Epilepsy & Behavior 137 (2022) 108955

Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Interim results of adaptive functioning and neurodevelopment in BUTTERFLY – An observational study of children and adolescents with Dravet syndrome

Joseph Sullivan [a,*], Elaine Wirrell [b], Kelly G. Knupp [c], Dillon Chen [d], Muhammad Zafar [e], Robert Flamini [f], James Stutely [g], Charlene Brathwaite [g], Pam Ventola [h], Javier Avendaño [g], Kimberly A. Parkerson [g], Nancy Wyant [g], Barry Ticho [g]

[a] University of California, San Francisco, CA, USA
[b] Mayo Clinic, Rochester, MN, USA
[c] Children'sHospital Colorado, Denver, CO, USA
[d] University of California, San Diego, CA, USA
[e] Duke University, Durham, NC, USA
[f] PANDA Neurology, Atlanta, GA, USA
[g] Stoke Therapeutics, Bedford, MA
[h] Cogstate Ltd., New Haven, CT, USA

ARTICLE INFO

Article history:
Received 12 April 2022
Revised 3 October 2022
Accepted 12 October 2022
Available online 1 November 2022

Keywords:
Dravet syndrome
Pediatrics
Cognition
Quality of life
Observational
BUTTERFLY study

ABSTRACT

Purpose: The purpose of this study is to evaluate adaptive functioning and neurodevelopment study assessments in a prospective study of patients with Dravet syndrome (DS). We present 3-month interim adaptive functioning and neurodevelopment data from the prospective, observational BUTTERFLY study in patients with DS aged 2–18 years.

Results: BUTTERFLY enrolled thirty-six patients divided 1:1:1 across three age groups (2–7: 8–12: and 13–18 years). Most enrolled patients were female (61.1%), white (94.4%), and non-Latino (83.3%) with a mean (standard deviation; SD) age of 10.8 (5.2) years and a mean (range) age of seizure onset of 0.4 (0.2–1.0) years. Patients used a mean (SD) of 3.5 (1.63) anti-seizure therapies at baseline. Regression analysis of the baseline Vineland Adaptive Behavior Scale – third edition (VABS-III) composite score indicated that the gap in adaptive function between patients with DS (n = 33) and neurotypical children widens with age. Similarly, developmental quotients calculated for patients who completed all Bayley Scales of Infant Development – third edition (BSID-III) subtests at baseline (n = 15) highlighted a gap in intellectual functioning between patients with DS and neurotypical children that widens with age. More patients in the two older age groups were able to validly complete the Wechsler Preschool and Primary Scale of Intelligence – fourth edition (WPPSI-IV) at baseline compared with the youngest age group. There were trends towards higher raw scores, albeit of low magnitude, in the oldest age group compared with the younger age two groups across multiple VABS-III domains and WPPSI-IV subtests. All three measures showed no significant change in the all-patients analyses and demonstrated relatively low intra-patient variability from baseline to Month 3.

Conclusions: Three-month interim data from BUTTERFLY demonstrated the feasibility of utilizing the VABS-III, BSID-III, and WPPSI-IV for the assessment of adaptive function and neurodevelopment in future clinical studies of DS. Moreover, many patients with DS appear to gain neurodevelopmental and adaptive function skills over time, although at a slower rate and lower magnitude than that seen in the neurotypical population.

Abbreviations: BSID-III, Bayley Scales of Infant Development – third edition; CFR, Code of Federal Regulations; CI, confidence interval; DQ, Development Quotient; DS, Dravet Syndrome; GCP, Good Clinical Practice; ICH, International Council for Harmonisation; LS, least-squared; QoL, quality of life; QOLCE-55, Quality of Life in Childhood Epilepsy – 55 item; SD, standard deviation; TESC, The Epilepsy Study Consortium; VABS-III, Vineland Adaptive Behavior Scale – third edition; WASI-II, Wechsler Abbreviated Scale of Intelligence – second edition; WPPSI-IV, Wechsler Preschool and Primary Scale of Intelligence – fourth edition.

* Corresponding author.
E-mail address: joseph.sullivan@ucsf.edu (J. Sullivan).

https://doi.org/10.1016/j.yebeh.2022.108955
1525-5050/

1. Introduction

Dravet syndrome (DS) is a severe, genetic, developmental, and epileptic encephalopathy [1] with a reported incidence of between 1 in 15,000 and 1 in 33,000 live births [2–5]. Patients with DS usually present with frequent, prolonged, and treatment-resistant seizures during their first year of life [1]. Approximately 85% of DS cases are caused by pathogenic heterozygous variants in the *SCN1A* gene, which encodes the neuronal voltage-gated sodium channel type 1 α subunit ($Na_v1.1$) [6–9]. At least 1257 different variants in the *SCN1A* gene have been reported in patients with DS including single nucleotide substitutions, small insertions or deletions, and whole gene deletions [8–11]. These pathogenic variants can result in the haploinsufficiency of the $Na_v1.1$ protein in select neurons in the brain [12].

In addition to seizures, over 90% of patients with DS also suffer from multiple significant non-seizure comorbidities that adversely affect their quality of life (QoL), including intellectual disability, ataxia/motor abnormalities, behavioral problems, speech impairment, sleep disturbances, disruptions of the autonomic nervous system, growth defects, mood disorders, and a high risk of sudden unexpected death [13,14]. The burden of DS is also significant for caregivers, who reported profound adverse effects on their QoL in a recent multinational cohort study [15].

To date, the primary goal of DS treatment has been to maximize the quality of life by reducing the frequency and severity of seizures [16,17]. However, DS is a highly treatment-resistant form of epilepsy, and many patients (>90%) continue to experience uncontrolled seizures despite the use of multiple anti-seizure therapies [2,17]. In addition, while also being sensitive to seizure-related comorbidities such as frequent injuries and hospitalizations, most clinical trials in patients with DS do not emphasize the impact of therapies on other non-seizure comorbidities. Therefore, there remains a significant unmet need for additional therapies that can further optimize seizure control and address the non-seizure comorbidities associated with DS and improve patient QoL [2].

As potential disease-modifying therapies emerge, it is important to understand the trajectory of the non-seizure comorbidities in patients with DS across childhood and adolescence and to establish optimal methods for reliably measuring the change in these parameters over time, particularly during the shortened timeframe of a clinical trial. Overall, prior studies suggest a period of normal development followed by stagnation or slowing of development beginning in the second year of life, though there remains relatively limited perspective, systematic data using repeat measures across childhood and adolescence [18–22]. BUTTERFLY aims to supplement this current knowledge base, and here we present 3-month interim adaptive function and neurodevelopment data from the prospective, observational BUTTERFLY study in patients with DS aged 2–18 years.

2. Methods

2.1. Study design and participants

BUTTERFLY is an ongoing multicenter, US-based, longitudinal, prospective observational study. A list of study sites is included in Supplementary Table 1. Patients were included in the study if they were 2–18 years old, had a confirmed diagnosis of DS, a documented pathogenic, likely pathogenic variant, or variant of uncertain significance in the *SCN1A* gene, and ≥2 convulsive seizures in the 4 weeks prior to screening. Patients were excluded from the study if they had specific variants of the *SCN1A* gene demonstrated to cause a gain of function if they were receiving treatment with a sodium channel blocker, if they had clinically significant unstable medical conditions other than epilepsy, or if they had a prior history of brain/spinal cord disease or bacterial meningitis. A full list of the study's inclusion and exclusion criteria is provided in Supplementary Table 2.

Patients participating in the BUTTERFLY study are assessed at baseline and 3-, 6-, 12-, 18-, and 24 months. All research participants, or their legally authorized representatives, granted informed consent or assent before undergoing any study-related activities.

2.2. Study procedures

Study assessments and procedures are summarized in Supplementary Fig. 1 and Supplementary Table 3. This interim analysis includes data available following the completion of Month 3 by all enrolled patients with a cut-off date of 01 April 2021. Data are only included for visits that occurred within the window specified in the protocol (Supplementary Fig. 1).

A patient could be withdrawn from the study for medical or administrative reasons including intercurrent illness, withdrawal of informed consent, or determination that the patient did not meet the criteria for DS diagnosis by The Epilepsy Study Consortium (TESC). This consortium comprises a group of scientific investigators from academic medical research centers who are dedicated to accelerating the development of new therapies in epilepsy to improve patient care.

2.3. Study objectives and assessment details

The primary objective of BUTTERFLY is to evaluate the neurodevelopmental status and change from baseline over 24 months according to four scales: the Vineland Adaptive Behavior Scale – third edition (VABS-III), the Bayley Scales of Infant Development – third edition (BSID-III), the Wechsler Preschool and Primary Scale of Intelligence – fourth edition (WPPSI-IV), and the Wechsler Abbreviated Scale of Intelligence – second edition (WASI-II). Only baseline and Month 3 data for the primary study objective are presented in this manuscript.

The VABS-III is used for the measurement of adaptive behavior across communication (subdomains: receptive, expressive, and written), daily living skills (subdomains: personal, domestic, and community), socialization (subdomains: interpersonal, relationships, play and leisure, and coping skills), motor skills (subdomains: gross motor and fine motor) and maladaptive behavior, and is designed for use from birth to age 90 years.

The BSID-III is used for the measurement of development across cognitive, language (receptive and expressive communication), and motor (fine and gross) domains and is designed for use from birth to 3 years and 6 months.

The WPPSI-IV is used for the assessment of verbal and nonverbal intellectual functioning and is designed for use from ages 2 years and 6 months to ages 7 years and 7 months via two versions dependent upon patient age. The WPPSI-IV, 2 years and 6 months to 3 years and 11 months version, evaluates verbal comprehension, visual-spatial intelligence, and working memory. The WPPSI-IV, 4 years and 0 months to 7 years and 7 months version, evaluates verbal comprehension, visual-spatial intelligence, fluid reasoning, working memory, and processing speed.

The WASI-II is used for the assessment of verbal and nonverbal intellectual functioning and is designed for use from age 6 years and 0 months to age 90 years and 11 months.

The BSID-III, WPPSI-IV, and WASI-II are completed by qualified raters, and the VABS-III is administered by a qualified rater via an interview with a parent or caregiver.

2.4. Sample size calculation and statistical methods

As this was an observational study, no formal sample size calculations were performed. However, an enrollment goal of thirty-six patients with a target ratio of 1:1:1 divided among three target age groups (2–7, 8–12, and 13–18 years) was established to try and ensure that a minimum of 30 patients would likely complete the study. All statistical analyses were carried out on the all-patients data set, which comprised every patient that was enrolled in the BUTTERFLY study. Some analyses were also performed on data sets split by patient age group.

Age-based standardized scores for VABS-III (ranging from 20 to 160, with higher scores indicating better adaptive function; mean neurotypical score of 100 with a standard deviation [SD] of 15), as well as raw scores for each domain and subdomain, were summarized and assessed for change from baseline to Month 3, overall and by age group. The VABS-III composite score was also calculated, not including the motor component for patients older than 9 years and 11 months.

Raw scores for each subtest in the WPPSI-IV and WASI-II were summarized and assessed for change from baseline at Month 3, overall, and by age group. Raw scores were also converted into standardized scores for patients whose chronological age fell into the testing age range; however, this evaluation was limited by sample size and not reported here.

Raw scores as well as age-equivalent scores for BSID-III cognitive, communication, and motor component subtests, were summarized and assessed for change from baseline at Month 3, overall and by age group. Development Quotient (DQ) scores for the BSID-III were calculated for patients who completed all BSID-III subtests as a ratio of the mean age-equivalent score of all subtests divided by the chronological age of the patient at the time of testing, multiplied by 100. These scores were summarized and assessed for change from baseline at Month 3, overall, and by age group.

For all primary endpoints, a mixed-effects model for repeated measures with the visit as a fixed effect and baseline score as a covariate was utilized to test for the overall group (all-patients data set) change from baseline to Month 3. Paired t-tests were also performed on the change from baseline to Month 3 for the overall group and by age group, to test the null hypothesis that the change was 0. If significant straying from the normal distribution was seen with a Shapiro–Wilk test, a signed rank test was used instead. Formal two-sided statistical analyses were carried out at the 5% level of significance on change from baseline in cognitive tests at Month 3.

2.5. Ethics and compliance

The procedures set out in this study protocol, pertaining to the conduct, evaluation, and documentation of BUTTERFLY, are designed to ensure that the Sponsor and Investigators abide by the International Council for Harmonisation (ICH) Guidelines. Compliance with these regulations and guidelines also constitutes compliance with the ethical principles described in the current version of the Declaration of Helsinki (2002).

3. Results

3.1. Patient disposition

Patient disposition for the BUTTERFLY study is shown in Supplementary Fig. 2. Briefly, of the 40 patients for whom signed informed consent for participation was obtained, 2 were lost to screen failure, and 2 discontinued early because they did not meet the criteria for DS defined by TESC. Therefore, a total of 36 patients were enrolled 1:1:1, divided across each of the three target age groups (2–7: 8–12: 13–18 years). Of these 36 patients, 35 completed both baseline and Month 3 visits, with 30 of these within the protocol-specified visit windows. At the time of the interim data cut-off (01 April 2021), 31 of these patients were still participating in the study; 5 had withdrawn (1 due to intercurrent illness and 4 had transferred to the MONARCH study [clinicaltrials.gov registration number NCT04442295]).

3.2. Patient baseline demographics

Patient baseline demographics are summarized in Table 1. The mean (SD) age of the 36 patients enrolled in BUTTERFLY was 10.8 (5.2) years; the minimum and maximum age of enrolled patients was 2.2 and 18.9 years, respectively. Most enrolled patients were female (61.1%), white (94.4%) and non-Latino (83.3%) with a mean (range) age of seizure onset (n = 36) of 0.4 (0.2–1.0) years. Patients were taking a mean (SD) of 3.5 (1.63) anti-seizure therapies (including non-pharmacological therapies, e.g., ketogenic diet, vagal nerve stimulator) during the baseline period; the most common was clobazam (67%, n = 24) followed by fenfluramine (42%, n = 15), stiripentol (36%, n = 13), and valproic acid (36%, n = 13). Across the 4-week baseline, the mean convulsive seizure frequency was 14.4 [95% confidence interval (CI) 8.10–20.68], n = 26), including 24 patients who had generalized tonic-clonic seizures with a mean of 9.1 [95% CI 5.54–12.69].

Overall, 35 of 36 caregivers completed at least some domains of the VABS-III at baseline, and 33 of 36 had adaptive behavior composite scores available, including 12 patients in the 2–7 years group, 12 patients in the 8–12 years group, and 9 patients in the 13–18 years group.

The number of patients completing the BSID-III, WPPSI-IV, and WASI-II by age group at baseline are shown in Supplementary Fig. 3. Most patients enrolled in BUTTERFLY attempted the WPPSI-IV, 4 years and 0 months to 7 years and 7 months version and/or the BSID-III based on the scale selection criteria in the study protocol, and data from these two scales are discussed further in the manuscript. Of the 31 patients who attempted one of these scales at baseline, 26 patients fully completed it, and 29 patients at least partially completed one of the scales. Two patients were unable to complete either of the scales and 1 patient went on to complete the WASI-II due to ceiling criteria on the WPPSI-IV. Only two patients completed the WPPSI-IV, 2 years and 6 months to 3 years and 11 months version; therefore, only data for the WPPSI-IV, 4 years and 0 months to 7 years and 7 months version, are presented in this manuscript. Since only 3 patients completed the WASI-II at baseline, and 2 of these represented protocol violations because of failure to attempt the WPPSI-IV first, WASI-II results are not discussed any further in this manuscript.

3.3. VABs-III

Regression analysis of the VABS-III adaptive behavior composite score at baseline (n = 33), indicated that the gap in adaptive function between patients with DS and neurotypical children widens with age (Fig. 1). This was also demonstrated in age-equivalent analyses of four selected key domains of VABS-III: receptive communication, expressive communication, interpersonal relationships, and gross motor domains (Fig. 2). Consistent with this, patients in the 2–7 years age group were found to perform closer to neurotypical children than patients in the older age groups across multiple VABS-III domains, and had a higher overall mean (95% [CI]) adaptive behavior composite score (ABC) at baseline (62.8 [52.25, 73.25] versus 37.9 [31.38, 44.46] for patients aged 8–12 years and 31.3 [19.34, 43.32] for patients aged 13–18 years,

Table 1
Patient baseline demographics.

Parameter	Patient group			
	2–7 years (n = 12)	8–12 years (n = 12)	13–18 years (n = 12)	All enrolled (N = 36)
Age at screening, years				
Mean (SD)	5.1 (2.0)	10.5 (1.5)	16.9 (1.6)	10.8 (5.2)
Min, Max	2.2, 7.8	8.2, 12.4	14.5, 18.9	2.2, 18.9
Age of seizure onset, months				
Mean (SD)	4 (0.9)	5 (1.7)	6 (2.5)	5 (2.0)
Min, Max	3, 6	2, 7	3, 11	2, 11
Age of initial onset of developmental delay, months				
Mean (SD)	17 (6.1)	18 (13.0)	25 (16.0)	20 (12.0)
Min, Max	6, 24	6, 48	6, 60	6, 60
Gender, n (%)				
Female	5 (41.7)	7 (58.3)	10 (83.3)	22 (61.1)
Male	7 (58.3)	5 (41.7)	2 (16.7)	14 (38.9)
Race, n (%)				
White	11 (91.7)	12 (100.0)	11 (91.7)	34 (94.4)
Asian	1 (8.3)	0 (0.0)	0 (0.0)	1 (2.8)
Other	0 (0.0)	0 (0.0)	1 (8.3)	1 (2.8)
Ethnicity, n (%)				
Non-Latino	9 (75.0)	10 (83.3)	11 (91.7)	30 (83.3)
Latino	3 (25.0)	2 (16.7)	0 (0.0)	5 (13.9)
Missing	0 (0.0)	0 (0.0)	1 (8.3)	1 (2.8)
Baseline convulsive seizure frequency	n = 9	n = 8	n = 9	n = 26
Mean (SD)	7.9 (6.2)	13.6 (8.4)	21.6 (23.5)	14.4 (15.6)
Min, Max	1.0, 19.0	0.0, 26.0	0.0, 65.0	0.0, 65.0

Baseline is defined as the first 4-week period after the Screening Visit. SD; standard deviation; Max, maximum, Min, minimum.

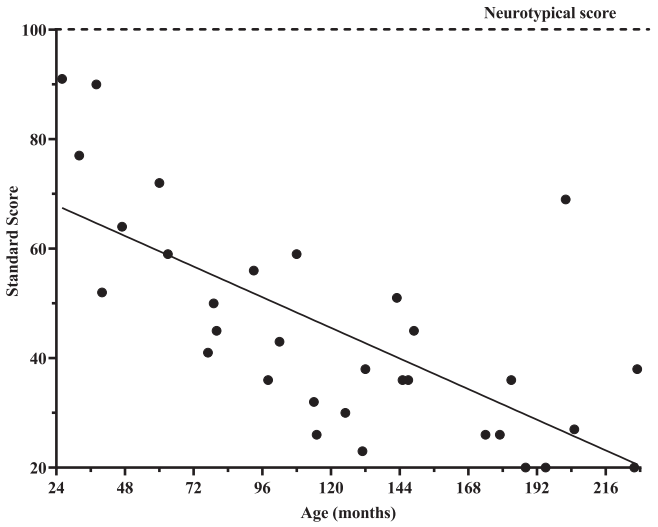

Fig. 1. Regression analysis of adaptive behavior composite score for individual patients with all VABS-III domains completed at baseline (n = 33/36). Regression analysis statistics: $F(1, 31) = 29.76$; $p \leq 0.0001$; $r^2 = 0.4898$. The patients with scores of 91 and 90 were 2 years and 2 months and 3 years and 2 months old, respectively, at screening. The adaptive behavior composite score only includes the motor component for patients aged 2 years to 9 years and 11 months. The neurotypical score is 100 ± 15 SD. VABS-III, Vineland Adaptive Behavior Scales – third edition.

respectively; Supplementary Table 4). There was a minimal relationship between baseline seizure frequency and VABS-III ABC score at baseline (n = 26, Supplementary Fig. 5).

Analysis of VABS-III age-equivalent scores in the all-patients group detected no significant change from baseline to Month 3 in any subdomain and indicated that intra-patient variability was relatively low over this time period (Table 2; Fig. 2). When split by patient age group, the only significant mean change in age-equivalent scores from baseline to Month 3 was in the gross motor subdomain of motor skills for those aged 2–7 years (n = 8, mean change: −1.8; 95% CI: −3.22, −0.28; P-value = 0.0256). Of note, for 18/23 (78%) patients with VABS-III age-equivalent scores at both baseline and Month 3, the same parent or caregiver was interviewed at both time points.

Mean raw scores were higher with increased age across some VABS-III subdomains at baseline, although the observed magnitude and rate of this were low compared to neurotypical children (Table 3). For example, the mean (95% CI) raw expressive communication score was 39.0 (25.88, 52.12) in patients aged 2–7 years, 47.6 (31.69, 63.48) in patients aged 8–12 years, and 61.2 (40.27, 82.10) in patients aged 13–18 years at baseline. Neurotypical children would be expected to achieve raw expressive communication scores of 46.0–95.0 when aged 2–7 years, 95.0–97.0 when aged 8–12 years, and 97.0–98.0 when aged 13–18 years (98.0 is the maximum score on the nonlinear scale, which is expected to be reached by age 16 years and 0 months) [23]. Therefore, based on these observed mean raw scores, participants in the 2–7 years group were performing at a neurotypical age equivalence of 1 year and 10 months, participants in the 8–12 years group were performing at a neurotypical age equivalence of 2 years and 0 months, and participants in the 13–18 years group were performing at a neurotypical age equivalence of 2 years and 5 months in their expressive communication [23].

The mean (95% CI) raw gross motor skills score was 46.8 (32.68, 60.82) in patients aged 2–7 years, 65.6 (54.54, 76.58) in patients aged 8–12 years, and 64.8 (47.75, 81.75) in patients aged 13–18 years at baseline. Neurotypical children would be expected to achieve a raw gross motor skills score of 60.0–83.0 when aged 2–7 years, 84.0–86.0 when aged 8–12 years, and 86.0 when aged 13–18 years (86.0 is the maximum score on the nonlinear scale, which is expected to be reached by the age of 8 years and 10 months) [23]. Therefore, based on these observed mean raw scores, participants in the 2–7 years group were performing at a

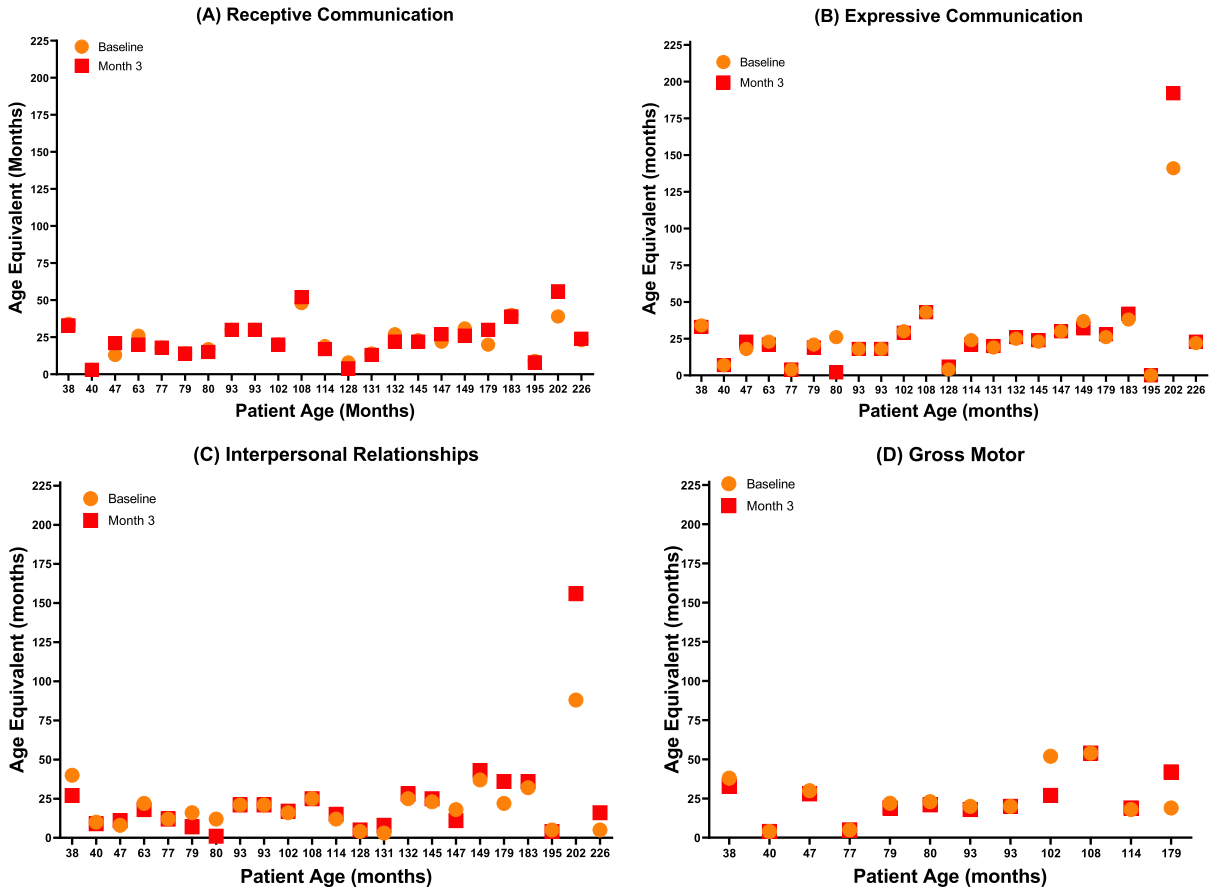

Fig. 2. Patient baseline and Month 3 age equivalent scores across the (A) Receptive communication; (B) Expressive communication; (C) Interpersonal relationships; and (D) Gross motor subdomains of the VABS-III. Only patients that have both baseline and Month 3 scores within the protocol-specified windows are included. Examples of components in the VABS-III receptive communication domain include: responding to a parent/caregiver's voice by looking, understanding at least 10 words, identifying at least three pictured objects, and paying attention to a television show for at least 30 minutes. Examples of components in the VABS-III expressive communication domain include: making sounds of pleasure, saying "yes", saying what he/she is doing in simple sentences, uses plural nouns. Examples of components in the VABS-III interpersonal relationships domain include: looking at the face of the parent/caregiver, smiling/ vocalizing when someone familiar approaches, checking to make sure the parent/familiar other is nearby, and participating in conversations on nonpreferred topics. Examples of components in the VABS-III gross motor domain include: sitting for at least one minute with his/her back supported, crawling backward downstairs or scooting on the bottom, running without falling though may be uncoordinated, pedaling a tricycle or similar vehicle for at least 6 feet. VABS-III, Vineland Adaptive Behavior Scales – third edition.

Table 2
Change from baseline in VABS-III age equivalent scores at Month 3.

VABS-III domain	VABS-III subdomain	n	LS mean change from baseline	95% CI	*P*-value
Communication	Receptive	23	0.70	−1.57, 2.96	0.5304
	Expressive	23	1.30	−1.98, 4.59	0.4179
	Written	23	−1.09	−3.67, 1.49	0.3911
Daily living skills	Personal	23	−1.70	−4.23, 0.84	0.1793
	Domestic	23	0.30	−2.52, 3.13	0.8250
	Community	23	0.09	−2.32, 2.49	0.9408
Socialization	Interpersonal	23	3.26	−1.52, 8.04	0.1709
	Play & Leisure	23	3.35	−1.68, 8.38	0.1807
	Coping	23	1.65	−1.04, 4.34	0.2158
Motor skills	Gross	12	−1.25	−7.39, 4.89	0.6595
	Fine	12	2.25	−1.74, 6.24	0.2376

CI, confidence interval; LS, least squared; VABS-III, Vineland Adaptive Behavior Scales – third edition.

neurotypical age equivalence of 1 year and 6 months, participants in the 8–12 years group were performing at a neurotypical age equivalence of 2 years and 6 months, and participants in the 13–18 years group were performing at a neurotypical age equivalence of 2 years and 5 months on assessment of gross motor skills [23].

3.4. WPPSI-IV (4 years to 7 years and 7 months)

There were higher numbers of patients in the older age groups (3/12 [25.0%] patients aged 8–12 years, and 7/12 [58.3%] patients aged 13–18 years, respectively) completing the WPPSI-IV at base-

Table 3
Baseline VABS-III raw scores across subdomains for patients enrolled in the BUTTERFLY study.

VABS-III domain	VABS-III subdomain	Patient group									Expected minimum score by age (years:months)*			Maximum score (expected age of achievement; years:months)*
		2–7 years			8–12 years			13–18 years						
		n	Mean	95% CI	*n*	Mean	95% CI	*n*	Mean	95% CI	2:0	8:0	13:0	
Communication	Receptive	12	44.4	34.70, 54.13	12	43.3	31.63, 54.87	11	47.7	35.34, 60.11	51.0	74.0	76.0	78 (18:3)
	Expressive	12	39.0	25.88, 52.12	12	47.6	31.69, 63.48	11	61.2	40.27, 82.10	46.0	95.0	97.0	98 (16:0)
	Written	10	5.7	2.76, 8.64	12	14.5	7.10, 21.90	11	20.1	8.41, 31.77	0.0±	54.0	71.0	76 (22:0)
Daily living skills	Personal	12	37.0	26.46, 47.54	12	40.1	32.04, 48.13	11	59.2	37.71, 80.65	33.0	97.0	103.0	110 (20:0)
	Domestic	10	4.5	1.39, 7.61	12	3.7	1.02, 6.31	11	13.2	4.18, 22.18	0.0†	34.0	51.0	60 (22:0)
	Community	10	6.5	3.26, 9.24	12	9.8	4.82, 14.68	11	21.4	7.14, 35.58	0.0¥	59.0	87.0	116 (22:0)
Socialization	Interpersonal	12	35.3	28.50, 42.00	12	32.0	23.81, 40.19	11	48.7	32.52, 64.93	41.0	79.0	83.0	86 (22:0)
	Play & Leisure	12	22.4	16.05, 28.78	12	19.3	9.81, 28.85	11	31.7	17.05, 46.41	27.0	59.0	67.0	72 (20:0)
	Coping	12	21.3	16.38, 26.29	12	15.0	8.82, 21.18	11	26.8	16.10, 37.53	24.0	51.0	62.0	66 (22:0)
Motor skills	Gross	12	46.8	32.68, 60.82	9	65.6	54.54, 76.58	8	64.8	47.75, 81.75	60.0	84.0	86.0	86 (8:10)
	Fine	12	29.3	19.86, 38.64	8	33.8	26.51, 40.99	8	42.4	25.38, 59.37	29.0	66.0	68.0	68 (9:10)

CI, confidence interval; VABS-III, Vineland Adaptive Behavior Scales – third edition.
* Information taken from the Vineland Adaptive Behavior Scales – third edition Appendixes B–E, Table B.2 (2016), available at: https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/vineland-3/vineland-3-manual-appendices-b-e.pdf (accessed November 3rd 2021).
± Expected written communication raw scores for children aged less than 3 years and 0 months range from 0.0 to 5.0.
† Expected domestic daily living skills raw scores for children aged less than 3 years and 0 months range from 0.0 to 6.0.
¥ Expected community daily living skills raw scores for children aged less than 3 years and 0 months range from 0.0 to 10.0.

Table 4
Baseline WPPSI-IV (4 years and 0 months to 7 years and 7 months) raw scores across subtests for patients enrolled in the BUTTERFLY study.

WPPSI-IV components	WPPSI-IV subtest	Patient group									Maximum score (developmental age of 7 years and 7 months)*
		2–7 years			8–12 years			13–18 years			
		n	Mean	95% CI	*n*	Mean	95% CI	*n*	Mean	95% CI	
Verbal Comprehension	Information	1	13.0	–	3	11.3	−1.17, 23.84	7	14.6	8.58, 20.56	29
	Similarities	1	4.0	–	3	5.3	−4.71, 15.37	6	16.0	4.28, 27.72	40
Visual Spatial	Block design	1	12.0	–	3	11.0	−10.66, 32.66	7	15.6	8.33, 22.81	34
Fluid Reasoning	Matrix reasoning	1	13.0	–	3	7.3	4.46, 10.20	6	11.8	6.34, 17.32	26
Working Memory	Picture memory	1	8.0	–	3	8.3	2.08, 14.58	7	10.9	5.45, 16.26	35
Processing Speed	Bug search	1	6.0	–	3	9.0	−3.91, 21.91	6	14.8	5.02, 24.64	66

CI, confidence interval; WPPSI-IV, Wechsler Preschool and Primary Scale of Intelligence – fourth edition.
* Apart from one patient in the 2–7 years group, all patients with valid scores on the WPPSI-IV had a chronological age above 7 years and 7 months.

line compared with the youngest age group (1/8 [12.5%] patients aged 2–7 years, 8 patients were at least 4 years old; Table 4). Consistent with this, patients aged 13–18 years achieved higher raw scores than those aged 8–12 years in some components of the WPPSI-IV including verbal comprehension, visual-spatial reasoning, and fluid reasoning, although the magnitude of the difference was low (Fig. 3).

Overall, analysis of WPPSI-IV raw scores in the all-patients group detected no significant change from baseline to Month 3 in any subtest and indicated that intra-patient variability was relatively low over this time period (Table 5; Fig. 3). No significant differences from baseline to Month 3 were detected in any of the WPPSI-IV subtests when the raw scores were further broken down by patient age group.

3.5. BSId-III

All patients aged 2 years to 2 years and 5 months, and older patients who were unable to validly complete the WPPSI-IV were assessed using the BSID-III. Calculation of a DQ for all patients

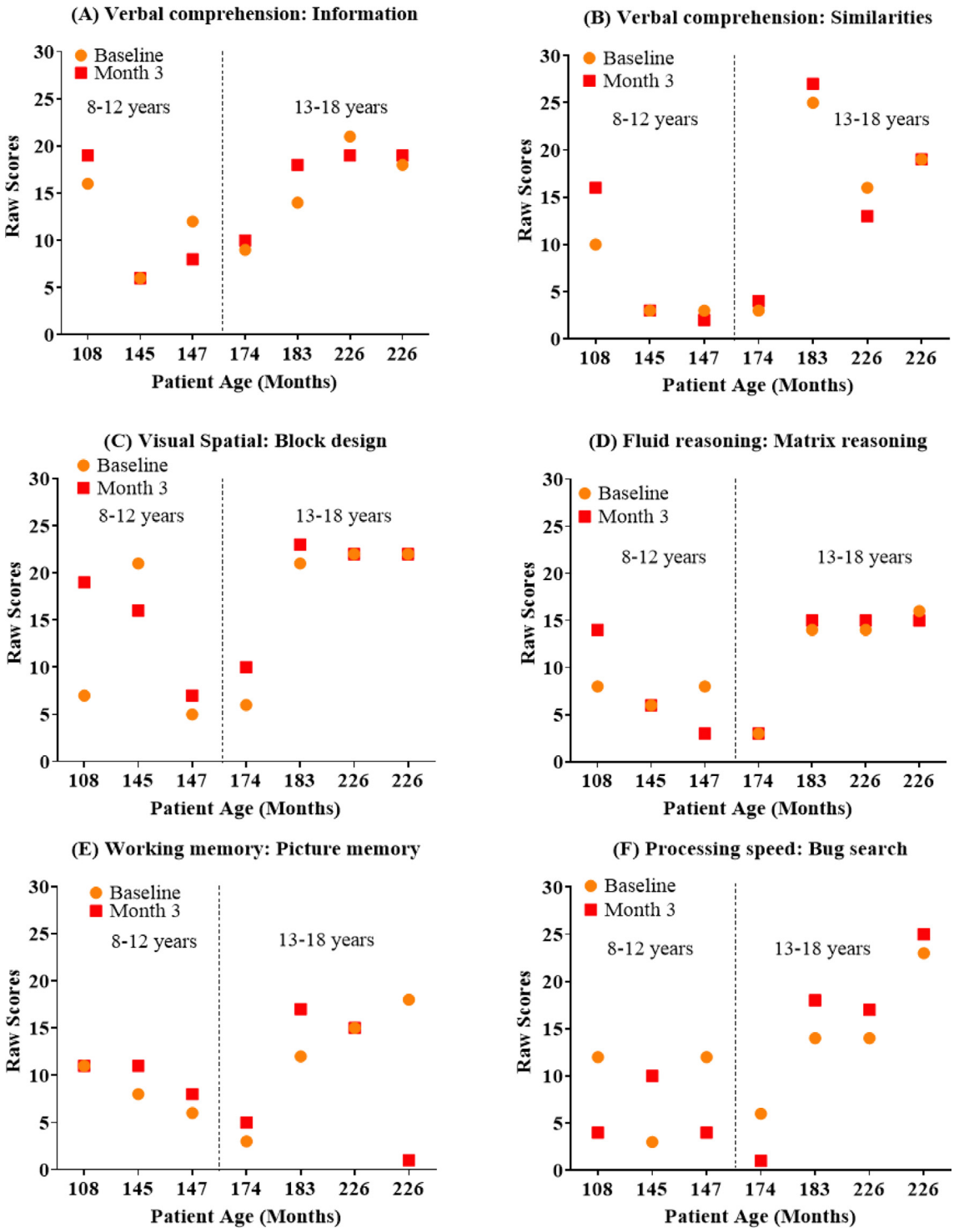

Fig. 3. WPPSI-IV (4 years to 7 years and 7 months) baseline and Month 3 raw scores across the subtests: (A) Information; (B) Similarities; (C) Block design; (D) Matrix reasoning; (E) Picture memory; and (F) Bug search. Only patients that have both baseline and Month 3 scores within the protocol-specified windows are included. The variable y-axis maximums represent the different maximum raw scores of the different subtests. The WPPSI-IV information subtest includes the following components: acquire, retain, and retrieve general factual knowledge (e.g., which animal in the picture barks?). Examples of components in the WPPSI-IV similarities subtest include: picking a picture from the same category as the other 2 pictures and verbally describing the similarity between 2 objects or concepts. Examples of components in the WPPSI-IV block design subtest include: recreating a pictured design using colored blocks within the timeframe. Examples of components in the WPPSI-IV matrix reasoning subtest include selecting 1 of 4 pictures that fit into a sequence or pattern of pictures shown. Examples of components in the WPPSI-IV picture memory subtest include: viewing pictures for the timeframe and then selecting the same pictures on the next page. Examples of components in the WPPSI-IV bug search subtest include a select bug in the group that matches the target bug in the timeframe. WPPSI-IV, Wechsler Preschool and Primary Scale of Intelligence – fourth edition.

who completed every BSID-III subtest at baseline (n = 15) indicated an overall decline with increasing age and highlighted a notable gap between these patients with DS and neurotypical children in their intellectual functioning (Fig. 4; Supplementary Fig. 4). There was a minimal relationship between baseline seizure frequency and BSID-III DQ score at baseline (n = 11, Supplementary Fig. 5).

Table 5
Change from baseline in WPPSI-IV (4 years and 0 months to 7 years and 7 months) raw scores at Month 3 (n = 7 for each subtest).

	LS mean change from baseline	95% CI	*P*-value
Information	0.43	−2.51, 3.37	0.7230
Similarities	0.71	−2.28, 3.70	0.5661
Block design	2.14	−2.05, 6.33	0.2456
Matrix reasoning	0.29	−3.17, 3.75	0.8403
Picture memory	−0.71	−6.62, 5.19	0.7685
Bug search	−0.71	−7.27, 5.84	0.7905

CI, confidence interval; LS, least squared; WPPSI-IV, Wechsler Preschool and Primary Scale of Intelligence – fourth edition.

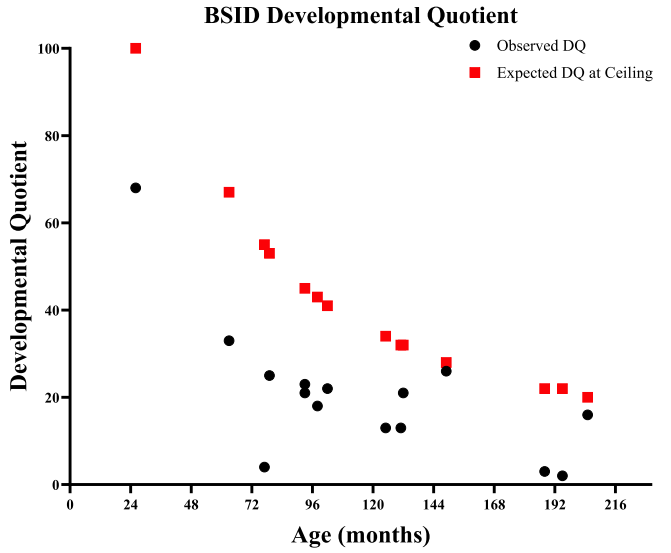

Fig. 4. BSID-III Developmental Quotient (DQ) at baseline (n = 15). The BSID-III DQ can only be calculated for patients who completed all BSID-III subtests and included 6 patients each in the 2–7 years, the 8–12 years groups, and 3 patients in the 13–18 years group. The expected DQ ceiling was calculated as the maximal DQ that could be achieved by patients above the age of the scale (i.e., >42 months old) using the following equation: 42 months divided by age in months at enrollment multiplied by 100. BSID-III, Bayley Scales of Infant Development – third edition; DQ, developmental quotient.

Unlike that noted in the analyses of raw scores of the VABS-III and the WPPSI-IV, mean raw scores were lowest in the oldest age group across multiple subtests for those patients assessed using the BSID-III, though only a minority of patients were assessed using the BSID-III in the oldest age group (Supplementary Table 5).

Overall, the analysis of BSID-III age-equivalent scores in the all-patients group detected no significant change from baseline to Month 3 in any subtest and indicated that intra-patient variability was relatively low over this time period (Table 6; Supplementary Fig. 4). No significant differences from baseline to Month 3 were detected in any of the BSID-III subtests when the age equivalent scores were further broken down by patient age group.

4. Discussion

DS is a severe, genetic, developmental, and epileptic encephalopathy that typically begins in the first year of life and is characterized by frequent, prolonged, and treatment-resistant seizures in combination with multiple other substantial comorbidities including cognitive disability, motor abnormalities, behavioral disturbances, speech impairment, sleep disturbances, and sudden death [1,13].

The observational BUTTERFLY study is the largest, prospective, longitudinal study of cognitive functioning in genetically confirmed DS patients. BUTTERFLY was designed to evaluate both seizure and non-seizure manifestations over 24 months in patients aged 2–18 years with a diagnosis of DS in order to provide systematic data on the trajectory of these patients across childhood and adolescence. Herein we show results from a 3-month interim analysis of BUTTERFLY, which demonstrates that the VABS-III, BSID-III, and WPPSI-IV are appropriate for use in assessing adaptive behavior and neurodevelopment in patients with DS. High baseline completion rates were achieved for these three scales. Our data indicate that the gap in adaptive function and overall intellectual development between patients with DS and neurotypical children widens with age from childhood to adolescence, and our data suggest this is typically not due to regression or loss of skills already gained, which is consistent with reports in the literature describing the evolution of DS over time [18–22]. Based on mean VABS-III and WPPSI-IV raw scores across age groups, and the higher numbers of patients in older age groups completing the WPPSI-IV, many but not all patients with DS appear to gain neurodevelopmental and adaptive function skills over time, although the rate of progression in these skills appears to be much slower and lower in magnitude than that seen in the neurotypical population [22]. Also noteworthy is that scores on the VABS-III socialization domain are higher than in communication (Supplementary Table 4) which is consistent with previous reports [21]. This indicates that neurodevelopmental skills are not equally affected in patients with DS.

The BSID-III, WPPSI-IV, and VABS-III measures detected no significant change from baseline to Month 3 across the all-patients group in the BUTTERFLY study. Further, measures showed relatively low intra-patient variability over 3 months which demonstrates the reliability of these assessments in this patient group. Overall, this interim analysis indicates that these three assessments are likely appropriate for the evaluation of patients with DS in clinical trials.

We identified some limitations to this study and the interim analysis. First, patients in the study may have had more than one caregiver; therefore, this may have resulted in an inconsistency in the way skills were scored between patients on the VABS-III. However, 78% of the patients had the same caregiver complete assessments at baseline and Month 3; therefore, this minimized intra-patient variability at the two-time points. Next, some patients were unable to complete some of the neurodevelopmental evaluations at both baseline and Month 3 leading to a small sample size in some cases. This was partly explained by the evolution of the COVID-19 pandemic, which created the requirement for con-

Table 6
Change from baseline in BSID-III age equivalent scores at Month 3.

	n	LS mean change from baseline	95% CI	*P*-value
Cognitive	9	−1.89	−4.59, 0.82	0.1425
Receptive communication	8	0.00	−4.78, 4.78	1.0000
Expressive communication	8	3.88	−4.20, 11.95	0.2849
Gross motor	6	0.33	−1.12, 1.78	0.5579
Fine motor	8	1.13	−2.78, 5.03	0.5075

BSID-III, Bayley Scales of Infant Development – third edition; CI, confidence interval; LS, least squared.

version from on-site to remote assessments, of which the BSID-III and WPPSI-IV were not allowed per the protocol and also led to some visits occurring outside of the protocol-specified windows. Further, this interim analysis did not evaluate the impact of specific anti-seizure therapies, type of *SCN1A* mutation, or educational level of the parents or caregiver type on adaptive function or neurodevelopment in these patients.

5. Conclusions

These data highlight that additional therapies are desperately needed to address comorbidities in DS. Consistent with prior literature, baseline data from the BUTTERFLY study indicate that patients with DS have profound deficits in cognitive functioning and that the gap in adaptive functioning and intellectual development between patients with DS and neurotypical children widens over childhood and adolescence. The majority of patients with DS are not regressing and many are not even plateauing in their development; rather they are gaining skills, just at a rate much slower than neurotypical children. Importantly, our interim analysis demonstrates the feasibility of utilizing the VABS-III, BSID-III, and WPPSI-IV scales for the assessment of adaptive function and neurodevelopment in patients with DS across childhood and adolescence. The high completion rates at baseline and relatively low intra-patient variability observed on these scales over 3 months highlight their reliability and potential for use in repeat assessments over time, as would be required in interventional clinical studies, in the studied patient population. While the BSID-III is designed for use from birth to 3 years and 6 months, and the WPPSI-IV discussed in this manuscript is designed for use from 4 years and 0 months to 7 years and 7 months, our interim data suggest that these two assessments are valid across a larger range of ages in patients with DS. Therefore, the adoption of these scales could be useful for the assessment of neurodevelopment and adaptive behavior in future clinical studies of patients with DS. Future publications of the remainder of the BUTTERFLY data will provide follow-up data for these patients for up to 24 months.

Acknowledgements

The authors thank the investigators, healthcare providers, research staff, patients, and caregivers who participated in the BUTTERFLY study. The medical writing support was provided by Jennifer Shepherd, PhD, of Porterhouse Medical Group and funded by Stoke Therapeutics, in accordance with Good Publication Practice (GPP3) guidelines.

