantitative and/or qualitative changes in metabolite and lipid abundances.
2. Performing structural characterization of novel metabolites or lipids identified in participants with rare and undiagnosed diseases.
3. Integrating metabolomics and lipidomics and genomics data with participant clinical phenotypes to provide mechanistic insight and aid diagnoses of rare and undiagnosed diseases.

B. General Workflow

If before, during, or after a UDN evaluation, a CS decides that a proband would benefit from one or more of the 'omics analyses available at the MC, then the CS completes a Metabolomics Request in the Gateway. Requests for metabolomics analysis are reviewed by the MC, with additional consultation with the referring CS as necessary. If no specific hypotheses are generated the standard approach with all samples is to perform untargeted metabolomics and/or lipidomics. A targeted approach to metabolomics and/or lipidomics is used if there are specific hypotheses regarding the participant's diagnosis that the MC wishes to explore. The resultant data is initially analyzed and reviewed by the MC, and if necessary, the data are reviewed by the Metabolomics Case Review Committee and requesting CS. If no diagnosis is identified, further rounds of analysis may be undertaken based on new hypotheses generated via discussions between the Metabolomics Case Review Committee and referring CS. The following figure outlines this workflow:

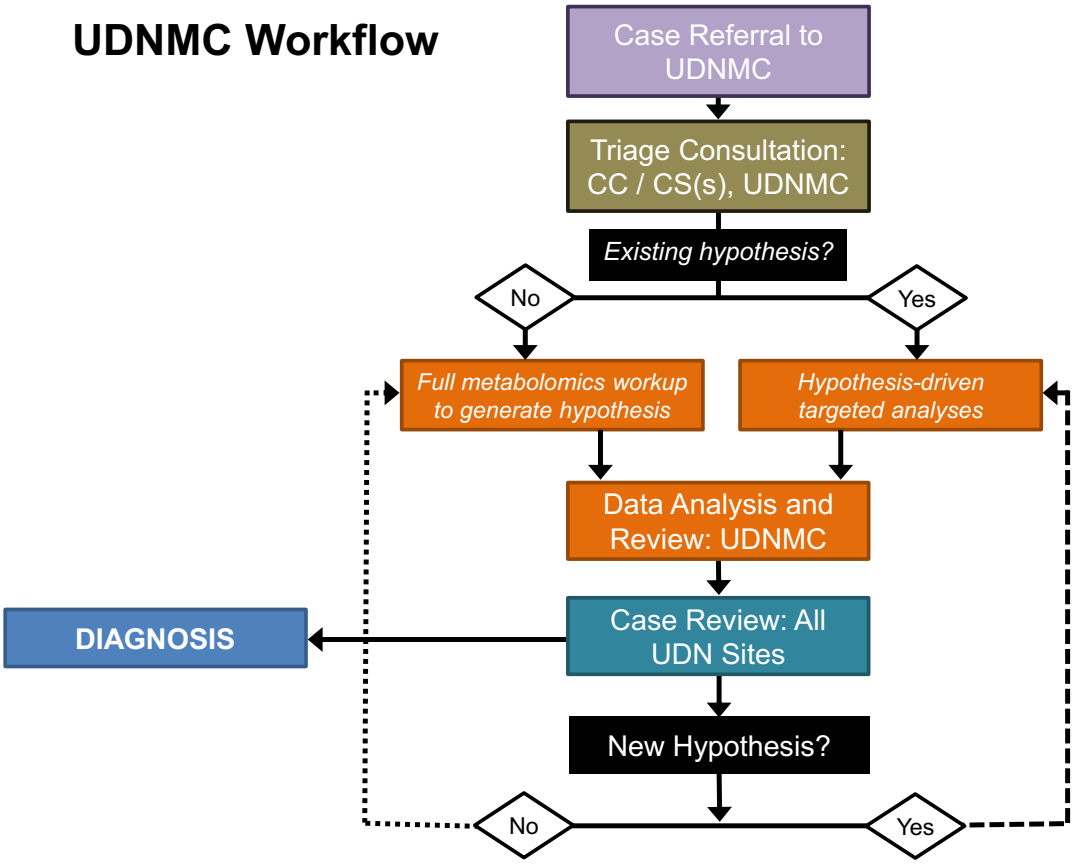


Figure 1: UDNMC workflow diagram

C. Requesting Metabolomics Analyses

To initiate a request for metabolomics analyses, a Metabolomics Request should be submitted in the Gateway. The Metabolomics Request serves two purposes, 1) to alert the MC of the request for services, 2) to provide a high-level summary of the clinical phenotype of the proband and other affected family members.

1. CS completes a request for metabolomics analysis of the proband via the Gateway
2. CS completes a request for metabolomics analysis for each additional family member to be evaluated via the Gateway

Metabolomics case review

1. Requests for metabolomics analysis are reviewed by the MC
2. Consultation with the referring CS is initiated by the MC for any questions or concerns regarding the participant's symptoms and/or analyses requested

D. Flow of Samples to the Metabolomics Core

Recommendations for sample collection, processing, tracking, storage, and shipping

Instructions for sample collection, processing, tracking, storage, and shipping can be found in the Biospecimens section of this manual. As outlined in the Biospecimens section, specimens

for metabolomics analysis should be collected on all probands evaluated in person, unless doing so would compromise participant safety, or they are refused by the proband.

E. Flow of Clinical and Sequencing Information to the Metabolomics Core

Following submission of the Metabolomics Request containing a high-level summary of the clinical phenotype of the proband and other affected family members by the CS, the MC reviews additional clinical and sequencing information available in the Gateway. The MC contacts the CS if more information is required. The CS notifies the MC via email if new information becomes available after the metabolomics request form is submitted.

F. Metabolomics, Lipidomics, and Oxylipid Analyses

Metabolomics and lipidomics – Pacific Northwest National Laboratory

This section covers sample intake, preparation, and analysis at Pacific Northwest National Laboratory (PNNL). This section describes the sample flow from plasma, urine, or CSF sample receipt to production of metabolomics and lipidomics, including appropriate quality control and assurance procedures.

Sample intake

Plasma, urine, or CSF samples are received at the Biological Sciences Facility (BSF). A visual inspection of the sample tubes is conducted to ensure that sample tube integrity has not been compromised. Sample tube label is compared with information in the sample submission form to ensure consistency with and completeness of the Gateway data, and proper sample labeling. Samples are then stored at -70°C. Samples are accepted for analysis if no discrepancies are found, sample labels match, and no tube damage is observed. If any of the above criteria is not met, PNNL notifies the referring CS.

Once accepted, samples are accessioned into the PNNL Laboratory Information Management System (LIMS), incorporating participant UDN ID and sample type.

Sample QC

Prior to processing, samples are thawed and inspected for unusual characteristics such as hemolysis, precipitation, and discoloration. The sample tube label is compared with information on the Metabolomics Request. Samples are accepted for analysis if no discrepancies are found, sample labels match, and no tube damage is observed. If any of the above criteria is not met, PNNL staff notify the referring CS. Any unusual findings are also recorded in the LIMS.

Sample processing

Samples are batched and processed according to which analyses are requested. For metabolomics or lipidomics analyses in plasma and CSF, a chloroform/methanol extraction is used to isolate metabolites and lipids. For metabolomics analyses in urine, a methanol extraction is used to isolate metabolites. Pooled QC samples are included in all sample batches. Internal reference standards are added to participant and QC samples. Details of these extraction protocols are given below.

Extraction of metabolites and lipids from plasma and CSF

For plasma, 50 µL is prepared. For CSF, 200 µL is prepared. Samples are transferred to Sorenson low-binding microcentrifuge tubes. To the plasma, 50 µL of GC-MS internal standards (malonic acid-d4, succinic acid-d4, glycine-d5, citric acid-d4, fructose $^{13}C_6$, L-tryptophan-d5, lysine-d4, alanine-d7, stearic acid-d35, benzoic acid-d5, octanoic acid-d15 at a final concentration of 1 µg/µL each) and 10 µL of LC-MS internal standards (SPLASH™ Lipidomix® Mass Spec Standard, Avanti Polar Lipids, Inc.) are added. To the CSF, 50 µL of GC-MS internal standards (fructose $^{13}C_6$, L-tryptophan-d5, lysine-d4, alanine-d7, stearic acid-d35, benzoic acid-d5, octanoic acid-d15 at a final concentration of 1 µg/µL each) and 10 µL of LC-internal standards ((PC(17:0/14:1) at 1 µg/µL, LPC(19:0) at 0.01 µg/µL, and TG(17:0/17:1/17:0)-d5 at 0.01 µg/µL) are added to the CSF. Cold (-20°C) chloroform/methanol (2:1, v/v) is added in a 5-fold excess to the sample volume. Samples are vortexed for 10 seconds to facilitate mixing of samples and solvent, allowed to sit at 4°C for 5 minutes, and then vortexed again for 10 seconds. Then, samples are centrifuged to facilitate separation of a hydrophilic layer containing polar metabolites and a hydrophobic layer containing lipids. The two liquid layers are removed and placed into new microcentrifuge tubes and evaporated to dryness in vacuo. Metabolite extracts are stored dry at -20°C until chemical derivatization (see below). Lipid extracts are stored in chloroform/methanol (2:1, v/v) until instrument analysis.

Extraction of metabolites from urine

For urine, 100 µL is prepared. Samples are transferred to Sorenson low-binding microcentrifuge tubes to which 50 µL of GC-MS internal standards (malonic acid-d4, fructose $^{13}C6$, L-tryptophan-d5, lysine-d4, alanine-d7, stearic acid-d35, benzoic acid-d5, octanoic acid-d15 at a final concentration of 1 µg/µL each) and 100 µL of a 1 mg/mL solution of urease prepared in water is added. The samples are incubated for 30 minutes at 37°C with mild shaking to deplete urea. Metabolites are then extracted with concomitant protein precipitation by addition of 1 mL of cold (-20°C) methanol with vortexing for 30 seconds, and precipitated proteins are isolated by centrifugation. The supernatants are transferred to glass autosampler vials and then dried in vacuo. Metabolite extracts are stored dry at -20°C until chemical derivatization (see below).

Chemical derivatization of metabolites

Polar metabolites are chemically derivatized prior to metabolomics analysis. Two post extraction standards are added (pentadecanoic acid-d3 and 3-hydroxymyristic acid-d5 at 1 µg/µL final concentration) to monitor instrument performance. To protect carbonyl groups and reduce the number of tautomeric isomers, 20 µL of methoxyamine in pyridine (30 mg/mL) is added to each sample, followed by vortexing for 30 seconds and incubation at 37°C with generous shaking for 90 minutes. At this point, the sample vials are inverted one time to capture any condensation of solvent at the cap surface, followed by a brief centrifugation. To derivatize hydroxyl and amine groups to trimethylsilyated (TMS) forms, 80 µL of N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA) with 1% trimethylchlorosilane (TMCS) is added to each vial, followed by vortexing for 10 seconds and incubation at 37°C with shaking for 30 minutes. Again, the sample vials are inverted one time, followed by centrifugation. The samples are allowed to cool to room temperature and are analyzed that day.

Sample analysis

Samples are analyzed using gas chromatography-mass spectrometry (GC-MS) for metabolomics and liquid chromatography-mass spectrometry (LC-MS) for lipidomics.

GC-MS-based metabolomics analyses

An Agilent GC 7890A coupled with a single quadrupole MSD 5975C is used to analyze chemically derivatized metabolites. The samples are batched and analyzed in random order. A HP-5MS column (30 m × 0.25 mm × 0.25 µm; Agilent Technologies, Inc) is used for untargeted metabolomics analyses. The sample injection mode is splitless and 1 µL of each sample is injected. The injection port temperature is held at 250°C throughout the analysis. The GC oven is held at 60°C for 1 minute after injection and the temperature is increased to 325°C by 10°C/min, followed by a 5-minute hold at 325°C. The helium gas flow rates are determined by the Agilent Retention Time Locking function based on analysis of deuterated myristic acid. Data are collected over the mass range 50 - 550 *m/z*. A mixture of fatty acid methyl esters (FAMES; C8-C28) are analyzed once per day together with the samples for retention index alignment purposes during subsequent data analysis.

LC-MS-based lipidomics analyses

Stored plasma lipid extracts are dried in vacuo and reconstituted in 200 µL of methanol containing post-extraction internal standards (PE(17:0/14:1) and PI(17:0/14:1) at a final amount of 0.05 µg and 0.01 µg, respectively. Stored CSF lipid extracts are dried in vacuo and reconstituted in 50 µL of methanol containing the same post-extraction internal standards at a final amount of 0.02 µg each. A Waters Acquity UPLC H class system interfaced with a Velos-ETD Orbitrap mass spectrometer is used for LC-ESI-MS/MS analyses. A Waters CSH column (3.0 mm x 150 mm x 1.7 µm particle size) is used to separate lipid molecular species over a 34-minute gradient (mobile phase A: ACN/H_2O (40:60) containing 10 mM ammonium acetate; mobile phase B: ACN/IPA (10:90) containing 10 mM ammonium acetate) at a flow rate of 250 µl/minute. Eluting lipids are introduced to the MS via electrospray ionization in both positive and negative modes, and lipids are fragmented using HCD (higher-energy collision dissociation) and CID (collision-induced dissociation) to obtain high coverage of the lipidome.

Data processing

Metabolomics and lipidomics raw data are processed using open-source, in-house developed, and commercial software tools to align data and identify detected molecules.

GC-MS-based metabolomics data processing

GC-MS raw data files are processed using the Metabolite Detector software, version 2.0.6 beta. Briefly, Agilent .D files are converted to netCDF format using Agilent Chemstation, followed by conversion to binary files using Metabolite Detector. Retention indices (RI) of detected metabolites are calculated based on the analysis of the FAMEs mixture, followed by their chromatographic alignment across all analyses after deconvolution. Metabolites are identified by matching experimental spectra to an augmented version of the Agilent Fiehn Metabolomics Retention Time Locked (RTL) Library, containing spectra and validated retention indices for over 700 metabolites, using a Metabolite Detector match probability threshold of 0.6 (combined retention index and spectral probability). All metabolite identifications are manually validated to reduce deconvolution errors during automated data-processing and to eliminate false identifications. The NIST 08 GC-MS library is also used to cross validate the spectral matching scores obtained using the Agilent library and to provide identifications of unmatched metabolites. The three most abundant fragment ions in the spectra of each identified metabolite are automatically determined by Metabolite Detector, and their summed abundances are integrated across the GC elution profile; fragment ions due to trimethylsilylation (i.e., *m/z* 73 and 147) are excluded from the determination of metabolite abundance. A matrix of identified metabolites, unidentified metabolite features (characterized by mass spectra and retention indices and assigned as 'unknown'), and their abundances are created for statistical analysis. Features resulting from GC column bleeding are removed from the data matrices.

LC-MS-based lipidomics data processing

LC-MS/MS lipidomics data are analyzed using the in-house developed software LIQUID (Lipid Informed Quantitation and Identification) (Kyle et al., *Bioinformatics*, 2017; PMID: 28158427). For global data analysis, the software iterates over each MS/MS spectrum to find the best lipid match and presents a graphical display of all observed and theoretical lines of evidence used for identification. The primary evidence shown is a stem plot of the MS/MS spectra including colors and labels for peaks that match to fragments of the identified lipid. A stem plot of the isotopic profile and a line plot of the extracted ion chromatogram are also provided to show the MS-level evidence of the identified lipid. Intensity, mass measurement error, and elution time are also provided. To align and gap-fill the mass spectrometry data, the identified lipid name, observed *m/z*, and the retention time from each analysis is used as a target database for feature identification across all LC-MS/MS runs. This is achieved by aligning all datasets (grouped by ionization type) and matching unidentified features to their identified counterparts using the open-source tool MZmine 2. Aligned features are manually verified and peak apex intensity values are exported for statistical analysis. A matrix of identified lipids, unidentified lipid features (characterized by m/z and retention times and assigned as 'Unidentified'), and their abundances are created for statistical analysis.

Oxylipid analysis – Lipidomics Core Facility (LCF), Wayne State University

This section covers sample intake, preparation, and analysis at the LCF at Wayne State University.

Sample intake

PNNL ships samples to the LCF for oxylipid analysis. Plasma samples shipped on dry ice are received by the LCF staff. Vials are removed and visually inspected to ensure that sample tube integrity has not been compromised. The sample tube label is compared with information on the Metabolomics Request and assigned an LCF sample number for the analysis. Samples are then stored at -70°C. Samples are accepted for analysis if no discrepancies are found, sample labels match, and no tube damage is observed. If any of the above criteria is not met, the LCF staff notify the referring CS.

Sample QC

Prior to processing, samples are thawed and inspected for unusual characteristics, such as hemolysis, precipitation, discoloration and appropriate volume. All findings are documented for each sample and entered into the master sample excel sheet. If there is extensive hemolysis the CS is asked to confirm they want to proceed with analysis.

