Nusinersen for treating spinal muscular atrophy: A Single Technology Appraisal

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Paul Tappenden, Reader, ScHARR, University of Sheffield, Sheffield, UK Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Eva Kaltenthaler, Professor, ScHARR, University of Sheffield, Sheffield, UK Emma Hock, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Andrew Rawdin, Research Assistant, ScHARR, University of Sheffield, Sheffield, UK Clara Mukuria, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Mark Clowes, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK Anita Simonds, Professor of Respiratory and Sleep Medicine, Royal Brompton & Harefield NHS Foundation Trust, London, UK Anne-Marie Childs, Consultant Paediatric Neurologist, Leeds Teaching Hospitals, Leeds, UK
Correspondence Author	Paul Tappenden, Reader, ScHARR, University of Sheffield, Sheffield, UK
Date completed	30th May 2018

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 16/134/08.

Declared competing interests of the authors

Dr Anne-Marie Childs has received travel expenses and a professional fee for attending an advisory board meeting held by Biogen in September 2017. None of the other authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Matt Stevenson, ScHARR, for providing comments on the draft report and Gill Rooney, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Tappenden P, Hamilton J, Kaltenthaler E, Hock E, Rawdin A, Mukuria C, Clowes M, Simonds A, Childs A. Nusinersen for treating spinal muscular atrophy: A Single Technology Appraisal. School of Health and Related Research (ScHARR), 2018.

Contributions of authors

Eva Kaltenthaler and Emma Hock summarised and critiqued the clinical effectiveness evidence reported within the company's submission. Paul Tappenden and Andrew Rawdin critiqued the health economic analyses submitted by the company and undertook the ERG's exploratory analyses. Jean Hamilton critiqued the statistical analyses presented in the company's submission. Clara Mukuria provided advice on the mapping analysis used to value health states. Mark Clowes critiqued the company's search strategy. Anne-Marie Childs and Anita Simonds provided clinical input to the ERG. All authors were involved in drafting and commenting on the final report.

CONTENTS

Abbreviations 1
1. SUMMARY 2
1.1 Critique of the decision problem in the company's submission 2
1.2 Summary of clinical effectiveness evidence submitted by the company 2
1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted 4
1.4 Summary of cost effectiveness submitted evidence by the company 4
1.5 Summary of the ERG's critique of cost effectiveness evidence submitted 6
1.6 ERG commentary on the robustness of evidence submitted by the company 6
1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG 7
2. BACKGROUND 8
2.1 Critique of company's description of the underlying health problem 8
2.2 Critique of company's overview of current service provision 10
3. CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM 14
3.1 Population 17
3.2 Intervention 17
3.3 Comparators 18
3.4 Outcomes 18
3.5 Economic analysis 19
3.6 Subgroups 19
3.7 Special considerations 19
4. CLINICAL EFFECTIVENESS 20
4.1 Critique of the methods of review(s) 20
4.2 Critique of studies of nusinersen for treating SMA 20
4.3 Conclusions of the clinical effectiveness section 55
4.4 Additional work undertaken by ERG 57
5. COST EFFECTIVENESS 58
5.1 Company's review of published cost-effectiveness studies 58
5.2 Model scope – early onset and later onset models 59
5.3 Early onset model – methods and results 61
5.4 Later onset model – methods and results 91
5.5 Critical appraisal of the company's health economic analyses 111
5.6 Exploratory analyses undertaken by the ERG 136
5.7 Discussion 146
6. END OF LIFE 149
7. OVERALL CONCLUSIONS 152
7.1 Implications for research 153
8. REFERENCES 154
9. APPENDICES 158
Appendix 1: Exclusion criteria for ENDEAR and CHERISH studies 158
Appendix 2: Overall survival and event free survival by disease duration subgroup 159
Appendix 3: Instruments used to inform transition probabilities within the company's models 161
Appendix 4: Methods for implementing the ERG's exploratory analyses 164

LIST OF TABLES

Table 1: Classification and subtypes of SMA (adapted from CS Table 3, based on Farrah *et al*[8]) 9
Table 2: Clinical management recommendations from the consensus statement by the SCC for SMA (reproduced from CS, Table 4) 11
Table 3: Company's statement of the decision problem (reproduced from CS Table 1) 15
Table 4: Nusinersen studies identified in the CS (adapted from CS, Figure 3) 21
Table 5: ENDEAR study characteristics (adapted from CS, Table 5 and Table 7) 22
Table 6: ENDEAR baseline demographics of the ITT population (adapted from CS, Table 11) 23
Table 7: Company and ERG quality assessment for ENDEAR (adapted from CS, Table 18) 25
Table 8: ENDEAR analysis sets (adapted from CS, Table 15) 27
Table 9: ENDEAR motor function outcomes (adapted from CS, Table 19) 27
Table 10: ENDEAR study ventilation and survival outcomes (adapted from CS, Table 19) 29
Table 11: Summary of study characteristics for CS3A (based on data reported in CS Appendix L, Table 20) 30
Table 12: CHERISH study characteristics (adapted from CS, Table 5 and Table 7) 32
Table 13: CHERISH baseline demographics in the ITT population (reproduced from CS, Table 12) 33
Table 14: Company and ERG quality assessment for CHERISH (adapted from CS, Table 18) 35
Table 15: CHERISH efficacy sets (adapted from CS, Table 16) 37
Table 16: CHERISH motor function outcomes (adapted from CS, Table 20) 38
Table 17: Study characteristics for additional late onset studies (adapted from CS Appendix L, Table 20) 39
Table 18: Summary of ongoing nusinersen studies (reproduced from CS, Table 30) 41
Table 19: NURTURE study characteristics (adapted from CS, Table 7 and NURTURE CSR[21]) 42
Table 20: Baseline characteristics for the NURTURE study (adapted from CS, Table 13, including additional data from the company's clarification response, question A9) 43
Table 21: Company and ERG quality assessment for NURTURE (adapted from information in CS, Appendix D, pages 21-26) using the Newcastle-Ottawa Scale[19] 44
Table 22 CHERISH study CGI-I assessment (investigator and caregiver) at month 15 (reproduced from CS, Table 22) 47
Table 23: ENDEAR subgroups analyses according to disease duration at screening (≤12 weeks, >12 weeks) 48
Table 24: ENDEAR subgroups analyses according to age at symptom onset (≤12 weeks, >12 weeks) 48
Table 25: AEs from integrated safety analysis (reproduced from CS, Table 27) 52
Table 26: SAEs and death summary from integrated safety analysis (reproduced from CS, Table 28) 54
Table 27: Scope of company's health economic analyses – early and later onset models 59

Table 28: Early onset model health states according to HINE-2 scoring (adapted from CS, Figure 31 footnotes and the company's clarification response) 62
Table 29: Evidence used to inform the company's early onset model 66
Table 30: Summary of survival models applied for extrapolation of OS 67
Table 31: Model fit statistics for parametric models fitted to ENDEAR OS data (adapted from CS Appendix P, Figure 30, Figure 31 and Figure 33) 70
Table 32: Model fit statistics for parametric models fitted to adjusted Gregoretti *et al* NRA OS data (adapted from CS Appendix P, Figure 63, 64, 66) 71
Table 33: Model fit statistics for parametric models fitted to Zerres *et al* 1997 Type II OS data (adapted from CS Appendix P, Figure 77, Figure 78 and Figure 80) 71
Table 34: Model fit statistics for parametric models fitted to reconstructed general population mortality data (adapted from CS Appendix P, Figures 68, 69 and 71) 72
Table 35: CHOP INTEND data used to inform transition probabilities beyond month 13 73
Table 36: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 1-64 (taken from company's model) 74
Table 37: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 65-183 (taken from company's model) 75
Table 38: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 184-302 (taken from company's model) 76
Table 39: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 303-394 (taken from company's model) 77
Table 40: Transition matrices for nusinersen (top) and sham (bottom), extrapolation based on CHOP INTEND score in ENDEAR trial, months 13-14 (taken from company's model) 78
Table 41: Transition matrices for nusinersen (top) and sham (bottom), extrapolation based on CHOP INTEND score in ENDEAR trial, all 4-month cycles after month 14 (taken from company's model) 79
Table 42: Scoliosis surgery parameters included in the early onset model 80
Table 43: Patient utilities used in the early onset model 81
Table 44: Parent/carer utilities used in the early onset model 82
Table 45: Estimated nusinersen administration costs 83
Table 46: Estimated annual costs by category of resource use in Type I, II and III SMA patients (reproduced from CS Table 41) 83
Table 47: Allocation of costs by resource classification 84
Table 48: Annual health state costs, early onset model 85
Table 49: Distributions used in company's PSA, early onset model 86
Table 50: Company's model results, early onset model (including patient health gains only) 87
Table 51: Company's model results, early onset model (including patient health gains and caregiver QALY losses) 87

Table 52: Scenario analysis results, early onset model 90
Table 53: Subgroup analysis results, early onset model 91
Table 54: Model health states according to HFMSE score (adapted from CS, Figure 43 footnotes) ... 92
Table 55: Evidence used to inform the company's later onset model 95
Table 56: Summary of survival models applied for extrapolation of overall survival 96
Table 57: HFMSE data used to inform transition probabilities after month 15 97
Table 58: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 1-92 98
Table 59: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 93-169 99
Table 60: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 170-274 100
Table 61: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 275-365 101
Table 62: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 366-456 102
Table 63: Transition matrices for nusinersen (top) and sham (bottom), extrapolation based on HFMSE score in CHERISH trial, all 4-month cycles after month 15 103
Table 64: Parent/carer utilities used in the later onset model 104
Table 65: Annual health state costs, later onset model 105
Table 66: Distributions used in company's PSA, later onset model 106
Table 67: Company's model results, later onset model (patient health gains only) 107
Table 68: Company's model results, later onset model (patient health gains and caregiver QALY losses) 107
Table 69: Scenario analysis results, later onset model 110
Table 70: Subgroup analysis results, later onset model 110
Table 71: Adherence of the company's economic analyses to the NICE Reference Case (early onset and later onset models) 112
Table 72: Comparison of the company's model and the ERG's double-programmed Markov traces, end of trial follow-up to end of time horizon (excludes the possibility of scoliosis surgery) 116
Table 73: Comparison of the company's model results and the ERG's estimated ICERs using the company's Markov traces 117
Table 74: EQ-5D utilities (parent proxy) reported by Bastida *et al* 134
Table 75: Elicited utilities from Lloyd *et al* vignette study 134
Table 76: Caregiver utilities reported by Bastida *et al*[37] 135
Table 77: Health utilities from vignette study applied in ERG's exploratory analyses 137
Table 78: Caregiver utilities applied in the ERG's exploratory analyses 138

Table 79: Clinical advisors' estimates of HRQoL associated with model health states 139
Table 80: ERG preferred analysis, early onset 141
Table 81: Additional exploratory analyses undertaken using the ERG preferred model, early onset 142
Table 82: ERG preferred analysis, later onset 144
Table 83: Additional exploratory analyses undertaken using the ERG preferred model, later onset . 145
Table 84: Evidence supporting the application of end of life criteria presented in the CS (adapted from CS, Table 31) 149
Table 85: Exclusion criteria for ENDEAR and CHERISH (adapted from Table 8, CS, page 34) 158
Table 86: HINE-2 classification (reproduced from CS Figure 5) 161
Table 87: CHOP INTEND domains and scoring 162
Table 88: HFMSE domains and scoring 162
Table 89: WHO motor milestones 163
Table 90: ERG analysis 1 - baseline distribution for early onset model 164
Table 91: ERG exploratory analysis 4 - caregiver utilities for early onset model (Bastida) 164
Table 92: ERG exploratory analysis 6a – patient utilities for early onset model (Bastida) 165
Table 93: ERG exploratory analysis 6b – patient utilities for early onset model (ERG's clinical advisors) 165
Table 94: Transition matrix for ERG exploratory analysis 8a - 5% of nusinersen-treated patients deteriorate to next worst state 166
Table 95: Transition matrix for ERG exploratory analysis 8b - 10% of nusinersen-treated patients deteriorate to next worst state 166
Table 96: Transition matrix for ERG exploratory analysis 8c - 20% of nusinersen-treated patients deteriorate to next worst state 166
Table 97: Transition matrix for ERG exploratory analysis 8d – all patients remain in the state achieved at the end of ENDEAR follow-up 167
Table 98: Transition matrix for ERG exploratory analysis 8e – all patients revert to no milestones state at the end of ENDEAR follow-up 167
Table 99: ERG analysis 1 - baseline distribution for later onset model 168
Table 100: ERG exploratory analysis 3 – patient utilities for later onset model (vignette) 168
Table 101: ERG exploratory analysis 4 - caregiver utilities for later onset model (Bastida) 168
Table 102: ERG exploratory analysis 6a – patient utilities for later onset model (Bastida) 169
Table 103: ERG exploratory analysis 6b – patient utilities for later onset model (ERG's clinical advisors) 169
Table 104: Transition matrix for ERG exploratory analysis 8a - 5% of nusinersen-treated patients deteriorate to next worst state 170
Table 105: Transition matrix for ERG exploratory analysis 8b - 10% of nusinersen-treated patients deteriorate to next worst state 170

Table 106: Transition matrix for ERG exploratory analysis 8c - 20% of nusinersen-treated patients deteriorate to next worst state 171
Table 107: Transition matrix for ERG exploratory analysis 8d – all patients remain in the state achieved at the end of CHERISH follow-up 171
Table 108: Transition matrix for ERG exploratory analysis 8e – all patients revert to no milestones state at the end of CHERISH follow-up 172

LIST OF FIGURES

Figure 1: Clinical care pathway with nusinersen (reproduced from CS, Figure 2) 13
Figure 2: ENDEAR Kaplan-Meier curves for EFS (A) and OS (B) (ITT population, final analysis) (reproduced from CS, Figure 15) 29
Figure 3: ENDEAR: Kaplan-Meier plots of time to death or permanent ventilation in the subgroup of infants below the median disease duration at screening (reproduced from CS, Figure 16) 49
Figure 4: ENDEAR: Kaplan-Meier plots of time to death or permanent ventilation in the subgroup of infants above the median disease duration at screening (reproduced from CS, Figure 17) 50
Figure 5: Change from baseline in total HFMSE score according to age (A) and disease duration (B) at screening (final analysis) (reproduced from CS, Figure 23) 51
Figure 6: Company's early onset model structure (reproduced from CS, Figure 31) 61
Figure 7: Kaplan-Meier estimates based on Gregoretti *et al*, adjusted for mean age of patients at the start of the ENDEAR trial (reproduced from CS Appendix P, Figure 55) 69
Figure 8: CEACs, early onset model, patient health gains only 88
Figure 9: Company's DSA tornado diagram, early onset model, patient health gains only 89
Figure 10: Company's later onset model structure (reproduced from CS, Figure 43) 91
Figure 11: CEACs, later onset model, patient health gains only 108
Figure 12: Company's DSA tornado diagram, later onset model, patient health gains only 109
Figure 13: Initial HINE-2 health state distribution of patients in the company's early onset model (based on ENDEAR) 118
Figure 14: Initial HFMSE health state distribution of patients in the company's later onset model (based on CHERISH) 119
Figure 15: Observed and assumed transitions between HINE-2 health states over time, early onset model 122
Figure 16: Observed and assumed transitions between HFMSE health states over time, later onset model 122
Figure 17: Health state occupancy over time, early onset model, nusinersen group 124
Figure 18: Health state occupancy over time, later onset model, nusinersen group 125
Figure 19: Fitted survival curves, early onset model 128

Figure 20: Fitted survival curves, later onset model 129
Figure 21: Kaplan-Meier OS estimates from adjusted Gregoretti *et al* and ENDEAR 130
Figure 22: Kaplan-Meier curve based on reconstructed IPD from Zerres *et al* (reproduced from CS, Figure 44) 131
Figure 23: Overall survival by disease duration (reproduced from CS, Appendix E, Figure 8) 159
Figure 24: Event free survival by disease duration subgroup (reproduced from CS, Appendix E, Figure 9) 160

Abbreviations

ACEND	Assessment of Caregiver Experience with Neuromuscular Disease
ADL	Activities of daily living
AE	Adverse event
AFT	Accelerated failure time
AIC	Akaike Information Criterion
ASO	Antisense oligonucleotide
BIC	Bayesian Information Criterion
BiPAP	Bi-level positive airway pressure
CADTH	Canadian Agency for Drugs and Technologies in Health
CEAC	Cost-effectiveness acceptability curve
CGI-I	Clinical Global Impression of Improvement
CHOP INTEND	Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
CI	Confidence interval
CMAP	Compound Muscle Action Potential
CPAP	Continuous positive airway pressure
CRD	Centre for Reviews and Dissemination
CS	Company's submission
CSR	Clinical study report
DSA	Deterministic sensitivity analysis
EFS	Event-free survival
EMA	European Medicines Agency
EQ-5D	Euroqol 5-Dimensions
EQ-5D-5L	Euroqol 5-Dimensions 5-Level
EQ-5D-Y	Euroqol 5-Dimensions Youth
ERG	Evidence Review Group
FDA	Food and Drug Administration
GP	General practitioner
HFMSE	Hammersmith Functional Motor Scale-Expanded
HINE-2	Hammersmith Infant Neurological Examination (Module 2)
HR	Hazard ratio
HRQoL	Health-related quality of life
IBS	Integrated Brier Score
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient-level data
ITT	Intention-to-treat
KM	Kaplan-Meier
LSM	Least squares mean
mg	Milligram
MI-E	Mechanical insufflation-exsufflation
mRNA	Messenger ribonucleic acid
MUNE	Motor Unit Number Estimation
mV	Megavolt
N/A	Not applicable
NG	Nasogastric
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIV	Noninvasive ventilation
NJ	Nasojejunal
NR	Not reported
NRA	Non-invasive respiratory aid
OLS	Ordinary least squares
ONS	Office for National Statistics

OS	Overall survival
OT	Occupational therapy
PedsQL	Paediatric Quality of Life Inventory
PedsQL NMM	Paediatric Quality of Life Inventory Neuromuscular Module
PH	Proportional hazards
pre-mRNA	Pre-messenger ribonucleic acid
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PT	Physiotherapy
QALY	Quality-adjusted life year
RCT	Randomised controlled trial
RULM	Revised Upper Limb Module
RWC	Real world care
SAE	Serious adverse event
SCC	International Standard of Care Committee
SMA	Spinal muscular atrophy
SMN	Survival motor neuron
SmPC	Summary of Product Characteristics
STA	Single Technology Appraisal
UK	United Kingdom
WHO	World Health Organization
WTP	Willingness-to-pay

1. SUMMARY

1.1 Critique of the decision problem in the company's submission

The company's submission (CS) assesses the clinical effectiveness and cost-effectiveness of nusinersen (Spinraza®) within its licensed indication for the treatment of 5q spinal muscular atrophy (SMA). The CS notes that nusinersen is the first and only approved disease-modifying treatment for SMA. The company's description of SMA and its management is generally appropriate. The decision problem addressed by the CS is partly in line with the final scope issued by the National Institute for Health and Care Excellence (NICE). The evidence presented within the CS relates to a narrower population than that defined in both the NICE scope and the marketing authorisation for nusinersen; specifically, the available evidence is limited to patients with pre-symptomatic and symptomatic early (infantile) onset and later onset SMA. No evidence is presented on the clinical effectiveness or cost-effectiveness of nusinersen in people with Type 0 or Type IV SMA. Despite the limited scope of the available evidence, the CS states that the anticipated place of nusinersen in therapy is as a first-line treatment for all SMA patients as soon as possible after diagnosis (in combination with usual symptomatic care).

The final NICE scope defines the comparator as best supportive care (BSC). The comparator within the randomised controlled trials (RCTs) of nusinersen is a sham procedure. The comparator considered within the company's health economic analysis is "real world care" (usual care), including respiratory, gastrointestinal, nutritional and orthopaedic care. The CS highlights that the differential use of life-extending symptomatic care, including permanent respiratory support, means that real world survival may not reflect that seen in clinical trials. The CS argues that nusinersen meets NICE's end-of-life criteria in the early onset (Type I) SMA population, but not the later onset (Types II and III) SMA population. The Evidence Review Group (ERG) notes that the company's model suggests that the mean predicted survival for patients with early onset SMA receiving usual care is 3.87 years.

1.2 Summary of clinical effectiveness evidence submitted by the company

The CS did not contain a systematic review of clinical effectiveness evidence; this is a requirement of the NICE Single Technology Appraisal (STA) process. Two key studies were presented in the CS: (i) the ENDEAR study, which recruited infantile onset SMA patients, and (ii) the CHERISH study, which recruited later onset SMA patients. Both studies were RCTs comparing nusinersen against a sham procedure control group. ENDEAR (n=122) was undertaken in 31 secondary care centres worldwide. CHERISH (n=126) was undertaken in 24 secondary care centres worldwide.

In the ENDEAR study, 80 participants received nusinersen, administered as a single intrathecal lumbar puncture injection with a scaled 12mg loading dose on study days 1, 15, 29 and 64 and maintenance dosing every 4 months (days 183 and 302), while 41 patients received the sham procedure. Overall, the

baseline characteristics of the two groups were similar, although patients in the nusinersen group were on average younger than those in the control group and had an earlier age of symptom onset. Primary outcomes were: proportion of motor milestone responders (measured using Module 2 of the Hammersmith Infant Neurological Examination [HINE-2]) and event-free survival (EFS, defined as time to death or permanent ventilation). ENDEAR included three analysis sets: (i) an interim analysis set; (ii) a final efficacy set and (iii) a final intention-to-treat (ITT) set. With regard to HINE-2, a significantly greater percentage of patients in the nusinersen group achieved motor milestone responses than the control group (41% vs 0% in the interim analysis and 51% vs 0% in the final efficacy set), although many patients in the nusinersen group could not be classified as responders (49% of patients in the final efficacy set). There was a statistically significant increase in EFS for the nusinersen group compared with the sham control group (ITT analysis set, p=0.005). ENDEAR was rated as being at low risk of bias in the CS; the ERG consider this study to be at moderate risk of bias due to concerns regarding the preservation of blinding, an imbalance in dropouts between groups, and the potential for incomplete reporting of outcomes.

The CHERISH study included 84 patients who received nusinersen administered as single intrathecal lumbar puncture injection, at single dose level of 12mg delivered in 4 doses over 9 months using a loading regimen (days 1, 29, 85) with a maintenance dose at 6 months (day 274). The control group was comprised of 42 patients who received the sham control. Overall, the two groups were similar, although there were imbalances between groups with respect to the proportions of patients who had ever achieved a motor milestone and in the median time from disease onset to study enrolment, with a longer delay in receiving therapy in the nusinersen group compared with the sham group. The nusinersen group had a slightly higher Hammersmith Functional Motor Scale-Expanded (HFMSE) total score at baseline. The CHERISH study included three analysis sets: (i) an interim analysis set; (ii) an efficacy set and (iii) an ITT set. The primary outcome measure in CHERISH was motor function as measured by the HFMSE instrument. The change in HFMSE from baseline was significant in both the interim analysis (least squares mean [LSM] change difference: 5.9; 95% confidence interval [CI] 3.7 to 8.1; p<0.001) and the final efficacy set analysis (LSM change difference: 4.9; 95% CI 3.1 to 6.7; p=0.0000001) for the nusinersen group compared with the control group. CHERISH was rated as being at low risk of bias in the CS; the ERG consider this study to be at moderate risk of bias due to concerns regarding the preservation of blinding and the potential for incomplete reporting of outcomes.

In the ENDEAR study, treatment effects for key outcome measures were evaluated for two pre-specified subgroups: disease duration at screening (≤12 weeks, >12 weeks) and age at symptom onset (≤12 weeks, >12 weeks). Overall, nusinersen demonstrated a benefit in all subgroups, except for the analysis of overall survival (OS) in the subgroup with age at onset of symptoms >12 weeks; however, the number of patients in this subgroup was small. For all outcomes, more pronounced treatment effects were

observed for infants with a disease duration ≤12 weeks at screening; however, statistical tests for a difference between subgroups were not provided.

An integrated safety analysis with data from eight completed or ongoing studies including a total of 260 patients was presented in the CS. In the integrated safety analysis, both nusinersen-treated patients and control group patients experienced adverse events (AEs). The most commonly reported AEs were those expected in patients with SMA or after lumbar puncture, such as headache, vomiting, back pain and post-lumbar puncture syndrome. Overall, there were fewer deaths in the nusinersen-treated patients than the control patients (19% vs 7%) and fewer serious adverse events (SAEs) in the nusinersen-treated patients compared with the control patients (39% vs 60%).

1.3 Summary of the ERG's critique of clinical effectiveness evidence submitted

Although no systematic review was presented in the CS, the ERG is confident that no relevant studies of nusinersen for SMA were missed. However, a systematic review of studies related to the BSC comparator was not presented. The quality assessment tools used to appraise the included studies was considered appropriate by the ERG. Most outcomes listed in the NICE scope were presented, with the exception of complications of SMA and stamina and fatigue.

1.4 Summary of cost effectiveness submitted evidence by the company

The company submitted two *de novo* model-based health economic evaluations of nusinersen: the first model relates to patients with early onset (Type I) SMA, whilst the second relates to patients with later onset (Type II/III) SMA.

Early onset model

The company's early onset model assesses the cost-effectiveness of nusinersen versus usual care for the treatment of patients with early onset SMA (initial age = 5.58 months), based on the ENDEAR trial. The incremental health gains, costs and cost-effectiveness of nusinersen are evaluated over a 60-year time horizon from the perspective of the NHS and Personal Social Services (PSS). The company's early onset model adopts a state transition approach, with health states defined by motor function milestones based on the HINE-2 instrument. The model parameters were largely informed by: HINE-2 and Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND) outcomes collected within ENDEAR; mortality outcomes from ENDEAR and other observational data (Gregoretti *et al*, Zerres *et al* and general population life tables); a mapping exercise to translate Paediatric Quality of Life Inventory (PedsQL) outcomes collected in the CHERISH trial to the Euroqol 5-Dimensions (EQ-5D); a cross-sectional study of the costs and caregiver health-related quality of life (HRQoL) impacts of SMA and standard costing sources. The model assumes that treatment using nusinersen will be discontinued for patients who do not achieve any milestones after 13 months, and

for patients undergoing scoliosis surgery who cannot subsequently receive nusinersen administration via lumbar puncture. The company's early onset model employs two key assumptions: (i) after month 13, nusinersen-treated patients who reach health states consistent with Type II/III SMA milestones gain an additional survival advantage, and (ii) after month 13, the motor function of nusinersen-treated patients cannot deteriorate, whilst the motor function of patients receiving usual care cannot improve.

Based on a re-run of the probabilistic version of the company's early onset model by the ERG, nusinersen is expected to generate an additional 5.29 quality-adjusted life years (QALYs) at an additional cost of £2,160,048 per patient; the corresponding incremental cost-effectiveness ratio (ICER) for nusinersen versus usual care is £408,712 per QALY gained. The inclusion of caregiver QALY losses leads to a slightly lower probabilistic ICER of £404,270 per QALY gained. The probability that nusinersen produces more net benefit than usual care at willingness-to-pay (WTP) thresholds below £337,000 per QALY gained is approximately zero. The company's subgroup analyses suggest that the cost-effectiveness profile for nusinersen may be improved in early onset SMA patients with shorter disease duration (≤12 weeks subgroup ICER≈£375,000 per QALY gained, ICER includes patient health gains only).

Later onset model

The company's later onset model assesses the cost-effectiveness of nusinersen versus usual care for the treatment of patients with later onset SMA (initial age = 43.71 months), based on the CHERISH trial. The incremental health gains, costs and cost-effectiveness of nusinersen are evaluated over an 80-year time horizon from the perspective of the NHS and PSS. The company's later onset model adopts a state transition approach, with health states defined by motor function milestones based on the HFMSE instrument and WHO criteria. The model parameters were largely informed by: HFMSE outcomes collected within CHERISH; mortality outcomes from CHERISH and other observational data (Zerres *et al* and general population life tables); and the same cost and HRQoL sources as those used in the early onset model (see above). The company's model assumes that treatment using nusinersen will be discontinued for patients who do not achieve milestones beyond the Sits without support but does not roll state after 15 months, and for patients undergoing scoliosis surgery who cannot subsequently receive nusinersen administration via lumbar puncture. The later onset model includes two key assumptions: (i) after month 15, patients in either treatment group who reach health states consistent with Type III SMA milestones gain an additional survival advantage, and (ii) after month 15, the motor function of nusinersen-treated patients cannot deteriorate, whilst the motor function of patients receiving usual care cannot improve.

Based on a re-run of the probabilistic version of the company's later onset model by the ERG, nusinersen is expected to generate an additional 2.28 QALYs at an additional cost of £2,938,441 per

patient: the corresponding ICER for nusinersen versus usual care is £1,286,149 per QALY gained. The inclusion of caregiver QALY losses leads to a markedly lower probabilistic ICER of £933,088 per QALY gained. The probability that nusinersen produces more net benefit than usual care is approximately zero even at WTP thresholds of £500,000 per QALY gained. The company's subgroup analyses are inconclusive with respect to whether the cost-effectiveness profile for nusinersen is improved for later onset SMA patients with shorter disease duration (<25 months).

1.5 Summary of the ERG's critique of cost effectiveness evidence submitted

The ERG critically appraised the company's economic analyses of early and later onset SMA and double-programmed: (a) simplified versions of the Markov traces from the company's models and (b) the remainder of the model structures based on the company's Markov traces. The ERG's critical appraisal identified a number of issues relating to the company's economic analyses and the evidence used to inform them. The most pertinent of these include: (i) the absence of economic evidence relating to Type 0 and Type IV SMA; (ii) the unnecessary complexity of the company's implemented models; (iii) highly favourable assumptions regarding the expected trajectory of nusinersen-treated patients through modelled motor milestone health states; (iv) highly favourable assumptions regarding the expected survival of nusinersen-treated patients; (v) poor face validity of patient utilities used in the models, and (vi) arbitrary calculations underpinning the caregiver disutilities used in the models.

1.6 ERG commentary on the robustness of evidence submitted by the company

1.6.1 Strengths

The two key RCTs of nusinersen were included in the CS; these studies included early onset and later onset SMA patients. The included studies were considered to be of moderate quality and included most outcomes of relevance for this appraisal.

The clinical advisors to the ERG considered that the structures of the company's health economic models were broadly appropriate and reflected some of the key outcomes associated with SMA.

Despite the unnecessary complexity of the company's models, the ERG's model verification exercise did not identify any significant programming errors.

1.6.2 Weaknesses and areas of uncertainty

The limitations of the clinical evidence review mainly concern the absence of a systematic review and the absence of a systematic review of studies relating to BSC, the comparator of interest in the NICE decision problem.

The long-term probabilities of achieving, maintaining and losing motor function for nusinersen-treated patients, the long-term survival advantage of nusinersen and the relationship between motor function milestones and HRQoL are all highly uncertain. The ERG notes that the use of less optimistic assumptions regarding the extrapolation of motor function and survival outcomes has the propensity to markedly increase the ICERs for nusinersen. However, the ERG also notes that given the acquisition cost of nusinersen, the level of decision uncertainty with respect to NICE's usual thresholds for cost-effectiveness is low.

1.7 Summary of exploratory and sensitivity analyses undertaken by the ERG

The ERG undertook eight sets of exploratory analyses using the deterministic version of the company's early onset and later onset models. The ERG's preferred analysis includes: (i) the use of a common initial distribution across health states for both treatment groups; (ii) the inclusion of end-of-life costs for the later onset population; (iii) the use of patient utilities from the vignette study (Lloyd *et al*) and (iv) the application of caregiver utilities by SMA type (from Bastida *et al*) to states relating to SMA milestones. Importantly, this analysis does not address the ERG's concerns regarding the lack of plausibility surrounding the company's modelled survival and motor function trajectories; as such, the ERG's "preferred" ICERs are very likely to be underestimated in both SMA populations. In order to address this uncertainty, additional sensitivity analyses were undertaken to explore the use of alternative patient utilities, the exclusion of mortality adjustments for better health states and the use of alternative long-term (post-trial) transition probabilities.

Early onset model

The ERG's preferred ICER for nusinersen versus usual care in the early onset population is estimated to be £421,303 per QALY gained (including patient health gains only). The inclusion of caregiver QALY losses increases the ICER to £631,583 per QALY gained. The ERG's additional exploratory analyses lead to ICERs ranging from £366,289 per QALY gained to dominated (the ERG notes that the upper limit of the ICER range reflects a particularly pessimistic scenario).

Later onset model

The ERG's preferred ICER for nusinersen versus usual care in the later onset population is estimated to be £408,769 per QALY gained (including patient health gains only). The inclusion of caregiver QALY losses increases the ICER to £632,850 per QALY gained. The ERG's additional exploratory analyses lead to ICERs ranging from £432,191 per QALY gained to in excess of £18.4million per QALY gained (again, the upper limit of the ICER range reflects a particularly pessimistic scenario).

2. BACKGROUND

This report provides a review of the evidence submitted by Biogen in support of nusinersen for the treatment of spinal muscular atrophy (SMA). It considers both the company's submission[1] (CS) received on 20th March 2018 and the subsequent responses to clarification questions supplied by the company.[2, 3]

2.1 Critique of company's description of the underlying health problem

The CS[1] (pages 15-17) provides a reasonable description of the underlying health problem; this is summarised briefly below.