Conflicts of interest

Joseph Sullivan is under contracted research with Zogenix, Stoke Therapeutics, Encoded Therapeutics, and Takeda. He has received consulting fees from Stoke Therapeutics, Encoded Therapeutics, Greenwich Biosciences, Epygenix Therapeutics, and Longboard Pharmaceuticals. Dr. Sullivan also is a member of the Scientific Advisory Board for Epygenix Therapeutics and holds stock options. He is a member of the Dravet Syndrome Foundation Board of Directors and Medical Advisory Board.

Elaine Wirrell has participated on data and safety monitoring boards for Encoded Therapeutics, Neurocrine Biosciences, Amicus Therapeutics, and Acadia Pharmaceuticals. Kelly G Knupp has received research funding from Zogenix, Encoded, Eisai, and West Pharmaceuticals. She has participated on data and safety monitoring boards for GW Pharmaceuticals and Epygenix, and received consulting funds from Biomarin, Zogenix, Encoded, Eisai, Stoke Therapeutics, and Biocodex.

Dillon Chen has nothing to declare.

Muhammad Zafar receives grant support from Eisai, LivaNova, UCB, SK Life Sciences, Marinus, Biopharm, Stoke Therapeutics, Jazz Pharmaceutical, and Neurocrine Biosciences. Muhammad Zafar has participated on data and safety monitoring boards for Xenon Pharmaceutical and the Advisory board for Eisai.

Robert Flamini is a speaker for Lundbeck, Eisai, and Supernus Pharmaceuticals and has received honoraria from Lundbeck and Eisai and research support from GW Pharma.

Pam Ventola is an employee of Cogstate.

James Stutely, Charlene Brathwaite, Javier Avendaño, Kimberly A Parkerson, Nancy Wyant, and Barry Ticho are employees of Stoke Therapeutics.

Funding source

Funding for the BUTTERFLY study and the development of this manuscript was provided by Stoke Therapeutics.

Clinical trial registration

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.yebeh.2022.108955.

References

[1] Dravet C. The core Dravet syndrome phenotype. Epilepsia 2011;52:3–9.
[2] Strzelczyk A, Schubert-Bast S. Therapeutic advances in Dravet syndrome: a targeted literature review. Expert Rev Neurother 2020;20:1065–79.
[3] Bayat A, Hjalgrim H, Møller RS. The incidence of SCN 1A-related Dravet syndrome in Denmark is 1: 22,000: A population-based study from 2004 to 2009. Epilepsia 2015;56(4):e36–9.
[4] Rosander C, Hallböök T. Dravet syndrome in Sweden: a population-based study. Dev Med Child Neurol 2015;57(7):628–33.
[5] Wu YW, Sullivan J, McDaniel SS, Meisler MH, Walsh EM, Li SX, et al. Incidence of Dravet syndrome in a US population. Pediatrics 2015;136:e1310–5.
[6] Scheffer IE. Diagnosis and long-term course of Dravet syndrome. Eur J Paediatr Neurol 2012;16:S5–8.
[7] Claes L, Del-Favero J, Ceulemans B, Lagae L, Van Broeckhoven C, De Jonghe P. De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy. Am J Hum Genet 2001;68(6):1327–32.
[8] Djémié T, Weckhuysen S, von Spiczak S, Carvill GL, Jaehn J, Anttonen A-K, et al. Pitfalls in genetic testing: the story of missed SCN1A mutations. Mol Genet Genomic Med 2016;4(4):457–64.
[9] Suls A, Claeys KG, Goossens D, Harding B, Luijk RV, Scheers S, et al. Microdeletions involving the SCN1A gene may be common in SCN1A-mutation-negative SMEI patients. Hum Mutat 2006;27(9):914–20.
[10] Zuberi SM, Brunklaus A, Birch R, Reavey E, Duncan J, Forbes GH. Genotype–phenotype associations in SCN1A-related epilepsies. Neurology 2011;76(7):594–600.
[11] Meng H, Xu H-Q, Yu L, Lin G-W, He N, Su T, et al. The SCN1A mutation database: updating information and analysis of the relationships among genotype, functional alteration, and phenotype. Hum Mutat 2015;36(6):573–80.
[12] Isom LL. Is targeting of compensatory ion channel gene expression a viable therapeutic strategy for Dravet syndrome? Epilepsy Curr 2019;19(3):193–5.
[13] Anwar A, Saleem S, Patel UK, Arumaithurai K, Malik P. Dravet syndrome: an overview. Cureus 2019;11(6):e5006.
[14] Lagae L, Brambilla I, Mingorance A, Gibson E, Battersby A. Quality of life and comorbidities associated with Dravet syndrome severity: a multinational cohort survey. Dev Med Child Neurol 2018;60(1):63–72.
[15] Lagae L, Irwin J, Gibson E, Battersby A. Caregiver impact and health service use in high and low severity Dravet syndrome: A multinational cohort study. Seizure 2019;65:72–9.
[16] Cross JH, Caraballo RH, Nabbout R, Vigevano F, Guerrini R, Lagae L. Dravet syndrome: treatment options and management of prolonged seizures. Epilepsia 2019;60:S39–48.
[17] Wirrell EC, Laux L, Donner E, Jette N, Knupp K, Meskis MA, et al. Optimizing the diagnosis and management of Dravet syndrome: recommendations from a North American consensus panel. Pediatr Neurol 2017;68:18–34.e3.

[18] Wolff M, Cassé-Perrot C, Dravet C. Severe myoclonic epilepsy of infants (Dravet syndrome): natural history and neuropsychological findings. Epilepsia 2006;47:45–8.
[19] Nabbout R, Chemaly N, Chipaux M, Barcia G, Bouis C, Dubouch C, et al. Encephalopathy in children with Dravet syndrome is not a pure consequence of epilepsy. Orphanet J Rare Dis 2013;8(1):176.
[20] Battaglia D, Chieffo D, Lucibello S, Marini C, Sibilia V, Mei D, et al. Multicenter prospective longitudinal study in 34 patients with Dravet syndrome: Neuropsychological development in the first six years of life. Brain Dev 2021;43(3):419–30.
[21] Villeneuve N, Laguitton V, Viellard M, Lépine A, Chabrol B, Dravet C, et al. Cognitive and adaptive evaluation of 21 consecutive patients with Dravet syndrome. Epilepsy Behav 2014;31:143–8.
[22] Ragona F. Cognitive development in children with Dravet syndrome. Epilepsia 2011;52:39–43.
[23] Sparrow SS, Cicchetti DV, Saulnier CA. Vineland Adaptive Behavior Scales - third edition, Appendices B-E, Table B.2. In: Vineland Adaptive Behavior Scales - third edition (https://www.pearsonassessments.com/content/dam/school/global/clinical/us/assets/vineland-3/vineland-3-manual-appendices-b-e.pdf): NCS Pearson, Inc.; 2016 (Accessed Nov 3 2021).