Maximizing the Research Value of Clinical Data

Steven Fiske

Director of Data Engineering

USF Health Informatics Institute

The Health Informatics Institute (HII)

- Our mission:

  - to establish and maintain expertise in biomedical science, statistics, clinical trial study design and coordination, software and data engineering, big data and high performance computing, and integrative bioinformatics.

  - Leverage this expertise to act as an integrative force in bringing together clinical and biological data with results from diverse fields, applications, and enterprises.

  - In doing so, we endeavor to support the investigation of disease etiology, prevention, and treatment in a comprehensive and transdisciplinary fashion.

The Health Informatics Institute (HII)

- Functions:
  - coordinate and conduct clinical trials and health services research
  - provide scientific guidance on research study design, statistical methodology, and data management, analysis, and interpretation
  - design, develop, and maintain applications, databases, data warehouses, and data governance policies for the full data lifecycle needs of clinical research
  - provide and maintain comprehensive reporting services for clinical operations
  - collaborate with analytical partners in the design, development, and implementation of analysis software pipelines
  - offer a suite of data sharing platforms and interfaces

HII Projects

IT Infrastructure

- Oracle Application Databases

- SQL Server Data Warehouses

- SAS Grid Computing Environment
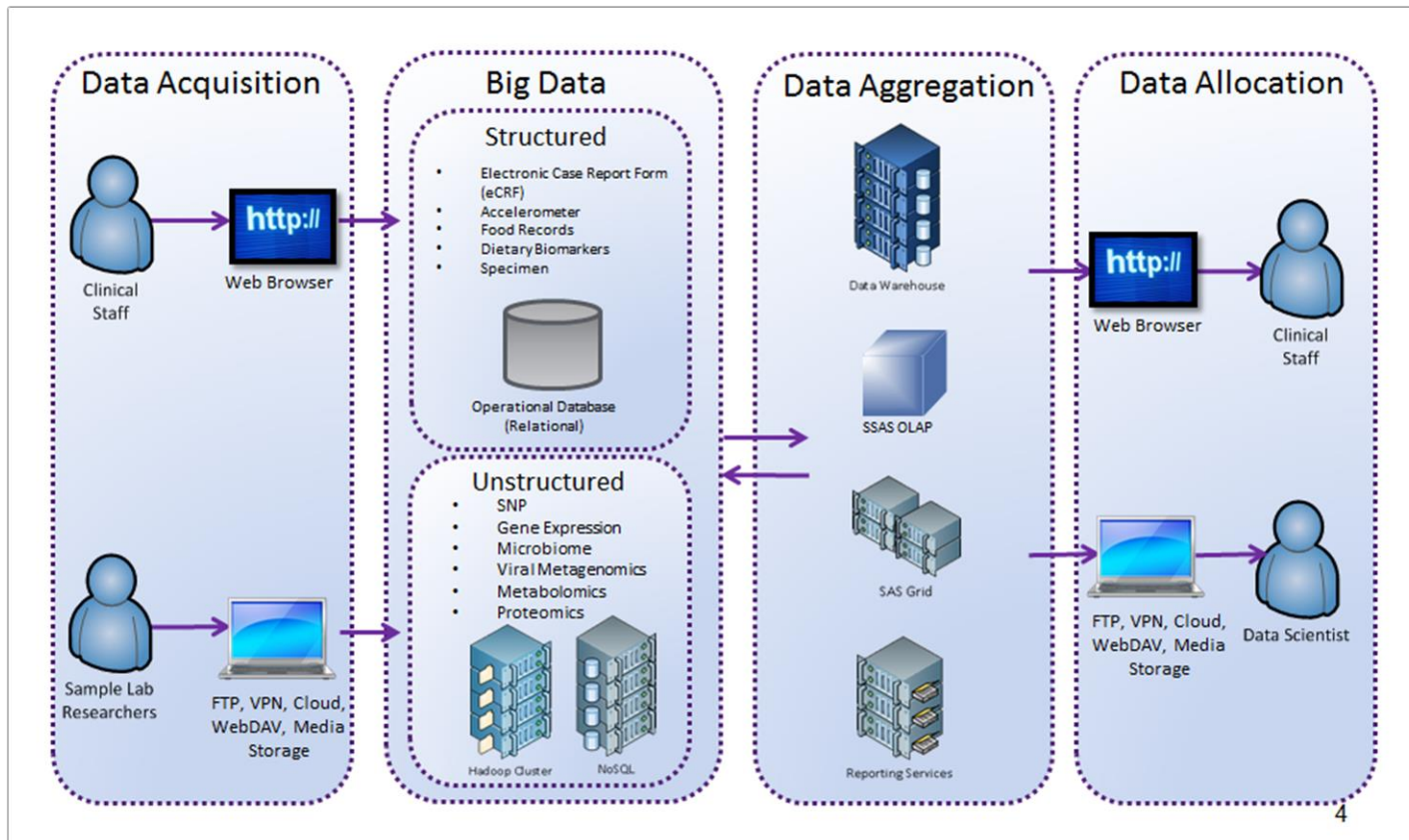
- Big Data (Hadoop Cluster, Hive DW)

- High Performance Computing Cluster

- Web Services and APIs

IT Infrastructure

IT Infrastructure

- ▶ All of this eventually feeds into reporting and analysis datasets

- ▶ This level of complexity is matched by the complexity of interactions in the networks we support

  - ▶ Challenge is exposing/leveraging the power of these data assets for consumption by study personnel

Access to Study Data

- Existing approaches to data sharing

- Demand from network investigators for access to study data

- Disorganization and lack of clarity of existing data sharing avenues

- Internal demand for ad hoc reporting of data

The LabKey Server Platform
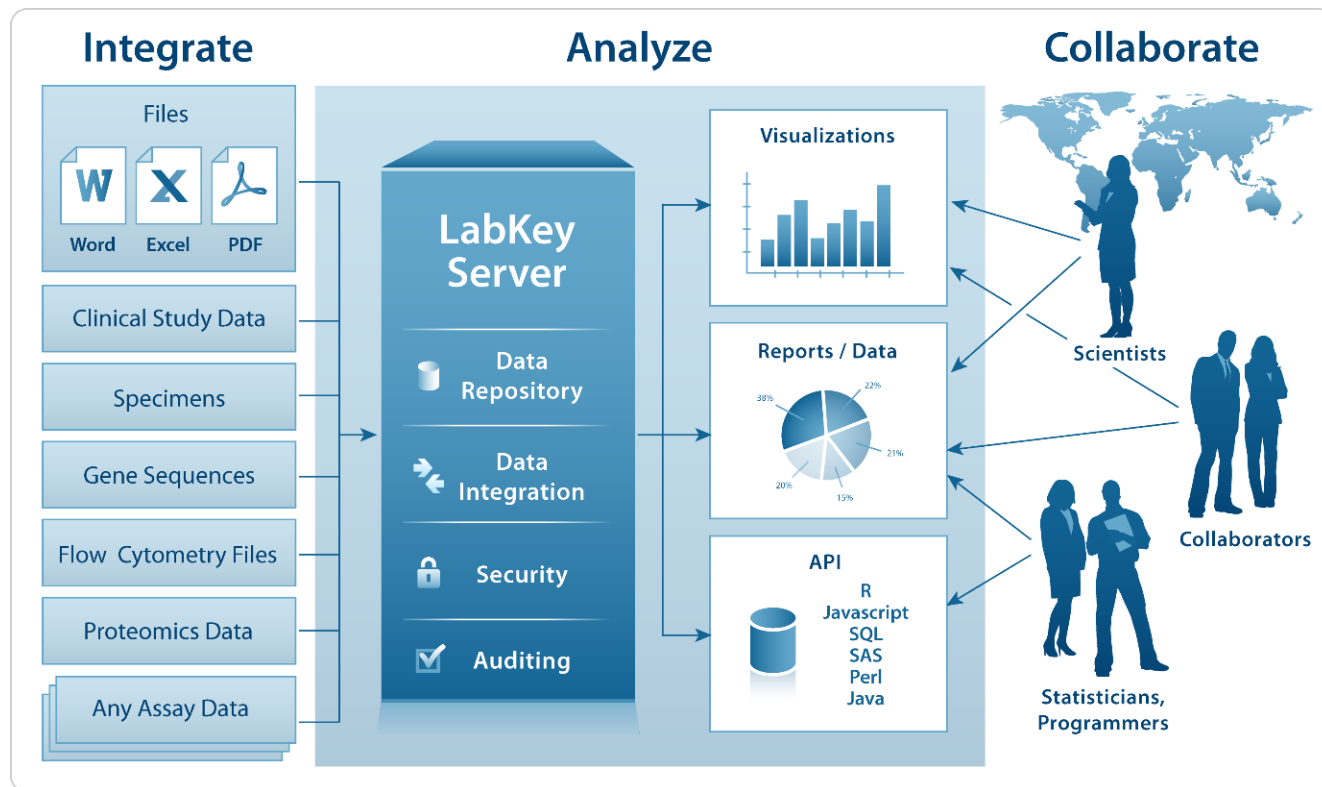
Implementation: Objectives

- To empower investigators and project staff to explore, analyze, and share study data in a convenient and accessible environment. This platform contains not only an expanding set of study data, but also a suite of built-in tools to support interaction with the data.