SMA is a progressive neuromuscular disease which results from mutations in chromosome 5q in the *SMN1* gene. The disease causes muscle weakness and progressive loss of movement and physical disability. As well as affecting patients' musculoskeletal system, SMA also impacts upon their respiratory and gastrointestinal systems.[1] SMA is rare and is recognised as an orphan disease by the European Medicines Agency (EMA).[4] SMA is recognised as the most common genetic cause of death in infants.[5]

SMA affects the motor neurons (the nerves from the brain and spinal cord that control muscle movements). Patients with SMA lack a protein called "survival motor neuron" (*SMN*) which is made by the *SMN1* and *SMN2* genes; this protein is essential for the normal functioning and survival of motor neurons. In the absence of this protein, the motor neurons deteriorate and eventually die, leading to muscle disuse, atrophy and weakness.[6]

SMA presents across a spectrum of subtypes (Types 0-IV) which are related to the age of onset (see Table 1). Younger age of onset is associated with greater severity of disease and poorer prognosis. The CS[1] defines Type I as early (infantile) onset SMA and Type II and III as later onset SMA, based on the age of onset and the level of motor function achieved. With the exception of Type 0 SMA, the disease usually involves a pre-symptomatic period followed by rapidly progressive functional loss and a later relatively static phase with slow progression.[7] Diagnosis of Type I SMA and more severe Type II SMA usually occurs during the first year of life. Most patients with Type II SMA are diagnosed in their second year of life, whilst Type III SMA is typically diagnosed at age 2-3 years, but may be later.

Table 1: Classification and subtypes of SMA (adapted from CS Table 3, based on Farrah *et al*[8])

SMA type	Age of onset	Maximal motor milestone	Motor ability and additional features	Prognosis‡
SMA Type 0	Before birth	None	Severe hypotonia; unable to sit and roll*	Respiratory insufficiency at birth: death within weeks
SMA Type I	2 weeks (Ia) 3 months (Ib) 6 months (Ic)	None	Severe hypotonia; unable to sit and roll†	Death/ventilation by 2 years
SMA Type II	6–18 months	Sitting	Proximal weakness: unable to walk independently	Survival into adulthood (typically >25 years)
SMA Type III	<3 years (IIIa) >3 years (IIIb) >12 years (IIIc)	Walking	May lose ability to walk	Normal life span
SMA Type IV	>30 years or 10–30 years	Normal	Mild motor Impairment	Normal life span

SMA - spinal muscular atrophy
** Need for respiratory support at birth; contractures at birth, reduced foetal movements*
† Ia joint contractures present at birth; Ic may achieve head control
‡ Prognosis varies with phenotype and supportive care interventions

Type I SMA (early onset)

Type I SMA has been reported to be the most common and severe form of the disease (accounting for approximately 45% of all cases of SMA), with an estimated incidence of 5.83 per 100,000 live births.[1, 9] Type I SMA is associated with a particularly poor prognosis and early mortality; most patients do not survive to their second birthday unless they receive ventilatory support.[8] Symptoms appear early (before 6 months) and include severe hypotonia (decreased muscle tone), inability to lift head/poor head control, and poor feeding.[1, 7] By definition, patients with Type I SMA never develop the ability to sit independently.[7] Patients suffer from a range of severe problems including pulmonary, nutritional and gastrointestinal complications. Despite these symptoms, cognitive ability is normal.

Type II/III SMA (later onset)

Type II and Type III SMA (accounting for around 50% of all cases of SMA) are less severe forms of the disease compared with Type I SMA. The incidence of Type II and Type III SMA is reported to be 2.66 and 1.20 per 100,000 live births, respectively.[1, 9] The age of onset is usually between 6 and 18 months for Type II SMA, and between 18 months and adulthood for Type III SMA.[7] Both Type II and Type III SMA are associated with a loss of motor function over time and numerous secondary complications. The severity of motor function impairment is highly variable between patients, with some patients with Type III SMA developing the ability to walk without assistance and others with Type II SMA being unable to sit without support.[1] Scoliosis is universally present in patients with Type II disease. Patients have an increased risk of respiratory disease and muscle weaknesses in the upper

chest make breathing and coughing more difficult, thereby leading to ineffective secretion clearance and an increased risk of chest infections.[1] Survival of patients with Type II SMA is typically greater than 25 years, and many patients live considerably longer as a consequence of more aggressive supportive care.[7] Survival of patients with Type III SMA is believed to be normal. As with more severe types of SMA, cognitive ability in these patients is normal.

The CS highlights the impact of the disease on patients' health-related quality of life (HRQoL), particularly with respect to physical disability, the inability to live independently, the high incidence of chronic pain, and the psychological burden associated with the progressive decline in health, including fear of losing independence, difficulties feeding and impaired breathing.[1] The CS also highlights the considerable economic and emotional burden affecting parents/caregivers as a consequence of giving up work to provide care, attending frequent hospital appointments and undertaking other SMA-related tasks.[1] Additional information relating to the impact of SMA on patients and caregivers is available within the submissions to NICE from clinical and patient groups.

2.2 Critique of company's overview of current service provision

The CS presents a useful overview of the current management of SMA. This is briefly described below.

There is no standard of care pathway for SMA and no guidance has been published by the National Institute for Health and Care Excellence (NICE). The CS notes that, excluding nusinersen, there is currently no effective disease-modifying therapy for SMA. Treatment requires a multidisciplinary approach and is focussed principally on respiratory and nutritional support, but also includes neuromuscular and orthopaedic care.

The CS refers to an SMA consensus statement released by the International Standard of Care Committee (SCC), which reports recommendations on the management of SMA according to physical functioning (non-sitters, sitters and walkers) rather than SMA type (Types 0 to IV).[10, 11] Non-sitters include patients who currently are not able to sit independently (i.e. the infantile Type I SMA patients). Sitters include those patients who can sit independently but cannot walk independently. Walkers can walk independently.[1] The guidelines from the SCC (summarised by the company) are reproduced in Table 2. Clinical advisors to the Evidence Review Group (ERG) noted that there has been a shift towards proactive/anticipatory respiratory care which is unlikely to be reflected within historical SMA natural history studies.

The CS highlights that for early onset patients (non-sitters), survival is very poor. [redacted].The ERG notes that the number of

patients, the disease subtype and the extent of ventilatory support provided is not clear within this survey sample. Whilst gastrostomy and ventilation can extend patient survival for early onset patients, these interventions do not impact upon motor function decline and their use in clinical practice is variable. With respect to later onset patients (sitters and walkers), symptoms may be highly variable between patients and the requirement for intensive nutritional and respiratory support may be less than for patients with early onset SMA. Later onset patients who are classed as sitters are more likely to develop scoliosis and subsequently require surgery, bracing and physical therapy.

Table 2: Clinical management recommendations from the consensus statement by the SCC for SMA (reproduced from CS, Table 4)

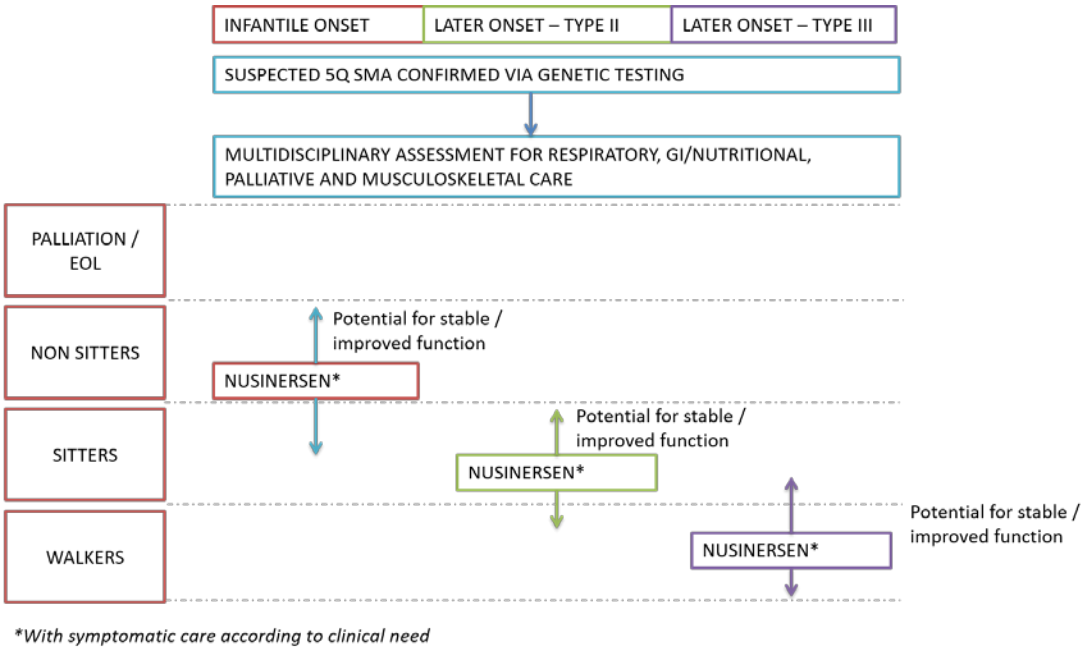
<table>
<tr><th>Type of care</th><th>NON-SITTERS</th><th>SITTERS</th><th>WALKERS</th></tr>
<tr><td colspan="4">Pulmonary care</td></tr>
<tr><td>Anticipatory respiratory care</td><td colspan="3">• Understanding the child’s baseline, deviations from his/her baseline, hypoventilation and intervention
• Acute illness management including rapid access to specialty medical care providers
• Nutrition and hydration
• A low threshold to start antibiotics
• Routine immunisations</td></tr>
<tr><td rowspan="3">Chronic respiratory management</td><td colspan="3">Airway clearance:
• Assisted cough (MI-E or manual)
• Secretion mobilisation techniques (chest physiotherapy, postural drainage)
• Oximetry to guide therapy
Respiratory support:
• NIV
CPAP (goal to transition to BiPAP)</td></tr>
<tr><td>• Option: Care without ventilation support
• Palliative care
• Tracheotomy</td><td>Airway clearance/ respiratory support, as needed</td><td rowspan="2">Airway clearance/ respiratory support not likely to be required until late into the disease course</td></tr>
<tr><td colspan="2">NIV with high span BiPAP, even for short daytime periods</td></tr>
<tr><td rowspan="2">Acute care management</td><td colspan="3">Airway clearance:
• Assisted cough (MI-E or manual), oral or airway suctioning
• Oximetry
• Chest physiotherapy
• Postural drainage
Respiratory support:
• Acute use of NIV
• Oxygen therapy</td></tr>
<tr><td colspan="2">Respiratory support:
• Daytime NIV with airway clearance
• Intubation and mechanical ventilation
• Palliative care</td><td>Respiratory support:
• NIV for home use</td></tr>
<tr><td colspan="4">Gastrointestinal and nutritional care</td></tr>
<tr><td>Feeding and swallowing difficulties</td><td colspan="3">• Changing food consistency
• Positioning and seating alterations and orthotic devices
• Nutritional supplementation through NG or NJ feeding
• Gastrostomy tube feeding</td></tr>
</table>

Type of care	NON-SITTERS	SITTERS	WALKERS
Gastrointestinal dysfunction	Management of gastroesophageal reflux: • Short term use of acid neutralisers and/or inhibitors of acid secretion • Prokinetic agents • Probiotics • Laparoscopic anti-reflux Nissen fundoplication		
Growth and under or over nutrition problems	• Monitoring of growth velocity (growth charts) • Dietician assessment of nutritional intake • Appropriate intake of calcium and vitamin D • Monitor pre-albumin levels		
Management of nutrition in acutely sick SMA patients	• Avoid prolonged fasting due to high risk of hypoglycaemia • Enteral and/or parenteral feeding to meet caloric needs within 4-6 hours of acute illness admission • Post-operative caloric supplementation		
Neuromuscular and musculoskeletal evaluation			
Managing musculoskeletal system problems and related functional impairments	• Assessments of strength and range of joint motion, relevant motor functional scales and timed tests to monitor those aspect of function that reflect activities of daily living		
Orthopaedic care and rehabilitation			
Managing problems caused by muscle weakness	• Wheelchair mobility • Environmental controls and home modifications		
	• Nutritional support • Posture management with supportive seating • Contracture management by splinting • Pain management • Therapy for ADL and assistive equipment • Limb orthotics	• Contracture management by stretching, bracing, serial casting, orthotics and supports/ slings • Regular exercise and standing with appropriate assistive devices and orthotics • Spine orthotics and surgery	• Contracture management and education • PT and OT • Regular exercise and walking with appropriate assistive devices and orthotics • Spine/limb orthotics and surgery
Orthopaedic surgery	Non-sitters do not benefit from surgery	• Hip subluxation and contractures • Scoliosis surgery	
Other care			
Perioperative care	Due to high risk for post-anaesthesia complications, respiratory status needs to be optimised and orthotic interventions need to be adjusted before surgery. After surgery, close monitoring, aggressive respiratory management, and rapid mobilisation, may be required.		

ADL - activities of daily living; BiPAP - bi-level positive airway pressure; CPAP - continuous positive airway pressure; NIV - non-invasive ventilation; NG - nasogastric; NJ - nasojejunal; MI-E - mechanical insufflation/exsufflation; PT - physiotherapy; OT - occupational therapy; SCC - International Standard of Care Committee; SMA - spinal muscular atrophy

The CS states that nusinersen is the first disease-modifying treatment for SMA. The anticipated place of nusinersen in therapy is as a first-line treatment for all SMA patients as soon as possible after diagnosis, in addition to existing symptomatic care (see Figure 1).[1] Nusinersen is currently available in England for patients with Type 1 SMA (subject to eligibility criteria) through an Expanded Access Programme; under this programme, the acquisition costs of nusinersen are reimbursed by NHS England.

Figure 1: Clinical care pathway with nusinersen (reproduced from CS, Figure 2)



With symptomatic care according to clinical need

3. CRITIQUE OF COMPANY'S DEFINITION OF THE DECISION PROBLEM

This chapter presents a summary and critique of the decision problem addressed by the CS.[1] A summary of the decision problem as outlined in the final NICE scope[12] and addressed in the CS[1] is presented in Table 3.

Table 3: Company's statement of the decision problem (reproduced from CS Table 1)

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Population	People with 5q SMA	Pre-symptomatic and symptomatic people with 5q SMA who have infantile onset (those who have or are most likely to develop type I) or later onset (those who have or are most likely to develop types II and III) SMA	The proposed population is narrower than the marketing authorisation (which includes all patients with 5q SMA) because the evidence base on nusinersen is limited to patients with pre-symptomatic and symptomatic infantile onset and later onset SMA
Intervention	Nusinersen	Nusinersen	N/A
Comparator(s)	Best supportive care	Sham procedure and standard of care treatment	Biogen consider that the most appropriate comparator is sham procedure (administered by lumbar puncture prick), as no disease-modifying therapies (other than nusinersen) are approved or routinely used in SMA
Outcomes	The outcome measures to be considered include: • Motor function (including, where applicable, age appropriate motor milestones) • Respiratory function • Complications of SMA (including, for example, scoliosis and muscle contractures) • Need for non-invasive or invasive ventilation • Stamina and fatigue • Mortality • Adverse effects of treatment • HRQoL	The outcome measures to be considered include: • Motor function (including, where applicable, age appropriate motor milestones) • Event-free survival (time to death or permanent assisted ventilation) and overall survival • Respiratory function • Need for non-invasive or invasive ventilation • Mortality • Adverse effects of treatment • HRQoL	Complications of SMA (including, for example, scoliosis and muscle contractures), and stamina and fatigue, are not included as these outcomes were not collected in the pivotal clinical trials

	Final scope issued by NICE	Decision problem addressed in the company submission	Rationale if different from the final NICE scope
Economic analysis	The reference case stipulates that the cost effectiveness of treatments should be expressed in terms of incremental cost per QALY. The reference case stipulates that the time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared. Costs will be considered from an NHS and personal social services perspective.	The economic analysis considers 2 *de novo* models to assess the cost-effectiveness of nusinersen using motor milestones health states – 1 relating to infantile onset SMA and the other to later onset SMA. The pre-symptomatic health state is being developed but could not be modelled in time for submission.	N/A
Subgroups to be considered	Consideration will be given to subgroups based on severity of disease (including considerations such as age of SMA onset, SMA type and genotype [including *SMN2* copy number]). Guidance will only be issued in accordance with the marketing authorisation. Where the wording of the therapeutic indication does not include specific treatment combinations, guidance will be issued only in the context of the evidence that has underpinned the marketing authorisation granted by the regulator.	The pivotal trials in infantile onset (ENDEAR) and later onset SMA (CHERISH) included pre-specified subgroups based on disease duration and age at symptom onset. For infantile onset SMA patients the economic analysis has evaluated the subgroups based on age at onset of SMA symptoms and disease duration (>12 weeks and ≤12 weeks) from the ENDEAR trial For later onset SMA patients, subgroup analysis has not been conducted in the economic analysis due to the small subgroup sample sizes within	N/A
Special considerations including issues related to equity or equality	NR	N/A	N/A

SMA - spinal muscular atrophy; HRQoL - health-related quality of life; QALY - quality-adjusted life year; SMN2 - survival motor neuron 2; N/A - not applicable; NR - not reported

3.1 Population

The population defined in the NICE scope[12] relates to people with 5q SMA. This is consistent with the marketing authorisation for nusinersen.[4] The evidence presented within the CS[1] relates to a population which is narrower than that defined in both the final NICE scope and the marketing authorisation for nusinersen. The available evidence for nusinersen is limited to patients with pre-symptomatic and symptomatic infantile onset and later onset SMA; no evidence is presented on the clinical effectiveness or cost-effectiveness of nusinersen in people with Type 0 or Type IV SMA.

3.2 Intervention

The intervention under appraisal is nusinersen (Spinraza®). Nusinersen is an antisense oligonucleotide (ASO) which increases the proportion of exon 7 inclusion in survival motor neuron 2 (*SMN2*) messenger ribonucleic acid (mRNA) transcripts by binding to an intronic splice silencing site (ISS-N1) found in intron 7 of the *SMN2* pre-messenger ribonucleic acid (pre-mRNA). By binding, the ASO displaces splicing factors, which normally suppress splicing. Displacement of these factors leads to retention of exon 7 in the *SMN2* mRNA and hence when *SMN2* mRNA is produced, it can be translated into the functional full length *SMN* protein.[4] The CS[1] states that the anticipated place of nusinersen in therapy is as a first-line treatment for all SMA patients as soon as possible after diagnosis (in combination with usual symptomatic care).

Nusinersen is available as a single vial containing 12mg of nusinersen solution. The current list price for a single vial of nusinersen is £75,000.[13]

The Summary of Product Characteristics[4] (SmPC) recommends that nusinersen treatment should be initiated as early as possible after diagnosis of SMA with four loading doses on days 0, 14, 28 and 63. A maintenance dose should be administered once every four months thereafter. This corresponds to an acquisition cost of £450,000 per patient in the first year of treatment, and £225,000 per patient for each subsequent year of treatment. It should be noted that this dosing regimen reflects the treatment schedule adopted within the ENDEAR study[14] (infant onset); however, a different treatment schedule was used in the CHERISH study[15] (later onset). The SmPC notes that there is no evidence relating to the long-term efficacy of nusinersen and that the need for continuation of nusinersen treatment should be reviewed regularly and considered on an individual basis depending on the patient's clinical presentation and response to the therapy.[4]

The SmPC[4] states that nusinersen has not been studied in patients with renal or hepatic impairment and there are no or limited data from the use of nusinersen in pregnant women. The SmPC also highlights a risk of adverse reactions occurring as part of the lumbar puncture procedure, which may be a problem particularly for very young children and those with scoliosis. According to the SmPC,

thrombocytopenia and coagulation abnormalities (including acute severe thrombocytopenia) and renal toxicity have been observed after the administration of other subcutaneously and intravenously administered ASOs.[4] The available data on adverse events (AEs) from the clinical study programme and post-marketing studies of nusinersen are presented in Chapter 3 of this report (see Section 3.2.7).

Contraindications to nusinersen include hypersensitivity to the active substance or to any of the excipients listed in the SmPC.[4]

3.3 Comparators

The final NICE scope[12] defines the comparator as best supportive care (BSC). The comparator within the randomised controlled trials (RCTs) of nusinersen is a sham procedure. The comparator considered within the company's health economic analyses is defined as "real-world care", including respiratory, gastrointestinal, nutritional and orthopaedic care. As noted in the CS,[1] the differential use of life-extending symptomatic care, including permanent respiratory support, means that real world survival may not reflect that seen in clinical trials.

3.4 Outcomes

The final NICE scope[12] lists the following outcomes:

- Motor function (including, where applicable, age appropriate motor milestones)
- Respiratory function
- Complications of SMA (including, for example, scoliosis and muscle contractures)
- Need for non-invasive or invasive ventilation
- Stamina and fatigue
- Mortality
- Adverse effects of treatment
- HRQoL.

The CS[1] includes evidence relating to all of these outcomes except for: (i) stamina and fatigue, and (ii) complications of SMA. These outcomes were excluded from the CS as these endpoints were not included in the pivotal clinical trials (ENDEAR[14] and CHERISH[15]). Clinical advisors to the ERG commented that measuring stamina and fatigue in younger children involves subjectivity and that there are no useful questionnaires available, hence this omission may be reasonable. However, one advisor noted that it is possible to record specific outcomes such as the length of time for which a particular motor skill can be maintained. The advisors also commented that scoliosis is an important marker for disease progression, particularly in older children. However, the advisors also noted that complications of SMA are long-term problems that would be difficult to measure in short-term trials.

3.5 Economic analysis

The CS[1] reports the methods and results of two *de novo* model-based health economic analyses to assess the incremental cost-effectiveness of nusinersen versus usual care for the treatment of patients with early onset (Type I) SMA and later onset (Types II and III) SMA. The company's health economic analyses are detailed and critiqued in Chapter 5.

3.6 Subgroups

The pivotal trials included in the CS (ENDEAR[14] and CHERISH[15]) included pre-specified subgroups based on disease duration and age at symptom onset. Clinical data relating to these subgroups are summarised in Section 4.2.6.

The company's health economic analysis includes subgroup analyses based on duration of disease (≤12 weeks, >12 weeks).[1] CS Table 1 states that subgroup analysis was also undertaken according to age of onset, however no results are presented in the CS for these subgroups. Table 1 of the CS states that subgroup analysis was not conducted for the later onset population due to the small subgroup sample sizes; however, this statement is inaccurate as CS Table 77 reports the results of subgroup analyses based on duration of disease (<25 months, ≥25 months). No subgroup analysis is presented for age of onset within the later onset economic analysis.

3.7 Special considerations

Table 1 of the CS[1] states that there are no equality issues relating to the use of nusinersen for the treatment of SMA. CS Section 1.4 notes that although the available RCT evidence relates specifically to infants and children, older patients may also benefit from nusinersen treatment. Despite the absence of evidence for older patients, the CS argues that it is important that all age groups and patient disabilities are considered regarding access to treatment.

The CS[1] argues that NICE's end-of-life criteria apply to the early onset SMA population, but not the later onset population. The evidence supporting this argument is presented and critiqued in Chapter 6.

4. CLINICAL EFFECTIVENESS

This chapter presents a summary and critique of the clinical evidence contained within the CS[1] for nusinersen for the treatment of SMA. Section 4.1 presents a critique of the methods used to identify and select evidence for inclusion in the CS. Section 4.2 presents a critique of the key studies included in the CS. Section 4.3 presents the conclusions relating to the clinical effectiveness evidence.

4.1 Critique of the methods of review(s)

The CS[1] did not include a systematic review of the clinical effectiveness evidence for nusinersen. No searches were reported, hence it is unclear whether all relevant studies of nusinersen were identified. However the ERG is confident that all relevant studies have been included in the CS. No searches were undertaken for studies of BSC, the comparator listed in the final NICE scope.[12] In response to a request for clarification regarding the absence of systematic review from the CS (see clarification response,[2] question A1), the company stated that a quarterly SMA bibliography is compiled by an external consultancy firm on behalf of Biogen to ensure that no relevant studies were overlooked. The company also stated in their clarification response that *"due to the availability of head-to head data, it was considered unnecessary to perform a systematic literature review to identify further comparator studies for an indirect comparison analysis"* (Company's clarification response,[2] question A2).

As part of the NICE Single Technology Appraisal (STA) process, it is a requirement for the company to present a systematic review of the clinical effectiveness evidence. The review should have addressed the decision problem set out in the NICE scope (see Table 3).

4.2 Critique of studies of nusinersen for treating SMA

4.2.1 Studies included in the submission

The company states that there are 10 studies in the nusinersen development programme. These studies are shown in Table 4, and include four patient groups: (i) pre-symptomatic; (ii) infantile onset; (iii) later onset and (iv) both infantile and later onset. Of the studies listed in Table 4, ENDEAR (CS3B), in infantile onset patients and CHERISH (CS4), in later onset patients, are the two studies presented as the key evidence in the CS.[1] The CS presents results for these two key studies together with additional results from the NURTURE study (pre-symptomatic patients), which is stated to be a supporting study.

Table 4: Nusinersen studies identified in the CS (adapted from CS, Figure 3)

Pre-symptomatic patients	Infantile onset	Both infantile and later onset	Later onset (Type I and Type II)
CS5 **NURTURE**: Phase II, open-label, target enrolment n=25	CS3B **ENDEAR**: Phase III, RCT n=122	CS7 **EMBRACE** Phase II, open-label, n=21 enrolled	CS4 **CHERISH**: Phase III RCT n=126
	CS3A: Phase II, open-label , n=21 enrolled	CS11 **SHINE:** Phase III, extension for CS3B, CS4 and CS12, open-label, target enrolment n=274	CS1: open-label, dose escalation, n=28
			CS10: extension for CS1, open-label, n=18
			CS2: open-label, dose-escalation, n=34
			CS12: extension for CS2 and CS10, n=47

RCT - randomised controlled trial; n - number

4.2.1 Critique of quality assessment

The CS[1] included quality appraisals of the ENDEAR, CHERISH and NURTURE studies. The company used the Centre for Reviews and Dissemination (CRD) checklist[16] to assess the study quality of ENDEAR and CHERISH; this checklist is appropriate for the assessment of RCTs and is recommended in the NICE guide for preparing company submissions.[17] In addition, a quality assessment checklist for quantitative intervention studies taken from the Methods for the Development of NICE Public Health Guidance[18] was provided in CS Appendix D. The ERG have not considered this checklist as the NICE guide for company submissions[17] recommends the use of the CRD checklist.[16] Quality assessment of NURTURE was undertaken using only the quality appraisal checklist for quantitative intervention studies[18] in the CS. The ERG has used the Newcastle-Ottawa Scale[19] for assessing the quality of NURTURE, as it is an appropriate and validated quality assessment tool for non-randomised studies. The CS does not provide details regarding the number of reviewers who undertook the quality assessments, nor does it state whether, if more than one reviewer was involved, they undertook quality appraisal independently from one another.

4.2.2 Early onset studies

The ENDEAR study is the main source of evidence for patients with infantile onset SMA. The key study characteristics of ENDEAR are presented in Table 5.

Table 5: ENDEAR study characteristics (adapted from CS, Table 5 and Table 7)

Study	Location (sites)	Design	Population	Interventions	Comparator	Primary outcome measure	Secondary outcome measures	Duration
ENDEAR	31 secondary care settings in Austria, Belgium, Canada, France, Germany, Italy, Japan, Korea, Spain, Sweden, Turkey, UK, USA	Phase III, randomised, double blind	Symptomatic infantile onset SMA,(n=122); those who have or are most likely to develop SMA Type 1	Nusinersen (n=80); administered as a single intrathecal lumbar puncture injection with a scaled 12mg loading dose on study days 1, 15, 29 and 64. Maintenance dosing every 4 months (days 183 and 302)	Sham procedure control (n=41)	Proportion of motor milestone responders (HINE-2) Event-free survival (EFS): Time to death or permanent ventilation	CHOP INTEND responders Proportion of CMAP responders Survival rate Participants not requiring permanent ventilation Time to death or permanent ventilation by disease duration subgroup	Unclear, study terminated early when at least 80 infants had been enrolled for at least 6 months, 27 months from date of first treatment to last patient visit[14]

CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CMAP - compound muscle action potential; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination

Patients

Patients in the ENDEAR study were infants with symptomatic infantile onset SMA. Infants enrolled in the study had:

- Signed informed consent of parent(s) or guardian(s)
- A genetic diagnosis of 5q-linked SMA due to homozygous gene deletion or compound heterozygote deletion/mutation of *SMN1*
- Two copies of the *SMN2* gene; younger than 6 months of age (180 days) at SMA symptom onset
- Younger than 7 months of age (210 days) at screening;
- Receiving adequate nutrition and hydration (with or without gastrostomy) in the opinion of the site investigator at the time of study entry
- Measuring to at least the third percentile in body weight using country-specific guidelines
- Adherence to the consensus statement for standard of care in SMA for medical care guidelines
- Gestational age of 37–42 weeks
- Live within a 9-hour ground travel time from a study centre
- Ability to complete all study procedures and parent/guardian has adequate psychosocial support.[1]

Exclusion criteria for the ENDEAR study can be found in Appendix 1. Table 6 presents the baseline characteristics of patients enrolled into the ENDEAR study.

Table 6: ENDEAR baseline demographics of the ITT population (adapted from CS, Table 11)

Characteristic	**Nusinersen (N=80)**	**Sham control (N=41)**
Female, n (%)	43 (54)	24 (59)
[redacted]	[redacted]	[redacted]
Mean (range) age at first dose, day	163 (52, 242)	181 (30, 262)
Mean (range) age at symptom onset, week	7.9 (2, 18)	9.6 (1, 20)
Mean (range) age at SMA diagnosis, week	12.6 (0, 29)	17.5 (2, 30)
Mean (range) disease duration at screening, week	13.2 (0, 25.9)	13.9 (0, 23.1)
[redacted]	[redacted]	[redacted]
SMA symptoms, n (%) Hypotonia Developmental motor delay Paradoxical breathing Pneumonia or respiratory symptoms Limb weakness Swallowing or feeding difficulties Other	 80 (100) 71 (89) 71 (89) 28 (35) 79 (99) 41 (51) 20 (25)	 41 (100) 39 (95) 27 (66) 9 (22) 41 (100) 12 (29) 14 (34)

Characteristic	Nusinersen (N=80)	Sham control (N=41)
Use of a ventilation support, n (%)	21 (26)	6 (15)
Use of a gastrointestinal tube, n (%)	7 (9)	5 (12)
Total HINE-2 score, mean (SD)	1.29±1.07	1.54±1.29
CHOP INTEND score at baseline, mean (SD)	26.63 (8.13)	28.43 (7.56)
CMAP amplitude, mV, mean (SD) Ulnar nerve Peroneal nerve	 0.226 (0.19) 0.371 (0.31)	 0.225 (0.12) 0.317 (0.29)

CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CMAP - compound muscle action potential; ITT – intention-to-treat; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; SD - standard deviation; SMA - spinal muscular atrophy; SMN – survival of motor neuron. Source: Finkel 2017[20]; ENDEAR CSR[14]

Overall, demographic and baseline disease characteristics and SMA history of the intention-to-treat (ITT) population in the ENDEAR study are consistent with a population highly likely to develop Type I SMA.[4] The groups were similar, although patients in the nusinersen group were on average younger than those in the control group and had an earlier age of symptom onset. Information on subgroups relating to age of onset of symptoms is provided in Section 4.2.7. There was an apparent imbalance with regard to SMA symptoms, with more infants in the nusinersen group (n=80) than the control group (n=41) having the following: history of paradoxical breathing (89% vs 66%), pneumonia or respiratory symptoms (35% vs 22%), ventilator support (26% vs 15%) and more swallowing or feeding difficulties than the control group (51% vs 29%), (see Table 6). The difference in symptoms implies a worse prognosis for the nusinersen group. The ERG's clinical advisors suggested that patients in ENDEAR had a lower use of ventilation and tubes than would be expected in this patient population.

Intervention and comparator

Nusinersen was administered in the ENDEAR study as a single intrathecal lumbar puncture injection on study days 1, 15, 29 and 64 followed by maintenance dosing once every four months (days 183 and 302). Dosage was adjusted for age in order to be equivalent to a 12mg dose in a person two years of age or older. The sham procedure was a small needle prick to the skin over the lumbar spine covered with a bandage. In response to a clarification request from the ERG regarding the use of sedation in the ENDEAR study, the company stated that *"in ENDEAR 6 (8%) of nusinersen treated patients and 2 (5%) of sham control patients received inhalation anaesthesia and 2 (3%) and 0 respectively received intravenous sedation"* (Company's clarification response,[2] question A6).

Quality assessment for ENDEAR

Table 7 compares the quality assessments of the ENDEAR study undertaken by the company and the ERG.