Sample processing and data analysis

Due to analyte instability, samples need to be processed within 30 days of the initial sample collection. Where possible, samples are batched into a single processing run. If that is not possible, they are prepared as single samples.

After thawing on ice, samples are centrifuged to clarify. The samples are then spiked with an internal standard mix that contains 5 ng each of PGE1-d4, RvD2-d5, LTB4-d4, 15-HETE-d8, and 14(15)-EpETrE-d11. The samples are then extracted for PUFA metabolites and analyzed by HPLC-MS following SOP for fatty acyl lipidomics.

Sample analysis

Scope and applicability

This method is used for the extraction of fatty acyl lipids such as prostaglandins, thromboxanes, leukotrienes, epoxy, hydroxy, hydroperoxy, or keto fatty acids, resolvins, protectins, etc. that are relatively hydrophobic and bind to C18 modified silica. The procedure is applicable to homogeneous liquid samples without particulates. Typically, these samples are tissue culture media, plasma, serum, urine, synovial fluids, lavage, and tissue homogenates free of cell debris.

Sample volume

Typical volume for extraction is 1 mL. If the volume of the sample is less, make up the volume to 0.5 - 1 mL with normal saline. If more volume of the sample has to be extracted due to low concentration of the analytes, the 1 mL cartridge (60 mg of adsorbant) can be used for up to 2 mL of the sample. Use a 3 mL cartridge (100 mg adsorbant) for samples above 2 mL but less than 4 mL. For samples above 4 mL, use the 6 mL cartridge (500 mg adsorbant).

Materials

Strata-X cartridges (Phenomenex, Strata-X, 1 mL, catalog# 8B-S100-TAK) – one per sample, 1.8 mL autosampler vials with Teflon-lined caps (Microsolv, vial: catalog# 9512S-0CV,

cap: catalog# 9502S-10C-B) – one per sample, LC-MS grade solvents (formic acid, methanol and hexane), and ultrapure water (e.g., MilliQ, nanoPure, etc.). Prepare 15% methanol in water (approximately 4 mL per sample) and 0.1% formic acid in methanol (about 1 mL per sample).

Procedure

1. Precondition the cartridges by washing first with 1 mL methanol followed by 1 mL of 15% methanol in water. Do not let the cartridge dry. Keep adding 15% methanol, if necessary.
1. Add methanol to sample to a final concentration of 15%, mix thoroughly, and let it stand at room temperature for 30-60 minutes. Centrifuge, if necessary, to remove any particulates or precipitated protein. Apply the clear sample to the preconditioned cartridge and pass the sample through. Rinse the sample tube twice with 1 mL of 15% methanol and add to the cartridge (i.e., two individual rinses with 1 mL each, centrifuge each time, if necessary, to sediment the particulates).
2. Wash the cartridge with 2 mL 15% methanol and dry the cartridge briefly using vacuum.
3. Wash the cartridge further with 2 mL hexane and dry again. Apply vacuum to dry the cartridge.
4. Remove any water droplets sticking to the cartridge by shaking them off or by blotting with a paper towel, DO NOT wipe or scrub.
5. Elute the cartridge with 0.5 mL of methanol containing 0.1% formic acid directly into a 1.8 mL autosampler vial. (Tip: Add the elution solvent along the sides of the cartridge at the top. This ensures recovery of any compound hydrophobically adsorbed to the plastic barrel.) The sample can be stored at -80°C in a tightly capped vial at this stage.
6. Evaporate the eluate under nitrogen to complete dryness. Using minimal volume of LC-MS grade methanol, wash the sides of the vial such that all residue is collected at the bottom of the vial. Evaporate the methanol again making sure that the extracted residue stays at the bottom of the vial. Immediately add LC-MS grade methanol, close with Teflon-lined cap, and store at 4°C or below. DO NOT leave the samples dry. (Volume of methanol to be added depends on the expected concentration of the metabolites. Typical volumes are 25 – 50 µL).
7. If the sample can be analyzed immediately, proceed to the next step. Otherwise, store the methanolic solution below -30°C (-80 °C for long term storage) in a frost-free freezer.
8. If methanol reconstituted sample was stored below 4°C, remove it from the freezer and let it warm to room temperature. Add an equal volume (equal to the volume of methanol used) of 25 mM ammonium acetate freshly prepared in MilliQ water. Observe the sample for any turbidity and/or precipitate. If the sample is clear, transfer the capped vial to the LC-MS autosampler and analyze. If the sample precipitated (even partially) or appears turbid, proceed to the next step.
9. Turbid or precipitated samples clog the LC-MS system. NEVER inject a sample that is not clear. If possible, dilute the sample with the same composition of the final solvent mixture until clear. A general practice is to add measured volumes of the organic solvent (e.g., methanol) first until the solution becomes clear and then add an equal volume (equal to the amount of additional organic solvent added) of 25 mM aqueous ammonium acetate. Avoid excessive dilution since the measurable analyte concentration per unit volume might become too low to be detectable. If necessary, filter the methanol:water reconstituted sample through a 0.45 µ syringe filter directly into a new LC-MS autosampler vial, as a last resort.
10. If the reconstituted samples are not analyzed immediately, DO NOT freeze the sample. Store the sample at 4°C if the analysis has to wait for 12 hours or more. Always keep the samples in dark.

LC-MS method:
LC conditions:
Solvent A: Water-Methanol-Acetonitrile-ammonium acetate (85:10:5:0.1%)
Solvent B: Methanol-Water-ammonium acetate (95:5:0.1%)
Equilibration solvent: 1:1 mix of solvents A and B
Column: Luna, C18(2), 3 µ, 2.1x150 mm (from Phenomenex)
Flow rate: 0.2 mL/minutes
Oven temperature: 35°C
Gradient Table:
Time (minutes), %B
0, 50
1, 50
8, 80
17, 95
17.1, 50
For the two column alternating regeneration method, equilibration starts at 13 minutes.

MS conditions:
Curtain gas: 35; CAD gas: Medium; IP: -2500 V; Source temperature: 600°C; GS1: 35; GS2: 35

Collision energies and other mass analyzer parameters are optimized for each fatty acyl lipid mediator MRM transition. At the beginning of each sample batch analysis, a set of standards consisting of RvE1, 6-keto PGF1a, TXB2, PGE2, PGF2a, RvD1, LTB4, 13-HODE, 15-HETE, and 4-HDoHE along with the Internal Standard mixture (PGE1-d4, RvD2-d5, LTB4-d4, 15-HETE-d8, and 14(15)-EpETrE-d11) are injected and verified that the retention times have not changed more than 10% of their historical average. In addition, the set of internal standards used are injected directly to compare recoveries during extraction.

G. Analysis

Overview

The steps in the analysis of metabolomics, lipidomics, and oxylipid data can generally be divided into 4 phases: normalization, statistical analysis, data integration, and identification of disease-related pathways and signaling networks.

Normalization

To facilitate the identification of potentially disease associated changes in the analyte profiles of participants a reference dataset of healthy individuals has been established. Batches of samples from participants are analyzed and compared to this reference population. Quality control (QC) processing of the reference dataset includes log2 transformation and the removal of any identified or unidentified features not present in at least 10% of the samples. Samples with missing or low abundance values and an uncorrelated pattern of expression by Pearson correlation and rMd-PAV (Matzke et al., *Bioinformatics*, 2011; PMID: 21852304) are assessed to determine whether the seemingly poor behavior is most likely due to biological or to technical/sample preparation issues. If biological issues appear to be the cause, the sample is retained in the current batch for further analysis; if technical issues appear to be the cause, the sample is omitted from further analyses. QC processing for the participant samples includes the

same steps as for the reference samples; however, participant samples require stronger evidence before removal than reference samples.

Normalization of the reference data and the participant data are performed in two steps. The first, Quality Control Robust LOESS Signal Correction (QC-RLSC), accounts for batch effects, and is performed on a per-batch basis (Dunn et al., *Nat Protoc*, 2011; PMID: 21720319). This requires identical QC samples to be run with each batch of samples (both the reference samples and the participant samples). QC-RLSC is followed by global median centering of each sample, in order to account for differences in the amount of sample material analyzed by GC-MS or LC-MS.

Statistical analysis

To identify unique features in the analyte profiles of participants, results are compared to those form the reference dataset derived from healthy individuals. First, a univariate approach is applied that compares the feature values of the participant to the mean and standard deviation of the feature values in the reference dataset using z-scores (Miller et al., *J Inherit Metab Dis*, 2015; PMID: 25875217). An absolute value z-score threshold is used to obtain a list of metabolites and/or lipids with potential diagnostic significance. Second, to help account for the complex interdependencies of these features, we repeat the univariate approach on ratios between all possible metabolite pairs (Altmaier, et al., *Endocrinology*, 2008; PMID: 18372322), to identify both individual metabolites and groups of metabolites that distinguish a participant from the reference population.

Data integration

Traditional classification methods work well for individual data streams. However, they do not integrate diverse data because each data stream requires a different classification algorithm for maximum predictive group accuracy. Therefore, statistical classification algorithms coupled with a Bayesian model integration method that identifies the contribution of each data type to overall probabilities of classification accuracy are used for quantitative integration of MS-based 'omics data. The approach for integrating clinical phenotyping data with model organism and protein interaction data for prioritizing variants is complementary to the 'omics data classification. Cross-validation approaches are used to select the features that most strongly contribute to the predicted class (e.g., disease phenotype, genotype) for each data type. A combination of features from several different technologies will likely be more predictive of 'treatment' effects than any single data type alone using this method. The molecular features (e.g., metabolite, lipid, gene) that contribute to the overall model accuracy are rank ordered, regardless of whether a molecular identity is known. This strategy may be modified in the context of undiagnosed diseases due to the very small sample sizes that are encountered as a result of evaluating phenotypes restricted to single families. Cluster-based approaches (e.g., hierarchical, k-means) are used on the molecular data types to identify significant groupings in the data. Permutation-based statistics are used to map these clusters to signatures of phenotypes as defined by the Monarch platform.

Identification of disease-related pathways and signaling networks

Metabolites with molecular identifications are integrated with existing pathway tools, reaction databases, and the integrated corpus of genotype-phenotype data within the Monarch platform for biological interpretation of disease etiology and biomarker signatures. Methods to

reconstruct signaling networks describing metabolic changes in the context of protein-protein interactions and integrated genomics data is also used.

In order to map changes in the metabolome and lipidome to the metabolic pathways from which they are most likely derived, analytes that show significant deviation from the reference dataset (absolute value z-score > 2) are evaluated using bioinformatics tools. An integrated corpus of KEGG, Reactome, and HumanCyc databases with EC numbers or SwissProt IDs are used. Integrated data are made available for computation in the public Monarch neo4J graph database, SciGraph (https://github.com/SciCrunch/SciGraph). SciGraph includes provisions for maximal query utility across the integrated data – such as standard queries for gene, metabolite, ortholog, organisms, and source, but also much more sophisticated queries that leverage subgraph comparisons to identify the most likely biological pathways in which metabolites are involved. Where DNA sequencing data is available to integrate with metabolomics and lipidomics data, a combination of the Monarch Exomiser tool and SciGraph queries described above are used, or the DAVID bioinformatics web portal—both approaches utilize several different annotation databases simultaneously and perform functional annotation clustering to group annotation categories by membership with statistical metrics of enrichment. Data types are integrated to identify sets of biomolecules that correspond to the same networks, presenting a weight of evidence approach that the identified pathways are significantly perturbed by the disease process. Metabolite biomarkers are also prioritized according to their known mechanism of action in disease pathogenesis, taking advantage of the cross-species data stores available within Monarch.

H. Return of Results of Metabolomics Analyses

The metabolomics data is initially processed and reviewed by the MC, and subsequently returned to the referring CS via the Gateway. If necessary, the data are also reviewed by the Metabolomics Case Review Committee in collaboration with the referring CS. If no diagnosis is identified, further rounds of metabolomics analysis, and/or other evaluations are undertaken based on new hypotheses generated via discussion between the Metabolomics Case Review Committee and referring CS.

Format of metabolomics results report

1. Summary of key findings
    a. Metabolomics
    b. Lipidomics
    c. Oxylipids
2. Primary diagnostic considerations
3. Spreadsheets of output data (see Figure 1)

Metabolite	KEGG	CAS	PubChem	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6
1,5-anhydrohexitol*	not found	154-48-5	219984	5098978.4	3317542.6	5143193.1	4500130.2	3757596.4	4936321.3
2,3-dihydroxybutanoic acid*	not found	1518-62-3	192742	1382803.0	583266.8	1137075.3	618951.7	651166.1	797528.7
2-hydroxybutyric acid	C05984	600-15-7	11266	958590.0	644536.0	1062866.6	492711.5	581209.0	604425.9
3,4-dihydoryxbutanoic acid*	C10833	1443-76-1	10742	79536.3	4753.2	60380.4	22144.7	14384.8	12307.8
3-hydroxybutyric acid	not found	300-85-6	441	331813.8	112207.3	324314.1	200807.9	173687.1	213907.9
3-phosphoglyceric acid	C00597	3443-58-1	724	104312.8	17735.4	173978.4	126185.4	71361.3	12416.6
alpha ketoglutaric acid	C00026	17091-15-5	51	342294.2	69001.7	539339.0	324638.5	263233.5	387701.0
alpha-hydroxyisovaleric acid*	not found	17407-55-5	99823	440149.0	208393.4	390610.2	563918.6	310641.8	251832.9
capric acid	C01571	111-27-3	2969	549021.8	59215.9	527871.8	201097.5	242959.0	104409.1
caprylic acid	C06423	124-07-2	379	250430.4	256446.9	344759.5	164024.9	275588.6	219791.7
carbonate ion	C01353	144-55-8	22639876	19134541.9	9467804.8	19589625.8	21001927.3	10678799.5	24385863.3
citric acid	C00158	12262-73-6	19782904	1496799.3	1080370.2	2506567.6	1864002.6	887105.6	2155784.2
creatinine	C00791	60-27-5	588	59681.1	6490.0	26029.7	49059.9	4821.0	43795.8
dehydroalanine (possibly from cysteine)	C02218	1948-56-7	123991	328106.6	18550.4	46583.0	64479.3	25931.6	44886.8
D-glucose	C00031	2280-44-6	5793	13203712.3	12382028.9	20419232.9	18225525.2	15478458.6	19564933.1
D-malic acid	C00497	6283-27-8	92824	128226.1	28148.7	255858.1	186997.6	55727.9	124479.4
D-maltose	C01971	133-99-3	6255	73831.3	70902.3	65011.7	104188.2	51359.0	165011.1
D-mannose	C00159	147-74-0	18950	252719.5	91348.1	113996.8	156205.9	197023.6	271538.9
D-ribose	C00121	24259-59-4	5779	340007.7	259022.4	72677.1	185242.7	164739.1	254049.5
D-sedoheptulose	not found	3019-74-7	5459879	7777557.4	6133462.2	15663664.8	12222851.7	15268102.5	16826447.5
D-threitol	C16884	2418-52-2	169019	803449.7	148628.9	252568.7	656278.9	219578.9	328717.9
EDTA	C00284	139-33-3	5232305	51492565.8	35978825.2	41533050.9	25846672.8	17125642.3	84139692.7
formylglycine*	not found	2491-15-8	75606	660903.1	595403.7	910077.0	415284.0	547109.0	385483.0
fructose	C00095	57-48-7	439163	3112797.1	603491.1	3088310.0	4908239.5	525770.9	924281.4
fumaric acid	C00122	110-16-7	21883788	79726.6	95374.2	160754.0	91451.6	88712.4	126839.1
glyceric acid	C00258	473-81-4	439194	365950.4	139873.2	320302.2	219867.0	158642.4	278861.1
glycerol	C00116	25618-55-7	753	16716104.2	12681486.6	15303876.8	16388251.7	16154121.4	21761724.3
glycine	C00037	112898-03-0	5257127	6115626.4	3457832.2	5812929.5	5603209.5	4470356.5	7600203.2
glycolic acid	C00160	1033720-48-7	3698251	52245.2	49312.0	77245.4	38885.7	46207.1	55305.9
hypoxanthine	C00262	146469-94-5	790	25870.0	29055.1	0.0	53058.2	132208.7	0.0
iminodiacetic acid	C19911	142-73-4	8897	1468343.2	1349956.3	1284695.2	1208729.9	1318160.9	728160.6
L-(+) lactic acid	C00186	1715-99-7	107689	30132825.3	26607224.2	40486911.0	43725453.7	45578741.1	42218218.6
L-alanine	C00041	115967-49-2	5950	8016365.7	3876145.8	9932099.8	5823638.1	5007088.7	9924360.9
L-asparagine	C00152	3130-87-8	6267	6160.5	4321.6	6890.6	27653.4	25767.0	40087.9
L-cysteine	C00097	4371-52-2	5862	164875.6	88251.4	106928.4	149788.6	55780.0	84133.2
L-cystine	C00491	24645-67-8	67678	412915.6	283493.5	398942.5	416830.4	138284.5	383301.4
L-glutamic acid	C00025	56-86-0	33032	1036824.0	445926.9	883096.9	1430834.5	681282.1	1036117.1
L-glutamine	C00064	32640-56-5	5961	3145983.9	1820278.9	3983462.2	3273861.1	2285053.3	4753457.4
L-isoleucine	C00407	24629-25-2	6306	1814139.7	509866.5	2502803.8	1488382.5	1454731.3	2585024.7

Figure 1: Example matrix of processed metabolomics data showing metabolite common name, database annotations, and measured abundances.

