



Long-Term Comparative Efficacy and Safety of Risdiplam and Nusinersen in Children with Type 1 Spinal Muscular Atrophy

Christos Kokaliaris · Rachel Evans · Neil Hawkins · Anadi Mahajan ·
David Alexander Scott · C. Simone Sutherland · Julian Nam ·
Gautam Sajeev

Received: December 18, 2023 / Accepted: March 13, 2024 / Published online: May 5, 2024
© The Author(s) 2024

ABSTRACT

Introduction: Spinal muscular atrophy (SMA) is a severe genetic neuromuscular disease characterized by a loss of motor neurons and progressive muscle weakness. Children with untreated type 1 SMA never sit independently and require increasing levels of ventilatory

support as the disease progresses. Without intervention, and lacking ventilatory support, death typically occurs before the age of 2 years. There are currently no head-to-head trials comparing available treatments in SMA. Indirect treatment comparisons are therefore needed to provide information on the relative efficacy and safety of SMA treatments for healthcare decision-making.

Methods: The long-term efficacy and safety of risdiplam versus nusinersen in children with type 1 SMA was evaluated using indirect treatment comparison methodology to adjust for differences between population baseline characteristics, to reduce any potential bias in the comparative analysis. An unanchored matching-adjusted indirect comparison was conducted using risdiplam data from 58 children in FIREFISH (NCT02913482) and published aggregate nusinersen data from 81 children obtained from the ENDEAR (NCT02193074) and SHINE (NCT02594124) clinical trials with at least 36 months of follow-up.

Prior Presentation: Some of the data included in this manuscript have been presented at the following congresses: International Congress on Neuromuscular Diseases, Brussels, Belgium, 5–9 July 2022; International Scientific Congress on Spinal Muscular Atrophy, Barcelona, Spain, 21–23 October 2022; International Society for Pharmacoeconomics and Outcomes Research Europe, Vienna, Austria, 6–9 November 2022; British Paediatric Neurology Association Annual Conference, Edinburgh, UK, 25–27 January 2023; Muscular Dystrophy Association Clinical and Scientific Conference, Dallas, TX, USA, 19–22 March 2023; European Paediatric Neurology Society Congress, Prague, Czech Republic, 20–24 June 2023; Cure SMA 2023 Research & Clinical Care Meeting, Orlando, FL, USA, 28–30 June 2023; International Annual Congress of the World Muscle Society, Charleston, USA, 3–7 October 2023; British Paediatric Neurology Association, 24–26 January 2024.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12325-024-02845-6>.

C. Kokaliaris (✉) · C. S. Sutherland · J. Nam
F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124,
4070 Basel, Switzerland
e-mail: christos.kokaliaris@roche.com

R. Evans · N. Hawkins · D. A. Scott
Visible Analytics, Oxford, UK

Results: Children with type 1 SMA treated with risdiplam had a 78% reduction in the rate of death, an 81% reduction in the rate of death or permanent ventilation, and a 57% reduction in the rate of serious adverse events compared with children treated with nusinersen. Children treated with risdiplam also had a 45% higher rate of achieving a Hammersmith Infant Neurological Examination, Module 2 motor milestone response and a 186% higher rate of achieving a ≥ 4 -point improvement in Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders compared with children treated with nusinersen.

Conclusion: Long-term data supported risdiplam as a superior alternative to nusinersen in children with type 1 SMA.

Video abstract available for this article.

PLAIN LANGUAGE SUMMARY

Risdiplam and nusinersen are two approved treatments for patients with type 1 spinal muscular atrophy (SMA). There are currently no head-to-head trials that compare the outcomes of these treatments in patients. This study conducted a statistical comparison of the efficacy and safety of risdiplam and nusinersen in children with type 1 SMA who received treatment for at least 36 months. Risdiplam data were collected from 58 children who participated in the FIREFISH trial (NCT02913482). Published combined data were collected from

81 children treated with nusinersen who participated in the ENDEAR (NCT02193074) and SHINE (NCT02594124) trials. Outcomes from the two studies were compared using matching-adjusted indirect comparison (MAIC) methodology. MAIC adjusts for differences in baseline characteristics between patients in two trials to make the populations more similar and reduce bias in the comparison. Results suggested that children with type 1 SMA treated with risdiplam had a 78% reduction in the rate of death and an 81% reduction in the rate of death or permanent ventilation compared with children treated with nusinersen. With risdiplam, children also had a higher rate of achieving motor function responses, and a longer time to the first serious adverse event compared with children treated with nusinersen. These results support risdiplam as a superior alternative to nusinersen in children with type 1 SMA over 36 months of follow-up. Access to long-term data beyond 36 months would allow for additional indirect comparisons between SMA treatments. These comparisons are key to guiding treatment decision-making in the absence of head-to-head trials.

Keywords: Clinical trial; Indirect treatment comparison; MAIC; Matching-adjusted indirect comparison; Propensity score matching; Neuromuscular disease; Nusinersen; Risdiplam; SMA; Spinal muscular atrophy

N. Hawkins
Institute of Health & Wellbeing, University of
Glasgow, Glasgow, UK

A. Mahajan
Bridge Medical Consulting Ltd, London, UK

G. Sajeew
Analysis Group, Inc, Boston, MA, USA

Key Summary Points

Why carry out this study?

Understanding the potential differences in the long-term efficacy and safety of disease-modifying treatments for spinal muscular atrophy (SMA) is important for treatment decision-making.

In the absence of head-to-head trials, indirect treatment comparisons (ITCs) may provide information on the relative efficacy and safety of SMA treatments.

This study conducted a matching-adjusted indirect comparison (MAIC) with assessment schedule matching to evaluate survival, motor function, and safety outcomes of risdiplam (FIREFISH) compared with nusinersen (SHINE-ENDEAR) in children with type 1 SMA over 36 months of treatment.

What was learned from the study?

Results suggested that children treated with risdiplam had improved survival, with a 78% reduction in the rate of death and an 81% reduction in the rate of death or permanent ventilation in children treated with risdiplam compared with those treated with nusinersen. These results also suggested that children treated with risdiplam had greater improvements in motor function, with higher rates of achieving responses on the HINE-2 and the CHOP-INTEND scales (45% and 186% higher rates, respectively) and a longer time to the first serious adverse event compared with children treated with nusinersen.

These findings suggest that children with type 1 SMA treated with risdiplam may see greater improvements compared with children treated with nusinersen over at least 36 months of follow-up.

The results of these analyses support risdiplam as a superior alternative to nusinersen in children with type 1 SMA.

DIGITAL FEATURES

This article is published with digital features, including a video to facilitate understanding of the article. To view digital features for this article, go to <https://doi.org/10.6084/m9.figshare.25398454>.

INTRODUCTION

Spinal muscular atrophy (SMA) is a severe genetic neuromuscular disease characterized by a loss of motor neurons [1]. Motor neuron degeneration in SMA is caused by insufficient levels of the survival of motor neuron (SMN) protein, due to homozygous deletion of or loss of function mutations within the *SMN1* gene [2]. Although a paralogous gene, *SMN2*, produces low levels of SMN protein [3], these are insufficient to compensate for the loss of *SMN1*, leading to progressive muscle weakness that affects respiratory function, swallowing, and motor function [1]. Without treatment, children with type 1 SMA fail to achieve major motor milestones, and never achieve the ability to sit independently [4]. As the disease rapidly progresses, these children require increasing levels of ventilatory and bulbar/feeding support [1, 5–9]. Without intervention and lacking ventilatory support, children with type 1 SMA typically succumb to the condition before the age of 2 years [10–13]. In recent years, the approval of disease-modifying treatments (DMTs) and advances in supportive care practices have changed the natural course of the disease in untreated children with type 1 SMA [13–15]; children with type 1 SMA are surviving longer and achieving motor milestones never before seen in natural history cohorts [7, 13].

There are currently three DMTs approved for the treatment of type 1 SMA: risdiplam (EVRYSDI®), an oral *SMN2* splicing modifier; nusinersen (SPINRAZA®), an intrathecally administered *SMN2*-targeting antisense oligonucleotide; and onasemnogene abeparvovec (ZOLGENSMA®), an intravenously administered gene therapy [16–21]. Although efficacy and safety have been demonstrated in type 1 SMA for these DMTs independently

[22–26], there are currently no head-to-head trials comparing available SMA treatments. Indirect treatment comparisons (ITCs), which compare treatments across individual clinical trials, are therefore needed to provide information on the relative efficacy and safety of SMA treatments for healthcare decision-making [27].

Many factors can affect clinical trial outcomes, including baseline characteristics such as age, genetic factors, and disease severity [28], which may differ between trials as a result of variations in recruitment criteria. There may also be differences in the timings of assessments. In ITCs, cross-trial differences may lead to bias if analyses are left unadjusted [27]. Consequently, ITCs require the use of population-adjustment methodologies that account for cross-trial differences in population baseline characteristics to reduce potential bias in relative effect estimates. One such adjustment methodology is matching-adjusted indirect comparison (MAIC) [29–31].

Previous ITCs based on 12- and 24-month clinical trial data indicated more favorable results for efficacy and safety outcomes with risdiplam compared with nusinersen in type 1 SMA [32, 33]. MAIC was not feasible in a comparison of risdiplam and nusinersen at 12 months in patients with types 2 and 3 SMA as a result of limited overlap between patient populations [32].

The aim of this study was to conduct an updated ITC comparing treatment outcomes of risdiplam versus nusinersen in type 1 SMA after at least 36 months of follow-up.

METHODS

Study Populations and Data Sources

A systematic literature review (SLR) was conducted across several bibliographic databases (Embase®, MEDLINE®, and Cochrane Central) from database inception (prior to 1966) to June 8, 2021, using a previously reported search strategy and inclusion criteria [32]. The SLR was reported according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (Fig. S1). The identified clinical trials on type 1

SMA are indicated in Table S1. At the time of performing these analyses, no additional clinical trials in patients with type 1 SMA had been published since the original SLR for the ITC conducted at 12 months [32].

Data for the comparison of risdiplam and nusinersen in type 1 SMA were obtained from three clinical trials (FIREFISH, ENDEAR, and SHINE). FIREFISH (NCT02913482) was a phase 2/3, single-arm, open-label, two-part study of risdiplam in children aged 1–7 months at enrollment [34]. FIREFISH Part 1 was a dose-finding study; Part 2 was the confirmatory study using the dose selected in Part 1. FIREFISH Parts 1 and 2 had distinct patient populations; patients in Part 1 did not participate in Part 2. This ITC included data from 17 patients from the high-dose (pivotal dose) cohort from FIREFISH Part 1 and all 41 patients from Part 2. ENDEAR (NCT02193074) was a phase 3, randomized, double-blind, multicenter, sham procedure-controlled study in children aged 1–7 months at enrollment [35]. SHINE (NCT02594124) was an open-label extension study that enrolled patients from ENDEAR (and from CS3A [NCT01839656], CS12 [NCT02052791], CHERISH [NCT02292537], and EMBRACE [NCT02462759]) [36]. Data were extracted only from children in SHINE who participated in ENDEAR.

Individual patient data (IPD) from children treated with risdiplam were obtained from the FIREFISH trial (provided by the sponsor) [34]. Aggregate comparator data from children treated with nusinersen in SHINE-ENDEAR were obtained from the publicly available submission dossier to the Federal Joint Committee (Gemeinsamer Bundesausschuss; the highest decision-making body of physicians, dentists, hospitals, and health insurance funds) in Germany [37]. Pseudo IPD for each time-to-event outcome were generated by digitizing the Kaplan–Meier curves available in the submission dossier [38].

Endpoints

The following survival outcomes were included in the analyses: overall survival, defined as time

to death, and event-free survival, defined as time to death or need for permanent ventilation, whichever occurred first. In FIREFISH, permanent ventilation was defined as ≥ 16 h of non-invasive ventilation/day or intubation for > 21 consecutive days in the absence of, or following the resolution of, an acute reversible event or tracheostomy. In SHINE-ENDEAR, permanent ventilation was defined as either tracheostomy or ≥ 16 h of ventilation/day continuously for > 21 days in the absence of an acute reversible event.

Motor function outcomes included in the analyses were time to achievement of a Hammersmith Infant Neurological Examination, Module 2 (HINE-2) response [defined as a ≥ 2 -point increase in the ability to kick (or achievement of the maximal score in that category), or a 1-point increase in head control, rolling, sitting, crawling, standing, or walking] and time to achievement of a Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP-INTEND) response (defined as a ≥ 4 -point improvement in CHOP-INTEND score from baseline).

Relative safety was evaluated according to the time of reporting the first serious adverse event (SAE), defined as any adverse event that is fatal, life-threatening, requires or prolongs inpatient hospitalization, results in persistent or significant disability/incapacity, or is a significant medical event in the investigator's judgment.

Endpoint definitions in FIREFISH and SHINE-ENDEAR are provided in Table S2.

Statistical Analysis Methods

The statistical analysis plan (SAP) is provided in the Supplementary SAP.

Primary Analysis: MAIC

As a result of the lack of a common comparator between FIREFISH and SHINE-ENDEAR, and because IPD were available from FIREFISH, an unanchored MAIC was performed using well-accepted methodology [29, 30, 39].

Unanchored MAICs (referred to as MAIC for the remainder of this paper) attempt to balance

prognostic factors (i.e., factors that affect the natural course of a disease) and effect modifiers (i.e., factors that affect the efficacy of a given treatment) to enable unbiased estimation of comparative effectiveness [40]. Prognostic factors and effect modifiers for type 1 SMA were identified from a previous literature review [28]. Matching baseline characteristics were determined on the basis of these identified prognostic factors and effect modifiers, the baseline characteristics available for comparison across FIREFISH and SHINE-ENDEAR, and feedback from medical experts.

Baseline characteristics selected as adjustment factors were mean age at first dose, disease duration at baseline, and baseline CHOP-INTEND score as an indicator of motor function.

Weighting of the IPD from FIREFISH was conducted on the basis of the baseline characteristics of the treated population in ENDEAR ($N = 80$) [35], as the SHINE cohort included one additional child who was never dosed in ENDEAR, but participated in SHINE ($N = 81$), and baseline data were not available for this individual. Weights were estimated via a logistic regression model that used the probability of being enrolled into FIREFISH based on the covariates of SHINE-ENDEAR [40], so that after weighting the populations were balanced on adjusted factors. The MAIC was bootstrapped ($N = 1000$ samples) and the uncertainty in the estimated weights was reflected in 95% confidence intervals (CIs). Outcomes from the FIREFISH data were then recalculated using the estimated weights. Hazard ratios (HRs) were estimated using Cox proportional-hazards models that allowed handling differences in follow-up time across studies and comparisons over time, and adjusted Kaplan–Meier curves were generated. In these comparisons, 95% confidence intervals (CIs) that did not span 1 indicated a statistically significant difference.

Proportional hazard assumption testing was also performed independently for each time-to-event analysis.

Unadjusted comparisons were also conducted.

Scenario Analyses: MAIC Including Additional Adjustment Factors

Analyses were conducted to evaluate the extent to which baseline characteristics other than those included in the primary analysis affected the FIREFISH effective sample size (ESS) and the HR estimates of the outcomes. These scenario analyses included, in addition to mean age at first dose, disease duration, and CHOP-INTEND score, the following adjustment factors (mean values at baseline): ventilatory support, nutritional support, both ventilatory and nutritional support, gender, HINE-2 score, and both gender and HINE-2 score.

Assessment Schedule Matching

In FIREFISH and SHINE-ENDEAR, motor function (HINE-2 and CHOP-INTEND) was evaluated only at scheduled visits. In the context of ITCs, this may lead to bias in favor of the trial with earlier or more frequently scheduled visits, as improvements could be documented earlier. SHINE-ENDEAR had earlier and more frequent HINE-2 assessments than FIREFISH, and FIREFISH had more frequent CHOP-INTEND assessments than SHINE-ENDEAR (Fig. S2).

Differences in the timing of scheduled visits for the assessment of motor function outcomes were adjusted for using assessment schedule matching (ASM) methodology [41]. This enabled the redistribution of IPD from FIREFISH for alignment with the scheduled visits of SHINE-ENDEAR. The adjustment of the FIREFISH visits is described in Fig. S2. Full details of the ASM methodology are provided in the Supplementary SAP.

Curve-fitting analyses were conducted to define the best-fit model of the relationship between scheduled visits for HINE-2 and CHOP-INTEND assessments across the two trials (Fig. S3). Log-normal was selected for use in the HINE-2 and CHOP-INTEND analyses. MAICs in combination with ASM were then conducted to determine the extent to which differences in scheduled visit timings between FIREFISH and SHINE-ENDEAR affected MAIC for motor function assessments.

This methodology was enhanced compared with that of the 12- and 24-month analyses [32, 42] which were conducted only on

contemporaneous visits that could be aligned between FIREFISH and SHINE-ENDEAR.

Ethical Approval

This study did not require ethical approval as data were obtained from published studies [23, 24, 37]. FIREFISH and SHINE-ENDEAR received institutional review board approval and were conducted in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained for patient participation in these studies. The data published in this article is owned by F. Hoffmann-La Roche. Roche has the necessary permissions to share with readers upon reasonable request.

RESULTS

IPD from ≥ 36 months of follow-up were available from a pooled dataset of 58 children from FIREFISH who received the pivotal dose of risdiplam (Part 1, $n = 17$; Part 2, $n = 41$). Aggregate data from ≥ 43 months of follow-up were available from 81 children in SHINE-ENDEAR who had received nusinersen [37].

Pre-matching Characteristics of FIREFISH and SHINE-ENDEAR Populations

FIREFISH and SHINE-ENDEAR had similar eligibility criteria (Table S3) and thus enrolled patients with relatively comparable disease burden, including pulmonary burden; neither trial would have included patients with invasive ventilation, awake non-invasive ventilation, hypoxemia, or active pulmonary hospitalization/infection at baseline. Pre-matching, the FIREFISH and SHINE-ENDEAR populations were relatively similar in mean age at first dose, disease duration, age at symptom onset, and age at diagnosis (Table 1). However, children in FIREFISH had on average lower baseline HINE-2 (0.9 vs. 1.3) and CHOP-INTEND scores (22.5 vs. 26.6). The need for ventilatory support was slightly higher in FIREFISH (29.3%) compared with SHINE-ENDEAR (26.0%), but the level of

Table 1 Comparison of baseline characteristics of FIREFISH and ENDEAR-SHINE both pre- and post-matching

	Pre-matching		Post-matching ^a	
	SHINE-ENDEAR ^b (N = 80)	FIREFISH ^c (N = 58)	Difference (SHINE-ENDEAR vs. FIREFISH)	Difference (SHINE-ENDEAR vs. FIREFISH) (ESS = 40.6)
Mean age at first dose, months (SD [range])	5.4 (NR [1.4–8.0])	5.3 (1.5 [2.2–7.0])	0.1	0
Mean disease duration, weeks (SD [range])	13.2 (NR [0–25.9])	13.0 (5.9 [1.0–23.3])	0.2	0
Mean CHOP-INTEND score (SD [range])	26.6 (8.1 [NR])	22.5 (6.8 [8.0–37.0])	4.1	0
Patients with invasive ventilation or NIV ^d , n (%)	21 (26.3)	17 (29.3)	– (3.0)	– (6.7)
Patients with nutritional support ^e , n (%)	7 (8.8)	5 (8.6)	– (0.2)	– (2.7)
Mean HINE-2 score (SD [range])	1.3 (1.1 [NR])	0.9 (1.0 [0–5.0])	0.4	0.1
Gender (female), n (%)	43 (53.8)	33 (56.9)	– (3.1)	– (12.9)
Mean age at symptom onset, months (SD [range])	2.0 (NR [0.5–4.5])	1.7 (0.7 [0.9–3.0])	0.3	0.3
Mean age at diagnosis, days (SD [range])	88.2 (NR [0–203.0])	88.9 (42.3 [28.0–185.0])	0.7	6.7
White race, n (%)	NR	31 (53.4)	NA	NA
Patients with two <i>SMN2</i> copies, n (%)	80 (100)	58 (100)	0	NA

BiPAP bilevel positive airway pressure, *CHOP-INTEND* Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders, *ESS* effective sample size, *GI* gastrointestinal, *HINE-2* Hammersmith Infant Neurological Examination, Module 2, *NA* not applicable, *NIV* non-invasive ventilation, *NR* not reported, *SD* standard deviation, *SMN* survival of motor neuron

^aMatching was conducted using MAIC methodology with the following adjustment factors: age at first dose, baseline CHOP-INTEND score, and disease duration

^bMatching involved weighting the baseline characteristics of 80 patients who were enrolled in SHINE-ENDEAR; one additional patient received their first dose of nusinersen in the SHINE trial, but baseline characteristics for this patient were not available [35]

^cIncludes patients from the "high-dose" (pivotal dose) cohort of FIREFISH Part 1 (*n* = 17) and all patients from FIREFISH Part 2 (*n* = 41)

^dSHINE-ENDEAR: undefined. FIREFISH: airway clearance or BiPAP support for < 16 h/day

^eSHINE-ENDEAR: the use of a GI tube. FIREFISH: the ability to swallow/GI tube feeding. Baseline data on GI tube feeding were not available for most patients in Part 1, as the questionnaire was only introduced 6 months after the start of the study. Ability to swallow was used as a proxy for tube feeding for these patients

Table 2 Baseline characteristics of FIREFISH post-matching with SHINE-ENDEAR: unadjusted, primary, and scenario analyses results

	SHINE-ENDEAR		FIREFISH ^a					
	Unadjusted	Primary analysis	Primary analysis + ventilatory support ^b	Primary analysis + ventilatory support ^b & nutritional support ^b	Primary analysis + ventilatory support ^b & nutritional support ^b	Primary analysis + gender	Primary analysis + gender + HINE-2	Primary analysis + gender & HINE-2
Sample size or ESS, <i>n</i>	58	40.6	39.9	40.2	39.7	38.5	39.6	38.2
Mean age at first dose, months	5.3	5.4	5.4	5.4	5.4	5.4	5.4	5.4
Mean disease duration, weeks	13.2	13.2	13.2	13.2	13.2	13.2	13.2	13.2
Mean CHOP-INTEND score	26.6	26.6	26.6	26.6	26.6	26.6	26.6	26.6
Patients with invasive ventilation or NIV, <i>n</i> (%)	21 (26.3)	8.0 (19.6)	10.4 (26.0)	8.6 (21.4)	10.3 (26.0)	8.1 (21.1)	8.3 (20.9)	7.5 (19.6)
Patients with nutritional support, <i>n</i> (%)	7 (8.8)	5 (8.6)	3.2 (7.9)	3.6 (9.0)	3.6 (9.0)	2.9 (7.4)	2.8 (7.3)	2.3 (6.1)
Mean HINE-2 score	1.3	1.2	1.2	1.2	1.2	1.2	1.3	1.3
Gender (female), <i>n</i> (%)	43 (53.8)	33 (56.9)	26.2 (65.6)	26.5 (65.8)	26.0 (65.3)	20.8 (54.0)	26.0 (65.6)	20.6 (54.0)
Mean age at symptom onset, months	2.0	1.7	1.7	1.7	1.7	1.7	1.7	1.7

Table 2 continued

SHINE-ENDEAR	FIREFISH ^a					
	Unadjusted	Primary analysis	Primary analysis + ventilatory support ^b	Primary analysis + ventilatory support + nutritional support ^b	Primary analysis + ventilatory support & nutritional support ^b	Primary analysis + gender & HINE-2
Mean age at diagnosis, days	88.9	94.9	93.7	94.2	93.5	95.6
						94.9
						95.5

BiPAP bilevel positive airway pressure, *CHOP-INTEND* Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders, *ESS* effective sample size, *GI* gastrointestinal, *HINE-2* Hammersmith Infant Neurological Examination, Module 2, *MAIC* matching-adjusted indirect comparison, *NIV* non-invasive ventilation

^aPrimary MAIC analyses were conducted with the following adjustment factors: age at first dose, baseline *CHOP-INTEND* score, and disease duration. In the scenario analyses, additional adjustment factors were included separately for each covariate and in combination, to evaluate the extent to which they affected the matching process individually and to prevent overfitting of the results

^bThe definition and assessment method for ventilatory and nutritional support were not clear for the SHINE-ENDEAR study [36, 37, 65], and thus it was difficult to determine whether definitions and assessment methods differed between FIREFISH and SHINE-ENDEAR. After careful consideration, the authors considered that the most likely comparators or proxies of ventilatory and nutritional support were the use of *BiPAP* (< 16 h per day) and the use of a *GI* tube. Therefore, ventilatory and nutritional support may only serve as proxies, and their use as adjustment factors in our study should be interpreted only as an investigational approach to assess the matching performance of our primary scenario

nutritional support was almost identical (8.6% and 8.8%, respectively).

Post-matching Baseline Characteristics

Matching of the FIREFISH IPD to the SHINE-ENDEAR aggregate data was successful. The distribution of rescaled weights is shown in Fig. S4 and Table S4, and the post-matching distributions of rescaled weights for the adjustment factors in FIREFISH is shown in Fig. S5. The analysis of rescaled weights highlighted large weights associated with a small number of children, and small weights (approaching zero) associated with even fewer children, which indicated a favorable overall matching process.

Post-matching, mean age at first dose, disease duration, and *CHOP-INTEND* score were balanced (Table 1). FIREFISH *ESS* was 40.6 (30% lower than the population enrolled in FIREFISH [$n = 58$]).

When additional adjustment factors were included, the post-matching distributions of weights for the adjustment factors in FIREFISH (Table S4) and the FIREFISH *ESS* (Table 2) remained similar to that of the primary analysis.

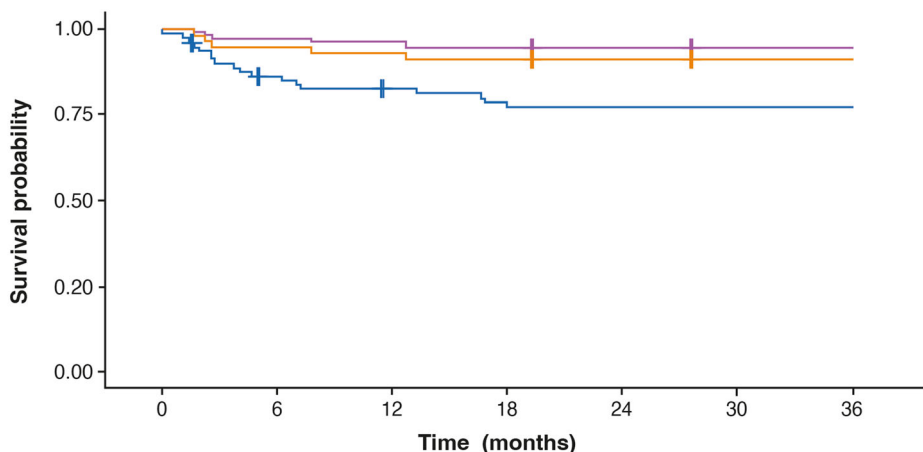
MAIC (Primary Analysis)

Survival and Permanent Ventilation

The number of deaths as well as death or permanent ventilation events are provided in Table S5. MAIC suggested that children treated with risdiplam had a 78% reduction in the rate of death compared with children treated with nusinersen (MAIC HR for overall survival [95% CI], 0.22 [0.04–0.47]; unadjusted HR [95% CI], 0.35 [0.13–0.95]; Fig. 1). Furthermore, results suggested that children treated with risdiplam had an 81% reduction in the rate of death or permanent ventilation compared with those treated with nusinersen (MAIC HR for event-free survival [95% CI], 0.19 [0.07–0.35]; unadjusted HR [95% CI], 0.24 [0.12–0.49]; Fig. 2).

Motor Function Outcomes

MAIC suggested that children treated with risdiplam had a 45% higher rate of achieving a



Treatment (study)	Number at risk						
Risdiplam	58	55	54	53	52	51	51
Risdiplam adjusted	41	39	39	38	37	38	38
Nusinersen	81	68	64	61	60	60	60

Fig. 1 Overall survival in patients with type 1 SMA treated with risdiplam and nusinersen (primary analysis). Unadjusted survival data from the pooled FIREFISH cohort (orange line) were plotted alongside the data from SHINE-ENDEAR (blue line). Patient baseline characteristics in FIREFISH were matched to the mean values of the baseline characteristics of the nusinersen arm in SHINE-ENDEAR using MAIC methodology, thus generating risdiplam-adjusted data (purple line).

Characteristics that were used as adjustment factors in the primary analysis were age at first dose, disease duration, and baseline CHOP-INTEND score. PH test: $p = 0.893$, not rejecting the null hypothesis. *CHOP-INTEND* Children’s Hospital of Philadelphia Infant Test of Neuromuscular Disorders, *MAIC* matching-adjusted indirect comparison, *PH* proportional hazard, *SMA* spinal muscular atrophy

HINE-2 motor milestone response (MAIC HR [95% CI], 1.45 [1.21–1.80]; unadjusted HR [95% CI], 1.15 [0.79–1.69]; Fig. 3) and an approximately threefold increase (186% higher) in the rate of achieving a ≥ 4 -point improvement on the CHOP-INTEND with risdiplam compared with nusinersen (MAIC HR [95% CI], 2.86 [2.18–4.48]; unadjusted HR [95% CI], 2.80 [1.92–4.07]; Fig. 4).

SAEs

MAIC suggested that children treated with risdiplam had a 57% reduction in the rate of SAEs compared with children treated with nusinersen (MAIC HR [95% CI], 0.43 [0.30–0.59]; unadjusted HR [95% CI], 0.45 [0.31–0.66]); Fig. 5).

MAIC (Scenario Analyses)

MAIC including additional adjustment factors provided results consistent with those of the

primary analysis, with similar FIREFISH ESS and HR estimates across all the scenarios explored (Tables 2 and 3). Alignment with the primary analysis indicated that the post-matching discrepancies in baseline characteristics between FIREFISH and SHINE-ENDEAR may be attributed to the inherent variability in CIs (Table 3).

Assessment Schedule Matching

Results from MAIC with ASM continued to support the suggestion that children treated with risdiplam had a higher rate of achieving a response on the HINE-2 and CHOP-INTEND compared with children who received nusinersen (Figs. S6 and S7, respectively). Compared with MAIC without ASM, the HR for the achievement of a HINE-2 motor milestone response increased to 1.58 (95% C 1.28–2.06) following MAIC with ASM. The HR for the time to achievement of a ≥ 4 -point improvement in CHOP-INTEND score decreased to 2.43 (95% C

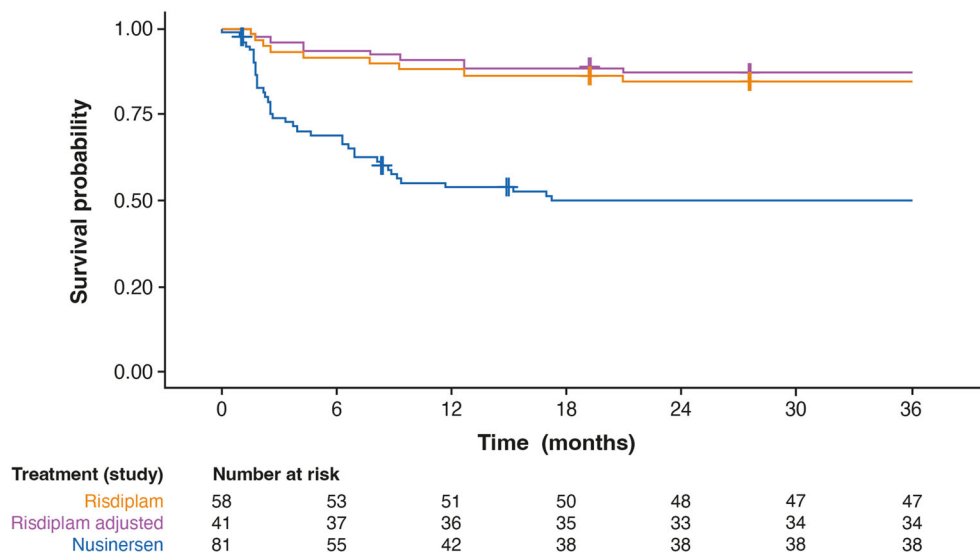


Fig. 2 Event-free survival in patients with type 1 SMA treated with nusinersen and risdiplam (primary analysis). Event-free survival is defined as alive without the need for permanent ventilation. In FIREFISH, permanent ventilation was defined as ≥ 16 h of NIV/day or intubation for > 21 consecutive days in the absence of, or following the resolution of, an acute reversible event or tracheostomy. In SHINE-ENDEAR, permanent ventilation was defined as either tracheostomy or ≥ 16 h of ventilation/day continuously for > 21 days in the absence of an acute reversible event. Unadjusted survival data from the pooled FIREFISH cohort (orange line) were plotted alongside the data from SHINE-ENDEAR (blue line). Patient baseline

characteristics in FIREFISH were matched to the mean values of the baseline characteristics of the nusinersen arm in SHINE-ENDEAR using MAIC methodology, thus generating risdiplam-adjusted data (purple line). Characteristics that were used as adjustment factors in the primary analysis were age at first dose, disease duration, and baseline CHOP-INTEND score. PH test: $p = 0.415$, not rejecting the null hypothesis. CHOP-INTEND Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders, MAIC matching-adjusted indirect comparison, NIV non-invasive ventilation, PH proportional hazard, SMA spinal muscular atrophy

1.90–3.61), but CIs overlapped, which was still suggestive of a superiority of risdiplam over nusinersen ($p > 0.05$).

DISCUSSION

Understanding potential differences in the long-term efficacy and safety of DMTs is important for patients, physicians, healthcare, and reimbursement authorities, and is a key requirement to optimize SMA treatment decision-making. In the absence of head-to-head trial comparisons, this study used MAIC methodology to conduct a balanced comparison of risdiplam and nusinersen in terms of survival, motor function, and the time to experiencing a first SAE. This methodology

accounted for differences in baseline characteristics between trial populations.

Outcomes in type 1 SMA were included in this ITC, which compares the longest possible follow-up data that have been published (at least 36 months) from three robust clinical trials (FIREFISH, ENDEAR, and SHINE).

MAIC suggested that children treated with risdiplam had prolonged survival and survival free of permanent ventilation compared with children treated with nusinersen over 36 months, and they had a higher rate of achieving HINE-2 and CHOP-INTEND responses. Children treated with risdiplam had a longer time to experiencing a first SAE when compared with nusinersen. These results were consistent with previous MAICs conducted at 12 and 24 months [32, 33].

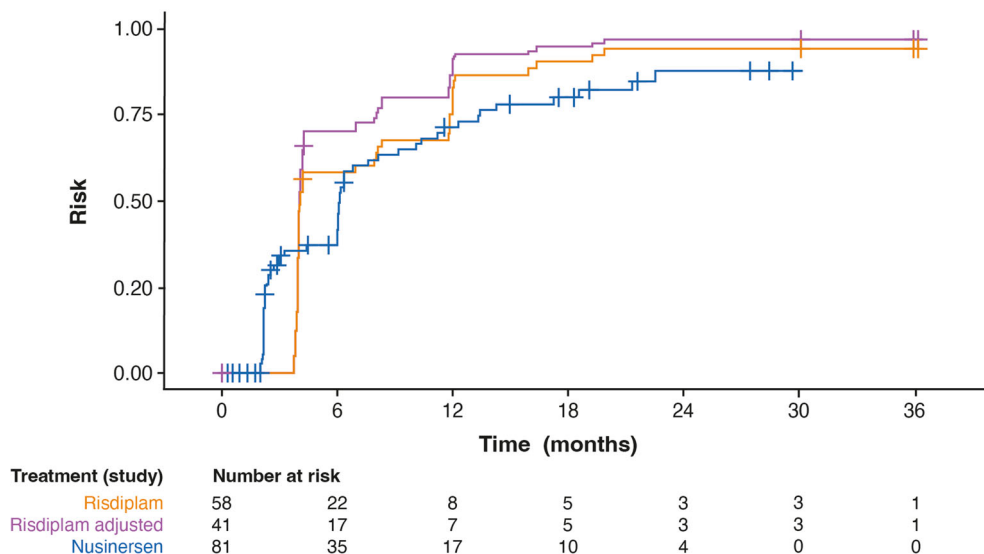


Fig. 3 Time to HINE-2 motor milestone response in patients with type 1 SMA treated with risdiplam and nusinersen (primary analysis). Children were classed as a HINE-2 responder if more motor milestones showed improvement than worsening. Improvement was defined as a ≥ 2 -point increase in the ability to kick (or maximal score) or a ≥ 1 -point increase in head control, rolling, sitting, crawling, standing, or walking. Worsening was defined as a ≥ 2 -point decrease in the ability to kick (or lowest score) or a ≥ 1 -point decrease in head control, rolling, sitting, crawling, standing, or walking. Unadjusted survival data from the pooled FIREFISH cohort (orange line) were plotted alongside the data from SHINE-

ENDEAR (blue line). Patient baseline characteristics in FIREFISH were matched to the mean values of the baseline characteristics of the nusinersen arm in SHINE-ENDEAR using MAIC methodology, thus generating risdiplam-adjusted data (purple line). Characteristics that were used as adjustment factors in the primary analysis were age at first dose, disease duration, and baseline CHOP-INTEND score. PH test: $p = 0.003$, rejecting the null hypothesis. CHOP-INTEND Children’s Hospital of Philadelphia Infant Test of Neuromuscular Disorders, HINE-2 Hammersmith Infant Neurological Examination, Module 2, MAIC matching-adjusted indirect comparison PH proportional hazard, SMA spinal muscular atrophy

Differences between risdiplam and nusinersen in their mode of administration may have contributed to the superiority of risdiplam for survival, motor function, and SAE outcomes suggested by our results. Risdiplam is an oral treatment, which enables systemic distribution throughout the bloodstream [43]. This has been proven to increase levels of functional SMN protein in both the central nervous system and peripheral tissues in animals [43, 44], and it is expected to do the same in humans. In addition, risdiplam crosses the blood–brain barrier [43], which is expected to lead to wide, homogeneous distribution of the drug along the spinal axis, particularly in areas innervating the upper limbs and respiratory muscles. In contrast, nusinersen is intrathecally administered [45], which is expected to lead to uneven drug

distribution. Indeed, higher concentrations of nusinersen have been reported in the lumbar and thoracic regions of the spinal cord [46, 47], which may potentially limit the clinical benefits of nusinersen in the upper limbs, and respiratory and bulbar muscles.

In the present study, the FIREFISH and SHINE-ENDEAR populations were effectively matched. MAIC suggested a significant overlap between the two trial populations, with a minor reduction in the FIREFISH ESS.

When selecting adjustment factors, we considered the CHOP-INTEND score to be more relevant than the HINE-2 score to indicate differences in baseline motor function in the very young and severely affected SMA population used in this study. This is because CHOP-INTEND was developed specifically for children

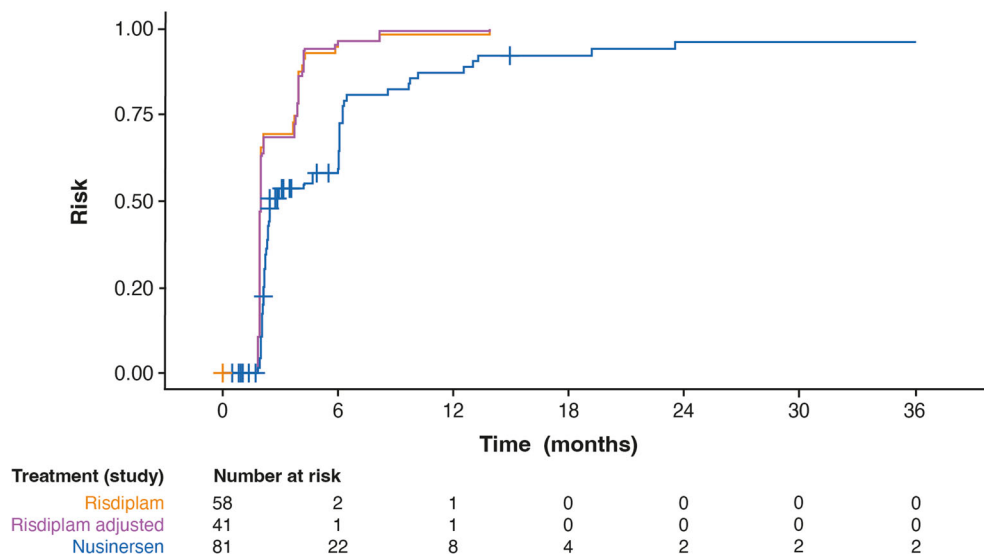


Fig. 4 Time to CHOP-INTEND response in patients with type 1 SMA treated with risdiplam and nusinersen (primary analysis). A CHOP-INTEND response was defined as a ≥ 4 -point improvement in CHOP-INTEND score from baseline. Unadjusted survival data from the pooled FIREFISH cohort (orange line) were plotted alongside the data from SHINE-ENDEAR (blue line). Patient baseline characteristics in FIREFISH were matched to the mean values of the baseline characteristics of the nusinersen arm in SHINE-ENDEAR using MAIC

methodology, thus generating risdiplam-adjusted data (purple line). Characteristics that were used as adjustment factors in the primary analysis were age at first dose, disease duration, and baseline CHOP-INTEND score. PH test: $p = 0.020$, rejecting the null hypothesis. *CHOP-INTEND* Children’s Hospital of Philadelphia Infant Test of Neuromuscular Disorders, *MAIC* matching-adjusted indirect comparison, *PH* proportional hazard; *SMA* spinal muscular atrophy

with type 1 SMA and is more granular in its scoring than HINE-2 [48, 49]. When the HINE-2 score was included as an additional adjustment factor, the resulting FIREFISH ESS was consistent with primary analysis results, indicating that the CHOP-INTEND score was sufficient to provide baseline assessment of motor function.

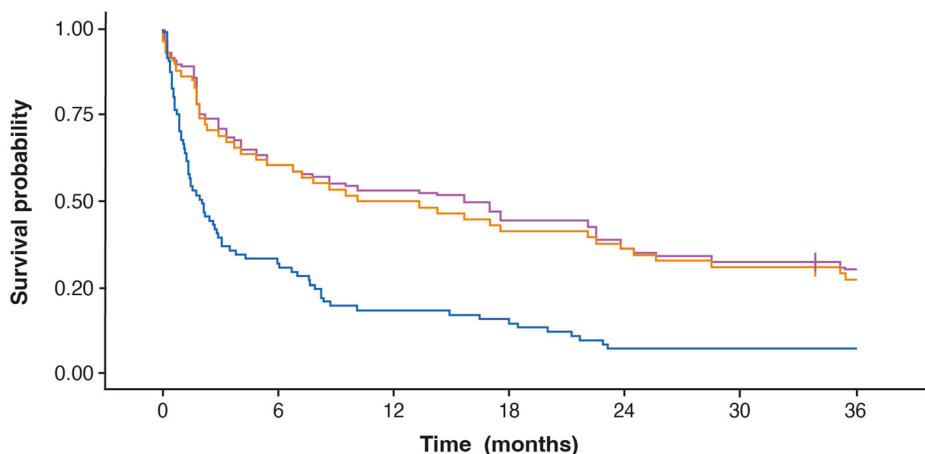
Although use of ventilatory (not “permanent ventilation”) and nutritional support are considered to have prognostic/predictive value for treatment outcomes in SMA [28], there were differences in how these were determined or reported across the trials. When these factors were included as additional adjustment factors in scenario analyses, results were consistent with the primary analysis results, indicating that imbalances in these factors did not affect MAIC outcomes.

Gender-related effects on SMA severity have recently been reported [50, 51] but no studies have yet examined the prognostic/predictive

value of gender on motor function in type 1 SMA [28]. Although the primary analysis revealed an imbalance in the percentage of female participants, the inclusion of gender as an adjustment factor yielded a FIREFISH ESS consistent with the primary analysis results, indicating that gender-based differences did not affect MAIC outcomes.

Taken together, results of the scenario analyses demonstrated that the mean age at first dose, disease duration, and CHOP-INTEND score were the main factors contributing to differences between the FIREFISH and SHINE-ENDEAR populations at baseline and that these were sufficient for use as adjustment factors in the population adjustment.

Mean age at symptom onset was excluded as an adjustment factor since it is a function of mean age at first dose and disease duration (already included in the matching algorithm). Furthermore, as both FIREFISH and SHINE-



Treatment (study)	Number at risk						
Risdiplam	58	35	29	24	21	18	15
Risdiplam adjusted	41	25	21	17	15	12	11
Nusinersen	81	25	15	12	6	6	6

Fig. 5 Time to first SAE in patients with type 1 SMA treated with nusinersen and risdiplam (primary analysis). Unadjusted survival data from the pooled FIREFISH cohort (orange line) were plotted alongside the data from SHINE-ENDEAR (blue line). Patient baseline characteristics in FIREFISH were matched to the mean values of the baseline characteristics of the nusinersen arm in SHINE-ENDEAR using MAIC methodology, thus generating risdiplam-adjusted data (purple line).

Characteristics that were used as adjustment factors in the primary analysis were age at first dose, disease duration, and baseline CHOP-INTEND score. PH test: $p = 0.263$, not rejecting the null hypothesis. *CHOP-INTEND* Children’s Hospital of Philadelphia Infant Test of Neuromuscular Disorders, *MAIC* matching-adjusted indirect comparison, *PH* proportional hazard, *SAE* serious adverse event, *SMA* spinal muscular atrophy

ENDEAR recruited only patients with two *SMN2* copies, it was not necessary to include copy number as an adjustment factor.

This study improved upon the methodology used in previous ITCs of risdiplam versus nusinersen [32, 33] by combining MAIC with ASM. This reduced potential biases from between-trial differences in the timing of scheduled visits for motor function assessments. ASM demonstrated that our MAIC findings on efficacy and safety were consistent regardless of the differing assessment schedules across FIREFISH and SHINE-ENDEAR.

The results from this study are applicable to patients with type 1 SMA and may not be generalizable across the SMA disease spectrum. Long-term data comparisons of risdiplam against nusinersen in other SMA types have not yet been conducted. Similarly, long-term data comparisons of risdiplam with other SMA treatments are not available to date. To our knowledge, only one other ITC has compared

risdiplam with onasemnogene abeparvovec [52]; it reported increased motor function outcomes with onasemnogene abeparvovec relative to risdiplam; however, this was conducted after 8 months of follow-up and without adjustment for known prognostic/predictive factors. In a previous MAIC conducted after 12 months of follow-up [32], differences in study characteristics could not be sufficiently controlled, making MAIC of risdiplam and onasemnogene abeparvovec unfeasible. No new published data were available, so comparison against onasemnogene abeparvovec was not feasible at 36 months.

A recent review by Jiang et al. [53] sought to provide perspectives on MAICs in SMA; we agree with the MAIC best practices highlighted by the authors, and thus we continue to apply the same practices throughout our works. Specifically, in the previous 12-month MAIC [32] and in the present study, comparisons were made for the treated populations in SHINE-

Table 3 HRs (95% CI) for the probability of each endpoint of interest in FIREFISH pre- and post-matching, including the primary and scenario analyses

Endpoints	Unadjusted	Primary analysis ^a	Primary analysis + ventilatory support ^b	Primary analysis + nutritional support ^b	Primary analysis + ventilatory & nutritional support ^b	Primary analysis + gender	Primary analysis + HINE-2	Primary analysis + gender & HINE-2
Overall survival	0.35 (0.13–0.95)	0.22 (0.04–0.47)	0.20 (0.07–0.60)	0.21 (0.07–0.64)	0.20 (0.07–0.61)	0.22 (0.07–0.69)	0.20 (0.07–0.62)	0.21 (0.07–0.67)
Event-free survival	0.24 (0.12–0.49)	0.19 (0.07–0.35)	0.18 (0.08–0.40)	0.19 (0.09–0.43)	0.19 (0.09–0.41)	0.21 (0.09–0.50)	0.17 (0.08–0.37)	0.20 (0.08–0.46)
Time to achievement of a HINE-2 response ^c	1.15 (0.79–1.69)	1.45 (1.21–1.80)	1.41 (0.98–2.03)	1.41 (0.98–2.03)	1.40 (0.97–2.03)	1.34 (0.93–1.95)	1.45 (1.01–2.09)	1.36 (0.94–1.97)
Time to achievement of a CHOP-INTEND response ^d	2.80 (1.92–4.08)	2.86 (2.18–4.48)	2.55 (1.62–4.01)	2.52 (1.60–3.96)	2.48 (1.56–3.95)	2.51 (1.59–3.97)	2.65 (1.73–4.05)	2.49 (1.58–3.92)
Time to first SAE ^e	0.45 (0.31–0.66)	0.43 (0.30–0.59)	0.34 (0.20–0.58)	0.35 (0.21–0.59)	0.34 (0.20–0.58)	0.40 (0.24–0.66)	0.35 (0.21–0.59)	0.40 (0.24–0.66)

BIPAP bilevel positive airway pressure, *CHOP-INTEND* Children’s Hospital of Philadelphia Infant Test of Neuromuscular Disorders, *CI* confidence interval, *GI* gastrointestinal, *HINE-2* Hammersmith Infant Neurological Examination, Module 2, *HR* hazard ratio, *MAIC* matching-adjusted indirect comparison, *SAE* serious adverse event

^aPrimary MAIC analyses were conducted with the following adjustment factors: age at first dose, baseline CHOP-INTEND score, and disease duration. In the scenario analyses, additional adjustment factors were included separately for each covariate and in combination, to evaluate the extent to which they affected the matching process individually and to prevent overfitting of the results

^bThe definition and assessment method for ventilatory and nutritional support were not clear for the SHINE-ENDEAR study [36, 37, 65], and thus it was difficult to determine whether definitions and assessment methods differed between FIREFISH and SHINE-ENDEAR. After careful consideration, the authors considered that the most likely comparators or proxies of ventilatory and nutritional support were the use of BIPAP (< 16 h per day) and the use of a GI tube. Therefore, ventilatory and nutritional support may only serve as proxies, and their use as adjustment factors in our study should be interpreted only as an investigational approach to assess the matching performance of our primary scenario

^cChildren were classed as a HINE-2 responder if more motor milestones showed improvement than worsening. Improvement was defined as a ≥ 2-point increase in the ability to kick (or maximal score) or a ≥ 1-point increase in head control, rolling, sitting, crawling, standing, or walking. Worsening was defined as a ≥ 2-point decrease in the ability to kick (or lowest score) or a ≥ 1-point decrease in head control, rolling, sitting, crawling, standing, or walking

^dA CHOP-INTEND response was defined as a ≥ 4-point improvement in CHOP-INTEND score from baseline

^eAn SAE is defined as any adverse event that is fatal (i.e., actually causes or leads to death); life-threatening (i.e., in the view of the investigator, places the patient at immediate risk of death); requires or prolongs inpatient hospitalization; results in persistent or significant disability/incapacity (i.e., results in substantial disruption of the patient’s ability to conduct normal life functions); or is a significant medical event in the investigator’s judgment (e.g., may jeopardize the patient or may require medical/surgical intervention to prevent one of the outcomes listed above)

ENDEAR and FIREFISH upon evaluation of inclusion/exclusion criteria and baseline characteristic data, demonstrating that both trials enrolled patients of comparable disease burden, including pulmonary burden. Post-matching for documented prognostic/predictive factors, these populations were even more similar.

ITCs and external comparisons are considered useful tools for treatment decision-making and are used by regulatory agencies and reimbursement authorities in SMA and other neuromuscular diseases [27, 40, 54–61]. We acknowledge that ITCs carry their own strengths and limitations and are not a replacement for high-quality, randomized clinical trials. Although Jiang et al. caution against the use of MAIC for drawing conclusions in health technology assessment (HTA) appraisals [53], the method has been used to assess the risk–benefit balance of leukemia treatment for marketing authorization [62], showing that it is accepted among healthcare decision-makers for the evaluation of treatment efficacy and safety.

HTA authorities commented explicitly on the previous 12-month MAIC [32], which was conducted in the same patient population as this 36-month analysis. While considering the limitations of the 12-month MAIC, HTA authorities stated that the study was justified in the absence of head-to-head trials [63], the propensity score matching resulted in reasonably balanced baseline characteristics [64], and the overall MAIC results were acceptable [55]. In this revised analysis, steps have been taken to address previous criticism. Baseline characteristics of children in FIREFISH (risdiplam) were matched to those of children from the nusinersen arm of SHINE-ENDEAR, excluding the best supportive care population. Furthermore, we conducted a series of sensitivity analyses on alternative adjustment factors to test matching stability and compare with the primary analysis results. ASM was also conducted to minimize potential bias from differences in the timing of scheduled assessments.

Study Limitations

Our comparisons were limited by the scope of the publicly available data. IPD were not available for nusinersen, but access to them would have increased the robustness of our comparisons as would the inclusion of additional treatment outcomes (e.g., swallowing, fatigue, or caregiver-reported outcomes). Safety outcomes other than time to first SAE were not considered in this study because of a lack of comparable data across trials.

Whilst our population adjustment was based on known prognostic/predictive factors [28], the choice of adjustment factors was limited by the availability of baseline characteristics reported in both studies. In addition, results may be confounded (in any direction) by unadjusted baseline differences derived from unreported prognostic factors or effect modifiers.

ASM accounted for differences in assessment schedules between FIREFISH and SHINE-ENDEAR; however, this involved simplifications and assumptions [41].

The sample size of the patient populations compared in our study was small. This was especially the case for the FIREFISH sample size, where the pre-matching population of 58 patients was further reduced to an ESS of 40.6 post-matching. Larger sample sizes would have allowed for more robust comparisons; however, it must be noted that these sample sizes are not unusual in clinical trials investigating a rare disease.

There may have been heterogeneity in standard of care (SoC) across clinical sites and over time. However, the two trials had contemporaneous periods: patients treated with nusinersen first enrolled in ENDEAR in August 2014 [65], and migrated into SHINE-ENDEAR in November 2016 [36], with FIREFISH starting in December 2016 [34]. In addition, both trials were conducted globally in many of the same clinical sites [24, 66], with a protocol that encouraged

site investigators to meet local/national healthcare considerations, especially those focused on the respiratory, gastrointestinal/nutritional, and physical therapy management of study participants [67–69]; specifically, 7/13 countries (Belgium, France, Italy, Japan, Spain, Turkey, and the USA) participated in both SHINE-ENDEAR and FIREFISH. Therefore, any differences we observed in the present study were not likely attributable to differences in SoC.

CONCLUSIONS

MAIC analyses suggested that children with type 1 SMA treated with risdiplam had improved survival and greater motor function responses and experienced a longer time to their first SAE compared with children treated with nusinersen. These results may assist patients, physicians, regulatory agencies, and payers when deciding on the optimal SMA treatment for patients with type 1 SMA.

Public disclosure of follow-up data beyond 36 months from FIREFISH and SHINE-ENDEAR would allow longer-term comparisons (including later-onset SMA), thus providing further conclusions on the efficacy and safety of these SMA treatments. Future ITCs would greatly benefit from the inclusion of safety data on risdiplam and nusinersen that are comparable across trials (beyond time to first SAE), enabling a more extensive evaluation of safety outcomes.

ACKNOWLEDGEMENTS

We would like to thank the individuals with SMA and their families, as well as the investigators and staff involved in the FIREFISH clinical trial. We would also like to thank our collaborators at PTC Therapeutics and the SMA Foundation for their contributions to the FIREFISH clinical trial.

Medical Writing and Editorial Assistance. Medical writing and editorial assistance were provided by Laura Pérez-Pachón, MSc, PhD, and Lauren Walmsley, PhD, of Nucleus Global, UK, in accordance with Good Publication Practice (GPP) 2022 guidelines (<https://www.ismpp.org/gpp-2022>). Medical writing and editorial assistance, and the creation of the video abstract, were funded by F. Hoffmann-La Roche Ltd, Basel, Switzerland.

Author Contributions. Christos Kokaliaris was responsible for formulating research goals and objectives, planning and coordinating the analysis, assisting in data extraction, supervising the manuscript writing, and interpreting the results. Rachel Evans was responsible for the data extraction and processing the analysis. Neil Hawkins was responsible for processing the analysis and providing important intellectual content on methodology. Anadi Mahajan was responsible for the study implementation, including development of the search strategy and contribution to the SLR and data extraction. David Alexander Scott was responsible for providing important intellectual content on the methodology and interpretation of results. C Simone Sutherland was responsible for providing important intellectual content on research goals, objectives, and interpreting the results. Julian Nam was responsible for providing important intellectual content on research goals and objectives, and interpreting the results. Gautam Sajeev was responsible for providing important intellectual content on the methodology and interpretation of results. All authors contributed to the drafting and revision of this manuscript and approved it for submission.

Funding. This study was funded by F. Hoffmann-La Roche Ltd, Basel, Switzerland. F. Hoffmann-La Roche Ltd, Basel, Switzerland also funded the Rapid Service and Open Access fees for the publication of this article.

Data Availability. All data generated or analyzed during this study are included in this article and its accompanying supplementary material, or available from the corresponding author upon reasonable request.

Declarations

Conflict of Interest. The SLR and data extraction were conducted by Bridge Medical Consulting Ltd., London, UK and funded by F. Hoffmann-La Roche Ltd., Basel, Switzerland. ITC analyses were designed by Visible Analytics, UK, funded by F. Hoffmann-La Roche Ltd. Christos Kokaliaris and Julian Nam are employees of F. Hoffmann-La Roche Ltd. Julian Nam owns stocks in Roche. Rachel Evans is a former employee of Visible Analytics Ltd., which designed the analysis and received consultancy fees and expenses from F. Hoffmann-La Roche Ltd. Neil Hawkins and David Alexander Scott are partners/employees of Visible Analytics Ltd., which designed the analysis and received consultancy fees and expenses from F. Hoffmann-La Roche Ltd. Anadi Mahajan is an employee of Bridge Medical Consulting Ltd., which conducted this SLR and received consultancy fees and expenses from F. Hoffmann-La Roche Ltd. C Simone Sutherland was employed by F. Hoffmann-La Roche Ltd and owns stocks in Roche. Gautam Sajeev is an employee of Analysis Group, Inc. which received consulting fees from Genentech Inc. (a member of the Roche group) for this research. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Ethical Approval. This study did not require ethical approval as data were obtained from published studies. FIREFISH and ENDEAR received institutional review board approval and were conducted in accordance with the principles of the Declaration of Helsinki. Informed consent was obtained for patient participation in these studies. The data published in this article is owned by F. Hoffmann-

La Roche. Roche has the necessary permissions to share with readers upon reasonable request.

Open Access. This article is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Mercuri E, Sumner CJ, Muntoni F, Darras BT, Finkel RS. Spinal muscular atrophy. *Nat Rev Dis Primers*. 2022;8(1):52.
2. Lefebvre S, Burglen L, Reboullet S, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. 1995;80(1):155–65.
3. Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci USA*. 1999;96(11):6307–11.
4. Munsat TL, Davies KE. International SMA consortium meeting. (26–28 June 1992, Bonn, Germany). *Neuromuscul Disord*. 1992;2(56):423–8.
5. Kaneko K, Arakawa R, Urano M, Aoki R, Saiito K. Relationships between long-term observations of motor milestones and genotype analysis results in childhood-onset Japanese spinal muscular atrophy patients. *Brain Dev*. 2017;39(9):763–73.
6. Wadman RI, Stam M, Gijzen M, et al. Association of motor milestones, SMN2 copy and outcome in spinal muscular atrophy types 0–4. *J Neurol Neurosurg Psychiatry*. 2017;88(4):365–7.

7. De Amicis R, Baranello G, Foppiani A, et al. Growth patterns in children with spinal muscular atrophy. *Orphanet J Rare Dis.* 2021;16(1):375.
8. van der Heul AMB, Cuppen I, Wadman RI, et al. Feeding and swallowing problems in infants with spinal muscular atrophy type 1: an observational study. *J Neuromuscul Dis.* 2020;7(3):323–30.
9. Choi YA, Suh DI, Chae JH, Shin HI. Trajectory of change in the swallowing status in spinal muscular atrophy type I. *Int J Pediatr Otorhinolaryngol.* 2020;130: 109818.
10. Mercuri E, Bertini E, Iannaccone ST. Childhood spinal muscular atrophy: controversies and challenges. *Lancet Neurol.* 2012;11(5):443–52.
11. Zerres K, Rudnik-Schoneborn S. Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications. *Arch Neurol.* 1995;52(5): 518–23.
12. Finkel RS, McDermott MP, Kaufmann P, et al. Observational study of spinal muscular atrophy type I and implications for clinical trials. *Neurology.* 2014;83(9):810–7.
13. Burr P, Reddivari AKR. Spinal muscle atrophy. 2023. <https://www.ncbi.nlm.nih.gov/books/NBK560687/>. Accessed Feb 2024.
14. Bach JR, Sinqee DM, Saporito LR, Botticello AM. Efficacy of mechanical insufflation-exsufflation in extubating unweanable subjects with restrictive pulmonary disorders. *Respir Care.* 2015;60(4): 477–83.
15. Bach JR, Reyna M, Hon A. Spinal muscular atrophy type 1: prolongation of survival by noninvasive respiratory aids. *Pediatr Asthma Allergy Immunol.* 2009;22(4):151–62.
16. Biogen. SPINRAZA® (nusinersen) [package insert] 2020. <https://www.spinraza.com/PI>. Accessed Feb 2024.
17. Biogen. SPINRAZA® (nusinersen) summary of product characteristics. https://www.ema.europa.eu/en/documents/product-information/spinraza-epar-product-information_en.pdf. Accessed Feb 2024.
18. Roche. EVRYSIDI® (risdiplam) prescribing information 2020. https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/213535s000lbl.pdf. Accessed Feb 2024.
19. Roche. EVRYSIDI® (risdiplam) summary of product characteristics. https://www.ema.europa.eu/en/documents/product-information/evrysdi-epar-product-information_en.pdf. Accessed Feb 2024.
20. Novartis. ZOLGENSMA® (onasemnogene abepar-ovec-xioi) [package insert] Bannockburn, IL2019. <https://www.fda.gov/media/126109/download>. Accessed Feb 2024.
21. Novartis. ZOLGENSMA® (onasemnogene abepar-ovec-xioi) summary of product characteristics 2020. https://www.ema.europa.eu/en/documents/product-information/zolgensma-epar-product-information_en.pdf. Accessed Feb 2024.
22. Darras BT, Farrar MA, Mercuri E, et al. An integrated safety analysis of infants and children with symptomatic spinal muscular atrophy (SMA) treated with nusinersen in seven clinical trials. *CNS Drugs.* 2019;33(9):919–32.
23. Masson R, Mazurkiewicz-Beldzinska M, Rose K, et al. Safety and efficacy of risdiplam in patients with type 1 spinal muscular atrophy (FIREFISH part 2): secondary analyses from an open-label trial. *Lancet Neurol.* 2022;21(12):1110–9.
24. Baranello G, Darras BT, Day JW, et al. Risdiplam in type 1 spinal muscular atrophy. *N Engl J Med.* 2021;384(10):915–23.
25. Mercuri E, Muntoni F, Baranello G, et al. Onasemnogene abeparovec gene therapy for symptomatic infantile-onset spinal muscular atrophy type 1 (STRIVE-EU): an open-label, single-arm, multicentre, phase 3 trial. *Lancet Neurol.* 2021;20(10):832–41.
26. Day JW, Finkel RS, Chiriboga CA, et al. Onasemnogene abeparovec gene therapy for symptomatic infantile-onset spinal muscular atrophy in patients with two copies of SMN2 (STRIVE): an open-label, single-arm, multicentre, phase 3 trial. *Lancet Neurol.* 2021;20(4):284–93.
27. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC good practice task force report. *Value Health.* 2014;17(2):157–73.
28. Baranello G, Gorni K, Daigl M, et al. Prognostic factors and treatment-effect modifiers in spinal muscular atrophy. *Clin Pharmacol Ther.* 2021;110: 1435–54.
29. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. *Pharmacoeconomics.* 2010;28(10):935–45.
30. Signorovitch JE, Sikirica V, Erder MH, et al. Matching-adjusted indirect comparisons: a new

- tool for timely comparative effectiveness research. *Value Health*. 2012;15(6):940–7.
31. Jiang Y, Ni W. Performance of unanchored matching-adjusted indirect comparison (MAIC) for the evidence synthesis of single-arm trials with time-to-event outcomes. *BMC Med Res Methodol*. 2020;20(1):241.
 32. Aponte Ribero V, Daigl M, Marti Y, et al. How does risdiplam compare with other treatments for types 1–3 spinal muscular atrophy: a systematic literature review and indirect treatment comparison. *J Comp Eff Res*. 2022;11(5):347–70.
 33. Marti-Gil Y, Aponte Ribero V, Gorni K, et al. CO135 Matching-adjusted indirect comparison of risdiplam versus nusinersen in type 1 spinal muscular atrophy: 2-year update. *Value Health*. 2022;25(12):S44.
 34. ClinicalTrials.gov. NCT02913482: a study to investigate the safety, tolerability, pharmacokinetics, pharmacodynamics and efficacy of RO7034067 in infants with type 1 spinal muscular atrophy (FIRE-FISH). <https://clinicaltrials.gov/ct2/show/NCT02913482>. Accessed Feb 2024.
 35. Finkel RS, Mercuri E, Darras BT, et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N Engl J Med*. 2017;377(18):1723–32.
 36. ClinicalTrials.gov. NCT02594124: a study for participants with spinal muscular atrophy (SMA) who previously participated in nusinersen (ISIS 396443) investigational studies. (SHINE). <https://clinicaltrials.gov/ct2/show/NCT02594124>. Accessed Feb 2024.
 37. Biogen. German benefit assessment dossier 2020. https://www.g-ba.de/downloads/92-975-4215/2020-12-01_Modul4A.4_Nusinersen.pdf. Accessed Feb 2024.
 38. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol*. 2012;12:9.
 39. Phillippo DM, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. Methods for population-adjusted indirect comparisons in health technology appraisal. *Med Decis Mak*. 2018;38(2):200–11.
 40. Phillippo AE, Dias S, Palmer S, Abrams KR, Welton N. NICE DSU technical support document 18: Methods for population-adjusted indirect comparisons in submissions to NICE NICE2016. Updated December. https://research-information.bris.ac.uk/ws/portalfiles/portal/94868463/Population_adjustment_TSD_FINAL.pdf. Accessed Feb 2024.
 41. Kapetanakis V, Prawitz T, Schlichting M, et al. Assessment-schedule matching in unanchored indirect treatment comparisons of progression-free survival in cancer studies. *Pharmacoeconomics*. 2019;37(12):1537–51.
 42. Hawkins N, Gorni K, Daigl M, et al. Matching-adjusted indirect comparison of risdiplam versus nusinersen in type 1 spinal muscular atrophy: 2-year update. 17th international congress on neuromuscular diseases, 5–9 July 2022. Brussels, Belgium.
 43. Poirier A, Weetall M, Heinig K, et al. Risdiplam distributes and increases SMN protein in both the central nervous system and peripheral organs. *Pharmacol Res Perspect*. 2018;6:e00447.
 44. Ratni H, Ebeling M, Baird J, et al. Discovery of risdiplam, a selective survival of motor neuron-2 (SMN2) gene splicing modifier for the treatment of spinal muscular atrophy (SMA). *J Med Chem*. 2018;61(15):6501–17.
 45. Hua Y, Sahashi K, Hung G, et al. Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes Dev*. 2010;24(15):1634–44.
 46. Sullivan JM, Mazur C, Wolf DA, et al. Convective forces increase rostral delivery of intrathecal radiotracers and antisense oligonucleotides in the cynomolgus monkey nervous system. *J Transl Med*. 2020;18(1):309.
 47. Ramos D, d'Ydewalle C, Gabbeta V, et al. Age-dependent SMN expression in disease-relevant tissue and implications for SMA treatment. *J Clin Invest*. 2019;129(11):4817–31.
 48. Haataja L, Mercuri E, Regev R, et al. Optimality score for the neurologic examination of the infant at 12 and 18 months of age. *J Pediatr*. 1999;135:153–61.
 49. Glanzman AM, Mazzone E, Main M, et al. The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. *Neuromuscul Disord*. 2010;20(3):155–61.
 50. Maggi L, Bello L, Bonanno S, et al. Adults with spinal muscular atrophy: a large-scale natural history study shows gender effect on disease. *J Neurol Neurosurg Psychiatry*. 2022;93(12):1253–61.
 51. Ar Rochmah M, Shima A, Harahap NIF, et al. Gender effects on the clinical phenotype in Japanese patients with spinal muscular atrophy. *Kobe J Med Sci*. 2017;63(2):E41–4.

52. Finkel RS, Dabbous O, Arjunji R, et al. P.060 time to treatment effect in spinal muscular atrophy type 1 (SMA1): an indirect comparison of treatments. *Can J Neurol Sci.* 2019;46:530.
53. Jiang T, Youn B, Paradis AD, Beckerman R, Barnieh L, Johnson NB. A critical appraisal of matching-adjusted indirect comparisons in spinal muscular atrophy. *Adv Ther.* 2023;40(7):2985–3005.
54. Liao S, Bohn J, de Moor C, Xaplanteris X, Farwell W, Padaris A. PR109 a cautionary tale for indirect treatment comparisons: an example from infantile-onset spinal muscular atrophy. *Value Health.* 2020;23:S709.
55. National Institute for Health and Care Excellence. Risdiplam for treating spinal muscular atrophy. Technology appraisal guidance [TA755]. <https://www.nice.org.uk/guidance/ta755>. Accessed Feb 2024.
56. Clemens PR, Rao VK, Connolly AM, et al. Safety, tolerability, and efficacy of viltolarsen in boys with duchenne muscular dystrophy amenable to exon 53 skipping: a phase 2 randomized clinical trial. *JAMA Neurol.* 2020;77(8):982–91.
57. US Food and Drug Administration. Clinical review: viltolarsen 2017. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2020/212154Orig1s000MedR.pdf. Accessed Feb 2024.
58. Mendell JR, Goemans N, Lowes LP, et al. Longitudinal effect of eteplirsen versus historical control on ambulation in Duchenne muscular dystrophy. *Ann Neurol.* 2016;79(2):257–71.
59. US Food and Drug Administration. Clinical review: eteplirsen 2016. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/206488_summary%20review_redacted.pdf. Accessed Feb 2024.
60. Swallow E, Pham T, Patterson-Lomba O, et al. Comparative efficacy and safety of ozanimod and ponesimod for relapsing multiple sclerosis: a matching-adjusted indirect comparison. *Mult Scler Relat Disord.* 2023;71:104551.
61. US Food and Drug Administration. September 7, 2022: Meeting of the peripheral and central nervous system drugs advisory committee meeting announcement 2022. <https://www.fda.gov/advisory-committees/advisory-committee-calendar/september-7-2022-meeting-peripheral-and-central-nervous-system-drugs-advisory-committee-meeting>. Accessed Feb 2024.
62. European Medicines Agency. Assessment report: Kymriah 2018. https://www.ema.europa.eu/en/documents/assessment-report/kymriah-epar-public-assessment-report_en.pdf. Accessed Feb 2024.
63. Haute Autorité de Santé. EVRYSDI (risdiplam) 2021. https://www.has-sante.fr/jcms/p_3289078/fr/evryydi-0-75-mg/ml-risdiplam#videotoc_1_1. Accessed Feb 2024.
64. Australian Government, Department of Health and Aged Care. RISDIPLAM, Powder for oral solution 0.75 mg per 1mL, 80 mL, Evryydi®, Roche Products Pty Ltd 2021. <https://www.pbs.gov.au/industry/listing/elements/pbac-meetings/psd/2021-03/files/risdiplam-psd-mar-2021.pdf>. Accessed Feb 2024.
65. ClinicalTrials.gov. NCT02193074: A study to assess the efficacy and safety of ISIS-SMNRx in infants with spinal muscular atrophy (ENDEAR). <https://clinicaltrials.gov/ct2/show/NCT02193074>. Accessed Feb 2024.
66. Darras BT, Masson R, Mazurkiewicz-Betdzińska M, et al. Risdiplam-treated infants with type 1 spinal muscular atrophy versus historical controls. *N Engl J Med.* 2021;385(5):427–35.
67. Wang CH, Finkel RS, Bertini ES, et al. Consensus statement for standard of care in spinal muscular atrophy. *J Child Neurol.* 2007;22(8):1027–49.
68. Mercuri E, Finkel RS, Muntoni F, et al. Diagnosis and management of spinal muscular atrophy: part 1: recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscul Disord.* 2018;28(2):103–15.
69. Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: part 2: pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. *Neuromuscul Disord.* 2018;28(3):197–207.