Table 7: Company and ERG quality assessment for ENDEAR (adapted from CS, Table 18)

Quality assessment question	Company's quality assessment	ERG's quality assessment
Was randomisation carried out appropriately?	Yes	Yes: performed using an interactive voice/web response system.[4, 21]
Was the concealment of treatment allocation adequate?	Yes	Yes: performed using an interactive voice/web response system.[4, 21]
Were the groups similar at the outset of the study in terms of prognostic factors?	Partly: Baseline demography was balanced between the nusinersen and control groups. Patients enrolled in the nusinersen treatment group showed greater disease severity compared with the sham-control group.	Unclear: It appears that patients randomised to receive nusinersen had earlier symptom onset and greater burden of disease than patients randomised to the control group.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Partly: Very few participants received sedation (see clarification response,[2] question A6), although participants' age may negate this. Outcome assessors may have been able to determine which participants received a lumbar puncture due to related AEs.
Were there any unexpected imbalances in drop-outs between groups?	No	Yes: A disproportionately high proportion of participants in the control group dropped out (17/41 - 41%) compared with the nusinersen group (15/80 - 19%), according to data on clinicaltrials.gov[22] and the clinical study report (CSR).[21] In most cases (16/41 and 13/80, respectively[21]), this was due to an AE.
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Unclear: In the protocol registered on clinicaltrials.gov, secondary outcome measures 8, 9, 10, 11 and 12 relating to specific types of AEs do not appear to be reported in the Finkel *et al* paper,[20] although these outcomes are reported on clinicaltrials.gov.[22]
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes: Participants who died or withdrew were counted as non-responders.[4]
Summary rating	**Low risk of bias**	**Moderate risk of bias**

AEs - adverse events; ITT – intention-to-treat

Overall, the CS[1] rated ENDEAR as a good quality study, with a low risk of bias. The ERG agrees with this in terms of randomisation, allocation concealment, and ITT analysis. The quality assessments undertaken by both the company and the ERG agree that there are differences between the nusinersen and control groups on some key variables at baseline. The CS and ERG differ in terms of ratings of:

- *Blinding:* The CS rated this item as "yes" (low risk of bias), however the ERG rated it as "partly" (moderate risk of bias) and noted that very few patients were sedated or received inhalational anaesthesia. However, due to patients' age, it is unlikely that patients would have been aware of which treatment they were receiving. Outcome assessors, however, may have been able to determine which participants had received a lumbar puncture according to which participants experienced AEs associated with lumbar puncture.
- *Unexpected imbalances in drop-outs between groups:* The CS rated this item as "no" (low risk of bias). However, the ERG noted an imbalance (as reported on the clinicaltrials.gov study record[22]), in that there were twice as many drop-outs in the control group compared with the nusinersen group (41% versus 19%, respectively); drop-outs were counted as non-responders, although it was not clear whether they improved or deteriorated.
- *Unreported outcome measures:* The CS rated the item "Is there any evidence to suggest that the authors measured more outcomes than they reported?" as "no" (low risk of bias). However, the ERG noted that some of the specific AE-related outcomes were pre-specified in the protocol on clincaltrials.gov, but results on these outcomes were not provided in the Finkel *et al* paper.[20] Findings relating to these outcomes are reported on clinicaltrials.gov.[22]

Results for early onset study (ENDEAR)

All of the outcomes listed in the final NICE scope[12] (see Table 3) are included in the CS for the ENDEAR study, except for complications (such as scoliosis and muscle contractures), stamina and fatigue and HRQoL. The clinical advisors to the ERG suggested that although scoliosis and muscle contractures are relevant outcomes for patients, they would be difficult to measure in short-term studies. Therefore, this omission was considered to be reasonable. As there are no validated questionnaires for stamina and fatigue for younger children, this omission was also considered to be reasonable. Results relating to AEs and HRQoL are presented in Sections 4.2.6 and 4.2.7.

The results of the ENDEAR study are presented in the CS using three different analyses sets (see Table 8). At the interim analysis for ENDEAR, the decision was made terminate the study early due to the benefit-risk assessment being in favour of nusinersen. Infants who completed the ENDEAR study were invited to enrol in the SHINE study, including those in the control arm.

Table 8: ENDEAR analysis sets (adapted from CS, Table 15)

Analysis	Number of patients	Description
Interim (15 June 2016)	Nusinersen: 51 Sham control: 27	Infants in the ITT set who were assessed at the day 183, 302, or 394 visit and had a time difference of at least 190 days between the date of first dose and the data cut-off date of the interim analysis
Final efficacy set (21 November 2016)	Nusinersen: 73; Sham control: 37	Infants in the ITT set who were assessed at the day 183, 302, or 394 visit and had a time difference of at least 190 days between the date of the first dose and the data cut-off date of the final analysis
Final ITT set (21 November 2016)	Nusinersen: 80; Sham control: 41	All infants who were randomised and received ≥1 dose of study drug

ITT – intention-to-treat

Motor function

Motor function was measured in the ENDEAR study using three measures: Module 2 of the Hammersmith Infant Neurological Examination (HINE-2 - the primary endpoint); the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND) and the Compound Muscle Action Potential (CMAP), an electrophysiological technique used to measure nerve function, were both secondary outcomes. Responders were infants with a greater number of motor milestone categories with improvement than worsening.[4] Motor function outcomes are shown in Table 9.

Table 9: ENDEAR motor function outcomes (adapted from CS, Table 19)

Outcome	Nusinersen	Control	Difference (95% CI) and *p*-value
Interim analysis (data cut-off 15 June 2016) (interim analysis set)			
HINE-2 proportion responders	21 (41%)	0 (0%)	41.18 (18.6, 61.20); $p<0.001$
Final analysis (data cut-off 21 November 2016) (efficacy analysis set)			
HINE- 2 proportion responders	37 (51%)	0 (0%)	[redacted]; $p<0.0001$
HINE -2 proportion with improvement in total score	49 (67%)	5 (14%)	
HINE -2 proportion with worsening in total score	1 (1%)	8 (22%)	
CHOP INTEND proportion with ≥ 4 point improvement	52 (71%)	1 (3%)	[redacted]; $p<0.001$
CHOP INTEND proportion with any improvement	53 (73%)	1 (3%)	
CHOP INTEND proportion with any worsening	5 (7%)	18 (49%)	
CMAP amplitude responders	26 (36%)	2 (5%)	$p=0.001$

CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CI - confidence interval; CMAP - compound muscle action potential; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination

As shown in Table 9, a significantly greater percentage of patients in the nusinersen group achieved motor milestone responses compared with the control group, although many patients in the nusinersen group (49%) could not be classified as responders. In the nusinersen group, 22% of infants achieved full head control, 10% were able to roll over, 8% were able to sit independently and 1% were able to stand. In the control group, no infants achieved these milestones.[1]

Respiratory function

The only measure of respiratory function reported from the ENDEAR study was the annualised rate of serious respiratory events; 2.836 events were reported in the nusinersen group versus 3.065 events in the control group in the interim analysis (95% confidence intervals [CIs] not reported).[4] [redacted]

Ventilation

The ENDEAR study reported the number of hours of ventilator support as a measure of ventilation. In the interim analysis, the median percentage of time on ventilator support was lower in the nusinersen group (27.1%) compared with the control group (43.0%).[4] [redacted]. Outcomes relating to the endpoints of use of permanent assisted ventilation and time to death or permanent ventilation are presented in Table 10. A higher percentage of nusinersen patients had no use of permanent ventilation compared with the control group, although the difference was not statistically significant (p=0.13).

Mortality

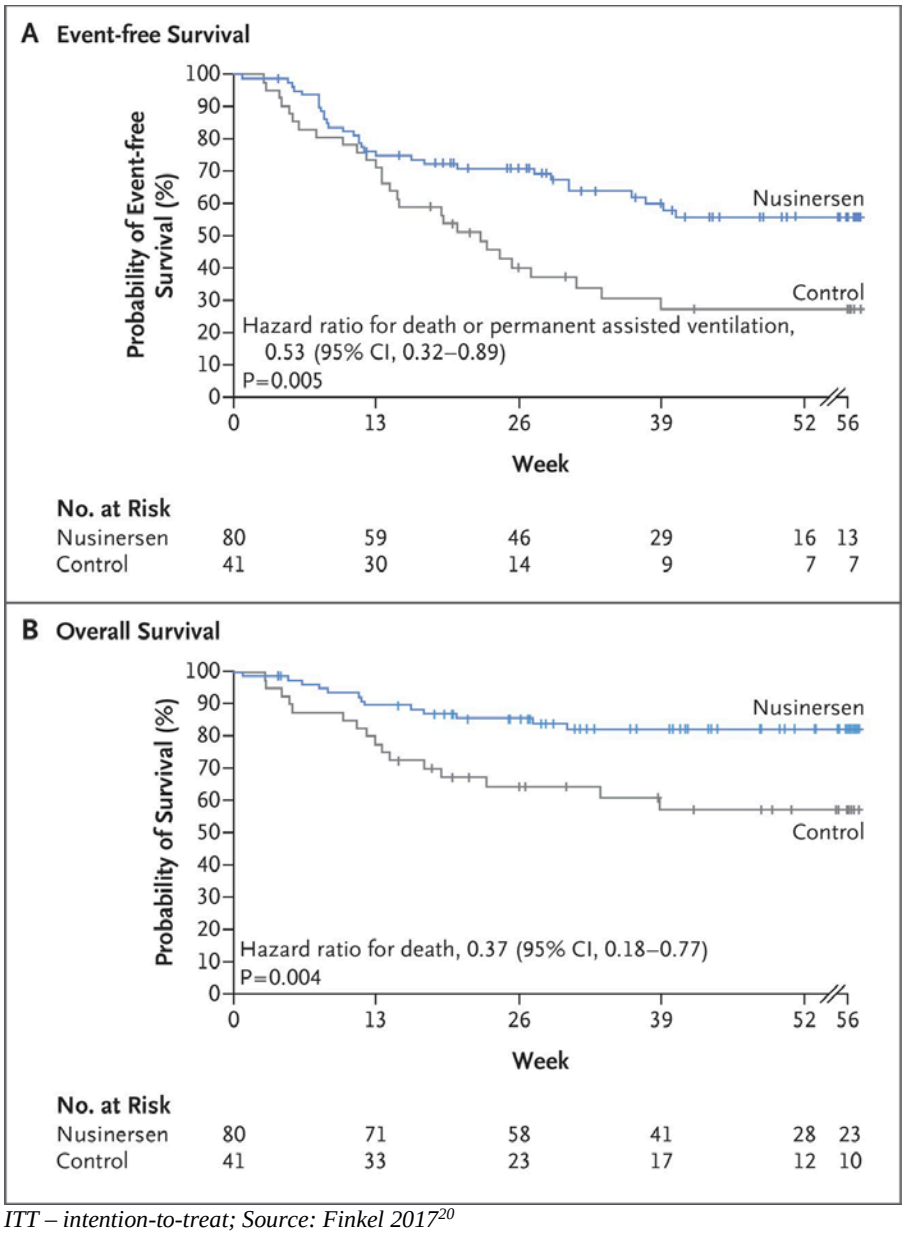
Measures of mortality within ENDEAR included event-free survival (EFS), defined as time to death or permanent ventilation (primary endpoint) and overall survival (OS); results for these outcomes are shown in Table 10. Statistically significant increases in both EFS (p=0.005) and OS (p=0.004) were observed for the nusinersen group. Figure 2 presents the associated Kaplan-Meier curves for these outcomes.

Table 10: ENDEAR study ventilation and survival outcomes (adapted from CS, Table 19)

Outcome	Nusinersen	Control	Difference (95% CI) or HR (95% CI) and *p*-value
No use of permanent assisted ventilation (ITT analysis set)	62 (78%)	28 (68%)	0.66 (0.32-1.37) *p*=0.13
EFS (ITT analysis set) (patients who had died or received permanent assisted ventilation)	31 (39%)	28 (68%)	HR: 0.53 (0.32, 0.89) *p*=0.005
OS (ITT analysis set) Dead Alive	 13 (16%) 67 (84%)	 16 (39%) 25 (61%)	HR: 0.37 (0.18, 0.77); *p*=0.004

ITT – intention-to-treat; EFS - event-free survival; HR – hazard ratio

Figure 2: ENDEAR Kaplan-Meier curves for EFS (A) and OS (B) (ITT population, final analysis) (reproduced from CS, Figure 15)



ITT – intention-to-treat; Source: Finkel 2017[20]

Number and length of hospitalisations

The number and length of hospitalisations was not included as an outcome in the NICE scope;[12] however, this outcome was included in the CS[1] (page 73) and is presented here for completeness. The adjusted annualised rates of hospitalisation in the nusinersen group were 4.378 (95% CI: 3.636 to 5.273) compared with 5.817 (95% CI: 3.636 to 5.273) hospitalisations/year in the control group (p=0.0959). Overall time spent hospitalised was significantly lower in the nusinersen group than the control group (LSM: 0.114 versus 0.207 [unit of time unclear from the CS]; LSM treatment difference: -0.093; 95% CI -0.151 to -0.034; p=0.0022).

Additional early onset study: CS3A

One additional early onset study, CS3A, was presented in the CS.[1] Table 11 below presents the study characteristics for CS3A.

Table 11: Summary of study characteristics for CS3A (based on data reported in CS Appendix L, Table 20)

Study ID	**CS3A**
Study objectives	Safety, tolerability, efficacy and PK
Study type/design	Phase II, open-label, multiple dose, single arm
Study population	Symptomatic, infantile onset SMA: 17 of 20 subjects (85%) had 2 copies of the *SMN2* gene (all 4 subjects in Cohort 1 and 13 subjects in Cohort 2); 2 subjects had 3 copies of the *SMN2* gene
Primary efficacy endpoint	Motor milestones (HINE Module 2)
Secondary efficacy endpoints	CHOP INTEND, OS and EFS
Intervention(s)	Nusinersen – Cohort 1: 6mg scaled equivalent loading dose, 12mg maintenance dose Cohort 2: 12mg scaled equivalent loading dose and 12mg maintenance dose Loading dose: days 1, 15, 85 Maintenance dose: day 253 and every 4 months thereafter
Number of patients dosed	TOTAL: 20 Cohort 1: 4 Cohort 2: 16 1 subject withdrew before dosing
Mean (median) age at baseline	141 (155) days (range 36–210 days)
Mean (median) age at symptom onset	60 (56) days

CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; HINE - Hammersmith Infant Neurological Examination; SMA – spinal muscular atrophy; SMN - survival of motor neurone; PK - pharmacokinetics

Key results for CS3A, as outlined in CS[1] Appendix L (pages 134-135) were:

- Change in HINE-2 score from baseline to last visit was significant for both cohorts combined (p=0.0002) and for participants in the 12mg dose group (p<0.0001).
- HINE-2 motor milestones increased steadily over time from a baseline mean score of 2.25 up to a mean increase of 9.40 milestones on day 694.
- CHOP-INTEND scores showed a mean increase of 11.5 points from baseline to last visit (p=0.0080, n=18)
- 15 of 20 subjects (75%) were alive and continuing the study at data cut-off.
- 13 subjects (65%) were free from permanent ventilation and continuing the study at data-cut-off.

4.2.3 Later onset studies

The CHERISH study is the main source of evidence for patients with later onset SMA. The characteristics of the CHERISH study are presented in Table 12.

Table 12: CHERISH study characteristics (adapted from CS, Table 5 and Table 7)

Study	Location (sites)	Design	Population	Interventions	Comparator	Primary outcome measure	Secondary outcome measures	Duration
CHERISH	24 centres in Canada, China, France, Germany, Italy, Japan, Korea, Spain, Sweden, USA	Phase III, randomised double-blind study in secondary care	Symptomatic later onset SMA (n=126); those who have or are most likely to develop SMA Type II or III	Nusinersen (n=84) administered as single intrathecal lumbar puncture injection. Single dose level 12mg delivered in 4 doses over 9 months using a loading regimen (days 1, 29. 85); maintenance dose given 6 months later (day 274)	Sham control (n=42)	HFMSE	≥ 3 point increase in HFMSE score WHO motor milestone Standing alone Walking with assistance RULM	Unclear; early termination of study after analysis of primary endpoint at the interim analysis; date from first treatment to last visit for last patient: 27 months[15]

HFMSE - Hammersmith Functional Motor Scale-Expanded; RULM - Revised Upper Limb Module; WHO - World Health Organization

Patients

Patients enrolled in the CHERISH study had later onset SMA with symptom onset after six months of age. The inclusion criteria for CHERISH were:

- Signed informed consent of parent(s) or guardian(s) and signed informed assent of child (if indicated per child's age and institutional guidelines)
- Genetic documentation of 5q-linked SMA due to homozygous gene deletion, mutation, or compound heterozygote of *SMN1*
- Onset of clinical signs and symptoms consistent with SMA at more than 6 months of age
- Age 2 to 12 years inclusive
- Able to sit independently but never had the ability to walk independently
- Hammersmith Functional Motor Scale-Expanded (HFMSE) score of 10 or higher and 54 or lower at screening
- Able to complete all study procedures, measurements, and visits and parent or guardian/child had adequately supportive psychosocial circumstances; estimated life expectancy more than 2 years from screening; met age-appropriate institutional criteria for use of anaesthesia/sedation if use was planned for study procedures
- For those individuals who may have reached reproductive maturity, females must have had a negative pregnancy test at screening and agree to employ adequate contraceptive measures for the duration of the study, and males were to be abstinent for the duration of the study.[1]

Exclusion criteria for the CHERISH study can be found in Appendix 1. Mercuri *et al*[23] state that one of the limitations of the study was the application of strict eligibility criteria (no severe contractures or scoliosis, outlying HFMSE scores, respiratory insufficiency or reliance on a gastric tube), which meant that the study population was more homogenous and younger than the population that is encountered in usual clinical practice. The baseline characteristics of the patients in the CHERISH study are shown in Table 13.

Table 13: CHERISH baseline demographics in the ITT population (reproduced from CS, Table 12)

Characteristic	**Nusinersen (N=84)**	**Sham-procedure control (N=42)**
Female, n (%)	46 (55)	21 (50)
White, n (%)	64 (76)	30 (71)
Median (range) age at screening, years	4.0 (2–9)	3.0 (2–7)
Median (range) age at symptom onset, months	10.0 (6–20)	11.0 (6–20)
Median (range) time from disease onset to enrolment, months	39.3 (8–94)	30.2 (10–80)
Median (range) age at SMA diagnosis, months	18.0 (0–48)	18.0 (0–46)

Characteristic	Nusinersen (N=84)	Sham-procedure control (N=42)
Median (range) time from diagnosis to enrolment, months	27.8 (2–86)	26.0 (2-72)
Median (range) disease duration, months	39.3 (8–94)	30.2 (10–80)
SMN2 copy number, 2/3/4/unknown, %	7/88/2/2	10/88/2/0
Children who have ever achieved motor milestone, n (%) Sat without support Walked with support Stood without support Walked ≥15 feet independently	 84 (100) 20 (24) 11 (13) 0	 42 (100) 14 (33) 12 (29) 0
Children using a wheelchair, n (%)	64 (76)	29 (69)
Mean (SD) HFMSE total score [a]	22.4 (8.3)	19.9 (7.2)
Mean (SD) WHO total score [a,b]	1.4 (1.0)	1.5 (1.0)
Mean (SD) RULM total score [a,c]	19.5 (6.2)	18.4 (5.7)

HFMSE - Hammersmith Functional Motor Scale-Expanded; ITT - intention-to-treat; RULM - Revised Upper Limb Module; SD - standard deviation; SMA - spinal muscular atrophy; SMN - survival motor neuron; WHO - World Health Organization; [a] Baseline is defined as the last non-missing value before the first dose of nusinersen or sham-procedure control. [b] If the baseline value as defined above was missing, then baseline was imputed as the median of the non-missing values of the stratum to which the child belongs: age<6 or ≥6 years. [c] One child had a missing value and this was imputed as the median baseline value of the child across all the multiply imputed datasets. Source: Mercuri 2018[23]

As stated in the CS,[1] overall, the groups were similar although there was an imbalance in the proportion of patients who had ever achieved a motor milestone and an imbalance in the median time from disease onset to study enrolment, with a longer delay in receiving therapy in the nusinersen group than the sham group. The nusinersen group had a slightly higher HFMSE total score at baseline, indicating slightly better motor function.

Intervention and comparator

Nusinersen was administered intrathecally as a single lumbar puncture injection using a loading dose on study days 1, 29 and 85, followed by maintenance dosing 6 months thereafter (starting on day 274). The sham control procedure was administered on days 1, 29, 85 and 274 using the same administration procedure as in the ENDEAR study. In the CHERISH study, however, if anaesthesia or sedation were used in a study site for the administration of nusinersen, then minimal sedation was used for the sham procedure.

Quality assessment for CHERISH

Table 14 presents the quality assessment of the CHERISH trial undertaken by the company and the ERG.

Table 14: Company and ERG quality assessment for CHERISH (adapted from CS, Table 18)

Quality assessment question	Company's quality assessment	ERG's quality assessment
Was randomisation carried out appropriately?	Yes	Yes: Performed using an interactive web response system.
Was the concealment of treatment allocation adequate?	Yes	Yes: Performed using an interactive web response system.
Were the groups similar at the outset of the study in terms of prognostic factors?	Partly: Baseline demography was balanced between the nusinersen and control groups. There was an imbalance in the proportion of patients who had ever achieved a milestone, with fewer patients in the nusinersen group than in the control group having stood without support, and having walked with support; more patients in the nusinersen group used a wheelchair than in the control group.	Unclear: Differences between groups were not examined statistically. It appears that fewer patients randomised to receive nusinersen had ever achieved a milestone, stood without support, and walked with support, and more nusinersen group patients using a wheelchair than among the control group. The nusinersen group had a slightly higher HFMSE total score at baseline.
Were the care providers, participants and outcome assessors blind to treatment allocation?	Yes	Partly: 51% of nusinersen and 57% sham patients received inhalational anaesthesia, and 86% nusinersen and 81% sham patients received intravenous sedation (see clarification response,[2] question A6). Therefore, as patients ranged from 2 to 9 years of age, some may not have been adequately blinded. Outcome assessors may have been able to determine which participants received a lumbar puncture due to related AEs.
Were there any unexpected imbalances in drop-outs between groups?	No	No
Is there any evidence to suggest that the authors measured more outcomes than they reported?	No	Unclear: Some of the specific AE-related outcomes were pre-specified in the protocol on clincaltrials.gov, but results on these outcomes were not provided in the Mercuri *et al* paper.[23] These outcomes are reported on clinicaltrials.gov.[24]
Did the analysis include an ITT analysis? If so, was this appropriate and were appropriate methods used to account for missing data?	Yes	Yes: The imputation methods are reasonable; sensitivity analysis using other imputation methods yielded similar results.
Summary rating	**Low risk of bias**	**Moderate risk of bias**

AEs, adverse events; HFMSE, Hammersmith Functional Motor Scale-Expanded; ITT- intention-to-treat

Overall, the CS[1] rated CHERISH as a good quality study, with a low risk of bias. The ERG agrees with this in terms of randomisation, allocation concealment, and ITT analysis. The quality assessments undertaken by both the company and the ERG agree that there are differences between the nusinersen and control groups on some key variables at baseline. The quality assessments differ in terms of ratings of:

- *Blinding:* The CS rated this item as "yes" (low risk of bias). However, the ERG rated it as "partly" (moderate risk of bias) and noted that not all patients received inhalational anaesthesia (51% nusinersen and 57% sham control) or intravenous sedation (86% nusinersen and 81% sham control) (see clarification response,[2] question A6), and participants' ages ranged from 2 to 9 years, therefore, some participants may have been aware of which treatment they received (nusinersen or sham). In addition, outcome assessors may have been able to determine which participants had received a lumbar puncture according to which participants experienced AEs relating to this procedure.
- *Unreported outcomes:* The CS rated the item, "Is there any evidence to suggest that the authors measured more outcomes than they reported?" as "no" (low risk of bias). However, the ERG noted that some of the specific AE-related outcomes were pre-specified in the protocol on clincaltrials.gov, but results for these outcomes were not provided in the paper reported by Mercuri *et al*.[23] The findings relating to these outcomes are reported on clinicaltrials.gov.[24]

Results for later onset study (CHERISH)

Motor function, AEs and HRQoL were collected in the CHERISH study and are presented in the CS.[1] However, outcomes relating to respiratory function, complications, ventilation, stamina, fatigue and mortality, which were included in the NICE scope,[12] were not collected. Three separate efficacy sets were used in the CHERISH study (see Table 15).

Table 15: CHERISH efficacy sets (adapted from CS, Table 16)

Population	Number of patients	Description
Interim efficacy set (31 August 2016)	Nusinersen: 35 Control: 19	A subset of the ITT set who had been assessed at month 15 (i.e. the day 456 visit), which included all children with a day 456 visit and all children with a time difference of at least 463 days (456 days plus a 7-day window) between the date of first dose and the data cut-off date for the interim analysis (August 31, 2016). Used for the main interim analysis of motor milestones and also as a supportive analysis for the primary endpoint and all other secondary efficacy endpoints.
Efficacy set (3 March 2017)	Nusinersen: 66 Control: 34	Subset of children in the ITT set who had the opportunity to be assessed at the day 456 visit (i.e., month 15), which included all children with a day 456 visit and all children with a time difference of at least 463 days (456 days plus a 7-day window) between the date of first dose and the date for the final analysis. Used for the analysis of WHO motor milestones.
ITT set (3 March 2017)	Nusinersen: 84 Control: 42	All patients who were randomised and received ≥1 dose of the study drug or control procedure. Children were analysed in the treatment group to which they were randomised. Used for the change from baseline to month 15 in HFMSE score, percentage of HFMSE responders, and change in RULM score.

HFMSE – Hammersmith Functional Motor Scale-Expanded; ITT – intention-to-treat; RULM - Revised Upper Limb Module; WHO – World Health Organization

Motor function

Motor function was measured using HFMSE scores, the World Health Organization (WHO) criteria motor milestones and the Revised Upper Limb Module (RULM) measure. Outcomes relating to these motor function endpoints are presented in Table 16. HFMSE is a validated tool to assess motor function in children with SMA; higher scores indicate better function. The WHO motor milestones are a set of six gross motor milestones (sitting without support, standing with assistance, hands and knees crawling, walking with assistance, standing alone, walking alone) that are expected to be attained by 24 months in healthy children. The RULM measure was used to assess upper limb functional abilities in people with SMA.

Table 16: CHERISH motor function outcomes (adapted from CS, Table 20)

Outcome	Nusinersen	Control	Difference (95% CI) and *p*-value
Interim analysis (data cut-off 31 August 2016)			
HFMSE score: change from baseline in HFMSE (95% CI)	4.0 (2.9, 5.1)	-1.9 (-3.8, 0.0)	LSM change difference: 5.9 (3.7, 8.1); $p<0.001$
Final analysis (data cut-off 3 March 2017)			
HFMSE score: change from baseline in HFMSE (95% CI)	3.9 (3.0, 4.9)	-1.0 (-2.5, 0.5)	LSM change difference: 4.9 (3.1, 6.7); $p=0.0000001$[a]
Proportion of children with change(%) in HMSE score of ≥3 points (95% CI)	57 (46, 68)	26 (12, 40)	Odds ratio: 6 (2, 15); $p<0.001$
Motor milestones at 15 months: % who achieved ≥1 new motor milestone (95% CI)	20 (11,31)	6 (1, 20)	Difference in proportions14 (-7, 34); $p=0.08$
WHO criteria motor milestones at 15 months: LSM number of new motor milestones achieved per child (95% CI)	0.2 (0.1, 0.3)	-0.2 (-0.4, 0.0)	LSM difference 0.4 (0.2, 0.7); $p=0.0001$
WHO criteria motor milestones at 15 months: % who achieved standing alone (95% CI)	2 (0, 8)	3(0, 15)	Difference in proportions: -1 (-22, 19); $p>0.9999$
WHO criteria motor milestones at 15 months: % who achieved walking with assistance (95% CI)	2 (0, 8)	0 (0, 10)	Difference in proportions: 1.5 (-19.1, 22.0); $p>0.9999$
RULM: change from baseline at 15 months (95% CI)	4.2 (3.4, 5.0)	0.5 (-0.6, 1.6)	LSM difference: 3.7 (2.3, 5.0); $p=0.0000001$

CI - confidence interval; HFMSE - Hammersmith Functional Motor Scale-Expanded; LSM - least squares mean; RULM - Revised Upper Limb Module; WHO - World Health Organization

[a] *Because the p-value for the primary endpoint was significant in the interim analysis, this endpoint was not formally tested for significance in the final analysis. The exploratory p-value is not reported in the full publication and is from Mercuri et al*[23]

As shown in Table 16, compared with the control group, patients in the nusinersen group showed significant improvement in HFMSE scores from baseline, an increase in the number of new motor milestones achieved per child according to the WHO criteria and improvement in RULM score from baseline.

Additional late onset studies: CS1, CS10, CS2 and CS 12

Four additional late onset studies are presented in the CS: CS1, CS10 (extension for CS1), CS2 and CS12 (extension for CS2 and CS10).[1] The characteristics of these studies are shown in Table 17.

Table 17: Study characteristics for additional late onset studies (adapted from CS Appendix L, Table 20)

Study ID	CS1	CS10	CS2	CS12
Study objectives	Safety, tolerability, dose finding, and efficacy	Safety, tolerability, efficacy, and PK	Safety, tolerability, efficacy, and PK	Safety, tolerability, efficacy, and PK
Study type/design	Phase I, open-label, escalating dose	Phase I, open-label, single dose	Phase I, open-label, dose escalation, multiple dose	Phase I, open-label, multiple dose, single arm
Study population	Symptomatic, later onset SMA: 15 subjects (54%) had Type II SMA and 13 (46%) had Type III SMA	Symptomatic, later onset SMA in patients who previously participated in CS1: 10 subjects (56%) had Type II SMA and 8 subjects (44%) had Type III SMA	Symptomatic, later onset SMA: 13 subjects (38%) had Type II SMA and 21 (62%) had Type III SMA	Symptomatic, later onset SMA in CS10 and CS2: 22 subjects (47%) had Type II SMA and 25 (53%) had Type III SMA
Primary efficacy endpoint	HFMSE	HFMSE	HFMSE	HFMSE
Secondary efficacy endpoints	PedsQL, CMAP, MUNE	PedsQL, CMAP, MUNE	PedsQL, CMAP, MUNE, ULM, myometry, 6MWT, ACEND	6MWT, ULM, CMAP, PedsQL, ACEND
Intervention(s)	Nusinersen 1mg, 3mg, 6mg and 9mg single dose	Nusinersen – Cohort 1: 6mg on day 1 Cohort 2: 9mg on day 1	Nusinersen – Cohort 1: 3mg on days 1, 29, 85 Cohort 2: 6mg on days 1, 29 and 85 Cohort 3: 9mg on days 1, 85 Cohort 4: 12mg on days 1, 29 and 85 Total duration: approximately 8 months	Nusinersen 12mg Doses on days 1, 169, 351, and 533 Total duration: approximately 1.5 years
Number of patients dosed	**TOTAL: 28** 1mg cohort: 6 3mg cohort: 6 6mg cohort: 6 9mg cohort: 10	**TOTAL: 18** Cohort 1: 4 Cohort 2: 14	**TOTAL: 34** Cohort 1: 8 Cohort 2: 8 Cohort 3: 9 Cohort 4: 9	**TOTAL: 47** 12mg: 47
Mean (median) age at baseline	6.1 years (range 2–14 years)	6.6 years (range –11 years)	7.4 years (range 2–15 years)	8 years (range 3–17 years)
Mean (median) age at symptom onset	Not summarised	Not summarised	Not summarised	Not summarised

6MWT - 6-minute walk test; ACEND - Assessment of Caregiver Experience with Neuromuscular Disease; CMAP - Compound Muscle Action Potential; HFMSE - Hammersmith Functional Motor Scale-Expanded; MUNE - Motor Unit Number Estimation; PedsQL - Paediatric Quality of Life Inventory; PK - pharmacokinetics; SMA - spinal muscular atrophy; ULM - Upper Limb Module; PK - pharmacokinetics

Results from later onset studies

Key findings from these later onset studies were as follows (CS Appendix L, pages 137-139):

CS1 & CS10

- Dose-dependent improvement in HFMSE total score with a mean increase from baseline of 3.1 points (17.6%) at day 85 at the highest dose evaluated (9mg; note – the licensed dose is 12mg)
- 7 of 10 subjects with 9mg dose exhibited improvement of ≥ 3 points in the HFMSE.

CS2 & CS12

- For patients with Type II SMA with up to three years of treatment, there were improvements observed in motor function over time as measured by HFMSE scores and ULM test
- One patient with Type II SMA gained the ability to walk independently
- Two patients with Type III SMA regained the ability to walk independently
- For patients with Type III SMA with up to three years of treatment, HFMSE scores were stable over time
- Increases were observed in 6MWT distances.[25]

Results were not presented separately for Study CS2 and Study CS10 either in the CS or in publications related to these studies.[25, 26]

4.2.4 Ongoing studies

Three ongoing studies are described in the CS: NURTURE (study completion January 2022), SHINE (study completion August 2022), and EMBRACE (study completion April 2019, see Table 18).[1] The CS includes results for the NURTURE study only as data were not available for SHINE and EMBRACE. The CS states that the NURTURE study, in pre-symptomatic infants, is a supportive study.

Table 18: Summary of ongoing nusinersen studies (reproduced from CS, Table 30)

Study	Study title	Design	Subject population	Treatment groups	Interim analyses	Ongoing /updated analyses
SHINE	A Study for Participants With Spinal Muscular Atrophy (SMA) Who Previously Participated in Nusinersen (ISIS 396443) Investigational Studies.	Open-label extension study	Infantile and later onset SMA patients from ENDEAR and CHERISH, CS12 and CS3A	Nusinersen	Estimated dates for interim analyses: Q1 2018 Data cut-off: 30 June 2017	Estimated study completion: August 1, 2022
NURTURE	A Study of Multiple Doses of Nusinersen (ISIS 396443) Delivered to Infants With Genetically Diagnosed and Pre-symptomatic Spinal Muscular Atrophy (NURTURE)	Open-label, Phase II	Genetically diagnosed and pre-symptomatic SMA	Nusinersen	Estimated dates for interim analyses: Q1/Q2 2018 Data cut-off: June 2017	Estimated study completion: January 26 2022
EMBRACE	A Study to Assess the Safety and Tolerability of Nusinersen (ISIS 396443) in Participants With Spinal Muscular Atrophy (SMA). (EMBRACE)	Phase II, randomised, double-blind, sham-procedure controlled study	Patients with SMA who are not eligible to participate in the clinical studies ENDEAR and CHERISH	Nusinersen and Sham	Estimated dates for interim analyses: Part 1: August 10, 2017	Estimated study completion: April 1, 2019

SMA - spinal muscular atrophy; Q - quarter

Patients

The study characteristics for the NURTURE study in pre-symptomatic infants are presented in Table 19. In response to a request for clarification from the ERG2 (question A4), the company stated that 25 out of 30 screened infants have been enrolled. The 25 enrolled infants were identified through diagnosis of an affected sibling (n= 18), a newborn screening programme (n=3), prenatal testing (n=3) and known carrier status (n=1). The company's response to clarification question A82 stated that patients with any clinical signs or symptoms strongly suggestive of SMA at screening or immediately prior to the first dosing were excluded. Table 19 below shows data presented in the CS, for the first 20 patients entered in the NURTURE study only.