XV. Model Organisms Screening

A. Goals

The goal of the Model Organisms Screening Center (MOSC) is to provide compelling data based on studies in flies (*Drosophila melanogaster*) or zebrafish (*Danio rerio*) that refute or support the disease causality of specific genetic variants identified in participants by the CSs and the SCs after genome or exome sequencing.

B. General Considerations

The MOSC began analyzing variants in early 2016. The MOSC had numerous discussions with the CSs about how to optimize the process of best diagnosing disease-associated variants in genes submitted by each site. MOSC investigators also had numerous interactions with human geneticists that helped shape approaches and logistics. The overall philosophy of the MOSC is that of a collaborative center for clinicians, human geneticists, and model organism researchers who work together to solve cases. This chapter of the Manual of Operations represents a summary of the general operations of the MOSC.

C. Drosophila Core Operations

The Baylor College of Medicine (BCM) *Drosophila* Core performs bioinformatic analysis of human and *Drosophila* databases. The team closely communicates with the Zebrafish Core to come to a unanimous decision on whether the MOSC should pursue the variant experimentally. Upon further discussion with the Clinical Site that submitted the case, the core directors and scientists of the fly core design experimental strategies to pursue in *Drosophila melanogaster*.

The *Drosophila* Core tailors experiments to each gene and variant using the optimal genetic strategy for each individual case. One strategy depends on the versatile engineered transposable element MiMIC (*Minos*-mediated integration cassette) (Venken *et al.*, 2011), which allows “humanization” of *Drosophila* genes (Bellen and Yamamoto, 2015). This pipeline allows for: (1) generation of strong loss of function (LOF) alleles in the fly gene of interest; (2) expression of the yeast GAL4 transcription factor driven by endogenous enhancers of the gene of interest; (3) functional replacement of the fly gene with the reference (wild-type) human gene to test for rescue of the fly mutant phenotype with the orthologous reference human cDNA expressed under control of GAL4 expressed in the proper spatial and temporal pattern; and (4) variant functional analysis through expression of the human disease allele variant in the same fly line. (1) and (2) are accomplished simultaneously by integrating a T2A-GAL4 cassette into a coding exon of a gene of interest (Chao *et al.*, 2017; Wangler *et al.*, 2017; Yoon *et al.*, 2017) via recombinase-mediated cassette exchange of MiMIC, or via CRISPR-mediated homology directed repair (CRIMIC). Other strategies can also be used depending on the availability of pre-existing reagents (*e.g.*, well characterized LOF alleles, experimentally verified RNAi lines), or when rescue with human cDNA has not been successful (Harel *et al.*, 2016). For example, in cases in which the variant affects an amino acid conserved between the fly and human orthologs, a variant can be engineered in the context of a fly genomic construct to perform rescue experiments using previously characterized alleles with scorable phenotypes (Luo *et al.*, 2017). Regardless of the exact methodology and phenotypes to be scored, the overarching goal of the *Drosophila* Core is to demonstrate whether the variant identified in the participant is

functionally important, knowledge that provides a molecular handle to further explore the molecular mechanism of disease.

D. Zebrafish Core Operations

The goal of the Zebrafish Core at the University of Oregon (UO) is to mimic in zebrafish the lesion found for the participant's candidate gene and to provide a preliminary characterization of the phenotype of the mutation in zebrafish. The Zebrafish Core initially performs extensive bioinformatic analyses that identify the zebrafish ortholog (or in a minority of cases, co-orthologs) of the human candidate gene. Next, the Core checks whether existing mutations or knockdowns have been characterized for the candidate gene in zebrafish and/or mouse, and if so, whether the described phenotypes support or fail to support the conclusion that the participant's clinical manifestation arises from a defect in the candidate gene. If no prior models exist, or if prior models have not been adequately investigated, the Zebrafish Core develops knockout mutations for the candidate gene using CRISPR/Cas9 technology. The goal is to make at least two alleles for each gene tested, ideally one has an early stop codon to block the formation of any functional protein, and a second one has a lesion in a known functional domain, often at or near the site of the lesion in the participant's DNA. Finally, the Zebrafish Core analyzes mutant phenotypes caused by the induced mutations and compares them to the participant's clinical findings. Results are discussed with the CSs upon completion of the phenotyping.

Future goals include experiments to test for variant function by learning whether the human wild-type cDNA or mRNA can rescue the zebrafish mutant phenotype but the human mutant cDNA cannot. In cases in which the participant's phenotype is thought to arise due to dominant gain-of-function mechanisms, experiments to determine whether over-expression of the variant in wild-type fish results in phenotypes that mimic the participant's symptoms can provide valuable mechanistic information. In addition, future studies may model the human mutation even better by making the precise homologous change in the zebrafish protein, a goal that requires methodological innovation for improving zebrafish knock-in technology.

E. Procedures

Gene variant submissions

Each CS submits cases to the MOSC for analysis. Within the Gateway, CSs can enter clinical information about the participant, submit candidate variants for the MOSC, and explain the rationale for prioritizing candidate variants that may contribute to disease phenotypes. The Gateway tracks MOSC submissions and allows communication about the prioritization and progress back to the site.

Review process

Using information submitted though the Gateway, the MOSC reviews clinical phenotypes of participants as documented by the CS during the visit, reviews the list of submitted variants and assesses which variants are suited for functional analyses in model organisms. This process is an important step because it shapes all downstream procedures. The MOSC carefully reviews each of the genes and variants that the CSs submit and explores the vast amounts of information that are available in public human and model organism databases. The MOSC performs expansive bioinformatic analyses to asses submitted genes and their variants. The

MOSC developed a systematic Internet-based database search engine that integrates many of the steps that the MOSC first performed manually: MARRVEL (Model organism Aggregated Resources for Rare Variant ExpLoration; http://marrvel.org/). This search tool is publicly available and greatly reduces time spent to aggregate information used to prioritize variants (Wang *et al.*, 2017).

Genes and variants that are being considered are available for review within the Gateway, providing all members of the UDN access to this information. Collaborative analyses are pursued as described (Wangler *et al.*, 2017), involving the MOSC variants and those being analyzed by the Canadian Rare Disease Models and Mechanisms Network (RDMM), the Canadian equivalent of the UDN (http://www.rare-diseases-catalyst-network.ca/). In addition, the MOSC is currently trying to expand this approach with the Japanese (Initiative on Rare and Undiagnosed Diseases) and Oceania (UDP Australia and New Zealand) consortia that attempt to diagnose rare or unknown diseases primarily to find matching cases. This communication avoids duplication of work and fosters collaborations to identify additional participants who have genotypes and symptoms similar to the UDN case. Similarly, the MOSC collaborates with the Baylor-Hopkins Center for Mendelian Genomics (CMG) and with Baylor Genetics (a clinical diagnostic sequencing facility) to attempt to identify additional patients who carry mutations in the same gene. A comparison of phenotypes of the additional patients may then lower or elevate the probability that the candidate variant is causative, and hence the priority with which the MOSC continues the characterization of gene and variant function *in vivo*. These collaborative searches and inquiries can exclude some variants or genes. In some cases, additional patients with similar phenotypes may be found that together lead to definitive diagnosis without the need of model organism data (Schoch *et al.*, 2017). In other cases, the discovery of one or more additional patients with similar phenotypes in the collaborative search can significantly elevate the priority of the gene/variant to be studied in the MOSC.

Decision communication

The MOSC summarizes bioinformatic analyses in a PowerPoint format and sends the information back to the CSs for discussion. Genes and variants are classified into one of three categories based on human and model organism data that guide the ability to perform the experiments.

Category A genes are considered to be good candidates for the diagnosis for the participant based on segregation in the family, examination of other candidates in the participant, presence/absence of the variant in numerous databases queried by MARRVEL.org, available model organism data and reagents, and a feasible set of strategies in the model organisms. ***Category B*** and ***C*** genes do not meet the requirements to be pursued in the MOSC, and are handled based on the following principles.

Category B genes are either not considered good candidates based on the above available criteria or include cases that are considered to be solved. Alternatively, the variant may not be directly testable because some key reagents are not available (*e.g.*, no full length human cDNAs available or affordable) or the nature of the variant cannot be mimicked in the model organism (*e.g.*, splicing variants). For these cases, however, the MOSC is making every effort to improve research infrastructure to facilitate the study of these genes/variants. For example, in collaboration with the Drosophila Gene Disruption Project (BCM, Harvard, Carnegie) and the Drosophila Human cDNA Working Groups (BCM, Lawrence Berkeley National Laboratory, Kyoto Institute of Technology), the MOSC will create a T2A-GAL4 line and generate UAS-human cDNA transgenic stocks if the reagents are available. These genes can therefore more

easily be pursued in the future if other variants in these genes are discovered, or if additional reagents become available to the community.

Furthermore, the MOSC collaborates with the Knockout Mouse Phenotyping Project (KOMP) at BCM to knockout the orthologous gene in the mouse if a null allele has not been published in the literature. These mouse knockout lines will undergo systematic phenotyping based on the standards of the International Mouse Phenotyping Consortium (IMPC; http://www.mousephenotype.org/). In addition, for specific genes in which experts in the Drosophila and zebrafish community have established functional assays and are willing to collaborate with the UDN, the MOSC will facilitate communication between the Clinical Site and the model organism investigator on a case-by-case basis. The goal of these efforts is to provide new tools, speed up future research on the gene for the model organism community, and ultimately to benefit the UDN by creating larger collections of models that were initiated from the UDN. A typical example for a ***Category B*** gene is a candidate gene the UDN has proposed for a participant that seems to be solved by another gene. However, if the gene has a strong ortholog candidate in *Drosophila* that is not well studied, the MOSC creates a T2A-GAL4 line via MiMIC or CRIMIC and make this reagent publically available. The MOSC also creates human cDNA transgenic lines if available. Therefore, in some cases, the MOSC creates reagents for human gene rescue experiments for ***Category B*** genes.

Finally, ***Category C*** genes are not accepted by the MOSC for in vivo studies on the same basis as ***Category B*** genes, but in addition, the construction of appropriate reagents cannot be done for technical reasons. The above criteria are independent of funds or the number of submissions per site, and could be reevaluated when experiments fail to validate the ***Category A*** candidate gene.

Once the MOSC PIs come to a unanimous decision about the genes/variants submitted, a conference call is then organized with the CS within two months of submission for ***Category A*** genes. The genes/variant findings are then summarized and recommendations are discussed using the PowerPoint file described above. Selections or exclusions are discussed with the clinicians that submitted the cases. These discussions provide both parties with a much better understanding of the process and selection. Once both parties agree on the gene and variant to pursue (***Category A*** genes), the gene is assigned to one of the Cores depending on available reagents, the presence of good ortholog candidates in *Drosophila* and zebrafish, and the interests of the CSs and cores in particular model organism questions.

In general, genes of ***Category A*** that are conserved in *Drosophila* (~65% of human genes) are assigned to the *Drosophila* Core, and the remaining (~35%) genes are assigned to the Zebrafish Core. For ***Category A***, the goal of the MOSC is to validate the variant with strong conclusive model organism data that meet criteria for variant causality (MacArthur *et al.*, 2014) and lead to collaborative publications that drive discovery from the UDN. Not infrequently, ***Category A*** genes require several phone conferences with the MOSC and CSs. Variant causality is to be conclusively tested and conclusions are made available in the Gateway.

The goal is to provide a feedback of the acquired model organism data within 6 months of the sharing of the PowerPoint file for the *Drosophila* Core and 12 months for the Zebrafish Core. During this time, the CSs are asked not to publish or widely disseminate their observations without a discussion with and approval of the MOSC. Because experimentation must be customized for all genes and each gene has different available reagents and pre-existing data (see core operations above), the MOSC faces a number of technical and biological challenges

(e.g., human cDNA not rescuing the mutant phenotype, no obvious scorable phenotype observed in mutants) that are unique to each gene. Therefore, the 6/12 months guideline should be considered as a turnaround time for communication of results obtained at that time, and additional time may be required to complete each project.

Communication with CSs after data generation

After generation of mutant models and characterization of mutant phenotypes, the MOSC contacts CSs to discuss observations. The CSs often raise questions and issues that may be addressed in subsequent experiments. Although the MOSC carries out characterization of the gene in mutant models, the CSs often pursue identification of additional patients to provide more robust evidence that the mutation observed in the participant is indeed associated with the observed phenotype. These data are integrated as part of the characterization to provide a final and compelling diagnosis.

Throughput and allocation

CSs submit approximately 200 variants per year. These variants can be subdivided in two groups: ~100 variants that fall in ***Category C*** and the remainder fall into about 30 to be analyzed in zebrafish and about 70 to be analyzed in *Drosophila*. The Zebrafish Core can produce and analyze ~15 variants per year for the ***Category A*** genes. The remainder of these 30 genes (***Category A*** and ***B***) are submitted to the BCM KOMP to create and phenotype mouse knock out strains. The *Drosophila* Core can pursue 30 genes per year of ***Category A*** for in-depth analysis and work on 40 genes per year for ***Category B*** to generate valuable reagents for the community. If the ***Category A*** genes exceed the MOSC's capacity, the genes are prioritized on a first come, first served basis. However, such genes are integrated into ***Category B*** and areprioritized for generation of reagents.

Distribution of reagents

All the fly stocks are deposited in the Bloomington Drosophila Stock Center at Indiana University (BDSC, http://flystocks.bio.indiana.edu/) and the fish resources are deposited at the Zebrafish International Resource Center (ZIRC, https://zebrafish.org/home/guide.php). These two facilities keep records of stocks they distribute, and these records document the usage of the MOSC materials. In addition, the vectors containing the UAS-human cDNAs are deposited in the Drosophila Genome Research Center (DGRC) in Indiana and are publicly available. All these distribution centers are supported by the NIH. Finally, the reagents (fish, flies and cDNAs) are or will be available for sharing with the CSs as soon as they are generated. Note that the MOSC is in the process of acquiring a human cDNA library containing about 40,000 unique cDNA clones assembled and generated in the past 6 years by Dr. Ken Scott at BCM.

F. References

1. Bellen, H. J., and Yamamoto, S. (2015). Morgan's Legacy: Fruit Flies and the Functional Annotation of Conserved Genes. Cell 163, 12-14.
2. Chao, H. T., Davids, M., Burke, E., Pappas, J. G., Rosenfeld, J. A., McCarty, A. J., Davis, T., Wolfe, L., Toro, C., Tifft, C., *et al.* (2017). A Syndromic Neurodevelopmental Disorder Caused by De Novo Variants in EBF3. Am J Hum Genet 100, 128-137.

3. Diao, F., Ironfield, H., Luan, H., Diao, F., Shropshire, W. C., Ewer, J., Marr, E., Potter, C. J., Landgraf, M., and White, B. H. (2015). Plug-and-play genetic access to drosophila cell types using exchangeable exon cassettes. Cell Rep 10, 1410-1421.
4. Harel, T., Yoon, W. H., Garone, C., Gu, S., Coban-Akdemir, Z., Eldomery, M. K., Posey, J. E., Jhangiani, S. N., Rosenfeld, J. A., Cho, M. T., *et al.* (2016). Recurrent De Novo and Biallelic Variation of ATAD3A, Encoding a Mitochondrial Membrane Protein, Results in Distinct Neurological Syndromes. Am J Hum Genet 99, 831-845.
5. Luo, X., Rosenfeld, J. A., Yamamoto, S., Harel, T., Zuo, Z., Hall, M., Wierenga, K. J., Pastore, M. T., Bartholomew, D., Delgado, M. R., *et al.* (2017). Clinically severe CACNA1A alleles affect synaptic function and neurodegeneration differentially. PLoS Genet 13, e1006905.
6. MacArthur, D. G., Manolio, T. A., Dimmock, D. P., Rehm, H. L., Shendure, J., Abecasis, G. R., Adams, D. R., Altman, R. B., Antonarakis, S. E., Ashley, E. A., *et al.* (2014). Guidelines for investigating causality of sequence variants in human disease. Nature 508, 469-476.
7. Ramoni, R. B., Mulvihill, J. J., Adams, D. R., Allard, P., Ashley, E. A., Bernstein, J. A., Gahl, W. A., Hamid, R., Loscalzo, J., McCray, A. T., *et al.* (2017). The Undiagnosed Diseases Network: Accelerating Discovery about Health and Disease. Am J Hum Genet 100, 185-192.
8. Schoch, K., Meng, L., Szelinger, S., Bearden, D. R., Stray-Pedersen, A., Busk, O. L., Stong, N., Liston, E., Cohn, R. D., Scaglia, F., *et al.* (2017). A Recurrent De Novo Variant in NACC1 Causes a Syndrome Characterized by Infantile Epilepsy, Cataracts, and Profound Developmental Delay. Am J Hum Genet 100, 343-351.
9. Venken, K. J., Carlson, J. W., Schulze, K. L., Pan, H., He, Y., Spokony, R., Wan, K. H., Koriabine, M., de Jong, P. J., White, K. P., *et al.* (2009). Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat Methods 6, 431-434.
10. Venken, K. J., Schulze, K. L., Haelterman, N. A., Pan, H., He, Y., Evans-Holm, M., Carlson, J. W., Levis, R. W., Spradling, A. C., Hoskins, R. A., and Bellen, H. J. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nat Methods 8, 737-743.
11. Wang, J., Al-Ouran, R., Hu, Y., Kim, S. Y., Wan, Y. W., Wangler, M. F., Yamamoto, S., Chao, H. T., Comjean, A., Mohr, S. E., *et al.* (2017). MARRVEL: Integration of Human and Model Organism Genetic Resources to Facilitate Functional Annotation of the Human Genome. Am J Hum Genet 100, 843-853.
12. Wangler, M. F., Yamamoto, S., Chao, H. T., Posey, J. E., Westerfield, M., Postlethwait, J., Hieter, P., Boycott, K. M., Campeau, P. M., and Bellen, H. J. (2017). Model Organisms Facilitate Rare Disease Diagnosis and Therapeutic Research. Genetics 207, 9-27.
13. Yoon, W. H., Sandoval, H., Nagarkar-Jaiswal, S., Jaiswal, M., Yamamoto, S., Haelterman, N. A., Putluri, N., Putluri, V., Sreekumar, A., Tos, T., *et al.* (2017). Loss of Nardilysin, a Mitochondrial Co-chaperone for alpha-Ketoglutarate Dehydrogenase, Promotes mTORC1 Activation and Neurodegeneration. Neuron 93, 115-131.

