This report may include approved and non-approved uses, formulations, or treatment regimens. The results reported may not reflect the overall profile of a product. Before prescribing any product mentioned in this report, healthcare professionals should consult local prescribing information for the product approved in their country.

SPONSOR is committed to publicly disclosing all medical research results that are significant to patients, health care providers or payers—whether favorable or unfavorable to the SPONSOR product—in an accurate, objective and balanced manner in order for our customers to make more informed decisions about our products.

Personally identifiable information (PII) within this document is either removed or redacted (i.e., specific content is masked irreversibly from view with a black bar) to protect personal privacy. Personally identifiable information includes:

- All named persons associated with the study
- Patient identifiers within text, tables, or figures
- By-patient data listings

Anonymized patient data may be made available subject to an approved research proposal submitted. Information which is considered intellectual property or company confidential was also redacted.
SYNOPSIS

Sponsor: Shire

Name of Finished Product: Not applicable

Name of Active Ingredient: Not applicable

Study Title: A Prospective, Longitudinal, Observational Study to Evaluate Neurodevelopmental Status in Pediatric Patients with Hunter Syndrome (MPS II)

Investigators and Study Centers: Multicenter

Publication (reference): see Appendix 12.1.11

Studied Period: 18 Jan 2013 (first subject enrolled) to 05 Oct 2016 (last subject completed)

Study Phase: Not applicable

Objectives

The primary objective of the study was to evaluate the neurodevelopmental status of pediatric subjects with mucopolysaccharidosis II (MPS II) over time.

The secondary objectives of the study were:

- To evaluate the incidence of adverse events (AEs) in pediatric subjects with MPS II
- To evaluate medication usage in pediatric subjects with MPS II
- To evaluate functional status in pediatric subjects with MPS II as measured by the Hunter Syndrome-Functional Outcomes for Clinical Understanding Scale (HS-FOCUS) instrument
- To evaluate health status in pediatric subjects with MPS II as measured by the EuroQol-5D (EQ-5D) instrument.

Methodology

Study HGT-HIT-090 was a multicenter, multinational, prospective, longitudinal, observational study designed to evaluate the neurodevelopmental status of pediatric subjects with MPS II over a 2-year period. Up to 100 pediatric subjects with MPS II were to be enrolled. The study methodology included standardized neurodevelopmental and adaptive function instruments and parent-reported measures.

Subject eligibility was based initially on age and a confirmed diagnosis of MPS II. Eligibility was not based on current or prior treatment with ELAPRASE. Once consented, the subject underwent additional screening assessments of physical, developmental, neurological, neurocognitive, and disease status. Subjects who successfully met all screening criteria were enrolled in the study.

During the study, subjects visited the site for 1 to 2 days once every 3 months for up to 2 years. Site visits occurred at baseline (day 0) and months 3, 6, 9, 12, 15, 18, 21, and 24.

Activities repeated once every 3 months: hearing, neurodevelopmental, and quality of life assessments; survey of medications, treatments, and procedures; and AE check

Activities repeated once every 6 months: physical and neurological examinations and vital sign check.

Any subject who withdrew consent or discontinued from the study was to complete the assessments assigned to the end of study (EOS)/month 24 visit within 14 days after discontinuation. Neurodevelopmental testing performed within the 8 weeks before the EOS visit was not repeated at the EOS visit.
Subjects were expected to participate in the study for approximately 25 months.

No investigational drug was administered in this observational study of MPS II.

Number of Subjects (Planned and Analyzed)
Up to 100 pediatric subjects were to be enrolled. In total, the required signed informed consent/assent was provided by the 55 subjects who enrolled in the study. The study was completed by 23 subjects (41.8%); 32 subjects (58.2%) discontinued. Of the 55 subjects who enrolled in the study, 25 subjects (45.5%) discontinued from the study and enrolled in study HGT-HIT-094.

Diagnosis and Main Criteria for Inclusion
Male pediatric subjects with a deficient in iduronate-2-sulfatase (I2S) enzyme activity (≤10% of the lower limit of the normal range as measured in plasma, fibroblasts, or leukocytes and (based on the reference laboratory’s normal range), and had either a documented mutation in the IDS gene that left the fragile X-mental retardation 1 (FMR1) and FMR2 genes intact or a normal enzyme activity level of 1 other sulfatase as measured in plasma, fibroblasts, or leukocytes (based on the reference laboratory’s normal range). Subjects were to have sufficient auditory capacity with or without hearing aids as needed.

Test Product, Dose and Mode of Administration, Lot Number: Not applicable.

Duration of Treatment: The planned overall duration of each subject’s participation in this study was approximately 25 months.

Reference Therapy, Dose and Mode of Administration, Lot Number: Not applicable.

Criteria for Evaluation

Disease Progression
Neurodevelopmental assessments were performed to assess cognitive and adaptive functions over time as follows:

- Cognitive Function: Standard scores in cluster areas of the Differential Ability Scales, Second Edition (DAS-II) or developmental quotient in the Bayley Scales of Infant Development®, Third Edition (BSID-III). Cognitive function was assessed using the DAS-II unless the subject was too young (<2 years 6 months of age) or unable to complete the assessment; in which case, the BSID-III was used.
- Functional status: HS-FOCUS form
- Health status: EQ-5D questionnaire

Safety
Safety was assessed through recording of AEs, including type and severity; medication usage; vital signs; physical examination; height, weight, and body mass index; and hearing assessments.

Statistical Methods

Disease Progression
As that this was an observational study, the statistical methodology supporting the study focused on descriptive rather than inferential approaches. Any statistical testing was viewed as exploratory. Descriptive statistics included n, mean, standard deviation (SD), median, minimum, and maximum for continuous variables, and frequencies and percentages for categorical variables. Two-sided 95% confidence intervals for mean, median, or percentages were presented as appropriate. Graphical techniques were also used to assess trends across time.

Cognitive Assessment Analyses
Descriptive statistics for the DAS-II and VABS-II neurodevelopmental outcomes, as well as the mean change from baseline, were provided for each time point. Box plots of the parameters by time on study were produced. Plots of the parameters by chronological age (ie, the age of the subject at the time of the test) were produced for individual
subjects.

The interdependence and relationships among the neurodevelopmental parameters (eg, the General Conceptual Ability (GCA) scores versus Special Nonverbal Composite (SNC) scores in DAS-II) were assessed by calculating correlations and other measures of association. Scatter plots of the paired variables and their correlation coefficients were produced.

No summary tables were produced for BSID-III data.

**Adaptive Behavior Assessment Analyses**

For subjects with VABS-II data, descriptive statistics were provided for observed values and changes from baseline at each visit. Box plots of VABS-II standard scores for each category by visit were produced. Additionally, plots of standard scores by chronological age were produced for individual subjects.

**Random Coefficient Modeling**

The analysis method for DAS-II GCA and VABS-II Adaptive Behavior Composite (ABC) scores was to fit a random coefficients model (ie, random slope and intercept for time) to assess the trends over time in each parameter. Quadratic and cubic terms for time could have been added to the model if significant or necessary to better characterize the relationship with outcome. Predicted curves (spaghetti plots) for individuals using the subject-specific intercept and slope estimates from the model were plotted. The fit of the model was assessed by examination of residuals and by assessing whether the model was significantly improved by the addition of higher-order interactions.

**Correlation Between Different Neurodevelopmental Assessment Domains**

The following correlation coefficients for observed values and change from baseline were estimated from a linear mixed-effects model that took into account the within-subject repeated measures):

- GCA standard score from DAS-II and the following scores: standard scores of Nonverbal, Spatial, Verbal, and SNC from DAS-II; ABC, Communication, Daily Living Skills, Socialization, and Motor Skills from VABS-II.
- ABC standard score from VABS-II and the following scores from VABS-II: Communication, Daily Living Skills, Socialization, and Motor Skills.

SAS Proc Mixed with restricted maximum likelihood estimation and an unstructured within-subject covariance structure for the random effects was used for the model.

Each pair of parameters was presented graphically as a scatter plot.

**Subgroup Analyses**

The prognostic importance of subgroup covariates was a predefined exploratory analysis. Subgroup analyses were performed for exploratory purposes for the change from baseline in GCA and ABC standard scores.

Descriptive statistics of the observed value and change from baseline in GCA and ABC standard scores were presented by visit week, both overall and within each subgroup. Subgroup analysis included baseline GCA composite score group (≤70 versus >70), age group (<7 years or ≥7 years), phenotype (if available), and genotype mutation category (if available).

**Exploratory Analyses**

Mixed-effects model for repeated measures (MMRM) analysis was performed separately to assess the effect of subgroup factors, including baseline GCA (≤70 versus >70), age group (7 years or ≥7 years), phenotype (if available) and genotype (if available), on the changes from baseline in the GCA and the ABC scores. The MMRM contained categorical effects for visit and subgroup and their interaction. For all models an unstructured within-subject covariance structure was used. If this model failed to converge, a first order autoregressive covariance structure was to be used. The Kenward-Roger approximation was used to estimate denominator degrees of freedom for tests of fixed effects. The estimated least squares means and standard errors at each visit were summarized and plotted by subgroups.

**Functional Assessment Analyses**

Items were averaged within each domain of the HS-FOCUS to derive the 5 functional domain scores.
(walking/standing, grip/reach, school/work, activities, and breathing). The response “does not apply” was treated as missing. The domain scores and change from baseline were summarized descriptively for each time point. The mean observed values for each dimension were plotted over time.

Health Status Analyses
The number and percentage of subjects with each level of response to the EQ-5D was presented by dimension at each visit. The visual analog scale score and change from baseline were summarized. Mean observed values for each dimension were plotted over time.

Safety
Adverse Events
All AEs were coded using the Medical Dictionary for Regulatory Activities, version 16.1. Only AEs reported from the time of informed consent until 30 days after the EOS visit were summarized.

The number and percentage of subjects experiencing each AE and the number of events were summarized by system organ class (SOC) and preferred term (PT). Adverse events were further summarized by the severity and nature of the event (related to conduct of a study procedure, to disease progression of underlying MPS II, or to ELAPRASE infusion). Serious AEs were also summarized.

If a subject had multiple occurrences of the same AEs (overall, and at the SOC and PT levels) with different severity or relationship, the subject was counted once with the maximum known severity or closest relationship, as applicable.

Vital Signs
Descriptive statistics were presented for vital sign observed values and changes from baseline at each time point. Box plots for each parameter at each visit were produced.

Physical Examinations
Any new or worsening abnormalities that occurred after screening or baseline were to be reported as AEs. No additional summary or listing was produced.

Height and Weight
Body mass index (BMI) was calculated as (weight [kg])/(height [m])^2. Height, weight, and BMI observed values and changes from baseline were summarized by scheduled time point.

Hearing Assessments
Data on the use of hearing aids (yes or no) by ear (right, left, or both), and whether the subject had sufficient hearing to participate in study assessments, were summarized by scheduled time point.

Concomitant Medications
Concomitant medications were coded using the WHO Drug Dictionary Enhanced version September 2013. Concomitant medication use was summarized by anatomic therapeutic class and preferred term. The types and incidence of concomitant medication usages were summarized.

Summary of Results
Demographics and Baseline Disease Characteristics
All subjects were male per the inclusion criteria. The majority of subjects were White and not Hispanic or Latino: 47 subjects (85.5%) and 36 subjects (65.5%), respectively. Subject mean (SD) age at informed consent was 5.00 (3.316) years; 42 subjects (76.4%) comprised the <7 years age group and 13 subjects (23.6%) comprised the ≥7 years age group. Subject mean (SD) height and weight were 111.58 (14.923) cm and 25.00 (10.514) kg, respectively.

Overall, 55 subjects (100%) were receiving ELAPRASE ERT. Except for enzymes, the most commonly reported concomitant medication by therapeutic class was anilides, which was used by 26 subjects (47.3%). The observed medication profile was generally consistent with that expected for patients with MPS II.

Disease Progression
Neurodevelopmental Assessments
Cognitive Assessment Analyses

At baseline (n=44), the mean (SD) GCA score was 78.4 (19.11), which is low (2-8 percentile). The median score was 80.0, which is below average (9-24 percentile). The minimum score was 43, which is very low (<2 percentile) and the maximum score was 122, which is high (91-97 percentile). At month 12 (n=31), the mean (SD) GCA score was 83.5 (22.10), which is below average (9-24 percentile) and the mean (SD) change from baseline (n=27) was -0.9 (9.39). The median score was 84.0, which is below average (9-24 percentile). The minimum score was 38, which is very low (<2 percentile) and the maximum score was 122, which is high (91-97 percentile). At month 24 (n=22), the mean (SD) GCA was 76.0 (27.47), which is low (2-8 percentile) and the mean (SD) change from baseline (n=20) was -3.8 (12.71). The median score of 76.0 was consistent with the mean.

Adaptive Behavior Assessment Analyses

At baseline (n=53), the ABC mean (SD) score was 83.7 (14.22), which indicates moderately low adaptive function (3-17 percentile). The minimum score of 57 corresponds to low adaptive function (≤2 percentile); while the maximum score of 126 corresponded to moderately high adaptive function (84-97 percentile). At month 12 (n=30), the ABC mean (SD) score of 83.3 (14.27) was similar to the baseline score. The mean (SD) change from baseline (n=29) was -2.4 (7.64). At month 24 (n=21), the ABC mean (SD) score of 82.1 (13.99) remained moderately low. The mean (SD) change from baseline (n=21) was -2.0 (8.07).

Correlation Between Different Neurodevelopmental Assessment Domains

The DAS-II GCA standard score and the VABS-II ABC score had a moderate, positive correlation (r=0.6871).