Implementation: TEDDY

- The Environmental Determinants of Diabetes (TEDDY)

  - A consortium of six clinical centers and a Data Coordinating Center (DCC) working to identify environmental triggers relating to Type I Diabetes Mellitus (T1DM) in genetically susceptible individuals.

  - Goal: the identification of infectious agents, dietary factors, or other environmental agents, including psychosocial factors which trigger T1DM in genetically susceptible individuals, or which protect against the disease.

Implementation: TEDDY

▶ DCC historical source of TEDDY analysis

▶ Data release

  ▶ NIDDK data sharing requirements

  ▶ Manuscript datasets

▶ Going further

  ▶ Bring these processes together

  ▶ But also, make critical study data available and accessible to investigators for exploratory analysis

TEDDY

Start Page | Study Data | Manuscripts | NIDDK Submissions | FAQ

Welcome

Welcome to TEDDY's Data Sharing Platform

TEDDY members can use this site to view and download data for The Environmental Determinants of Diabetes in the Young (TEDDY) study.

Data are categorized and archived into the following areas: **Study Data, Manuscripts,** and **NIDDK Submissions.** To view the data, please select one of the tabs in the top right corner.

Help

This site is hosted by the **TEDDY Data Coordinating Center** at the **University of South Florida**. For assistance, contact Dena Garcia or Sarah Austin-Gonzalez.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

Study Data	Manuscripts	
Antibodies	**Ancillary**	**Immune**
• GAD, IA2A, and MIAA, and T1D status	• MP89	• MP37
Demographics	**Celiac**	• MP78
• Demographics	• MP72	• MP127
TGA Endpoints	• MP74	**Infectious**
• Marsh scores, TGA, and Celiac Disease status	• MP85	• MP76
	Diet	• MP86
	• MP55	**Maternal**
	• MP100	• MP20
	Genetics	**Psychosocial**
	• MP56	• MP105
	• MP73	**Other**
		• MP88

NIDDK Submissions

Screening	3 Month Forms	Publications
• Screening Form	• 3 Month Interview	• Pub10_Sjohnson (MP10)
Enrollment		• Pub14_BLernmark (MP11)
• Enrollment Form	• Physical Exam Form	• Pub12_YSterner (MP8)
First Questionnaires		
• Mother Questionnaire		
• Father Questionnaire		
• Primary Caretaker Questionnaire		

TEDDY 🧸

| Start Page | Study Data | Manuscripts | NIDDK Submissions | FAQ |

Study Data Home

Welcome to TEDDY Study Data!

For variable definitions, see the corresponding data dictionary file.

Study Data ▾

☐ Mine

Name

☐ **Antibodies**

    ▦ Antibody Endpoints

    ▦ Antibody Endpoints Data Dictionary

☐ **Demographics**

    ▦ Demographics

    ▦ Demographics Data Dictionary

☐ **TGA**

    ▦ TGA Endpoints

    ▦ TGA Endpoints Data Dictionary

TEDDY

| Start Page | Study Data | Manuscripts | NIDDK Submissions | FAQ |

Antibody Endpoints

VIEWS ▾   CHARTS ▾   EXPORT ▾   PRINT   PAGE SIZE ▾

View: Antibody Endpoints

Data Cut Date	Mask Id	HLA Category	T1D	T1D Visit	Confirmation Visit	Confirmed Positive GAD	Confirmed Positive IA2A	Confirmed Positive MIAA	Confirmation Visit GAD	Confirmation Visit IA2A	Confirmation Visit MIAA	Persistent-confirmed GAD	Persistent-confirmed GAD Visit	Persistent-confirmed IA2A	Persistent-confirmed IA2A Visit	Persistent-confirmed MIAA
2015-04-30	568435	1	0													
2015-04-30	750921	4	0													
2015-04-30	534380	1	0													
2015-04-30	534288	1	0													
2015-04-30	850752	9	0													
2015-04-30	760612	1	0													
2015-04-30	824788	1	0													
2015-04-30	397929	2	0													
2015-04-30	574372	9	0													
2015-04-30	896599	9	0													
2015-04-30	621667	2	0													
2015-04-30	843090	1	0													
2015-04-30	800991	9	0													
2015-04-30	293377	1	0													
2015-04-30	852733	1	0													
2015-04-30	612066	4	0													
2015-04-30	393395	4	0	60	0	1	1		60	60	0		1	60		
2015-04-30	761319	2	0		0	0	0									
2015-04-30	664170	4	0													
2015-04-30	721144	9	0													
2015-04-30	297836	9	0													

Antibody Endpoints Data Dictionary

VIEWS ▾   CHARTS ▾   EXPORT ▾   PRINT   PAGE SIZE ▾   1 - 23 of 23

View: Antibody Endpoints Data Dictionary   ✕

Variable Name	Variable Label	Definition	Source Form
conf_visit	Confirmation Visit	the first visit at which subject had confirmed antibodies (any antibody)	
conf_visit_GAD	Confirmation Visit GAD	visit at which the GAD antibody was confirmed	
conf_visit_IA2A	Confirmation Visit IA2A	visit at which the IA2A antibody was confirmed	
conf_visit_MIAA	Confirmation Visit MIAA	visit at which the MIAA antibody was confirmed	
confpos_ab	Confirmed Antibody Positive	indicates whether or not the subject has confirmed positive antibodies (any antibody) – 0=no, 1=yes (confirmed positive is defined as the same sample being positive in both labs for the same antibody)	
confpos_GAD	Confirmed Positive GAD	indicates whether subject's confirmed positive antibody was GAD – 0=no, 1=yes (this data is only provided on those subject's that have a confirmed antibody)	
confpos_IA2A	Confirmed Positive IA2A	indicates whether subject's confirmed positive antibody was IA2A – 0=no, 1=yes (this data is only provided on those subject's that have a confirmed antibody)	
confpos_MIAA	Confirmed Positive MIAA	indicates whether subject's confirmed positive antibody was MIAA – 0=no, 1=yes (this data is only provided on those subject's that have a confirmed antibody)	
HLA_Category	HLA Category	subject's HLA category – this variable is numeric for analysis purposes -1='HLA*Results*Pending' 0='Not*Eligible' 1='DR4*030X/0302*DR3*0501/0201' 2='DR4*030X/0302*DR4*030X/0302' 3='DR4*030X/0302*DR4*030X/020X' 4='DR4*030X/0302*DR8*0401/0402' 5='DR4*030X/0302*DR1*0101/0501' 6='DR4*030X/0302*DR13*0102/0604' 7='DR4*030X/0302*DR4*030X/0304' 8='DR4*030X/0302*DR9*030X/0303' 9='DR3*0501/0201*DR3*0501/0201' 10='DR3*0501/0201*DR9*030X/0303' 99='Results*Under*Review'	
Mask_ID	Mask ID	a unique masked subject identifier	
matabexp	Maternal Antibody Exposure	the subject was exposed to maternal antibodies at some point in time - 1 = exposed, 0 = no exposure (includes unknown but possible)	
persist_conf_ab	persistent-confirmed antibody positive	indicates whether or not the subject has persistent confirmed positive antibodies (any antibody - GAD, IA2A or MIAA) 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)	
persist_conf_gad	persistent-confirmed GAD	indicates whether or not the subject has a persistent confirmed GAD antibody - 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)	
persist_conf_gad_visit	persistent-confirmed GAD Visit	visit at which the subject is considered persistent confirmed for GAD (this would be the first visit of the 2 consecutive visits)	
persist_conf_ia2a	persistent-confirmed IA2A	indicates whether or not the subject has a persistent confirmed IA2A antibody - 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)	
persist_conf_ia2a_visit	persistent-confirmed IA2A Visit	visit at which the subject is considered persistent confirmed for IA2A (this would be the first visit of the 2 consecutive visits)	
persist_conf_miaa	persistent-confirmed MIAA	indicates whether or not the subject has a persistent confirmed MIAA antibody - 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)	
persist_conf_miaa_visit	persistent-confirmed MIAA Visit	visit at which the subject is considered persistent confirmed for MIAA (this would be the first visit of the 2 consecutive visits)	

Manuscript Home

Welcome to TEDDY Manuscripts!

In the **Studies** section below, you will find links to manuscript datasets that are currently available to you. Datasets will only become available once a signed Data and Materials Distribution Agreement (DMDA) is received by the DCC. For inquiries on data access, please contact teddypublications@epi.usf.edu.