XVI. Appendices

APPENDIX 1: NIH UDP Protocols

Screening (30 inquiries each week)

A Patient Care Coordinator (PCC), selected for having pleasing but firm interpersonal skills, provides a central point for all inquiries that range in specificity from direct physician-to-physician referrals to cold calls to NIH Call Center (866-444-8806) from patients or family members seeking to learn more about the UDP. The NIH Call Center refers these calls to the PCC. Whatever the source of inquiries, the PCC mails the potential participant (or family, in the case of pediatric patients) an invitation package that includes a cover letter and an attached frequently asked questions document. A second letter is sent that the patient can share with his/her physician with an attached form for listing contact information of the current attending physician, a list of prior hospitalizations, and specialists that have been involved in the patient's care. This is often followed by phone exchanges with the PCC to clarify goals and structure of the program and the information required for further evaluation. Descriptions of the detailed flow of participants prior to CRC admission are below.

Substantial delays are often encountered at this phase of patient recruitment as families often request medical records from multiple institutions, reflective of the long diagnostic odyssey. The UDP believes it is essential to obtain a physician recommendation letter in order to provide a clear, current picture of the patient's illness and to ensure follow-up care after completion of the UDP evaluation.

As detailed later, initial UDP medical review requires complete records of previous care and evaluations. Patients and physicians may encounter problems with collection of results of prior blood work, imaging, and special tests as they negotiate retrieval of these materials from various health care facilities. A series of form letters are used to remind potential participants of documentation required, but not yet received, including prior phenotyping and a physician's recommendation letter.

Clarification of the goals of the UDP sometimes results in withdrawal of applicants who have been interested only in a 'second opinion' process. Potential participants who fail to provide the necessary phenotyping data, or for whom there is no physician recommendation letter, will not be further considered. Approximately two-thirds of patients who were initially interested in learning more about the UDP or in participating fail to complete the information gathering process trimming the 30 per week who express an interest in the program to 10 who remain interested and whose records can be gathered and reviewed. For pediatric patients only approximately one-third of families fail to complete the process or are found ineligible, usually because they already have a diagnosis (e.g., they have an unbalanced chromosome translocation with multiple malformation syndrome but the family does not think this is the answer).

Creation and careful review of a CRC medical record (10 applicants each week)

The next step in the recruitment process is to carry out a detailed review of each candidate's medical record, including the study recommendation letter from the current personal physician or physician- extender summarizing the salient features of the applicant's symptoms, with reports detailing already collected phenotyping. These reports might include personal and family

health history, physical examination, blood work and urine analysis/chemistries, imaging, and special studies such as CSF findings, electromyography (EMG), photos of skin lesions, and videos that display abnormalities of balance, gait, and strength. If biopsy or surgical procedures have been performed, biopsy slides may be reviewed by CRC pathologists if this appears to be essential for a decision. Prior imaging, especially CT and MR imaging, is extremely important in the review process, and every effort is made to obtain the images themselves, and not simply reports. The clinical records available vary across applicants, since some have had extensive prior evaluations by skilled physicians and others have had only a limited approach to finding a diagnosis.

The completed file is assigned to UDP team members and/or consultants to evaluate the likelihood that a rare or yet-to-be described disease is present and that the focused, systematic UDP approach might lead to a diagnosis. Useful indicators include other affected family members, objective physical findings, abnormalities found in blood work and/or imaging or other clues pointing to the presence of significant disease. A further consideration is whether, depending on family size and the availability of blood specimens on additional affected family members, the UDP's diagnostic armamentarium, especially SNP arrays or whole exome sequencing, could be useful in providing an answer. The review is physician intensive, and because records are often very extensive, the review process may be lengthy. Moreover, it may prove necessary to request additional information, or the advice of other UDP consultants. While the principal goal of this review process is to select participants for UDP evaluation, there are other potential results. Some applicants may be more suitable for referral to other open NIH research protocols. If, in the judgment of the review panel, there has been incomplete evaluation, the panel may choose to return the applicant to the referring physician with suggestions for further diagnostic approaches or recommend referral to an appropriate academic medical center.

The decision to invite applicants to travel, expense-free, to Bethesda, MD for a 5-day admission to the NIH CRC is made by Program Directors after detailed discussion with consultants and other members of the UDP team. The goal of this review process is to ensure, to the extent possible, that the accepted applicants are appropriate and match the resources of the UDP. A criterion for acceptance to the UDP is that the applicant is safe to travel. The pediatric applicants in particular are often medically fragile, medical clearance for commercial travel must be documented by the referring physician before applicants can be accepted and scheduled for evaluation. Pediatric applicants must have clearance from their physician one week prior to making the trip. In some cases, visits need to be rescheduled if the applicant is too sick to travel. UDP does not have the ability to pay for hospital-to-hospital transports, nor can they carry out these transports.

Preparation for the 5-day evaluation

Fitting the required diagnostic efforts into a 5-day evaluation requires careful planning to complete thorough phenotyping and place the findings in context for anxious participants and their families. This planning is complex and involves scheduling heavily used imaging resources and other diagnostic tests and insuring that initial evaluations by sub-specialists can be performed in a timely fashion. Participant-specific time constraints must also be considered.

Overall approach to phenotyping and specific data gathering in common subgroups

More than half of the participants accepted into the UDP have a neurological phenotype, and in children in particular this leads to a common phenotyping framework that includes intracranial

imaging (magnetic resonance imaging (MRI) and magnetic resonance spectroscopy (MRS)), neurologic consultation, electroencephalogram (EEG), EMG/nerve conduction velocity (NCV), lumbar puncture for CSF neurotransmitters and other special testing, skin biopsy both for fibroblast culture for functional verification of new candidate genes and for immunohistochemistry and electron microscopy, ophthalmologic exam under anesthesia, physiatry consultation, and neurocognitive testing. In adult participants, CSF is also obtained for immunologic studies.

UDP metrics

The UDP has had much success, as can be seen from the following metrics:

a) Metrics in the 4 years following the establishment of the UDP:
   - 6,300 inquiries evaluated
   - 2,300 physician referral letters with applicant medical records reviewed
   - 450 participants admitted to the NIH-CRC (Clinical Research Center)

b) Weekly metrics:
   - 30 new inquiries
   - 10 applicants with completed referral letters and results from prior diagnostic efforts evaluated
   - 3 participants/families admitted for work-up at the CRC

c) Diagnostic metrics:
   - Approximately 100 participants (20-25%) were diagnosed with rare to extremely rare diseases
   - Two participants were found with diseases unknown to medicine
   - 15 genes not previously associated with human disease were discovered and tentatively related to disease phenotypes

Summary

The NIH UDP initial approach to identifying and evaluating participants with undiagnosed diseases has been refined and focused over nearly five years. It seeks to identify participants who are most likely to have a rare or unknown undiagnosed disease. The 5-day admission to the NIH is designed to define the underlying pathophysiology by careful phenotyping and to identify settings in which genomics may prove useful.

References

1. Gahl WA, Tifft CJ. The NIH Undiagnosed Diseases Program: Lessons learned. JAMA 305:1904-5, 2011.
2. Gahl WA, Markello TC, Toro C, Fuentes Fajardo K, Sincan M, Gill F, Carlson-Donohoe H, Gropman A, Pierson TM, Golas G, Wolfe L, Groden C, Godfrey R, Nehrebecky M, Wahl C, Yang S, Madeo A, Boerkoel CF, Tifft CJ, Adams D. The NIH Undiagnosed Diseases Program: Insights into rare diseases. Genet Medicine 14:51-9, 2012.
3. Gahl WA, Boerkoel CF, Boehm M. The NIH Undiagnosed Diseases Program: Bonding scientists to clinicians. Dis Model Mech 5:3-5, 2012.
4. St. Hilaire C, Ziegler SG, Markello T, Brusco A, Groden C, Gill F, Carlson-Donohoe H, Lederman RJ, Chen MY, Yang D, Siegenthaler MP, Arduino C, Mancini C, Freudenthal B, Stanescu HC, Zdebik A, Chaganti K, Nussbaum R, Kleta R, Gahl WA, Boehm M. *NT5E* mutations and arterial calcifications. N Engl J Med 364:432-42, 2011.

Figure 1: NIH UDP participant flow

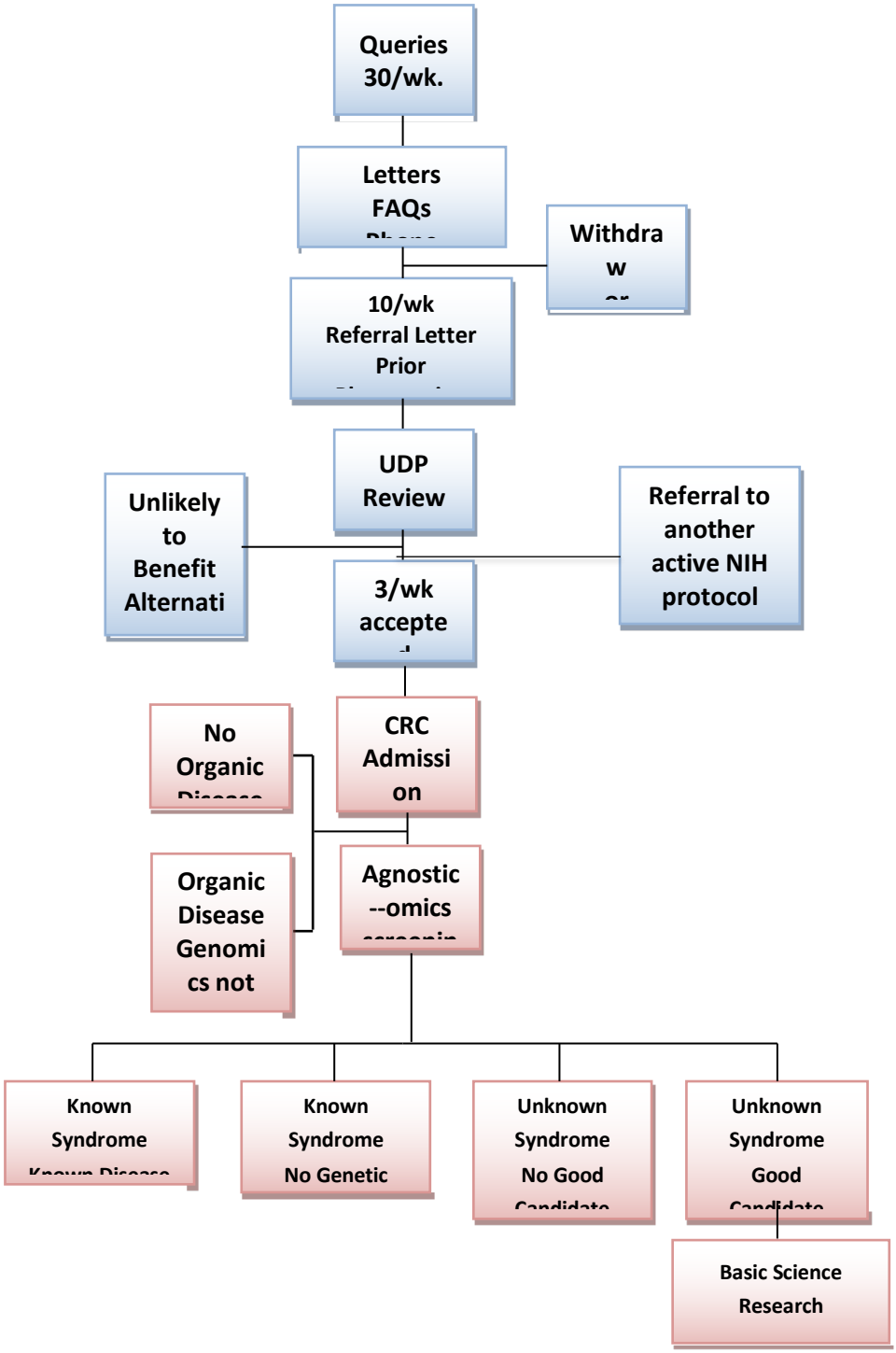
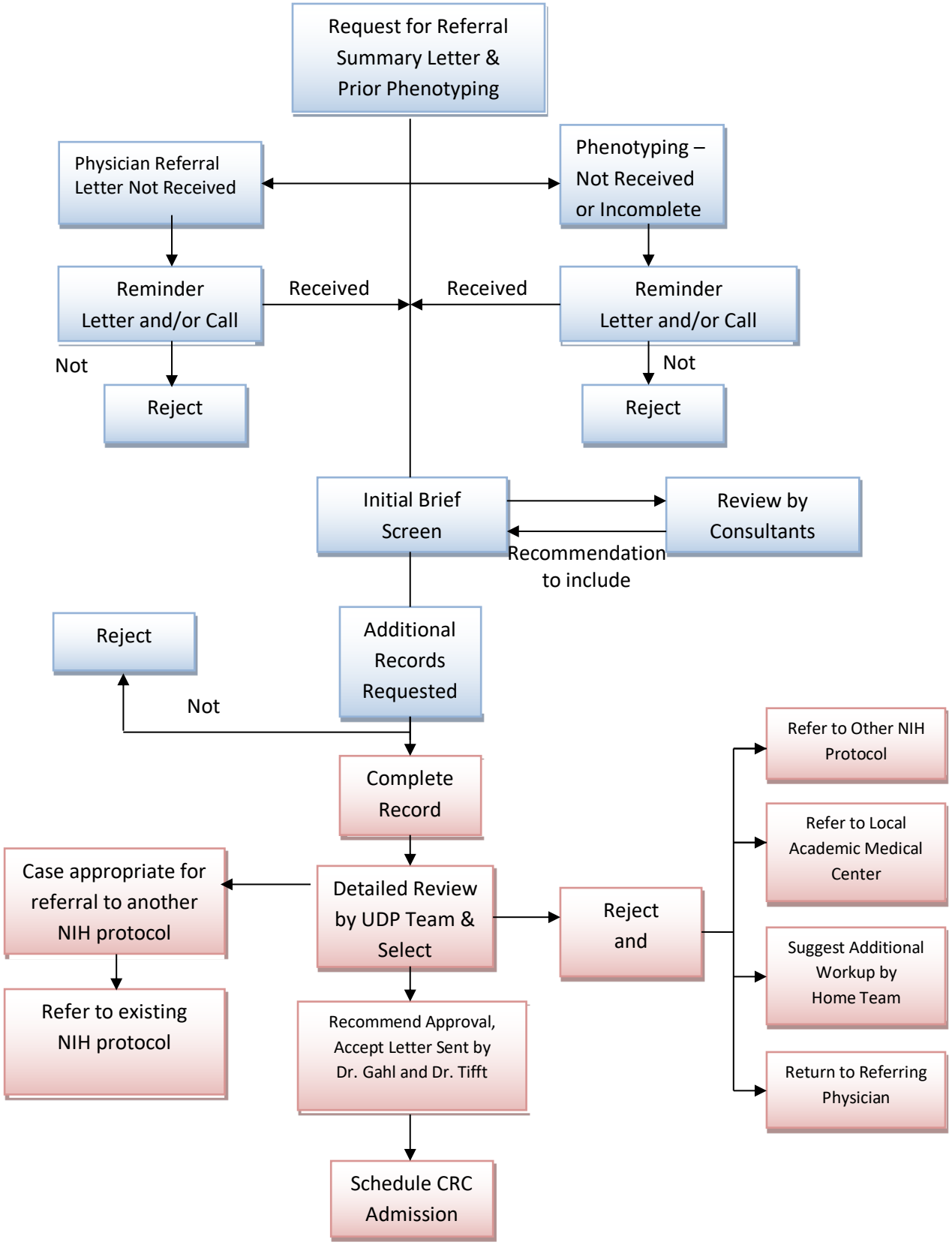


Figure 2: NIH UDP pre-CRC admission



APPENDIX 2: Example Study Recommendation Letters

Example Study Recommendation Letter #1 (pediatric)

Hospital/clinic logo here

To Whom it May Concern:

We are writing to you to request consideration of siblings, [patient names], for enrollment in the Undiagnosed Diseases Network (UDN). [Patient names] are followed by multiple specialists at [hospital name]. They are also followed by local pediatrician [physician name] for routine pediatric care.