Safety

During the study, 497 AEs were reported in 49 subjects (89.1%). Of these, 20 events in 6 subjects (10.9%) were considered related to ELAPRASE ERT treatment; 230 events in 41 subjects (74.5%) were attributed to disease progression; and 12 events in 5 subjects (9.1%) were related to procedures performed in this study.

The most commonly reported PT was pyrexia with 20 events in 14 subjects (25.5%). Upper respiratory tract infections were also commonly reported with 19 events in 12 subjects (21.8%).

The majority of AEs were mild in severity. Moderate events were reported by the largest proportion of subjects. Overall, 19 subjects (34.5%) reported 430 mild events; 26 subjects (47.3%) reported 61 moderate events; and 4 subjects (7.3%) reported the following 6 severe events: 4 PTs of carpal tunnel were reported in 2 subjects (3.6%) and 1 PT of headache and knee deformity each were reported in 1 subject (1.8%). All AEs except for headache were related to disease progression. The knee deformity event was an SAE as it required hospitalization. There were 9 subjects (16.4%) with 13 SAEs; no SAE was considered to be life-threatening. No deaths occurred during the study.

No notable trends were seen in vital sign measurements during the study.

Hearing aids were used by 9 subjects (16.4%) at baseline and 10 subjects (18.2%) at month 24.

CONCLUSIONS

Study HGT-HIT-090 enrolled a population of pediatric MPS II subjects with heterogeneous baseline characteristics. The subjects ranged in age from approximately 2 to 17 years with a majority (76%) being younger than aged 7 years. All subjects were males, per study eligibility criteria and consistent with the X-linked inheritance of MPS II. Most subjects were white (86%) and identified as not Hispanic or Latino (66%). Subjects presented with a variety of IDS genotypes including missense, nonsense, frameshift, splice site, truncation, and large deletion mutations. The most common IDS genotype was missense mutation (60%).

Neurodevelopmental status varied widely among subjects in HGT-HIT-090 with a range of individual GCA scores (43 to 122 points) at baseline consistent with very low to high levels of cognitive ability. Approximately 33% of subjects had a mean baseline GCA score ≤70, indicative of very low cognitive function (<2 percentile), while nearly half of subjects (47%) had baseline GCA scores >70. Overall, for the study population as a whole, cognitive function at baseline was low (2-8 percentile) with a mean (SD) GCA score of 78.4 (19.11). The mean GCA scores at baseline were similar between age groups (baseline age <7 years [n=32] and baseline age ≥7 years [n=12]) in the low (2-8 percentile) range. The mean (SD) GCA scores at baseline for the subgroups with baseline GCA ≤70 (n=18) and baseline GCA>70 (n=26) were in the very low (<2 percentile) and in the average (25-74 percentile) ranges,
respectively.

Modest mean changes from baseline in GCA and its component cluster scores were observed for the overall population over the 2-year observational period. For example, at months 12 and 24, the mean (SD) change from baseline in GCA score was -0.9 (9.39) and -3.8 (12.71) points, respectively. These results may reflect heterogeneity within the study population with respect to the extent of neurological involvement and disease course. Indeed, individual changes from baseline in GCA scores at 24 months spanned a wide range (-36 to 23 points). Additionally, intersubject variability with respect to an individual’s state in relation to his particular trajectory of disease progression would not be unexpected, given that children with neuronopathic MPS II experience plateauing of cognitive and adaptive abilities prior to the onset of rapid neurodevelopmental decline (Holt et al 2011). The greatest mean changes from baseline in GCA at 24 months were observed for the subgroups of subjects with baseline GCA ≤70 (mean [SD] GCA -12.1 (14.51), range -36 to 4) and subjects with baseline age <7 years (mean [SD] GCA -4.8 (14.98), range -36 to 23). In general, declines in cognitive and adaptive functions appeared correlated.

It is noteworthy that a sizeable percentage (44%) of subjects in this observational study discontinued early in order to participate in an ongoing Shire treatment study (HGT-HIT-094) enrolling pediatric subjects with neuronopathic MPS II. The cumulative loss of data for segments of the HGT-HIT-090 population may have had the effect of appearing to stabilize overall mean GCA scores as the population became enriched over time for subjects who may have had reduced propensity for further change in neurodevelopmental status.

There were no notable safety findings in this observational study. The majority of AEs were mild in severity and attributable to disease progression. There were few SAEs; none was life threatening. Concomitant medication use was generally consistent with that expected in the pediatric MPS II population. All subjects received treatment with ELAPRASE as standard of care.

**Date of final report:** 18 Jan 2018