Studies

MP100 - Beyerlein, et al. (Diet)
Dietary intake of soluble fiber and risk of islet autoimmunity by 5 years of age: Results from the TEDDY study

MP105 - McIver, et al. (Psychosocial)
Cross-cultural comparisons of physical activity among 5-year-old children who are at risk for development of Type I diabetes

MP127 - Bonifacio, et al. (Immune)
Joint and functional modeling of longitudinal islet autoantibody trajectories

MP20 - Törn, et al. (Maternal)
Autoantibody positive mothers to children enrolled in the The Environmental Determinants of Diabetes in the Young (TEDDY)

MP37 - Vehik, et al. (Immune)
Islet Autoantibody Levels from Seroconversion to Time of Type 1 Diabetes Diagnosis

MP55 - Riikonen, et al. (Diet)
Milk feeding during the first year of life

MP56 - Hagopian, et al. (Genetics)
Overlap between Type 1 diabetes and Celiac Disease is primarily explained by shared HLA risk alleles- results from a prospective cohort study

MP72 - She, et al. (Celiac)
Two regions associated with celiac disease in the prospective TEDDY cohort in regions associated with autoimmunity from ImmunoChip genotyping

MP73 - She, et al. (Genetics)
Combinations of non-HLA genotypes modify the HLA-based genetic risk for islet autoimmunity and type 1 diabetes in the first 5 years of life in the prospective TEDDY cohort

MP74 - Koletzko, et al. (Celiac)
Risk of celiac specific autoimmunity and celiac disease in relation to mode of delivery– the results from a multi-center screening study

MP76 - Hyöty, et al. (Infectious)
Molecular epidemiology of enteroviruses in different populations – international comparison in the TEDDY study

MP78 - Achenbach, et al. (Immune)
Modeling of longitudinal islet autoantibody profiles

MP85 - Triplett, et al. (Celiac)
The Effect of Antibiotic Treatment in Early Childhood on the Development of Celiac Autoimmunity

MP86 - Triplett, et al. (Infectious)
The effect of antibiotic treatment in early childhood on the development of type 1 diabetes autoimmunity

MP88 - Qian, et al. (Other)
A Personalized Prognostic Model for Type-I Diabetes by Identifying and Synthesizing Risk-Predictive Rule

MP89 - Hyöty, et al. (Ancillary)

Help

This site is hosted by the **TEDDY Data Coordinating Center** at the **University of South Florida**. For assistance, contact Dena Garcia or Sarah Austin-Gonzalez.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

Overview   Datasets   Data Dictionary

MP127 - Overview

MP127: Joint and functional modeling of longitudinal islet autoantibody trajectories

Authors:
Ezio Bonifacio, Meike Köhler, Andreas Beyerlein, Sonja Greven, Kendra Vehik, Marian Rewers, Bill Hagopian, Jin-Xiong She, Åke Lernmark, Jorma Toppari, Olli Simell, Beena Akolkar, Jeffrey Krischer, Anette-G Ziegler, and the TEDDY study group

Dataset releases:

1. April 10, 2015 (initial release of assay, demographics, and SNP datasets to Meike Köhler)

- mp127_assay_data.csv - contains antibody results from the harmonized assay for GAD and IA2A and the TEDDY assay for GAD, IA2A and MIAA
- mp127_demo_data.csv - contains demographics and persistent confirmed antibody status.
- mp127_snp_data.csv - contains SNP results

Statistician:
Kendra Vehik

Manuscript Proposal

MP127 - Discussion Board

NEW

Showing: all comments

No comments

Help

Data can be viewed in the **Datasets** tab at the top right of the page.

Variable definitions can be found in the **Data Dictionary** tab at the top right of the page.

TEDDY

Overview   **Datasets**   Data Dictionary

MP127 - Datasets

☐ Mine

Name	Details	Access
⊟ **Demographics**		
▦ MP127_Demo_Data	☰	
⊟ **SNPs**		
▦ MP127_SNP_Data	☰	
⊟ **Test Results**		
▦ MP127_Assay_Data	☰	

TEDDY

Overview | Datasets | Data Dictionary

Dataset: MP127_Assay_Data, All Visits

Contains up to one row of MP127_Assay_Data data for each Participant/Visit/row_id combination.

Data Cut Date: 12-31-2013

FILTER | VIEWS ▾ | CHARTS ▾ | PARTICIPANT GROUPS ▾ | EXPORT ▾ | PAGE SIZE ▾ | VIEW SPECIMENS

1 - 100 of 121,304 Next › Last »

Mask ID	DUE NUM	Age Coll Months	Ref Result Gad	Ref Outcome Gad	Ref Assay Gad	Rep Result Gad	Rep Outcome Gad	Rep Assay Gad	Ref Result Miaa	Ref Outcome Miaa	Ref Assay Miaa	Rep Result Miaa	Rep Outcome Miaa	Rep Assay Miaa	Ref Result Ia2a	Ref Outcome Ia2a	Ref Assay Ia2a	Rep Result Ia2a	Rep Outcome Ia2a	R
10001	12	12	2.446	Neg	9932				0.0	Neg	9932				1.78	Neg	9932			
10001	15	15	2.742	Neg	9932				0.0	Neg	9932				2.02	Neg	9932			
10001	18	18	3.074	Neg	9932				0.0	Neg	9932				1.68	Neg	9932			
10001	21	21	4.536	Neg	9932				0.01	Neg	9932				1.66	Neg	9932			
10001	24	25	4.076861324	Neg	9932				1.0E-4	Neg	9932				1.25	Neg	9932			
10001	27	28	4.209	Neg	9932				0.0	Neg	9932				2.52	Neg	9932			
10001	3	4	2.059	Neg	9932				0.0	Neg	9932				1.46	Neg	9932			
10001	30	31	3.846	Neg	9932				0.11	Neg	9932				1.9	Neg	9932			
10001	33	33	4.344	Neg	9932				0.08	Neg	9932				2.02	Neg	9932			
10001	36	36	4.25	Neg	9932				0.0	Neg	9932				1.79	Neg	9932			
10001	39	39	3.61	Neg	9932				0.0	Neg	9932				2.34	Neg	9932			
10001	45	44	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	48	48	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	54	53	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	6	6	2.774	Neg	9932				0.0	Neg	9932				1.5	Neg	9932			
10001	60	60	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	66	66	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	72	72	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	78	79	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	84	85	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	9	9	2.202158073	Neg	9932				0.05	Neg	9932				1.470379825	Neg	9932			
10001	90	90	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10001	96	94	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10002	108	107	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10002	12	12	4.842118908	Neg	9932				0.0	Neg	9932				1.111855811	Neg	9932			
10002	15	16	3.067682406	Neg	9932				0.0	Neg	9932				0.884027526	Neg	9932			
10002	21	20	6.298	Neg	9932				0.0	Neg	9932				2.26	Neg	9932			
10002	3	3	2.606	Neg	9932				0.0	Neg	9932				1.304	Neg	9932			
10002	6	6	5.808529314	Neg	9932				0.0	Neg	9932				1.292452622	Neg	9932			
10002	84	83	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10002	9	10	4.136292114	Neg	9932				0.0	Neg	9932				1.885709941	Neg	9932			
10002	96	95	0.0	Neg	9933				0.0	Neg	9932				0.0	Neg	9933			
10003	12	12	5.767	Neg	9932				0.0	Neg	9932				1.8	Neg	9932			
10003	15	15	5.203	Neg	9932				0.28	Neg	9932				1.15	Neg	9932			
10003	18	18	3.834	Neg	9932				0.05	Neg	9932				1.02	Neg	9932			
10003	24	24	3.422	Neg	9932				0.047575532	Neg	9932				1.364033626	Neg	9932			
10003	3	4	2.121	Neg	9932				0.0	Neg	9932				2.23	Neg	9932			

TEDDY 🧸🌎

Overview   Datasets   **Data Dictionary**

MP127 - Data Dictionary ▾

Data dictionary for MP127: Joint and functional modeling of longitudinal islet autoantibody trajectories.

Freeze Date: December 31, 2014

Datasets provided for this manuscript: mp127_demo_data.csv mp127_assay_data.csv and mp127_snp_data.csv

Datasets created April, 2015 by M. Butterworth using data as of December 31, 2014. The demo_data dataset includes all enrolled TEDDY subjects, the assay_data dataset includes harmonized and TEDDY assay data, and the snp_data dataset includes data for requested SNPs.