[Patient name] is now a [age] year old [gender] with a history of dysmorphic features, failure to thrive, and hepatomegaly of unknown origin. Due to cryptogenic cirrhosis, liver transplant was performed at [age] months of age. Pathology results of [patient name]'s previous liver biopsy was suspicious for a [condition], specifically [specific condition], however, molecular testing for the [gene name] was negative. [Condition] enzyme screening and [condition] screen for the explanted liver sample came back in the low ranges, but not in the deficiency range usually seen. Additional extensive workup was unrevealing.

[Patient name] is now a [age] month old [gender] noted prenatally to have holoprosencephaly via fetal MRI at [time] weeks gestation. Brain MRI performed on DOL [time] was consistent with [description of MRI]. [He/she] was admitted at [age] months of life for evaluation of liver steatosis, microcephaly, and failure to thrive. At [age] months of age, [patient name] was identified to have new onset hepatomegaly in [month] with vomiting. A liver biopsy from [date] identified [results of liver biopsy]. Due to persistent FTT, G tube was placed in [month] with subsequent fungal peritonitis, now post-[time] day course of [medication]. [Patient name] continues to have daily emesis. [Patient name] is currently evaluated for liver transplantation (persistently elevated transaminases and synthetic dysfunction).

Of note, both siblings have a history of IUGR with failure to thrive, improved for [patient name] following liver transplantation. [Patient name] has a history of developmental delays, making significant progress with therapies, and now within normal limits. [Patient name] continues to have developmental delays and facial features similar to [his/her] [brother/sister] during infancy. Given the similarities in the presentation of these two siblings with an unremarkable family history (parents are not consanguineous), whole exome sequencing was obtained for [patient name] and identified a heterozygous mutation in the [gene name], which in the homozygous state is associated with [condition]. Subsequent deletion/duplication testing via the MitoMet oligonucleotide array returned as normal. Mitochondrial genome testing via

massively parallel sequencing was obtained for [patient name] and was unrevealing.

At this time, we are unable to identify a specific genetic etiology that would explain the findings seen in [patient name] and [patient name]. The presentation of two siblings with similar features, however, is suggestive of a possible autosomal recessive condition, which remains undiagnosed at this time. Parents are interested in identifying a diagnosis, and are also interested in having a third child. We would like to refer these siblings to the Undiagnosed Diseases Network for further evaluation to try and identify a diagnosis. Thank you for your review and consideration for acceptance into the program. Please do not hesitate to contact our office at [phone number] if you have any questions or require any additional materials.

Sincerely,

Healthcare provider signature

[Healthcare provider name]

Example Study Recommendation Letter #2 (pediatric)

Hospital/clinic logo here

To Whom it May Concern:

I have followed [patient name] since [age] months of age. [He/she] has a history of significant global developmental delay, [he/she] is nonverbal, has hyperoral behavior, macrocephaly, small stature, [further description]. [He/she] has a great disposition and visually interacts with [his/her] environment. [He/she] has continued to make very slow but steady motor development but has never developed speech. [He/she] has never had seizures or developmental regression. Significant genetic, metabolic, and neurodiagnostic evaluation (as listed below) has yet to yield an underlying diagnosis. I am referring [him/her] to the Undiagnosed Diseases Network in attempts to find a unifying diagnosis for [his/her] multitude of symptoms. I truly feel that there is an underlying metabolic or genetic cause for [his/her symptoms that our testing thus far has not uncovered. [He/she] has been seen by numerous other specialists across the country.

[His/her] evaluation to date includes:

Normal or negative metabolic studies:
Urine organic acids
Serum amino acids
Creatinine guanidinoacetate
Etc.

Normal or negative genetic studies:
Routine chromosomes
Chromosomal microarray [year]
mtDNA point mutations and deletions
GeneDx 101 mitochondrial nuclear gene panel
Etc.

Neuroimaging/neurodiagnostics:
[year]--- MRI showed [results]
[year] CT showed [results]
Etc.

Normal or negative CSF studies:
Neurotransmitters
Biopterin

I truly appreciate your consideration for evaluation for [patient name]. This family has been on a very long quest to find a diagnosis and would be grateful for the opportunity to have [him/her] evaluated through the UDN.

Sincerely,

Healthcare provider signature

[Healthcare provider name]

Example Study Recommendation Letter #3 (pediatric)

Hospital/clinic logo here

Dear Undiagnosed Diseases Network,

I wholeheartedly recommend [patient name] to be evaluated by the Undiagnosed Diseases Network. [He/she] is a [age] year old with persistent myalgias, dyspnea, [description of condition] of unknown etiology. There are several other family members who are less severely affected with similar symptoms, suggesting a genetic etiology.

I recently met [patient name] to evaluate him for endocrinologic involvement of [his/her] presentation. While I did not find any endocrine pathology, I wanted to take the opportunity to refer him to the UDN. The notes from [his/her] neurologist Dr. [name] will have more details on his history, but I will describe the summary of what I learned.

[patient name] currently presents with [symptoms]. [Patient name]'s family reports that [patient name]'s symptoms initially began at [age] years of age when he began complaining of leg pain out of proportion to those expected for his age. He was evaluated at [age] years of age by a rheumatologist at [hospital], and then by Neurology where a deltoid biopsy was performed and reportedly normal. Additionally, other genetic testing for different forms of [condition] was negative. [He/she] was then referred to Dr. [name] at [hospital]. An EMG was normal, but a quadriceps biopsy showed a predominance of [finding] of unclear significance.

[Patient name]'s symptoms have all progressed over time. [He/she] complains of significant exercise intolerance and weakness in all muscles that have been slowly worsening over time. [His/her] weakness is particularly extreme after activity. Additionally, [patient name] has pain in [his/her] legs, around [his/her] neck, and lower back that is present all the time, although also worsened with activity. [He/she] has seen some improvement in the pain, especially in [area], after starting [medication]. [He/she] occasionally tries [medication] without much relief. The pain is particularly bad [time of day] whereas [his/her] other symptoms seem to be more extreme [time of day]. [His/her] [parent] notes that [he/she] also has some ptosis and [symptom] on several mornings when [he/she] wakes up that sometimes persists later in the day. Initially, this was one-sided, predominantly on the [side], but now appears to be bilateral. [Patient name] walks when [he/she] is at home but uses a wheelchair for transportation of further distances. [He/she] also appears to have [symptoms]. [His/her] motor strength and reflexes, however, are typically normal when [he/she] is evaluated in the neuromuscular clinic, suggesting that [patient name] has more trouble with fatigue than baseline muscle weakness. [He/she] also has a normal serum CK level. [He/she] has had evaluations for [syndromes] that were negative. [He/she] also had an empiric

trial of [medication] that did not improve [his/her] symptoms. Dr. [name] most recently requested a [test] given the [symptoms].

[Patient name] recently developed [symptom] on [his/her] back, which particularly precipitated the referral to my clinic. I did not feel any sign of excess [hormone]. [He/she] was also evaluated by dermatology who felt these to be [condition]. Additionally, [he/she] has a [birthmark].

[Patient name] has been seen by several other specialists. [He/she] follows with Dr. [name] at [hospital] for pulmonary and has been noted to have [symptom]. [He/she] also was briefly followed by Dr. [name] in [state] at [hospital] for some time, but no further diagnoses were noted. [He/she] has been evaluated by Cardiology with a normal echo and EKG. [He/she] has also been evaluated by Physical Therapy, who did not think that [he/she] would benefit from their intervention due to [his/her] exercise intolerance.

[Patient name]'s family history is of particular interest. [His/her] [Parent] is healthy other than migraines and is of [ethnicity] background. [His/her] [Parent] is healthy and of [ethnicity] descent. There is no consanguinity in the family. [Patient name] has [number] siblings. [His/her] oldest sibling is [age] years old with some slight degree of muscle weakness as well. [He/she] has [number] healthy child and is currently pregnant with no complications. [Patient name]'s oldest brother is [age] years old, and his next sibling is [age] years old. Both of them are healthy except for some asthma and allergies. [Patient name] has an [age]-year old sibling who has joint and muscle problems that are not as severe as [patient's].

Thank you for your consideration of [patient's] application.

Sincerely,

Healthcare provider signature

[Healthcare provider name]

Example Study Recommendation Letter #4 (adult)

Hospital/clinic logo here

Dear Undiagnosed Diseases Network Team,

I propose my patient [name] for your special protocol in the Undiagnosed Diseases Network. When I learned of your protocol, I immediately thought of [him/her]. [He/she] seems an ideal participant in your program.

Symptoms & History: [Name] suffers from an excruciating and bizarre illness that has devastated [his/her] life and gone undiagnosed for [number] years despite exhaustive workups at [institution] and here at [institution]. [He/she] has consulted over 100 medical specialists of whom many are at the pinnacle of their fields. [name] is a pleasant, intelligent [man/woman] and a motivated, cooperative patient.

- [Name] is a fair-skinned [age]-year old [man/woman] who has been disabled for the last [number] of years by burning facial pain and flushing of elusive etiology. [His/her] entire face and ears are involved; they are inflamed, red, and hot to the touch.
- Onset was rapid and for no apparent reason. Prior to the illness, [he/she] was in excellent health, a parent with a healthy child and successful professional who worked full-time.
- The facial pain requires [name] to remain nearly all the time in a cold room with a fan blowing directly on [his/her] face. [more explanation]
- While [name]'s face and ears are chronically hot, the rest of [his/her] body [description].
- [Name] has anhidrosis over 90% of [his/her] body. However, sweating that cannot be elicited by heat can sometimes be elicited with [system] stimulation.
- [He/she] developed [eye condition] in [his/her] [age], since remedied surgically.
- Other major symptoms include: [additional symptoms]

Diagnostics & Etiology: [Name]'s case is a medical mystery cutting across many organ systems/braches of medicine. One might describe it functionally as a putative sympathetic neurologic disorder of the thermoregulatory system that especially affects the vasculature and skin of the head. The origin of the proposed neuropathy could be genetic, autoimmune, infectious, toxicological, or some combination.

There are a number of tantalizing but unexplained clues including:

1. [He/she] is a carrier of one copy of the gene for the rare recessive genetic disease [condition], of which [his/her] relative died. But the [condition] experts have never seen

symptoms manifested in a [condition] carrier.

2. [Protease] levels are chronically high, but not high enough for [condition]/
3. [He/she] tests relatively normal on most blood and urine diagnostics, but with some curious exceptions: high on [tests]. Low on [tests].
4. [Medication] has a minor positive effect on [his/her] symptoms and [he/she] takes it on an ongoing basis. This is the most helpful of the 100 or so medications that have been tried.
5. [He/she] has idiosyncratic negative reactions to many medications, often responding to "subclinical" doses.
6. [Name] was on a course of the medication [medication name] when [his/her] illness started, but there are no other documented cases of such a reaction to this medication.
7. A number of surgical sympathetic blocks have been implemented on a temporary basis, sometimes with great beneficial effect and sometimes the opposite.
8. Her illness bears some similarity to [condition], itself a rare and largely unexplained disease. However, [condition] affects the feet and sometimes the hands, and there is little or no reference in the literature to a similar disease affecting only the face.

Records: [Name] has carefully retained and organized the voluminous diagnostics and reports on [his/her] condition over [time] years seeking a diagnosis and treatment. This should be helpful to your efforts.

My role: While I am a [specialist] in private practice, I have served as [his/her] primary physician since very early in the illness. I would be pleased to support your efforts and provide follow-up. I understand that several other physicians that regularly see [name] are also in support of [his/her] application and would be available to communicate with you if requested.

Patient's perspective: [Name] has been exhaustive and courageous in seeking an explanation for this illness. [He/she] read about your program in [magazine]. [He/she] fully understands that your program is primarily for research purposes and that the chances of significant benefits from participating are rather small. Please consider [him/her] for your program. My contact information and [his/hers] appears below.

Sincerely,

Healthcare provider signature

[Healthcare provider name]

APPENDIX 3: Applicant Review Form (completed by Clinical Sites)

Applicant name: _____

UDN identifier: _____

Date of birth: _____

Date application submitted: _____

**Auto-populates from Gateway application*

UDN site: _____ ***Auto-populates from site assignment***
Name of primary reviewer(s): _____
Category of primary condition (*drop down list)***:**

- Allergy/immunology
- Cardiology and vascular conditions
- Dentistry and craniofacial
- Dermatology
- Endocrinology
- Fibromyalgia/chronic fatigue syndrome
- Gastroenterology
- Gynecology and reproductive
- Hematology
- Infectious disease
- Musculoskeletal and orthopedics
- Nephrology
- Neurology
- Oncology
- Psychiatry
- Pulmonary
- Rheumatology
- Multiple pediatric (multiple congenital anomalies)
- Other
- None of the above

Please provide a narrative summary (150-200 words) of the applicant's condition. If applicable, please include:

- History of present illness
- Date symptoms first noted
- Past medical history
- Previous diagnoses/comorbidities (using ICD terms if possible)
- Prior procedures and surgeries.

Height:
Weight:
Head circumference:
Please indicate the applicant's pertinent prior evaluations. If applicable, please include:

- Prior positive or negative test results
- Prior genetic testing (especially whole exome sequencing)

Provisional diagnosis/working plan:

Other family members affected: Yes/No

- If yes:
  - How many affected? _____
  - How many available for analysis? Unknown/Some/All/None

Applicant images: *Attach files*

Other files: *Attach files*

Category 1: Inclusion/Exclusion Criteria

Inclusion Criteria

- ☐ Does Not Have Diagnosis Explaining Objective Findings
- ☐ Agrees to Storage and Sharing of Information & Biomaterials

Exclusion Criteria

- ☐ Has Diagnosis Explaining Objective Findings
- ☐ Diagnosis Suggested Based on Record Review; Further Evaluation Unnecessary
- ☐ Too Ill to Travel Safely to UDN site

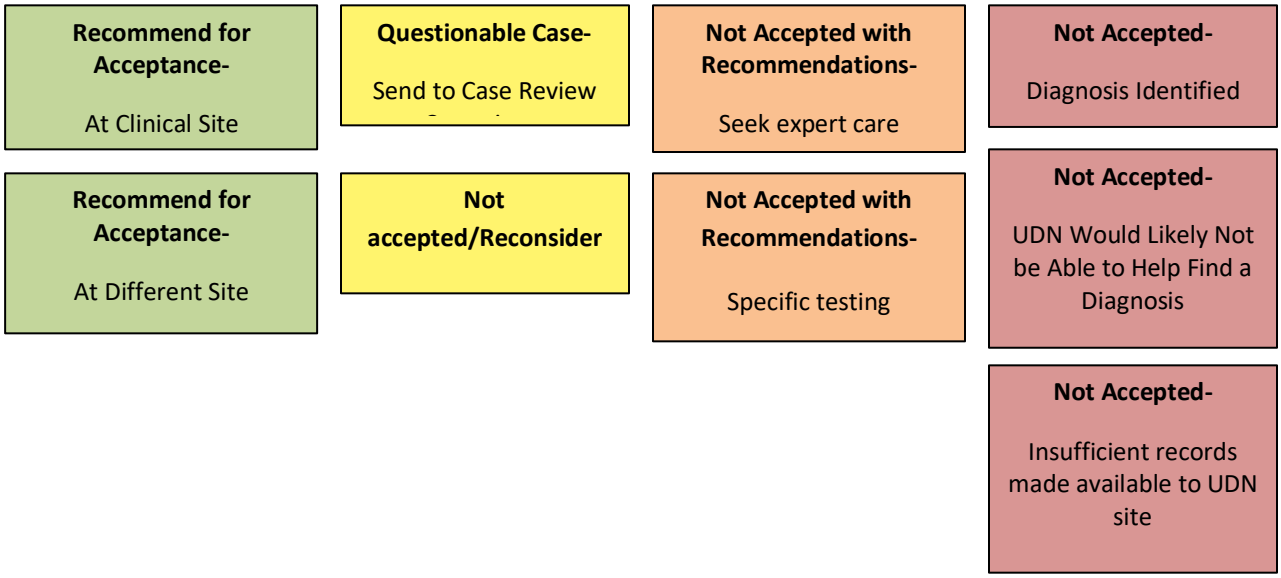
Category 2: Strengths (≥3 Recommended)

- ☐ Objective Abnormal Finding(s)
- ☐ Unique Clinical Presentation
- ☐ Multiple Systems Affected
- ☐ Family History of Condition
- ☐ Relevant Family Members Available for Testing
- ☐ High Likelihood of Genetic Diagnosis
- ☐ Local Patient
- ☐ Relevant to Other UDN Participants
- ☐ Can Offer Sequencing
- ☐ Can Offer Additional Clinical Workup

☐ Other

Category 3: Limitations (<1 Recommended)

☐ No Relevant Family Members Available for Genetic Testing
☐ UDN Resources Not Appropriate for Case
☐ High Likelihood of Not Solving Case at Present
☐ Proband Likely to Refuse Certain Tests/Procedures
☐ No Objective Clinical Findings
☐ Other



APPENDIX 4: Template Letters for Patients and Health Care Providers

Description: *Information for patients*

Date:
Address:

Dear [patient]:

Thank you for your interest in the Undiagnosed Diseases Network (UDN). Participants accepted into this program will be part of a clinical research study aimed at answering questions about medical conditions that have eluded diagnosis. We hope to advance medical knowledge in ways that can help improve health care for everyone. The study will be conducted at Baylor College of Medicine, Boston Children's Hospital/Brigham and Women's Hospital/Massachusetts General Hospital, Duke University, the National Institutes of Health, Stanford University, University of California-Los Angeles, and Vanderbilt University Medical Center.

