Colorado TEDDY Newsletter

TEDDY TALK

DATA

By: Marissa Reback

For my 8th grade capstone project I wanted to serve the TEDDY community. I decided to help with creating this TEDDY Talk about data, because I wanted to learn more about how my participation helps the study. I had the chance to interview Dr. Vehik and ask her some of my questions about her work.

Dr. Kendra Vehik (PhD, MPH) works with TEDDY to understand risk factors and triggers of type 1 diabetes, and has written publications about them.

From a Researcher

Q&A with Dr. Kendra Vehik

Q: How do you figure out which data to use to look for what may trigger type 1 diabetes (T1D), and what does the data from the questionnaires tell you about the research?

A: Many of the researchers involved in TEDDY are part of earlier studies, like DAISY (The Diabetes Autoimmunity Study in the Young) in Colorado. From those studies, we learned a lot about diabetes and were able to utilize that information to help us determine the types of data that we wanted to collect. In terms of that data, it really comes down to trying to understand what an important marker (or **biomarker**) is for recognizing those who are at high risk to progress to diabetes and those that actually get the disease.

We designed the questionnaires to gather information on relevant exposures that may be triggers of disease, like infections, life events, diet, allergies, etc. We hope to gather as much information as we can on the questionnaires, so when we do have a specific question we will be able to answer it. An example of this is once we have a question, we identify measures or questions that can help answer the question and then apply different methods to determine who is more likely to have a higher risk than others.

Q: How many cases of type 1 diabetes are not caused by genetics?

A: We're not really sure what causes type 1 diabetes; we do know it's highly genetic, though. Genetics contribute about 50-60% to the risk of type 1 diabetes. It could be particularly one gene family, the **HLA complex**, these genes contribute about 50% of risk. Other non-HLA genes contribute about 10%. What we're doing, with the help of all these **biomarkers** that we've been collecting through our stool and blood samples, is to figure out how different genes interact with some of the exposures we have collected.

Right now, we are able to identify the kids at highest risk primarily using genetics. Our group is trying to understand how those genes behave by focusing on interactions of genes with the exposures we have collected. For example, a child may have a specific virus early in their life that we've shown to be associated with specific genes that may interact and increase their chances of developing type 1 diabetes. Once we can understand this more, we can begin to evaluate all the factors', such as dietary elements or family history, involvement in developing type 1 diabetes.

Q&A continued on back…

TEDDY Stats

Samples/Data Collected by TEDDY:

- 12,234 saliva samples
- 135,691 nasal swab samples
- 184,413 blood (serum) samples
- 198,898 stool (poop) samples
- 33,486 toenail samples
- 56,659 urine samples
- 34,113 water samples
- 92,253 diet records
- 19,893 activity meters
- 187,952 TEDDY book interviews

Did you know?

- The Environmental Determinants of Diabetes in the Young (TEDDY) is turning 18 years old in 2022.
- 424,788 babies were screened at birth worldwide
- 8,676 participants enrolled in the follow-up study.
- There are 6 research centers across 4 countries:

Sweden, Germany, Finland, Colorado, Washington and Georgia/ Florida

WHAT HAPPENS TO THE DATA?

1. Clinician asks TEDDY questions

2. Staff give specific code for consistency across 4 languages

3. Staff enter data into TEDDY database

4. Other staff check data entries

5. TEDDY Data Coordinating Center checks to see if all the data is present

6. All the data is stored in a safe **repository**

7. TEDDY data managers write code to bring together data from all 6 centers

8. Biostatisticians and epidemiologists analyze data and answer the research questions

9. The analysis is shared with scientists who present the findings

Data Word Search:

S	R	E	K	R	A	M	O	I	B
I	N	T	T	U	S	I	B	W	S
I	N	C	J	W	L	G	N	H	A
P	E	T	A	B	Y	T	E	D	Q
A	O	C	E	C	W	G	K	D	H
N	Y	C	U	R	O	O	N	J	L
Z	Z	E	L	U	F	N	U	W	A
K	J	A	U	N	T	A	C	O	X
U	Z	M	Y	S	X	H	C	P	K
Y	R	O	T	I	S	O	P	E	R

Petabyte

Interface

Biomarkers

Repository

HLA

Contact Us

Barbara Davis Center
for Childhood Diabetes

1775 Aurora Court F527
Aurora, Colorado 80045

Phone: 303-724-7577

Email:
teddy.study@ucdenver.edu

Facebook:
facebook.com/teddycolorado

Q&A cont.

Q: What happens to the data once the study ends?
A: I will tell you that the data is not going to disappear. Certain types of data we have to submit regularly to the National Institutes of Health (NIH), who fund us to do this research. The NIH has a huge data **repository**, and we submit all the data to this repository and include dictionaries. These dictionaries are just like you would think a dictionary is, but it is for each variable. If we collect age we have to tell how we collected it, so whoever uses that data can then use it by the right definition. Data is submitted all the time, so when TEDDY officially ends, all of the data will be in this huge repository. It will probably be mined or used for information for a very long time. It is one of the largest-ever dataset that is very comprehensive, from birth to 15 years, looking at children at risk for type 1 diabetes.

Data Vocabulary

Petabyte: 1,000 Terabytes or 500 billion pages of standard printed text.

Interface: a device or program enabling a user to communicate with a computer.

Biomarkers: a measurable substance in an organism that can show signs of disease, infection, or environmental exposure.

Repository: a place where something is stored.

HLA Complex: a complex of genes that make protein.