Data dictionary for demo_data.csv:

mask_id – the masked ID for the subject

status – subject's current status in the study

cc – subject's clinical center – 1=COL, 2=GEO, 3=WAS, 4=FIN, 5=GER, 6=SWE, 133=NBD, 134=CHP

country – subject's country – 1=US, 2=FIN, 3=GER, 4=SWE

BabyBirthType – was subject a singleton, twin or triplet birth

sex – subject's gender (Male, Female)

fdr – subject's FDR status – 0= GEN POP (also includes unknown), 1=FDR

family_mem – indicates which family members are FDR – 0=none (GEN POP), 1=all, 2=mom & dad, 3=mom & sibling, 4=dad & sibling, 5=mom, 6=dad, 7=sibling

family – indicates which family member is FDR – .=GEN POP, 1=mom, 2=dad, 3=sibling. This was done in a priority order to establish only one family member for each subject. Rank from highest to lowest is mom, dad, sibling. So in other words if family_mem=2 then family would=1 based on the fact that mom is priority in ranking.

race_ethnicity – subject's race and ethnicity combined into 5 groups: 1=All Hispanics regardless of race, 2=White Non-Hispanic, 3=African Americans (non-Hispanic), 4=All Other Races (non-Hispanic), 5=Missing or Unknown race and ethnicity

indeterminate – indicates if subject's current antibody status is unable to be determined – 0=no, 1=yes

t1d – indicates if subject has developed T1D – 0=no, 1=yes

t1d_visit – visit at which subject was diagnosed with T1D – since diagnosis is not always in the TEDDY clinic at a visit this is estimated based on the date of diagnosis

age_t1d_months – age subject developed T1D (in months)

persist_conf_gad – indicates whether or not the subject has a persistent confirmed GAD antibody – 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)

persist_conf_gad_visit – visit at which the subject is considered persistent confirmed for GAD (this would be the first visit of the 2 consecutive visits)

age_persist_conf_gad_mos – age in months at the draw date of the persistent GAD antibody (this would be the first sample of the 2 consecutive samples)

persist_conf_ia2a – indicates whether or not the subject has a persistent confirmed IA2A antibody – 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)

persist_conf_ia2a_visit – visit at which the subject is considered persistent confirmed for IA2A (this would be the first visit of the 2 consecutive visits)

age_persist_conf_ia2a_mos– age in months at the draw date of the persistent IA2A antibody (this would be the first sample of the 2 consecutive samples)

persist_conf_miaa – indicates whether or not the subject has a persistent confirmed MIAA antibody – 0=no, 1=yes (persistent confirmed is defined as the antibody being confirmed positive on 2 consecutive visits)

TEDDY

NIDDK Submissions Home

Welcome to TEDDY NIDDK Submissions!

Data Submitted	Details
3 Month Interview	contains data for 8,663 participants who filled out the primary caretaker interview 3 Month Clinical Visit form.
3 Month Physical Exam Form	contains data for 8663 participants who filled out the 3 Month Clinic Visit Physical Examination form.
Enrollment Form	contains data for 21,575 participants who completed TEDDY enrollment forms. Data include birth date, clinical center, consent/exclude/refusal status.
Father Questionnaire	contains data for 8,030 participants who filled father form in TEDDY. Data include father's reactions to the baby's generic rest result and experience in the TEDDY study.
Mother Questionnaire	contains data for 8,529 participants who filled mother form in TEDDY. Data include mother's latest pregnancy and when the mother pregnant with the TEDDY child. It also includes illness, medications, diet, smoking, alcohol.
Primary Caretaker Questionnaire	contains data for 22 participants who filled Primary care Taker form in TEDDY. Data include mother's history of diabetes, and reactions to baby's test results and experience in the TEDDY study.
Screening Data	contains basic indicators corresponding with 424,788 participants screened for the TEDDY Study. Indicators include HLA screening genotype; any family member has type 1 diabetes (Y/N);mother, father, or sibling has type 1 diabetes (Y/N for each); and country of clinical center. The dataset does not contain information about subject enrollment.
Pub10_Sjohnson (MP10)	contains data for 4,897 unique eligible participants who enrolled in TEDDY. Data include parent characteristics, family history of type 1 diabetes, and reasons for early withdrawal from the study.
Pub14_BLernmark (MP11)	contains 18,129 records which constitute data for 16,435 unique participants. Data include basic demographics, family history of type 1-diabetes, and reasons for ineligibility and/or refusal to participate in the study.
Pub12_YSterner (MP8)	contains data for 5,461 unique eligible participants who enrolled in TEDDY with birth weight ranged from 2 to 6 kg and birth length from 40 to 60 cm. Data include baby's gender, race, length and weight, mother's Gestational age, Maternal Age (in Years), smoking status, height and weight, father's height. The data also include Number of Alcoholic Drinks, HLA, Country, and Delivery Complications.

Files

PARENT FOLDER

	Name	Size
☐ 📁	3 Month Interview	
☐ 📁	3 Month Physical Exam	
☐ 📁	Enrollment Data	
☐ 📁	Father Questionnaire	
☐ 📁	Mother Questionnaire	

Welcome to TEDDY NIDDK Submissions!

Data Submitted	Details
3 Month Interview	contains data for 8,663 participants who filled out the primary caretaker interview 3 Month Clinical Visit form.
3 Month Physical Exam Form	contains data for 8663 participants who filled out the 3 Month Clinic Visit Physical Examination form.
Enrollment Form	contains data for 21,575 participants who completed TEDDY enrollment forms. Data include birth date, clinical center, consent/exclude/refusal status.
Father Questionnaire	contains data for 8,030 participants who filled father form in TEDDY. Data include father's reactions to the baby's generic rest result and experience in the TEDDY study.
Mother Questionnaire	contains data for 8,529 participants who filled mother form in TEDDY. Data include mother's latest pregnancy and when the mother pregnant with the TEDDY child. It also includes illness, medications, diet, smoking, alcohol.
Primary Caretaker Questionnaire	contains data for 22 participants who filled Primary care Taker form in TEDDY. Data include mother's history of diabetes, and reactions to baby's test results and experience in the TEDDY study.
Screening Data	contains basic indicators corresponding with 424,788 participants screened for the TEDDY Study. Indicators include HLA screening genotype; any family member has type 1 diabetes (Y/N);mother, father, or sibling has type 1 diabetes (Y/N for each); and country of clinical center. The dataset does not contain information about subject enrollment.
Pub10_Sjohnson (MP10)	contains data for 4,897 unique eligible participants who enrolled in TEDDY. Data include parent characteristics, family history of type 1 diabetes, and reasons for early withdrawal from the study.
Pub14_BLernmark (MP11)	contains 18,129 records which constitute data for 16,435 unique participants. Data include basic demographics, family history of type 1-diabetes, and reasons for ineligibility and/or refusal to participate in the study.
Pub12_YSterner (MP8)	contains data for 5,461 unique eligible participants who enrolled in TEDDY with birth weight ranged from 2 to 6 kg and birth length from 40 to 60 cm. Data include baby's gender, race, length and weight, mother's Gestational age, Maternal Age (in Years), smoking status, height and weight, father's height. The data also include Number of Alcoholic Drinks, HLA, Country, and Delivery Complications.

Files

PARENT FOLDER

	Name	Size
☐	05212010_441_TE_Interview3Month.pdf	62.3 KB
☐	niddk_submission_form.xls	31 KB
☐	three_month_int_file_data_dictionary.rtf	1.4 MB
☐	three_month_int_form.sas7bdat	542.5 MB

TEDDY

Frequently Asked Questions

Can I export and download data from LabKey?

LabKey provides a variety of methods for exporting data. You can export data in a grid view to an Excel file, a TSV text file, an Excel Web Query, or a variety of different scripts, which can recreate the data grid. Click the **Export** button above any grid view and use the left-hand tabs to choose between Excel, Text and Script options. You can also choose whether to export the entire set of data or only selected rows.

How can I access LabKey from SAS?

You must first configure your SAS installation to use the SAS/LabKey interface. For more information, please visit LabKey Documentation

Why can't I see any data under the Manuscripts tab?

Only authorized users who have signed a DMDA can view manuscript data.

Who should I contact if I have questions about the site?