Please discuss your participation in this program with your primary healthcare provider. Important considerations include:

- Many cases accepted will NOT result in a diagnosis.
- A referral by a healthcare provider is required.
- The provider who refers you will be asked to provide your medical information.
- The UDN will communicate the decision on accepting your case for evaluation in writing to you and your referring healthcare provider.
- If your case is accepted for UDN evaluation, the UDN will provide information from the evaluation to you and to your healthcare provider.
- Your healthcare provider will be responsible for your medical care after you have been evaluated in the UDN.

Details about the information needed from your referring provider are on the attached letter. Please insert information where requested and give the *Information Sheet for Referring Healthcare Providers* to your provider.

A UDN staff member will notify you when the information from your provider has been received. Once all materials are received, UDN review is expected to take about six to eight weeks.

Again, thank you for your interest in the UDN. Medical advances depend on individuals like you who volunteer as partners in medical discovery. More information about this new program is on line at https://undiagnosed.hms.harvard.edu.

The Undiagnosed Diseases Network Team

Description: *Information for healthcare providers*

Information for healthcare providers

Your patient has contacted the Undiagnosed Diseases Network (UDN) about participating in this research study. Patient participants will be evaluated using the unique combination of scientific and medical expertise and resources. Participants must have a condition that has not been diagnosed following a thorough medical evaluation. More information about the UDN can be found here: https://undiagnosed.hms.harvard.edu.
If your patient decides to apply for this study, he/she will need to answer a series of brief questions through our application gateway (https://gateway.undiagnosed.hms.harvard.edu) related to the following:

- Contact information
- Demographic information
- Primary licensed healthcare provider information (name, address, phone, FAX, email)
- Brief medical history questions
- Previous evaluations
- Travel limitations

A study recommendation letter from a licensed healthcare provider is also necessary in order to submit the online application to the UDN. The patient will be asked to upload this letter to the online application. The study recommendation letter must include:

- A summary of the applicant's medical problems
- Date when symptoms were first noticed
- Previous diagnoses
- History of evaluations and testing
- History of treatments and medications
- Current medications
- Family history
- Healthcare provider's diagnostic impressions
- For pediatric patients, prenatal and birth history should also be provided.

After an application is submitted, it is then sent to one of the UDN clinical sites to review. The UDN clinical sites typically contact applicants within one month of application submission to request additional information, such as:

- Medical records
- Laboratory studies
- MRI, X-ray, or CT images (can be sent on CD)
- Photographs or videos
- Pathology slides and reports
- For pediatric patients: growth curves

In order to help with the application review process, the patient can begin to collect these records now. However, medical records should not be sent until specific instructions to do so

are received from a UDN clinical site.

After all of the necessary information is received by a UDN clinical site, it typically takes 6-8 weeks for the network to make a decision whether to accept an applicant.

If your patient's case is accepted for UDN evaluation, the UDN will provide information from the evaluation to you and to your patient. You will be responsible for your patient's follow up medical care.

Thank you for considering this opportunity to consult with the UDN on your patient. We appreciate your commitment to providing the best possible care for your patients in ways that help advance medical knowledge and discovery.

The Undiagnosed Diseases Network Team

Undiagnosed Diseases Network

Your institutional logo here

Description: *Medical record request- adult*

Date:

Address:

Dear [patient]:

Thank you for your interest in the Undiagnosed Diseases Network (UDN). You have been assigned to the [site name] in [location] for consideration. In order to properly review your case, we will need the following medical records:

_____ Admission and discharge summaries from hospitalizations related to your condition
_____ Consultant notes (such as cardiologist, gastroenterologist, neurologist, geneticist)
_____ Diagnostic laboratory studies
_____ Biopsy reports and slides
_____ Imaging studies (MRI, ultrasound, X-rays).
_____ One or more photos of your face and anything that is relevant to your medical condition
_____ Family history or drawn family tree
_____ Other:

Please also complete and send to us contact information for each of your healthcare providers.

You can mail records to: [address].

Allow at least 8 weeks from the time we receive all of these materials for us to review your records and consult with relevant specialists. All records will be reviewed by our team, but not all cases will be accepted into our research study.

When the team has made a decision about your case you will be notified by mail.

Thank you again for your interest in the Undiagnosed Diseases Network research study.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].

Undiagnosed Diseases Network

Your institutional logo here

Description: *Medical record request- pediatric*

Date:

Address:

Dear [patient's parent/guardian]:

Thank you for your interest in the Undiagnosed Diseases Network (UDN). Your child has been assigned to the [site name] in [location] for consideration. In order to properly review your child's case, we will need the following medical records:

_____ Admission and discharge summaries from hospitalizations related to your child's condition
_____ Consultant notes (such as cardiologist, gastroenterologist, neurologist, geneticist)
_____ Diagnostic laboratory studies
_____ Biopsy reports and slides
_____ Imaging studies (MRI, ultrasound, X-rays).
_____ One or more photos of your child's face and anything that is relevant to his/her medical condition
_____ Family history or drawn family tree
_____ Birth/Newborn records
_____ Growth curves (height, weight, and head circumference)
_____ Other:

Please also complete and send to us contact information for each of your child's healthcare providers.

You can mail records to: [address].

Allow at least 8 weeks from the time we receive all of these materials for us to review your child's records and consult with relevant specialists. All records will be reviewed by our team, but not all cases will be accepted into our research study.

When the team has made a decision about your child's case you will be notified by mail.

Thank you again for your interest in the Undiagnosed Diseases Network research study.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].

Undiagnosed Diseases Network

Your institutional logo here

Description: *Records received*

Date:

Address:

Dear [patient]:

Thank you for your interest in the Undiagnosed Diseases Network (UDN). We are writing to inform you that your records have been received. Please allow 60 days from this time for us to review your records. Our team reviews all records, however, not all cases are accepted into the UDN.

When the team has made a decision about your case you will be notified.

Thank you again for your interest in our program.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].

Undiagnosed Diseases Network

Your institutional logo here

Description: *Incomplete records*

Date:

Address:

Dear [patient]:

Thank you for your interest in the Undiagnosed Diseases Network (UDN). We are writing to inform you that some of your records have been received, however, they are not complete. When you send records: (1) please make sure there are not multiple copies of the same report, (2) organize the reports by subspecialty and date seen (i.e., genetics, gastroenterology, neurology etc.), and (3) Do NOT send double-sided copies. This will greatly speed processing and timely review of your case. Specifically, we are requesting:

_____ Summary letter from healthcare provider
_____ Medical records
_____ Labs
_____ Biopsy reports and slides
_____ All imaging on CD, including brain
_____ Anesthesiology records
_____ Seizure medication levels within 30 days of the admission
_____ Birth/Neonatal records **for pediatric patients**
_____ Growth curves **for pediatric patients**
_____ Photos **for pediatric patients**
_____ Other:

Please allow 60 days from the time of our receipt of these materials for us to review your records. Our team reviews all records, however, not all cases are accepted into the UDN. When the team has made a decision about your case you will be notified.

Thank you again for your interest in our program.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions or have difficulty requesting your medical records, please contact [Site Coordinator] at [phone number and email address].

Undiagnosed Diseases Network

Your institutional logo here

Description: *Partial application*

Date:

Address:

Dear [patient]:

You have expressed interest in the Undiagnosed Diseases Network (UDN) and have submitted a partial application. We requested medical records on [date], however, these records have not been received. Since three months has passed since our initial request, we assume that you are no longer interested in being evaluated in the network and will be removing your name from our active rolls.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions or have difficulty requesting your medical records, please contact [Site Coordinator] at [phone number and email address].
Cc: [referring provider]

Undiagnosed Diseases Network

Your institutional logo here

Description: *Acceptance Letter (Pediatric) to Healthcare provider*

Date:

Address:

Dear [healthcare provider]:

Your patient _____ [DOB] has applied to the Undiagnosed Diseases Network (UDN). As part of a stringent 2 step review process, your patient's case has completed the first level of review and is being considered for evaluation at [your institutional name; city, state]. Consent to the UDN as a research study is the next step for enrollment in the UDN. If enrolled, participants in the Network will be examined using the unique combination of scientific and medical expertise and resources at [your institution]. This evaluation will require a 2-5 day visit for inpatient and/or outpatient care. There is no need to order additional tests or procedures for the purpose of preparing your patient for this evaluation.

Travel, meals, and lodging expenses will be covered for research participants according to our policies, to the extent allowed by law. A representative of the Network will contact the participant within the next few weeks. If the parents of your patient are separated or divorced we will need to receive the court paperwork specifying who is legally able to consent for the child to participate in medical research. We may also request blood specimens for DNA isolation from one or both parents prior to scheduling the child's visit to the [your institutional name].

Not all admissions will result in a diagnosis, but the evaluations should yield valuable information that medical researchers will use to: (1) help identify previously unrecognized rare diseases; (2) suggest new ways to treat and prevent common illnesses; and (3) determine promising options for continued medical research.

[Your institution] will provide information from the evaluation to you and to your patient as a part of this Network. You will be responsible for your patient's follow up medical care. Selected patients may be eligible for other ongoing research studies.

Thank you for consulting with the Undiagnosed Diseases Network on your patient. We appreciate your commitment to providing the best possible care for your patients in ways that help advance medical knowledge and discovery.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].

Cc: [participant]

Undiagnosed Diseases Network

Your institutional logo here

Description: *Acceptance Letter (Adult) to Healthcare provider*

Date:

Address:

Dear [healthcare provider]:

Your patient _____ [DOB] has applied to the Undiagnosed Diseases Network (UDN). As part of a stringent 2 step review process, your patient's case has completed the first level of review and is being considered for evaluation at [your institutional name; city, state]. Consent to the UDN as a research study is the next step for enrollment in the UDN. If enrolled, participants in the program will be examined using the unique combination of medical and scientific expertise and resources at [your institution]. This evaluation will require a 2-5 day visit for inpatient and/or outpatient care. There is no need to order additional tests or procedures for the purpose of preparing your patient for this evaluation.

Travel, meals, and lodging expenses will be covered for research participants according to our policies, to the extent allowed by law. A representative of the Network will contact the participant within the next few weeks.

Not all admissions will result in a diagnosis. In addition to contributing to the diagnosis of individual participants, UDN evaluations should yield valuable information that medical researchers will use to: (1) help identify previously unrecognized rare diseases; (2) suggest new ways to treat and prevent common illnesses; and (3) determine promising options for continued medical research.

[Your institution] will provide information from the evaluation to you and to your patient as a part of this Network. You will be responsible for your patient's follow-up medical care. Selected patients may be eligible for other ongoing research studies.

Thank you for consulting with the Undiagnosed Diseases Network on your patient. We appreciate your commitment to providing the best possible care for your patients in ways that help advance medical knowledge and discovery.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].

Cc: [participant]

Undiagnosed Diseases Network

Your institutional logo here

Description: *Acceptance Letter to Participant (Optional)*

Date:

Address:

Dear [participant/caregiver]:

As part of a stringent 2-step review process, [your/your child's] case has completed the first level of review and is being considered for evaluation at [your institutional name; city, state]. Consent to the UDN as a research study is the next step for enrollment in the UDN. If enrolled, participants in the program will be examined using the unique combination of medical and scientific expertise and resources at [your institution]. This evaluation will require a 2-5 day visit for inpatient and/or outpatient care.

We would like to schedule the evaluation for [approximate time] from now. Our team will be in contact with you regarding arranging travel to [Institution] and lodging. The research study will cover the cost of travel, meals, and lodging for [you/your child] and one caregiver. Any additional travelers will be at your expense. The research study will not cover additional expenses, like pay-per-view movies, room service, and mini-bar items.

Once scheduled, you may not reschedule the evaluation, except in the case of an emergency. The evaluation will be designed by the [Institution] research team and will be unique to [you/your child] & requires effort and care to plan. Visits may only be rescheduled in cases of adverse weather conditions, acute illness that would interfere with the planned evaluation, or otherwise at the discretion of the study investigators.

It is our hope that the care and evaluation [you/your child] will receive at [institution] will yield long awaited answers. Although [you/your child] may not initially receive a diagnosis, research into [your/your child's] condition will continue even after you return home.

Additionally, UDN evaluations should yield valuable information that medical researchers will use to: (1) help identify previously unrecognized rare diseases; (2) suggest new ways to treat and prevent common illnesses; and (3) determine promising options for continued medical research.

We look forward to meeting you [and your child] soon!

Sincerely,
The UDN team at [Institution]

If you have any questions, please contact [Travel Coordinator] at [phone number and email address].

Undiagnosed Diseases Network

Your institutional logo here

Description: *Not Accept Letter to Healthcare provider*

Note: required sections are in ***bold*** *below.*

Date:

Address:

Dear [healthcare provider]:

Your patient [name] [DOB] has applied to the Undiagnosed Diseases Network (UDN). After a stringent review process, your patient was not accepted for evaluation by the UDN. The Network's goals are to provide answers to patients with mysterious conditions that have long eluded diagnosis and to advance medical knowledge about rare and common diseases. The medical team bases its decisions on whether or not there is a reasonable chance to achieve these goals. Review of each application considers the Network's resources as a whole.

[Insert one of these or another scenario below depending on whether you have additional suggestions for work up and whether or not you would be willing to reconsider the patient if the suggested work up comes back normal.]

[If no recommendations given] Upon extensively reviewing [name of patient]'s records, we do not believe we can improve on the comprehensive work-up [he/she] already received.

[If recommendations given] Members of the Case Review Committee had a few thoughts for your consideration.

> *[Mitochondrial/Metabolic Recommendations]* Pursuit of a mitochondrial and metabolic evaluation, including plasma and urinary amino acids, urinary organic acids, plasma lactate, pyruvate, carnitine (free and total), leukocyte CoQ, and acylcarnitine profile. For a complete metabolic and mitochondrial disease work up, you may consider contacting the Medical Genetics Laboratories at the Baylor College of Medicine: www.bcm.edu/geneticlabs/.
>
> *[Pitt-Hopkins syndrome]* Testing for Pitt-Hopkins syndrome by looking at mutations or even a deletion of the *TCF4* gene. For testing center sites, visit GeneTests at: http://www.ncbi.nlm.nih.gov/sites/GeneTests/.
>
> *[Congenital Disorders of Glycosylation]* Congenital disorders of glycosylation (CDG) testing at the Mayo Medical Laboratories. Results are reported as the mono-oligosaccharide/di-oligosaccharide transferrin ratio, the a-oligosaccharide/di-oligosaccharide transferrin ratio, the tri-sialo/di-oligosaccharide transferrin ratio, the apolipoprotein CIII-1/apolipoprotein CIII-2 ratio and the apolipoprotein CIII-0/apolipoprotein CIII-2 ratio. For more information on how to send this test, visit: http://www.mayomedicallaboratories.com/test-catalog/Overview/89891

[PGM1 testing] *PGM1* full gene sequencing, which is offered at multiple laboratories that can be identified at: http://www.ncbi.nlm.nih.gov/sites/GeneTests/?db=GeneTests.

[Lysosomal enzyme screening] Lysosomal enzyme screening. Information about this screening can also be found at the Emory Genetics Laboratory. Their website is: http://genetics.emory.edu/egl/.