The inclusion criteria for NUTURE were:

- Age ≤6 weeks at first dose
- Genetic documentation of 5q SMA homozygous gene deletion or mutation or compound heterozygous mutation
- Genetic documentation of 2 or 3 copies of *SMN2*
- CMAP ≥1 mV at baseline
- Gestational age of 37–42 weeks for singleton births; gestational age of 34–42 weeks for twins.[1]

Table 19: NURTURE study characteristics (adapted from CS, Table 7 and NURTURE CSR[21])

Study	NURTURE (CS5)
Location (sites)	20 study sites in 10 countries including UK
Design	Phase II, open-label, multicentre, single arm study
Population	Pre-symptomatic infants genetically diagnosed with SMA (likely to develop infantile or later onset) (target enrolment: N= 25)
Interventions	Nusinersen (n= 20)
Comparator	None
Primary outcome measure	Respiratory intervention or death
Secondary outcome measures	• Proportion of patients developing clinically manifested SMA as defined by: o Age-adjusted weight <5th percentile or decrease of ≥2 major weight growth curve percentiles (3rd, 5th, 10th, 25th, or 50th) or a percutaneous gastric tube placement for nutritional support o Failure to achieve age-appropriate attainment of the 6 WHO motor milestones • OS, i.e. proportion of patients alive • Percentage of participants who attained motor milestones assessed as part of HINE-2 • Attainment of motor milestones as assessed by WHO criteria • Change from baseline in CHOP INTEND motor function scale • Change in baseline in growth parameters
Duration	Ongoing

CHOP-INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; SMA - spinal muscular atrophy WHO- World Health Organisation

The baseline characteristics for the NURTURE study are presented in Table 20.

Table 20: Baseline characteristics for the NURTURE study (adapted from CS, Table 13, including additional data from the company's clarification response, question A9)

Characteristic	2 SMN2 copies **N=13** [a] ***(n=15)***	3 SMN2 copies **N=7** ***(n=10)***	Total **N=20** ***(n=25)***
Age at first dose, days, n			
≤14 >14 to ≤28 >28 Range	6 5 2 3–41	2 3 2 10–42	8 *(n=9)* 8 *(n=12)* 4 *(n=4)* 3–42
Mean CHOP INTEND total score Median (range) [b]	48.0 50.0 (25–60) c	53.8 56.0 (40–60) d	49.6 54.0 (25–60) e
Mean HINE total motor milestones Median (range) [b]	2.5 3.0 (0–5) c	4.2 4.0 (2–7) d	3.0 3.0 (0–7) e
Mean ulnar CMAP amplitude Median (range), mV [b]	2.62 2.15 (1.0–6.7) c	3.96 4.00 (2.7–4.9) d	2.99 2.85 (1.0–6.7) e
Mean peroneal CMAP amplitude Median (range), mV [b]	2.47 2.65 (0.2–4.2) f	4.88 4.40 (4.0–7) d	3.27 3.20 (0.2–7.0) g
Male, %			55
Region, n North America Europe Asia-Pacific			 13 4 3

CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; CMAP - compound muscle action potential; HINE - Hammersmith Infant Neurological Examination
NURTURE study interim analysis data cut-off date: October 21, 2016. [a] Included 1 set of twins each with 2 SMN2 copies; [b] Based on efficacy set of patients who completed the day 64 visit or longer (N=18) ; [c] N=13. [d] N=5. [e] N=18. [f] N=10. [g] N=15
Source: Crawford 2017[27] Numbers in italics are from clarification response.

In the NURTURE study, 13 patients had 2 SMN2 copies and would therefore be expected to develop a more severe SMA phenotype than subjects with 3 *SMN2* copies, although other genetic modifying factors will affect the type of SMA an individual will develop. Most patients were male, younger than one month and from the US.

Intervention and comparators

Nusinersen was administered intrathecally (12mg equivalent dose) by lumbar puncture with loading doses on days 1, 15, 29 and 64 and maintenance doses on days 183, 302, 421, 540, 659 and 778.

Quality assessment for NURTURE

Table 21 presents the quality assessment of the NURTURE study undertaken by the company and the ERG, based on the Newcastle-Ottawa Scale.[19]

Table 21: Company and ERG quality assessment for NURTURE (adapted from information in CS, Appendix D, pages 21-26) using the Newcastle-Ottawa Scale[19]

Quality assessment question	Company's quality assessment	ERG's quality assessment
Representativeness of the exposed cohort	[redacted]	Unclear
Selection of the non-exposed cohort	[redacted]	N/A (single-arm study)
Ascertainment of exposure	[redacted]	Patients were administered nusinersen as an intervention within the study. Administration was monitored (CS[1] page 48; CSR[21] pages 30-31).
Demonstration that outcome of interest was not present at start of study	Not assessed in CS	Primary outcome is time to respiratory intervention or death, which was not present at baseline (CS,[1] page 49; CSR,[21] page 37). WHO motor milestones were not achieved at baseline due to age (CS,[1] page 52; CSR,[21] page 76).
Comparability of cohorts on the basis of the design or analysis	[redacted]	N/A
Assessment of outcome	[redacted]	Standard clinician-assessed outcome measurements used (CS,[1] pages 28-19 and pages 85-88; CSR,[21] pages 37-40), open-label (CS,[1] pages 28, 48, and 97; CSR[21] page 27).
Was follow-up long enough for outcomes to occur?	[redacted]	Treatment occurred over 778 days (CS,[1] page 32; CSR,[21] page 31), and followed up to interim data cut-off for 421 days (CSR,[21] page 68), during which time motor outcomes occurred, but not death or ventilation, however the median time to death or permanent ventilation is 10.5 months in those with 2 copies of *SMN2* and 13 months overall (CSR[21] page 115), and therefore follow-up should have been long enough.
Adequacy of follow up of cohorts	[redacted]	No withdrawals as of the recent interim analysis (cut-off date 31st October 2016) in CS[1] page 49 and CSR[21] page 56.
Stars total	[redacted]	**5**

[redacted]. The ERG agrees with this in terms of ascertainment of exposure, assessment of outcome, whether follow-up was long enough for outcomes

to occur, and adequacy of follow-up cohort. The quality assessments undertaken by the company and the ERG differ in terms of ratings of:

- *Representativeness of cohort:* [redacted] however, the ERG rated this as "unclear", as information demonstrating how the NURTURE cohort compared with the wider SMA population was not presented in the CS.
- *Demonstration that outcome of interest was not present at start of study:* This item was not assessed in the CS, and the ERG judged NURTURE as "good" on this item.

Results for supportive study in pre-symptomatic infants (NURTURE)

Those infants assessed in the interim analysis had been in the study for a median of 317.5 days (range 2-524 days).

Motor function

Motor function was measured in the NURTURE study by HINE, CHOP INTEND and WHO motor milestones. HINE motor milestones were achieved in 16 of 18 subjects in the efficacy set (89%). At the data cut-off, 12 subjects achieved sitting independently, 9 subjects achieved standing with or without support and 6 subjects achieved walking with or without support.

From baseline, 16 of 18 subjects (89%) achieved and maintained improvements in the CHOP INTEND total score. An increase of ≥4 points in the CHOP INTEND total score from baseline, the chosen definition of a responder in the CS, was seen in 61% of subjects (n=11/18).

With regard to WHO motor milestones, at the last observed visit, 71% of patients had achieved sitting without support, 59% achieved standing with assistance, 29% walking with assistance, 18% standing alone and 12% walking alone. In response to a request for clarification from the ERG[2] (question A9), the company provided the following information: 22 (100%) of infants achieved the WHO motor milestone sitting without support and 8/13 (62%) achieved walking alone, among infants with enough observation time. It is unclear how "enough observation time" was determined.

Mortality and ventilation

All infants were alive and none required invasive ventilation, tracheostomy or non-invasive ventilation (NIV) for ≥6 hours/day continuously for ≥7 days. The company's clarification response[2] (question A9) reported that as of 5th July 2017, all infants were alive and none required tracheostomy or permanent ventilation. Two of 15 infants (13%) with 2 *SMN2* copies required respiratory intervention for ≥6 hours/day continuously for ≥7 days during an acute, reversible viral infection. One additional infant

with 2 *SMN2* copies required respiratory support for ≥6 hours/day continuously for ≥1 day but less than 7 days.

Information on other ongoing studies

The ERG requested further information relating to the SHINE and EMBRACE studies during the clarification stage of the appraisal. Interim results (data cut-off 30th June 2017) from SHINE (infantile onset patients only from ENDEAR) reported additional improvements in total and specific HINE-2 motor milestones and general motor function as measured by CHOP INTEND. Median time to death or permanent ventilation was 73 weeks. Those patients who were in the control group in ENDEAR and who began nusinersen in SHINE showed improvements in total HINE-2 motor milestones and CHOP INTEND scores. No data were presented for later onset patients (from CHERISH) taking part in the SHINE study. The CSR for SHINE was not provided by the company.

Information on the inclusion criteria for the EMBRACE study was provided in the company's clarification response[2] (question A15). Patients included in EMBRACE had genetic documentation of 5q SMA; onset of symptoms ≤6 months with 3 *SMN2* copies or onset of symptoms ≤6 months and aged >7 months at screening with 2 *SMN2* copies or onset of SMA symptoms >6 months and aged ≤18 months at screening with 2 or 3 *SMN2* copies. They did not meet the inclusion criteria for ENDEAR: symptom onset ≤6 months and aged ≤7 months at screening with 2 *SMN2* copies or CHERISH: symptom onset >6 months and aged 2-12 years at screening.

4.2.5 HRQoL

Three measures of HRQoL were assessed in the CHERISH study: the Paediatric quality of life inventory (PedsQL), the Clinical Global Impression of Improvement (CGI-I) and the Assessment of Caregiver Experience with Neuromuscular Disease (ACEND).

PedsQL score

PedsQL is a modular self-report and parent proxy report approach to measuring HRQoL in children and adolescents 2-18 years of age. [redacted]

CGI-I

The CGI-I is a clinician reported outcome measuring patient's global functioning after initiating treatment and uses a seven point ordinal scale from 1 (very much improved) to 7 (very much worse). Table 22 presents the CGI-I assessments for both investigator and caregiver in the CHERISH study.

[redacted]

Table 22 CHERISH study CGI-I assessment (investigator and caregiver) at month 15 (reproduced from CS, Table 22)

Outcome	Investigator assessment		Caregiver assessment	
CGI assessment N (%)	Nusinersen (N=66)	Sham control (N=34)	Nusinersen (N=64)	Sham control (N=34)
Very much improved	[redacted]	[redacted]	[redacted]	[redacted]
Much improved	[redacted]	[redacted]	[redacted]	[redacted]
Minimally improved	[redacted]	[redacted]	[redacted]	[redacted]
No change	[redacted]	[redacted]	[redacted]	[redacted]
Minimally worse	[redacted]	[redacted]	[redacted]	[redacted]
Much worse	[redacted]	[redacted]	[redacted]	[redacted]
Very much worse	[redacted]	[redacted]	[redacted]	[redacted]

CGI-I - Clinical Global Impression of Improvement; N - number

ACEND

ACEND quantifies the caregiver burden experienced by parents of children affected by severe muscular diseases including children with SMA.

[redacted]

4.2.6 Subgroups

The decision problem set out in the final NICE scope[12] states that subgroups to be given consideration are based on severity of disease and should include the following:

- Age of SMA onset
- SMA type
- SMA genotype (including *SMN2* copy number).

In the ENDEAR study, treatment effects for key outcome measures were evaluated for two pre-specified subgroups as well as above and below median disease duration:

- disease duration at screening (≤12 weeks, >12 weeks)
- age at symptom onset (≤12 weeks, >12 weeks)

Table 23 presents ENDEAR subgroups by disease duration at screening; Table 24 presents ENDEAR subgroups by age at symptom onset. With regard to time to death or permanent ventilation in patients below the median disease duration, the hazard ratio (HR) was 0.24 (95% CI 0.10 to 0.58, $p<0.001$), whilst for those above the median disease duration, the HR was 0.84 (95% CI 0.43 to 1.67, $p=0.4$).[1] (see Figure 3 and Figure 4).

Table 23: ENDEAR subgroups analyses according to disease duration at screening (≤12 weeks, >12 weeks)

Outcome (source)	**≤ 12 weeks**			**>12 weeks**		
	Control (n=18)	**Nusinersen (n=34)**	***p*-value**	**Control (n=23)**	**Nusinersen (n=46)**	***p*-value**
HINE-2, % responders (CS, Appendix E, Figure 6)	0 (n=16)	75 (n=32)	$p<0.0001$	0 (n=21)	32 (n=41)	$p=0.0026$
CHOP INTEND, % improvement ≥ 4 points) (CS, Appendix E, Figure 7)	0 (n=16)	88 (n=32)	$p<0.0001$	5 (n=21)	59 (n=41)	$p<0.0001$
CHOP INTEND % worsening >= 4 points (CS, Appendix E, Figure 7)	50 (n=16)	0 (n=32)	NR	43 (n=21)	5 (n=41)	$p<0.0001$
OS (CS, Appendix E, Table 1 and Figure 8)	-	-	[redacted]	[redacted]	[redacted]	[redacted]
EFS (CS, Appendix E, Figure 9)	-	-	HR:0.158 ($p=0.0004$)	-	-	HR=0.816, $p=0.5325$

CHOP-INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; EFS - event-free survival; HINE-2 - Module 2 of the Hammersmith infant Neurological Examination: OS - overall survival

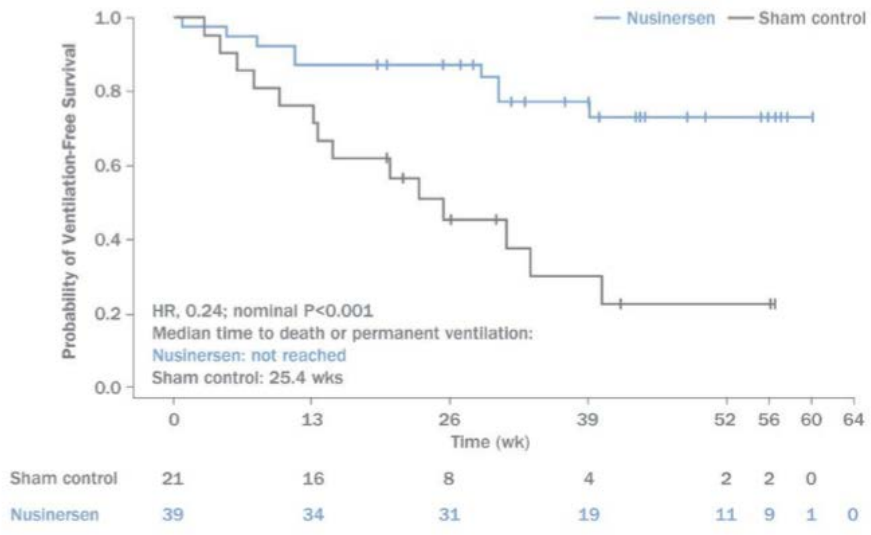
Table 24: ENDEAR subgroups analyses according to age at symptom onset (≤12 weeks, >12 weeks)

Outcome (source)	**≤ 12 weeks**			**>12 weeks**		
	Control (n=32)	**nusinersen (n=72)**	***p*-value**	**Control (n=8)**	**nusinersen (n=9)**	***p*-value**
OS (CS, Appendix E, Table 1)	-	-	HR:0.261 (95% CI: 0.1154-0.5919)	-	-	HR: 3.275 (95% CI: 0.509-21.3746)

OS - overall survival

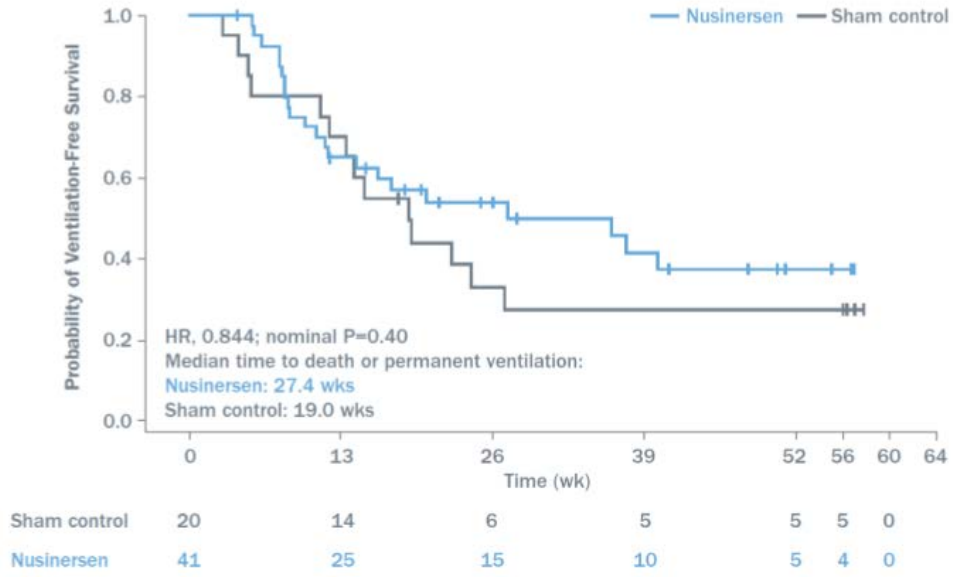
Overall, nusinersen demonstrated a benefit in all subgroups, apart from OS in the subgroup with age at onset of symptoms >12 weeks; however, the number of patients in this subgroup was small. For all outcomes, more pronounced treatment effects were observed for infants with disease duration ≤12 weeks at screening, however statistical tests for a difference between subgroups were not provided. In response to a request for clarification from the ERG[2] (question B6), the company provided results of Cox proportional hazards models and indicated that disease duration did not have a statistically significant effect on OS, while age of onset did have a statistically significant effect. Age was included as a continuous covariate and the company stated that "*it appears that survival in the sham arm is poor if age of onset is less than around 10 or 12 weeks, whereas survival on nusinersen may not be affected by age of onset.*" [2] Figure 3 and Figure 4 present Kaplan-Meier plots for time to death or permanent ventilation for subgroups by median disease duration at screening.

Figure 3: ENDEAR: Kaplan-Meier plots of time to death or permanent ventilation in the subgroup of infants below the median disease duration at screening (reproduced from CS, Figure 16)



HR - hazard ratio
Note: HR <1 indicates lower risk of event for the nusinersen group. The HR is calculated based on Cox regression adjusted for each infant's disease duration at screening; Source: Finkel 2017[20]

Figure 4: ENDEAR: Kaplan-Meier plots of time to death or permanent ventilation in the subgroup of infants above the median disease duration at screening (reproduced from CS, Figure 17)

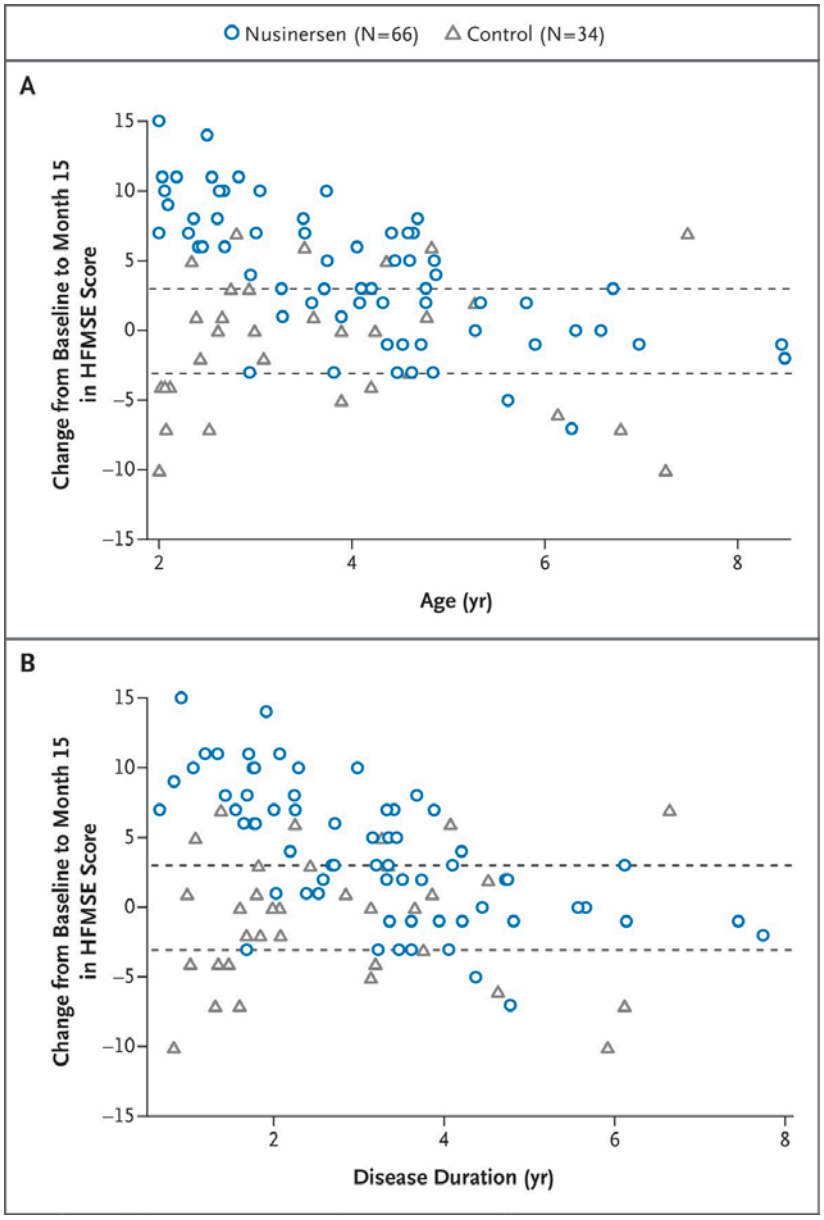


HR - hazard ratio
Note: HR <1 indicates lower risk of event for the nusinersen group. The HR is calculated based on Cox regression adjusted for each infant's disease duration at screening. Source: Finkel 2017[20]

Kaplan-Meier plots for OS and EFS by disease duration can be found in Appendix 2.

With regard to the CHERISH study, Figure 5 shows the change from baseline in HFMSE score according to age and disease duration. This illustrates that younger children who received treatment earlier in their disease course tended to have greater improvements. The treatment effects for each subgroup were not reported in the CS.

Figure 5: Change from baseline in total HFMSE score according to age (A) and disease duration (B) at screening (final analysis) (reproduced from CS, Figure 23)



HFMSE- Hammersmith Functional Motor Scale-Expanded; Disease duration is a child's age at screening minus the age at symptom onset. The analyses included children in the ITT population who did not have missing data for the 15-month assessment (66 in the nusinersen group and 34 in the control group). Dotted lines represent a ±3-point change in HFSME score, which is considered to be clinically meaningful. Source: Mercuri 2018[23]

The CS also included waterfall plots for HFMSE and RULM at 15 months (CS, Appendix E, Figures 10 and 11) again showing that younger children and those who received treatment earlier in their disease course tended to have greater improvements.

4.2.7 Safety and tolerability

Adverse events

The CS[1] presents an integrated safety analysis with data from eight completed or ongoing studies including a total of 260 patients (see Table 25). The studies with infantile onset patients included ENDEAR and CS3A; later onset studies included CHERISH and CS1, CS2, CS10 and CS12, while the pre-symptomatic group was from the NURTURE study only. Overall, the most commonly reported AEs in nusinersen-treated patients were either consistent with events occurring in the natural history of SMA, consistent with common conditions in the general population, consistent with common age-appropriate events or consistent with events observed in the context of lumbar puncture.[4]

Table 25: AEs from integrated safety analysis (reproduced from CS, Table 27)

N (%)	**Nusinersen-treated patients**				**Sham-control-treated patients**
	Infantile onset SMA	**Later onset SMA**	**Pre-symptomatic SMA**	**All nusinersen-treated patients**	
	ENDEAR & CS3A (N=100)	**CHERISH & CS1, 2, 10 & 12 (N=140)**	**NURTURE (N=20)**	**ENDEAR, CHERISH, NURTURE, CS1, 2, 3A, 10 & 12 (N=260)**	**ENDEAR & CHERISH (N=83)**
Summary of AEs					
AEs leading to discontinuation [a]	16 (16)	0 (0)	0 (0)	16 (6)	16 (19)
Treatment-related AEs	0 (0)	1 (<1)	0 (0)	1 (<1)	0 (0)
Common AEs No. of events No. of patients	 1,627 97 (97)	 1,187 134 (96)	 141 16 (80)	 2,955 247 (95)	 909 82 (99)
AEs by preferred term, with an incidence of >10% in nusinersen-treated patients					
Pyrexia	59 (59)	49 (35)	5 (25)	113 (43)	39 (47)
Upper respiratory tract infection	36 (36)	50 (36)	8 (40)	94 (36)	25 (30)
Nasopharyngitis	21 (21)	33 (24)	4 (20)	58 (22)	15 (18)
Vomiting	22 (22)	33 (24)	0 (0)	55 (21)	8 (10)
Headache	0 (0)	51 (36)	0 (0)	52 (20)	0 (0)
Constipation	37 (37)	0 (0)	2 (10)	50 (19)	14 (17)
Back pain	0 (0)	44 (31)	0 (0)	45 (17)	0 (0)
Cough	15 (15)	26 (19)	3 (15)	44 (17)	17 (20)
Pneumonia	30 (30)	0 (0)	2 (10)	41 (16)	14 (17)
Respiratory distress	28 (28)	0 (0)	0 (0)	31 (12)	12 (14)
Scoliosis	11 (11)	18 (13)	0 (0)	29 (11)	0 (0)
Diarrhoea	16 (16)	0 (0)	0 (0)	27 (10)	7 (8)
Respiratory failure	26 (26)	0 (0)	0 (0)	27 (10)	16 (19)
Post-lumbar puncture syndrome	0 (0)	26 (19)	0 (0)	26 (10)	0 (0)

AE - adverse event; SMA - spinal muscular atrophy; [a] All AEs leading to study discontinuation were events with fatal outcomes; Source: Mercuri et al[28]

In the integrated safety analysis, both nusinersen-treated patients and control patients experienced AEs. The most commonly reported AEs were those expected in patients with SMA or after lumbar puncture, such as headache, vomiting, back pain and post-lumbar puncture syndrome. Other common AEs occurring in ≥20% patients were (nusinersen versus control): pyrexia (43% vs 47%), upper respiratory infections (36% vs 30%) and nasopharyngitis (22% vs 18%). NURTURE (in pre-symptomatic infants) reported fewer AEs compared with symptomatic infants as would be expected with their healthier baseline condition.[4]

Within the ENDEAR study (CS,[1] Appendix F, Table 2), the incidence of AEs in nusinersen group and the control group was similar. However, the following AEs occurred more frequently in the nusinersen group (n=80) than in the control group (n=41): constipation (35% vs 22%), upper respiratory infection (30% vs 22%) and pneumonia (29% vs 17%). With regard to the CHERISH study (CS,[1] Appendix F, Table 3), again the incidence of AEs was similar in the nusinersen and control groups, except for the following AEs which occurred more frequently in the nusinersen group (n=84) than the control group (n=42): headache (29% vs 7%), vomiting (29% vs 12%), back pain (25% vs 0%) and epistaxis (7% vs 0%).

Serious adverse events and death

Serious adverse events (SAEs) and death were also presented from the integrated safety analysis including the same studies as described above (see Table 26).

Table 26: SAEs and death summary from integrated safety analysis (reproduced from CS, Table 28)

N (%)	Nusinersen-treated patients				Sham-control-treated patients
	Infantile onset SMA	Later onset SMA	Pre-symptomatic SMA	All nusinersen-treated patients	
	ENDEAR & CS3A (N=100)	CHERISH & CS1, 2, 10 & 12 (N=140)	NURTURE (N=20)	ENDEAR, CHERISH, NURTURE, CS1, 2, 3A, 10 & 12 (N=260)	ENDEAR & CHERISH (N=83)
Patient death	17 (17)	0 (0)	0 (0)	17 (7)	16 (19)
Incidence of SAEs	77 (77)	19 (14)	6 (30)	102 (39)	50 (60)
SAEs					
Respiratory, thoracic, and mediastinal disorders	63 (63)	4 (3)	2 (10)	69 (27)	33 (40)
Infections and infestations	60 (60)	13 (9)	4 (20)	77 (30)	29 (35)
Cardiac disorders	12 (12)	0 (0)	0 (0)	12 (5)	7 (8)
Metabolism and nutrition disorders	10 (10)	0 (0)	2 (10)	12 (5)	7 (8)
Gastrointestinal disorders	7 (7)	1 (<1)	1 (5)	9 (3)	7 (8)
General disorders and administrative site conditions	7 (7)	1 (<1)	1 (5)	9 (3)	1 (1)
Injury, poisoning, and procedural complications	3 (3)	3 (2)	0 (0)	6 (2)	3 (4)
Investigations	3 (3)	0 (0)	0 (0)	3 (1)	3 (4)
Nervous system disorders	3 (3)	0 (0)	0 (0)	3 (1)	0 (0)
Vascular disorders	2 (2)	0 (0)	0 (0)	2 (<1)	0 (0)
Immune system disorders	0 (0)	1 (<1)	0 (0)	1 (<1)	-
Musculoskeletal and connective tissue disorders	1 (1)	0 (0)	0 (0)	1 (<1)	-
Skin and subcutaneous tissue disorders	1 (1)	0 (0)	0 (0)	1 (<1)	0 (0)

SAE - serious adverse event

Overall, there were fewer deaths in the nusinersen-treated patients compared with the control patients (19% vs 7%) and fewer SAEs in the nusinersen patients compared with the control patients (39% vs 60%). Common SAEs affecting >20% of patients were respiratory, thoracic and mediastinal disorders (27% for nusinersen patients and 40% for control) and infections and infestations (30% in nusinersen

patients and 35% in control). The SAEs reported are consistent with those expected in patients with SMA and are not necessarily related to treatment.

Within the ENDEAR study (CS,[1] Appendix F, Table 2), there were fewer SAEs in the nusinersen group than the control group; however, specific AEs that occurred more frequently in the nusinersen group (n=80) than the control group (n=41) were: respiratory distress (26% vs 20%); pneumonia (24% vs 12%) and atelectasis (18% vs 10%). Within the CHERISH study (CS,[1] Appendix F, Table 3), all reported SAEs were higher in the control group than in the nusinersen group.

Additional safety issues

In the post-marketing setting, cases of meningitis have been noted following the administration of nusinersen, although numbers were not reported in the CS. [redacted] The US Food and Drug Administration (FDA) medical review for nusinersen[29] highlights some AEs with potentially severe consequences and recommends warnings regarding the risk of thrombocytopenia, coagulation abnormalities, renal toxicity, hyponatremia, decreased growth, rash and possible vasculitis, and hepatotoxicity.

4.3 Conclusions of the clinical effectiveness section

4.3.1 Completeness of the CS with regard to relevant clinical studies and relevant data within those studies

The CS[1] did not contain a systematic review as would be expected within a submission to the NICE STA process. As such, it is not entirely certain that all nusinersen studies have been identified, although the ERG is confident that all relevant studies of nusinersen for SMA have been included in the CS. In addition, a systematic review of studies relating to BSC, listed as the comparator in the NICE scope,[12] was not presented in the CS.

The studies included in the CS were well presented and included studies in three patient groups: (i) early onset (ENDEAR); (ii) late onset (CHERISH) and (iii) pre-symptomatic SMA (NURTURE), with ENDEAR and CHERISH being the key studies.

4.3.2 Interpretation of treatment effects reported in the CS in relation to relevant population, interventions, comparator and outcomes

The ERG is content that the relevant populations and intervention have been included in the CS, that is, infantile and later onset patients treated with nusinersen. However, the appropriate comparator, BSC

was not included. All relevant outcomes were included in the CS, apart from complications, stamina and fatigue.

In the ENDEAR study, the primary outcome measures were the proportion of motor milestone responders (HINE-2) and EFS (defined as time to death or permanent ventilation). With regard to HINE-2, a significantly greater percentage of patients in the nusinersen group achieved motor milestone responses than the control group. The proportion of HINE-2 responders in the interim analysis was 41% in the nusinersen group and 0% in the control group [difference: 41.18% (95% CI 18.6% to 61.20%, $p<0.001$). In the final efficacy set, the proportion of HINE-2 responders was 51% in the nusinersen group compared with 0% in the control group (difference=50.68%; 95% CI 31.81% to 66.48%, $p<0.0001$), although many patients in the nusinersen group (49% in final efficacy set) could not be classified as responders. For EFS (ITT analysis set), there was a statistically significant increase for the nusinersen group compared with the sham control group (HR=0.53; 95% CI 0.32 to 0.89; $p=0.005$).

In the CHERISH study, the primary outcome measure was motor function as measured by HFMSE. The change in HFMSE from baseline was significant in both the interim analysis (LSM change difference=5.9; 95% CI 3.7 to 8.1; $p<0.001$) and in the final efficacy set analysis (LSM change difference=4.9; 95% CI 3.1 to 6.7; $p=0.0000001$) for the nusinersen group compared with the control group.

The company's integrated safety analysis showed that both nusinersen-treated patients and control patients experienced AEs. The most commonly reported AEs were those expected in patients with SMA or after lumbar puncture, such as headache, vomiting, back pain and post-lumbar puncture syndrome. Overall, there were fewer deaths in the nusinersen treated patients compared to the control patients (19% vs 7%) and fewer SAEs in the nusinersen patients compared with the control patients (39% vs 60%).