Questions about site navigation or technical issues can be directed to Dena Garcia or Sarah Austin-Gonzalez. All data and access-related questions should be directed to teddypublications@epi.usf.edu.
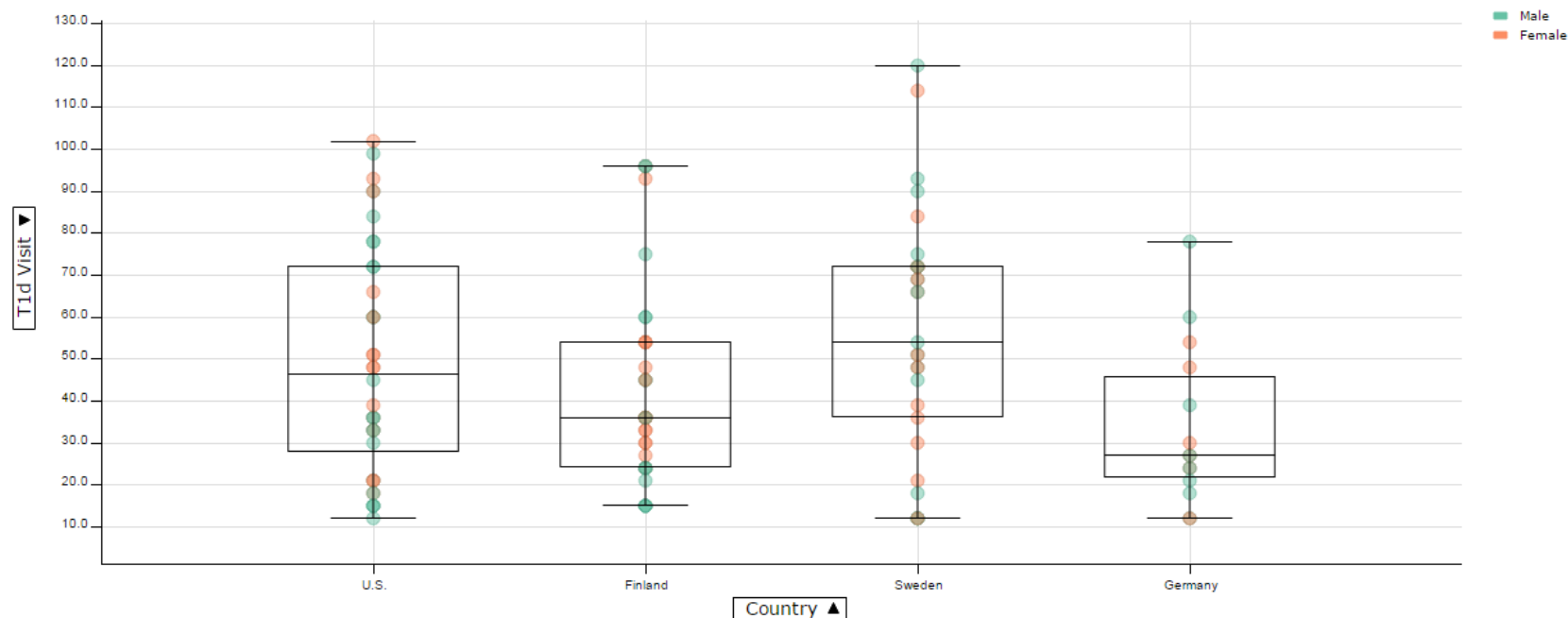
TEDDY

Box Plot Report

Box Plot Report

VIEW DATA   EXPORT ▾   OPTIONS   GROUPING   HELP   SAVE
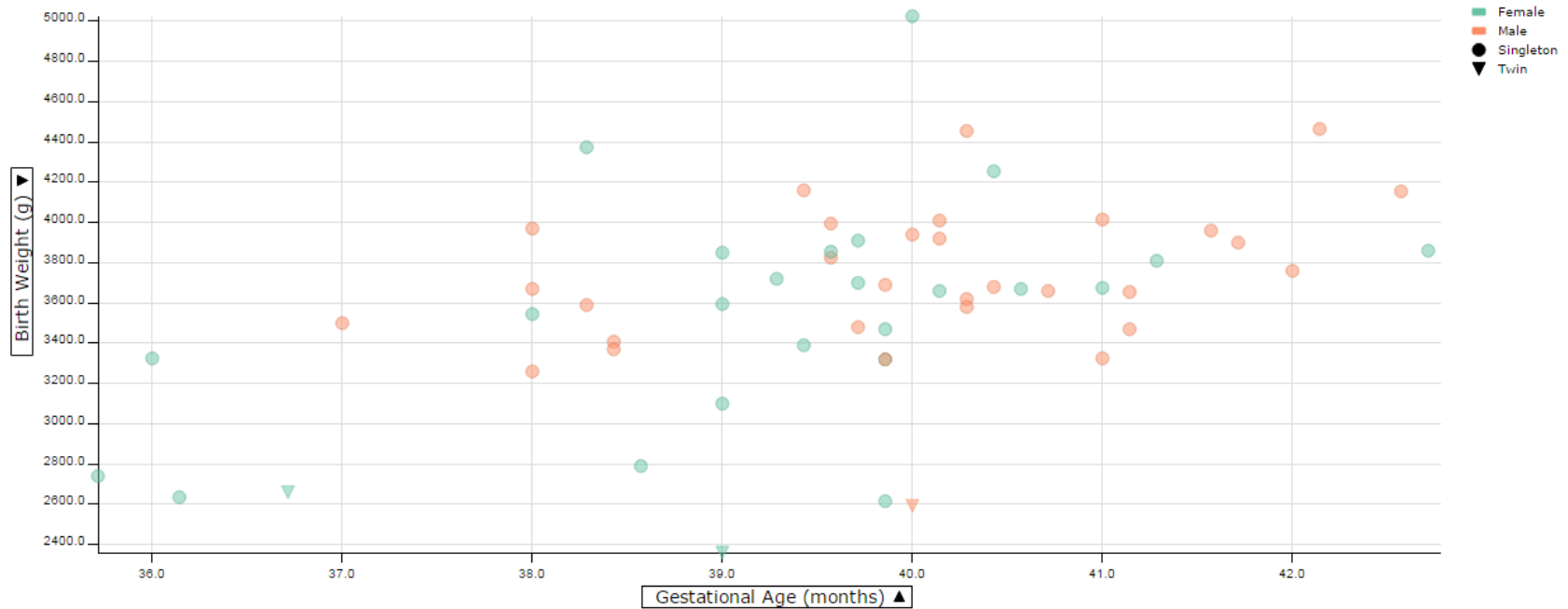

T1D Diagnosis Visit by Country ▼

Legend:
- Male
- Female

Y-axis: T1d Visit ▶ (values: 10.0 to 130.0)

X-axis: Country ▲ (U.S., Finland, Sweden, Germany)

TEDDY

Scatter Plot Report

Scatter Plot Report

VIEW DATA   EXPORT ▾   OPTIONS   GROUPING   HELP   SAVE

Birth Weight by Gestational Age ▾

- ■ Female
- ■ Male
- ● Singleton
- ▼ Twin

Birth Weight (g) ▶

Gestational Age (months) ▲

TEDDY
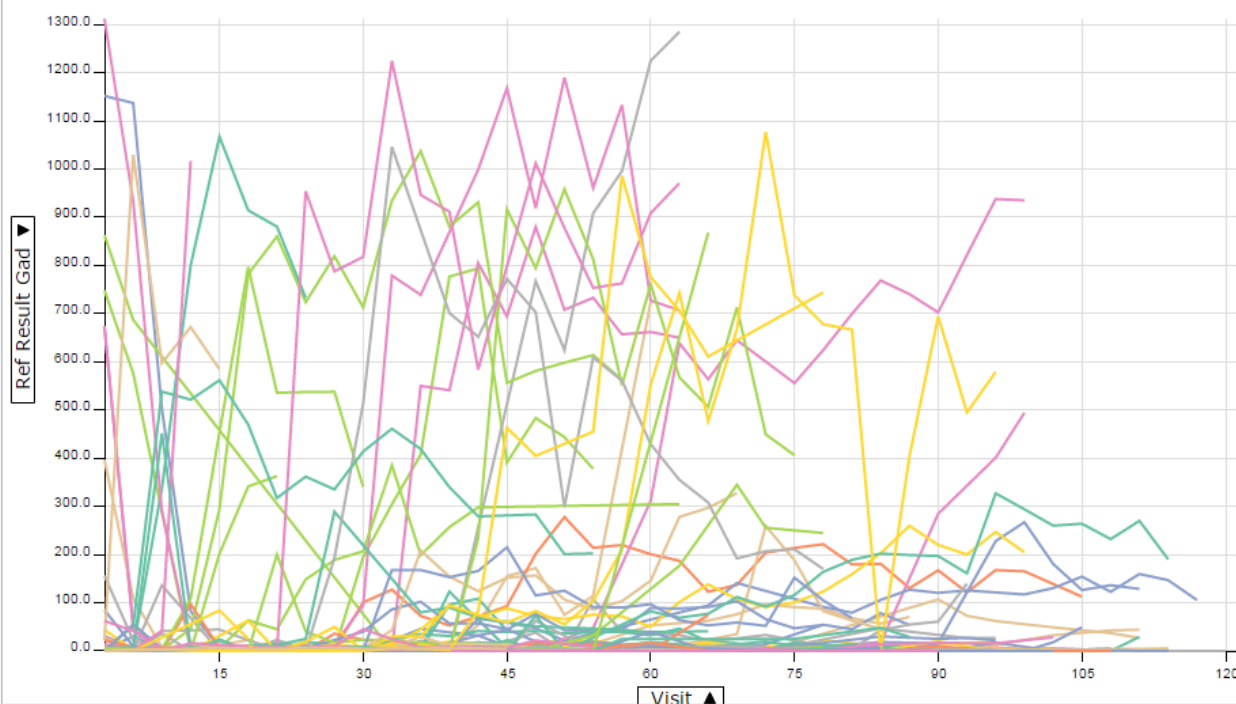
Search

Time Chart Wizard

VIEW DATA   EXPORT ▾   MEASURES   GROUPING   OPTIONS   SAVE

Filters »

The data limit for plotting has been reached. Consider filtering your data.