[Required if recommendations given] **These comments are for consideration by [name of patient]'s medical team. The UDN does not coordinate medical care outside of the UDN evaluation, and all decisions about further workup are in the hands of the patient's local caregivers and their consultants.**

[Required if patient would be reconsidered] **If a diagnosis remains elusive after further work up, we would be willing to reconsider [name of patient] for admission to the UDN. Once further work up is complete, [name of patient] can send an updated letter with the results of the work up to [contact information]. At that time, we will review the letter and previously submitted application.**

Again, thank you for consulting with the UDN on your patient. We appreciate your commitment to providing the best possible care for your patients in ways that help advance medical knowledge and discovery.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].

Cc: [participant]

Undiagnosed Diseases Network

Your institutional logo here

Description: *Welcome packet to patient*

Welcome to the Undiagnosed Diseases Network

We want your visit to the UDN clinical site at [institution] to be as comfortable and productive as possible. We have created a package of information for you to review prior to your admission so you will know what to expect. The following are included in the package:

1. **[institution] map and interior layout map of [institution]**
2. **Cafeterias [information]**
3. **Shuttle Service [information]**
4. **Airport transportation [information]**
5. **Security Procedures [information]**
6. **Patient Library [information]**
7. **Chapel [information]**
8. **Hospitality Services [information]**
9. **Gift Shops [information]**

For further information, please visit the following websites:
http://www

Please be sure to bring the following with you:

- **List of medications and dosages in their original containers**- Please discuss any questions you may have regarding medications and/or equipment with your admitting Nurse Practitioner or Physician Assistant.
- **Any assistive devices that you use daily**, e.g., wheelchair, cane, walker, braces
- **Complete list of current physicians and their contact information**, i.e., addresses and phone numbers

A UDN staff member will contact your shortly to make travel arrangements and book local hotel accommodations. An additional member of the UDN team will contact you to review the admission process and ask additional questions about your medical history.

In the event you are unable to keep your scheduled visit, please contact your admitting Clinical Site Coordinator. You may be asked to re-schedule your visit.

Please feel free to contact our team if you have any questions or concerns. Generally, phone messages and e-mails are responded to promptly. We are looking forward to meeting you at [your institution].

[team contact information- names, emails, phone numbers]

Undiagnosed Diseases Network

Your institutional logo here

Description: *Directions for remote blood draw*

DIRECTIONS: Remote Blood Draw

Date:
Address:
Dear [patient],
In order for your blood to be processed appropriately, please follow the directions below carefully. If possible, please arrange the blood draw through your healthcare provider.

Directions:

1) Draw [number] tubes of blood in the lavender topped tubes enclosed.
2) While talking to a UDN team member, review, sign, and date the consent form(s) enclosed and send the signed copy of the consent form(s) with your blood sample. **If we do not receive your signed consent form(s), we cannot process your blood.**
3) Ship the blood overnight (Monday-Thursday) to the UDN clinical site at [institution].
4) Call [number] or email [email] on the day that you ship the blood and provide the FedEx Tracking number.

Please contact us with any questions or concerns.

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].
Cc: [referring provider]

Undiagnosed Diseases Network

Your institutional logo here

Description: *No Diagnosis Following Evaluation*

Date:

Address:

Dear [patient]:

Thank you for your participation in the Undiagnosed Diseases Network (UDN) evaluation at [institution]. At this time, the clinical testing and evaluation phase is complete and has not yielded a definitive diagnosis. We will continue to pursue leads as they arise based upon ongoing research and new ideas that are generated among our expert consultants. We will keep the valuable information and biological samples collected during your visit in the hope that future research studies will be able to shed light on the medical problems that brought you to the UDN. If new prospects for investigation appear, we will contact you.

We very much appreciate your involvement in the UDN and your commitment to our joint goal of helping to advance medical knowledge, scientific discovery, and optimal care.

Sincerely,

[name]
Principal Investigator at [Institution]

If you have any questions, please contact [Site Coordinator] at [phone number and email address].
Cc: [referring provider]

APPENDIX 5: Wrap-up Template

Name:
DOB:
Dates of Visit:
Primary Clinician:
Attending Physician:
Presenting Symptoms/short summary of case:
Testing/Recommendations by System (as applicable):

GASTROENTEROLOGY

Consultant: **Testing/Results:** **Recommendations:**

NUTRITION

Consultant: **Testing/Results:** **Recommendations:**

NEUROLOGY

Consultant: **Testing/Results:** **Recommendations:**

PULMONOLOGY

Consultant: **Testing/Results:** **Recommendations:**

CARDIOLOGY

Consultant: **Testing/Results (EKG/ECHO):** **Recommendations:**

IMMUNOLOGY/INFECTIOUS DISEASE

Consultant: **Testing/Results:** **Recommendations:**

HEMATOLOGY

Consultant: **Testing/Results:** **Recommendations:**

ENDOCRINOLOGY

Consultant: **Testing/Results:** **Recommendations:**

DENTAL

Consultant:
Testing/Results:
Recommendations:

DERMATOLOGY

Consultant:
Testing/Results:
Recommendations:

OPHTHALMOLOGY

Consultant:
Testing/Results:
Recommendations:

AUDIOLOGY

Consultant:
Testing/Results:
Recommendations:

PHYSICAL MEDICINE & REHABILITATION

Consultant:
PT:
OT:
Speech:
Recommendations:

NEPHROLOGY

Consultant:
Testing/Results:
Recommendations:

RADIOLOGY

Test:
Results:
Test:
Results:

SPECIAL STUDIES

EMG/NCV:
Sleep Study:
Metabolic Cart:
Other:

OTHER

Consultant:
Testing/Results:
Recommendations:

SEND OUT AND PENDING TEST RESULTS

Test Name	Date Sent	Date Resulted Rev'd	Testing Lab	Result/Interp	Normal Range	Copies sent (date and person)
CSF studies						
CDT/N-glycan screen						
Urine Oligosaccharides						
WBC CoQ level						
Lysosomal Screen						
WBC Buffy Coat						
Urine Sulfocysteine						
MitoGEN						
POLG						
Muscle mtDNA Content						
Muscle ETC Enzymology & CoQ content						

PLEASE NOTE:
You will be contacted by the UDN for surveys to tell us about your experience. The first contact will be a week or so after you get home.

Participant Webpages Project:
UDN participants have the option of having a page created on the UDN website to try to find and connect with others with the same or similar condition.
The participant pages include the following information: genetic changes, symptoms, medical history, treatments, procedures, medications, gender, race, ethnicity, and photographs/videos (optional).
Example pages can be found here: https://undiagnosed.hms.harvard.edu/updates/participant-pages/

Please indicate if this participant would be interesting in participating in this project:
☐ Yes
☐ No

APPENDIX 6: Participant Follow-up Surveys

7-14 days post visit survey: Interpersonal Processes of Care Survey: Short Form IPC-18 – UDN CC revised.

These questions are about your experience at _____*[CS]* during your evaluation. When we ask "how often…?" we are referring to the many specialists you saw during your UDN visit. Please think about your entire UDN visit as you answer these questions.

	Never	Rarely	Sometimes	Usually	Always	N/A
1. How often did health care providers speak too fast or use words that were hard to understand?						
2. How often did health care providers listen carefully to what you had to say?						
3. How often did health care providers take your health concerns seriously?						
4. How often did health care providers explain your test results such as blood tests or X-rays?						
5. At the end of the visit, did the health care providers summarize the visit and results and answer your questions?						
6. How often were health care providers concerned about your feelings?						
7. How often did health care providers really respect you as a person?						
8. How often did health care providers treat you as an equal?						
9. How often did health care providers pay less attention to you because of your race or ethnicity?						

Open-ended questions:

10. How was your experience with the logistics of their UDN visit, including travel to and from the medical center, quality of their overnight accommodations, organization of the clinic visits/procedures?
11. What part/s of the UDN experience worked particularly well for you?
12. What part/s of the UDN experience did not work well?
13. What could we do to make the UDN experience better?
14. Did you feel the UDN visit was burdensome? If so, please explain.
15. Were changes to your medications suggested at the end of the UDN visit? If so, what were they and do you intend to follow the recommendations?

16. Were recommendations made for next steps (e.g., additional testing, referrals, etc.)? If so, what were they and do you intend to follow the recommendations?

6 months post-visit survey: Participant-specific clinical and research outcomes.

Please note:

1. *There will be "yes", "no", "not applicable" and "refuse to answer" checkboxes for each question, in addition to text boxes.*
2. *"You" should be replaced by "your child", "your spouse", or "your family member" depending on who is answering the survey for the participant, i.e., the participant's parent, spouse, or other family member, respectively.*

Topic:	**Question:**	*If yes*:
Diagnosis	Did you receive a diagnosis at, or following, the UDN visit?	Please explain when you were diagnosed, and how this information was provided to you
Follow-up	Has follow-up with a health care provider occurred since the UDN visit?	Please describe the follow-up, including if the medical provider is at or in communication with the UDN site.
	Have you received results from tests that were pending at the time of the evaluation	Please describe who communicated to you and how those results were communicated.
Treatment	Was a new treatment started at the UDN visit?	Please describe the new treatment.
	Was a previous treatment discontinued at the UDN visit?	Please describe the discontinued treatment.
Research	Did the health care providers at the UDN site discuss research options at, or following, the visit?	Please describe the types of research discussed.
	Are you currently participating in research initiated by the UDN?	Please describe the research.
	Are you currently participating in research initiated outside of the UDN?	Please describe the research.
Support	Did the health care providers at the UDN site discuss support services (e.g., support groups, local and national resources, counseling) at, or following, the UDN visit?	Please describe the support services discussed.
	Are you participating in a support group (either in person or online)?	Please describe the support group.
Continuing contact	Have you been in contact with a member of the UDN since the time of the visit?	Please describe your communication with the UDN since the time of the visit.
	Do you have contact information for a team member, in case you would like to contact them later?	Please describe what type of contact information you have.
Additional comments	Is there anything that you would change about your UDN experience?	Please explain what you would change about your experience.

	Have you noticed any changes to yourself or your family members as a result of your UDN experience?	Please explain these changes.
	Are you having any issues with insurance billing for services from the UDN visit, such as uncovered expenses?	Please explain.

Yearly post-visit survey.

Please note that there will be "yes", "no", "not applicable" and "refuse to answer" checkboxes for each question, in addition to text boxes.

Topic:	**Question:**	*If yes*:
Diagnosis	If you did not have a diagnosis at the time of the last survey, have you received a diagnosis since the last survey?	Please explain how this information was provided to you.
Follow-up	Has follow-up with a health care provider occurred since the UDN visit?	Please describe the follow-up, including if the medical provider is at or in communication with the UDN site.
	Have you received the results from tests that were pending at the time of the last survey, including results from the genomics sequencing?	Please describe who communicated to you and how those results were communicated.
Treatment	Has a new treatment been started since the UDN visit?	Please describe the new treatment. What have been the effects of this treatment over the past year?
	Has a previous treatment been discontinued since the UDN visit?	Please describe the discontinued treatment.
Research	Have you discussed research options with the health care providers at the UDN site since the last survey?	Please describe the types of research discussed.
	Are you currently participating in research initiated by the UDN?	Please describe the research.
	Are you currently participating in research initiated outside of the UDN?	Please describe the research.
Support	Are you participating in a support group (either in person or online)?	Please describe the support group.
Continuing contact	Have you been in contact with a member of the UDN since the time of the visit?	Please describe your communication with the UDN since the time of the visit.
	Do you have contact information for a team member, in case you would like to contact them later?	Please describe what type of contact information you have.
Additional comments	Is there anything that you would change about your UDN experience?	Please explain what you would change about your experience.

	Have you noticed any changes to yourself or your family members since the last survey as a result of your UDN experience?	Please explain these changes.
	Are you having any issues with insurance billing for services from the UDN visit, such as uncovered expenses?	Please explain.

APPENDIX 7: Feature Request Form

Feature Requester Name(s):

Feature Requester Contact Information:
Name:
Institution:
Email:

Detailed Description of Feature:
Please provide a description of the feature and types of users that will interact with the feature and how they will access and use the feature. Describe the workflow from the perspective of each of these users. Screen shots of the Gateway where the new feature will exist are also preferable.

Importance to the UDN and Justification:
□ Critical
□ Major
□ Minor

Justification:

APPENDIX 8: Research Concept Sheet

Thank you for your interest in collaborating with the UDN. Please complete the following application.
(To complete online, go to https://hms.az1.qualtrics.com/jfe/form/SV_3CyZOKOuiEvnCx7)

Type of submission:
☐ Research Project*
☐ Data Analysis*
☐ Manuscript
☐ Grant Application

*Distinction between research project and data analysis- research projects will involve the collection of new data while data analysis will involve new analyses of existing data.

Is this research project network wide?
☐ Yes
☐ No

Primary investigator contact information:

Name:
Institution:
Phone:
Email:

Are you a UDN Member?

☐ Yes
☐ No; Please list the name of your UDN contact:

Additional/Senior investigator(s) (specify institution):

Project title:

If the submission is a manuscript, please provide the following information:

Intended date (mm/dd/yyyy) and place of manuscript submission or meeting presentation:

Overall manuscript outline:

If the submission is a research project, data analysis, or grant application, please provide the following information:

Specific aims:
Funding plan:
Subject eligibility criteria:
Access to data and burden on subjects:

Consent required?

☐ Yes; Please describe:
☐ No

Methods:

Which (if any) UDN Working Groups are (or will be) involved? Include primary Working Group first.

Data management and statistical support:

☐ Required
☐ Not required

If required, please provide contact information for your local statistician:

Name:
Phone:
Email:

Biospecimens required:

Other needs:

Timeline:

Does this research project involve administering a survey to participants?

☐ Yes; Please attach a copy of the instrument to this application.
☐ No

If yes, is this research project under the UDN NHGRI Central IRB (falls under the UDN consent form) or will there be a separate IRB protocol and consent form through a local IRB?

☐ Will be under UDN Central IRB
☐ Will be under a separate local IRB

If the survey will be under the UDN Central IRB, the UDN Survey Committee will use the information requested below to make a recommendation as to whether or not this survey should be approved by the UDN. The goal is to ensure that approved surveys are consistent with the purpose of the UDN and that patients are not overburdened.

A. Survey Details

1. Survey Instrument
2. Construct measured
3. Description of validation of survey instrument
4. Rationale for using this particular survey instrument (as opposed to other instruments that measure the construct)
5. Translation of measure into other languages

B. Survey Population

1. UDN sites included in the study
2. Participant inclusion/exclusion criteria
3. Sample size

C. Study Logistics

1. Amount of time to administer
2. Method of administration (online [Qualtrics, other?], phone, mail, etc.)
3. Individuals administering survey (study staff, self, etc.)
4. # times survey is administered and when
5. Timeline for study: When will survey commence, for how long will you continue the survey, when will it end, when will data be analyzed

D. Please provide a plan for follow up if a problem or concern is identified on responses to the measure.

If the survey will be under a separate local IRB, the UDN Survey Committee will not be involved in determining whether or not this survey should be approved. However, the committee is requesting the following information in order obtain a complete picture of the surveys being administered to UDN patients at various sites.

A. Survey Details

1. Survey Instrument
2. Construct measured
3. Rationale for using this particular survey instrument (as opposed to a different instrument that measures the same construct)

B. Survey Population

1. UDN sites included in the study
2. Participant inclusion/exclusion criteria
3. Sample size

C. Study Logistics

1. Amount of time to administer
2. Method of administration (online, phone, mail, etc.)
3. # times survey is administered and when
4. Timeline for study: When will survey commence, for how long will you continue the survey, when will it end, when will data be analyzed

D. Please provide a plan for follow up if a problem or concern is identified on responses to the measure.

E. Data Sharing – who will have access to the data?

APPENDIX 9: Annual Survey Review Worksheet

Directions: Please provide the following information for the Survey Committee annual review. Please email the completed form to the chair of the Survey Committee: Ingrid A. Holm, MD, MPH- Ingrid.holm@childrens.harvard.edu

Due Date:

Date:
Date of initial IRB approval:
Name of study:
Name of survey instrument/s:
Names of UDN members/group responsible for conducting survey research:
Main contact person:
Name:
Site:
Email address:

Summary of survey progress (*response should be no more than 2 pages*)

1. Survey implementation summary (include any changes to the survey or its implementation):
2. Number of UDN participants who qualified for the survey:
3. Number of UDN participants who participated/completed the survey:
4. Administration time:
5. The nature and amount of missing data:
6. Preliminary results (if interim analyses are explicitly allowed as part of the study design):

7. Impact on clinical care, if appropriate:
8. Describe any findings that impact the effectiveness of the survey or its implementation:

9. Any new knowledge gained based on the results to date:

APPENDIX 10: PEER Charter and Application

Participant Engagement and Empowerment Resource (PEER) Charter

PEER Mission Statement: The purpose of the UDN PEER is to support participants and family members in part by creating and sharing resources, and to provide the participant and family perspective on UDN research goals and participant experience. The PEER provides a "post-UDN visit voice". Their goal is to advocate for participants, improve participant experience, "get the word out" about the UDN, and facilitate interactions between participants and the UDN.

