Journal of Pediatric Psychology, 2016, 373–383
doi: 10.1093/jpepsy/jsv092
Advance Access Publication Date: 27 September 2015
Original Research Article

Predicting Later Study Withdrawal in Participants Active in a Longitudinal Birth Cohort Study for 1 Year: The TEDDY Study

Suzanne Bennett Johnson,[1] PhD, Kristian F. Lynch,[2] PhD, Judith Baxter,[3] MA, Barbro Lernmark,[4] PhD, Roswith Roth,[5] PhD, Tuula Simell,[6] PhD, Laura Smith,[2] PhD, and The TEDDY Study Group[7]

[1]Department of Behavioral Science and Social Medicine, Florida State University, [2]Health Informatics Institute, University of South Florida, [3]Barbara Davis Center for Childhood Diabetes, University of Colorado, [4]Department of Clinical Sciences, Lund University, [5]Helmholtz Center and Institute for Psychology, Graz University, [6]Department of Pediatrics, University of Turku, and [7]see Supplemental Data

All correspondence concerning this article should be addressed to Suzanne Bennett Johnson, PhD, Department of Behavioral Science and Social Medicine, Florida State University College of Medicine, 1115 West Call Street, Tallahassee FL 32306-4300, USA. E-mail: suzanne.johnson@med.fsu.edu

Received April 13, 2015; revisions received August 23, 2015; accepted August 28, 2015

Abstract

Objective To identify predictors of later study withdrawal among participants active in The Environmental Determinants of Diabetes in the Young (TEDDY) for 1 year. **Methods** Multiple logistic regression was used to discriminate 3,042 children active in TEDDY for the first 3 years from 432 children who withdrew in Years 2 or 3. Predictor variables were tested in blocks—demographic, maternal lifestyle behaviors, stress and child illness, maternal reactions to child's increased diabetes risk, in-study behaviors—and a final best model developed. **Results** Few demographic factors predicted study withdrawal. Maternal lifestyle behaviors, accuracy of the mother's risk perception, and in-study behaviors were more important. Frequent child illnesses were associated with greater study retention. **Conclusions** Demographic measures are insufficient predictors of later study withdrawal among those active in a study for at least 1 year; behavioral/psychological factors offer improved prediction and guidance for the development of retention strategies.

Key words: adherence; diabetes; genetics and genetic disorders; longitudinal research; prevention/control; research design and methods.

In 2001, the Human Genome Project reported fully sequencing the human genome. However, few diseases are caused by single genes with complete penetrance. Consequently, the field quickly moved to consideration of gene–environment interactions, which seem to underlie many of our greatest health care challenges—cancer, heart disease, and diabetes (Baccarelli & Ghosh, 2012; Barouki, Gluckman, Grandjean, Hanson, & Heindel, 2012; Brisson, Alves, & Pombo-de-Oliveira, 2015; Nadeem et al., 2015; Sharafeldin et al., 2015). For many diseases, early or lifelong exposures may be critical to the development or prevention of disease. Consequently, large longitudinal birth cohort studies are being conducted in Europe and the United States (e.g., The Environmental Health Risks in European Birth Cohorts Project, the National Children's Study) in an attempt to elucidate these gene–environment interactions (Coughlin, 2014).

Type 1 diabetes is one of the most common chronic diseases of childhood and is increasing worldwide (DIAMOND Project Group, 2006; Patterson, Dahlquist, Gyurus, Soltesz, & EURODIAB Study Group, 2009; SEARCH for Diabetes in Youth Study Group, 2006). The disease has long been known to

have a genetic component, and recent scientific advances have identified the human leukocyte antigen (HLA) locus as the critical area for identifying genetic markers for Type 1 diabetes. Nevertheless, most children with high-risk genes for Type 1 diabetes never develop the disease, and only half of monozygotic twins are concordant for the disease (Johnson, 2011; Metcalfe et al., 2001). Consequently, it is widely believed that environmental exposures play a critical role in disease onset among genetically at-risk children.

Funded by the National Institutes of Health, the Environmental Determinants of Diabetes in the Young (TEDDY) study seeks to identify possible environmental triggers of Type 1 diabetes mellitus in children at increased genetic risk for the disease (The TEDDY Study Group, 2007). Over 420,000 newborns were screened for HLA-conferred Type 1 diabetes genetic risk in sites across Finland, Germany, Sweden, and the USA; 21,589 were identified as HLA-eligible and 8,676 joined the TEDDY study. Most participants (89%) came from the general population, with no first-degree Type 1 diabetes relative.

This longitudinal natural history study will follow these genetically at-risk children for up to 15 years. Because both the identification of TEDDY eligible infants and their participation in TEDDY are time-consuming and expensive for the investigators, families, and the funders, loss of these valuable participants from TEDDY is a major concern. Significant loss of study participants over time will reduce statistical power, further threatening the investigators' ability to meet the study objectives.

In an effort to address this important issue, we initially sought to identify characteristics of the TEDDY sample—collected at study onset—that would successfully predict withdrawal from TEDDY among general population families during the first year of the study (Johnson et al., 2011). However, longitudinal studies of potential gene–environment interactions, like TEDDY, ask participants to remain in the study for many years. Consequently, concerns about study withdrawal are not limited to the first year of the study, but remain important year after year, as the study continues. We were surprised to learn that few studies have examined predictors of retention in long-term observational studies of newborns; the few studies that do exist focused primarily on demographic variables and none examined a comprehensive set of psychosocial and behavioral predictors (Aylward, Hatcher, Stripp, Gustafson, & Leavitt, 1985; Constantine, Haynes, Spiker, Kendall-Tackett, & Constantine, 1993; Greene, Greenland, Olsen, & North, 2011; Howe, Tilling, Golobardes, & Lawlor, 2013; Kotecha et al., 2010; Turner & le Souef, 2003; Wolke et al., 2009).

While some variables may be associated with both early and later study withdrawal, it is also possible that other factors become important in predicting later withdrawal. Families who stay in the study for more than a year are likely different from those who leave a study in the first year. Further, we have important additional information on families who have been in TEDDY for a full year that may prove useful to understanding who leaves the study and who stays 2 and 3 years after study inception. Perhaps most importantly, many of the behavioral factors collected in the first year of TEDDY are potentially modifiable, offering guidance for future efforts to enhance study retention.

We conceptualized possible predictors of late withdrawal in blocks, from those most removed from the target behavior—study withdrawal—and least modifiable to those most likely to be related to study withdrawal and potentially modifiable. Our first block included demographic factors that are relatively easy to collect and have been tied to study withdrawal in other studies. Maternal lifestyle behaviors were considered next, as maternal smoking, alcohol use, working outside the home during pregnancy proved predictive of study withdrawal during the first year of TEDDY (Johnson et al., 2011). Our third block of variables included measures of life stress and child illness collected during the first year of TEDDY. We hypothesized that mothers who report more life stress and more illness in their child might find the demands of TEDDY more difficult. Our fourth block included mothers' reactions to the child's increased risk for Type 1 diabetes. We knew from prior work that maternal anxiety about the child's increased risk and the accuracy of her perceptions of the child's risk at study inception were important predictors of study withdrawal in the first year (Johnson et al., 2011). We also knew that maternal anxiety about a child's diabetes risk often declines over time and her risk perception accuracy may decline over time as well (Johnson, 2011). Consequently, her reactions to her child's increased diabetes risk at study inception may not be as important predictors of later withdrawal as her reactions after a year of being in the TEDDY study. In addition, we had measures of maternal depression, not available at study inception. Although no previous long-term observational studies of newborns have examined maternal depression as a predictor of study retention, two short-term intervention studies conducted with mothers of young children found depression to be associated with study dropout (Bigatti, Cronan, Anaya, 2001; Moser, Dracup & Doering, 2000). Our final block included measures of in-study behavior. We expected that mothers who had the most difficulty attending study visits, who had little father support for the study, and who were more dissatisfied with the TEDDY experience would be more likely to leave the study in years 2 or 3. We considered both study compliance and satisfaction as

potentially modifiable. Although we did not consider father support as necessarily modifiable, we did consider the larger concept of social support as a factor that could be modified.

The TEDDY study is unique in that it offers a broad array of psychological and behavioral variables, in addition to demographic variables, that may prove important predictors of study withdrawal as late as 2 and 3 years after study inception. Behavioral and psychological variables have been largely ignored in the few studies that have attempted to examine study withdrawal in long-term birth cohort studies. Unlike demographic variables, which are largely unmodifiable, behavioral and psychological variables may prove important not only for their predictive value but because many of these factors are potentially modifiable, providing direction to retention efforts. This guidance is important not only for TEDDY but for other pediatric multiyear longitudinal studies that face similar retention challenges.

Methods

The TEDDY Study

Parents of TEDDY eligible babies were fully informed of the infant's increased genetic risk as well as the demands of study participation; all infants were required to join the study before 4.5 months of age. The TEDDY protocol is demanding with study visits every 3 months during the first 4 years of the child's life and every 6 months thereafter. Clinic visits involve blood draws and other biologic sample collection. Blood samples are regularly tested for the development of autoantibodies, indicating possible progression toward Type 1 diabetes. Parents are informed of all test results and are asked to keep detailed records of the child's diet, illnesses, life stresses, and other environmental exposures between study visits. The TEDDY study was approved by relevant institutional review board at each TEDDY site (The TEDDY Study Group, 2007).

Study Sample

We first identified general population families who were active in TEDDY during the first year of the TEDDY study and who were at least 39 months of age as of June 30, 2011 ($n = 4{,}138$). We excluded children with a first-degree Type 1 diabetes relative because few of these children leave TEDDY; only 22 of 458 children with a Type 1 diabetes relative withdrew in years 2–3 of the TEDDY study. We also excluded any child who developed Type 1 diabetes during the second and third year of TEDDY ($n = 54$). The remaining study sample was placed into three groups: Active—defined as currently in the TEDDY study with a visit within the last 6 months ($n = 3{,}042$); Late Withdrawal—withdrew from the TEDDY study between 15 months and 39 months of age ($n = 432$); and Inactive—listed as active in the study database but with no data collected in the last 6 months ($n = 152$). This analysis focused on a comparison of the Actives and the Late Withdrawals. Inactives were removed from the analysis because their status was unclear. We also examined whether children who became autoantibody positive during the first year of TEDDY were more or less likely to stay in TEDDY during Years 2–3; no significant effect was found, so these children were retained in the current analysis. None of the children included in this analysis received the retention intervention developed from our prior studies (Johnson et al., 2011, 2014), which was initiated with families joining TEDDY in January of 2009 or later; all of the families used in this analysis joined TEDDY before March 1, 2008.

Predictors of Study Withdrawal

We organized possible predictors of study withdrawal into five blocks, from those most removed from the target behavior of interest—study withdrawal—and least modifiable to those most likely to be related to study withdrawal and potentially modifiable.

Block 1: Demographic

Demographic information was obtained from maternal interviews and questionnaires and included the following: country of residence; child gender; child ethnic-minority status (United States: the TEDDY child's mother's first language is not English or the mother was not born in the United States or the child is a member of an ethnic minority group—yes/no; Europe: the child's mother's first language or country of birth is other than that of the TEDDY country in which the child resides—yes/no); whether the TEDDY child is an only child (yes/no); maternal age at the child's birth; maternal marital status (married or living together vs. single parent); maternal education (primary education or high school, trade school or some college, graduated from college); and crowding (number of persons in the household divided by the number of rooms in the house). Because the crowding variable was skewed, it was rescored to normalize the distribution ($0 = 0$–0.49; $1 = 0.50$–0.59; $2 = 0.60$–0.75; $3 = 0.76$–1.00; $4 = >1.00$).[1]

Block 2: Maternal Lifestyle Behaviors

Information on maternal lifestyle behaviors during pregnancy and during the first year of TEDDY was obtained from interviews and questionnaires and included the following: smoking during pregnancy or

1 In follow-up analyses, we reran all statistical models using standard scores and raw scores; there was no change in the results.

during the first year of TEDDY (yes/no); reducing work hours or not working at all during pregnancy (yes/no); complete abstinence from alcohol during the last trimester of the mother's pregnancy (yes/no); working outside of the home in the first year of TEDDY (yes/no).

Block 3: Stress and Child Illnesses

Every 3 months, the TEDDY child's mother was interviewed about whether a serious negative life event (illness, injury, hospitalization, death, divorce, or violence) had occurred in the past 3 months. We totaled the number of serious life events across all data collection points in the first year of TEDDY; because the data were bimodal, we created a yes/no variable as to whether a serious negative life event had occurred any time during the first year of TEDDY. Every 3 months, we also collected information on any illnesses the child had during the preceding 3 months. These data were summed across all data collection points in the first year of TEDDY. Because these data were skewed, the data were rescored to create a more normal distribution (1 = 0–2 illnesses; 2 = 3–4 illnesses; 3 = 5–7 illnesses; 4 = 8–10 illnesses; 5 = >10 illnesses).[2]

Block 4: Maternal Reactions to the TEDDY Child's Type 1 Diabetes Risk

In this block, we included two measures specifically addressing the child's increased diabetes risk: maternal risk perception accuracy and maternal anxiety about the child's increased diabetes risk. We also included two measures of maternal depression.

Maternal risk perception accuracy (accurate—indicating that the mother perceived the TEDDY child's Type 1 diabetes risk as higher or much higher than other children's risk; underestimate—indicating that the mother perceived the child's risk as the same, somewhat lower, or much lower than other children's Type 1 diabetes risk) was assessed by questionnaire at study inception and at study visits scheduled when the child was 6 and 15 months of age.

Maternal anxiety about the child's Type 1 diabetes risk was assessed using a six-item measure adapted from the state component of the State-Trait Anxiety Inventory (Hood, Johnson, Baughcum, She, & Schatz, 2006; Johnson et al., 2011; Speilberger, Gorsuch, & Luschene, 1970) and given at the first TEDDY visit and at two subsequent visits when the child was 6 and 15 months of age. The coefficient alphas for maternal State Anxiety Inventory scores in this sample were excellent ($\alpha = .895$ at first study visit; $\alpha = .901$ at the 6-month study visit; $\alpha = .904$ at the 15-month study visit).

Preliminary analyses indicated that risk perception accuracy was highly correlated across time; the same was true for the State Anxiety Inventory scores. Consequently, we elected to use the 15-month measures in our predictive models. In cases where the 15-month measure was missing, we used the 6-month measure.

Maternal depression was assessed using the Edinburgh Postnatal Depression scale (Cox, Holden, & Sagovsky, 1987) collected at the child's 6-month study visit (coefficient $\alpha = .844$)[3] and the depression subscale of Bradley's Well Being scale (Bradley & Lewis, 1990) administered at the child's 15-month study visit (coefficient $\alpha = .664$). Responses on this subscale were not normally distributed and were rescored using a 3-point scale to normalize the distribution (0 = 0–1; 1 = 2–4; 2 = ≥5).[4]

Block 5: In-Study Behaviors

Because TEDDY participants were recruited over several years, we included the year of recruitment as an in-study behavior variable given that recruitment behaviors by study staff may have changed over the years (Year 1: September 2004–August 2005; Year 2: September 2005–August 2006; Year 3: September 2006–August 2007; Year 4: September 2007–February 2008). Other in-study behaviors included: TEDDY father participation, study compliance, and maternal study satisfaction.

The TEDDY child's father was asked to complete a brief survey at study inception, and at TEDDY visits scheduled when the child was 6 months and 15 months of age. We examined how many fathers completed zero, one, two, or all three surveys during the first year of TEDDY. The distribution was bimodal. Consequently, we scored this variable 0/1 as to whether the father had completed all three surveys.

Study compliance was measured by the number of missed clinic visits and missed blood draws in the first year of TEDDY. Each child was placed into one of three groups: completed all visits and all blood draws, completed all visits but missed one or more blood draws, missed one or more clinic visits.

Maternal study satisfaction was assessed by questionnaire at TEDDY study visits scheduled when the child was 6 and 15 months of age. Mothers were asked: Overall, how do you feel about having your child participate in the TEDDY study? (scored:

2 In follow-up analyses, we reran all statistical models using standard scores and raw scores; there was no change in the results.

3 In a follow-up analysis, we replaced the Edinburgh Postnatal Depression score with a categorical variable indicating whether the score was above the clinical cutoff (≥13); there was no change in the results.

4 In follow-up analyses, we reran all statistical models using standard scores and raw scores; there was no change in the results.

0 = *like it a lot*, 1 = *like it a little*, 2 = *it is ok or dislike it*); Do you think your child's participation in TEDDY was a good decision? (scored: 0 = *a great decision*, 1 = *a good decision*, 2 = *an ok decision or bad decision*); Would you recommend the TEDDY study to a friend? (scored: 0 = *yes*, 1 = *maybe*, 2 = *no*). The items were significantly correlated and consequently were summed to create a total satisfaction score. Because the total satisfaction scores were not normally distributed and skewed toward 0, for analysis purposes, we placed each respondent into one of four groups: very satisfied (total score = 0); satisfied (total score = 1 or 2); somewhat satisfied (total score = 3 or 4); neutral/dissatisfied (total score = 5 or 6). Because maternal study satisfaction was highly correlated at 6 and 15 months, we selected the 15-month measure for inclusion in the block. If the 15-month measure was missing, we used the 6-month measure.

Data Analysis

Multiple logistic regression was used to identify significant predictors of withdrawal from TEDDY in Years 2–3 of the study. We considered predictor variables in blocks: demographic, maternal lifestyle behaviors, stress and child illness, maternal reactions to the child's Type 1 diabetes risk, and in-study behaviors. We initially created prediction models for each block of variables. We then retained significant predictors ($p < .10$) from each block in a full model. The full model was further trimmed by eliminating through backward selection any predictor with a p value of $> .05$. Finally, receiver operating characteristic (ROC) curves were used to evaluate the discriminatory power of the models to identify families who withdrew compared with those who did not. The area under the ROC curve (AUC) was used to quantify the predictive accuracy. An area of 1 represented a perfect prediction; an area of 0.5 indicated random chance. We calculated ROC curves for the model that included only the significant demographic variables in Block 1 and the final model that included behavioral and psychological variables. Analyses were conducted using the Statistical Analysis Software (Version 9.4, SAS Institute, Cary NC), and the figure was created using GraphPad PRISM 5.03 (GraphPad Software Inc., San Diego, CA).

Results

Table I provides the results of the multiple logistic regression for each of the five blocks of variables. For categorical measures, the sample size and percent of participants withdrawing from TEDDY are provided. For continuous variables, the mean and standard deviation are provided. For all variables, the odds ratio, 95% confidence interval (CI), and p value are provided. Sample size differs slightly across blocks because only subjects with no missing data for that block were retained.

In the Demographic block, only-child status, younger maternal age at the child's birth, less maternal education, and household crowding were all associated with leaving the TEDDY study. In the Maternal Lifestyle Behavior block, not working or reducing work hours during pregnancy and smoking were associated with study withdrawal. In the Stress and Child Illness block, child illness was associated with study withdrawal but in the opposite direction expected; children with more illness were more likely to remain in the TEDDY study. In the Maternal Reactions to the Child's Risk for Type 1 Diabetes block, maternal risk perception accuracy was predictive; mothers who underestimated their child's risk were more likely to leave the TEDDY study. All of the variables in the In-Study Behavior block were significant predictors of study withdrawal. Participants recruited in the earliest years of TEDDY, those whose fathers did not complete all study surveys, those who missed blood draws or study visits, and mothers least satisfied with the TEDDY study were more likely to withdraw.

Table II provides the final best predictive model, incorporating variables from all five blocks. All of the variables that were statistically significant in their own block were retained in the final model. Maternal education was dropped because it did not meet the $p < .05$ criteria.

Figure 1 provides the ROC curves for the null model (no predictors), the demographics-only model (only-child status, maternal age, maternal education, crowding), and the final model depicted in Table II. For the demographics-only model, the AUC was 0.61, 95% CI = 0.58–0.64. Prediction substantially improved when behavioral and psychological variables were added to the model, AUC = 0.73, 95% CI = 0.71–0.76.

Discussion

In this sample of participants who had been active in TEDDY for a full year, two demographic variables—only-child status and younger maternal age—were associated with leaving the TEDDY study in Year 2 or 3. Young maternal age was an important predictor of study withdrawal in Year 1 of TEDDY (Johnson et al., 2011) and has been associated with loss to follow-up in a prior long-term birth cohort observational study (Greene et al., 2011). Younger moms are less experienced and may find study demands more challenging. Only-child status appears to be rarely used in studies of dropout. Wolke et al. (2009) found large family size to be predictive of study dropout but did not assess the impact of only-child families. Turner and le Souef (2003) did not assess only-child status per

Table I. Predictors of TEDDY Study Withdrawal in Years 2–3 by Block: Demographic, Maternal Lifestyle Behaviors, Stress and Child Illness, Maternal Reactions to Child's Type 1 Diabetes Risk, and In-Study Behaviors

Blocks	Factors	Total	Withdrawals	Odds ratio for withdrawal as estimated from multivariate logistic regression within each block		
		N or mean (SD)	%	OR	95% CI	*p*-value
Demographic	Country of residence					
	Finland	845	11.7	1.00	Ref	
	Sweden	1,230	12.2	1.08	0.82–1.44	
	Germany	112	18.8	1.70	1.00–2.89	
	United States	1,141	11.7	1.12	0.82–1.53	0.28
	Gender of child					
	Boy	1,716	11.9	1.00	Ref	
	Girl	1,612	12.3	1.04	0.84–1.29	0.70
	Child ethnic/minority status					
	No	2,880	11.6	1.00	Ref	
	Yes	448	15.4	1.17	0.85–1.61	0.33
	Only child					
	No	1,871	11.2	1.00	Ref	
	Yes	1,457	13.2	1.30	1.03–1.65	0.03
	Maternal age at child's birth	30.7 (5.1)		0.97	0.95–1.00	0.03
	Parents married or living together					
	No	140	20.0	1.45	0.93–2.27	0.11
	Yes	3,188	11.8	1.00	Ref	
	Mother's education					
	Primary education	661	15.7	1.33	0.99–1.79	
	Trade school/ some college	790	14.8	1.32	1.02–1.72	
	Graduated college	1,877	9.7	1.00	Ref	0.06
	Crowding (persons/room)[a]	2.02 (1.18)		1.21	1.09–1.34	0.0004
Maternal Lifestyle Behaviors	Working during pregnancy					
	Reduced work/did not work at all	1,657	14.3	1.49	1.20–1.86	0.0004
	Worked all trimesters	1,692	9.9	1.00	Ref	
	Alcohol consumption during pregnancy					
	Total abstinence in third trimester	2,721	12.4	1.13	0.86–1.50	0.38
	Occasional drink in third trimester	628	10.7	1.00	Ref	

(*continued*)

Table I. Continued

Blocks	Factors	Total	Withdrawals	Odds ratio for withdrawal as estimated from multivariate logistic regression within each block		
		N or mean (SD)	%	OR	95% CI	*p*-value
	Smoking (pregnancy or during first year of TEDDY)					
	No	2,873	11.2	1.00	Ref	
	Yes	476	17.2	1.55	1.19–2.02	0.001
	Working outside the home in the first year of TEDDY					
	No	2,218	12.1	1.13	0.90–1.42	0.31
	Yes	1,131	12.1	1.00	Ref	
Stress and Child Illness	Serious negative life event in the first year of TEDDY					
	No	1,805	12.4	1.00	Ref	
	Yes	1,651	12.5	1.06	0.86–1.29	0.61
	Number of child illnesses[b] in the first year of TEDDY	2.1 (1.2)		0.78	0.72–0.86	<0.0001
Maternal	Risk perception accuracy (15mo)					
Reactions to	Underestimate	1,279	14.5	1.84	1.46–2.31	<0.0001
Child's Type 1	Accurate	1,851	8.5	1.00	Ref	
Diabetes Risk	State Anxiety Inventory score (15 mo)	33.6 (9.4)		1.01	0.99–1.02	0.30
	Edinburgh Postnatal Depression score (6 mo)	6.2 (4.3)		0.99	0.97–1.02	0.66
	Bradley Depression subscale (15 mo)[c]	2.24 (1.13)		1.07	0.96–1.19	0.24
In-Study Behaviors	TEDDY recruitment cohort					
	Year 1	699	13.6	2.04	1.37–3.02	
	Year 2	1,018	13.3	1.99	1.37–2.89	
	Year 3	1,189	12.5	1.60	1.11–2.32	
	Year 4	516	8.3	1.00	Ref	0.001
	Father's participation in first year of TEDDY					
	Completed all 3 surveys	2,647	10.2	1.00	Ref	
	Completed 0, 1, or 2 surveys	775	19.6	1.67	1.31–2.13	<0.0001
	Study compliance in first year of TEDDY					
	No missed study visits or blood draws	2,706	9.4	1.00	Ref	
	Missed $\geq$ 1 blood draws but not a study visit	545	19.4	2.15	1.66–2.78	
	Missed $\geq$ 1 study visits	171	35.7	4.14	2.85–6.02	<0.0001
	Maternal study satisfaction (15 months)					
	Very satisfied (0)	1,496	7.9	1.00	Ref	
	Somewhat satisfied (1–2)	939	13.2	1.70	1.29–2.23	
	Somewhat dissatisfied (3–4)	698	16.2	2.14	1.61–2.84	
	Very dissatisfied (5–6)	289	23.2	3.52	2.50–4.96	<0.0001

[a]Because the crowding variable was skewed, the data were rescored to normalize the distribution (0 = 0–0.49; 1 = 0.50–0.59; 2 = 0.60–0.75; 3 = 0.76–1.00; 4 = > 1.00)
[b]Because the illness data were skewed, the data were rescored to normalize the distribution (1 = 0–2 illnesses; 2 = 3–4 illnesses; 3 = 5–7 illnesses; 4 = 8–10 illnesses; 5 = > 10 illnesses).
[c]Because the Bradley depression subscale scores were not normally distributed, they were rescored to normalize the distribution (0 = 0–1; 1 = 2–4; 2 = $\geq$ 5).

Table II. Final Logistic Regression Model Predicting TEDDY Study Withdrawal in Years 2–3

Factors	*N* or Mean (SD)	OR	95%CI	*p*-value
Only child				
No	1,862	1.00	Ref	
Yes	1,471	1.28	1.01–1.64	0.04
Maternal age (at child's birth)	30.7 (5.0)	0.98	0.95–1.00	0.03
Household crowding (rooms/person)[a]	2.0 (1.2)	1.16	1.04–1.28	0.007
Worked during pregnancy				
Reduced hours or did not work at all	1,648	1.37	1.09–1.72	
Worked all trimesters	1,685	1.00	Ref	0.007
Maternal smoking				
No	2,862	1.00	Ref	
Yes	471	1.39	1.04–1.86	0.03
Number of child illnesses[b]	2.1 (1.2)	0.88	0.80–0.97	0.01
Maternal risk perception accuracy				
Underestimate	1,364	1.43	1.15–1.79	
Accurate	1,969	1.00	Ref	0.002
TEDDY recruitment cohort				
Year 1	687	1.98	1.30–3.00	
Year 2	990	2.12	1.43–3.13	
Year 3	1,155	1.69	1.15–2.50	
Year 4	501	1.00	Ref	0.002
Father's participation in first year of TEDDY				
Completed all 3 surveys	2,610	1.00	Ref	
Completed 0, 1, or 2 surveys	723	1.58	1.21–2.04	0.0006
Study compliance in first year of TEDDY				
No missed study visits or blood draws 000	2,670	1.00	Ref	
Missed $\geq$ 1 blood draw but not a study visit	516	1.98	1.51–2.60	
Missed $\geq$ 1 study visit	147	3.30	2.18–4.99	<0.0001
Maternal study satisfaction (15 months)				
Very satisfied	1,469	1.00	Ref	
Satisfied	909	1.89	1.42–2.51	
Somewhat satisfied	677	2.45	1.82–3.29	
Neutral/dissatisfied	278	4.19	2.93–6.00	<0.0001

[a]Because the crowding variable was skewed, the data were rescored to normalize the distribution (1 = 0–0.49; 2 = 0.50–0.59; 3 = 0.60–0.75; 4 = 0.76–1.00; 5 = > 1.00).

[b]Because the illness data were skewed, the data were rescored to normalize the distribution (1 = 0–2 illnesses; 2 = 3–4 illnesses; 3 = 5–7 illnesses; 4 = 8–10 illnesses; 5 = > 10 illnesses).

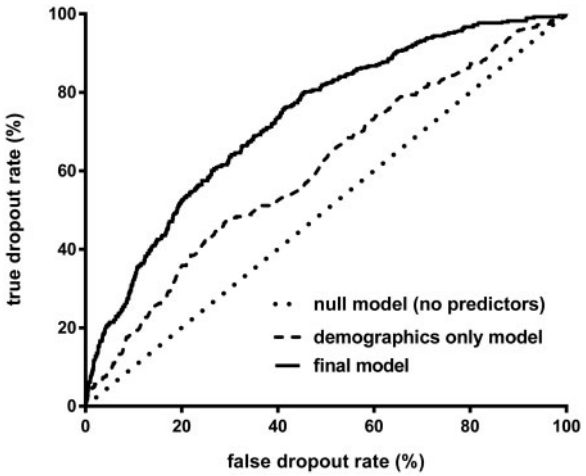

Figure 1. ROC curves for null, demographics only, and final model.

se, but did find children who were first born or were the only child in the study were more likely to leave their study. Dickinson et al. (2012) failed to find an association between only-child status and dropout in their 5-year follow-up study of older children with cerebral palsy. Only-child status may be particularly important for longitudinal studies of newborns, as mothers are often more stressed with their first child (Zinardo et al., 2009). This may translate into greater difficulty dealing with the demands of the TEDDY protocol. Prior pediatric cohort studies have found maternal education to be an important predictor of study retention (Constantine et al., 1993; Howe et al., 2013; Turner and le Souef, 2003; Wolke et al., 2009). We found it to be associated with dropout in the current study, but the association was so weak that it was not retained in the final model. Its importance may have been diminished by the fact that this was a highly educated sample; 56% of participating mothers were college graduates. Crowding was the demographic variable most strongly associated with leaving TEDDY; families living in more crowded households were more likely to withdraw. Crowding is often linked to low socioeconomic status, which may underlie this association (Curtis, Corman, Noonan, & Reichman, 2010), as low socioeconomic status has

been linked to study attrition in prior long-term observational birth cohort studies (Greene et al., 2011; Howe et al., 2013; Kotecha et al., 2010; Wolke et al., 2009).

Two maternal lifestyle behaviors were predictive of study dropout in Year 2 or 3: maternal smoking and not working or reducing work hours during pregnancy. Both variables were also predictive of TEDDY dropout in Year 1 (Johnson et al., 2011). Parental smoking has been associated with study attrition in several previous long-term birth cohort studies (Greene et al., 2011; Howe et al., 2013; Kotecha et al., 2010; Turner & le Souef, 2003; Wolke et al., 2009). We have speculated that maternal smokers may be less health conscious and therefore more likely to leave the TEDDY study. Maternal work status during the first year of TEDDY did not predict study withdrawal in Years 2–3. However, maternal work status during pregnancy did predict; if a mother worked all three trimesters during her pregnancy, she was more likely to remain in TEDDY in Year 1 and in Years 2–3. We can only speculate as to what may underlie this highly reliable association; perhaps mothers who choose to work throughout their pregnancy have strong coping skills and are able to respond to life's demands—including pregnancy and the demands of the TEDDY protocol—with greater ease than those who choose to reduce their work hours or not work at all when they are pregnant.

We did not find that stress was associated with study withdrawal, and the frequency of child illnesses was associated with withdrawal in a direction opposite to what was expected; children with more reported illnesses in the first year of TEDDY were more likely to remain in the study. Because all children in TEDDY are at increased risk for Type 1 diabetes, participating mothers may view frequent illnesses as a sign of impending diabetes and as a consequence remain in TEDDY. Alternatively, mothers who are concerned about their child's Type 1 diabetes risk may become hypervigilant about the child's health, reporting more child illnesses while remaining in the TEDDY study. Because we have biomarkers of the child's illness episodes from stool and saliva samples, in subsequent studies we will be able to determine whether these mothers are overreporting their child's illnesses or the children are, in fact, ill more often.

We had several measures of maternal reactions to the child's increased diabetes risk, including measures of risk perception, anxiety, and depression. Only risk perception proved to be a significant predictor of study withdrawal. Mothers who underestimated the child's risk for Type 1 diabetes were more likely to leave the TEDDY study both in Year 1 (Johnson et al., 2011) and in Years 2 and 3. This finding is important not only because of its reliability but because so many mothers underestimate the child's risk; 39% of mothers underestimated the child's risk at study inception, 40% at 6 months and 41% at 15 months. Risk underestimation occurred despite the fact that families were informed of the child's increased risk both orally and in writing when they were recruited into TEDDY. Other studies have reported high rates of risk underestimation among parents of children genetically at risk for Type 1 diabetes (Carmichael, et al., 2003; Hood et al., 2006). Despite our best efforts to fully inform parents of their child's diabetes risk, many parents do not completely understand or accept this information. This, in turn, has implications for study retention. In this study, 10% of those with accurate risk perceptions and who were active in TEDDY for 1 year left the study in Year 2 or 3 compared with 15% of those who underestimated the child's risk. We did not find maternal depression to be a predictor of study dropout in Years 2–3. Maternal depression may be a better predictor of leaving a study early, in the first few months of a study protocol (Bigatti et al., 2001; Moser et al., 2000).

Consistent with our hypothesis, In-Study Behaviors proved to be the most important predictors of study withdrawal. All variables in this block were significant and the block as a whole showed the strongest association to study dropout. Families who were recruited into TEDDY in the fourth year of the study had lower Year 2–3 withdrawal rates than families recruited in earlier years. We suspect that the TEDDY staff may have developed greater skill at relating to TEDDY families, leading to greater study retention. It is also possible that study staff became more selective in their recruitment efforts over time. The analysis of father participation as a predictor of study withdrawal is unique and its reliable association with study withdrawal in both Year 1 (Johnson et al., 2011) and Years 2–3 of TEDDY suggests this is a variable that needs greater attention in pediatric longitudinal studies. Although 77% of fathers of children who were in TEDDY for at least 1 year completed all study surveys, the withdrawal rate for families whose fathers failed to complete all of the surveys was nearly twice as high (19%) as the withdrawal rate for families whose fathers completed all surveys (10%). Compliance with study visits during Year 1 of TEDDY was also a predictor of study withdrawal in Years 2–3. Most TEDDY parents (81%) who remained in TEDDY for 1 year did not miss a single visit or blood draw, and only 9% of these families left TEDDY in Years 2–3. In contrast, 19% of those who missed a blood draw and 36% of those who missed a study visit left TEDDY in Years 2–3. It appears that measures of study compliance are good indicators of those families who will ultimately leave a study. A long-term study of quality of life of children with cerebral palsy reported similar results (Dickinson et al., 2012). Maternal dissatisfaction with the study also

proved to predict leaving TEDDY in Years 2–3. Only 8% of families exhibiting the highest maternal satisfaction left TEDDY compared with 23% of those expressing the most dissatisfaction.

Although studies of participant dropout have historically focused on demographic variables, the results of this study's multiple logistic regression and ROC curve analyses suggest that behavioral and psychological variables are important and underutilized predictors. The ultimate goal of any effort to predict study withdrawal is to develop strategies to enhance study retention. Most demographic predictors are not amenable to change. In contrast, many behavioral and psychological variables are amenable to change. In this study, several significant predictors of study withdrawal—risk underestimation, study compliance, and study satisfaction—are potentially modifiable. There is an evidence-based literature on risk communication (Trevena et al., 2013) that could be used to develop new strategies to improve parents' understanding of their child's Type 1 diabetes risk. Tailored interventions could be developed targeting families who are missing study visits. Structures could be put in place to assess and address any dissatisfaction families may be experiencing in TEDDY. Other important predictors—cohort year and father participation—could be further explored to discern potential avenues for intervention. For example, retention efforts by study staff may be examined to discern those that are most effective. While father participation per se may not be easily modifiable, efforts to increase social support in families with poor father participation may be one strategy. Of course, whether such efforts result in improved retention rates must be subject to empirical scrutiny.

Study limitations include its focus on a highly educated sample of families from the general population whose infants were identified as genetically at-risk for Type 1 diabetes. Consequently, the results may not be applicable to other populations. Although the inclusion of behavioral and psychological variables markedly improved the prediction model over one limited to demographic variables, the ROC curve for the final model clearly indicates that other unmeasured variables remain to be identified to further improve predictive power. Further, this study focused on withdrawal in Year 2 or 3; the study is projected to continue until the child is 15 years of age. Study retention will continue to be a major concern, and the predictors of retention and withdrawal may well change as the child becomes older.

Nevertheless, the current study's findings will guide retention efforts in TEDDY and may prove useful to others conducting pediatric multiyear longitudinal birth cohort studies. Long-term studies of gene–environment interactions present formidable challenges in both identification of at-risk children at a young age and intensive monitoring of environmental exposures. Retention of participants is critical if study investigators hope to meet their scientific objectives. However, the literature on factors associated with study withdrawal and retention is sparse, has rarely focused on studies with long follow-ups and has been primarily limited to demographic variables. Our findings suggest that the inclusion of psychological and behavioral measures may improve our scientific understanding of why participants join studies and stay or leave them. Perhaps equally important, such variables offer possible avenues for protocol modification to improve study retention.

Supplementary Data

Supplementary data can be found at: http://www.jpepsy.oxfordjournals.org/. TEDDY Study Group members are provided at (list web link).

Funding

This study was funded by U01 DK63829, U01 DK63861, U01 DK63821, U01 DK63865, U01 DK63863, U01 DK63836, U01 DK63790, UC4 DK63829, UC4 DK63861, UC4 DK63821, UC4 DK63865, UC4 DK63863, UC4 DK63836, UC4 DK95300, and UC4 DK100238, and Contract No. HHSN267200700014C from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases (NIAID), National Institute of Child Health and Human Development (NICHD), National Institute of Environmental Health Sciences (NIEHS), Juvenile Diabetes Research Foundation (JDRF), and Centers for Disease Control and Prevention (CDC). This work was supported in part by the NIH/NCATS Clinical and Translational Science Awards to the University of Florida (UL1 TR000064) and the University of Colorado (UL1 TR001082).

Conflicts of interest: None declared.

References

Aylward, G., Hatcher, R., Stripp, B., Gustafson, N. & Leavitt, L. (1985). Who goes and who stays: subject loss in a multicenter longitudinal follow-up study. *Journal of Developmental and Behavioral Pediatrics*, *6*, 3–8.

Baccarelli, A., & Ghosh, S. (2012). Environmental exposures, epigenetics, and cardiovascular disease. *Current Opinion in Clinical Nutrition and Metabolic Care*, *15*, 323–329.

Barouki, R., Gluckman, P. D., Grandjean, P., Hanson, M., & Heindel, J. J. (2012). Developmental origins of non-communicable disease: Implications for research and public health. *Environmental Health*, *11*, 42. doi: 10.1186/1476-069X-11-42

Bigatti, S., Cronan, T., & Anaya, A. (2001). The effects of maternal depression on the efficacy of a literacy intervention program. *Child Psychiatry and Human Development*, *32*, 147–162.

Bradley, C., & Lewis, K. S. (1990). Measures of psychological well-being and treatment satisfaction developed from responses of people with tablet-treated diabetes. *Diabetes Medecine*, *7*, 445–451.

Brisson, G. D., Alves, L. R., & Pombo-de-Oliveria, M. S. (2015). Genetic susceptibility in childhood acute leukaemias: A systematic review. *Ecancermedialscience*, *9*, 539. doi: 10.3332/ecancer.2015.539

Carmichael, S., Johnson, S. B., Baughcum, A., North, K., Hopkins, D., Dukes, M. G., … Schatz, D. (2003). Prospective assessment in newborns of diabetes autoimmunity (PANDA): maternal understanding of infant diabetes risk. *Genetic Medicine*, *5*, 77–83.

Constantine, W., Haynes, C., Spiker, D., Kendall-Tackett, K. & Constantine, N. (1993). Recruitment and retention in a clinical trial for low birth weight, premature infants. *Journal of Developmental and Behavioral Pediatrics*, *14*, 1–7.

Coughlin, S. (2014). Toward a road map for global-omics: A primer on –omic technologies. *American Journal of Epidemiology*, *180*, 1188–1196.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry*, *150*, 782–786.

Curtis, M. A., Corman, H., Noonan, K., & Reichman, N. E. (2010). Effects of child health on housing in the urban U.S. *Social Science and Medicine*, *71*, 2049–2056.

DIAMOND Project Group. (2006). Incidence and trends of childhood type 1 diabetes worldwide 1990-1999. *Diabetic Medicine*, *23*, 857–866.

Dickinson, H. O., Rapp, M., Arnaud, C., Carlsson, M., Colver, A. F., Fauconnier, J., … Parkinson, K. (2012). Predictors of drop-out in a multi-centre longitudinal study of participation and quality of life of children with cerebral palsy. *BioMed Central Research Notes*, *5*, 300. doi:10.1186/1756-0500-5-300

Greene. N., Greenland, S., Olsen, J., & Norh, E. (2011). Estimating bias from loss to follow-up in the Danish National Birth Cohort. *Epidemiology*, *22*, 815–822.

Hood, K., Johnson, S. B., Baughcum, A. E., She, J., & Schatz, D. A. (2006). Maternal understanding of infant diabetes risk: The differential effects of maternal anxiety and depression. *Genetic Medicine*, *8*, 665–670.

Howe, L., Tilling, K., Golobardes, B., & Lawlor, D. (2013). Loss to follow-up in cohort studies: Bias in estimates of socioeconomic inequalities. *Epidemiology*, *24*, 1–9.

Johnson, S. B. (2011). Psychological impact of screening and prediction in type 1 diabetes. *Current Diabetes Reports*, *11*, 454–459.

Johnson, S. B., Lee, H., Baxter, J., Lernmark, B., Roth, R., & Simmel, T. for the TEDDY Group. (2011). The Environmental Determinants of Diabetes in the Young (TEDDY) Study: predictors of early withdrawal among participants with no family history of type 1 diabetes. *Pediatric Diabetes*, *12*, 165–171.

Johnson, S. B., Lynch, K., Lee, H., Baxter, J., Lernmark, B., Roth, R., … Smith, L. for the TEDDY Study Group. (2014). At high risk for early withdrawal: Using a cumulative risk model to increase retention in the first year of the TEDDY study. *Journal of Clinical Epidemiology*, *67*, 609–611.

Kotecha, S., Watkins, J., Heron, J., Henderson, J., Dunstan, F., & Kotecha, S. (2010). Spirometric lung function in school-age children. *American Journal of Respiratory and Critical Care Medicine*, *181*, 969–974.

Metcalfe, K. A., Hitman, G. A., Rowe, R. E., Hawa, M., Huang, X., Stewart, T., & Leslie, R. D. G. (2001). Concordance for type 1 diabetes in identical twins is affected by insulin genotype. *Diabetes Care*, *24*, 838–842.

Moser, D., Dracup, K., & Doering, L. (2000). Factors differentiating dropouts from completers in a longitudinal, multicenter clinical trial. *Nursing Research*, *49*, 109–116.

Nadeem, A., Mumtaz, S., Naveed, A. K., Aslam, M., Siddiqui, A., Lodhi, G. M., & Ahmed, R. (2015) Gene-gene, gene-environment, gene-nutrient interactions and single nucleotide polymorphisms of inflammatory cytokines. *World Journal of Diabetes*, *6*, 642–647.

Patterson, C. C., Dahlquist, G. G., Gyurus, E., Soltesz, G., & EURODIAB Study Group. (2009). Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: A multicentre prospective registration study. *Lancet*, *373*, 2027–2033.

SEARCH for Diabetes in Youth Study Group. (2006). The burden of diabetes mellitus among US youth: prevalence estimates from the SEARCH for Diabetes in Youth Study. *Pediatrics*, *118*, 1510–1518.

Sharafeldin, N., Slattery, M. L., Liu, Q., Franco-Villalobos, C., Caan, B. J., Potter, J. D., & Yasui, Y. (2015). A candidate-pathway approach to identify gene-environment interactions: Anayses of colon cancer risk and survival. *Journal of the National Cancer Institute*, *107*, 9. doi:10.1093/jnci/djv160

Speilberger, C. D., Gorsuch, R. L., & Lushene, R. (1970). *Test manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

The TEDDY Study Group. (2007). The Environmental Determinants of Diabetes in the Young (TEDDY) study: study design. *Pediatric Diabetes*, *8*, 286–298.

Trevena, L., Zikmund-Fisher, B., Edwards, A., Gaissmaier, W., Galesic, M., Han, P., … Woloshin, S. (2013). Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. *BioMed Central Medical Informatics and Decision Making*, *13* (Suppl 2) S7. doi:10.1186/1472-6947-13-S2-S7

Turner, S., & le Souef, P. (2003). Is patient dropout from a longitudinal study of lung function predictable and reversible? *Pediatric Pulmonology*, *35*, 29–33.

Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T., & Lambers, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behavior disorders. *British Journal of Psychiatry*, *195*, 249–256.

Zinardo, V., Gasparetto, S., Giustardi, A., Suppiej, A., Trevisanuto, D., Pascoli, I., & Freato, F. (2009). Impact of anxiety in the puerperium on breast-feeding outcomes: Role of parity. *Journal of Pediatric Gastroenterology and Nutrition*, *49*, 631–634.