Participants —

MP127_Assay_Data ▾

☑ **All**
☑ 10001
☑ 10002
☑ 10003
☑ 10004
☑ 10005
☑ 10006
☑ 10007
☑ 10008
☑ 10009
☑ 10010
☑ 10011
☑ 10012
☑ 10013
☑ 10014
☑ 10015
☑ 10016
☑ 10017
☑ 10018
☑ 10019
☑ 10020
☑ 10021
☑ 10022
☑ 10023
☑ 10024
☑ 10025
☑ 10026
☑ 10027
☑ 10028

Legend: 10001, 10002, 10003, 10004, 10005, 10006, 10007, 10008, 10009, 10010, 10011, 10012, 10013, 10014, 10015, 10016, 10017, 10018, 10019, 10020, 10021, 10022, 10023, 10024, 10025, 10026, 10027, 10028, 10029, 10030, 10031, 10032, 10033, 10035

Y-axis: Ref Result Gad ▸

X-axis: Visit ▲

Implementation: TrialNet

- Type 1 Diabetes TrialNet

  - An international network of researchers who are exploring ways to prevent, delay and reverse the progression of type 1 diabetes.

  - Conducting clinical trials at 18 Clinical Centers in the United States, Canada, Finland, United Kingdom, Italy, Germany, Australia, and New Zealand.

  - More than 150 medical centers and physician offices participate in the network.

  - Studies available for people newly diagnosed with type 1 diabetes, and relatives of people with type 1 diabetes who are at greater risk of developing the disease.

Implementation: TrialNet

- Access to comprehensive Pathway to Prevention dataset
  - current 1 month (August 31, 2015)
- View and download datasets
- Generate reports and perform basic data plots
- Create customize views for each dataset(s)
- Save custom settings to user profile

Implementation: TrialNet

▶ Access to data enables informed analyses, characterization of population, and study feasibility planning

▶ Highly customizable to individual needs and preferences

▶ User-defined views, reports, participant groups, and data plots can be saved and shared with other users

▶ Diverse tools accommodate a variety of data types (form entries vs. lab results)

▶ Flexible configuration; data exports in Excel, Text, Script

▶ Sharing functions maximize collaboration

▶ Investigators can build off each other's work by sharing user-generated reports and analyses

Type1
Diabetes
TrialNet

TrialNet

Welcome

Welcome to TrialNet Data Explorer

TrialNet members can use this data sharing site to view and download clinical trial data from the following studies:

Observational	New Onset	Pilot
• **TN01 Pathway to Prevention**	• **TN02 MMF/DZB** • **TN05 Anti-CD20 (Rituxumab)** • **TN08 GAD Vaccine** • **TN09 CTLA-4 Ig (Abatacept)** • **TN14 Anti IL-1 Beta (Canakinumab)**	• **TN06 NIP Diabetes Pilot**

Under the "**Studies**" heading below, click on each trial name to access posted data. Datasets are categorized as **Test Results** or **Forms.** Users can create and save custom views that incorporate fields from multiple datasets.

Datasets and Tools

On each study's homepage, Click on the **Test Results and Forms** tab at the top right to see a list of available datasets, categorized as Test Results or Forms. Access data tools from the **Site Navigation** box.

Study Overview

Provides a calendar-based view of all study datasets.

- Each **dataset** is listed as a row in the Overview.
- Each **visit** is displayed as a column.
- The numbers in each square indicate the **participant count**, or the number of participants for which data is available for each dataset for that particular visit. Use the checkboxes to switch to a **row count**; the number of data entries for each dataset for that visit.
- To display information collected at a particular visit, click the number at the intersection of the dataset and visit you are interested in. All data collected for this particular dataset at this particular visit are displayed.
- The "**All**" column displays the cumulative sum of participants (or rows) at each visit for all applicable visits in a dataset.

Participant Reports

Provide data on one or more individual participants for selected measures (data fields). Users may select data fields from all available datasets.

Plot Data Over Time

Search TrialNet

Guest User

Pathway to Prevention Home

TN01 Pathway to Prevention: Natural History Study of the Development of Type 1 Diabetes

Datasets are categorized as **Test Results** or **Forms.** Each TN01 dataset is described below. Users can create and save custom views that incorporate fields from multiple datasets.

Test Result Datasets	**Form Datasets**

Autoantibody Test Results

- GAD65, ICA512/IA-2, mIAA, ZnT8, ICA titer values

HbA1c Test Results

- HbA1c value (%)

OGTT Test Results

- Glucose, insulin, c-peptide results at -10, 0, 30, 60, 90, & 120 minutes
- Overall glucose tolerance interpretation

HLA Test Results

- Absence/presence of DQA1*0102, DQB1*0602 protective allele
- Absence/presence of DR3 and DR4
- HLA Haplotypes α & β (DQA1, DQB1, DRB1 typing)

Screening Form

- Demographics
- Proband details
- How participant heard about TrialNet
- Consent data

Family History Form

- Data on participant's relatives with T1D

Baseline Risk Assessment Form

- Eligibility and consent
- Height, weight, vital signs
- Medical and medication history

Follow-up Risk Assessment Form

- Changes in eligibility and consent
- Height, weight, vital signs
- Changes in medical and medication history

Diabetes Onset Form

- Date and location of T1D diagnosis
- Circumstances of diagnosis (symptoms, lab results, hospitalization)

Change of Status Form

- Inactivation or reactivation of participant
- Reason for status change
- Willingness to be contacted while inactive

Site Navigation

STUDY OVERVIEW

CREATE PARTICIPANT REPORT

PLOT DATA OVER TIME

TRIALNET HOME

Help

This site is hosted by the **TrialNet Coordinating Center** at the **University of South Florida**. For assistance, contact Sarah Muller.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

Pathway To Prevention Home | Test Results and Forms | Specimen Data

Overview:TN01 - Pathway to Prevention

Guest User

VIEWS ▸   **SPECIMENS ▸**   **SHOW ALL DATASETS ▸**

Participant's current cohort: [All ▾]   ☑ Participant Count   ☐ Row Count

	All Visits	Screening	Rescreen - 1 Year	Rescreen - 2 Years	Rescreen - 3 Years	Rescreen - 4 Years	Rescreen - 5 Years	Rescreen - 6 Years	Rescreen - 7 Years	Rescreen - 8 Years	Rescreen - 9 Years	Rescreen - 10 Years	Rescreen - 11 Years	Rescreen - 12 Years	Rescreen - 13 Years
Test Results															
Autoantibody Test Results?	236260	149260	19833	13871	9466	6391	4016	2458	1297	677	344	130	10		
HbA1c Test Results?	15018														
HLA Test Results?	4343														
OGTT Test Results?	14303														
Forms															
Screening?	154285	154285													
Family History?	151988	151988													
Diabetes Onset?	1168														
Change of Status?	10970														
Follow-Up Risk Assessment?	11257														
Baseline Monitoring Risk Assessment?	4171														
DPT-1 Entry?	116	116													
Control Entry?	102	102													

Type 1
Diabetes
TrialNet

TrialNet

Search TrialNet

Pathway To Prevention Home | Test Results and Forms | Specimen Data

Participant Report

Guest User

Participant Report ✎

Add Measure...   ✖

Choose a data measure: Autoantibody Test Results : MIAA (index units)

Filter: [Search]   ☐ Show all

	Dataset/Query	Measure	Description
☐	Autoantibody Test Results	Draw Date	
☐	Autoantibody Test Results	GAD65 (NIDDK units/mL)	
☐	Autoantibody Test Results	GAD65H (NIDDK units/mL)	
☑	Autoantibody Test Results	IA-2H (NIDDK units/mL)	
☑	Autoantibody Test Results	ICA (5*2^N JDF units)	
☑	Autoantibody Test Results	ICA512 (NIDDK units/mL)	
☑	Autoantibody Test Results	MIAA (index units)	
☐	Autoantibody Test Results	Visit	The ID of the visit.
☐	Autoantibody Test Results	ZNT8 (index units)	
☐	Baseline Monitoring Risk Ass...	Addison's Disease Within Las...	
☐	Baseline Monitoring Risk Ass...	Allergies Specify	
☐	Baseline Monitoring Risk Ass...	Allergies Within Last Year	
☐	Baseline Monitoring Risk Ass...	Alopecia Within Last Year	
☐	Baseline Monitoring Risk Ass...	Asthma Within Last Year	
☐	Baseline Monitoring Risk Ass...	Beta Blockers	
☐	Baseline Monitoring Risk Ass...	Beta Blockers Specify	
☐	Baseline Monitoring Risk Ass...	Cancer Within Last Year	
☐	Baseline Monitoring Risk Ass...	Celiac Disease Within Last Y...	
☐	Baseline Monitoring Risk Ass...	Colitis or Colon Problems Wit...	

SELECT   CANCEL

EXPORT ▾   TRANSPOSE   Showing 0 Results

CANCEL   SAVE

Type1 Diabetes TrialNet

TrialNet

🔍 Search TrialNet

| Pathway To Prevention Home | **Test Results and Forms** | Specimen Data |
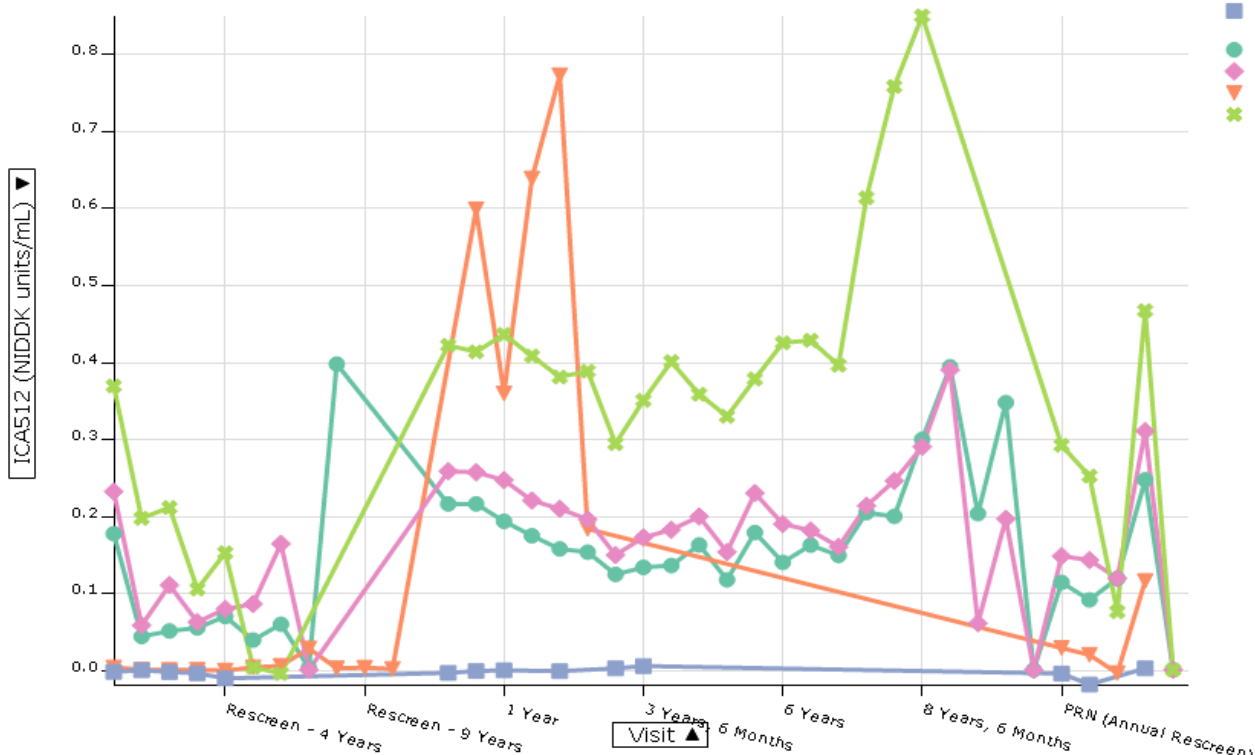
Test Results and Forms ▾

🔍 name, category, etc.   ☐ Mine

Name	Data Cut Date
⊟ **Test Results**	
▦ Autoantibody Test Results	2015-06-30
▦ HbA1c Test Results	2015-06-30
▦ HLA Test Results	2015-06-30
▦ OGTT Test Results	2015-06-30
⊟ **Forms**	
▦ Screening	2015-06-30
▦ Family History	2015-06-30
▦ Diabetes Onset	2015-06-30
▦ Change of Status	2015-06-30
▦ Follow-Up Risk Assessment	2015-06-30
▦ Baseline Monitoring Risk Assessment	2015-06-30
▦ DPT-1 Entry	2015-06-30
▦ Control Entry	2015-06-30

Guest User

Site Navigation

STUDY OVERVIEW

CREATE PARTICIPANT REPORT

PLOT DATA OVER TIME

PATHWAY TO PREVENTION HOME

TRIALNET HOME

Help

This site is hosted by the **TrialNet Coordinating Center** at the **University of South Florida**. For assistance, contact Sarah Muller.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

TrialNet

| Pathway To Prevention Home | **Test Results and Forms** | Specimen Data |

Dataset: Autoantibody Test Results, All Visits

Guest User

Autoantibody Test Results

Data Cut Date: 2015-06-30

FILTER   VIEWS ▾   CHARTS ▾   PARTICIPANT GROUPS ▾   EXPORT ▾   PAGE SIZE ▾   VIEW SPECIMENS

1 - 100 of 240,750   Next ›   Last »

	Participant ID (Masked) ▲	Visit	Draw Date	GAD65 (NIDDK units/mL)	GAD65H (NIDDK units/mL)	ICA512 (NIDDK units/mL)	IA-2H (NIDDK units/mL)	MIAA (index units)	ZNT8 (index units)	ICA (5*2^N JDF units)
☐	200000	Screening	2007-06-29	-0.0020		0.0030		-0.0010		
☐	200014	Screening	2009-10-02	0.0180		0.0010		0.0010		
☐	200014	Rescreen - 1 Year	2010-09-10		0		0	0.0010		
☐	200023	Screening	2011-04-01		0		0	0.0050		
☐	200023	Rescreen - 3 Years	2014-03-08		0		0	0.0010		
☐	200024	Screening	2006-03-21	0.0020		0.0070		0.0010		
☐	200025	Screening	2005-07-17	-0.0090		0.0060		0.0010		
☐	200027	Screening	2009-01-12	-0.0010		-0.0020		-0.0040		
☐	200030	Screening	2007-09-15	-0.0060		-0.0020		0.0010		
☐	200041	Screening	2014-07-23		0		0	0.0010		
☐	200045	Screening	2011-03-02		2		0	0.0030		
☐	200054	Screening	2008-05-14	-0.0080		-0.0030		-0.0020		
☐	200058	Screening	2013-05-04		3		0	0.0040		
☐	200061	Screening	2012-05-30		1		0	0.0000		
☐	200064	Screening	2009-06-26	0.0060		0.0010		0.0010		
☐	200065	Screening	2007-07-03	-0.0100		-0.0010		-0.0080		
☐	200065	Rescreen - 1 Year	2008-07-07	0.0030		-0.0210		-0.0100		
☐	200065	Rescreen - 2 Years	2009-07-15	0.0050		0.0020		0.0070		
☐	200065	Rescreen - 3 Years	2010-07-05		0		0	0.0010		
☐	200065	Rescreen - 4 Years	2011-07-18		6		0	-0.0040		
☐	200065	Rescreen - 5 Years	2012-07-19		0		0	0.0010		
☐	200074	Screening	2011-06-29		2		1	0.0020		
☐	200074	Rescreen - 3 Years	2014-05-23		1		0	0.0000		

Implementation: RDCRN

▶ Rare Diseases Clinical Research Network (RDCRN)

   ▶ Established by the Office of Rare Diseases Research (ORDR), which is now a part of the National Center for Advancing Translational Sciences (NCATS). The RDCRN began in 2003.

   ▶ Over 200 diseases are studied by the 22 research groups (consortia) of the RDCRN.

   ▶ Over 253 clinical centers have participated in over 76 research studies conducted through the RDCRN.

   ▶ Funding and/or programmatic support for the RDCRN has been provided by: Office of Rare Diseases Research, National Center for Advancing Translational Sciences (ORDR, NCATS), National Cancer Institute (NCI), National Heart, Lung, and Blood Institute (NHLBI), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), National Institute of Dental and Craniofacial Research (NIDCR), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Mental Health (NIMH), National Institute of Neurological Disorders and Stroke (NINDS)

Implementation: RDCRN

- ▶ Currently implemented for 2 of 22 consortia
  - ▶ LDN and UCDC

- ▶ Users may have access to one or more consortia
  - ▶ Links to consortia folders conditionally displayed based on permissions

- ▶ Dedicated landing page for each consortium
  - ▶ Homepage wiki with access, study, and data documentation
  - ▶ eCRF and test result data
  - ▶ Datasets refreshed via scheduled study reload process from SAS produced extracts
    - ▶ Built into existing production processes
    - ▶ Necessary XML files generated by SAS production code

RARE DISEASES
CLINICAL RESEARCH NETWORK
Initiative of the National Center for Advancing Translational Sciences (NCATS)

🔍 Search

Guest User

Welcome

Welcome to RDCRN Data Explorer

Members of the Rare Diseases Clinical Research Network can use this data sharing site to view and download clinical trial data from the following consortia:

- **LDN** Lysosomal Disease Network

- Urea Cycle Disorders Consortium

Consortia available to you are listed under the **"Consortia"** heading below. Click on a consortium name to access posted studies.

Consortia

Lysosomal Disease Network

Urea Cycle Disorders Consortium

Help

This site is hosted by the **Health Informatics Institute** at the **University of South Florida**. For assistance, contact explore@epi.usf.edu.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

Lysosomal Disease Network

Available protocols are listed below in the "Studies" section.

On each study's homepage, Click on the **Test Results and Forms** tab at the top right to see a list of available datasets, categorized as Test Results or Forms. Access data tools from the **Site Navigation** box. Users can create and save custom views that incorporate fields from multiple datasets.

Datasets and Tools

Study Overview:

Provides a calendar-based view of all study datasets.

- Each **dataset** is listed as a row in the Overview.
- Each **visit** is displayed as a column.
- The numbers in each square indicate the **participant count**, or the number of participants for which data is available for each dataset for that particular visit. Use the checkboxes to switch to a **row count**; the number of data entries for each dataset for that visit.
- To display information collected at a particular visit, click the number at the intersection of the dataset and visit you are interested in. All data collected for this particular dataset at this particular visit are displayed.
- The "**All**" column displays the cumulative sum of participants (or rows) at each visit for all applicable visits in a dataset.

Participant Reports:

Provide data on one or more individual participants for selected measures (data fields). Users may select data fields from all available datasets.

Plot Data Over Time:

Time Charts provide time-based visualizations for datasets. The X-axis shows a calculated time interval or visit series, while the Y-axis shows one or more numerical measures of your choice.

- Select which study participants, cohorts, or groups appear in the chart.
- Refine your chart by defining data dimensions and groupings.

LDN
Lysosomal Disease Network

Help

This site is hosted by the **Health Informatics Institute** at the **University of South Florida**. For assistance, contact explore@epi.usf.edu.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

Studies

p6703: Longitudinal Studies of Brain Structure and Function in MPS Disorders
p6703: Longitudinal Studies of Brain Structure and Function in MPS Disorders tracks data in 19 datasets over 7 visits. Data is present for 139 Participants.

Chester Whitley, Ph.D., M.D.

p6703 Home

Longitudinal Studies of Brain Structure and Function in MPS Disorders

Datasets are categorized as **Test Results** or **Forms.** Each dataset is described below. Users can create and save custom views that incorporate fields from multiple datasets.
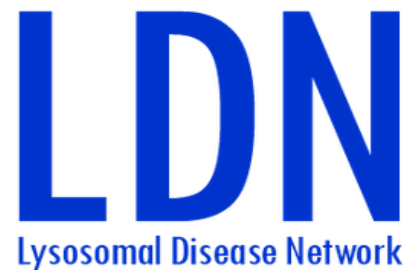
Test Result Datasets

Biological Laboratory Results include:

- Heparin

Form Datasets

Data Collection Forms include:

- Adverse Events
- Concomitant Medications
- Eligibility
- Diagnosis
- Quality of Life
- Behavioral Measures
- Treatment History
- Medical History
- Conclusion of Study Participation
- Death Record
- Demographics
- Registration

Data Dictionaries:

Protocol_6703.Data Dictionary.rtf

Study Tools

STUDY OVERVIEW
CREATE PARTICIPANT REPORT
PLOT DATA OVER TIME
RDCRN HOME
LDN HOME

Help

LDN Lysosomal Disease Network

Files

	Name	Last Modified	Size	Created By	Description	Usages
	export					
	folder_load_2015-07-31_14-45-25.log	2015-07-31 14...	9.9 KB			
	folder_load_2015-07-31_14-56-55.log	2015-07-31 14...	18.2 KB			

Data Views ▾

☐ Mine

Name	Details	Access
⊟ **Uncategorized**		
concomitant_meds	☰	
eligibility	☰	
p6703_14071_labs	☰	
diagnosis	☰	
ldn_14090_mri	☰	
adaptive_qol	☰	
behavioral_measures	☰	
cognitive_lt4	☰	
cognitive_4to6	☰	
cognitive_6to8	☰	
treatment_history	☰	
medical_history	☰	
cognitive_gt8	☰	
p6703_19914_protocold	☰	
adverse_event	☰	
conclusionofstudyparticipation	☰	
deathrecord	☰	
demographics	☰	
registration	☰	
Demo - Demographics Form	☰	public

Study Tools

STUDY OVERVIEW

CREATE PARTICIPANT REPORT

PLOT DATA OVER TIME

RDCRN HOME

LDN HOME

Help

This site is hosted by the **Health Informatics Institute** at the **University of South Florida**. For assistance, contact explore@epi.usf.edu.

This application uses LabKey Server. For general information, visit labkey.org:

- Documentation
- Tutorials

Dataset: concomitant_meds, All Visits

concomitant_meds

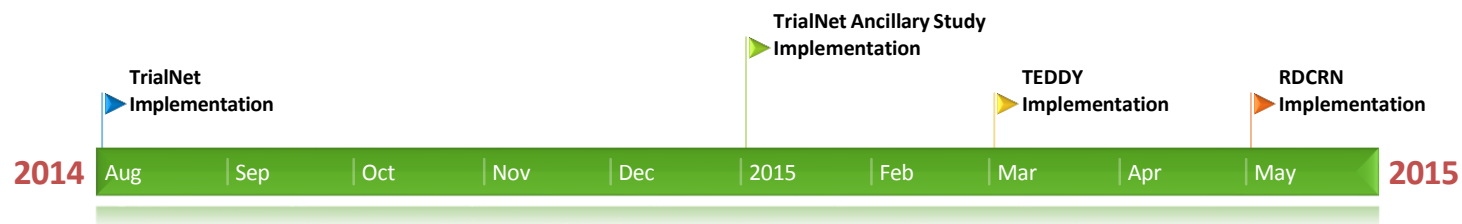FILTER   VIEWS ▾   CHARTS ▾   PARTICIPANT GROUPS ▾   EXPORT ▾   PAGE SIZE ▾   VIEW SPECIMENS   1 - 100 of 590

	Participant ID	Visit	assess_date	AssessDateDay	AssessDateYear	start_date	StartDateDay	StartDateYear	stop_date	StopDateDay	StopDateYear	User ID of Interviewer	User ID of Data Entry Personnel	Id
☐	107225	Baseline	2010-06-29	29	2010			2007			2012	3318	3318	MP
☐	107225	Baseline						2009				3318	3318	MP
☐	107225	Baseline						2007			2012	3318	3318	MP
☐	107225	Baseline						2010				3318	3318	MP
☐	107225	Baseline						2010				3318	3318	MP
☐	107225	Baseline						2010				3318	3318	MP
☐	107225	Baseline						2007			2009	3318	3318	MP
☐	107225	Baseline						2004			2008	3318	3318	MP
☐	107225	Baseline										3318	3318	MP
☐	107226	Baseline	2011-10-20	20	2011			2009				3318	3318	MP
☐	107226	Baseline						2008				3318	3318	MP
☐	107226	Baseline						2011				3318	3318	MP

Implementation: Results

▶ Data made available on our LabKey server ranges from eCRFs, to assay results, genetic (SNP) data, activity monitor data.

▶ [IMPRESSIVE NUMBERS]

    ▶ 75+ active users

    ▶ 3 projects with 30+ "studies"

    ▶ Data from over 170,000 subjects

    ▶ 300+ datasets including from 1 to 13,615,779 records each

        ▶ 50+ GB of data overall

Implementation: Timeline

TrialNet Ancillary Study Implementation

TrialNet Implementation

TEDDY Implementation

RDCRN Implementation

2014 | Aug | Sep | Oct | Nov | Dec | 2015 | Feb | Mar | Apr | May | 2015

Implementation: Challenges

- Technical
  - Data volume
    - Number of subjects in TrialNet, number of records in TEDDY, number of datasets in RDN
  - Integrating data from multiple sources
    - External SQL server schemas, manual uploads, SAS uploads
  - Managing structure and permissions for multiple different projects in single LabKey instance
  - Column limits
- Regulatory
  - Data materials agreements
  - Appropriate use guidelines
  - Access restriction

Implementation: Support

- Scheduled weekly support calls
- Demos/User training sessions
- Contracted development support
- Documentation

Implementation: Future

▶ RedCap Integration

▶ Specimen Inventory Management

▶ 'Omics Big Data

▶ Leveraging the LabKey APIs for security management

▶ Ancillary study support

Acknowledgements

Brian Hays,
Sr. Data Engineer, Project Lead

Christopher Shaffer,
Sr. Data Engineer

Dena Garcia, Data
Engineering Administrator

Jennifer Garmeson,
Data Architect

Kenneth Young,
Chief Information
Officer

Laura Gandolfo,
Data Engineer

Steven Fiske
Associate Director,
Data Engineering

Louie Ona,
Sr. Systems
Administrator

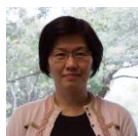Timothy Adams,
Associate Director,
Software Engineering

Sarah Muller,
Assistant Director

Rosalie Holland,
Clinical Research
Administrator

Amy Holbert,
Auditor

Belinda Hsiao, Test
Engineer

Bhushan Pawar,
Graduate Assistant

Maryouri
Avendano,
Graduate Assistant