Responsibilities of PEER Members: PEER members are expected to participate actively in the PEER and to respect each other's views and opinions. Each team member is responsible for bringing and sharing experiences, and is expected to volunteer for tasks, be open, and be willing to talk.

PEER Membership: UDN Participant/Parent/Guardian

New member recruitment:

- Interest in joining the PEER is assessed on an ongoing basis by the Coordinating Center (CC) on the Transitions of Care calls.
- The CC sends interested members an application.
- By November 1 of each year, the PEER will review their membership status for openings, turnover of participants, need for new members, etc.
- After November 1 when the PEER has decided on how many new members are needed, applications undergo an initial review by the CC and the site coordinator at the UDN participant's site.
- Applications passing the initial review are reviewed by the PEER, which votes on new member(s) and 1-2 back up choices if the first choice(s) decline. Each PEER member has one vote. The PEER will consider diversity (see below) and applications will be ranked in preference by the PEER based on the following attributes:
    - Initial application review feedback by CC and site coordinator
    - Application response content and length
    - Diversity of applicant in relation to current PEER members
- The CC contacts the invited members and back-up invited members if applicable.
- Once the new members are confirmed, the CC will contact the rest of the applicants to let them know they were not chosen but will be kept in the pool for the next year if desired.

PEER term: 1 year, may be renewed annually

Number of PEER participant/parent/guardian members: up to 11

PEER members who are inactive: Members who miss 2 consecutive meetings and/or we are not able to contact will be removed from the PEER.

Termination of membership: CC reserves the right to terminate membership of any member who does not act in accordance with the charter.

Diversity of members: The PEER prioritizes diversity in representation and will consider diversity in the acceptance process based on the following criteria:

1. Major criteria
    a. UDN clinical sites (location of clinical evaluation)
    b. Member type (Participant vs parent/guardian)
2. Minor criteria
    a. Geographic location
    b. Gender, race, ethnicity, urban/rural

PEER Membership: *Ex-officio* Members

Ex-officio members will include:

- 1 CC representative
- 2 site representatives from Site Operations Working Group and/or Genetic Counseling and Testing Working Group

Ex-officio members will be selected by the UDN Coordinating Center and UDN leadership and be rotated on a yearly basis.

Ex-officio member roles:

- *Ex-officio* members will participate on all calls and meetings.
- *Ex-officio* members will have one vote each on all votes except election of PEER leaders.

PEER Leadership:

- The PEER will nominate, agree, and vote on 1-2 co-chairs every year.
    - Co-chairs will be voted on during the annual in-person meeting.
    - Co-chair term: 1 year, January 1 – December 31; may be renewed annually.
    - Tasks of the co-chair(s): Engage the members, lead discussions, bring forward ideas, represent the PEER as needed.
- The PEER will nominate, agree, and vote on 1 secretary every six months.
    - Secretary will be voted on during conference calls.
    - Secretary term: 6 months; may be renewed every 6 months.
    - Tasks of the secretary: draft meeting agendas, take meeting minutes, circulate approved meeting minutes.
- The role of the UDN CC is to facilitate and provide the infrastructure to support the leader(s) and the PEER, and to serve as a liaison to the UDN Steering Committee

The Relationship Between the PEER and the UDN Steering Committee:

- The PEER will interface between the UDN Steering Committee and the UDN participants. The PEER will develop an annual report on activities to be shared with the Steering Committee. The UDN Steering Committee will send feedback on the annual report to the CC representative *ex-officio* member to be shared with the PEER. The

Coordinating Center will keep the UDN Steering Committee apprised of the PEER activities.
- Any concerns, comments, or questions between annual reports will be reported via the same mechanism.
- The PEER will respond to feedback from the UDN Coordinating Center within one month.
- Minutes from PEER meetings will be shared with the UDN Coordinating Center.

Conference Calls and Meeting:

1. Conference calls every 3 months (4 times a year) and ad hoc as needed.
2. Annual in-person meeting – cost is covered for up to 11 PEER members.
3. The agenda for meetings will be developed by the PEER with input from the UDN CC.
4. The PEER will set priorities with input from the UDN CC.
5. All meetings' minutes will be kept confidential within the UDN.
6. The PEER will create groups to work on specific projects as needed.

Application Form: UDN Participant Engagement and Empowerment Resource (PEER)

Overview of the UDN Participant Engagement and Empowerment Resource (PEER)

Purpose: The purpose of the UDN Participant Engagement and Empowerment Resource (PEER) is to provide the participant and family perspective on UDN research goals and participant experience. The PEER will engage with UDN investigators in the development and assessment of participant-oriented materials and identify best practices for receiving participant input in research.

Activities: The PEER will be responsible for its structure and activities. Activities may include, but are not limited to:
- Providing input regarding various research questions, eligibility criteria, and recruitment and informed consent processes
- Identifying unmet participant needs
- Contributing perspectives on risk/benefits of research project
- Connecting families with one another and to support groups
- Collecting participant and family experiences with the UDN from participants
- Providing support for families when they are visiting a site far from home
- Being a resource for families who have questions or concerns
- Developing educational materials
- Organizing participant conferences
- Leading awareness efforts

Membership: The PEER includes participants and family members. Members of the PEER must be willing to engage in thoughtful conversation about the positive and negative aspects of the research process and respect the perspectives of others. There will be adult, adolescent, and pediatric participant and family member representation.

Terms: Terms are one year in duration and are renewable on a yearly basis.

Meetings: Conference calls are expected to be held quarterly. In-person meetings are expected to be held annually. Travel expenses will be covered for the in-person meeting and the Steering Committee meeting as needed.

Acceptance into the PEER: Factors taken into consideration in the acceptance process include the answers on the application, diversity of sites represented on the PEER, and diversity of UDN participant age (PEER members who are an adult UDN participant or the parent of a child participant).

The UDN PEER is an equal opportunity organization and does not discriminate on based on any criteria including gender, race, ethnicity, or age.

Application Form:

Thank you for your interest in joining the UDN Participant Engagement and Empowerment Resource (PEER). If you have questions about this application, please contact the UDN Coordinating Center at 1-844-746-4836 or UDNPEER@hms.harvard.edu. Completed applications can be submitted to UDNPEER@hms.harvard.edu.

Application Deadline: November 1

Contact Information:
Last Name: First Name: UDN ID:
Street Address:
City: State: Zip Code:
Phone Number:
E-mail Address:

I am a:
☐ UDN participant
☐ Family member of UDN participant
☐ Other, please specify:

Areas of Interest:

Please describe in 50-100 words why you want to join the PEER:

Please indicate areas that you would be interested in (check all that apply):

☐ Providing input regarding various research questions, eligibility criteria, and recruitment and informed consent processes
☐ Identifying unmet participant needs;
☐ Contributing perspectives on risk/benefits of research project;
☐ Connecting families with one another and to support groups;
☐ Collecting participant and family experiences with the UDN from participants;
☐ Providing support for families when they are visiting a site far from home;

☐ Being a resource for families who have questions or concerns;
☐ Developing educational materials;
☐ Organizing participant conferences;
☐ Leading awareness efforts.

Summary of UDN Experience:

UDN Clinical Site:
Name of Primary UDN Clinical Site Contact:
Month of UDN Evaluation (MM/YYYY):
Please describe your overall experience with the UDN, including what went well and what can be improved:

Based on your experience with the UDN, please describe in 50-100 words what **went well**:

Based on your experience with the UDN, please describe in 50-100 words what you feel **could be improved**:

Is there anything that you feel is important for us to know about you? (limit to 50-100 words):

Availability:

Please list times when you are able to attend meetings:

☐Daytime:
☐Evening:
☐Weekend:

APPENDIX 11: Standard Operating Procedures for Not Accepted Applicants

This Standard Operating Procedure (SOP) document was developed in November 2016 as a guideline for UDN investigators interacting with applicants not accepted or enrolled into the UDN study. Please note that this SOP is not intended to override institutional policy, but rather provide additional guidance for UDN investigators.

Stages of enrollment

Until an applicant has signed a consent form and enrolled in the UDN study, the term "accepted" should not be utilized in any communications with the applicant. Removing the use of the term "accepted" will help with confusion amongst applicants if any unforeseen reason prevents an individual from enrolling in the study between the time of application review by the CS and consent. This will lead to stages of enrollment before an applicant becomes a participant in the UDN study.

- Application Stage – from time of first inquiry or application start in the UDN Gateway through submission of an application to the UDN
- Review Stage – from submission of an application through medical records review and determination by the CS that an applicant should be "accepted" and referred to the consent stage
- Consent Stage – from decision of the CS to "accept" through consent of the applicant to participate in the UDN including second level review by the Case Review Committee and any other continuing review by the CS during the consent process
- Enrolled – participants in the UDN that are fully-consented to the study

Applicants that have completed the review stage may be informed that their application has completed the first level of review. Applicants should also be informed that they are not enrolled in the UDN study until consent is completed. These stages of enrollment will also be outlined on the UDN webpage FAQs.

Not accepted applicants

Applications that are not accepted during the review stage require the following information to be entered in the UDN Gateway:

- Applicant Review Form, Case Decision field – provides the reason for application to be not accepted
    - Not Accepted with Recommendations - Specific Testing
    - Not Accepted with Recommendations - Seek Expert Care
    - Not Accepted - Diagnosis Identified
    - Not Accepted - UDN would likely not be able to find a diagnosis
    - Not Accepted - Insufficient records made available to UDN site
- Not Accepted Letter – should be uploaded into the Gateway when sent to the applicant and referring provider

CC helpdesk communications

The applicant helpdesk email and phone number is staffed by personnel at the CC. When the helpdesk receives an applicant inquiry that rises above the level of general questions or

concerns, the CS involved is notified. Such inquiries will also be brought to the attention of the Executive Committee and CIRB. If an applicant expresses intent to contact Congress or any portion of the NIH directly, the Program Director will also be notified. A general log of applicant inquiries is kept at the CC.

Media inquiries

The CC will not comment on any applications to the media without the permission of the CS to which the participant has been assigned. All media inquiries related to specific applicants and participants will be referred to the CSs and logged by the CC. If a site receives a media inquiry directly, contact Kimberly LeBlanc who will log the inquiry and notify the Site Operations group. An example of language developed for media inquiries related to not accepted applicants is provided below:

"The Undiagnosed Diseases Network is not able to see all who apply. Many of our applicants have come to us as part of a long diagnostic odyssey in search of answers. With that in mind, we have to make difficult, often agonizing, decisions about whom to accept."

Congressional inquiries

If a congressional communication related to an applicant or participant is received, the UDN Program Director should be notified as soon as possible so that a coordinated response can be initiated. Most congressional offices will reach out to the NIH directly as the federal agency involved; however, if a request for information is received by a CS or the CC, general guidelines on replies are detailed below.

Timeline for response: Technically, there is no timeline to respond to a letter from a personal office, but a response should be made as soon as possible. For an email inquiry, a response email should be sent immediately acknowledging receipt and noting that the requested information will be provided as soon as possible.

Example email language:

Dear XXX,

Thank you for your email. We are writing to acknowledge that we have received your letter and are working to pull together all of the information needed in order to address your questions. We will be in touch again as soon as we are able to provide an informed response.

Sincerely,
XXX

If the request was received via email, an identical postal mail can be sent if desired. If the request was received over the phone, a call to acknowledge receipt with similar information to the draft email above is appropriate.

Individuals to be contacted: The CS Principal Investigator, CC, and UDN Program Director should be contacted regarding any communications with a congressional office.

Official response: The NIH will notify the CS or CC when an official response can be distributed. The language below should only be sent with approval from the NIH, but is included as a reference.

Example reply language:

Dear XXX (the individual who contacted you from each congressional office),

Thank you for reaching out to us regarding XXX. The Undiagnosed Diseases Network (UDN) is a research study funded by the National Institutes of Health (NIH). It aims to improve the diagnosis and care of patients with undiagnosed diseases. Its goals are two-fold—to come up with a diagnosis for each participant and to learn more about rare and previously undiagnosed diseases. While one goal of the UDN is to successfully diagnose each participant in the study, it is important to note the current programmatic rate of diagnosis is under one-third of patients. Furthermore, the UDN does not provide treatment for participants; this responsibility rests with the participant and his or her healthcare provider.

As a research study with a defined set of resources, the UDN is not able to accommodate all applicants who want to be part of the study. Researchers at the XXX UDN site determined XXX was [insert reason for not acceptance here]. It is important to note that decisions made by one site in the UDN are binding for the entire network; and the UDN does not accept appeals related to application or enrollment decisions.

Please feel free to contact us if you have any further questions regarding this matter.

Sincerely,
XXX

Identification of a UDN participant

As general practice, no identifiable information should be shared with individuals not on the IRB protocol unless they have been directly contacted by an applicant or participant. If identifiable information needs to be shared with individuals not on the IRB protocol, sites should contact Kimberly LeBlanc at the CC.

Appeal process

The UDN will not accept appeals related to application or enrollment decisions. Language that there is no appeal process and that the UDN decision to not accept is binding should be communicated to applicants during the application process and in the FAQs online. Questionable cases should be sent to the CRC for review and will be prioritized for discussion.

APPENDIX 12: Diagnosis Coding Tool v1.0

Patient UDN ID
Diagnosis Name
Date Diagnosis Entered
Diagnosis Entered By
Phenotype MIM Number
Gene Name
Gene MIM Number
ICD Number
Additional Diagnosis Notes

Category 1: Certainty

The overall certainty of the diagnosis.

- ☐ Certain
- ☐ Highly Likely
- ☐ Tentative
- ☐ Low (still on differential diagnosis, but lacking clear evidence)

More information:

Category 2: Chief Method Used to Achieve Diagnosis

The single most important piece of evidence supporting the diagnosis. □

- ☐ Genome-scale sequencing including genome sequencing and exome sequencing (select all that apply)
  - ☐ Reanalysis of prior clinical exome data
  - ☐ Reanalysis of prior research exome data
  - ☐ Reanalysis of prior clinical genome data
  - ☐ Reanalysis of prior research genome data
  - ☐ Clinical site analysis of Baylor exome data
  - ☐ Clinical site analysis of HudsonAlpha genome data
  - ☐ Baylor – pathogenic
  - ☐ Baylor – likely pathogenic
  - ☐ Baylor – VUS
  - ☐ Baylor – research finding
  - ☐ HudsonAlpha – pathogenic
  - ☐ HudsonAlpha – likely pathogenic
  - ☐ HudsonAlpha – VUS
  - ☐ HudsonAlpha – research finding
  - ☐ Known prior to enrollment in UDN
  - ☐ Other
- ☐ Non-sequencing, genome-wide diagnostic assay (for example, SNP array, oligo array or karyotype)

- ☐ Microarray
- ☐ Karyotype
- ☐ Known prior to enrollment in UDN
- ☐ Other

☐ Directed clinical testing based on phenotype (may be molecular, including gene panels; biochemical; radiologic; or, other test type)

- ☐ Panel testing
- ☐ Single gene testing
- ☐ Del/dup testing
- ☐ Known prior to enrollment in UDN
- ☐ Other

☐ Diagnosis made primarily on clinical grounds, including aggregate assessment of non-specific test results

☐ Other – please explain below

More information:

Category 3: Mechanistic Characterization

The degree to which the underlying mechanisms of disease pathology and causation are established.

☐ Biologically well-characterized disorder, e.g., high-specificity testing available; or, the gene, pathway, and/or cell biology is understood.

☐ Descriptive (no definitive testing available, specific mechanism unknown).

☐ Other – please explain below

More information:

Category 4: Timing and Source of Diagnosis

The source and timing of the diagnosis relative to evaluation by the UDN. □

☐ Diagnosis made before UDN application

☐ Diagnosis made by non-UDN entity after application but before UDN evaluation

☐ Diagnosis made by UDN after application but before evaluation

☐ Diagnosis made during UDN in-person evaluation

☐ Diagnosis made after UDN in-person evaluation by UDN

☐ Diagnosis made after UDN in-person evaluation by non-UDN entity

☐ Other, including misdiagnosis – please explain below

More information:

Category 5: Degree to Which This Diagnosis Explains Major Components of Phenotype

☐ Explains all major features of illness, including asserted phenotype expansion

☐ Explains some major features of illness, including asserted phenotype expansion

☐ Diagnosis does not explain any major features of illness

☐ Other – please explain below

More information:

Category 6: Consequence of Diagnosis

Was the diagnosis important in the future management of the study participant or family's overall health?

- ☐ The diagnosis had a major role in the subsequent management of the participant's and/or family's overall health
- ☐ The diagnosis had some role in the subsequent management of the participant's and/or family's overall health
- ☐ The diagnosis had no significant role in the subsequent management of the participant's and/or family's overall health
- ☐ Other – please explain below

More information:

Is there a readily available treatment?

- ☐ Yes
- ☐ No

If yes, please provide a brief description of the treatment:

More information:

Narrative: