CENTER FOR DRUG EVALUATION AND RESEARCH

APPLICATION NUMBER:

761183Orig1s000

STATISTICAL REVIEW(S)

STATISTICAL REVIEW MEMO OF
RESPONSE TO COMPLETE RESPONSE LETTER CLINICAL COMMENT

BLA #: 761183	SDN/eCTD # 59/0059	Stamp date: February 17, 2022
Applicant	Provention Bio, Inc.	
Name of drug	Teplizumab injection	
Indication(s)	Prevention of type 1 diabetes in at-risk individuals	
EDR link	\\CDSESUB1\EVSPROD\bla761183\0059	
Statistical Team	Yu Wang, Yoonhee Kim	
Clinical Team	Lauren Wood Heickman, Michael Nguyen	
Biometrics Division	DBII	
OND Division	DDLO	
Project Manager	Supendeep Dosanjh	

1. Background

In FDA's BLA 761183 *Complete Response Letter* (CRL, issue date July 2, 2021), FDA had the following additional clinical comments (not approvability issues).

> *We were not able to confirm your exploratory analysis of the relationship between ADA (anti-teplizumab drug antibodies) status at Month 3 (positive, negative, or missing) and efficacy in TN-10. We note that your SAS code for TN-10 CSR (memo note: CSR is the abbreviation for Clinical Study Report) Figure 2.5.4.1 contains the relevant code, which utilized SDTM.PC as the source dataset. To facilitate our analysis, please provide the following:*
>
> - *A description of the derivation algorithm for this variable (ADA status),*
> - *A SAS code that is used to generate this variable (ADA status), and*
> - *A SAS dataset that contains the variable (ADA status).*

2. The Applicant's Response – Description of the ADA Status Derivation Algorithm

The applicant's responses to FDA's CRL additional comments (not approvability issues) are contained in *Provention Bio's Responses to the CRL* (eCTD 0059, Module 1.2). In Section 2.1 of the *Responses to the CRL*, the applicant described the ADA status derivation algorithm applied in its generation of EOT (end of text) Figure 2.5.4.1 (Figure 1 reproduced in the review).

> *ADA status at each visit was tested in the following way – First, ADA screening was done; if that result was positive, then confirmation test was carried out. ADA result at a visit was considered positive when confirmative test was positive; otherwise the result was considered negative. If the screening test result was missing, then the result was missing.*

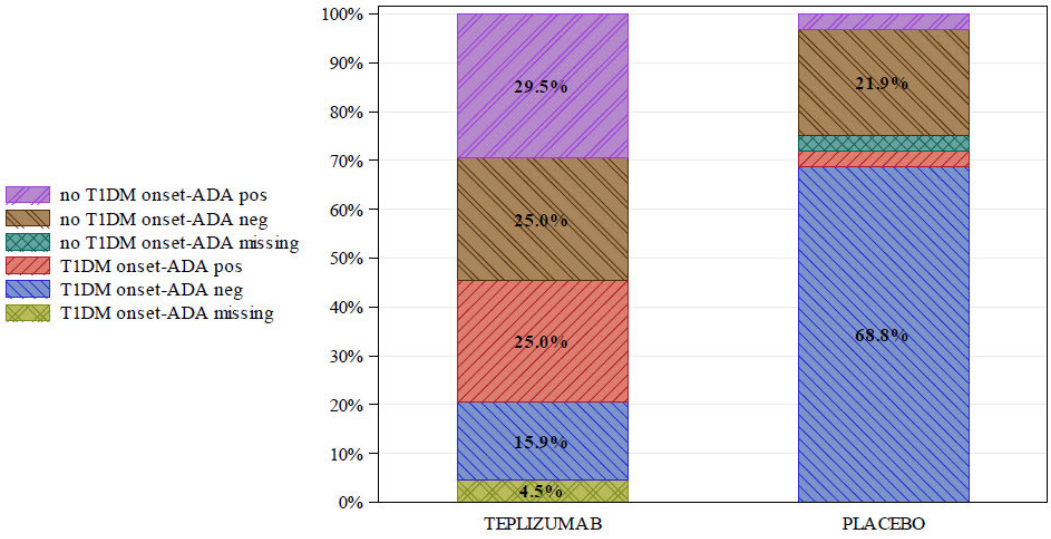
Criteria for diabetes onset (T1DM) were based on glucose testing or the presence of unequivocal hyperglycemia with acute metabolic decompensation (diabetic ketoacidosis).

Final variable is all the combinations of those 2 variables values.

For the figure, we are looking at ADA test done at 3 months after treatment start.

Figure 1. The Applicant's EOT Figure 2.5.4.1

Figure 2.5.4.1 Bar plot - anti-teplizumab response, onset of T1DM and treatment (ITT population)



Note: T1DM - Type 1 Diabetes Mellitus.

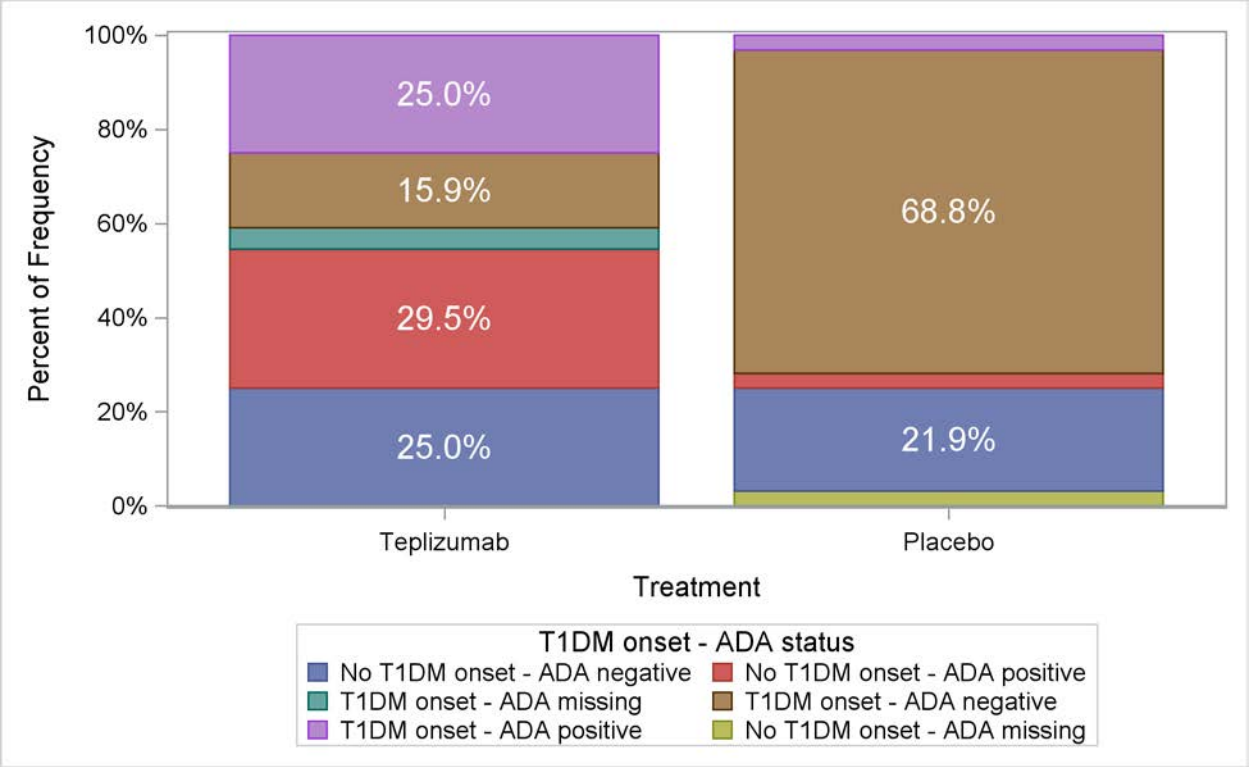
Source: The applicant's PRV-031 TN-10 Figures Final 07SEP2020, Figure 2.5.4.1, quoted in the Applicant's CSR section 11.4.4.3. Association of Anti-teplizumab Antibodies and Development T1D Diagnosis.
Source Link: \\CDSESUB1\evsprod\BLA761183\0003\m5\53-clin-stud-rep\535-rep-effic-safety-stud\preventionofdiabetest1dminrelativesatrisk\5351-stud-rep-contr\tn-10-addendum.

3. FDA Statistical Reviewer's Analysis Results

Under the original submission, in TN-10 CSR description, anti-teplizumab drug antibodies were analyzed on blood samples obtained at the 3-month visit. The ADA data is contained in the SDTM.PC dataset (\\CDSESUB1\evsprod\bla761183\0003\m5\datasets\tn-10-addendum\tabulations\sdtm\pc.xpt).

Following the applicant's *Response to CRL* description, this reviewer has generated Figure 2 and Table 1 below by independent programming. The reviewer's results confirmed the correctness of the applicant's results displayed in EOT Figure 2.5.4.1.

Figure 2. Bar plot – anti-teplizumab response, onset of T1DM and treatment (ITT population)



Source: FDA statistical reviewer.
Abbreviations: T1DM, type 1 diabetes mellitus; ADA, Anti-drug Antibody.

Table 1. Anti-teplizumab response, onset of T1DM and treatment (ITT population)

T1DM onset – ADA Status	Teplizumab (N=44)	Placebo (N=32)
No T1DM onset - ADA missing, n (%)	0	1 (3.1%)
No T1DM onset - ADA negative, n (%)	11 (25.0%)	7 (21.9%)
No T1DM onset - ADA positive, n (%)	13 (29.5%)	1 (3.1%)
T1DM onset - ADA missing, n (%)	2 (4.5%)	0
T1DM onset - ADA negative, n (%)	7 (15.9%)	22 (68.8%)
T1DM onset - ADA positive, n (%)	11 (25.0%)	1 (3.1%)

Source: FDA statistical reviewer.
Abbreviations: T1DM, type 1 diabetes mellitus; ADA, Anti-drug Antibody.

4. Statistical Review Comments

The statistical reviewer has confirmed the correctness of the applicant's description of the ADA status derivation algorithm and its subsequent generation of EOT Figure 2.5.4.1.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

YU WANG
09/14/2022 02:59:53 PM

YOONHEE KIM
09/14/2022 03:06:05 PM

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL MEMO

CLINICAL STUDIES

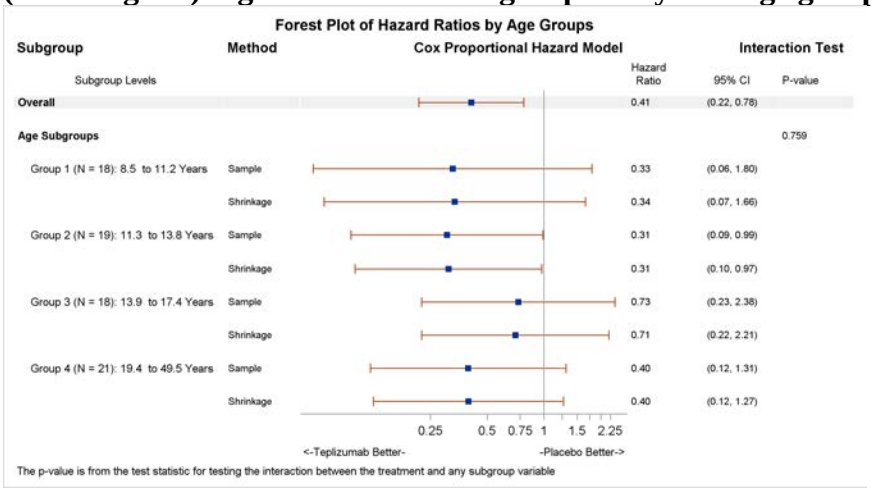
BLA #:	761183
Drug Name:	TZIELD (teplizumab)
Applicant:	Provention Bio, Inc.
Biometrics Division:	Division of Biometrics II
Statistical Reviewer:	Yu (Jade) Wang
Concurring Reviewers:	Yun Wang
Medical Division:	Division of Diabetes, Lipid Disorders, and Obesity (DDLO)
Clinical Team:	Lauren Wood Heickman
Project Manager:	Supendeep Dosanjh

The following memo was written as an addendum to the statistical review for teplizumab signed on 28 June 2021. The main purpose for this memo is to correct the errors with shrinkage estimates of subgroup treatment effects using a Bayesian hierarchical model. The erroneous estimates resulted from a SAS coding error: a by subgroup statement was inadvertently included in the PROC MCMC step that was used to generate the posterior summary statistics. In addition, the variance of the flat prior was increased from 4 to 16; and the half-Cauchy distribution was used to replace the inverse-Gamma distribution for τ^2 as recommended by Gelman (Gelman, 2006).

In our analyses with Cox proportional hazard regression model, while we reported results as hazard ratios (HR), we performed calculations on the log(HR) scale.

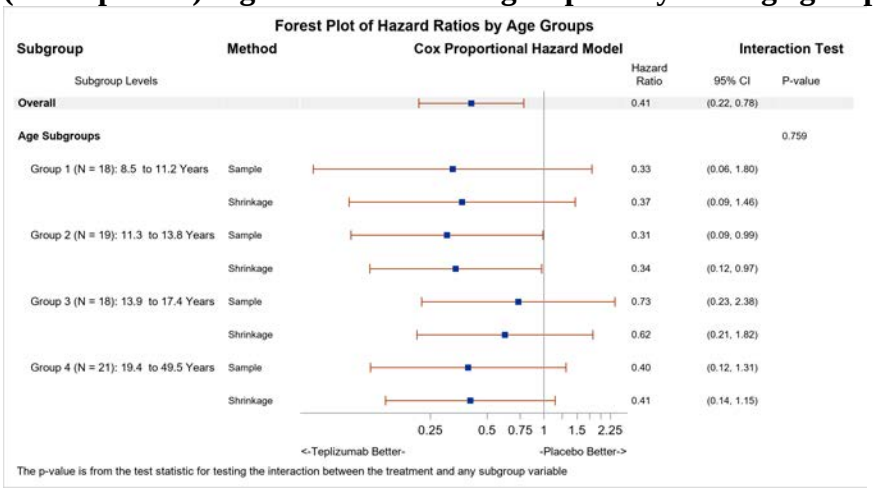
The Original Figure 5 to Figure 10 of the review, are replaced with the Updated Figures 5 to Figure 10.

(The Original) Figure 1. TN-10: Subgroup Analysis – Age groups



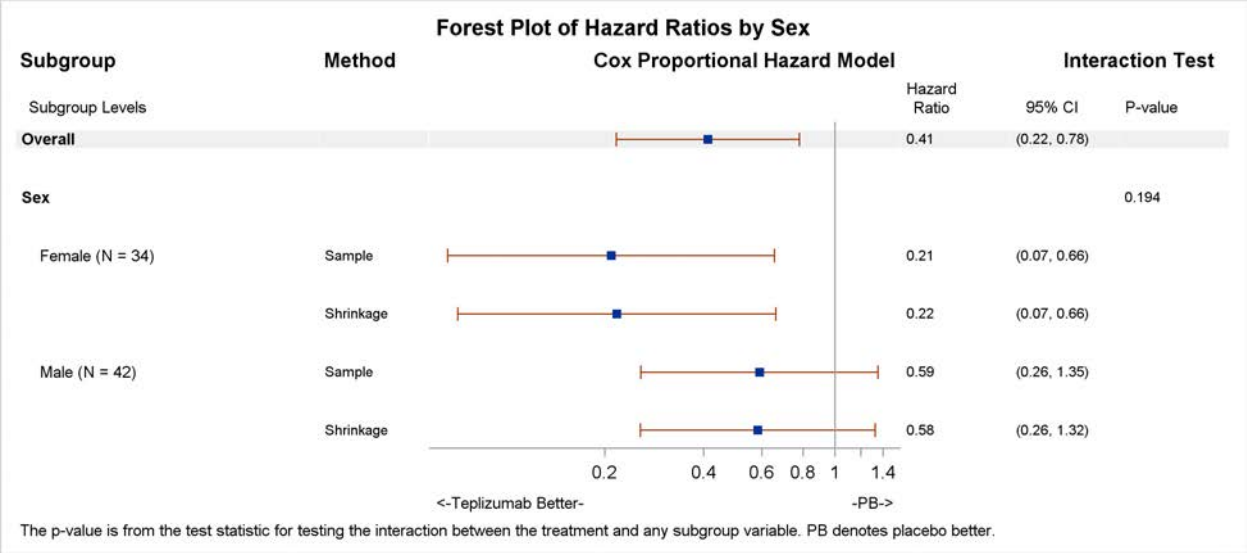
Source: FDA statistical reviewer

(The Updated) Figure 2. TN-10: Subgroup Analysis – Age groups



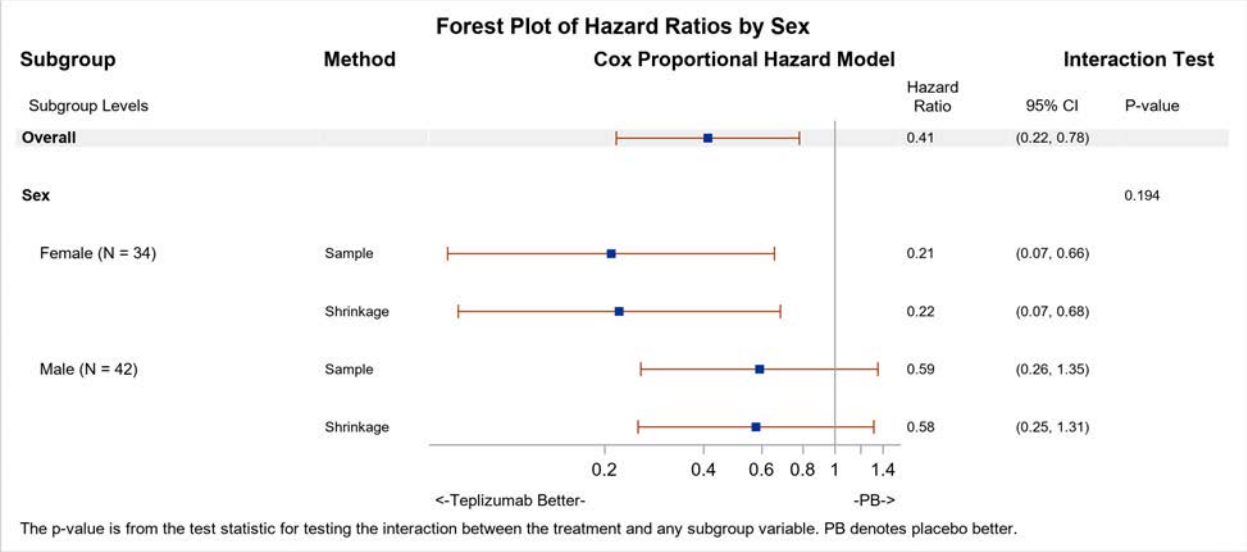
Source: FDA statistical reviewer

(The Original) Figure 3. TN-10: Subgroup Analysis – Sex



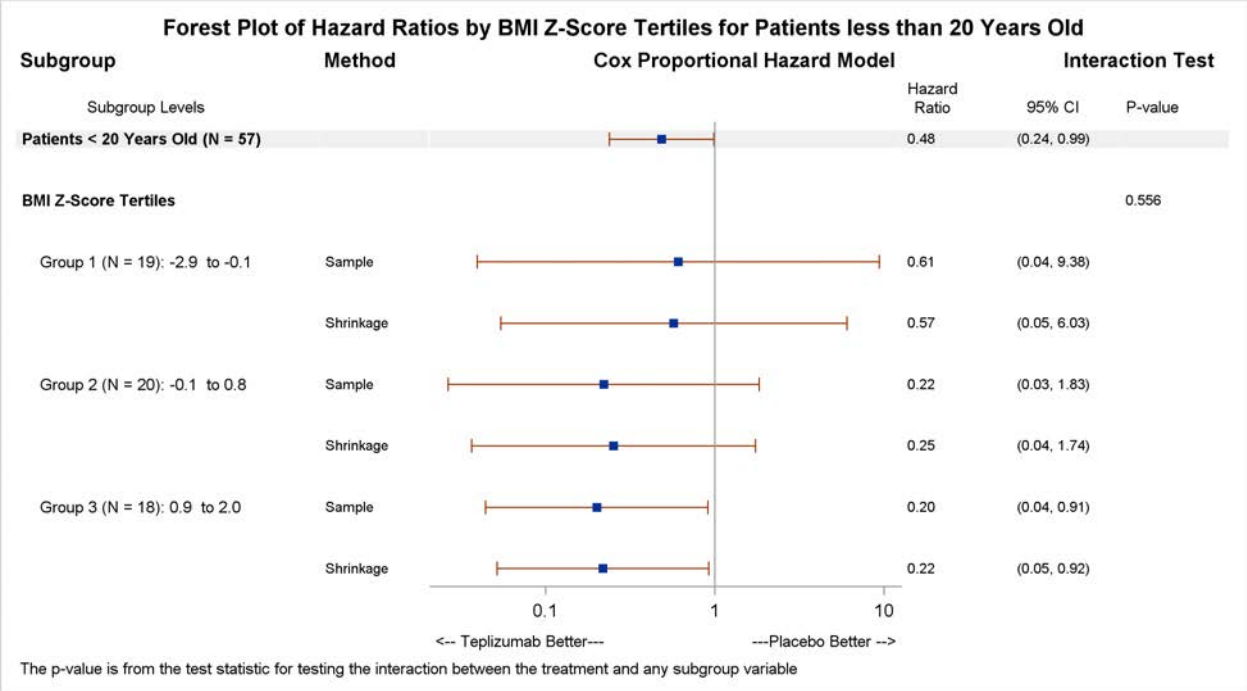
Source: FDA statistical reviewer

(The Updated) Figure 4. TN-10: Subgroup Analysis – Sex



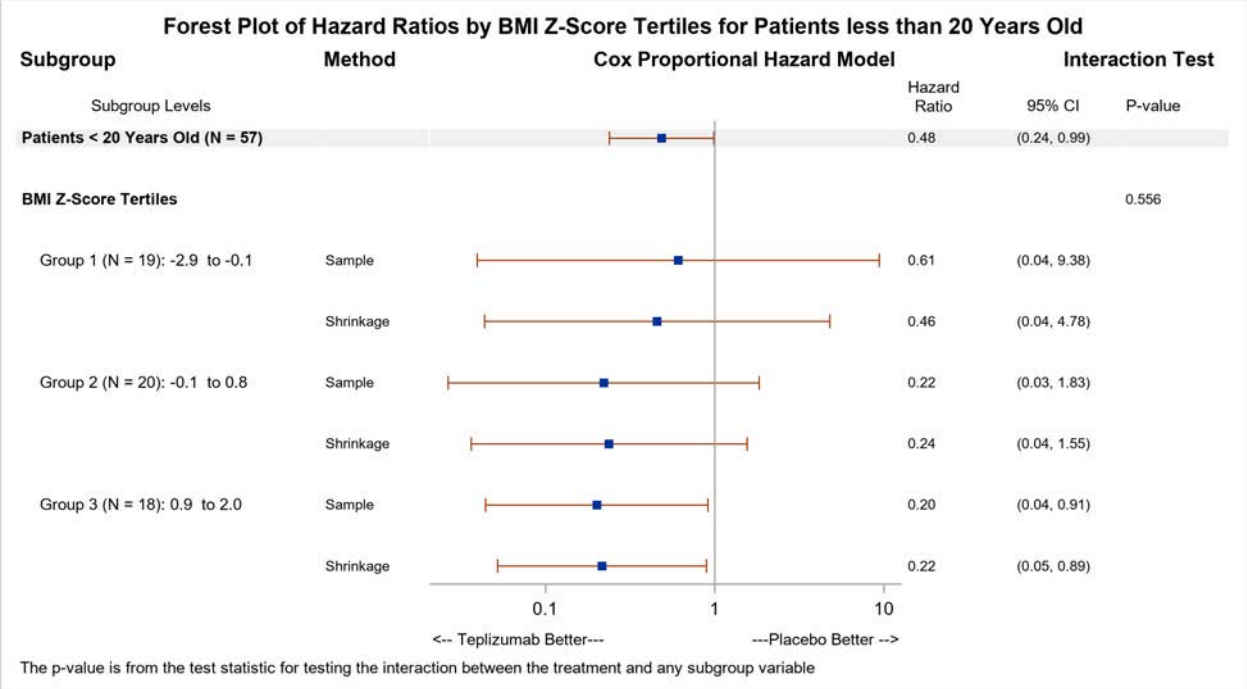
Source: FDA statistical reviewer

(The Original) Figure 5. TN-10: Subgroup Analysis – BMI-Z scores



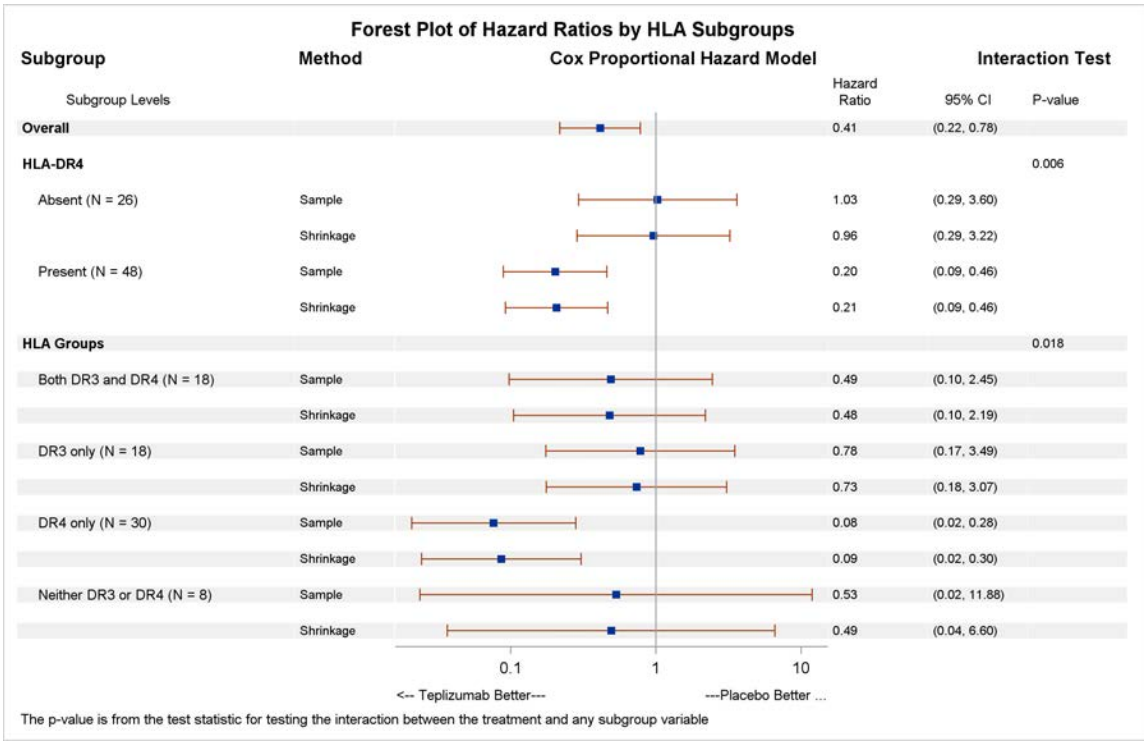
Source: FDA statistical reviewer

(The Updated) Figure 6. TN-10: Subgroup Analysis – BMI-Z scores



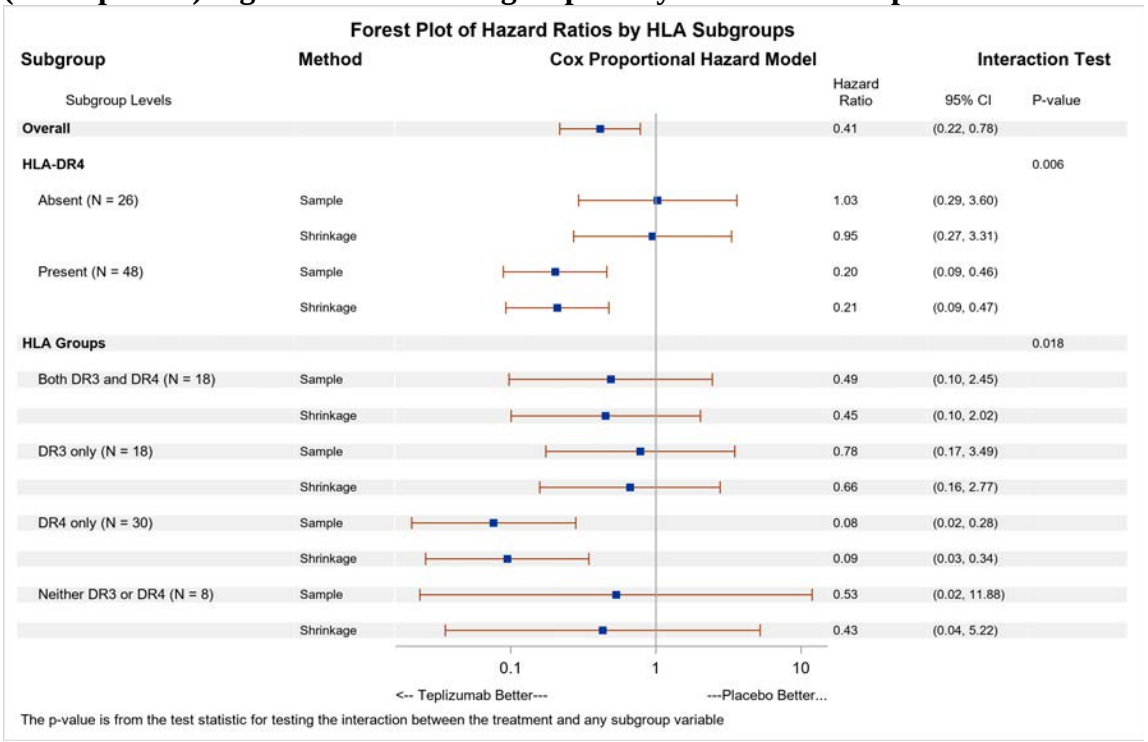
Source: FDA statistical reviewer

(The Original) Figure 7. TN-10: Subgroup Analysis – HLA-Groups



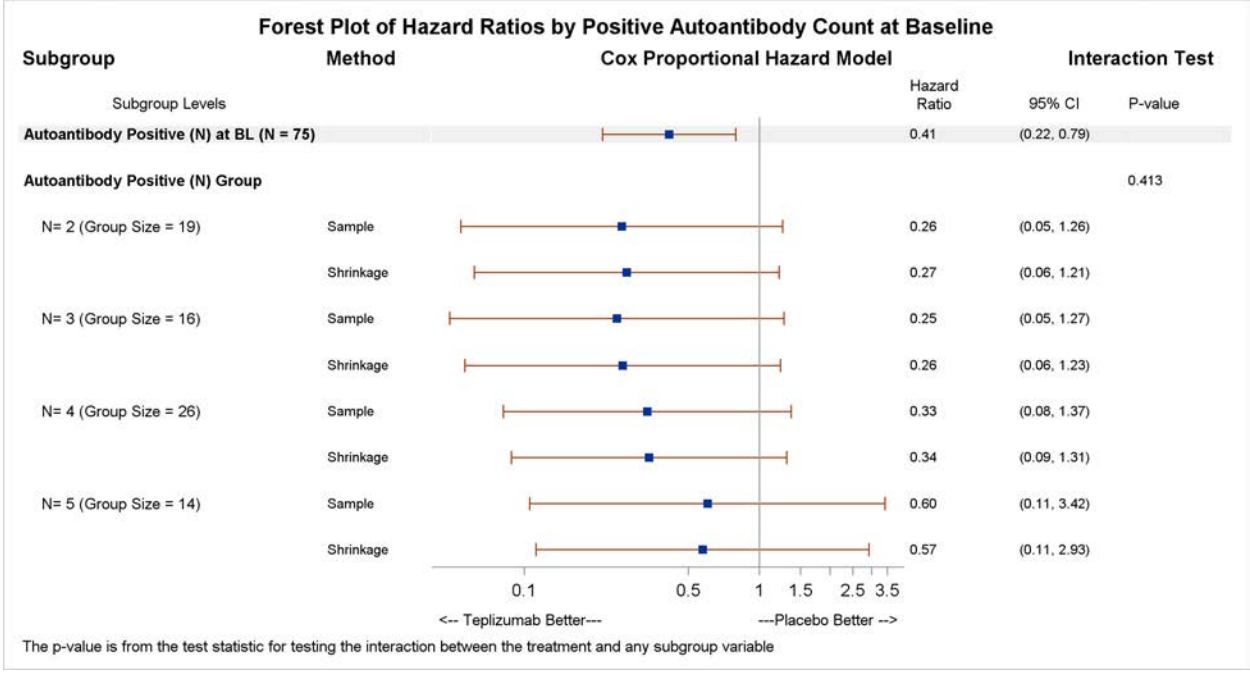
Source: FDA statistical reviewer

(The Updated) Figure 8. TN-10: Subgroup Analysis – HLA-Groups



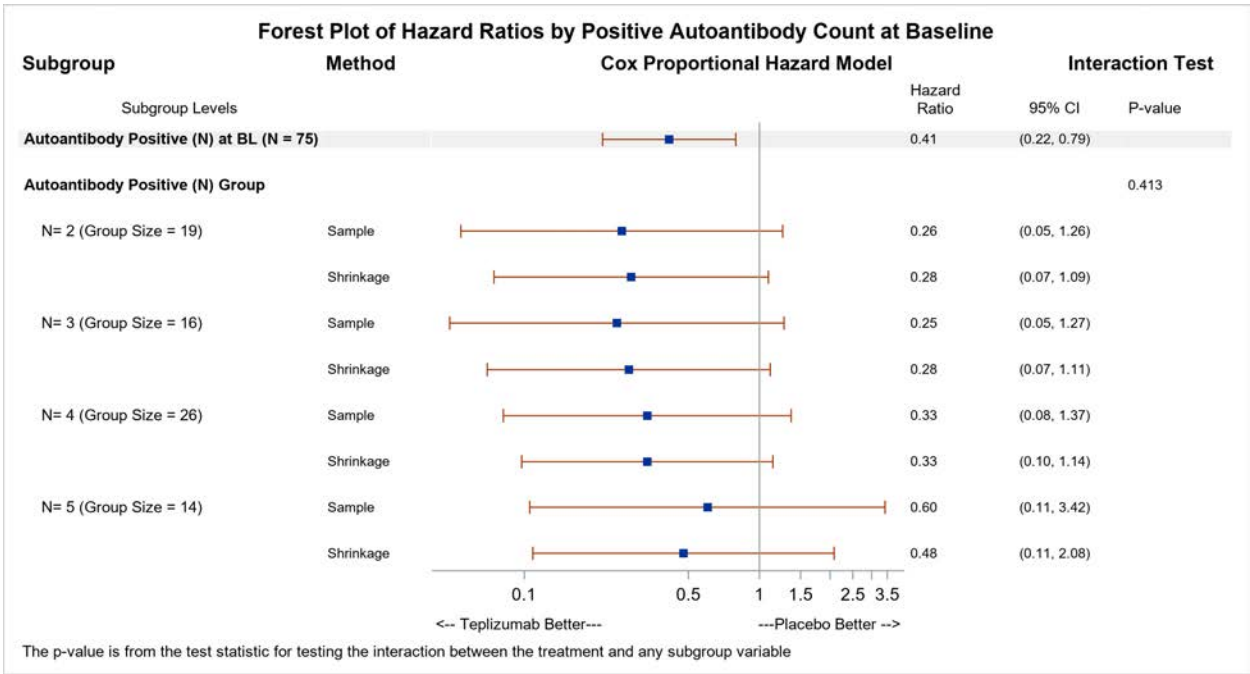
Source: FDA statistical reviewer

(The Original) Figure 9. TN-10: Subgroup Analysis – Baseline Autoantibody Positive Count



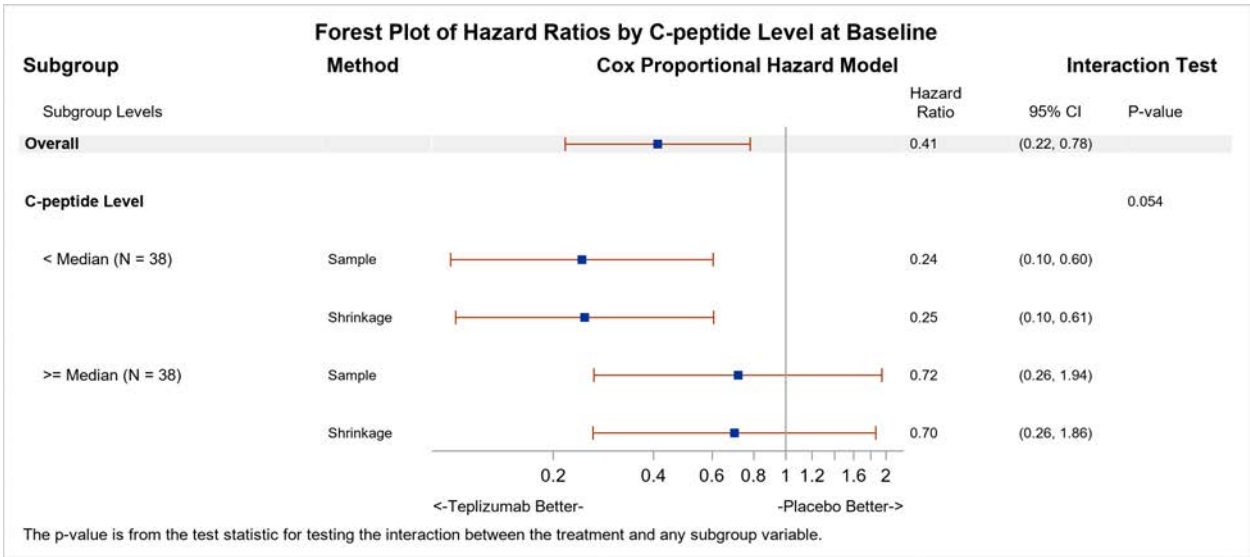
Source: FDA statistical reviewer

(The Updated) Figure 10. TN-10: Subgroup Analysis – Baseline Autoantibody Positive Count



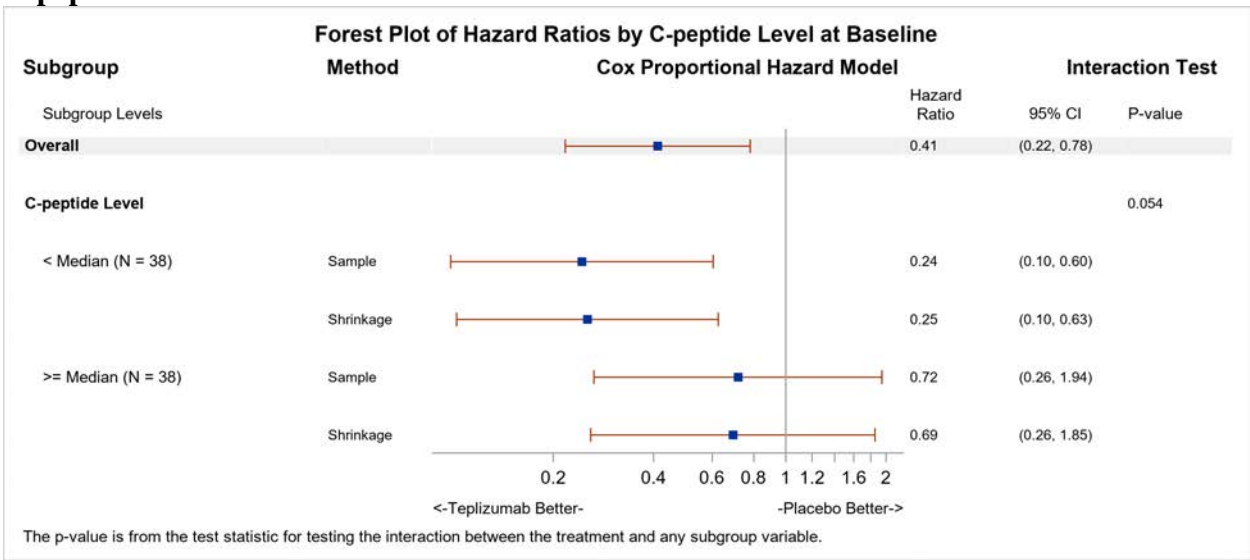
Source: FDA statistical reviewer

(The Original) Figure 11. TN-10: Subgroup Analysis – Baseline OGTT Mean AUC C-peptide Level



Source: FDA statistical reviewer

(The Updated) Figure 12. TN-10: Subgroup Analysis – Baseline OGTT Mean AUC C-peptide Level



Source: FDA statistical reviewer

References

Gelman, A. (2006). Prior distribution for variance parameters in hierarchical models. *Bayesian Analysis*, 515-533.

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

YU WANG
09/01/2021 09:49:10 AM

YUN WANG
09/01/2021 03:57:43 PM

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #:Supplement #:	BLA 761183
Drug Name:	TZIELD (teplizumab)
Indication(s):	The delay of clinical type 1 diabetes mellitus (T1D) in at-risk (Stage 2) individuals
Applicant:	Provention Bio, Inc.
Date(s):	Primary BLA Review Due Date: April 2, 2021
Review Priority:	Priority
Biometrics Division:	DBII
Statistical Reviewer:	Yu Wang
Concurring Reviewers:	Yun Wang
Medical Division:	Division of Diabetes, Lipid Disorders, and Obesity (DDLO)
Clinical Team:	Primary reviewer: Lauren Wood Heickman Medical division director: Lisa Yanoff
Project Manager:	Elisabeth Hanan

Keywords:

time to Type 1 diabetes diagnosis, C-peptide, delay of Type 1 diabetes, meta-analysis

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1 EXECUTIVE SUMMARY

The Applicant submitted one adequate and well-controlled (AWC) study TN-10 to support the proposed indication for teplizumab, delaying Type 1 Diabetes (T1D) diagnosis in high-risk first relatives of T1D patients. Confirmatory evidence consisted of meta-analyses on the treatment effect of teplizumab in slowing the rate of C-peptide decline in Stage 3 new onset T1D patients (see Table 1 for a description of studies included in the marketing application).

1.1 Study TN-10

The event driven Study TN-10 followed-up 76 Stage 2 at-risk patients with varied lengths: ranging from 9 months to 88 months, with a median follow-up time being 50.6 months. The primary analyses showed that the adjusted hazard ratio for teplizumab versus placebo over this follow-up was 0.41 with 95% CI: 0.22 to 0.79 and p = 0.0066. The median time to T1D diagnosis was 49.5 months and 24.9 months in the teplizumab and placebo groups, respectively.

Key statistical issues and findings that may impact the demonstration of efficacy included:

- TN-10 was an investigator-initiated trial with small sample size.
- Evaluation of the treatment effect of teplizumab on C-peptide area under curve (AUC) change from baseline was not possible because of study design issues.
- Imbalances in prognostic factors at randomization.

Among them, the impact of imbalances in prognostic factors at randomization on the robustness of the efficacy in delaying the onset of T1D were assessed by additional sensitivity adjusted analyses. These sensitivity analyses produced results similar to those in the primary analysis in effect size and significance level.

In summary, the treatment effect was statistically significant and robust to sensitivity analyses under a range of conservative assumptions for missing data, as well as to sensitivity analyses adjusting for prognostic factors at baseline.

1.2 The Meta-analyses

The FDA review team identified the following statistical issues in the meta-analyses that may impact its strength in providing the confirmatory evidence as intended.

- Change in C-peptide AUC is not a validated surrogate endpoint for clinical benefit, e.g. delay of progression of T1D, but literature reviews and inputs from some key opinion leaders[1] suggest it may likely predict clinical benefit. Scientific knowledge and clinical considerations both played crucial roles in how C-peptide AUC data was interpreted.

[1] Scientific workshop on June 15 – 16, 2021, Design of clinical trials in new onset Type I Diabetes: regulatory considerations.

- The meta-analysis was based on clinical T1D patients rather than at-risk patients. The applicability of evidence of C-peptide preservation among clinical T1D patients to support a delay in onset of T1D is conjectured but unconfirmed.
- The meta-analysis was proposed post hoc.
- There was notable heterogeneity in the estimated treatment effect in C-peptide preservation across the five trials included in the meta-analyses, with the two largest trials Protégé and Encore showing a small treatment effect.

Among them, the heterogeneity issue was addressed by applying a random-effects meta-analysis model in estimating the true treatment effect.

In summary, meta-analyses of five studies conducted in Stage 3 new onset T1D patients demonstrated a numerically significant treatment effect of teplizumab vs. comparator on C-peptide AUC change from baseline at 1-year after accounting for variability in treatment effects across studies; the estimated treatment effect at 2 years from the meta-analysis was associated with a wide confidence interval and was not statistically significant.

1.3 Recommendations

According to FDA's evidentiary standard for effectiveness[2], the one adequate and well-controlled trial (TN-10) provided statistically significant evidence in delaying T1D onset in Stage 2 at-risk patients; the confirmatory evidence consisted of the post-hoc meta-analyses of 5 studies in new-onset T1D patients showed numerically significant treatment effect of teplizumab vs. comparator on C-peptide AUC change from baseline at 1-year, but not 2-year. The application was brought to an advisory committee on May 27, 2021, where the panel voted 10-7 for approval. Taking into consideration the clinical review team's conclusion, we recommend approval of the proposed indication: the delay of clinical T1D in at-risk (Stage 2) individuals.

[2] The FDA guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products* (May 1998).

2 INTRODUCTION

2.1 Overview

2.1.1 Class and Indication

Teplizumab is a humanized monoclonal antibody that targets the cluster of differentiation 3 (CD3) antigen, which is co-expressed with the T-cell receptor (TCR) on the surface of T lymphocytes. This Biologics License Application (BLA) is for the delay or prevention of clinical T1D in at-risk (Stage 2) individuals. Scientific studies to date show that genetic, immunologic, and metabolic factors are all associated with heightened risk of developing type 1 diabetes (Stage 3) T1D. Teplizumab is also being developed for the treatment of newly diagnosed Stage 3 T1D.

2.1.2 History of Drug Development

2.1.2.1 Program development history

TN-10 was first brought to the FDA after study completion or near completion. Reviewer's knowledge on protocol changes are all based on the applicant's narrative.

At the BLA planning stage, for its marketing application, the Applicant intended to rely on data from the TN-10 study only for demonstration of efficacy for the proposed indication. While FDA agreed that TN-10 may be considered an adequate and well controlled trial that supports the efficacy of teplizumab, FDA noted that confirmatory evidence would also be needed to provide substantial evidence of effectiveness.

The Applicant proposed to leverage available data from five studies on teplizumab in patients with new-onset T1D to show that teplizumab protects beta cells, which provides confirmatory evidence that teplizumab may slow down T1D progression throughout the disease. Beta cell function was estimated by the biomarker C-peptide.

2.1.3 Studies in the clinical program, and specific studies reviewed

2.1.3.1 Studies in the Clinical Program

The teplizumab for the delay and prevention of Type 1 Diabetes program consisted of one pivotal study – Study TN-10. Supportive meta-analyses on treatment effect of teplizumab in slowing the rate of C-peptide decline in Stage 3 T1D patients were carried out based on 5 studies. Key design elements, primary endpoints and results, and C-peptide response data collection timepoints are summarized in Table 1.

Table 1. List of All Studies Included in the Clinical Program

Clinical Study Phase & Dates	Study Population & Sample Size	Study Design/ Type of Control	Dosing Regimen	Treatment & Follow-Up Duration
Stage 2 T1D Population				
TN-10 (At-Risk Study) Phase 2 Jul 2011 to Nov 2018	Age ≥8 years Relatives of patients with T1D at high risk for development of clinical T1D based on the presence of at least two positive autoantibodies and abnormal OGTT Enrolled total: 76 (44:32) Follow-up was still ongoing at the clinical cutoff date of TN-10	R, DB, PBO controlled, 2-arm, multicenter study Randomization ratio: 1:1	Full 14-day regimen; total dose ~9034 μg/m^2	Single-course Defined by time to protocol-specified number of T1D events observed; with median duration approximately 2 years
Stage 3 T1D Population				
Study 1 Phase 1/2 May 1999 to Aug 2001	Age 7.5 to 30 years New onset T1D (within 6 weeks of dx) Enrolled total: 42 (21:21)	R, OL, 2-arm, standard of care control Randomization ratio: 1:1	14-day regimen; total dose ~495 μg/m^2 12-day regimen; total dose ~19,500 μg/m^2	Single course 2 years
Protégé Phase 2/3 Feb 2007 to Jun 2011	Age 8-35 years New-onset T1D (within 12 weeks of dx) Enrolled total: Segment 1: 38 (38:0) Segment 2: 5131 (207:102:106:98)	Segment 1: OL Segment 2: R, DB, PBO controlled, multicenter, multinational, 4-arm, dose-ranging study Randomization ratio: 2:1:1:1	Segment 1: Full 14-day regimen Segment 2: Arm 1: Full 14-day; total dose ~9034 μg/m^2 Arm 2: 1/3 dose 14-day; total dose ~2985 μg/m^2 Arm 3: 6-day; total dose ~2426 μg/m^2 Arm 4: Placebo 14-day course	Segments 1 & 2: Two identical courses, 26 weeks apart 2 years
Encore Phase 3 Sep 2009 to Apr 2012	Age 8-35 years New-onset T1D (within 12 weeks of dx) Enrolled total: 254 (63:66:63:62)	R, DB, PBO controlled, multinational, 4-arm, dose-ranging study Randomization ratio: 1:1:1:1	Arms 1-4: Same as Protégé	Two identical courses, 26 weeks apart 89.8% completed 1 year and 29.1% completed 2 years

Clinical Study Phase & Dates	Study Population & Sample Size	Study Design/ Type of Control	Dosing Regimen	Treatment & Follow-Up Duration
AbATE Study 4 Sep 2005 to Apr 2011	Age 8-30 years New-onset T1D (within 8 weeks of dx) Enrolled total: 77 (52:25)	R, OL, 2-arm, multicenter study, standard of care control Randomization ratio: 2:1	Full 14-day regimen; total dose ~9034 μg/m^2	Two courses, 1 year apart 2 years
Delay2 Phase 2 Sep 2006 to Aug 2011	Age 8-30 years Recent T1D (within 4-12 months of dx) Enrolled total: 58 (32:28)	R, DB, PBO controlled, multicenter study, with OL extension Randomization ratio: 1:1	Full 14-day regimen; total dose ~9034 μg/m^2	Single course 1 year

Source: Clinical reviewer generated table, adapted from ISS and ISE
1 Protégé: 516 randomized, 513 received at least one dose of study drug
2 18 Delay subjects initially randomized to the control arm, who were later also elig ble for the open-label course 2 administration of teplizumab, are counted once for each treatment arm and once in the total column of subject count.
Abbreviations: IV, intravenous; T1D, type 1 diabetes; OGTT, oral glucose tolerance test; R, randomized; DB, double blind; OL, open-label; PBO, placebo; dx, diagnosis

2.2 Data Sources

Data sources include all materials reviewed, e.g. applicant study reports, data sets analyzed, and literatures referenced. Analysis datasets could be accessed at \\CDSESUB1\evsprod\BLA761183\0003\m5\datasets; analysis programs could be accessed at \\CDSESUB1\evsprod\BLA761183\0012\m5\datasets\tn-10-addendum\analysis\adam\programs; documents and literatures referenced could be accessed at \\CDSESUB1\evsprod\BLA761183\0003\m5.

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The submitted datasets were of acceptable quality in general and were adequately documented.

However, we noted there were some minor data quality issues in Study TN-10 that may have been caused by insufficient planning at the study design stage. In TN-10, T1D onset diagnosis information was all captured by one case report form (CRF) page, Diabetes Onset. The information on the Diabetes Onset page were transcribed into the Standard Data Tabulation Model (SDTM) ZD domain, which was subsequently used as the main source for deriving the primary analysis dataset ADTTE. However, this process lacked an important step of cross validation across different source domains for both data accuracy and data consistency; FDA statistical reviewer identified data inconsistency issues between the source datasets, ZD and ADLBOG. The insufficiency at the planning stage was also reflected in the Applicant's substitution of the "old" version of the Diabetes Onset CRF page with a "current" version at a

very late stage of the trial with the purpose to enable collection of pertinent diagnosis data. FDA's information request on these issues (IR, sent on June 22, 2021) and the Applicant's Response to the IR and supporting documents could be located under \\CDSESUB1\evsprod\BLA761183\0054\m1\us.

We noted that Study TN-10 was an investigator-initiated trial, and as a first time-to-event trial for delay of T1D, it is understandable that TN-10 had these insufficiencies in trial design and the subsequent analysis dataset generation. Also, the FDA medical reviewer's manual checking of the source datasets assured us of the quality of the data in supporting the primary analysis. While we considered these issues did not impact the interpretation of the overall study results, we document these findings here for future reference.

3.2 Evaluation of Efficacy – TN-10

3.2.1 Study Design and Endpoints

Study TN-10 was a randomized, placebo-controlled, double-blind, multicenter study designed to evaluate the treatment effect of teplizumab compared to placebo on the delay of diagnosis of clinical T1D in at-risk individuals.

3.2.1.1 Patient Population and Enrollment

The trial was conducted in relatives of T1D patients aged 8 years and older who were considered at high risk for developing T1D based on the presence of two or more positive autoantibodies and an abnormal oral glucose tolerance test (OGTT). The ongoing TrialNet Natural History/Pathway to Prevention study (TN-01) was a key source at the time of inception of study TN-10; it provided TN-10 with potential study subjects and facilitated TN-10 study subject identification and recruitment.

3.2.1.2 Efficacy Endpoints

3.2.1.2.1 Primary Efficacy Endpoints

The primary efficacy endpoint was the elapsed time from randomization to diagnosis of T1D or time of last contact.

Criteria for T1D onset were based on glucose testing or the presence of unequivocal hyperglycemia with acute metabolic decompensation, such as diabetic ketoacidosis (DKA). One of the following criteria had to be met on 2 occasions as soon as possible but ≥1 day apart:

- Symptoms of diabetes plus a casual plasma glucose concentration >200 mg/dL (11.1 mmol/L). Casual is defined as any time of day without regard to time since last meal. The classic symptoms of diabetes include polyuria, polydipsia, and unexplained weight loss.

- Fasting plasma glucose ≥126 mg/dL (7 mmol/L). Fasting was defined as no caloric intake for at least 8 hours.
- 2-hour plasma glucose ≥200 mg/dL (11.1 mmol/L). The test had to be performed using a glucose load containing the equivalent of 1.75 g/kg body weight to a maximum of 75 g anhydrous glucose dissolved in water.

3.2.1.2.2 Secondary Efficacy Endpoints

Stimulated C-peptide parameters are derived from C-peptide measurement assessed through 2-hour OGTT. The parameters include area under curve (AUC), mean AUC, and peak C-peptide. In this review, we will focus on mean C-peptide AUC as it is the parameter the applicant used.

The objective was to determine whether treatment with teplizumab was superior to placebo with respect to C-peptide responses to oral glucose, as obtained from repeated collections during longitudinal tests. However, it was specified in the study protocol that subjects were not to be followed up for 2-hour OGTT monitoring upon T1D diagnosis. The clinical study report (CSR) explained that due to both the varying lengths of patient's participation in this time-to-event study setting and the discontinuation of C-peptide AUC follow-up upon T1D diagnosis, the C-peptide level variables could not be used to support an appropriate intent-to-treat analysis.

The study was not designed to fully support these assessments, because C-peptide was not collected once a patient developed T1D, we performed exploratory analyses utilizing collected data and fully acknowledged restrictions on interpretation posed by the study design limitations. Therefore, we need to remind readers that analyses of C-peptide data collected in Study TN-10 are exploratory in nature.

These design limitations caused informative censoring and rendered analysis results of available C-peptide data uninterpretable.

3.2.1.3 Sample Size and Power

TN-10 was an event-driven study with the primary objective being evaluation of treatment effect in delay of T1D onset. The original study protocol planned to enroll 144 subjects over 2 to 3 years. Because of slower than expected accrual rates, the protocol was revised to reduce the sample size to 71 and to follow all enrolled subjects until 40 subjects were diagnosed with T1D. The study was targeted to detect an effect size measured by a hazard ratio of 0.4 for teplizumab vs. placebo, with 80% power at a one-sided alpha level of 0.025.

3.2.1.4 Randomization

Randomization was stratified by TrialNet study site and whether the subject was less than 18 years of age or 18 years and older.

The TrialNet Coordinating Center (TNCC) generated the randomization numbers (CSR Appendix 16.1.6). At each participating site, stratified by the age, subjects were randomized in a 1:1 ratio to either teplizumab or placebo.

3.2.1.5 Interim Analysis

The study protocol pre-planned to conduct interim analyses periodically during the study. The Lan and DeMets alpha spending function approach (Lan and DeMets 1983), a group sequential method that controls type I error level while allowing for the number and exact timing of the interim analyses to remain unspecified a prior, was used to protect the type I error probability for the primary outcome analyses, and to assess the significance of the interim results periodically during the trial. The spending function that approximates the O'Brien-Fleming boundaries was used:

$$\alpha_1(t^*) = 2 - 2\Phi\left[\frac{Z_{\alpha/2}}{\sqrt{t^*}}\right]$$

where t^* is the information fraction ($0 < t^* \leq 1$) and α is the fixed-sample type I error (i.e., 0.025).

One interim analysis was conducted when 18 patients were diagnosed with T1D. The Applicant applied the Lan-DeMets decision rule and spent type I error of p = 0.00083, one-sided.

$$\alpha_1(t^*) = 2 - 2\Phi\left[\frac{Z_{\alpha/2}}{\sqrt{t^*}}\right] = 2 - 2\Phi\left[\frac{Z_{0.025/2}}{\sqrt{18/40}}\right] = 0.00083$$

Therefore, the final analysis should be conducted at 0.0242 level to preserve type I error. Given this small difference from 0.025, the Applicant reported p-values associated with the primary and sensitivity analyses of the primary endpoint as if a fixed-sample size test was conducted. We consider this handling reasonable.

The TrialNet Data and Safety Monitoring Board (DSMB) reviewed the results for assessment of effectiveness and safety. TN-10 Protocol *Section 3.6.1 Procedures for Unmasking*, *Section 9.2 Participating Centers*, and *Section 9.4 Study Subject Confidentiality* contain description of procedures implemented to ensure study blinding and integrity.

FDA review team issued an information request on June 24, 2021 requesting submission of the interim analysis DSMB meeting minutes and clarification on who had access to the unblinded interim analysis results. In response, the Applicant submitted the *TrialNet DSMB Meeting Minutes* which also contains the DSMB roster and listing of personnel who had access to the unblinded interim analysis results (\\CDSESUB1\evsprod\BLA761183\0055\m1\us).

After reviewing these documents, we conclude that the Applicant' conduct and analysis in the interim analysis process followed the pre-planned procedures in protecting type I error and ensuring study integrity.

3.2.2 Statistical Methodologies

3.2.2.1 Analysis Sets

Two analysis sets were defined to support TN-10 efficacy analyses: the intention-to-treat (ITT) population included all subjects who were randomized; the full analysis set (FAS) included all subjects who received any study treatment (including control). Analyses on ITT or FAS sets were conducted by assigned treatment, regardless of treatment adherence.

3.2.2.2 Analysis Methods – Time to Onset of T1D

The Kaplan-Meier (KM) plot was used to show the proportion of patients free of TID by time since randomization. A Cox proportional hazards (PH) model was used to analyze the time to T1D diagnosis, stratified by age and OGTT status at randomization.

The Applicant also conducted three prospectively planned sensitivity analyses as follows:

- Sensitivity analysis #1 (SA1) set time-to-diagnosis of T1D to 1 year for placebo patients whose time to event was within 1 year. By delaying event time in a portion of patients in the placebo arm, this approach favors placebo to an extent depending on the actual data.
- Sensitivity analysis #2 (SA2) assumed that subjects who were censored before study cutoff had been diagnosed with T1D at their last evaluation. SA2 penalizes missing data without favoring either treatment group at the analysis planning stage. The actual data will decide which arm will benefit from this approach.
- Sensitivity analysis #3 (SA3) was a combination of SA1 and SA2.

The three pre-planned sensitivity analyses (SAs) were all reasonable in handling missing data. However, method SA2 penalizes missing data without differentiation, and depending on the real missing pattern, it may not be sufficiently conservative to evaluate the treatment effect of teplizumab on delay of T1D diagnosis. We therefore performed a fourth sensitivity analysis that was similar to SA2, but only applied the imputation (T1D diagnosis) to teplizumab patients who were censored before study cutoff.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

A total of 146 subjects were screened for inclusion, and 76 eligible patients were randomly assigned to the two treatment groups: 44 (58%) to the teplizumab group and 32 (42%) to the placebo group. With respect to the numerical imbalance between treatment groups, the Applicant postulates that with stratified randomization, although near balance was reached within strata, imbalance for the total trial occurred due to chance.
All patients received at least one dose of the randomized treatment; actual/received treatment aligned with assigned treatment (Table 2).

Table 2. TN-10: Analysis Sets (All Screened Patients)

Analysis Set	Teplizumab	Placebo	Total
All Screened, N			146
Screen Failure, N			70
All Randomized Patients, N	44	32	76
Intent-To-Treat Population, n (%)	44 (100)	32 (100)	76 (100)
Full Analysis Set, n (%)	44 (100)	32 (100)	76 (100)

Source: FDA Statistical Reviewer

Table 3 summarizes treatment completion status and study completion status. Three patients discontinued study treatment during the 14-day treatment course: one (2.3%) from the teplizumab group, and two (6.3%) from the placebo group. The reason for treatment discontinuation was protocol-specified drug withholding for all three patients (discussed further in the safety summary section of the Clinical Review of this BLA). To prevent missing data, these patients were followed-up in the same fashion as patients who completed the treatment course.

As prespecified, the clinical cutoff of the study occurred when more than 40 patients were diagnosed with T1D. At that time, patients who had experienced a T1D diagnosis henceforth met the protocol-specified withdrawal criterion and were considered as having "completed study." Study completion status for the 30 patients who were still T1D-free and continued to be followed-up were categorized as "ongoing." 6 patients discontinued the study before the clinical cutoff: 3 were diagnosed with T1D, and 3 were censored at their last contact date.

Table 3. TN-10: Patient Dispositions (ITT Population)

		Teplizumab	Placebo	Total
		N=44 n (%)	N=32 n (%)	N=76 n (%)
Treatment Completion Status	**Completed**	43 (97.7)	30 (93.8)	73 (96.1)
	Discontinued	1 (2.3)	2 (6.3)	3 (3.9)
	Reason for discontinuation of treatment			
	Protocol specified drug withholding	1 (2.3)	2 (6.3)	3 (3.9)
Study Completion Status	**Completed the study**	19 (43.2)	21 (65.6)	40 (52.6)
	Ongoing: Follow-up still ongoing at cutoff date	22 (50.0)	8 (25.0)	30 (39.5)
	Discontinued from the study	3 (6.8)	3 (9.4)	6 (7.9)
	Reason for discontinuation from study			
	Discontinued and had T1D onset time recorded			
	Onset of Diabetes	0	2 (6.3)	2 (2.6)
	Patient chose to no longer participate	1 (2.3)	0	1 (1.3)
	Discontinued with missing T1D onset time censored at last contact date			
	Lost to follow-up	0	1 (3.1)	1 (1.3)
	Participant's decision	2 (4.5)	0	2 (2.6)

Source: FDA Statistical Reviewer
Abbreviations: ITT, intent-to-treat; T1D, type 1 diabetes

Table 4 summarizes demographics by treatment group. Almost 75% of randomized patients were younger than 18 years of age. We further subgrouped the population by age to better study the potential impact of age on treatment effect. Based on literature review, the FDA Clinical Review of the BLA explained that there is an inverse association between age and onset of T1D, and disease incidence peaks in childhood between ages 10 to 14 years (Maahs, et al. 2010). This association has been shown in several cohorts of individuals at risk of T1D (Ziegler, et al. 2013; Steck, et al. 2015; Gorus, et al. 2017), including stage 1 and 2 T1D (Jacobsen, et al. 2020). The FDA Clinical Review further pointed out that it is unclear whether age impacts the risk of progression to T1D once metabolic disturbances (such as the dysglycemia seen in stage 2 T1D) are present. Without proper adjustment, the observed baseline age imbalance between teplizumab and placebo groups may impact the results in favor of teplizumab because the expected rate of progression to T1D would be slower in adults than in children. However, we note that the primary analysis had already adjusted for Age and OGTT strata, so we think there is no need for additional statistical procedures to address impact of imbalance in age.

Table 4. TN-10: Baseline Demographics (ITT Population)

		Teplizumab	Placebo	Total
		N=44	N=32	N=76
Age (Years)	Mean (SD)	19.2 (11.9)	17.5 (11.1)	18.5 (11.5)
	Median (Minimum, Maximum)	14 (9, 49)	13 (9, 45)	14 (9, 49)
Age Group /Strata, n (%)	>= 18	15 (34.1)	6 (18.8)	21 (27.6)
	< 18	29 (65.9)	26 (81.3)	55 (72.4)
	<18 OGTT Confirm	24 (54.5)	23 (71.9)	47 (61.8)
	<18 OGTT not Confirm	5 (11.4)	3 (9.4)	8 (10.5)
Age quartiles to support age subgroup analysis, n (%)	Pediatrics 1: 8.5 to 11.2 Years	10 (22.7)	8 (25.0)	18 (23.7)
	Pediatrics 2: 11.3 to 13.8 Years	10 (22.7)	9 (28.1)	19 (25.0)
	Pediatrics 3: 13.9 to 17.4 Years	9 (20.5)	9 (28.1)	18 (23.7)
	Adults: 19.4 to 49.5 Years	15 (34.1)	6 (18.8)	21 (27.6)
Sex, n (%)	Female	19 (43.2)	15 (46.9)	34 (44.7)
	Male	25 (56.8)	17 (53.1)	42 (55.3)
Race, n (%)	Asian	0	1 (3.1)	1 (1.3)
	Multiple	0	1 (3.1)	1 (1.3)
	White	44 (100)	30 (93.8)	74 (97.4)
Country, n (%)	Canada	2 (4.5)	0	2 (2.6)
	Germany	1 (2.3)	1 (3.1)	2 (2.6)
	USA	41 (93.2)	31 (96.9)	72 (94.7)

Abbreviations: ITT, intent-to-treat; OGTT, oral glucose tolerance test; SD, standard deviation
Source: FDA Statistical Reviewer

A slightly higher proportion (55.3%) of the patient population was male. The trial consisted predominantly of White patients (97.4%) and patients from the US (94.7%). Treatment groups were balanced for sex, and race.

Important baseline characteristics are summarized in Table 5. Treatment groups were balanced for human leukocyte antigen (HLA) subtypes and antibody type.

Table 5. TN-10: Baseline Characteristics (ITT Population)

		Teplizumab	Placebo	Total
		N=44	N=32	N=76
BMI (kg/m^2)	Mean (Standard Deviation)	21.9 (6.4)	22.1 (4.4)	22.0 (5.6)
	Median (Minimum, Maximum)	20 (15, 44)	22 (16, 35)	21 (15, 44)
BMI Z-score Group, n (%)	Group 1: Patient < 20 Yrs and BMI Z-score (-2.9, -0.1)	17 (38.6)	2 (6.3)	19 (25.0)
	Group 2: Patient < 20 Yrs and BMI Z-score (-0.1, -0.8)	5 (11.4)	15 (46.9)	20 (26.3)
	Group 3: Patient < 20 Yrs and BMI Z-score (0.9, 2.0)	9 (20.5)	9 (28.1)	18 (23.7)
	Group 4: Patient ≥20 Yrs	13 (29.5)	6 (18.8)	19 (25.0)
HLA-DR4, n (%)	DR4 Information Missing	2 (4.5)	0	2 (2.6)
	Absent	15 (34.1)	11 (34.4)	26 (34.2)
	Present	27 (61.4)	21 (65.6)	48 (63.2)
HLA-DR3/DR4, n (%)	Both DR3 and DR4	11 (25.0)	7 (21.9)	18 (23.7)
	DR3 only	10 (22.7)	8 (25.0)	18 (23.7)
	DR4 only	16 (36.4)	14 (43.8)	30 (39.5)
	Missing	2 (4.5)	0	2 (2.6)
	Neither DR3 or DR4	5 (11.4)	3 (9.4)	8 (10.5)
Autoantibodies Positive (N), n (%)	1	1 (2.3)	0	1 (1.3)
	2	12 (27.3)	7 (21.9)	19 (25.0)
	3	11 (25.0)	5 (15.6)	16 (21.1)
	4	12 (27.3)	14 (43.8)	26 (34.2)
	5	8 (18.2)	6 (18.8)	14 (18.4)

Source: FDA statistical reviewer
Abbreviations: BMI, body mass index; HLA, human leukocyte antigen; ITT, intent-to-treat; SD, standard deviation; Yrs: Years.

The applicant prespecified in the study protocol that analyses based on past data suggests that responses to teplizumab differ based on genetic characteristics of the participants. The absence of one type 1 diabetes-associated major histocompatibility complex (MHC) allele, HLA haplotype

HLA-DR3, the presence of HLA-DR4, as well as the absence of anti-ZnT8 antibodies identified the persons most likely to have a response.

Because of its prognostic potential on the clinical outcome (higher risk of progression with a larger number of positive autoantibodies at baseline), the number of baseline positive autoantibodies was summarized by treatment arm. There was a baseline imbalance in treatment allocation across different numbers of positive autoantibodies. This difference is demonstrated by a lower proportion of patients with 4 or 5 islet autoantibodies at randomization in the teplizumab arm (20/44 [45.5%]) compared with (20/32 [62.5%]) in the placebo arm. Overall, there were more teplizumab patients who had 2 or 3 autoantibodies at baseline and fewer teplizumab patients who had 4 or more at baseline. This difference may impact the results in favor of teplizumab because the expected rate of conversion to clinical T1D may be slower in the group with fewer positive autoantibodies at baseline.

Body mass index (BMI) z-score, used in children to control for the effects of age, and sex on BMI, was calculated by the clinical reviewer for patients younger than 20 years of age. While BMI was mostly balanced across groups, for patients younger than 20 years of age, the BMI z-score, was lower at baseline amongst patients treated with teplizumab. As explained in the FDA Clinical Review, there is some evidence from at-risk cohorts, including the TrialNet pathway to prevention cohort, linking differences in BMI z-score to differences in the risk of onset of clinical T1D, with higher BMI z-score predicting more rapid T1D onset (Barker, et al. 2007; Viner, et al. 2008). The lower BMI z-score at baseline among teplizumab patients may impact the results in favor of teplizumab, as placebo patients (with higher BMI z-score) would be at a higher risk of progression. Similar to age, however, it is unclear whether BMI or BMI z-score affects the risk of progression to T1D in multiple-antibody-positive individuals once metabolic disturbances are present. Therefore, we carried out an adjusted analysis by adding BMI at baseline into the primary analysis model as a stratification factor, and performed a subgroup analysis to assess the effect of BMI z-score on treatment effect for patients younger than 20 years of age.

3.2.4 Results and Conclusions

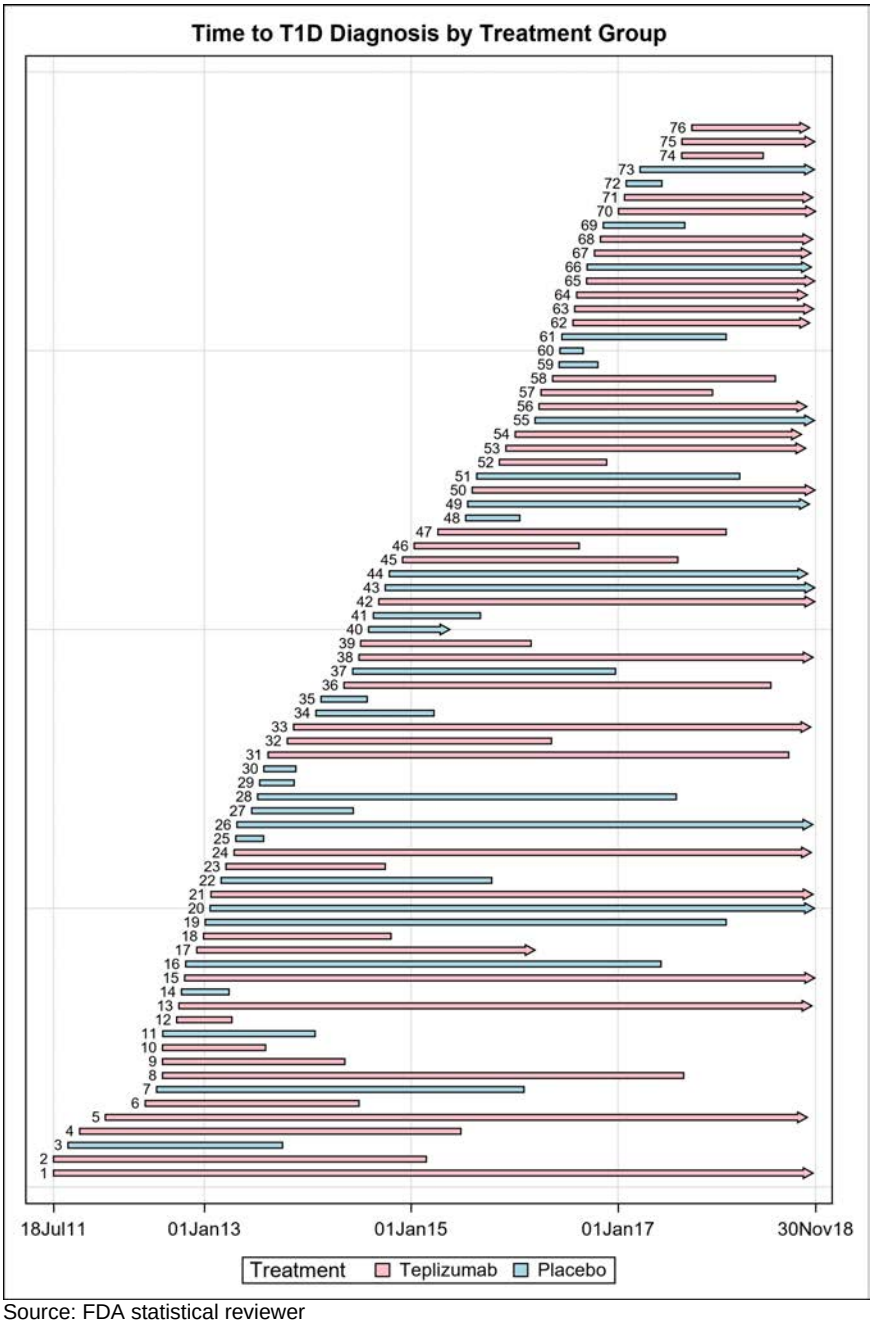
3.2.4.1 The primary endpoint: time to T1D diagnosis

To better understanding the time to event data collection and structure, a “swimmer plot” (Figure 1) was adopted for the TN-10 study design setting to show the duration of follow-up and disease status at last contact or study cutoff for each patient. The horizontal axis is the time axis, the trial started in April 2011 and ended on November 30, 2018. Vertically, each horizontal bar represented a patient, ordered according to the time they were randomized to the study and indexed by an order number at the left end of each bar. At the right end of each bar, arrow ends denote censoring due to study discontinuation or censoring at the end of the trial (the clinical cutoff); square ends denote T1D diagnoses.

The Kaplan-Meier plot (Figure 2) shows the declining percentage of patients remaining T1D free over time. The numbers at risk in each group are shown at 12-month intervals of the follow-up.

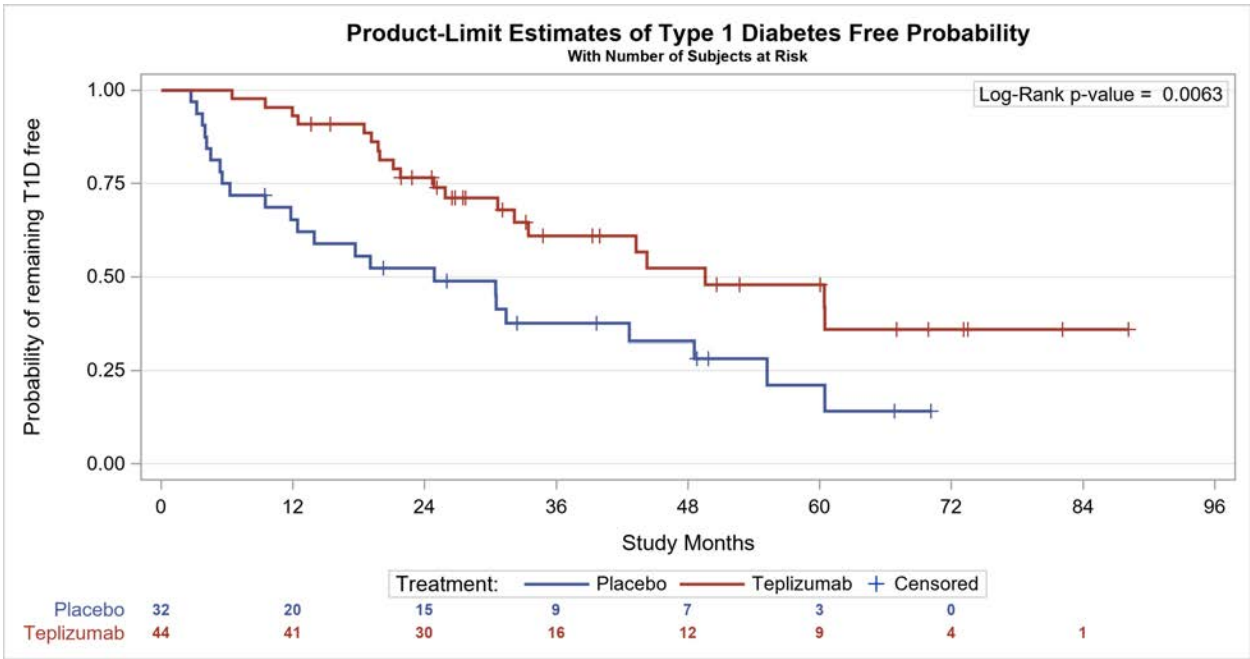
Since the trial ended when 43 patients were diagnosed with T1D, the lengths of follow-up varied among patients, ranging from 9 months to 88 months. The median follow-up time was similar between the two treatment arms, and the overall median follow-up time was 50.6 months. The curves diverge in favor of teplizumab; the difference in T1D free probability is greatest at 18 months and becomes smaller after that. The two curves stay almost parallel for the rest of the follow-up.

Figure 1. TN-10: Swimmer plot of time to T1D diagnosis and patient participation (ITT Population)



Source: FDA statistical reviewer

Figure 2. TN-10: Kaplan-Meier curves for time to T1D diagnosis



Source: FDA statistical reviewer

3.2.4.1.1 Primary Analysis and Sensitivity Analyses to Missing Data

The first part of Table 6 summaries T1D events and censoring information. At the clinical cutoff, 43 subjects were known to have had T1D onset, 3 subjects' T1D status was missing due to early discontinuation from study, and 30 patients were T1D-free. The numbers of patients having T1D were 20 (45%) and 23 (72%) in the teplizumab and placebo groups, respectively. The followed-up 76 Stage 2 at-risk patients with varied lengths: ranging from 9 months to 88 months, with a median follow-up time being 50.6 months. The median time to T1D diagnosis was 49.5 months and 24.9 months in the teplizumab and placebo groups, respectively. In the primary analysis of time to T1D diagnosis, the hazard ratio was 0.41 with 95% CI: 0.22 to 0.79 and p = 0.0066.

The total missing data on time to T1D diagnosis were few (three missing in total: two in the teplizumab group, one in the placebo group). All three pre-planned sensitivity analyses (Table 7) produced similar, although slightly higher, hazard ratios compared with the primary analysis results.

We performed a sensitivity analysis that was similar to, but more conservative than SA2. Instead of treating all three missing data points as events as per SA2, we censored the single placebo arm dropout and imputed the two teplizumab dropouts as events (we refer to this sensitivity analysis as SA4). As expected, the hazard ratio from SA4 was slightly larger than the others (0.44), but the nominal p-value remained low (0.01).

Table 6. TN-10: Primary Efficacy Endpoint Analysis Results (ITT Population)

	Teplizumab	Placebo	Total
	N=44	N=32	N=76
Event Numbers			
Number of patients diagnosed with T1D	20 (45%)	23 (72%)	43 (57%)
Number of censored patients	24 (55%)	9 (28%)	33 (43%)
Censored due to reaching trial end: follow-up still ongoing at cutoff date	22 (50%)	8 (25%)	30 (39%)
Censored due to discontinuation from study for reasons other than diagnosis	2 (5%)	1(3%)	3 (4%)
Follow-up Time			
Median follow-up time (month)	50.6	49.9	50.6
Time to Event Summary Statistics			
Median time to T1D diagnosis (month) (95% CI)	49.5 (32.1, NE)	24.9 (9.5, 48.6)	42.6 (25.9, 55.2)
Primary Analysis			
Hazard ratio (95% CI)	0.41 (0.22, 0.78)		
P-value	0.0066		

Abbreviations: T1D, Type 1 Diabetes; CI, confidence intervals, TTE, time to event; NE, not estimable.
Source: FDA statistical reviewer.

Table 7. TN-10: Primary Efficacy Endpoint Sensitivity Analysis for Impact of Missing Data (ITT Population)

Applicant's Pre-planned Sensitivity Analyses (SA)		
Method	**HR (95% CI), Nominal P-value**	**Delay in Median (Months)**
SA1: Set TTE to 1 year for placebo patients whose TTE < 1 year	0.43 (0.22, 0.81), 0.0089	49.5 – 24.9 = 24.6
SA2: Set TTE to last evaluation for early dropouts who were T1D free	0.42 (0.22, 0.78), 0.0065	44.3 – 22.0 = 22.3
SA3: A combination of SA1 and SA2	0.43 (0.23, 0.81), 0.0086	44.3 – 22.0 = 22.3
FDA Reviewer's Sensitivity Analysis		
SA4: Set TTE to last evaluation only for early teplizumab dropouts who were T1D free	0.44 (0.24, 0.83), 0.0113	44.3-24.9=19.4

Abbreviations: T1D, Type 1 Diabetes; CI, confidence intervals, TTE, time to event.
Source: FDA statistical reviewer.

3.2.4.1.2 Sensitivity Analyses to Imbalances in Prognostic Factors at Randomization

To address the concern that the imbalance at allocation between treatment groups in autoantibody count at randomization, which has a prognostic potential, may have impacted the results in favor of teplizumab, we performed an adjusted analysis by adding autoantibody count at randomization into the primary analysis model as a stratification factor. The adjusted analysis estimates the effect of teplizumab after adjusting for the covariate. The hazard ratio, 0.44 (95% CI: 0.22 to 0.87, p=0.0175), is numerically higher than what was estimated in the primary analysis, but still statistically significant. Similarly, we also carried out an adjusted analysis by adding BMI at baseline into the primary analysis model as a stratification factor. The hazard ratio, 0.41 (95% CI: 0.21 to 0.78, p=0.006), is very close to what was estimated in the primary analysis.

To address the concern that each of the above described adjusted analyses only assesses the impact of imbalance one at a time, we also performed an additional analysis by adjusting for both baseline positive autoantibody count and baseline BMI. The resulted hazard ratio was 0.42 (95% CI: 0.21 to 0.84, p=0.0136), again, very close to what was estimated in the primary analysis. With these being said, we also admit that factors such as the small sample size and the substantial correlation between BMI and Age may limit the ability of these analyses in detecting any impact caused by the imbalances.

Table 8. TN-10: Primary Efficacy Endpoint Sensitivity Analysis for Impact of Imbalances in Prognostic Factors at Randomization (ITT Population)

FDA Reviewer's Sensitivity Analysis to Imbalances in Prognostic Factors at Randomization		
Sensitivity analysis for assessing the impact of imbalances in baseline positive autoantibody counts	**Hazard ratio (95% CI)**	0.44 (0.22, 0.87)
	Nominal P-value	0.0176
Sensitivity analysis for assessing the impact of imbalances in baseline BMI	**Hazard ratio (95% CI)**	0.41 (0.21, 0.78)
	Nominal P-value	0.0064
Sensitivity analysis for assessing the impact of both imbalance in baseline positive autoantibody counts and imbalance in baseline BMI	**Hazard ratio (95% CI)**	0.42 (0.21, 0.84)
	Nominal P-value	0.0137

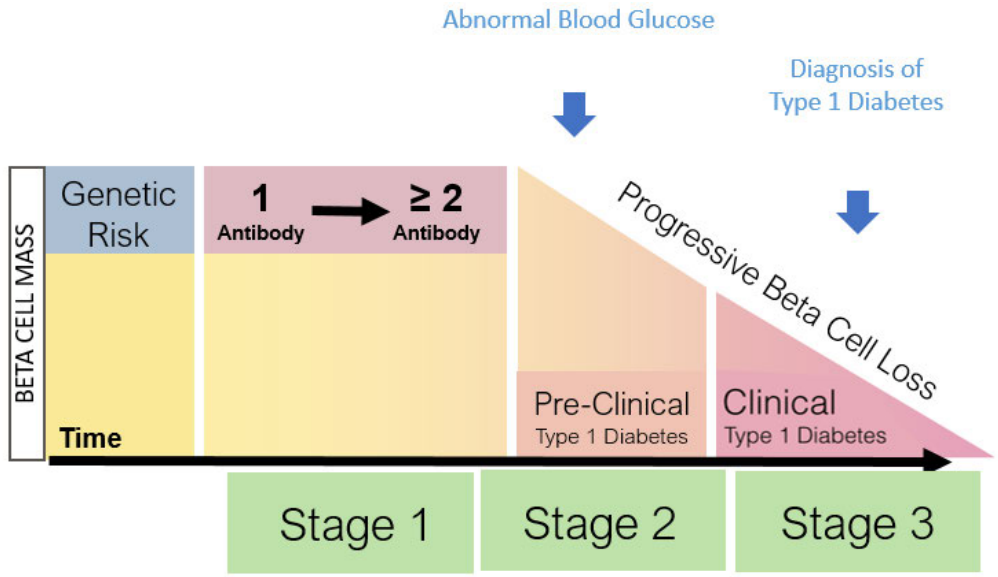
Source: FDA statistical reviewer

In summary, because the amount of missing data was small, sensitivity analyses performed using different missing data handling approaches show that the observed efficacy by the primary analysis method is sufficiently robust to withstand conservative approaches in handling missing data. In addition, the results of the sensitivity analysis, adjusting for autoantibody count at randomization, and BMI show that the baseline imbalances in these characteristics did not change the primary analysis conclusion.

3.3 Evaluation of Efficacy – Meta-analysis of Stage 3 Studies C-peptide Data

For its marketing application, the Applicant intended to rely on data from the TN-10 study for demonstration of efficacy for the proposed indication. While FDA agreed that TN-10 may be considered an AWC trial that supports the efficacy of teplizumab, FDA noted that confirmatory evidence would also be needed to provide substantial evidence of effectiveness. The Applicant proposed to leverage available data from teplizumab studies in new-onset T1D as supportive evidence and justified the use of C-peptide in these patients in the context of the T1D disease progression model shown in Figure 3.

Figure 3. Meta-analyses: Proposed Type 1 Diabetes Disease Model



Source: Created by FDA clinical reviewer; adapted from Insel 2015

3.3.1 Post Hoc Analysis Plan and Primary Endpoint

3.3.1.1 Meta-analysis study selection – A Post Hoc Analysis Plan

For supportive efficacy, meta-analyses of the change from baseline in C-peptide AUC were performed with the inclusion of data from AbATE (Phase 2), Delay (Phase 2), Protégé (Phase 2/3), Encore (Phase 3), and Study 1 (Phase 1/2). Study inclusion in the meta-analyses was based on the patient population (newly diagnosed T1D, within 6 weeks of diagnosis), the study design (i.e., randomized, controlled or open-label standard of care), and assessment of C-peptide as measured by a 4-hour mixed-meal tolerance test (MMTT). Subject-level data from five new-onset T1D studies were selected to be pooled for the meta-analyses. The studies are outlined in Table 1 and described in more detail below.

3.3.1.1.1 Protégé

The Protégé study was a dose-ranging study designed to assess the safety and efficacy of three teplizumab dosing regimens (a full 14-day course, a one-third dose 14-day course, and a 6-day course) of daily intravenous (IV) infusions compared with placebo (IV) given as two identical courses, 26 weeks apart. The study enrolled patients with newly diagnosed T1D who had experienced disease onset within the past 12 weeks, and subjects underwent a 4-hour MMTT every 6 months for 2 years. The primary endpoint for Protégé was the proportion of subjects who had both a total daily insulin dose of less than 0.5 units/kg/day and an HbA1c <6.5% at one year after randomization. The study was terminated early due to futility on its primary endpoint.

3.3.1.1.2 Encore

The Encore study was intended as a replicate study of Protégé. The primary endpoint for Encore was the proportion of subjects who had both a total daily insulin dose <0.25 units/kg/day and HbA1c <7.0% at one year after randomization. The Encore study was terminated prematurely with 254 subjects (64% of the planned 400 subjects) enrolled after Protégé failed to demonstrate efficacy on the primary endpoint.

3.3.1.1.3 Study 1

Study 1 was a randomized, controlled, open-label study to test the safety and efficacy of a single course of teplizumab on the loss of insulin production over 2 years in patients with newly diagnosed T1D (within 6 weeks) compared to standard of care. Patients returned for 4-hour MMTT every 6 months. The primary endpoint was change from baseline in the mean C-peptide 4-hour AUC in response to an MMTT at Month 24.

3.3.1.1.4 AbATE

The AbATE study was a randomized, open-label study that enrolled individuals diagnosed with T1D within 8 weeks of randomization. The study drug treatment group received two courses of teplizumab: at randomization and at 12 months. Patients underwent a 4-hour MMTT every 6 months for 2 years. The parallel group control received standard of care. The primary endpoint was change from baseline in the mean C-peptide 4-hour AUC in response to an MMTT at Month 24.

3.3.1.1.5 Delay

The Delay study was a randomized, double-blind, placebo-controlled study that enrolled patients 4 to 12 months after their T1D diagnosis. Patients received one course of teplizumab. Patients underwent a 4-hour MMTT every 6 months for 1 year. The primary endpoint was C-peptide AUC in response to an MMTT at Month 12.

3.3.1.2 Primary endpoint for the MA analysis – C-peptide

C-peptide was assessed in a similar manner in each study. 4-hour MMTTs were performed and measurements were taken at baseline, 15 minutes, and then every 30 minutes thereafter for 4 hours. Three of the five studies assessed C-peptide up to 2 years; the Delay study was designed for only 1 year of follow-up, and Encore was terminated early to provide C-peptide AUC data up to 1 year.

Table 1 summarizes key design features of the five studies, including the dosing regimen of teplizumab and C-peptide assessment schedule. To support the meta-analyses, the original study visits from the individual studies were unified to provide common windows.

Because AUC is typically log-normally distributed, these values were transformed to log(AUC+1) for analyses. The primary endpoint for the meta-analyses was the change from baseline in the C-peptide AUC. Analyses were performed on the log-transformed data, log(AUC+1).

3.3.2 Statistical Methodologies – A Post Hoc Analysis Plan

3.3.2.1 Individual Study Analyses

For each study, after multiple imputation (MI) was used to fill in missing values, an analysis of covariance (ANCOVA) was used to estimate the treatment effect. The dependent variable was the change from baseline in log (AUC+1) at the last study visit (2 years for Protégé, AbATE, Study 1 in one meta-analysis, and 1 year for all five studies in another meta-analysis). Covariates were treatment, log (baseline C-peptide AUC+1), and age.

3.3.2.2 Meta-analyses

By pooling individual participant data from the five completed studies, a one-stage fixed-effects model that combined and re-analyzed all studies simultaneously was used by the Applicant. On the basis of the individual study analysis ANCOVA model, the meta-analysis model included two additional terms: a fixed-effect term indicating study identities, and a term for treatment-by-study interaction.

In their original plan for the one-stage meta-analyses, the Applicant proposed a random-effects approach analysis. The Applicant explained, however, that because all five completed randomized studies to date were included in the meta-analyses, the purpose of the analyses was to assess the effect of these specific studies on the outcome, such that fixed-effects models would be more appropriate for the meta-analyses. The fixed-effects approach assumes that the between-study variance is zero, which is too strong an assumption that rarely holds. Instead, for each analysis time frame, FDA carried out a two-stage random-effects meta-analysis to estimate the overall treatment effect which accounts for the between-study variances.

3.3.2.3 Handling of Missing Data and Sensitivity Analysis

The amount, pattern, and reasons for missing data were to be described for each study. Multiple imputation was to be used for the primary analysis. The pattern mixture MI for missing data was based on two missing data patterns. For patients with missed assessment(s) due to administrative reasons, the missing data were imputed based on data from completers. For patients with missing data due to non-administrative reasons such as adverse events (AE), MI was based on data from retrieved dropouts. Retrieved dropouts are subjects in the same study who discontinued treatment but completed the studies and had C-peptide measured at the scheduled 1- or 2-year timepoints.

After MI had filled in missing values based on the approach described above, the modified dataset was to be used in analyses. Rubin's rules for combining data from imputed datasets were used.

3.3.3 Patient Disposition, Demographics, and Baseline Characteristics

Table 9 summarizes disposition by individual and pooled studies. Data from a total of 609 patients were included in the meta-analyses. Among them, 92.9% of patients completed planned treatment course 1, and 66.7% of patients completed planned treatment course 2.

Table 9. Meta-Analyses: Study Disposition

	Teplizumab N=375 n (%)	Control N=234 n (%)	Total N=609 n (%)
Number of Patients by study			
AbATE	52 (13.9)	25 (10.7)	77 (12.6)
Delay	32 (8.5)	28 (12.0)	60 (9.9)
Encore	63 (16.8)	62 (26.5)	125 (20.5)
Protege	207 (55.2)	98 (41.9%)	305 (50.1)
Study1	21 (5.6)	21 (9.0)	42 (6.9)
Treatment Status in the Pooled Study			
Completed planned treatment course 1	337 (89.9)	229 (97.9)	566 (92.9)
Patients assigned to a two-course regimen	322	185	507
Completed planned treatment course 2	211 (65.5)	127 (68.6)	338 (66.7)

Source: FDA statistical reviewer.

Table 10 displays demographics in the meta-analysis population. There were multiple countries involved in the five studies, so we grouped other countries into one category. Overall, baseline characteristics were balanced across treatment groups.

Table 10. Meta-Analyses: Pooled Demographics

		Teplizumab	Control	Total
		N=375	N=234	N=609
Age (Years)	Mean (SD)	16.9 (7.1)	16.6 (7.2)	16.8 (7.1)
	Median (Minimum, Maximum)	15 (7, 35)	14 (8, 35)	14 (7, 35)
Sex, n (%)	Female	145 (38.7)	86 (36.8)	231 (37.9)
	Male	227 (60.5)	148 (63.2)	375 (61.6)
	Unknown	3 (<1)	0	3 (<1)
Race, n (%)	American Indian or Alaska Native	2 (<1)	0	2 (<1)
	Asian	87 (23.2)	50 (21.4)	137 (22.5)
	Black or African American	4 (1.1)	4 (1.7)	8 (1.3)
	Other	5 (1.3)	1 (<1)	6 (<1)
	Unknown	2 (<1)	0	2 (<1)
Country, n (%)	US	186 (49.6)	122 (52.1)	308 (50.6)
	Canada	6 (1.6)	1 (<1)	7 (1.1)
	Other	183 (48.8)	111 (47.4)	294 (48.3)

Source: FDA statistical reviewer.
Abbreviations: SD, standard deviation

3.3.4 Results and Conclusions

For the pooled analysis, Table 11 summarizes mean AUC 4-hour MMTT C-peptide (nmol/L) data availability at Months 12 and 24. Because of the early termination of Encore, only 27.2% of the Encore study population provided Month 12 C-peptide data; approximately 80% of the pooled population provided Month 12 C-peptide data. At the end of Year 2, 64% of C-peptide data were followed-up in study Protégé; 70% of the pooled population (for the 2-year analysis) provided C-peptide data up to Month 24.

APPEARS THIS WAY ON ORIGINAL

Error! Reference source not found. displays both individual study analysis results and the meta-analysis results based on the imputed data.

For the individual studies, an ANCOVA model included the change from baseline C-peptide log(AUC+1) at the 1-year time point (Day 364) or 2-year time point (Day 728) as the dependent

variable and baseline C-peptide log(AUC+1), age, and treatment as covariates. At 1-year, Study 1 and AbATE observed larger treatment effects than Delay. Protégé and Encore failed to establish efficacy in C-peptide preservation.

Table 11. Meta-Analyses: C-peptide Data Availability by Unified Visit Windows

Study	Baseline N	Month 12 n (%)	Month 24 n (%)
Study 1	42	42 (100)	33 (78.6)
AbATE	77	72 (93.5)	70 (90.9)
Protégé	305	279 (91.5)	195 (63.9)
AbATE	77	72 (93.5)	70 (90.9)
Delay	60	58 (96.7)	
Encore	125	34 (27.2)	
Pooled population for 1-year analysis	609	485 (79.6)	
Pooled population for 2-year analysis	424		298 (70.3)

Source: The Applicant. FDA statistical reviewer replicated the analysis.

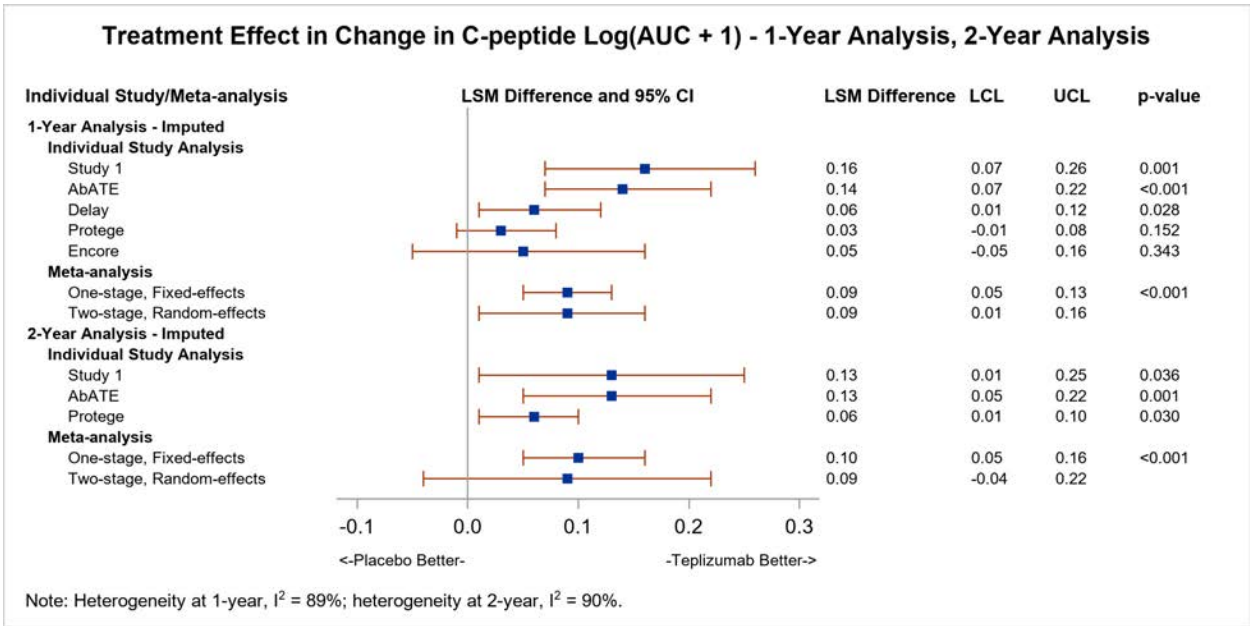
In the Applicant's one-stage fixed-effects meta-analyses, data from the five studies including 1-year follow-up data or three studies including 2-year follow-up data were pooled together. Two additional terms were added to the individual study model: a study ID variable and a study by treatment interaction term. The Applicant's one-stage fixed-effects meta-analyses were reproduced in **Error! Reference source not found.**: at the 1-year follow-up, compared with the control arm, the LS mean difference in change from baseline log(Mean C-peptide AUC+1) was 0.09 (95% CI: 0.05 to 0.13, p<0.0001), favoring teplizumab. At the 2-year follow-up, compared with the control arm, the LS mean difference in change from baseline was 0.10 (95% CI: 0.05 to 0.16, p<0.0001), favoring teplizumab. Analyses based on the observed data show similar results (see Figure 11 in Appendix A). More detailed analysis results are summarized in Table 12 to Table 15 in Appendix A.

APPEARS THIS WAY ON ORIGINAL

We observed high heterogeneity among the individual studies, in both the 1-year and 2-year analyses. The statistics I^2, which is often used to measure heterogeneity, was 89% for the 1-year meta-analysis; and 90% for the 2-year meta-analysis. To account for this unignorable between-study variability, for each analysis time frame, we performed a two-stage random-effects meta-analysis to estimate the overall treatment effect. The two-stage random-effects meta-analyses calculated the weighted average of treatment effect estimates from the individual studies, where the weights were given by the inverses of the sum of within study and between-study variances components. At both 1-Year and 2-Year, the overall treatment effects were similar in size with the one-stage fixed effect model estimates; however, the associated confidence intervals were wider and the estimated overall treatment effect at 2-year was not statistically significant. Specifically, the 1-year random effect meta-analysis results suggest that teplizumab improves preservation of C-peptide by 0.09 (95% CI: 0.01 to 0.16) on the scale of log (Mean C-peptide AUC+1), as compared with control. At 2 years, because there were only three studies, and

because of nonignorable variation in effects between studies, the estimation of the overall treatment effect was associated with a large variance and the 2-year overall treatment effect, 0.09 (95% CI: -0.04 to 0.22), was not statistically significant.

Figure 4. Meta-Analyses: Change from Baseline in Log (Mean C-peptide AUC +1) – Imputed Data



Source: FDA statistical reviewer.
Abbreviations: AUC, area under curve; CI, confidence interval; LSM, least squares mean; LCl, lower confidence limit; UCl, upper confidence limit; MMTT, mixed meal tolerance test

3.4 Evaluation of Safety

Please refer to clinical review by Dr. Lauren Wood Heickman for details on safety analysis results for teplizumab.

3.5 Benefit-Risk Assessment

Given statistically significant treatment effect of teplizumab in delaying T1D onset in at-risk patients, the numerical trend in slowing the rate of C-peptide decline in Stage 3 new onset T1D patients, and comparable safety profile between teplizumab and placebo, we think the benefit-risk assessment for teplizumab is favorable from statistical perspective.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Study TN-10

The Applicant pre-planned subgroup analyses for the primary efficacy endpoint, with subgroups based on age (8 to 17, and 18 to 45 years), sex, race/ethnicity, autoantibody type and titer at baseline, OGTT in the diabetic range at baseline (yes/no), and HLA risk categories. Forest plots are used to display the extent to which the estimated treatment effect differed across various subgroups. For each subgroup factor, the model was adapted from the pre-specified primary efficacy analysis model. The difference in the treatment effect between/among subgroups was tested using an interaction analysis, performed by including the subgrouping variable and subgrouping variable-by-treatment interaction as covariates to the primary analysis model. When a covariate in the model was the subgrouping variable (e.g., age), it was replaced with the categorical version of itself when needed. By-subgroup hazard ratios were estimated to illustrate the treatment effects under each subgroup.

There was some random highs and random lows in sample estimates of subgroup treatment effects due to small sample size and large variability for some subgroups. Therefore, we also derive shrinkage estimates of subgroup treatment effects using a Bayesian hierarchical model based on summary sample estimates. The total variability in the sample estimates is the sum of the within subgroup variability of the sample estimator and the across subgroups variability in underlying/true parameter values. A shrinkage estimate of the subgroup treatment effect, which borrows information from the other subgroups while estimating the treatment effect for a specific subgroup, is a "weighted" average of the sample estimate and overall estimate. We used similar flat priors to derive shrinkage estimates for the subgroups. The Bayesian hierarchical model assumptions are:

For $i = 1, 2\ldots$ Y_i represents the observed sample estimate of treatment effect in a subgroup level i, assume $Y_i \sim N(\mu_i, \sigma_i^2)$ where Y_i denotes log (HR_i) and

- σ_i^2 are the observed variance for sample estimates
- $\mu_i \sim N(\mu, \tau^2)$
- $\mu \sim N(C, 4)$, note the selection of the prior mean C for μ depends on the overall population based on which the subgroup analysis was carried out

- $1/\tau^2$ ~ Gamma (0.001, 0.001)

For shrinkage analyses performed with respect to Age and Sex, a flat prior with mean -0.887 and variance of 4 was chosen for μ. Note that -0.887 is the estimated log(HR) from the primary analysis model based on the overall population.

A flat prior with mean -0.882 and variance of 4 was chosen for shrinkage analyses with respect to positive autoantibody count at baseline as the overall treatment effect was based on patients who had at least 2 positive autoantibodies at baseline (N=75).

A flat prior with mean -0.724 and variance of 4 was chosen for shrinkage analyses with respect to BMI z-score tertile groups as the overall treatment effect was based on patients who were less than 20 years of age (N=57).

The results of the sample estimates and the shrinkage estimates of treatment effects in the same subgroups, are presented in Figure 5 to Figure 9, respectively.

4.1.1 Gender, Race, Age, and Geographic Region

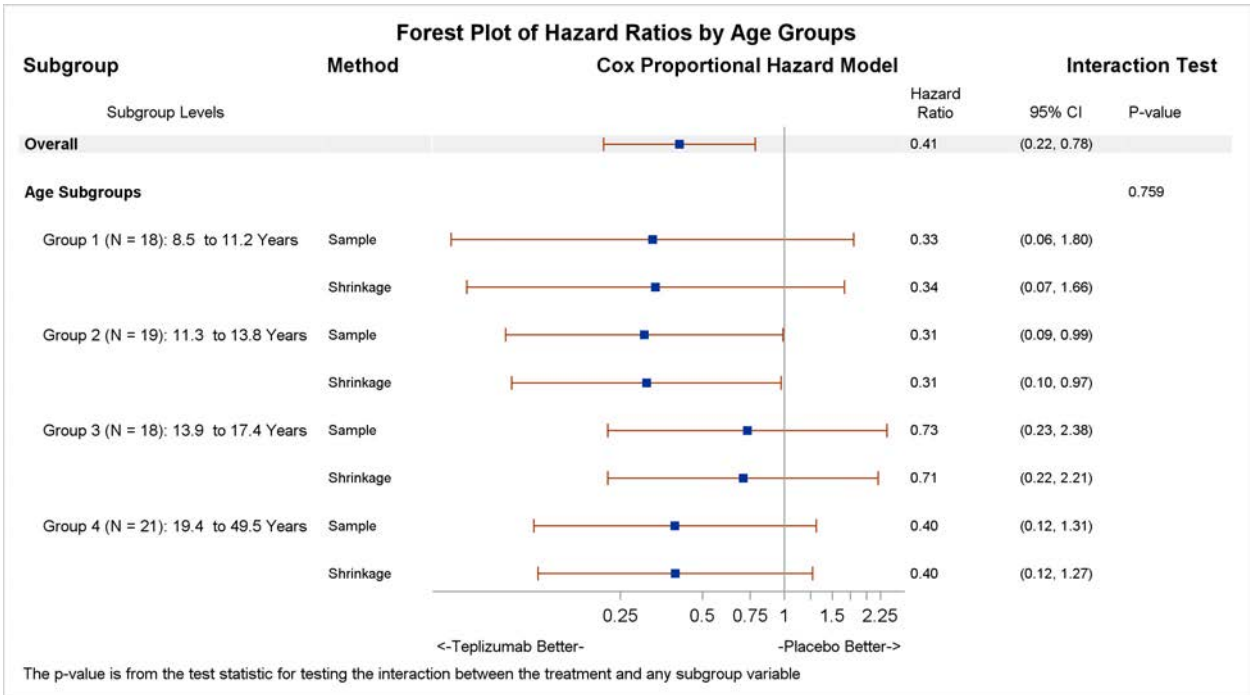
There were only two patients whose race was not White and four patients who participated at sites out of the US, hence subgroup analyses on race or geographical region were not carried out.

Pediatric patients (8 to 17 years of age) made up the majority (72%) of the TN-10 population. Therefore, we further divided the pediatric patient population into three groups using tertiles of the study pediatric population age distribution (see **Error! Reference source not found.** for patient counts and distributions among the groups).

In Figure 5, for the overall population and each of the four age subgroups, the estimated hazard ratio of T1D onset, representing treatment effect for teplizumab compared to placebo, and its 95% CI is plotted. The label indicates that points to the left of the vertical line (hazard ratio = 1) favor teplizumab, while those to the right favor placebo. All age subgroup CIs have substantial overlap with the overall point estimate, and CIs at the top indicate a consistency of findings across subgroups. No significant interaction was found between treatment and age subgroups at the 5% level (interaction test, p=0.758).

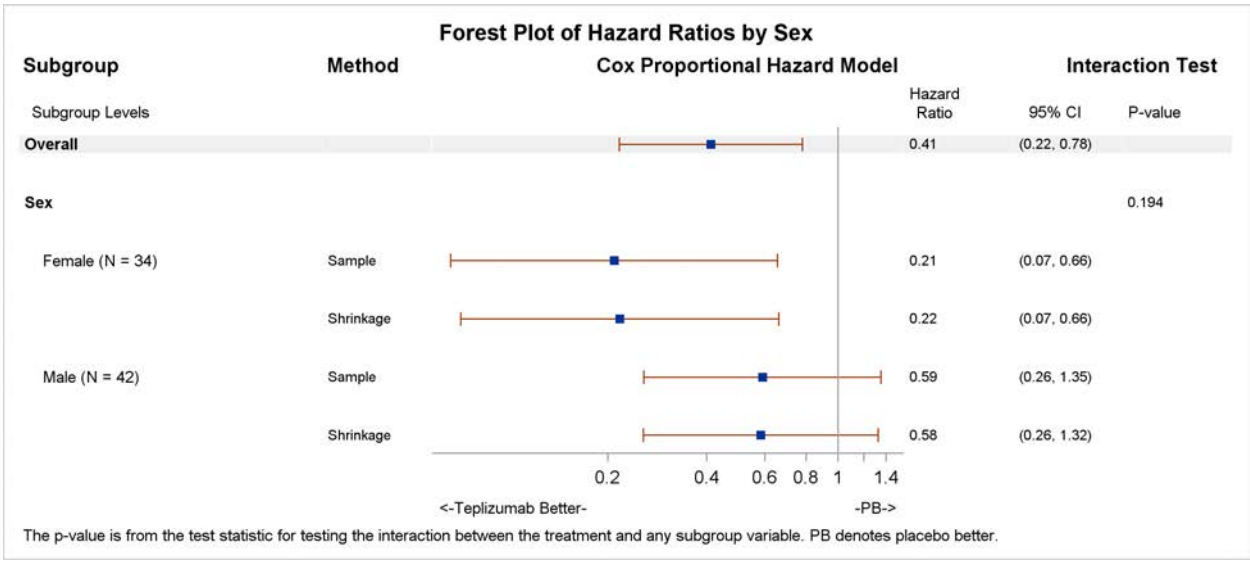
In Figure 6, no significant interaction was found between treatment and sex at the 5% level (interaction test, p=0.24), indicating a consistency in effect size between males and females.

Figure 5. TN-10: Subgroup Analysis – Age groups



Source: FDA statistical reviewer

Figure 6. TN-10: Subgroup Analysis – Sex



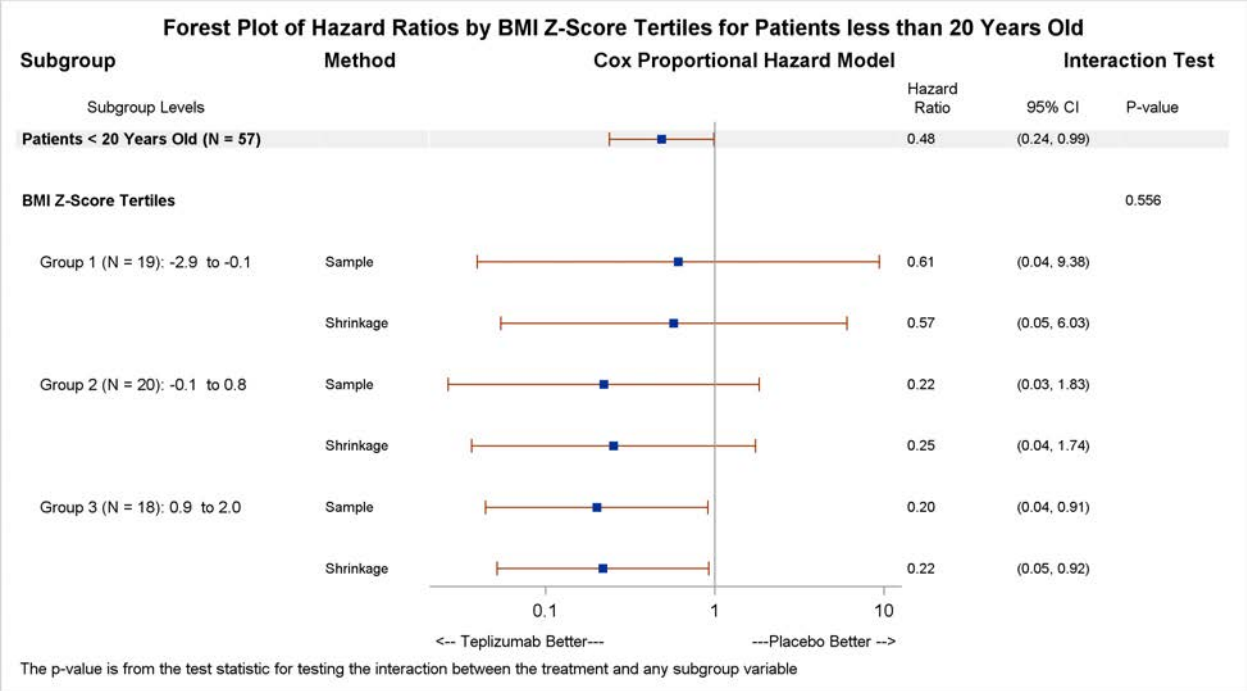
Source: FDA statistical reviewer

4.1.2 Other Special/Subgroup Populations

In light of the pre-existing suggestion of increased risks of T1D-related baseline characteristics, subgroup analyses by BMI z-score scores for children under 20 years old, baseline positive autoantibodies, and HLA risk categories are also of importance to inform the potential effect of these factors on time to T1D onset. We have also examined the modulating potential of baseline beta cell reserve on time to T1D onset through a subgroup analysis with subgroups defined by the median of OGTT C-peptide mean AUC at baseline.

BMI z-scores were derived by the clinical reviewer of this trial. The less than 20 years old group was divided by tertiles. Figure 7 shows that while the CI overlap with each other, there is a trend that larger effect size (smaller HR, 0.17) is associated with larger BMI z-scores. No significant interaction was found between treatment and BMI z-score subgroups at the 5% level (interaction test p-value = 0.481). The three CIs overlap suggesting there is insufficient evidence to claim an interaction between treatment and BMI z-score group.

Figure 7. TN-10: Subgroup Analysis – BMI-Z scores



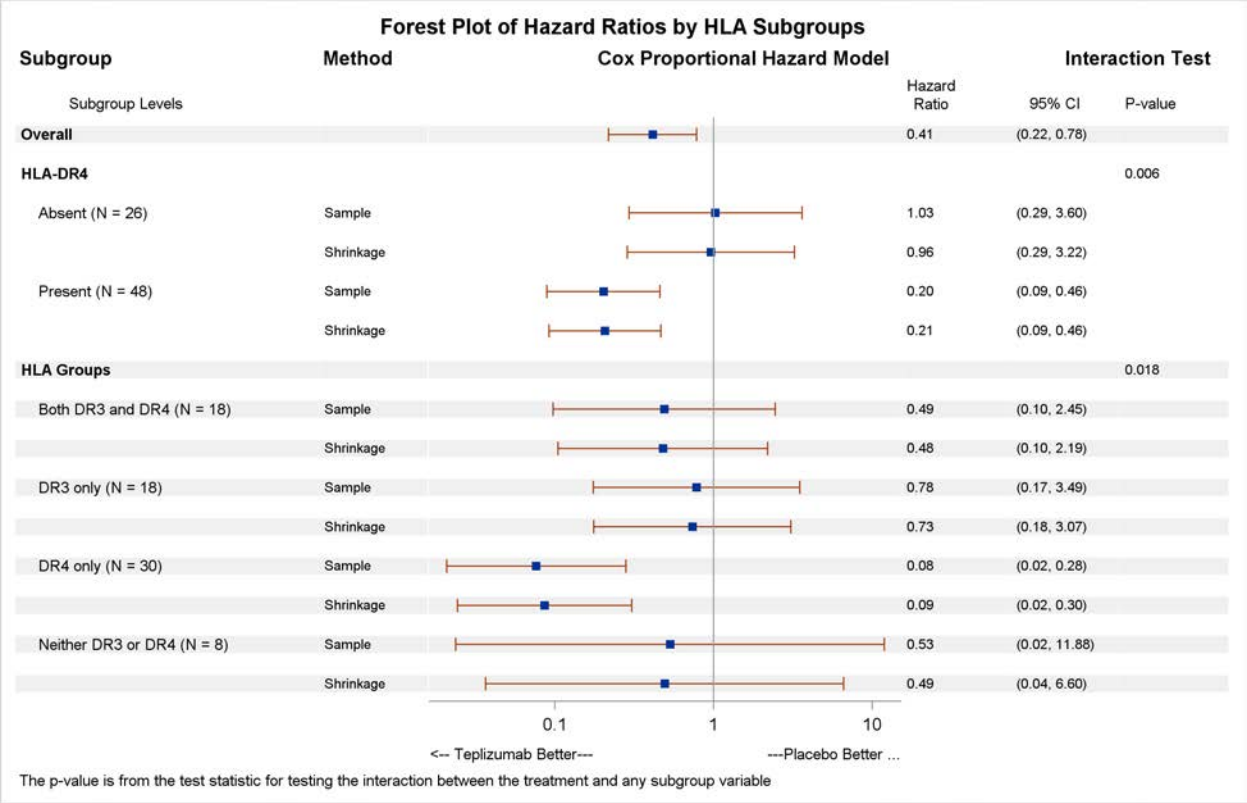
Source: FDA statistical reviewer

In the subgroup analysis by HLA-DR4 (top part of Figure 8), the two subgroups are HLA-DR4 present or absent. Presence of HLA-DR4 is associated with a bigger treatment effect (HR=0.2, 95% [0.09, 0.46], interaction test p-value = 0.006). Note that while the by sample analysis is indicative of no treatment benefit in HLA-DR4 absent patients (HR=1.03, 95% [0.29, 3.6]), the shrinkage method has an HR = 0.96 for the HLA-DR4 absent patients; the shrinkage analysis shows that treatment benefit is largely consistent across the subgroups and consistent with the overall treatment effect.

The p-value (0.018) is also numerically significant for the HLA risk categories. This significant interaction test indicates some quantitative differences in treatment effects across the HLA subgroups.

The treatment effects for subjects with and without HLA DR3 were similar and not plotted.

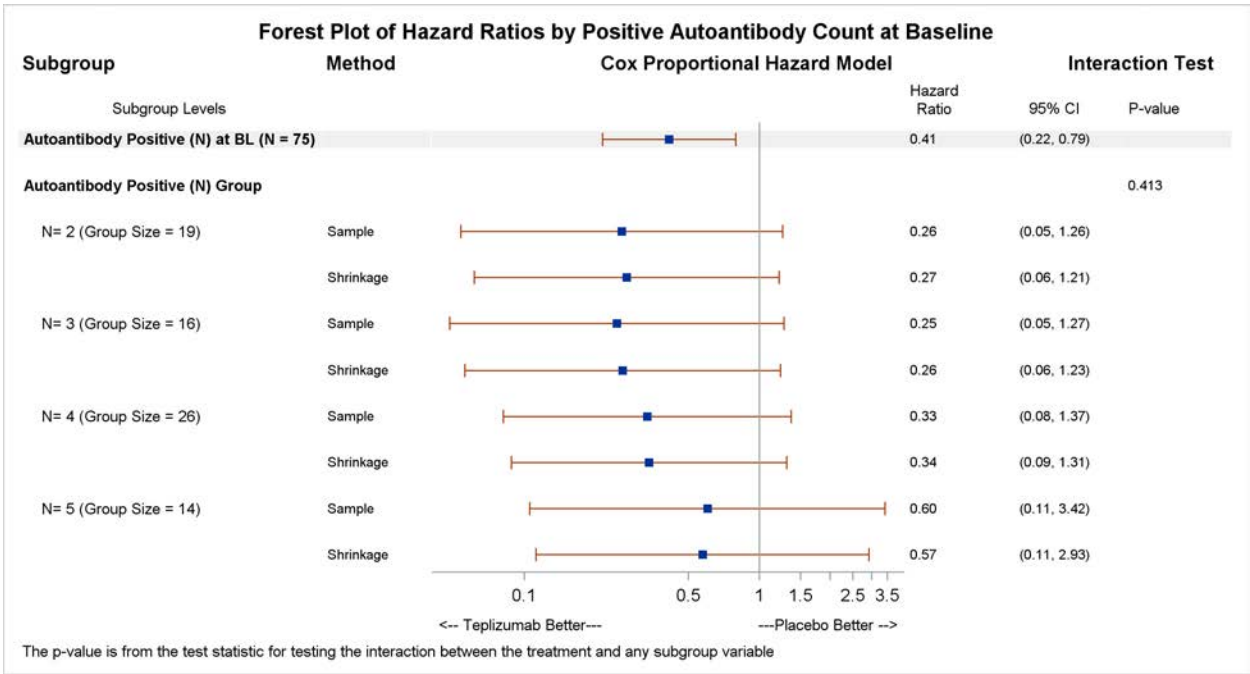
Figure 8. TN-10: Subgroup Analysis – HLA-Groups



Source: FDA statistical reviewer

The treatment effect by positive autoantibody count at baseline shows a trend that larger effect size is associated with subjects with fewer positive autoantibodies at baseline. However, the confidence intervals overlap each other and the p-value for the interaction test was insignificant, which suggests lack of evidence that treatment effects differ across various positive autoantibody count at baseline.

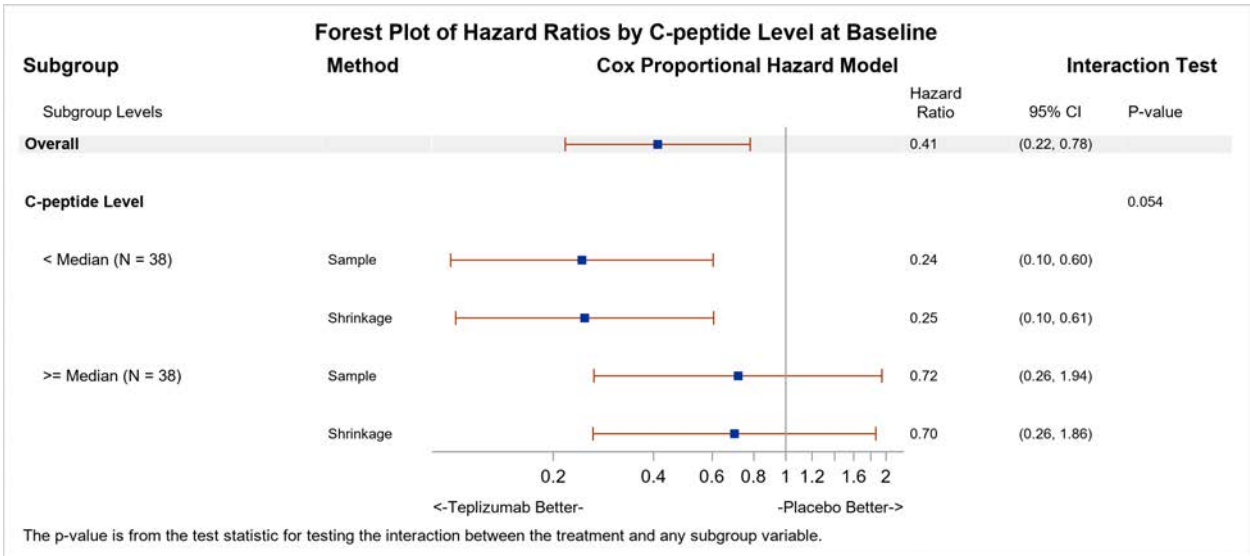
Figure 9. TN-10: Subgroup Analysis – Baseline Autoantibody Positive Count



Source: FDA statistical reviewer

In the subgroup analysis by OGTT C-peptide mean AUC level at baseline, there is a trend showing that lower level of C-peptide may be associated with a bigger treatment effect (HR=0.24, 95% CI [0.10, 0.60], interaction test p-value = 0.054) while the estimated hazard ratio for the higher C-peptide level group is 0.72 (95% CI [0.26, 1.94]).

Figure 10. TN-10: Subgroup Analysis – Baseline OGTT Mean AUC C-peptide Level



Source: FDA statistical reviewer

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

5.1.1 TN-10

The review team identified the following statistical issues in Study TN-10.

- TN-10 was an investigator-initiated trial with small sample size.
- Evaluation of the treatment effect of teplizumab on C-peptide AUC change from baseline was not possible because of study design issues.
- Imbalances in prognostic factors at randomization.

FDA conducted adjusted analyses to address the imbalances in prognostic factors at randomization.

5.1.1.1 Imbalances in Prognostic Factors at Randomization

To assess the impact of imbalances in baseline BMI and positive autoantibody count, two factors that may have prognostic potential, we conducted a covariate adjustment analysis for each factor by adding the factor as an additional covariate to the primary analysis model. Both the two sensitivity analyses produced results similar to those in the primary analysis in effect size and significance level.

In summary, the treatment effect was still statistically significant after adjusting for the impact imbalances in baseline prognostic factors.

5.1.2 Meta-Analyses

The FDA review team identified the following statistical issues in the meta-analyses that impact its strength in providing the confirmatory evidence as intended.

- Change in C-peptide AUC is not a validated surrogate endpoint for clinical benefit, e.g. delay of progression of T1D, but literature reviews and inputs from some key opinion leaders suggest it may likely predict clinical benefit. Scientific knowledge and clinical considerations both played crucial roles in how C-peptide AUC data was interpreted.
- The meta-analysis was based on clinical T1D patients rather than at-risk patients. The applicability of evidence of C-peptide preservation among clinical T1D patients to support a delay in onset of T1D is conjectured but unconfirmed.
- The meta-analysis was proposed post hoc.

- There was notable heterogeneity in the estimated treatment effect in C-peptide preservation across the five trials included in the meta-analyses, with the two largest trials Protégé and Encore showing a small treatment effect.

Among them, the heterogeneity issue was addressed by applying a random-effects meta-analysis model in estimating the true treatment effect.

5.1.2.1 Heterogeneity in Effect Sizes between Studies

In FDA's review of the meta-analyses on the treatment effect of teplizumab in slowing the rate of C-peptide decline in Stage 3 new onset T1D patients, to account for the identified high heterogeneity between individual effect sizes, FDA conducted two-stage random-effects meta-analyses in estimating the true treatment effects at both one-year and two-year.

The random-effects approach estimated treatment effects were similar in size with the one-stage fixed-effects model estimates; however, the associated confidence intervals were wider and the estimated overall treatment effect at 2-year was not statistically significant.

5.2 Collective Evidence

Study TN-10, an adequate and well-controlled trial, successfully demonstrated statistically significant treatment effect of teplizumab in delaying T1D onset in at-risk patients. The estimated treatment effect was robust to both sensitivity analyses to evaluate impact of missing data and imbalances in prognostic factors at randomization. Treatment effect was in general consistent among baseline demographic and disease characteristic subgroups with quantitative difference observed among HLA subgroups.

Literature reviews and inputs from some key opinion leaders suggest change in C-peptide AUC may likely predict clinical benefit and provide confirmatory evidence for efficacy of a study treatment in delaying T1D onset in at risk patients. However, statistical issues identified in the meta-analyses of change in C-peptide AUC for teplizumab vs. comparator (Section 5.1.2) may impact its strength in providing the confirmatory evidence in this specific case. The meta-analyses of five studies conducted in Stage 3 new onset T1D patients demonstrated a numerically significant treatment effect of teplizumab vs. comparator on C-peptide AUC change from baseline at 1-year after accounting for variability in treatment effects across studies; the estimated treatment effect at 2 years from the meta-analysis was associated with a wide confidence interval and was not statistically significant.

5.3 Conclusions and Recommendations

According to FDA's evidentiary standard for effectiveness, the one adequate and well-controlled trial (TN-10) provided statistically significant evidence in delaying T1D onset in Stage 2 at-risk patients; the confirmatory evidence consisted of the post-hoc meta-analyses of 5 studies in new-

onset T1D patients showed numerically significant treatment effect of teplizumab vs. comparator on C-peptide AUC change from baseline at 1-year, but not 2-year. The application was brought to an advisory committee on May 27, 2021, where the panel voted 10-7 for approval. Taking into consideration the clinical review team's conclusion, we recommend approval of the proposed indication: the delay of clinical T1D in at-risk (Stage 2) individuals.

5.4 Labeling Recommendations

There are issues identified by other line functions including clinical pharmacology and product quality, which will likely result in an action of complete response for this submission. The review team decided not to review and negotiate label in this round of submission. Therefore, we will not provide any labeling recommendation in this statistical review.

References

Allison, Paul D. 1995. *Survival Analysis Using SAS: A Practical Guide.* Cary, NC: SAS.

Barker, Jennifer M, Kim McFann, Leonard C Harrison, Spiros Fourlanos, Jeffrey Krischer, David Cuthbertson, H Peter Chase, and George S Eisenbarth. 2007. "Pre-type 1 diabetes dysmetabolism: maximal sensitivity achieved with both oral and intravenous glucose tolerance testing." *The Journal of Pediatrics* 31-36.e6.

Borenstein, Michael, Larry V. Hedges, Julian P.T. Higgins, and Hannah R. Rothstein. 2010. "A basic introductin to fixed-effect and random-effecs models for meta-analysis." *Research Synthesis Methods* 97-111.

Gorus, Frans K, Eric C Balti, Anissa Messaaoui, Simke Demeester, Annelien Van Dalem, Olivier Costa, Harry Dorchy, et al. 2017. "Twenty-year progression rate to clinical onset according to autoantibody profile, age, and HLA-DQ genotype in a registry-based group of children and adults with a first-degree relative with type 1 diabetes." *Diabetes Care* 1065-1072.

Insel, Richard A. 2015. "Staging Presymptomatic Type 1 Diabetes: A Scientific Statement of JDRF, the Endocrine Society, and the American Diabetes Association." *Diabetes Care* 1964-1974.

Jacobsen, Laura M, Laura Bocchino, Carmella Evans-Molina, Linda DiMeglio, Robin Goland, Darrell M Wilson, Mark A Atkinson, et al. 2020. "The risk of progression to type 1 diabetes is highly variable in individuals with multiple autoantibodies following screening." *Diabetologia* 588-596.

Lan, K. K. Gordon, and David L. DeMets. 1983. "Discrete sequential boundaries for clinical trials." *Biometrika* 659-63.

Maahs, David M, Nancy A West, Jean M Lawrence, and J Elizabeth Mayer-Davis. 2010. "Epidemiology of type 1 diabetes." *Endocrinology and Metabolism Clinics of North America* 481-497.

Steck, Andrea K, Kendra Vehik, Ezio Bonifacio, Ake Lernmark, Anette-G Ziegler, William A Hagopian, JinXiong She, et al. 2015. "Predictors of progression from the appearance of islet antoantibodies to early childhood diabetes: the environmental determinants of diabetes in the young (TEDDY)." *Diabetes Care* 808-813.

Viner, R M Viner, P C Hindmarsh, B Taylor, and T J Cole. 2008. "Childhood body mass index (BMI), breastfeeding and risk of Type 1 diabetes: findings from a longitudinal national birth control." *Diabetic Medicine* 1056-1061.

Ziegler, Anette G, Marian Rewers, Olli Simell, Tuula Simell, Johanna Lempainen, Andrea Steck, Christiane Winkler, et al. 2013. "Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children." *JAMA* 2473-9.

APPENDICES

Appendix A

Table 12. Meta-Analyses: Change from Baseline in Log (Mean C-peptide AUC +1) at the End of Year One – Imputed Data

Study	Treatment	Baseline			Change from baseline at the end of Year 1				Comparison vs. Control				
		Number Analyzed	Mean	SD	LS Mean	SE	95% CI		LSM Difference	SE	95% CI		p-value
							Lower	Upper			Lower	Upper	
Individual Study Analysis													
Study 1	Control	21	0.39	0.03	-0.18	0.03	-0.24	-0.11					
	Teplizumab	21	0.39	0.03	-0.02	0.03	-0.08	0.05	0.16	0.05	0.07	0.26	0.001
AbATE	Control	25	0.50	0.03	-0.23	0.03	-0.29	-0.17					
	Teplizumab	52	0.53	0.02	-0.08	0.02	-0.12	-0.04	0.14	0.04	0.07	0.22	<0.001
Delay	Control	28	0.44	0.04	-0.14	0.02	-0.18	-0.11					
	Teplizumab	32	0.47	0.03	-0.08	0.02	-0.12	-0.04	0.06	0.03	0.01	0.12	0.028
Protege	Control	98	0.47	0.02	-0.08	0.02	-0.12	-0.04					
	Teplizumab	207	0.47	0.02	-0.04	0.01	-0.07	-0.01	0.03	0.02	-0.01	0.08	0.152
Encore	Control	62	0.43	0.03	-0.12	0.04	-0.19	-0.05					
	Teplizumab	63	0.47	0.04	-0.07	0.03	-0.14	0.00	0.05	0.05	-0.05	0.16	0.343
Meta-analysis: One-stage, Fixed-effects													
5-Studies	Control	234	0.45	0.01	-0.14	0.01	-0.17	-0.11					
	Teplizumab	375	0.47	0.01	-0.05	0.01	-0.08	-0.03	0.09	0.02	0.05	0.13	<0.001

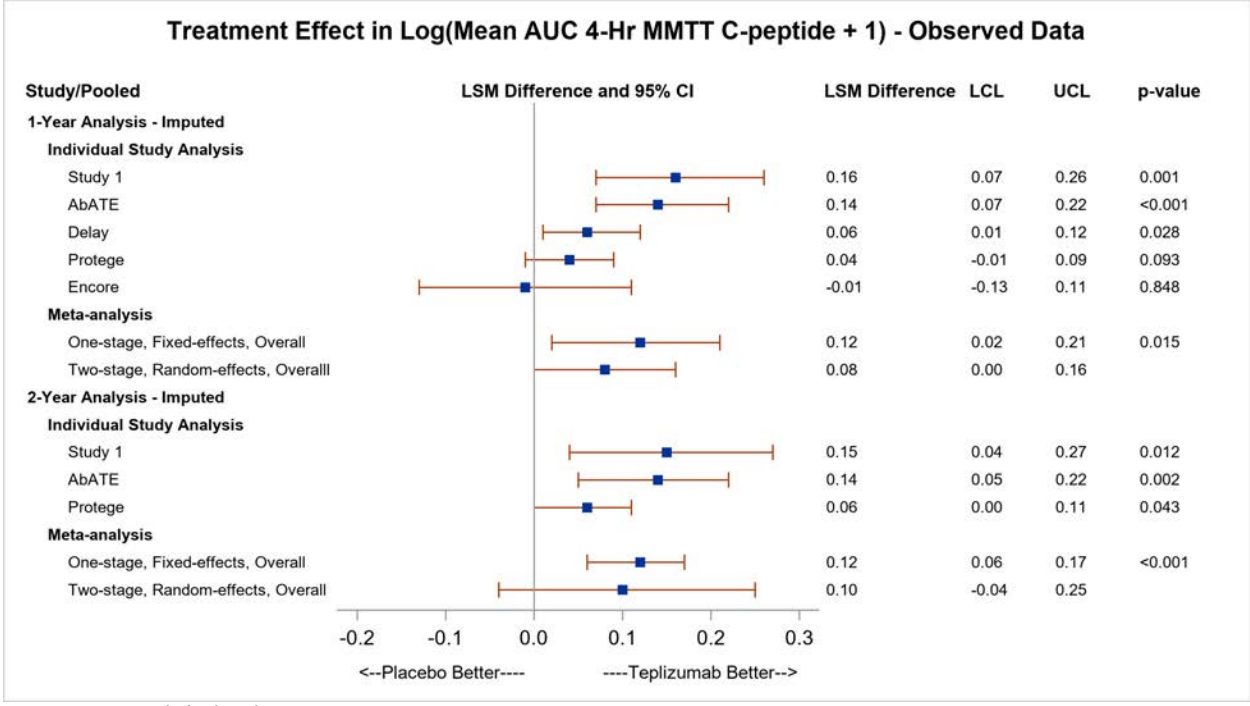
Source: FDA statistical reviewer.
Abbreviations: AUC, area under curve; CI, confidence interval; LSM, least squares mean; SE, standard error

Table 13. Meta-Analyses: Change from Baseline in Log (Mean C-peptide AUC +1) at the End of Year Two – Imputed Data

Study	Treatment	Baseline			Change from baseline at the end of Year 2				Comparison vs. Control				
		Number Analyzed	Mean	SD	LS Mean	SE	95% CI		LSM Difference	SE	95% CI		p-value
							Lower	Upper			Lower	Upper	
Individual Study Analysis													
Study 1	Control	21	0.39	0.03	-0.25	0.05	-0.34	-0.16					
	Teplizumab	21	0.39	0.03	-0.12	0.04	-0.20	-0.04	0.13	0.06	0.01	0.25	0.036
AbATE	Control	25	0.50	0.03	-0.33	0.04	-0.40	-0.26					
	Teplizumab	52	0.53	0.02	-0.20	0.02	-0.24	-0.15	0.13	0.04	0.05	0.22	0.001
Protege	Control	98	0.47	0.02	-0.19	0.02	-0.23	-0.15					
	Teplizumab	207	0.47	0.02	-0.14	0.01	-0.16	-0.11	0.06	0.03	0.01	0.10	0.030
Meta-analysis: One -stage, Fixed-effects													
3-Studies	Control	144	0.46	0.02	-0.25	0.02	-0.29	-0.21					
	Teplizumab	280	0.47	0.01	-0.14	0.02	-0.18	-0.11	0.10	0.03	0.05	0.16	<0.001

Source: FDA statistical reviewer.
Abbreviations: AUC, area under curve; CI, confidence interval; LSM, least squares mean; SE, standard error

Figure 11. Meta-Analyses: Change from Baseline in Log (Mean C-peptide AUC +1) – Observed Data



Source: FDA statistical reviewer.
Abbreviations: AUC, area under curve; CI, confidence interval; LSM, least squares mean; LCl, lower confidence limit; UCl, upper confidence limit; MMTT, mixed meal tolerance test

Table 14. Meta-Analyses: Change from Baseline in Log (Mean C-peptide AUC +1) at the End of Year One – Observed Data

Study	Treatment	Baseline			Change from baseline at the end of Year 1				Comparison vs. Control				
		Number Analyzed	Mean	SD	LS Mean	SE	95% CI		LSM Difference	SE	95% CI		p-value
							Lower	Upper			Lower	Upper	
Individual Study Analysis													
Study 1	Control	21	0.39	0.03	-0.18	0.03	-0.24	-0.11					
	Teplizumab	21	0.39	0.03	-0.02	0.03	-0.08	0.05	0.16	0.05	0.07	0.26	0.001
AbATE	Control	22	0.50	0.04	-0.22	0.03	-0.29	-0.16					
	Teplizumab	50	0.53	0.02	-0.08	0.02	-0.12	-0.04	0.14	0.04	0.07	0.22	<0.001
Delay	Control	28	0.44	0.04	-0.14	0.02	-0.18	-0.10					
	Teplizumab	30	0.47	0.03	-0.08	0.02	-0.12	-0.04	0.06	0.03	0.01	0.12	0.028
Protege	Control	91	0.47	0.03	-0.07	0.02	-0.11	-0.03					
	Teplizumab	188	0.46	0.02	-0.03	0.01	-0.06	0.00	0.04	0.02	-0.01	0.09	0.093
Encore	Control	17	0.52	0.05	-0.09	0.04	-0.17	0.00					
	Teplizumab	17	0.43	0.06	-0.10	0.04	-0.18	-0.02	-0.01	0.06	-0.13	0.11	0.848
Meta-analysis: One -stage, Fixed-effects													
5-Studies	Control	179	0.46	0.02	-0.14	0.02	-0.17	-0.11					
	Teplizumab	306	0.47	0.01	-0.06	0.01	-0.08	-0.03	0.08	0.02	0.04	0.12	<0.001

Source: FDA statistical reviewer.
Abbreviations: AUC, area under curve; CI, confidence interval; LSM, least squares mean; SE, standard error

Table 15. Meta-Analyses: Change from Baseline in Log (Mean C-peptide AUC +1) at the End of Year Two – Observed Data

2-Year Follow-up	Treatment	Baseline			Change from baseline at the end of Year 2				Comparison vs. Control				
		Number Analyzed	Mean	SD	LS Mean	SE	95% CI		LSM Difference	SE	95% CI		p-value
							Lower	Upper			Lower	Upper	
Individual Study Analysis													
Study 1	Control	14	0.43	0.03	-0.25	0.04	-0.34	-0.16					
	Teplizumab	19	0.38	0.03	-0.10	0.04	-0.17	-0.02	0.15	0.06	0.04	0.27	0.012
AbATE	Control	21	0.49	0.04	-0.34	0.04	-0.41	-0.27					
	Teplizumab	49	0.53	0.02	-0.20	0.02	-0.25	-0.15	0.14	0.04	0.05	0.22	0.002
Protege	Control	64	0.46	0.03	-0.18	0.02	-0.22	-0.13					
	Teplizumab	131	0.45	0.02	-0.12	0.02	-0.15	-0.09	0.06	0.03	0.00	0.11	0.043
Meta-analysis: One -stage, Fixed-effects													
3-Studies	Control	99	0.46	0.02	-0.25	0.02	-0.29	-0.21					
	Teplizumab	199	0.46	0.02	-0.13	0.02	-0.17	-0.10	0.12	0.03	0.06	0.17	<0.001

Source: FDA statistical reviewer.
Abbreviations: AUC, area under curve; CI, confidence interval; LSM, least squares mean; SE, standard error

---

This is a representation of an electronic record that was signed electronically. Following this are manifestations of any and all electronic signatures for this electronic record.

/s/

YU WANG
06/28/2021 05:09:11 PM

YUN WANG
06/28/2021 06:49:54 PM