Nusinersen appears to provide significant clinical benefit to patients and the safety profile reported in the studies was acceptable and generally more favourable than that for the sham control group. The patient groups in the study arms for the ENDEAR and CHERISH studies were broadly similar although the nusinersen groups had more severe symptoms and longer duration of treatment.

4.3.3 Uncertainties surrounding the reliability of the clinical effectiveness evidence

There are several areas of uncertainty in the clinical evidence. The dosage of nusinersen in the CHERISH study was different from the licensed dose in that the number of and timings for loading dose days were different as well as the timing for the maintenance dose (every 4 months for ENDEAR and every 6 months for CHERISH).

The use of three different analysis sets in both the ENDEAR and CHERISH studies made it difficult to interpret the study findings. Although, most outcomes listed in the NICE scope[12] were presented, with the exception of complications and stamina and fatigue, some outcomes were presented in only one study. Information on subgroups as set out in the NICE scope was provided although the data were limited.

The follow-up period in the studies was relatively short and no data were provided for patients in the post-marketing setting. The lack of long-term data means that it is unknown whether the effect size will change as the disease progresses and patients grow older. There is also a lack of data on the need for dose adjustments as patients grow older.

There are multiple phenotypes for SMA and no data were presented for patients with inborn symptoms (Type 0) or mild, adult onset SMA (Type IV). There is no information about how decisions should be made regarding treatment taking into account disease severity, duration and progression along with patient benefit. In addition, there is no information on the optimal dose of treatment. It is unclear when untreated pre-symptomatic patients, who are genetically diagnosed, would develop symptoms or how severe symptoms would be. Therefore, decisions regarding treatment are challenging in this patient group.

4.4 Additional work undertaken by ERG

No additional work on the clinical effectiveness section was undertaken by the ERG.

5. COST EFFECTIVENESS

This chapter presents a summary and critique of the company's health economic analyses of nusinersen for the treatment of early onset and later onset SMA. Section 5.1 presents a summary and critique of the results of the company's review of existing cost-effectiveness analyses. Section 5.2 summarises the scope of the company's *de novo* health economic analyses. Sections 5.3 and 5.4 detail the methods and results of the company's early onset and later onset models, respectively. Section 5.5 presents a critique of both health economic analyses. Section 5.6 presents the results of exploratory analyses undertaken by the ERG. Section 5.7 presents a discussion of the available economic evidence.

5.1 Company's review of published cost-effectiveness studies

The company conducted a combined search to identify studies of cost-effectiveness, HRQoL and resource use in relation to SMA (CS,[1] Appendices G, H and I; Sections 5, 6 and 7). The company's searches did not identify any economic evaluations of treatments for SMA.

During the clarification process[2] (question B1), the ERG queried the origin of the search filters which had been used to identify studies of each type. In their response, the company clarified that whilst they had not used any validated filters (e.g. those developed by McMaster University https://hiru.mcmaster.ca/hiru/HIRU_Hedges_MEDLINE_Strategies.aspx#Costs), the searches had been developed with input from an information specialist and terms were either based on the MeSH controlled vocabulary or drawn from previously published guidelines. The ERG notes that there is some overlap between the search terms used to identify economic evaluations and those for cost and resource use studies (CS,[1] Appendix G, Section 5.2.3), however, the ERG considers those included to be broadly fit for purpose.

With respect to the company's review of HRQoL studies, the search terms were grouped into two different sets - "utility studies" (including specific measures such as SF-36 and EQ-5D) and "human burden" in which broader terms such as "QoL" and "HRQL" were included. Most of the essential terms the ERG would expect to see in an HRQoL filter were included, although some published filters (e.g., the filter produced by the Canadian Agency for Drugs and Technologies in Health [CADTH] www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#health) include a wider variety of measures. Unusually, the "human burden" terms were searched for only in article titles; the ERG notes that it is more conventional to search for terms in multiple fields such as abstracts and index terms. However, independent searches undertaken by the ERG did not identify any further published studies reporting on EQ-5D utilities in patients with SMA.

5.2 Model scope – early onset and later onset models

As part of its submission to NICE,[1] the company submitted two fully executable health economic models programmed in Microsoft Excel®. The scope of the company's health economic analyses is summarised in Table 27. The company's models assess the cost-effectiveness of nusinersen versus "real-world care" (hereafter referred to as usual care) in two populations: (i) patients with early onset (Type 1) SMA, based on the ENDEAR study,[14] and (ii) patients with later onset (Type II and III) SMA, based on the CHERISH study.[15] Both models evaluate the incremental health gains, costs and cost-effectiveness of nusinersen versus usual care from the perspective of the UK NHS and Personal Social Services (PSS). The early onset model adopts a 60-year time horizon, whilst the later onset model adopts an 80-year time horizon (see footnotes to Table 27). Cost-effectiveness is expressed in terms of the incremental cost per quality-adjusted life year (QALY) gained. All health outcomes and costs are discounted at a rate of 3.5% per annum. Unit costs are valued at 2015/16 prices.

Table 27: Scope of company's health economic analyses – early and later onset models

	Early onset model	Later onset model
Population	ITT population of the ENDEAR study[14] (Type I SMA). Mean starting age=5.58 months (0.47 years)	ITT population of the CHERISH study[15] (Types II and III SMA). Mean starting age=43.71 months (3.64 years)
Time horizon	60 years*	80 years
Intervention	Nusinersen	
Comparator	Usual care	
Outcome	Incremental cost per QALY gained	
Perspective	NHS and PSS	
Discount rate	3.5%	
Price year	2015/2016	

** Whilst the CS states that a 40-year time horizon was adopted for the early onset model, all results presented for this population in the CS relate to a 60-year time horizon. The company's clarification response[2] confirms that a 60-year time horizon was intended*

5.2.1 Population

Early onset model

The population within the early onset model (Type I SMA) reflects the ITT population enrolled into the ENDEAR study.[14] The mean age of the cohort at baseline in ENDEAR was 5.58 months (0.47 years); this is taken as the patient start age within the model. The initial distributions of patients within the modelled intervention and control groups are defined according to baseline HINE-2 scores for the nusinersen and sham groups within ENDEAR, respectively (note – the ERG considers the use of treatment-specific initial distributions to reflect an error, see Section 5.5).

Later onset model

The population within the later onset model (Type II and III SMA) reflects the ITT population enrolled into the CHERISH study.[15] The mean age of the cohort at baseline in CHERISH was 43.71 months

(3.64 years); this is taken as the patient start age within the model. The initial distributions of patients within the intervention and control groups are defined according to baseline HFMSE scores for the nusinersen and sham groups within CHERISH, respectively (note – again, the ERG considers the use of treatment-specific initial distributions to reflect an error, see Section 5.5).

The licensed indication for nusinersen is for the treatment of 5q SMA.[4] As discussed in Chapter 3, the company's economic analyses do not include patients with Type 0 or Type IV SMA; as such, the populations captured within the company's early onset and later onset models are narrower than the marketing authorisation for nusinersen. Despite this absence of evidence, the CS[1] states that the anticipated place of nusinersen is as a first-line treatment for all SMA patients as soon as possible after diagnosis.

5.2.2 Intervention

The intervention within both the early onset and later onset models is nusinersen administered as an intrathecal bolus injection via lumbar puncture.

Early onset model

Within the early onset model, nusinersen is assumed to be given as four loading doses on days 0, 14, 28 and 63, followed by one maintenance dose every four months thereafter. Each loading/maintenance dose is assumed to consist of 12mg nusinersen. This is based on the treatment schedule within the ENDEAR study[14] and is consistent with the marketing authorisation for nusinersen.[4] The company's model assumes that nusinersen will be discontinued either if the patient has achieved no motor milestones (or all milestones previously achieved are lost) by the end of month 13 (the end of study follow-up within ENDEAR) or if the patient undergoes scoliosis surgery and cannot subsequently undergo lumbar puncture.[1]

Later onset model

Within the later onset model, nusinersen is assumed to be given as four loading doses on days 1, 30, 60 and 90, followed by one maintenance dose every four months thereafter. Each loading/maintenance dose is assumed to consist of 12mg nusinersen. This treatment schedule differs from that used in the CHERISH study,[15] whereby loading doses were administered on days 1, 29 and 85, with subsequent maintenance doses on day 274 and every 6 months thereafter. Both the modelled treatment schedule in the later onset model and the treatment schedule applied in the CHERISH study differ from the dosing regimen specified in the marketing authorisation[4] (as detailed above). With reference to this issue, the CS states that *"as the use of the modelled dosing regimen could lead to greater benefit in clinical practice, the modelled results may represent a conservative estimate of treatment effect"* (CS,[1] page 167). The company's model assumes that nusinersen will be discontinued either if the patient has

achieved no milestones (or all milestones previously achieved are lost) by the end of month 15 (the end of study follow-up within CHERISH) or if patient undergoes scoliosis surgery and cannot subsequently undergo lumbar puncture.

Comparator

Within both the early onset and later onset SMA models, the comparator is assumed to be usual care; this includes respiratory, nutritional, gastrointestinal and orthopaedic interventions.[1]

5.3 Early onset model – methods and results

5.3.1 Model structure and logic – early onset model

The company's early onset model adopts a state transition approach, based on health states defined according to the HINE-2 instrument[30] (see Figure 6). The early onset model includes eight health states: (i) No milestones achieved; (ii) Mild milestones achieved; (iii) Moderate milestones achieved; (iv) Sits without support; (v) Stands with assistance; (vi) Walks with assistance; (vii) Stands/walks unaided and (viii) Dead. The HINE-2 scoring system is presented in Appendix 2; the company's classification of HINE-2 health states according to this scoring system is summarised in Table 28).

Figure 6: Company's early onset model structure (reproduced from CS, Figure 31)

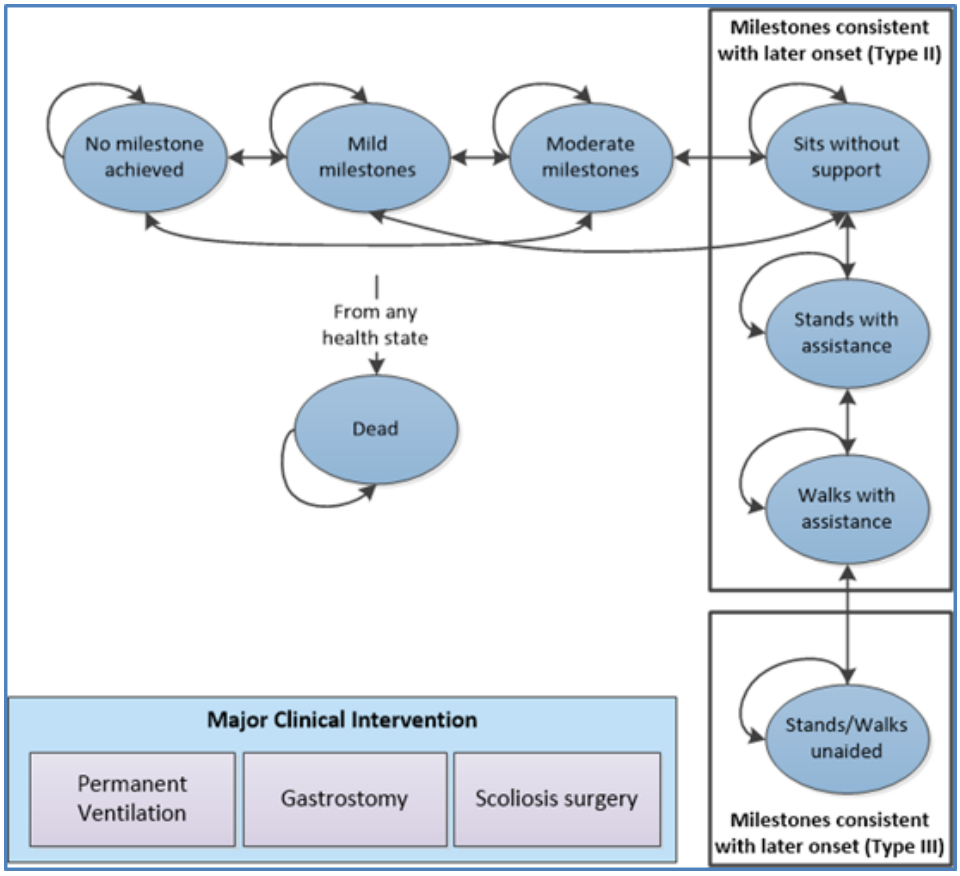


Table 28: Early onset model health states according to HINE-2 scoring (adapted from CS, Figure 31 footnotes and the company's clarification response)

Model health state	HINE-2 criteria for model health state
(i) No milestones	Patients have a score of 0 in all HINE-2 items. Voluntary grasp any score
(ii) Mild milestones	Patients have a score of 1 in at least one of the following items: head control, ability to kick, or crawling. Patients have a score of 0 in other items. Voluntary grasp any score.
(iii) Moderate milestones	Patients have any of the following scores in at least one of the following items: head control = 2; sitting = 1; ability to kick = 2 or 3; rolling = 1 or 2; crawling = 2; standing = 1; walking = 1.* Voluntary grasp any score.
(iv) Sits without support	Patients have a score of 2 or 3 or 4* in sitting ability and a score <2 in standing ability. Any score in other items except walking.
(v) Stands with assistance	Patients have a score of 2 in standing ability. Any score in other items except walking.
(vi) Walks with assistance	Patients have a score of 2 in walking. Any score in other items.
(vii) Stands/walks unaided	Patients have a score of 3 either in standing or walking ability. Any score in other items.

** Corrected by the company following clarification[2] (question B2)*

Model logic

The logic of the company's early onset model is described in the following sections.

Nusinersen group

Patients enter the model according to the observed baseline HINE-2 health state distribution for the nusinersen group in the ENDEAR study.[14] During the first four model cycles (up to the end of month 13), mortality risk is modelled using the predicted cumulative survival probabilities derived from a 1-knot Royston-Parmar spline model fitted to the observed survival data for the nusinersen group in ENDEAR. From model entry until the end of month 13, transitions between the seven HINE-2-based health states are governed by four cycle-specific transition matrices derived from observed count data within ENDEAR; these transition probabilities are then adjusted (normalised) during each cycle to account for the error between the predicted mortality probability from the spline model and the observed mortality probability within the nusinersen group of ENDEAR. From month 14 to the end of month 58, mortality is modelled using an exponential function estimated using survival outcomes for patients receiving non-invasive respiratory aid (NRA) from a retrospective chart review study reported by Gregoretti *et al;*[31] these data are adjusted to match the age of the ENDEAR population (see Section 5.3.3).

Mortality risk in all subsequent model cycles is based on an HR-adjusted Gompertz function fitted to general population mortality data[32] (HR=5,184.81). This time-dependent 3-step Type I SMA mortality function (1-knot spline [ENDEAR]→exponential [Gregoretti]→HR-adjusted Gompertz [general population, HR=5,184.81]) is applied to all patients in the three worst three health states (State [i] No milestones, State [ii] Mild milestones and State [iii] Moderate milestones). Beyond the end of follow-

up in ENDEAR, mortality risk for patients in the four best health states (State [iv] Sits without support, State [v] Stands with assistance, State [vi] Walks with assistance and State [vii] Stands/walks unaided) is adjusted to reflect an assumption of improved survival associated with Type II SMA, based on a 2-knot Royston-Parmar spline model fitted to data from an observational study reported by Zerres *et al*.[33] Beyond the end of follow-up within Zerres *et al*[33] (the end of month 622), this improved mortality for Type II SMA is modelled using a separate HR-adjusted Gompertz function fitted to general population mortality data (HR=26.41). Within these better health states, patients are allocated 90% of the mortality risk from the Type II mortality model (2-knot spline [Zerres]→HR-adjusted Gompertz [general population, HR=26.41]) and 10% of the mortality risk from the Type I mortality model (1-knot spline [ENDEAR]→exponential [Gregoretti]→HR-adjusted Gompertz [general population, HR=5,184.81]). From the end of month 13 onwards, all health state transitions are governed by two transition matrices: one corresponding to the cycle from the end of month 13 to the end of month 14, and one corresponding to all subsequent 4-monthly cycles. These two transition matrices were estimated using CHOP INTEND scores observed within the ENDEAR trial[14] and Study CS3A;[34] both matrices permit nusinersen-treated patients to either remain in their current state or to move to the next best health state; they do not allow for the deterioration of any patient's motor function from this timepoint onwards.

Patients are assumed to discontinue nusinersen if they do not achieve any milestones by the end of month 13 or if they undergo scoliosis surgery (at year 12 for non-ambulatory patients and at year 15 for ambulatory patients) and cannot subsequently undergo administration of nusinersen via lumbar puncture. Patients who discontinue nusinersen due to lack of efficacy are assumed to remain State (i) (No milestones) until death. Patients who discontinue nusinersen following scoliosis surgery are assumed to subsequently follow the final transition matrix for the sham group.[3]

Usual care group

Patients enter the model based on the baseline HINE-2 health state distribution for the sham group in the ENDEAR study.[14] During the first four cycles (up to the end of month 13), mortality risk is modelled using the predicted cumulative survival probabilities from a 1-knot Royston-Parmar spline model fitted to the survival data for the sham group within ENDEAR. Transitions between the seven HINE-2-based health states are governed by four cycle-specific matrices derived from observed count data within ENDEAR; these transition probabilities are then adjusted (normalised) during each cycle to account for the error between the predicted mortality probability from the spline model and the observed mortality probability within the sham group of ENDEAR.

From month 14 onwards, mortality is modelled using the same exponential (Gregoretti *et al*[31]) and HR-adjusted general population Gompertz (HR=5,184.81) function as that used in the nusinersen group (see above). In contrast to the assumptions applied to the nusinersen group, no mortality adjustment is

applied for patients in States (iv) to (vii) in the usual care group. After the end of month 13, all health state transitions are governed by two transition matrices: one corresponding to the cycle from the end of month 13 to the end of month 14, and one corresponding to all subsequent 4-monthly cycles. These matrices were estimated using CHOP INTEND scores observed within the ENDEAR trial[14] and Study CS3A;[34] these matrices permit patients on usual care to either remain in their current state or move to the next worst health state; they do not allow for the improvement of any patient's motor function from this timepoint onwards. A proportion of patients are assumed to undergo scoliosis surgery at year 10 if non-ambulant and at year 15 if ambulant; however, this event does not impact on the patient's health state occupancy, HRQoL or costs.

Estimation of health outcomes, costs and cost-effectiveness

Separate utilities are applied to each modelled health state. QALYs accrued by patients in each group are estimated by applying a vector of health utilities to the probability of being in each state during each model cycle. QALY losses for caregivers are estimated conditional on the patient's health state and include a QALY loss for bereavement on carers. The CS reports separate analyses including/excluding caregiver QALYs.

The early onset model includes the following cost components: (i) acquisition and administration costs for nusinersen; (ii) health state costs, including respiratory, gastrointestinal, nutritional and orthopaedic care (conditional on motor milestones achieved) and (iii) end-of-life care (applied as a once-only cost at the point of death).

Incremental cost-effectiveness is calculated in a pairwise fashion based on the difference in costs divided by the difference in QALYs for nusinersen and usual care.

5.3.2 Structural assumptions – early onset model

The company's early onset SMA model makes the following assumptions:

(i) Treatment using nusinersen will be discontinued if no milestones are achieved after 13 months (see clarification response,[2] question B27). The ERG notes that this assumption is applied only once, as patients receiving nusinersen are assumed never to transit to State (i) No milestones after this timepoint (see assumption [iv] below).

(ii) A proportion of patients discontinue nusinersen following scoliosis surgery.

(iii) After the end of month 13, an adjustment is applied to reflect improved survival for nusinersen patients in State (iv) Sits without support, State (v) Stands with assistance, State (vi) Walks with assistance and State (vii) Stands/walks unaided. These patients are allocated 90% of the mortality risk for Type II SMA and 10% of the mortality risk for Type I SMA. This adjustment

is not applied to patients reaching these states in the usual care group; instead, all patients in the usual care group are allocated 100% of the Type I mortality risk.

(iv) After the end of month 13, patients receiving nusinersen are assumed never to transit to a worse health state; rather, during any model cycle, patients can either remain in their current health state or transit to the next best health state. Beyond month 13, transition probabilities are based on the mean rate of improvement in CHOP INTEND score within ENDEAR and the mean CHOP INTEND scores within each HINE-2 state for the nusinersen group over the course of the ENDEAR study (supplemented using data from Study CS3A for State [v] Stands with assistance and State [vi] Walks with assistance). The rate of improvement in CHOP INTEND score is assumed to be constant with respect to time and monotonic across health states.

(v) After month 13, patients receiving usual care are assumed never to transit to an improved health state; rather, during any model cycle, patients can either remain in their current health state or transit to the next worst health state. Beyond month 13, transition probabilities are based on the mean rate of worsening in CHOP INTEND score within ENDEAR and the mean CHOP INTEND scores within each model health state for the sham group over the course of the ENDEAR study (supplemented using data from Study CS3A for State [v] Stands with assistance and State [vi] Walks with assistance). The rate of decline in CHOP INTEND score is assumed to be constant with respect to time and monotonic across health states.

(vi) After month 13, the probability of transiting from State (vi) Walks with assistance to State (vii) Walks/stands unaided is assumed to be the same as the probability of transiting from State (v) Stands with assistance to State (vi) Walks with assistance. The company considers this to be a conservative assumption.[2]

(vii) A proportion of ambulant patients undergo scoliosis surgery after 15 years.

(viii) The CS[1] states that the model assumes that a proportion of non-ambulant patients undergo scoliosis surgery at 12 years. Whilst this assumption is correctly implemented in the nusinersen group of the company's model, the model assumes that scoliosis surgery for non-ambulant patients receiving usual care occurs at 10 years. As separate costs and utility changes for scoliosis surgery are not included in the model, this does not impact on the model results.

(ix) Treatment costs are grouped according to milestones consistent with Type I SMA (State [i] No milestones, State [ii] Mild milestones and State [iii] Moderate milestones), Type II SMA (State [iv] Sits without support, State [v] Stands with assistance and State [vi] Walks with assistance) and Type III SMA (State [vii] Stands/walks unaided).

(x) The model does not include additional HRQoL impacts or costs associated with AEs. The CS[1] notes that no treatment-related AEs were observed in ENDEAR.

5.3.3 Evidence used to inform model parameters – early onset model

The main groups of parameters for the early onset model and the evidence used to inform these are summarised in Table 29. These are discussed in further detail in the subsequent sections.

Table 29: Evidence used to inform the company's early onset model

Parameter group	Evidence source
Initial HINE-2 health state distribution - nusinersen	Observed initial HINE-2 distribution in the nusinersen group of ENDEAR[14]
Initial HINE-2 health state distribution – usual care	Observed initial HINE-2 distribution in the sham group of ENDEAR[14]
Overall survival – nusinersen	1-knot Royston-Parmar spline model fitted to nusinersen group data from ENDEAR[14] switching (after month 13) to an exponential model fitted to adjusted NRA group data from Gregoretti *et al*[31] switching (after month 58) to an HR-adjusted general population Gompertz model (HR=5,184.81).[32] For States (iv) to (vii), after month 13, an adjustment of 0.90 is applied to reflect improved survival for Type II SMA based on a 2-knot spline fitted to data reported by Zerres *et al*[33] switching (after month 622) to an HR-adjusted general population Gompertz model (HR=26.41).[32]
Overall survival – usual care	1-knot Royston-Parmar spline model fitted to sham group of ENDEAR[14] switching (after month 13) to an exponential model fitted to adjusted NRA group data from Gregoretti *et al* switching (after month 58) to an HR-adjusted general population Gompertz model (HR=5,184.81). No adjustment is applied to reflect Type II SMA outcomes.
Transition probabilities – nusinersen (up to month 13)	Observed HINE-2 count data from ENDEAR[14] (without imputation)
Transition probabilities – usual care (up to month 13)	Observed HINE-2 count data from ENDEAR[14] (without imputation)
Transition probabilities – nusinersen (month 14 onwards)	Estimated mean rate of improvement in CHOP INTEND and mean CHOP INTEND scores by HINE-2 state in ENDEAR[14] for nusinersen group (supplemented using data from Study CS3A[34])
Transition probabilities – usual care (month 14 onwards)	Estimated mean rate of worsening in CHOP INTEND and mean CHOP INTEND scores by HINE-2 state in ENDEAR[14] for sham group (supplemented using data from Study CS3A[34])
Probability of undergoing surgery for scoliosis and age at time of surgery	Surgery probability based on assumption.[1] Timing of surgery loosely based on Haaker and Fujak[35]
Probability of discontinuing nusinersen after surgery for scoliosis	Assumption[1]
Patient utilities	PedsQL data collected in CHERISH[15] mapped to the EQ-5D using a published algorithm reported by Khan *et al*[36]
Baseline caregiver utility	Baseline caregiver utility based on Bastida *et al.*[37] Caregiver disutilities by health state estimated using Ara and Brazier[38] and mapped patient utilities from CHERISH. [15]
Nusinersen acquisition cost	CS[1]
Nusinersen administration costs	NHS Reference Costs 2015/16[39]
Health state costs	Bastida *et al*[37]
End-of-life care costs	NICE Guideline 61[40]

HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; HR – hazard ratio; CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; NRA - Non-invasive respiratory aid; EQ-5D – Euroqol 5-Dimensions; CS – company's submission

Overall survival – early onset SMA

Company's methods for estimating overall survival

As outlined in Section 5.3, OS is modelled using a piecewise approach with separate sources to inform different sections of the modelled time horizon. Extrapolation on the basis of external data was considered by the company to be "*more appropriate than extrapolating the survival models fitted to the observed trial period alone*" (CS,[1] page 122). The overall modelling approach is summarised in Table 30; further details of the external data and the modelling approach are provided below.

Table 30: Summary of survival models applied for extrapolation of OS

Survival interval	Treatment group		
	Usual care (N=41)	Nusinersen (N=80)	
		States (i) to (iii)	States (iv) to (vii)
OS time period 1 Month 0 to Month 13	ENDEAR sham arm* **1-knot spline combined model**	ENDEAR nusinersen arm† **1-knot spline combined model**	
OS time period 2 Month 14 to Month 58	Adjusted Gregoretti *et al*[31] NRA‡ **Exponential model**	Adjusted Gregoretti *et al*[31] NRA‡ **Exponential model**	Adjusted Gregoretti *et al*[31]‡ and Zerres *et al*[33] § **Gregoretti exponential (weight 0.1) Zerres 2-knot spline (weight 0.9)**
OS time period 3a Month 59 to Month 622	UK general population mortality data **HR-adjusted Gompertz (HR=5184.8)**	UK general population mortality data **HR-adjusted Gompertz (HR=5184.8)**	UK general population mortality and Zerres *et al*[33]§ **HR-adjusted Gompertz (HR=5184.8, weight 0.1) Zerres 2-knot spline (weight 0.9)**
OS time period 3b Month 623 to Month 720			UK general population mortality data **HR-adjusted Gompertz (HR=5184.8, weight 0.1) and HR-adjusted Gompertz (HR=26.4, weight 0.9)**

** Observed trial data, N=41*
† observed trial data, N=80
‡ N= 26, Type I SMA replicated from KM, adjusted
§ N=240, Type II SMA replicated from KM, no adjustment

Survival models

The company considered a range of common parametric models: exponential, Weibull, Gompertz, log normal, log logistic, generalised gamma and Royston-Parmar cubic splines fitted on the hazard scale with 1, 2, and 3 knots (described in the CS[1] as flexible spline-based Weibull models). Hybrid survival models were also considered for some situations but were not found to be appropriate (see CS,[1] Appendix P). For extrapolation based on UK general population mortality data, the company considered only the Weibull, Gompertz and Royston-Parmar cubic spline models.

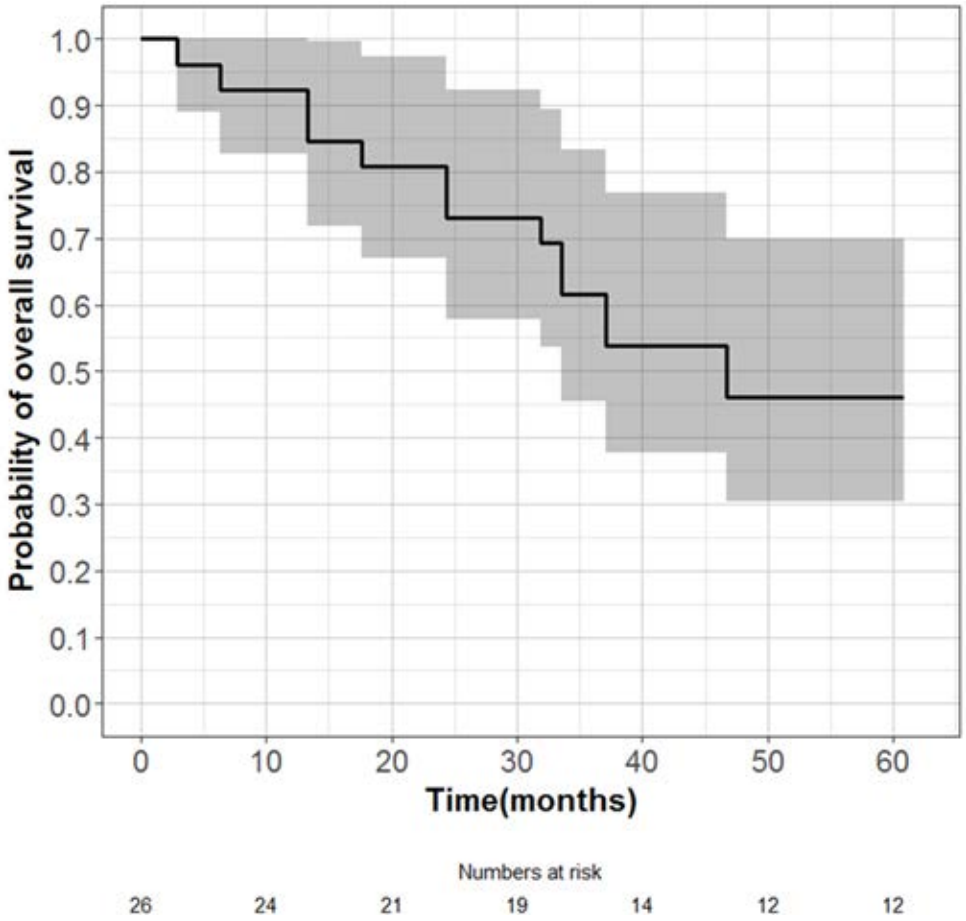
For the ENDEAR data,[14] two approaches were considered to account for differential survival probabilities in the nusinersen and sham arms: (i) a *combined model* with treatment group included as a covariate (described as *unstratified models* in the CS) and (ii) *stratified models* whereby all parameters are allowed to differ by treatment. The latter approach is equivalent to fitting separate models to each treatment group.

Models were fitted in R[41] using either the *eha* package (exponential, Weibull, log normal and log logistic models) or the *flexsurv* package (Gompertz, generalised gamma and Royston-Parmar spline models). Complementary log-log plots were produced to assess the proportional hazards assumption, and smoothed non-parametric estimates of the observed hazard rates were produced. Fit of the models was considered based on the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the Integrated Brier Score (IBS)[42] through bootstrap cross-validation., together with visual inspection of fit and consideration of the clinical plausibility of the extrapolated portion of the survival curves.

External data sources

The study reported by Gregoretti *et al*[31] is a retrospective chart review of 194 infantile onset SMA patients followed by 4 Italian centres between October 1, 1992 and December 31, 2010. Subgroup data on the 31 infants receiving non-invasive respiratory muscle aid (NRA) were deemed by the company to be the most reflective of current standard care. Individual patient-level data (IPD) were reconstructed from the published Kaplan-Meier curve. Gregoretti *et al* present data from birth whereas the mean age of patients at the start of treatment in ENDEAR[14] was 5.56 months (note – a slightly higher value of 5.58 months is assumed in the company's model). The company adjusted the reconstructed IPD by subtracting 5.56 from all event times, resulting in 5 individuals with negative event times who were excluded from the dataset (see clarification response,[2] question B11). The resulting adjusted survival curve is shown in Figure 7. Data from Zerres and Rudnik-Schöneborn[43] were also considered in a scenario analysis (see Section 5.3.5).

Figure 7: Kaplan-Meier estimates based on Gregoretti *et al*, adjusted for mean age of patients at the start of the ENDEAR trial (reproduced from CS Appendix P, Figure 55)



NRA - non-invasive respiratory aid; NT - no treatment; TV - tracheotomy and invasive mechanical ventilation.

For patients in States (iv) to (vii) in the nusinersen group, the company considered that motor milestones characteristic of later onset patients would be achieved, hence survival would be between that of Type I and Type II SMA patients. Type II mortality in time periods 2 and 3a was modelled based on the SMA Type II population of Zerres *et al*[33] and is briefly described in the CS[1] (Section 4.3.1, page 168 and Appendix P, page 212). This natural history study included 240 patients with Type II SMA, recruited from 1960 onwards. The company reconstructed IPD from the published Kaplan-Meier curve without performing any adjustment.

Beyond the end of follow-up in Gregoretti *et al.*[31] OS was modelled based on general population mortality data from the Office for National Statistics (ONS),[32] using average life tables for males and females. IPD were reconstructed using the algorithm reported by Guyot *et al.*[44] CS Appendix P (page 208) states *"Since only in survival after 19 years was of interest, infant mortality (children <3 years old) was removed from these data."* The ERG is unclear regarding the relevance and appropriateness of this statement.

Survival modelling results - early onset model

OS time period 1: ENDEAR follow-up (up to the end of month 13)

Model fit statistics for all parametric models fitted to the ENDEAR trial data[14] are summarised in Table 31. The predicted survival probabilities are illustrated in CS,[1] Appendix P, Figures 28 and 29. The company selected the combined model Royston-Parmar cubic spline with 1 knot for the base case, as it provided a good fit to the data and preserves the assumption of proportional hazards, which the company considered to be appropriate for the data. The CS states that the combined Royston-Parmar cubic spline with 2 knots and the combined Gompertz models also provided a good fit.

Table 31: Model fit statistics for parametric models fitted to ENDEAR OS data (adapted from CS Appendix P, Figure 30, Figure 31 and Figure 33)

Model	Combined/stratified	AIC	BIC	IBS
Cubic spline 1-knot	Combined (PH)	**251.4**	**256.9**	0.1556
Cubic spline 2-knot	Combined (PH)	**251.6**	**258.5**	0.1558
Gompertz	Combined (PH)	**251.9**	**256.0**	0.1556
Cubic spline 3-knot	Combined (PH)	253.5	261.7	NR
Gompertz	Stratified	253.8	259.2	0.1555
Cubic spline 1-knot	Stratified	255.2	263.4	NR
Log normal	Combined (AFT)	255.8	259.9	0.2192
Log normal	Stratified	256.5	262.0	0.1561
Cubic spline 2-knot	Stratified	256.9	267.8	NR
Generalised gamma	Combined (AFT)	257.2	262.7	NR
Log logistic	Combined (AFT)	257.7	261.8	0.2424
Cubic spline 3-knot	Stratified	257.9	271.6	NR
Weibull	Combined (PH)	259.2	263.4	0.2606
Log logistic	Stratified	259.3	264.8	0.1565
Exponential	Combined (PH)	259.8	262.5	0.2560
Weibull	Stratified	261.2	266.6	0.1558

AIC – Akaike Information Criterion; BIC – Bayesian Information Criterion; IBS - Integrated Brier Score, lower numbers are favourable; PH – proportional hazards; AFT – accelerated failure time; NR - not reported
Numbers in bold relate to highest rank (lowest AIC/BIC) or within 3 of lowest AIC/BIC

OS time period 2: From end of ENDEAR follow-up to end of Gregoretti et al follow-up

Model fit statistics for all parametric models fitted to the adjusted Gregoretti NRA data[31] are summarised in Table 32; fitted survival curves are provided in CS[1] Appendix P, Figures 57-62. The company selected the exponential model for the base case. The company considered that all models gave a good visual fit to the observed data but that only the exponential, Weibull and hybrid models gave plausible long-term predictions. The exponential model was considered to give the best fit; predicted hazard rates from this model were applied to the sham (usual care) group from month 14 to month 58.

OS for the treatment group was also informed by Zerres *et al*.[33] Model fit statistics for all parametric models fitted to the reconstructed Zerres *et al* data[33] are summarised in Table 33; predicted survival

probabilities are provided in CS[1] Appendix P, Figures 74 and 75. The combined Royston-Parmar spline model with 2 knots was selected for use in the company's base case as this model gave the best fit in terms of the AIC and the BIC.

For patients in model health states (iv) to (vii) who are receiving treatment using nusinersen, OS was then assumed to be between that of the survival prediction from Zerres *et al*[33] and Gregoretti *et al*[31] according to the weighting given in Equation [i].

$$S_{Nusinersen}(t) = 0.9\, S_{Zerres}(t) + 0.1\, S_{Gregorretti}(t) \quad \text{[i]}$$

Justification of these weightings was provided through reference to an advisory board meeting on SMA held by the company,[45] although the ERG notes that the documentation provided by the company does not report the values of the weights applied in the model. The predicted hazards from the weighted combination of survival functions were applied to the treatment arm from months 14 to 58.

Table 32: Model fit statistics for parametric models fitted to adjusted Gregoretti *et al* NRA OS data (adapted from CS Appendix P, Figure 63, 64, 66)

Model	AIC	BIC	IBS
Exponential	**152.2**	**152.8**	0.16733
Log logistic	**153.2**	**154.5**	0.16952
Log normal	**153.3**	**154.5**	0.17039
Weibull	**153.8**	**155.1**	0.17053
Gompertz	**154.2**	**155.5**	0.17057
Generalised gamma	**155.2**	157.2	NR
Cubic spline 1-knot	155.6	157.5	0.17072
Cubic spline 2-knot	157.3	159.8	NR
Cubic spline 3-knot	157.5	160.7	NR
Cubic spline 4-knot	159.1	162.9	NR

AIC – Akaike Information Criterion; BIC – Bayesian Information Criterion; IBS - Integrated Brier Score, lower numbers are favourable; NR - not reported Numbers in bold relate to highest rank (lowest AIC/BIC) or within 3 of lowest AIC/BIC

Table 33: Model fit statistics for parametric models fitted to Zerres *et al* 1997 Type II OS data (adapted from CS Appendix P, Figure 77, Figure 78 and Figure 80)

Model	AIC	BIC	IBS
Cubic spline 2-knot	**1563.5**	**1574.3**	0.16246
Weibull	1583.2	1588.6	0.16303
Cubic spline 1-knot	1585.2	1592.6	0.16340
Gompertz	1587.2	1593.3	0.16338
Log logistic	1590.7	1596.1	0.16359
Log normal	1592.6	1598.0	0.16351
Exponential	1644.5	1647.2	0.16920

AIC – Akaike Information Criterion; BIC – Bayesian Information Criterion; IBS - Integrated Brier Score, lower numbers are favourable; NR - not reported Numbers in bold relate to highest rank (lowest AIC/BIC) or within 3 of lowest AIC/BIC

OS time period 3: From end of Gregoretti et al follow-up to end of time horizon, usual care arm and nusinersen states (i)-(iv)

Model fit statistics for all parametric models fitted to the general population mortality data are summarised in Table 34; fitted survival probabilities are provided in CS[1] Appendix P, Figure 67. The Gompertz model was preferred by the company on account of the theoretical justification of this distribution to model healthy populations and the simplicity of the model. The predicted hazards were compared with those from the exponential model fitted to the Gregoretti *et al* data,[31] providing an HR of 5184.8. This HR was applied to the Gompertz model derived from the general population in order to "*adjust the survival curve for a population matching that of Gregoretti*" (CS,[1] Section 3.3.4.1, page 126).

Table 34: Model fit statistics for parametric models fitted to reconstructed general population mortality data (adapted from CS Appendix P, Figures 68, 69 and 71)

Model	AIC	BIC	IBS
Cubic spline 2-knot	**85927.1**	**85921.8**	0.070502
Cubic spline 1-knot	85935.0	85931.1	0.70496
Gompertz	85961.6	85958.7	0.070488
Weibull	86295.9	86293.3	0.70716

AIC – Akaike Information Criterion; BIC – Bayesian Information Criterion; IBS - Integrated Brier Score, lower numbers are favourable; NR - not reported Numbers in bold relate to highest rank (lowest AIC/BIC) or within 3 of lowest AIC/BIC

OS time period 3a: From end of Gregoretti et al follow-up to end of Zerres et al, nusinersen states (iv)-(vii)

For nusinersen-treated patients reaching states (iv) to (vii), there is an additional stage due to differing durations of follow-up within Gregoretti *et al*[31] and Zerres *et al.*[33] After the end of Gregoretti *et al*, the portion of the OS informed by Gregoretti *et al* is replaced by the HR-adjusted general population survival shown in Equation [ii].

$$S_{Nusinersen}(t) = 0.9\, S_{Zerres}(t) + 0.1\, (S_{genpop}(t))^{1/5184.8} \qquad \text{[ii]}$$

OS time period 3b: From end of Zerres et al, nusinersen states (iv)-(vii)

Beyond the end of the Zerres *et al*[33] follow-up period, the portion of Equation [ii] informed by the Zerres *et al* survival curve is then replaced by the adjusted general population mortality resulting in a weighted combination of two Gompertz models shown in Equation [iii]. To estimate the adjustment factor, predicted hazard rates at 53 years were compared. The Gompertz model fitted to the general population gave a hazard rate of 0.00028, whilst the exponential model fitted to the Zerres *et al*[33] data gave a hazard rate of 0.00745; the estimated HR was 26.4.

$$S_{Nusinersen}(t) = 0.9(S_{genpop}(t))^{1/26.4} + 0.1\, (S_{genpop}(t))^{1/5184.8} \qquad \text{[iii]}$$

Transition probabilities

Transition probabilities were estimated using different approaches for the observed period of ENDEAR[14] and for subsequent cycles. Within the observed period, transitions were estimated directly using observed HINE-2 count data for each treatment group. Separate matrices were calculated for each of four cycles (day 1-64, day 65-183, day 184-302 and day 303-394). In response to a request for clarification,[3] the company stated that matrices were generated using the efficacy dataset without imputation of missing data; however, the ERG notes that these matrices contain count data for a larger number of patients than were included in the efficacy set.

Beyond the end of follow-up in ENDEAR (after the end of month 13), two transition matrices are applied: the first is applied for the interval from the end of month 13 to the end of month 14, whilst the second is applied to all subsequent 4-monthly cycles. These matrices were estimated by calculating the mean rate of change in CHOP INTEND score in each treatment group and the mean CHOP INTEND score within each HINE-2 model health state within each treatment group over the duration of ENDEAR. Data on mean CHOP INTEND score by HINE-2 state from Study CS3A[34] were used for State (v) Stands with assistance and State (vi) Walks with assistance due to limited data in ENDEAR (see Table 35). Transition probabilities for patients in the nusinersen and usual care groups were calculated using Equation [iv] and Equation [v], respectively.

$$TP\ (nusinersen)\ = MIN[1{,}1 + \left(\frac{Rate\ CHI\ increase\ (per\ month)\ .\ cycle\ length\ (months)}{Mean\ CHI\ next\ best\ state - Mean\ CHI\ current\ state}\right) \quad \text{[iv]}$$

$$TP(usual\ care)\ = MIN[1{,}1 + \left(\frac{Rate\ CHI\ decrease\ (per\ month)\ .\ cycle\ length\ (months)}{Mean\ CHI\ current\ state - Mean\ CHI\ next\ worst\ state}\right) \quad \text{[v]}$$

CHI - CHOP INTEND score

Table 35: CHOP INTEND data used to inform transition probabilities beyond month 13

MEAN CHOP INTEND SCORE			
HINE-2 health state	**Nusinersen**	**Sham**	**Source**
No milestones	24.59	20.19	ENDEAR[14]
Mild milestones	32.98	26.83	
Moderate milestones	41.45	37.11	
Sits without support	46.67	48.00	
Stands with assistance	52.67	52.67	Study CS3A[34]
Walks with assistance	63.00	63.00	
Walks unaided	-	-	
RATE OF IMPROVEMENT/WORSENING			
	Nusinersen	**Sham**	
Monthly CHI rate	[redacted]	[redacted]	ENDEAR[14]

Estimated transition matrices for the first four model cycles (based on the observed count data from ENDEAR) are shown in Table 36, Table 37, Table 38 and Table 39. Table 40 and Table 41 present the transition matrices applied after the end of month 13.

Table 36: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 1-64 (taken from company's model)

NUSINERSEN GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								
SHAM GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								

** No observed transitions from state during cycle; Blank cells indicate zero probability*
n - number

Table 37: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 65-183 (taken from company's model)

NUSINERSEN GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								
SHAM GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								

** No observed transitions from state during cycle; Blank cells indicate zero probability*
n - number

Table 38: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 184-302 (taken from company's model)

NUSINERSEN GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								
SHAM GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								

** No observed transitions from state during cycle; Blank cells indicate zero probability*
n - number

Table 39: Transition matrices for nusinersen (top) and sham (bottom), HINE-2 observed count data, ENDEAR trial, days 303-394 (taken from company's model)

NUSINERSEN GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								
SHAM GROUP (patients alive with data n=)								
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided	Dead
No milestones								
Mild milestones								
Moderate milestones								
Sits without support								
Stands with assistance								
Walks with assistance								
Stands/walks unaided								
Dead								

** No observed transitions from state during cycle; Blank cells indicate zero probability*
n - number

Table 40: Transition matrices for nusinersen (top) and sham (bottom), extrapolation based on CHOP INTEND score in ENDEAR trial, months 13-14 (taken from company's model)

NUSINERSEN GROUP (patients alive with data n=n/a)							
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]						
Mild milestones		[redacted]	[redacted]				
Moderate milestones			[redacted]	[redacted]			
Sits without support				[redacted]	[redacted]		
Stands with assistance					[redacted]	[redacted]	
Walks with assistance						[redacted]	[redacted]
Stands/walks unaided							[redacted]
SHAM GROUP (patients alive with data n=n/a)							
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]						
Mild milestones	[redacted]	[redacted]					
Moderate milestones		[redacted]	[redacted]				
Sits without support			[redacted]	[redacted]			
Stands with assistance				[redacted]	[redacted]		
Walks with assistance					[redacted]	[redacted]	
Stands/walks unaided						[redacted]	[redacted]

Blank cells indicate zero probability
n – number; n/a - not applicable

Table 41: Transition matrices for nusinersen (top) and sham (bottom), extrapolation based on CHOP INTEND score in ENDEAR trial, all 4-month cycles after month 14 (taken from company's model)

NUSINERSEN GROUP (patients alive with data n=n/a)							
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones							
Mild milestones							
Moderate milestones							
Sits without support							
Stands with assistance							
Walks with assistance							
Stands/walks unaided							
SHAM GROUP (patients alive with data n=n/a)							
From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones							
Mild milestones							
Moderate milestones							
Sits without support							
Stands with assistance							
Walks with assistance							
Stands/walks unaided							

Blank cells indicate zero probability
n – number; n/a - not applicable

Probability of undergoing surgery for scoliosis and age at time of scoliosis surgery

The company's assumptions regarding scoliosis surgery are summarised in Table 42. The company's early onset model assumes that within the nusinersen group, 1% of surviving patients will undergo scoliosis surgery at year 12 if non-ambulant or at year 15 if ambulant. Twenty percent of patients receiving nusinersen who undergo scoliosis surgery are assumed to subsequently discontinue treatment. Within the usual care group, the model assumes that 1% of surviving patients will undergo scoliosis surgery at year 10 if non-ambulant or at year 15 if ambulant. The probabilities of undergoing scoliosis surgery and subsequently discontinuing nusinersen were based assumptions.[1] The timing of surgery appears to have been loosely based on a paper describing SMA and its management by Haaker and Fujak.[35]

Table 42: Scoliosis surgery parameters included in the early onset model

Parameter	Nusinersen	Usual care	Source
Percentage of patients undergoing surgery for scoliosis	1%	1%	Assumption[1]
Percentage of patients discontinuing nusinersen following scoliosis surgery	20%	n/a	Assumption[1]
Time of surgery since model start (non-ambulant)	12 years	10 years	Haaker and Fujak[35]
Time of surgery since model start (ambulant)	15 years	15 years	Haaker and Fujak[35]

HRQoL - patient utilities

Neither the ENDEAR trial nor the CHERISH trial included the use of a preference-based instrument to assess HRQoL. In addition, the company's review of published HRQoL studies[1] did not identify any suitable studies. The CS[1] highlights that the derivation of HRQoL estimates for patients with SMA is challenging due to the nature of the condition and the age of the population.

Initially, the company explored the use of a *de novo* case vignette study (Lloyd *et al*[46]) to estimate health utilities associated with each of the health states within the model. The company held interviews with five clinical experts to draft case studies representing each of the modelled health states. The company subsequently held further interviews with five clinical experts to value the health states using the EQ-5D-Y (using the adult EQ-5D tariff) and the PedsQL Neuromuscular Module. However, the valuations produced negative utility scores for most of the states and the CS states that some of the rankings of health state valuations were counterintuitive. Further details are given in the documentation relating to the expert advisory board meeting,[45] although the ERG notes that the issues relating to counterintuitive rankings relate to states which are not used in the company's final models (see clarification response,[1] question B5). As a consequence of the reservations raised by several of the clinical experts consulted, the company decided not to use these utilities within either the early onset or later onset models.

The clinical experts consulted by the company expressed a preference to instead use the PedsQL data collected as part of the CHERISH study in later onset SMA patients.[15] These data were mapped onto the EQ-5D using an algorithm published by Khan *et al.*[36] The PedsQL to EQ-5D mapping algorithm was estimated using data from a cross-sectional survey conducted in four secondary schools in England amongst children aged 11–15 years of age. The selected ordinary least squares (OLS) mapping algorithm was calculated relative to the EQ-5D utility for the general population based on a predictive equation with coefficients on age, the square of age and sex. The resulting mapped EQ-5D utility values were assumed to apply for later onset patients and were adapted for the early onset model based on an assumed correspondence of health states between early onset and later onset models (see Table 43).

Table 43: Patient utilities used in the early onset model

Early onset SMA model (HINE-2-based health states)	Later onset SMA model (HFMSE-based health states)	Mapped utility value
No milestones	Sits without support but does not roll	[redacted]
Mild milestones	Sits and rolls independently	
Moderate milestones	Sits and rolls independently	
Sits without support	Sits and crawls on hands and knees	
Stands with assistance	Stands or walks with assistance	
Walks with assistance	Stands without assistance	
Stand or walks without assistance	Walks without assistance	

SMA – spinal muscular atrophy; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; HFMSE - Hammersmith Functional Motor Scale-Expanded

HRQoL – caregiver disutilities

The early onset model includes disutilities for caregivers of SMA patients; these are assumed to be dependent on the motor milestones achieved by the patient during each model cycle. Caregiver disutilities were calculated using: (i) the mean caregiver EQ-5D score reported in a cross-sectional study of patients with SMA in France, Germany, Spain and United Kingdom (Bastida *et al*[37]); (ii) an estimate of the mean utility of the general population[38] (assuming a constant age of 30.88 years, 80% female) and (iii) the mapped patient utilities estimated for each health state (see Table 43). The company first estimated health utilities for caregivers conditional on the patient's health state by subtracting the difference in patient utilities between selected HINE-2 health states from the baseline caregiver utility reported by Bastida *et al*.[37] Caregiver disutilities were then calculated by subtracting the caregiver utility estimate from the mean general population utility estimate. The derivation of each health state-specific disutility is shown in Table 44.

Table 44: Parent/carer utilities used in the early onset model

HINE-2 health state	Patient utility	Caregiver utility	Caregiver disutility*	Calculation and assumptions
(i) No milestones	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (ii) Mild milestones and State (i) No milestones states
(ii) Mild milestones	[redacted]	[redacted]	[redacted]	Based directly on Bastida *et al*[37] caregiver utility
(iii) Moderate milestones	[redacted]	[redacted]	[redacted]	Based directly on Bastida *et al*[37] caregiver utility
(iv) Sits without support	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (ii) Moderate milestones and State (iv) Sits without support
(v) Stands with assistance	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (iii) Moderate milestones states and State (v) Stands with assistance
(vi) Walks with assistance	[redacted]	[redacted]	[redacted]	Assumed to be the same as State (v) Stands with assistance
(vii) Stand or walks without assistance	[redacted]	[redacted]	[redacted]	Assumed to be the same as State (v) Stands with assistance
Baseline parameters				
Bastida *et al*[37] caregiver utility		[redacted]	-	-
General population utility[38]		0.92	-	Caregiver age=30.88 years, 80% female
Bereavement		-	-0.04	-

** Calculated as general population utility minus caregiver utility*

Resource use and costs

The company's early onset model includes the following cost components: (i) nusinersen acquisition and administration costs; (ii) health state costs and (iii) end-of-life costs.

Drug acquisition and administration costs

The acquisition cost for nusinersen is £75,000 per vial.[1, 13]

The company's model assumes that nusinersen is administered via lumbar puncture. Forty percent of all nusinersen administrations are assumed to be given in an inpatient setting, 30% are assumed to be given in an outpatient setting and the remaining 30% are assumed to be given in a day case setting. The costs for lumbar puncture were taken from NHS Reference Costs 2015/2016[39] using HRG codes HC72A (Diagnostic Spinal Puncture, 19 years and over), HC72B (Diagnostic Spinal Puncture, between 6 and 18 years) and HC72C (Diagnostic Spinal Puncture, 5 years and under). The company calculated weighted mean administration costs of £1,359 for patients aged 5 years and under, £1,295 for those aged between 6 and 18 years and £606 for those aged 19 years and over (see Table 45).

Table 45: Estimated nusinersen administration costs

Description	Mean cost	NHS Reference Costs 2015/16 code[39]
Age 5 years and under		
Inpatient	£1,690	EL - HC72C
Outpatient	£577	OPROC - HC72C (service code 421)
Day case	£1,700	DC - HC72C
Weighted mean cost	**£1,359**	
Age 6 to 18 years		
Inpatient	£1,658	EL - HC72B
Outpatient	£560	OPROC - HC72B (service code 421)
Day case	£1,546	DC - HC72B
Weighted mean cost	**£1,295**	
Age 18 years and over		
Inpatient	£918	EL - HC72A
Outpatient	£204	OPROC - HC72A (service code 400)
Day case	£593	DC - HC72AB
Weighted mean cost	**£606**	

EL - elective inpatient; OPROC - outpatient procedures; DC - day case

Health state costs

Health state costs were based on data from the cross-sectional SMA study reported by Bastida *et al.*[37] Within this study, the main caregivers of children/adolescents diagnosed with SMA completed a self-administered questionnaire providing information related to sociodemographics, the costs of professional private care, the need for informal care, expenditure and resource utilisation related to SMA.[1] The company took the health state costs data from Bastida *et al*[37] (reported Euros, year 2014) and converted these values to Pounds Sterling (year 2016) using an exchange rate also provided in Bastida *et al*[37] and changes in consumer prices between 2014 and 2016 (see Table 46).

Table 46: Estimated annual costs by category of resource use in Type I, II and III SMA patients (reproduced from CS Table 41)

Description	Type I SMA		Type II SMA		Type III SMA	
	€ 2014	£ 2016	€ 2014	£ 2016	€ 2014	£ 2016
Drugs	[redacted]					
Medical tests						
Medical visits						
Hospitalisations						
GP & emergency						
Health material						
Social services						
Total						

SMA – spinal muscular atrophy

The data provided by Bastida *et al*[37] were divided into a number of resource classifications: drugs; medical tests; medical visits; hospitalisations; general practitioner (GP) & emergency visits, health material and social services; a brief description of what is included in each of these classifications is provided below:

- Drugs - costs for drugs such as creatine, gabapentin, hydroxyurea, vitamin supplements and calcium.
- Medical tests - costs associated with blood tests, urinalysis, electrocardiogram, magnetic resonance imaging, range of motion tests, spirometry and x-rays of the chest, back and hip.
- Medical visits - costs associated with home visits and hospital outpatient appointments with urologist, neurologist, psychiatrist, dermatologist, nephrologist, respiratory consultant, nutritionist, occupational therapist, traumatologist, specialists in palliative care and respiratory physiotherapist.
- Hospitalisations - costs associated with any hospital inpatient treatment.
- GP & emergency - costs related to appointments with GP, practice nurses or emergency treatments.
- Health material - costs associated with the provision of orthosis, prosthesis, wheelchairs, adjustable beds, shower chairs, humidifiers, portable oxygen, food supplements and gastric feeding cannulas, pulse oximetry and communication aids.
- Social services - costs associated with care provided by a day centre or occupational centre, respiratory physiotherapists, occupational physiotherapists, psychosocial care for the family are respite care in residential centres.[1]

The company then divided each of these costs according to four main therapy areas using proportions based either on expert medical opinion or assumptions (see Table 47).

Table 47: Allocation of costs by resource classification

Description	Respiratory care	Gastrointestinal care	Nutritional care	Orthopaedic care
Drugs	50%	50%	0%	0%
Medical tests	25%	25%	25%	25%
Medical visits	[redacted]	[redacted]	[redacted]	[redacted]
Hospitalisations	[redacted]	[redacted]	[redacted]	[redacted]
GP & emergency	[redacted]	[redacted]	[redacted]	[redacted]
Health material	25%	25%	25%	25%
Social services	25%	25%	25%	25%

This was done for each of the three types of SMA. The company applied the estimated costs for each SMA type to health states describing outcomes consistent with those SMA types (see Table 48).

Table 48: Annual health state costs, early onset model

Cost component	Milestones consistent with Type I SMA (State [i] No milestones; State [ii] Mild milestones; State [iii] Moderate milestones)	Milestones consistent with Type II SMA (State [iv] Sits without support; State [v] Stands with assistance; State [vi] Walks with assistance)	Milestones consistent with Type III SMA (State [vii] Stands/walks unaided).
Respiratory care	[redacted]	[redacted]	[redacted]
Gastrointestinal care	[redacted]	[redacted]	[redacted]
Nutritional care	[redacted]	[redacted]	[redacted]
Orthopaedic care	[redacted]	[redacted]	[redacted]
Total	[redacted]	[redacted]	[redacted]

The ERG notes that the company's approach to breaking down the costs by type of care is irrelevant as the sum of the costs shown in Table 48 (after manipulation) is the same as the sum of the costs presented in Table 46 (before manipulation).

End-of-life costs

The company's early onset model includes a once-only end-of-life cost of £11,839. The source of this cost is not described in the CS;[1] text contained in the executable model indicates that this value was informed by NICE Guideline 61.[40]

5.3.4 Methods for model evaluation

The CS[1] presents the results of the early onset model in terms of the incremental cost per QALY gained for nusinersen versus usual care. Separate results are presented for: (i) analyses including patient health gains only and (ii) analyses including patient health gains and caregiver QALY losses. The company's base case incremental cost-effectiveness ratios (ICERs) are based on the deterministic version of the model. The CS also includes the results of probabilistic sensitivity analysis (PSA), deterministic sensitivity analyses (DSAs), scenario analyses and subgroup analyses. The results of the PSA are presented in the form of a cost-effectiveness plane and cost-effectiveness acceptability curves (CEACs), based on 1,000 Monte Carlo simulations. The probabilistic ICER is also presented. The distributions applied in the company's PSA are summarised in Table 49. The results of the DSAs are presented in the form of a tornado diagram for specified model parameters. Scenario analyses were undertaken to explore the impact of alternative time horizons, and alternative assumptions surrounding mortality risk, transition probabilities, costs and HRQoL.

Table 49: Distributions used in company's PSA, early onset model

Parameter group	Distribution	ERG comment
Initial HINE-2 health state distribution - nusinersen	Fixed	These parameters are subject to uncertainty. Given the multinomial nature of the data, a Dirichlet distribution (applied to the combined ENDEAR population) would be appropriate.
Initial HINE-2 health state distribution – usual care	Fixed	
Overall survival – nusinersen early onset SMA	Multivariate normal	The adjustment factor for Type II mortality in the better states is fixed at its mean value.
Overall survival – usual care early onset SMA	Multivariate normal	-
Transition probabilities – nusinersen (up to month 13)	Dirichlet	Priors are included for some but not all unobserved transitions.
Transition probabilities – usual care (up to month 13)	Dirichlet	
Transition probabilities – nusinersen (month 14 onwards)	Dirichlet	
Transition probabilities – usual care (month 14 onwards)	Dirichlet	
Probability of undergoing surgery for scoliosis	Beta	Inappropriately characterised using treatment-specific parameters.
Age at time of surgery	Normal	Inappropriately characterised using treatment-specific parameters.
Probability of discontinuing nusinersen after surgery for scoliosis	Beta	-
Patient utilities	Beta	All utilities sampled using the same random number, thereby inducing over-correlation between states.[47]
Baseline caregiver utilities	Beta	No uncertainty is included in the Bastida *et al*[37] baseline caregiver disutility.
Nusinersen acquisition cost	Fixed	-
Nusinersen administration costs	Normal (cost) and Dirichlet (administration setting)	-
Health state costs	Gamma	-

HINE-2 – Module 2 of the Hammersmith Infant Neurological Examination; ERG – Evidence Review Group

5.3.5 Company's model results – early onset model

This section presents the results of the company's early onset model, evaluated over a 60-year time horizon.

Central estimates of cost-effectiveness – early onset model

Table 50 presents the central estimates of cost-effectiveness derived from the company's model (including health gains accrued by patients only). Based on a re-run of the probabilistic version of the model by the ERG, nusinersen is expected to generate an additional 5.29 QALYs at an additional cost

of £2,160,048 per patient; the corresponding ICER for nusinersen versus usual care is £408,712 per QALY gained. The deterministic version of the model produces a similar ICER of £407,605 per QALY gained for nusinersen versus usual care. The inclusion of caregiver QALY losses (see Table 51) leads to a slightly lower probabilistic ICER of £404,270 per QALY gained; the deterministic ICER is estimated to be £402,361 per QALY gained.

Table 50: Company's model results, early onset model (including patient health gains only)

Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	7.73	£2,229,863	5.29	£2,160,048	**£408,712**
Usual care	2.45	£69,814.82	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	7.86	£2,258,852	5.37	£2,187,311	**£407,605**
Usual care	2.49	£71,540	-	-	-

Inc. - incremental; QALY - quality-adjusted life year

Table 51: Company's model results, early onset model (including patient health gains and caregiver QALY losses)

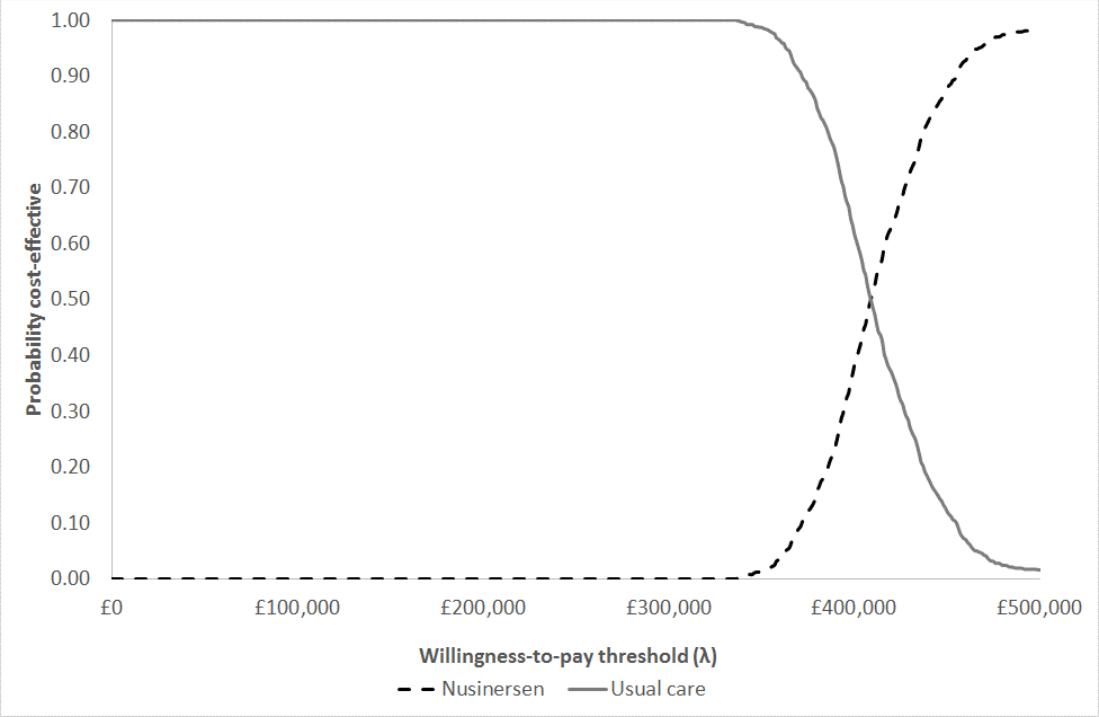
Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	7.49	£2,229,863	5.34	£2,160,048	**£404,270**
Usual care	2.14	£69,814.82	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	7.61	£2,258,852	5.44	£2,187,311	**£402,361**
Usual care	2.17	£71,540	-	-	-

Inc. - incremental; QALY - quality-adjusted life year

Company's probabilistic sensitivity analysis - early onset model

Figure 8 presents CEACs for nusinersen and usual care for the early onset population. As shown in the figure, the probability that nusinersen produces more net benefit than usual care at willingness-to-pay (WTP) thresholds below £337,000 per QALY gained is approximately zero.

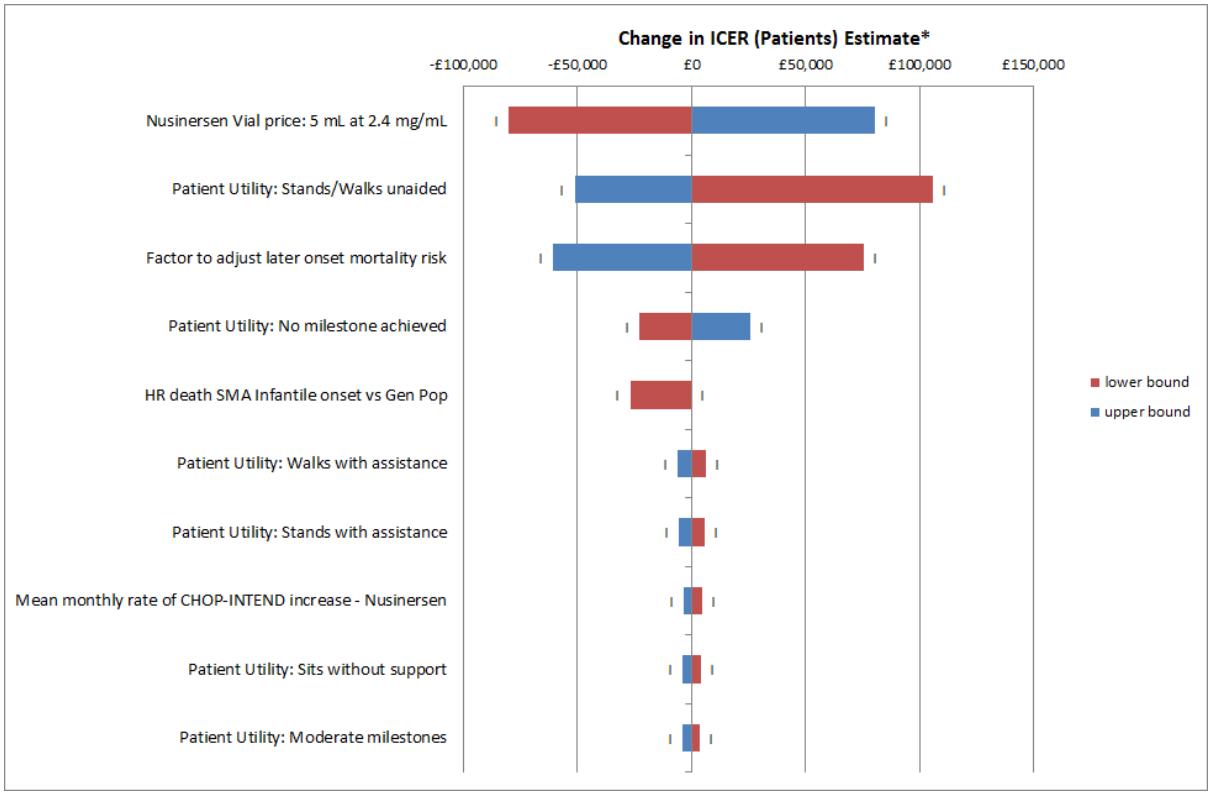
Figure 8: CEACs, early onset model, patient health gains only



Company's deterministic sensitivity analyses - early onset model

Figure 9 presents the results of the company's DSAs in the form of a tornado diagram (change in ICER from baseline). As shown in the figure, the most influential model parameters relate to the acquisition cost of nusinersen, the health utility associated with State (vii) Stands/walks unaided, and the Type II SMA mortality adjustment factor applied to the better health states. The lowest ICER generated from the company's one-way DSAs is £327,347 per QALY gained (nusinersen vial price=£60,000) whilst the highest ICER is £513,324 per QALY gained (health utility State [vii] Stands/walks unaided = [redacted]).

Figure 9: Company's DSA tornado diagram, early onset model, patient health gains only



CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; ICER – incremental cost-effectiveness ratio; HR – hazard ratio

Scenario analysis results - early onset model

Table 52 details the results of the company's scenario analyses. As shown in the table, the early onset model is very sensitive to the assumptions regarding the Type II mortality adjustment applied to States (iv) to (vii). The lowest ICER for nusinersen versus usual care is estimated to be £347,082 per QALY gained when only patient health gains are considered, and £345,578 per QALY gained when caregiver QALY losses are included (mortality adjustment factor=1.00). The highest ICER for nusinersen versus usual care is estimated to be £872,257 per QALY gained, when only patient health gains are considered, and £802,469 per QALY gained when caregiver QALY losses are included (mortality adjustment factor=0.00).

Table 52: Scenario analysis results, early onset model

Scenario	ICER (patient health gains only)	ICER (patient health gains and caregiver QALY losses)
Base case (deterministic)	£407,605	£402,361
Time horizon=10 years	£564,659	£543,695
Time horizon=20 years	£436,278	£428,375
Time horizon=30 years	£410,888	£405,315
Do not apply higher long-term risk of death based on SMA Type I - adjusted general mortality rates	£380,658	£376,357
OS beyond trial follow-up based on Zerres 1995 + 2 knots & 60-year time horizon	£379,804	£376,289
OS treatment effects - taper HR to 1.0 over 12 months	£405,766	£400,680
Apply discontinuation to State (i) No milestones and State (ii) Mild milestones	£406,096	£402,138
Do not apply Type II mortality rates from Zerres *et al* to patients in motor milestones characteristic of later onset	£872,257	£802,469
Mortality risk factor=0.50	£578,554	£556,339
Mortality risk factor=1.00	£347,082	£345,578
Assumption that proportion of patients on treatment reach a plateau (0% worsen)	£417,355	£412,445
Assumption that proportion of patients on treatment reach a plateau (10% worsen)	£421,445	£417,806
Source for usual care arm CHOP INTEND rate of decline - Finkel *et al*. 2012	£407,315	£402,328
All nusinersen administration inpatient	£409,438	£404,170
All nusinersen administration day case	£409,015	£403,752
Health state costs include costs of major clinical events only	£442,838	£437,140
Cost source – Klug *et al*	£405,194	£399,980
Patient utility based on vignettes	£421,703	£394,298
Patient utility based on Bastida upper bound	£450,353	£476,009
Patient utility based on Bastida lower bound	£503,295	£788,019
Patient utility based on PedsQL type 2 (<25 months disease duration)	£387,628	£364,333

SMA - spinal muscular atrophy; OS - overall survival; CHOP INTEND - Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders; PedsQL - Paediatric Quality of Life Inventory; ICER – incremental cost-effectiveness ratio

Subgroup analysis - early onset model

Table 53 presents the results of the company's subgroup analysis based on disease duration (≤12 weeks and >12 weeks). It should be noted that the results of the subgroup analyses presented in the CS[1] are incorrect and should be disregarded; the results presented in Table 53 are based on additional information provided by the company following the clarification process.[3]

Table 53: Subgroup analysis results, early onset model

Subgroup	ICER (patient health gains only)	ICER (patient health gains and caregiver QALY losses)
ITT population, each arm (base case)*	£407,605	£402,361
ITT population, both arms (base case)†	£409,235	£404,015
≤12 weeks disease duration each arm*	£375,237	£370,915
≤12 weeks disease duration both arms†	£375,775	£371,458
>12 weeks disease duration each arm*	£484,614	£473,247
>12 weeks disease duration both arms†	£485,766	£474,355

** "thresholds" defining HINE-2 health states based on mean CHOP INTEND scores in each treatment group;*
† "thresholds" defining HINE-2 health states based on mean CHOP INTEND scores across both treatment groups
ITT – intention-to-treat; ICER – incremental cost-effectiveness ratio; QALY – quality-adjusted life year

5.4 Later onset model – methods and results

5.4.1 Model structure and logic – later onset model

The company's later onset model follows a conceptual design which is broadly similar to the early onset model described in the previous section (see Figure 10). The later onset model adopts a state transition approach based on health states defined according to the HFMSE instrument.[48] The later onset model includes seven health states: (i) Sits without support but does not roll; (ii) Sits and rolls independently; (iii) Sits and crawls with hands and knees; (iv) Stands/walks with assistance; (v) Stands unaided; (vi) Walks unaided and (vii) Dead. The domains of the HFMSE are presented in Appendix 3. The classification of health states within the company's later onset model according to HFMSE scores is summarised in Table 54.

Figure 10: Company's later onset model structure (reproduced from CS, Figure 43)

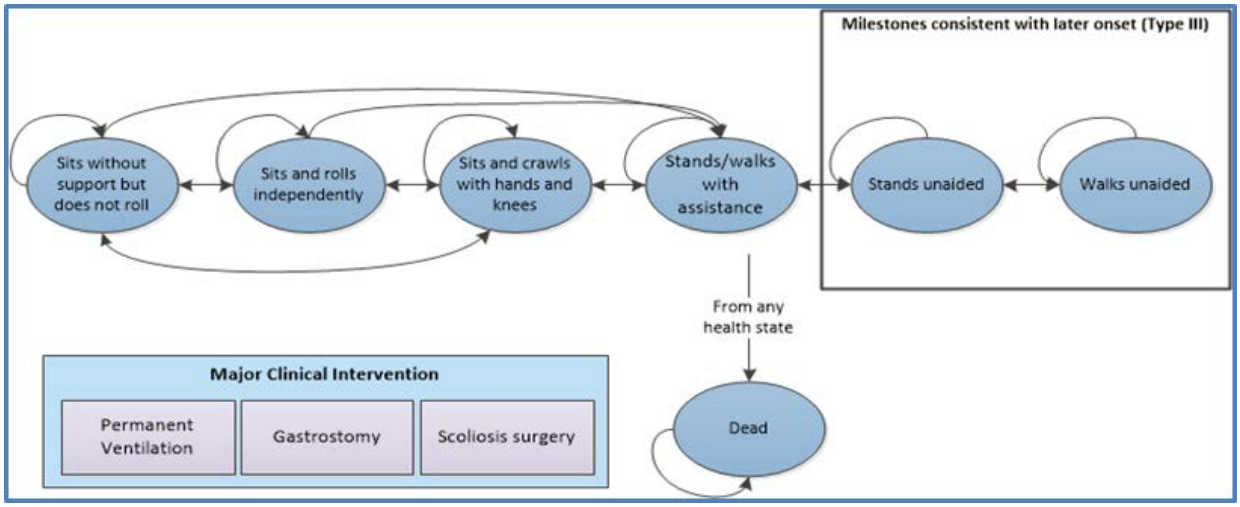


Table 54: Model health states according to HFMSE score (adapted from CS, Figure 43 footnotes)

Model health state	HFMSE criteria for model health state
(i) Sits without support but does not roll	Patients sit according to the WHO criteria and have a score <2 in Rolls Prone to Supine right and left in HFMSE score
(ii) Sits and rolls independently	Patients sit according to the WHO criteria and have a score of 2 in Rolls Prone to Supine right or Rolls Prone to Supine left in HFMSE score
(iii) Sits and crawls with hands and knees	Based on WHO criteria (see Appendix 3)
(iv) Stands/walks with assistance	
(v) Stands unaided	
(vi) Walks unaided	

WHO – World Health Organization

Model logic

The logic of the company's later onset model is described in the sections below.

Nusinersen group

Patients enter the model based on the baseline HFMSE health state distribution for the nusinersen group in the CHERISH study.[15] During the first five model cycles (up to the end of month 15), mortality risk is assumed to be zero, based on the observed number of deaths within the nusinersen group of CHERISH. From model entry until the end of month 15, transitions between the six HFMSE-based health states are governed by five cycle-specific transition matrices derived from observed count data within CHERISH.

From the end of month 15 to the end of month 623, mortality is modelled using a 2-knot Royston-Parmar spline model fitted to survival data for Type II patients reported by Zerres *et al*[33] (the same data used in the early onset model); beyond this timepoint, mortality is modelled using an HR-adjusted Gompertz function fitted to general population mortality data[32] (HR=26.41). This time-dependent 3-stage Type II SMA mortality function (zero risk [CHERISH]→2-knot spline [Zerres]→ HR-adjusted Gompertz[general population, HR=26.41]) is applied to all patients in the four worst health states (State [i] Sits without support but does not roll; State [ii] Sits and rolls independently; State [iii] Sits and crawls with hands and knees and State [iv] Stands/walks with assistance). Mortality risk for patients in the two better health states (State [vi] Stands unaided and State [vii] Walks unaided) is adjusted by a factor of 0.50 to reflect an assumption of improved survival associated with Type III SMA based on a Gompertz model fitted to general population mortality data (without HR adjustment).[32] After the end of month 15, all health state transitions are governed by a single transition matrix estimated using the HFMSE scores observed within the CHERISH trial,[15] Study CS2 and Study CS12.[25] This matrix permits nusinersen-treated patients to either remain in their current state or move to the next best health state, but does not allow for the deterioration of any patient's motor function from this timepoint onwards. Patients are assumed to discontinue nusinersen if they do not achieve milestones better than State (i) Sits without

support but does not roll by the end of month 15, or if they undergo scoliosis surgery (at year 12 for non-ambulatory patients and year 15 for ambulatory patients) and cannot subsequently undergo administration of nusinersen via lumbar puncture. Patients who discontinue nusinersen due to lack of efficacy are assumed to remain in State (i) Sits without support but does not roll state until death. Patients who discontinue nusinersen following scoliosis surgery are assumed to subsequently follow the post-trial transition matrix for the sham group.[3]

Usual care group

Patients enter the model based on the baseline HFMSE health state distribution for the sham group in CHERISH.[15] During the first five model cycles (up to the end of month 15), mortality risk is assumed to be zero, based on the observed number of deaths within the sham group of CHERISH. From model entry until the end of month 15, transitions between the six HFMSE-based health states are governed by five cycle-specific transition matrices derived from observed count data within CHERISH.

From month 15 onwards, mortality is modelled using the same data and assumptions as those applied within the nusinersen group, including the survival advantage assumed for States [v] and [vi]. After the end of month 15, all health state transitions are governed by a single transition matrix estimated using the HFMSE scores observed within CHERISH,[15] Study CS2 and Study CS12.[25] This matrix permits patients on usual care to either remain in their current state or to transit to the next worst health state, but does not allow for the improvement of any patient's motor function from this timepoint onwards (hence the survival advantage in the better two states only applies to those already in those states by the end of month 15). A proportion of patients are assumed to undergo scoliosis surgery at year 10 if non-ambulant and at year 15 if ambulant; however, this does not impact on the patient's health state occupancy, HRQoL or costs.

Estimation of health outcomes, costs and cost-effectiveness

Separate utilities are applied to each modelled health state. QALYs accrued by patients in each group are estimated by applying a vector of health utilities to the probability of being in each state during each model cycle. QALY losses for caregivers are estimated based on the patient's health state (including a QALY loss for bereavement). Analyses are presented separately which include/exclude caregiver QALY losses.

The model includes the following cost components: (i) acquisition and administration costs for nusinersen and (ii) health state costs, including respiratory, gastrointestinal, nutritional and orthopaedic care (conditional on motor milestones). In contrast with the early onset model, end-of-life care costs are not included.

Incremental cost-effectiveness is calculated in a pairwise fashion based on the difference in costs divided by the difference in QALYs for nusinersen and usual care.

5.4.2 Structural assumptions – later onset model

(i) Treatment using nusinersen is assumed to be discontinued if the patient has not progressed beyond State (i) Sits without support but does not roll state after 15 months. As with the early onset model, this assumption is applied only once as patients receiving nusinersen are assumed never to transit to this state after this timepoint (see assumption [v]).

(ii) A proportion of patients discontinue nusinersen following scoliosis surgery.

(iii) Patients cannot die in either treatment group until after month 15.

(iv) After month 15, an adjustment is applied to reflect improved survival for patients in State (v) Stands unaided and State (vi) Walks unaided. These patients are allocated 50% of the mortality risk for Type III SMA and 50% of the mortality risk for Type II SMA. Unlike the early onset model, this adjustment is applied to both the nusinersen and usual care groups.

(v) After month 15, patients receiving nusinersen are assumed never to transit to a worse health state; rather, during any model cycle, they can either remain in their current health state or transit to the next best health state. Beyond this timepoint, transition probabilities are based on the mean rate of improvement in HFMSE score within CHERISH and the mean HFMSE score within each model health state for the nusinersen group over the course of the CHERISH trial and Studies CS2 and CS12. The rate of improvement in HFMSE score is assumed to be constant with respect to time and monotonic across health states.

(vi) After month 15, patients receiving usual care are assumed never to transit to an improved health state; rather, during any model cycle, they can either remain in their current health state or transit to the next worst health state. Beyond this timepoint, transition probabilities are based on the mean rate of worsening in HFMSE score within CHERISH and the mean HFMSE scores within each model health state for the usual care group over the course of the CHERISH trial and Studies CS2 and CS12. The rate of worsening in HFMSE score is assumed to be constant with respect to time and monotonic across health states.

(vii) A proportion of ambulant patients undergo scoliosis surgery after 15 years.

(viii) The CS[1] states that the model assumes that a proportion of non-ambulant patients undergo scoliosis surgery at 12 years. However, the implemented model assumes that scoliosis surgery may occur at 10 years for the usual care group. As separate costs and utility changes for scoliosis surgery are not included in the model, this does not impact on the model results.

(ix) Treatment costs are grouped according to milestones consistent with Type II SMA ([i] Sits without support but does not roll; [ii] Sits and rolls independently; [iii] Sits and crawls with hands and knees; [iv] Stands/walks with assistance) and Type III SMA ([v] Stands unaided; [vi] Walks unaided).

(x) The model does not include additional HRQoL impacts or costs associated with AEs. The CS notes that the ENDEAR trial did not observe any treatment-related AEs.

5.4.3 Evidence used to inform model parameters – later onset model

The main groups of parameters for the later onset model and the evidence used to inform these are summarised in Table 55. These are discussed in further detail in the subsequent sections.

Table 55: Evidence used to inform the company's later onset model

Parameter group	Evidence source
Initial HFMSE health state distribution – nusinersen	Observed initial HFMSE distribution in the nusinersen group of CHERISH[15]
Initial HFMSE health state distribution – usual care	Observed initial HFMSE distribution in the sham group of CHERISH[15]
Overall survival – nusinersen	Zero risk (based on CHERISH) switching (after month 15) to a 2-knot Royston-Parmar spline model fitted to data reported by Zerres *et al*[33] switching (after month 623) to an HR-adjusted general population Gompertz model (HR=26.41).[32] For States (v) and (vi), after month 15, an adjustment of 0.50 is applied to reflect improved survival for Type III SMA based an unadjusted general population Gompertz model.[32]
Overall survival – usual care	Health state-dependent mortality probabilities are the same as those for the nusinersen group
Transition probabilities – nusinersen (up to month 15)	Observed HFMSE count data from CHERISH[15] (without imputation)
Transition probabilities – usual care (up to month 15)	Observed HFMSE count data from CHERISH[15] (without imputation)
Transition probabilities – nusinersen (month 16 onwards)	Estimated mean rate of improvement for nusinersen group in HFMSE and mean HFMSE scores in CHERISH[15] (supplemented using data from Study CS2 and Study CS12[25])
Transition probabilities – usual care (month 16 onwards)	Estimated mean rate of worsening for sham group in HFMSE and mean HFMSE scores in CHERISH[15] (supplemented using data from Study CS2 and Study CS12[25])
Probability of undergoing surgery for scoliosis and age at time of surgery	Probability based on estimate for scoliosis surgery in Type II SMA reported by Bladen *et al*.[49] Timing of surgery loosely based on Haaker and Fujak.[35]
Probability of discontinuing nusinersen after surgery for scoliosis	Assumption[1]
Patient utilities	PedsQL data collected in CHERISH[15] mapped to the EQ-5D using a published algorithm reported by Khan *et al*[36]
Baseline caregiver utilities	Baseline caregiver utility based on Bastida *et al*.[37] Caregiver disutilities by health state estimated using Ara and Brazier[38] and mapped patient utilities from CHERISH.[15]
Nusinersen acquisition cost	CS[1]
Nusinersen administration costs	NHS Reference Costs 2015/16[39]
Health state costs	Bastida *et al*[37]
End-of-life care costs	Not included in model

HFMSE - Hammersmith Functional Motor Scale-Expanded; HR – hazard ratio; SMA – spinal muscular atrophy; EQ-5D – Euroqol 5-Dimensions; CS – company's submission

Overall survival –later onset SMA

OS was modelled using similar approach to that adopted for the early onset model with separate sources used to inform each of the two different sections of the modelled time horizon. The overall modelling approach is summarised in Table 56. The company assumed that mortality risk for patients achieving State (v) Stands unaided and State (vi) Walks unaided would be between that of Type II SMA patients and the general population.

Model fit statistics for the Zerres *et al* data[33] and general population mortality data have been previously described in Section 5.2.4.

Table 56: Summary of survival models applied for extrapolation of overall survival

Time period	Both treatment groups	
	States (i) to (iv)	**States (v) and (vi)**
Month 0 to Month 15	CHERISH **No deaths**	
OS time period 1 Month 16 To Month 623	Zerres *et al*[33*] **2-knot spline**	Zerres *et al*[33*] **2-knot spline (weight 0.5)** UK general population mortality **unadjusted Gompertz** **(weight 0.5)**
OS time period 2 Month 623 to Month 960	UK general population mortality **HR-adjusted Gompertz** **(HR=26.4)**	UK general population mortality **unadjusted Gompertz** **(weight 0.5)** **HR-adjusted Gompertz** **(HR=26.4, weight 0.5)**

** N= 240, Type II SMA replicated from KM, no adjustment*
OS – overall survival; HR – hazard ratio

Transition probabilities

Similar to the early onset model, transition probabilities for the later onset model were estimated using different approaches for the observed period of CHERISH[15] and for subsequent cycles. Within the observed period, transitions were based directly on observed HFMSE count data for each treatment group. Separate matrices were calculated for five cycles (day 1-92, day 93-169, day 170-274, day 275-365 and day 366-456). All patients remained alive and none were lost to follow-up over the course of the trial.

Beyond the end of study follow-up, a single treatment-specific transition matrix is applied for all subsequent 4-monthly cycles. In contrast to the early onset model which attempts to map from the HINE-2 to CHOP INTEND, the later onset model uses HFMSE data from CHERISH to estimate milestone achievement/loss within the unobserved period (additional data from Study CS2 and CS12 were also used for State [vi] Walks unaided). Transition probabilities for patients in the nusinersen and

usual care groups were calculated using Equation [vi] and Equation [vii], respectively. The data used to estimate these transition probabilities are shown in Table 57.

$$TP\ (nusinersen)\ = MIN[1{,}1 + \left(\frac{Rate\ HFMSE\ increase\ (per\ month)\ .\ cycle\ length\ (months)}{Mean\ HFMSE\ next\ best\ state - Mean\ HFMSE\ current\ state}\right) \quad \text{[vi]}$$

$$TP(usual\ care)\ = MIN[1{,}1 + \left(\frac{Rate\ HFMSE\ decrease\ (per\ month)\ .\ cycle\ length\ (months)}{Mean\ HFMSE\ current\ state - Mean\ HFMSE\ next\ worst\ state}\right) \quad \text{[vii]}$$

Table 57: HFMSE data used to inform transition probabilities after month 15

MEAN HFMSE SCORE			
HFMSE health state	**Nusinersen**	**Sham**	**Source**
Sits without support but does not roll	17.7	15.9	CHERISH[15]
Sits and rolls independently	24.6	24.0	
Sits and crawls with hands and knees	34.5	26.7	
Stands/walks with assistance	38.4	26.7	
Stands unaided	40.3	31.5	
Walks unaided	51.0	38.8	CHERISH,[15] CS2 and CS12[25]
RATE OF IMPROVEMENT/WORSENING			
	Nusinersen	**Sham**	
Monthly HFMSE rate	[redacted]	[redacted]	CHERISH[15]

HFMSE - Hammersmith Functional Motor Scale-Expanded

Estimated transition probabilities for the first five model cycles (based on the observed count data from CHERISH) are shown in Table 58, Table 59, Table 60, Table 61 and Table 62. Table 63 presents the transition matrices applied for each 4-month cycle after the end of month 15.

Table 58: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 1-92

NUSINERSEN GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]		[redacted]			
Sits and rolls independently	[redacted]	[redacted]					
Sits and crawls with hands and knees	[redacted]		[redacted]				
Stands/walks with assistance		[redacted]		[redacted]	[redacted]		
Stands unaided					[redacted]		
Walks unaided						[redacted]	
Dead							[redacted]
SHAM GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]					
Sits and rolls independently	[redacted]	[redacted]					
Sits and crawls with hands and knees		[redacted]	[redacted]				
Stands/walks with assistance				[redacted]	[redacted]		
Stands unaided				[redacted]	[redacted]	[redacted]	
Walks unaided						[redacted]	
Dead							[redacted]

** No observed transitions from state during cycle; Blank cells indicate zero probability*
N - number

Table 59: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 93-169

NUSINERSEN GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]	[redacted]	[redacted]			
Sits and rolls independently	[redacted]	[redacted]	[redacted]	[redacted]			
Sits and crawls with hands and knees			[redacted]				
Stands/walks with assistance				[redacted]	[redacted]		
Stands unaided					[redacted]		
Walks unaided						[redacted]	
Dead							[redacted]
SHAM GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]		[redacted]			
Sits and rolls independently		[redacted]					
Sits and crawls with hands and knees			[redacted]				
Stands/walks with assistance	[redacted]			[redacted]		[redacted]	
Stands unaided				[redacted]	[redacted]	[redacted]	
Walks unaided					[redacted]		
Dead							[redacted]

** No observed transitions from state during cycle; Blank cells indicate zero probability*
N - number

Table 60: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 170-274

NUSINERSEN GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]					
Sits and rolls independently	[redacted]	[redacted]	[redacted]	[redacted]			
Sits and crawls with hands and knees			[redacted]	[redacted]			
Stands/walks with assistance			[redacted]	[redacted]			
Stands unaided					[redacted]		
Walks unaided						[redacted]	
Dead							[redacted]
SHAM GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]					
Sits and rolls independently	[redacted]	[redacted]					
Sits and crawls with hands and knees		[redacted]	[redacted]				
Stands/walks with assistance	[redacted]			[redacted]			
Stands unaided					[redacted]	[redacted]	
Walks unaided				[redacted]		[redacted]	
Dead							[redacted]

** No observed transitions from state during cycle; Blank cells indicate zero probability*

N - number

Table 61: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 275-365

NUSINERSEN GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]					
Sits and rolls independently	[redacted]	[redacted]					
Sits and crawls with hands and knees			[redacted]	[redacted]			
Stands/walks with assistance			[redacted]	[redacted]			
Stands unaided					[redacted]	[redacted]	
Walks unaided						[redacted]	
Dead							[redacted]
SHAM GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]						
Sits and rolls independently	[redacted]	[redacted]					
Sits and crawls with hands and knees	[redacted]	[redacted]					
Stands/walks with assistance				[redacted]	[redacted]		
Stands unaided					[redacted]		
Walks unaided						[redacted]	
Dead							[redacted]

** No observed transitions from state during cycle; Blank cells indicate zero probability*
N - number

Table 62: Transition matrices for nusinersen (top) and sham (bottom), HFMSE observed count data, CHERISH trial, days 366-456

NUSINERSEN GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]					
Sits and rolls independently	[redacted]	[redacted]	[redacted]				
Sits and crawls with hands and knees	[redacted]	[redacted]	[redacted]	[redacted]			
Stands/walks with assistance			[redacted]	[redacted]			
Stands unaided					[redacted]		
Walks unaided						[redacted]	
Dead							[redacted]
SHAM GROUP (patients alive with data n=[redacted])							
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided	Dead
Sits without support but does not roll	[redacted]	[redacted]					
Sits and rolls independently	[redacted]	[redacted]					
Sits and crawls with hands and knees			[redacted]				
Stands/walks with assistance	[redacted]			[redacted]			
Stands unaided					[redacted]		
Walks unaided					[redacted]	[redacted]	
Dead							[redacted]

** No observed transitions from state during cycle; Blank cells indicate zero probability*
N - number

Table 63: Transition matrices for nusinersen (top) and sham (bottom), extrapolation based on HFMSE score in CHERISH trial, all 4-month cycles after month 15

NUSINERSEN GROUP (patients alive with data n=n/a)						
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll						
Sits and rolls independently						
Sits and crawls with hands and knees						
Stands/walks with assistance						
Stands unaided						
Walks unaided						
SHAM GROUP (patients alive with data n=n/a)						
From\To state	Sits without support but does not roll	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll						
Sits and rolls independently						
Sits and crawls with hands and knees						
Stands/walks with assistance						
Stands unaided						
Walks unaided						

Blank cells indicate zero probability
n/a - not applicable

Probability of undergoing surgery for scoliosis and age at time of surgery

The assumptions regarding the timing of scoliosis surgery and the probability of discontinuing nusinersen treatment within the later onset model are the same as those for the early onset model (see Section 5.3.3). One clinical advisor to the ERG noted that patients with Type II SMA would typically undergo scoliosis surgery earlier than assumed in the model. The later onset model assumes that 43% of patients undergo scoliosis surgery at each assumed surgery timepoint, based on a survey-based study reported by Bladen *et al.*[49]

HRQoL - patient utilities

The source and derivation of the health state utility values in the later onset model are the same as those for the early onset model, albeit based on different health state descriptions (see Table 43).

HRQoL - caregiver utilities

Within the later onset model, caregiver disutilities were estimated using a similar approach and the same data as those used in the early onset model. The derivation of each health state-specific disutility is shown in Table 64.

Table 64: Parent/carer utilities used in the later onset model

HFMSE health state	Patient utility	Caregiver utility*	Caregiver disutility	Calculation and assumptions
Sits without support but does not roll	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (ii) Sits and rolls independently and State (i) Sits without support but does not roll
Sits and rolls independently	[redacted]	[redacted]	[redacted]	Based on weighted mean of Type II and Type III caregiver utility reported by Bastida *et al*[37]
Sits and crawls with hands and knees	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (ii) Sits and rolls independently and State (iii) Sits and crawls with hands and knees
Stands/walks with assistance	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (ii) Sits and rolls independently and State (iv) Stands/walks with assistance
Stands unaided	[redacted]	[redacted]	[redacted]	Assumed to be the same as State (iv) Stands/walks with assistance
Walks unaided	[redacted]	[redacted]	[redacted]	Bastida *et al*[37] baseline caregiver utility minus difference between State (ii) Sits and rolls independently and State (vi) Walks unaided. Disutility constrained at zero.
Baseline parameters				
Bastida *et al*[37] caregiver utility		[redacted]	-	-
General population utility[38]		0.92	-	Caregiver age=30.88 years, 80% female
Bereavement		-	-0.04	-

** Calculated as general population utility minus caregiver utility*

Resource use and costs

The company's later onset model includes the following cost components: (i) nusinersen acquisition and administration costs and (ii) health state costs. End-of-life care costs are not included in the later onset model.

Drug acquisition and administration costs

As with the early onset model, the cost of nusinersen is assumed to be £75,000 per vial. As noted in Chapter 3, the model assumes that nusinersen is given as four loading doses during the first 3-month cycle, with 4-monthly maintenance doses thereafter, based on the licensed treatment schedule[4] rather than the treatment schedule used in CHERISH.[15] Nusinersen administration costs are based on the same age-based calculations as those used in the early onset model (see Section 5.3.3).

Health state costs

Consistent with the early onset model, health state costs are based on estimates reported in Bastida *et al*[37] (see Table 65).

Table 65: Annual health state costs, later onset model

Cost component	Milestones consistent with Type II SMA ([i] Sits without support but does not roll; [ii] Sits and rolls independently; [iii] Sits and crawls with hands and knees; [iv] Stands/walks with assistance)	Milestones consistent with Type III SMA ([v] Stands unaided; [vi] Walks unaided)
Respiratory care	[redacted]	[redacted]
Gastrointestinal care	[redacted]	[redacted]
Nutritional care	[redacted]	[redacted]
Orthopaedic care	[redacted]	[redacted]
Total	[redacted]	[redacted]

SMA – spinal muscular atrophy

5.4.4 Methods for model evaluation

The CS[1] presents the results of the later onset model in terms of the incremental cost per QALY gained for nusinersen versus usual care. Results are presented separately for: (i) analyses including patient health gains only and (ii) analyses including patient health gains and caregiver QALY losses. The company's base case ICERs are based on the deterministic version of the model. The CS also includes the results of PSA, DSAs, scenario analyses and subgroup analyses. The results of the PSA are presented in the form of a cost-effectiveness plane and CEACs, based on 1,000 Monte Carlo simulations. The probabilistic ICER is also presented. The distributions applied in the company's PSA are summarised

in Table 66. The results of the DSAs are presented in the form of a tornado diagram for specified model parameters. Scenario analyses were undertaken to explore the impact of alternative time horizons and alternative assumptions surrounding mortality risk, transition probabilities and costs; no scenario analyses are presented around HRQoL estimates.

Table 66: Distributions used in company's PSA, later onset model

Parameter group	Distribution	ERG comment
Initial HFMSE health state distribution – nusinersen	Fixed	The initial distributions are subject to uncertainty. Given the multinomial nature of the data, a Dirichlet distribution (applied to the combined CHERISH population) would be appropriate.
Initial HFMSE health state distribution – usual care	Fixed	
Overall survival – nusinersen	Multivariate normal	-
Overall survival – usual care	Multivariate normal	-
Transition probabilities – nusinersen (up to month 15)	Dirichlet	Priors are included for some but not all unobserved transitions.
Transition probabilities – usual care (up to month 15)	Dirichlet	
Transition probabilities – nusinersen (month 16 onwards)	Dirichlet	
Transition probabilities – usual care (month 16 onwards)	Dirichlet	
Probability of undergoing surgery for scoliosis	Beta	Inappropriately characterised using treatment-specific parameters.
Age at time of surgery	Normal	Inappropriately characterised using treatment-specific parameters.
Probability of discontinuing nusinersen after surgery for scoliosis	Beta	-
Patient utilities	Beta	All utilities sampled using the same random number, thereby inducing over-correlation between states.[47]
Baseline caregiver utilities	Beta	No uncertainty is included in the Bastida *et al*[37] baseline caregiver disutility
Nusinersen acquisition cost	Fixed	-
Nusinersen administration costs	Normal (cost) and Dirichlet (administration setting)	-
Health state costs	Gamma	-

HFMSE - Hammersmith Functional Motor Scale-Expanded; ERG – Evidence Review Group

5.4.5 Company's cost-effectiveness results – later onset model

This section presents the results of the company's later onset model.

Central estimates of cost-effectiveness – later onset model

Table 67 presents the central estimates of cost-effectiveness derived from the company's updated model (including patient health gains only). Based on a re-run of the probabilistic version of the model by the ERG, nusinersen is expected to generate an additional 2.28 QALYs at an additional cost of £2,938,441 per patient: the corresponding ICER for nusinersen versus usual care is £1,286,149 per QALY gained. The deterministic version of the model produces a slightly lower ICER of £1,252,991 per QALY gained for nusinersen versus usual care. The inclusion of caregiver QALY losses leads to a markedly lower probabilistic ICER of £933,088 per QALY gained (see Table 68); the deterministic ICER is lower at £898,164 per QALY gained.

Table 67: Company's model results, later onset model (patient health gains only)

Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	16.85	£3,120,835	2.28	£2,938,441	**£1,286,149**
Usual care	14.56	£182,394	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	16.88	£3,148,754	2.37	£2,964,442	**£1,252,991**
Usual care	14.52	£184,312	-	-	-

Inc. – incremental; QALY – quality-adjusted life year

Table 68: Company's model results, later onset model (patient health gains and caregiver QALY losses)

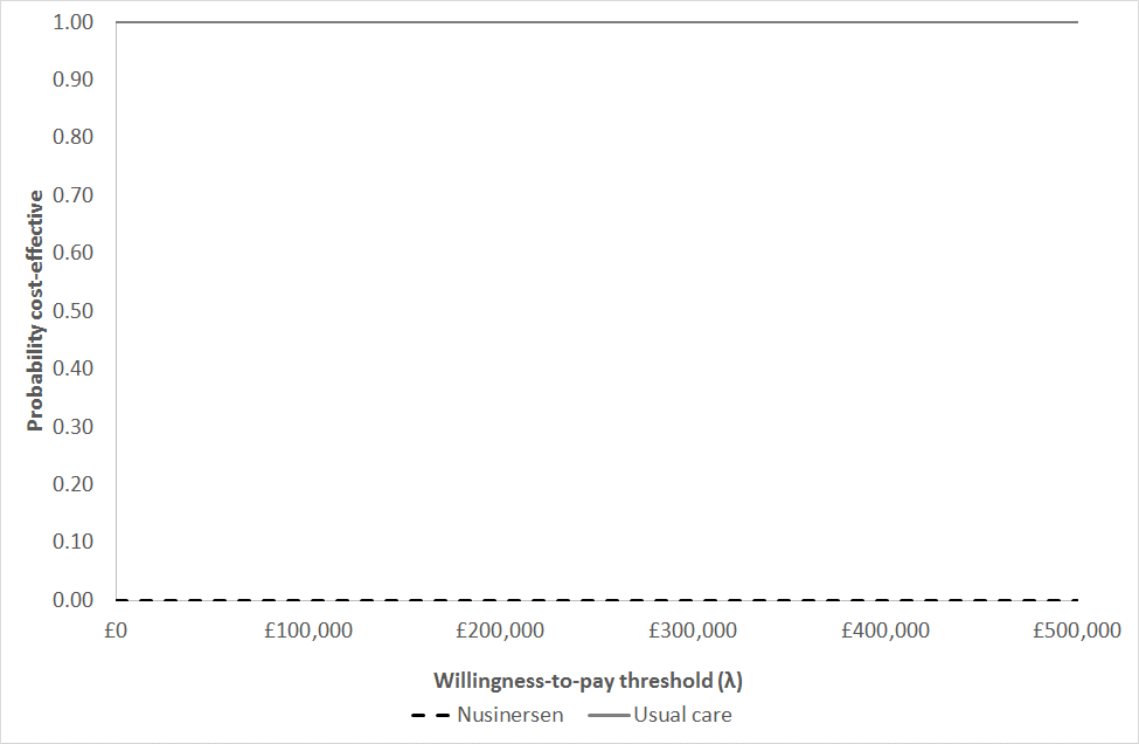
Probabilistic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	15.65	£3,120,835	3.15	£2,938,441	**£933,088**
Usual care	12.50	£182,394	-	-	-
Deterministic model					
Option	**QALYs**	**Costs**	**Inc. QALYs**	**Inc. Costs**	**Incremental cost per QALY gained**
Nusinersen	15.66	£3,148,754	3.30	£2,964,442	**£898,164**
Usual care	12.36	£184,312	-	-	-

Inc. – incremental; QALY – quality-adjusted life year

Company's probabilistic sensitivity analysis - later onset model

Figure 11 presents CEACs for nusinersen and usual care for the later onset population. As shown in the figure, the probability that nusinersen produces more net benefit than usual care is approximately zero even at WTP thresholds of £500,000 per QALY gained.

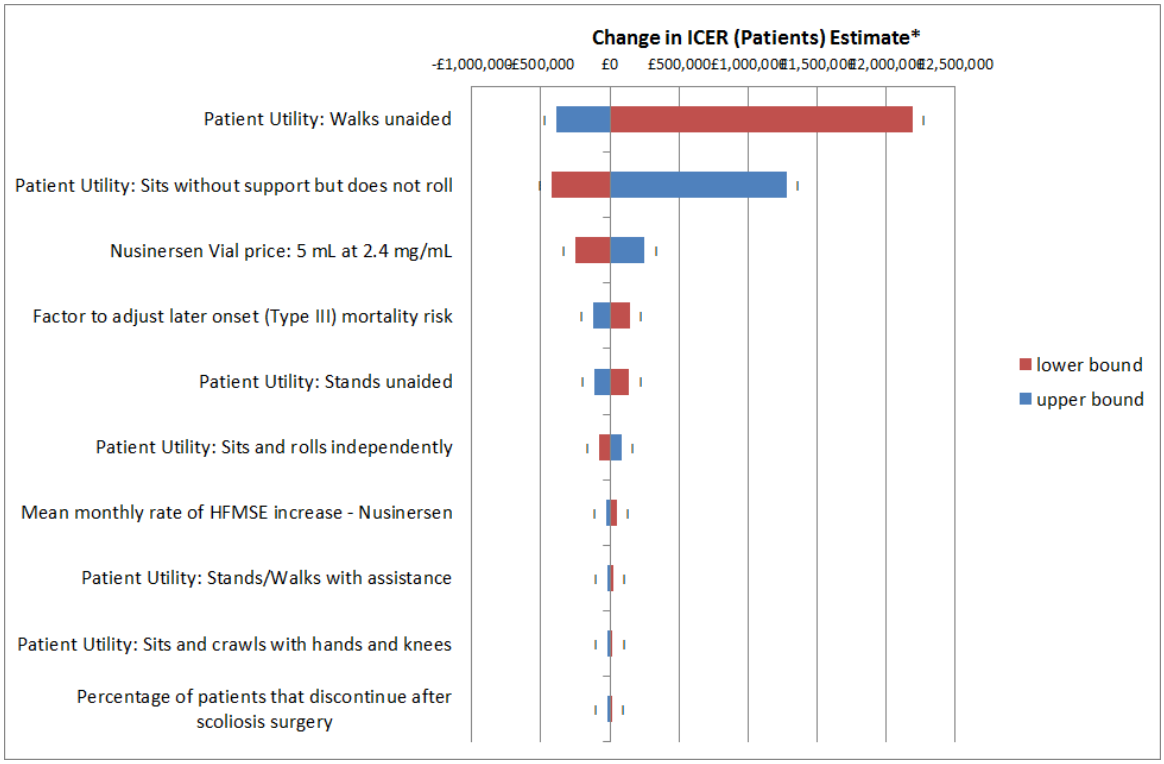
Figure 11: CEACs, later onset model, patient health gains only



Company's deterministic sensitivity analyses - later onset model

Figure 12 presents the results of the company's DSAs in the form of a tornado diagram (change in ICER from baseline). As shown in the figure, the most influential model parameters relate to the patient utility values for State (vi) Walks unaided and for State (i) Sits without support but does not roll. The lowest ICER generated from the company's one-way DSAs is £832,517 per QALY gained (patient utility for State [i] Sits without support but does not roll = [redacted]) whilst the highest ICER is £3,445,079 per QALY gained (patient utility for State [vi] Walks unaided = [redacted]).

Figure 12: Company's DSA tornado diagram, later onset model, patient health gains only



HFMSE - Hammersmith Functional Motor Scale-Expanded; ICER – incremental cost-effectiveness ratio; HR – hazard ratio

Scenario analysis results - later onset model

Table 69 presents the results of the company's scenario analyses. As shown in the table, the ICER for nusinersen is highly sensitive to the assumptions regarding mortality risk in the best two health states and the model time horizon. The lowest ICER for nusinersen versus usual care is estimated to be £734,749 per QALY gained when only patient health gains are included, and £614,044 per QALY gained when caregiver QALY losses are included in the analysis. These ICERs relate to the scenario in which general population mortality risk is attributed to all patients in States (v) and (vi) (mortality adjustment factor = 1.00). The highest ICER for nusinersen versus usual care is estimated to be £2,394,639 per QALY gained when only patient health gains are included, and £1,473,743 per QALY gained when caregiver disutilities are included in the analysis; these ICERs relate to the scenario in which the time horizon is truncated at 20 years.

Table 69: Scenario analysis results, later onset model

Scenario	ICER (patient health gains only)	ICER (patient health gains and caregiver QALY losses)
Base case (deterministic)	£1,252,991	£898,164
Time horizon=20 years	£2,394,639	£1,473,743
Time horizon=40 years	£1,528,733	£1,027,641
Time horizon=60 years	£1,280,983	£911,120
Societal cost perspective	£1,150,976	£825,038
Do not apply higher long-term risk of death based on SMA Type II adjusted general mortality rates	£1,227,736	£886,694
Do not apply general population mortality rates to patients in motor milestones characteristic of later onset (Type III) patients	£2,324,278	£1,285,987
Mortality risk factor=0.75	£969,170	£753,553
Mortality risk factor=1.00	£734,749	£614,044
Assumption a proportion of patients on treatment reach a plateau; 0% of those reaching an improvement plateau start getting worse	£1,371,100	£983,437
Assumption a proportion of patients on treatment reach a plateau; 10% of those reaching an improvement plateau start getting worse	£1,393,262	£997,921
Usual care arm HFMSE rate of decline based on Kaufmann *et al*[50]	£1,268,258	£911,947
All nusinersen administration inpatient	£1,258,656	£902,225
All nusinersen administration day case	£1,255,928	£900,269
Health state costs includes costs of major clinical events only	£1,276,308	£914,878
Cost estimates based on Klug *et al*[51]	£1,258,136	£901,852

SMA – spinal muscular atrophy; HFMSE - Hammersmith Functional Motor Scale-Expanded; ICER – incremental cost-effectiveness ratio

Subgroup analysis - later onset model

Table 70 presents the results of the company's subgroup analyses for the later onset model (disease duration <25 months or ≥25 months). The results suggest that the ICER for nusinersen versus usual care is less favourable in patients with longer disease duration (≥25 months).

Table 70: Subgroup analysis results, later onset model

Subgroup	ICER (patient health gains only)	ICER (patient health gains and caregiver QALY losses)
ITT population, each arm (base case)*	£1,252,991	£898,164
ITT population, both arms (base case) †	£1,265,944	£924,891
<25 months disease duration, each arm*	£1,263,457	£892,985
<25 months disease duration, both arms†	£1,201,673	£863,535
≥25 months disease duration, each arm*	£1,712,437	£1,220,287
≥25 months disease duration, both arms†	£1,615,299	£1,165,000

** "thresholds" defining HFMSE health states based on mean scores in each treatment group; † "thresholds" defining HFMSE health states based on mean scores across both treatment groups*

5.5 Critical appraisal of the company's health economic analyses

This section presents a critical appraisal of the health economic analyses of nusinersen for the treatment of early onset and later onset SMA presented within the CS.[1] Section 5.5.1 details the methods used by the ERG to interrogate and critically appraise the company's submitted health economic analyses. Section 5.5.2 discusses the extent to which the company's analyses adhere to the NICE Reference Case. Section 5.5.3 presents a detailed critique of the ERG's main issues and concerns relating to the company's analyses.

5.5.1 Methods for reviewing the company's health economic analyses

The ERG adopted a number of approaches to explore, interrogate and critically appraise the company's submitted economic analyses and the underlying health economic models upon which these were based. These included:

- Consideration of key items contained within published economic evaluation and health economic modelling checklists[52, 53] to critically appraise the company's models and analyses.
- Scrutiny of the company's models by health economic modellers and discussion of issues identified amongst the members of the ERG.
- Double-programming of the deterministic version of the company's models to fully assess the logic of the company's model structures, to draw out any unwritten assumptions and to identify any apparent errors in model implementation.
- Examination of the correspondence between the description of the models reported within the CS[1] and the executable models.
- Replication of the base case results, PSAs, DSAs and scenario analyses presented within the CS.[1]
- Where possible, checking of parameter values used in the company's models against their original data sources.
- The use of expert clinical input to judge the credibility of the company's assumptions underpinning the company's models.

5.5.2 Adherence of the company's economic analyses to the NICE Reference Case (early and later onset models)

The company's economic analyses of nusinersen for the treatment of early onset and later onset SMA are partially in line with the NICE Reference Case.[54] The ERG notes that the analyses exclude patients with Type 0 and Type IV SMA; patients with these SMA types are included in the marketing authorisation and the final NICE scope.[12] In addition, the evidence used to inform the clinical effectiveness evidence for nusinersen and the longer-term prognosis of patients with SMA are not based on formal systematic reviews. These issues are discussed in further detail in Section 5.5.3.

Table 71: Adherence of the company's economic analyses to the NICE Reference Case (early onset and later onset models)

Element	Reference case	ERG comments
Defining the decision problem	The scope developed by NICE[12]	The company's economic analyses relate to the ITT populations of the ENDEAR study[14] (Type I SMA) and the CHERISH study[15] (Types II and III SMA). Taken together, this population is narrower than the population defined in the final NICE scope and the marketing authorisation for nusinersen (people with 5q SMA). No economic evidence is presented for patients with Type 0 or Type IV SMA. The ERG notes that the model states are defined according to motor function milestones which may not fully capture the impact of other outcomes defined in the NICE scope[12] (e.g. respiratory function and the requirement for ventilation).
Comparator(s)	As listed in the scope developed by NICE	The company's economic analyses define the comparator as real world care (symptomatic or usual care), based on the sham arms of the ENDEAR and CHERISH trials[14, 15] and use observational data to inform survival outcomes beyond trial follow-up. The scope defines the comparator as BSC. The ERG and its clinical advisors consider this to be reasonable but note that there may be variation in how Type I SMA patients are managed, which may lead to differences between observed and predicted survival estimates. The ERG's clinical advisors commented that in the real world, the ability to provide BSC and the choices made by families may differ from a clinical trial situation. They also noted that families entering into trials are likely to be more motivated in seeking proactive support for their infants/children than many in routine clinical care.
Perspective on outcomes	All direct health effects, whether for patients or, when relevant, carers	Health gains accrued by patients are valued in terms of QALYs gained. Additional analyses are presented including QALY losses for caregivers.
Perspective on costs	NHS and PSS	The analysis adopts an NHS and PSS perspective.
Type of economic evaluation	Cost-utility analysis with fully incremental analysis	The results of the analyses are presented in terms of the incremental cost per QALY gained for nusinersen versus usual care.
Time horizon	Long enough to reflect all important differences in costs or outcomes between the technologies being compared	The early onset model adopts a 60-year time horizon. The later onset model adopts an 80-year time horizon. Within both models, approximately 100% of patients have died by the end of the modelled time horizon.
Synthesis of evidence on health effects	Based on systematic review	The company did not undertake a systematic review of clinical effectiveness evidence. The model is informed by the pivotal RCTs of nusinersen[14, 15] as well as observational data.[31-33] The methods for identifying these observational studies are unclear from the CS.[1]

Element	Reference case	ERG comments
Measuring and valuing health effects	Health effects should be expressed in QALYs. The EQ-5D is the preferred measure of HRQoL in adults.	Patient utilities were derived by mapping the PedsQL data from the CHERISH trial[15] to the EQ-5D using an algorithm reported by Khan *et al.*[36] Health utilities for the early onset model were based on an assumed correspondence between the HFMSE and HINE-2 defined health states. The mapping algorithm was derived using valuations from healthy schoolchildren. The ERG has concerns regarding the validity of these estimates and notes that alternative sources are available, although these are also subject to issues concerning face validity. Caregiver utilities were based on a single estimate from Bastida *et al*[37] and a large number of assumptions using the mapped patient utilities from CHERISH.[15]
Source of data for measurement of HRQoL	Reported directly by patients and/or carers	
Source of preference data for valuation of changes in HRQoL	Representative sample of the UK population	
Equity considerations	An additional QALY has the same weight regardless of the other characteristics of the individuals receiving the health benefit	No additional equity weighting is applied to estimated QALY gains. The CS[1] argues that nusinersen meets NICE's end-of-life criteria within the early onset population.
Evidence on resource use and costs	Costs should relate to NHS and PSS resources and should be valued using the prices relevant to the NHS and PSS	Resource components included in the company's models reflect those relevant to the NHS and PSS. Unit costs were valued at 2015/16 prices.
Discount rate	The same annual rate for both costs and health effects (currently 3.5%)	Costs and health effects are discounted at a rate of 3.5% per annum.

SMA - spinal muscular atrophy; ITT - intention-to-treat; QALY - quality-adjusted life year; HRQoL - health-related quality of life; PSS - Personal Social Services; EQ-5D - Euroqol 5-Dimensions; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; HFMSE - Hammersmith Functional Motor Scale-Expanded; NICE - National Institute for Health and Care Excellence; ERG - Evidence Review Group

5.5.3 Main issues identified within the critical appraisal

Box 1 summarises the main issues identified within the ERG's critical appraisal of the company's economic analyses. These issues are discussed in further detail in the subsequent sections.

Box 1: Main issues identified within the critical appraisal undertaken by the ERG

(1) Absence of economic evidence relating to Type 0 and Type IV SMA
(2) Model verification, errors and complexity of programming approach
(3) Concerns regarding model structures which focus only on motor milestones
(4) Highly favourable assumptions regarding the expected trajectory of nusinersen-treated patients through modelled motor milestone health states
(5) Highly favourable assumptions regarding the expected survival of nusinersen-treated patients
(6) Issues relating to estimated patient utilities
(7) Arbitrary calculations underpinning caregiver disutilities
(8) Issues relating to health state costs
(9) Representation of uncertainty

(1) Absence of economic evidence relating to Type 0 and Type IV SMA

The marketing authorisation for nusinersen states that treatment is indicated for the treatment of 5q SMA.[4] This population is also defined in the final NICE scope.[12] The company's early onset model relates to patients with Type I SMA, whilst the later onset model relates to patients with Types II and III SMA. The CS does not present any economic analyses for patients with Type 0 or Type IV SMA. With respect to this issue, the CS states: *"Patients with type 0 and type IV (adult onset) SMA are omitted from the submission, despite market authorisation,(1) as there is no clinical evidence for nusinersen in type 0 and type IV that meets the requirements for technology appraisal at the current time"* (CS,[1] page 9). However, the CS[1] (page 21) also states that the anticipated place of nusinersen in therapy is as first-line treatment for all SMA patients. The ERG's clinical advisors stated that they would not treat Type 0 SMA patients with nusinersen, except in the context of clinical trials. The advisors also stated that they would not treat Type IV SMA patients using nusinersen as it is unlikely that these patients would obtain benefit from treatment.

(2) Model verification, errors and complexity of programming approach

Concerns regarding complexity of the company's model implementation

The company's models were programmed in such a complex way that the key formulae (including the Markov trace) were largely impenetrable to the ERG. This caused significant problems for the ERG not only in terms of verifying that the model had been implemented as intended and without error, but more fundamentally in terms of understanding what assumptions had been applied within the models. The extent of these issues is evident from a single Markov trace calculation in the nusinersen group of the early onset model (see Box 2). This formula includes 14 =IF() statements, 38 =SUMPRODUCT() functions and 73 =TRANSPOSE() functions. The early onset model includes several hundred similar equations to calculate the Markov trace for the nusinersen group. The trace calculations for patients who have undergone scoliosis surgery are approximately twice as long as the example given in Box 2. The later onset model is also subject to similar complicated programming issues.

Box 2: Example formula from a single cell of the company's early onset model trace

*=IF(txt_disc=2,IF(os_f_type2=2,SUMPRODUCT(TRANSPOSE($F25:$M25),IF($C26<=$GF$7,F$419:F$426,F$433:F$440))*(1-$DE26),SUMPRODUCT(TRANSPOSE($F25:$H25),IF($C26<=$GF$7,F$419:F$421,F$433:F$435))*(1-$DE26)+SUMPRODUCT(TRANSPOSE($I25:$L25),IF($C26<=$GF$7,F$422:F$425,F$436:F$439))*(1-$DA26)+$M25*IF($C26<=$GF$7,F$426,F$440)*(1-$DE26)),IF($C26<=$GO$12,IF(os_f_type2=2,(SUMPRODUCT(TRANSPOSE($F25:$M25),TRANSPOSE($BH25:$BO25),F$419:F$426)+SUMPRODUCT(TRANSPOSE($F25:$M25),TRANSPOSE($BQ25:$BX25),F$458:F$465))*(1-$DE26),(SUMPRODUCT(TRANSPOSE($F25:$H25),TRANSPOSE($BH25:$BJ25),F$419:F$421)+SUMPRODUCT(TRANSPOSE($F25:$H25),TRANSPOSE($BQ25:$BS25),F$458:F$460))*(1-$DE26)+(SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BK25:$BN25),F$422:F$425)+SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BT25:$BW25),F$461:F$464))*(1-$DA26)+$M25*($BO25+$BX25)*F$426*(1-$DE26)),IF(HS_Stop_txt=1,IF(os_f_type2=2,(SUMPRODUCT(TRANSPOSE($G25:$M25),TRANSPOSE($BI25:$BO25),F$420:F$426)+SUMPRODUCT(TRANSPOSE($G25:$M25),TRANSPOSE($BR25:$BX25),F$459:F$465))*(1-$DE26)+$F25*($BH25+$BQ25)*F$419*(1-$DC26),(SUMPRODUCT(TRANSPOSE($G25:$H25),TRANSPOSE($BI25:$BJ25),F$420:F$421)+SUMPRODUCT(TRANSPOSE($G25:$H25),TRANSPOSE($BR25:$BS25),F$459:F$460))*(1-$DE26)+(SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BK25:$BN25),F$422:F$425)+SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BT25:$BW25),F$461:F$464))*(1-$DA26)+$M25*($BO25+$BX25)*F$426*(1-$DE26)+$F25*($BH25+$BQ25)*F$419*(1-$DC26)),IF(HS_Stop_txt=2,IF(os_f_type2=2,(SUMPRODUCT(TRANSPOSE($H25:$M25),TRANSPOSE($BJ25:$BO25),F$421:F$426)+SUMPRODUCT(TRANSPOSE($H25:$M25),TRANSPOSE($BS25:$BX25),F$460:F$465))*(1-$DE26)+(SUMPRODUCT(TRANSPOSE($F25:$G25),TRANSPOSE($BH25:$BI25),F$419:F$420)+SUMPRODUCT(TRANSPOSE($F25:$G25),TRANSPOSE($BQ25:$BR25),F$458:F$459))*(1-$DC26),(SUMPRODUCT(TRANSPOSE($H25),TRANSPOSE($BJ25),F$421)+SUMPRODUCT(TRANSPOSE($H25),TRANSPOSE($BS25),F$460))*(1-$DE26)+(SUMPRODUCT(TRANSPOSE($F25:$G25),TRANSPOSE($BH25:$BI25),F$419:F$420)+SUMPRODUCT(TRANSPOSE($F25:$G25),TRANSPOSE($BQ25:$BR25),F$458:F$459))*(1-$DC26)+(SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BK25:$BN25),F$422:F$425)+SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BT25:$BW25),F$461:F$464))*(1-$DA26)+$M25*($BO25+$BX25)*F$426*(1-$DE26)),IF(os_f_type2=2,(SUMPRODUCT(TRANSPOSE($G25:$L25),TRANSPOSE($BI25:$BN25),F$420:F$425)+SUMPRODUCT(TRANSPOSE($G25:$L25),TRANSPOSE($BR25:$BW25),F$459:F$464))*(1-$DE26)+(SUMPRODUCT(TRANSPOSE($F25),TRANSPOSE($BH25),F$419)+SUMPRODUCT(TRANSPOSE($F25),TRANSPOSE($BQ25),F$458))*(1-$DC26)+$M25*($BO25+$BX25)*F$426*(1-$DC26),(SUMPRODUCT(TRANSPOSE($G25:$H25),TRANSPOSE($BI25:$BJ25),F$420:F$421)+SUMPRODUCT(TRANSPOSE($G25:$H25),TRANSPOSE($BR25:$BS25),F$459:F$460))*(1-$DE26)+(SUMPRODUCT(TRANSPOSE($F25),TRANSPOSE($BH25),F$419)+SUMPRODUCT(TRANSPOSE($F25),TRANSPOSE($BQ25),F$458))*(1-$DC26)+(SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BK25:$BN25),F$422:F$425)+SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BT25:$BW25),F$461:F$464))*(1-$DA26)+$M25*($BO25+$BX25)*F$426*(1-$DC26)))))+IF(os_f_type2=2,SUMPRODUCT(TRANSPOSE($F25:$M25),TRANSPOSE($AY25:$BF25),F$445:F$452)*(1-$DC26),SUMPRODUCT(TRANSPOSE($F25:$H25),TRANSPOSE($AY25:$BA25),F$445:F$447)*(1-$DC26)+SUMPRODUCT(TRANSPOSE($I25:$L25),TRANSPOSE($BB25:$BE25),F$448:F$451)*(1-$DA26)+$M25*$BF25*F$452*(1-$DC26)))*

The ERG sought clarification regarding the justification for the company's programming approach (see clarification response,[2] question B20). In response, the company acknowledged that the formulae are unnecessarily complicated, but noted that: (a) the model had to be developed "from scratch" due to the absence of existing economic models of treatments for SMA, and (b) the model was developed iteratively and became more complex due to the inclusion of elements such as scoliosis surgery and different model extrapolation approaches. The ERG does not consider that either of these explanations presents a sufficient justification for the complicated programming approach adopted.

Double-programming of the company's early onset and later onset models

During the early stages of the appraisal, the ERG raised concerns with NICE regarding the complex implementation of the company's models. In response, the company held a tutorial telephone call with the ERG and NICE which helped to clarify the intended logic and assumptions of the models. Subsequently, the ERG was able to double-program simplified versions of the Markov traces for both treatment groups in the early and later onset models (excluding the possibility of scoliosis surgery, thereby reducing the complexity of both models). In addition, the ERG was able to use the Markov traces generated from the company's models to replicate the remaining model structure and to estimate ICERs for both the early and later onset SMA populations. The results of these two double-programming exercises are shown in Table 72 and Table 73.

Table 72: Comparison of the company's model and the ERG's double-programmed Markov traces, end of trial follow-up to end of time horizon (excludes the possibility of scoliosis surgery)

Health state	Mean health state sojourn time (years, from month 13-end of time horizon)			
	Nusinersen		Usual care	
	Company's Markov trace	ERG's double-programmed Markov trace	Company's Markov trace	ERG's double-programmed Markov trace
Early onset model				
No milestone achieved	2.55	2.55	9.35	9.35
Mild milestone	0.19	0.19	0.06	0.06
Moderate milestone	0.48	0.48	0.05	0.05
Sits without support	0.78	0.78	0.00	0.00
Stands with assistance	1.45	1.45	0.00	0.00
Walks with assistance	1.44	1.44	0.00	0.00
Stands/walks unaided	29.47	29.47	0.00	0.00
Later onset model				
Sits without support but does not roll	14.56	14.56	31.21	31.22
Sits and rolls independently	0.77	0.77	2.61	2.61
Sits and crawls with hands and knees	0.47	0.47	0.45	0.45
Stands/walks with assistance	0.29	0.29	0.03	0.03
Stands unaided	1.86	1.86	0.53	0.53
Walks unaided	22.39	22.38	0.20	0.20

ERG - Evidence Review Group

Table 73: Comparison of the company's model results and the ERG's estimated ICERs using the company's Markov traces

	Company's model			ERG's double-programmed model		
Early onset model						
Option	**Inc. QALYs**	**Inc. costs**	**ICER**	**Inc. QALYs**	**Inc. costs**	**ICER**
Nusinersen	7.86	£2,258,362	**£407,679**	7.86	£2,272,097	**£410,240**
Usual care	2.49	£71,540	-	2.49	£71,540	-
Later onset model						
Option	**Inc. QALYs**	**Inc. costs**	**ICER**	**Inc. QALYs**	**Inc. costs**	**ICER**
Nusinersen	16.88	£3,148,754	**£1,252,991**	16.88	£3,299,874	**£1,315,176**
Usual care	14.52	£184,312	-	14.52	£188,309	

Inc. - incremental; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio

On the basis of these double-programming exercises, the ERG is broadly satisfied that the company's base case analyses have been implemented correctly and without significant error. The only potential exception relates to discontinuation following scoliosis surgery; due to the programming approach, the ERG had difficulty in understanding exactly how this is applied. This may explain the discrepancies between the company's model results and those generated from the ERG's double-programming exercise. The ERG notes that the replicated model traces and the replicated cost and QALY calculations implemented by the ERG are very straightforward.

Through a combination of model scrutiny and the ERG's double-programming exercise, the ERG identified the following errors in the company's early onset and later onset models:

(i) *Inconsistent assumptions regarding end-of-life costs between the early and later onset models.* End-of-life costs are included in the early onset model but not in the later onset model. Given that all patients die, the ERG considers the inclusion of these costs to be largely irrelevant, as the only way in which this parameter could impact on the ICER is through discounting these costs at different death times between treatment groups. The company's clarification response[2] (question B31) shows that the inclusion of end-of-life costs has only a negligible impact on the ICER for nusinersen; within this analysis, the ICER for nusinersen is reduced by £236.

(ii) *Discrepancies between the company's model traces and the ERG's double-programmed model traces.* The ERG's double-programmed Markov traces are very similar but not identical to those generated using the company's models. It is unclear whether these discrepancies are the result of rounding errors or minor programming errors in the company's models. The ERG considers that these discrepancies are likely to have a negligible impact on the ICER for nusinersen.

(iii) *Ambiguity regarding intended model time horizon in the early onset model.* The CS[1] states that a 40-year time horizon was used for the early onset model; however, the submitted model and all results presented in the CS correspond to a 60-year time horizon. In response to a request

for clarification (see clarification response,[2] question B32) the company stated that they had intended to use a 60-year time horizon. The ERG notes that the impact of using a 40-year or 60-year time horizon has a minimal impact on the ICER for nusinersen as almost all patients have died within 40 years.

(iv) *Use of different initial distributions between treatment groups in both the early and later onset models.* The initial health state distribution at model entry is based on the treatment-specific distributions in the ENDEAR and CHERISH studies[14, 15] (see Figure 13 and Figure 14). In response to a request for clarification from the ERG (see clarification response,[2] question B22), the company stated that this approach was taken *"to ensure that the model followed the trial data more accurately."* However, the ERG considers this to represent an error that introduces a potential selection bias whereby the patients' initial health state is prognostic of outcomes. The ERG believes that it would have been more appropriate to apply a common initial distribution based on the overall health state distribution within each trial. This issue is tested in the ERG's exploratory analyses and is shown not to significantly impact upon the ICER for nusinersen (see Section 5.6).

Figure 13: Initial HINE-2 health state distribution of patients in the company's early onset model (based on ENDEAR)

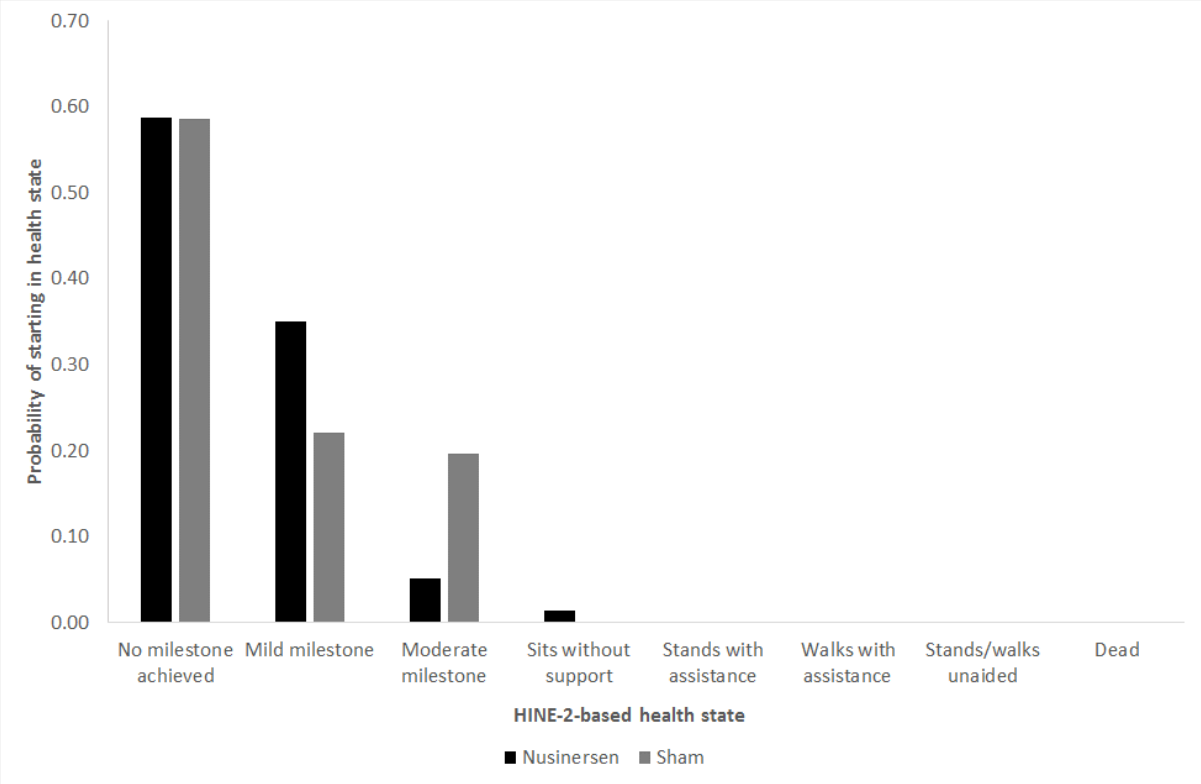
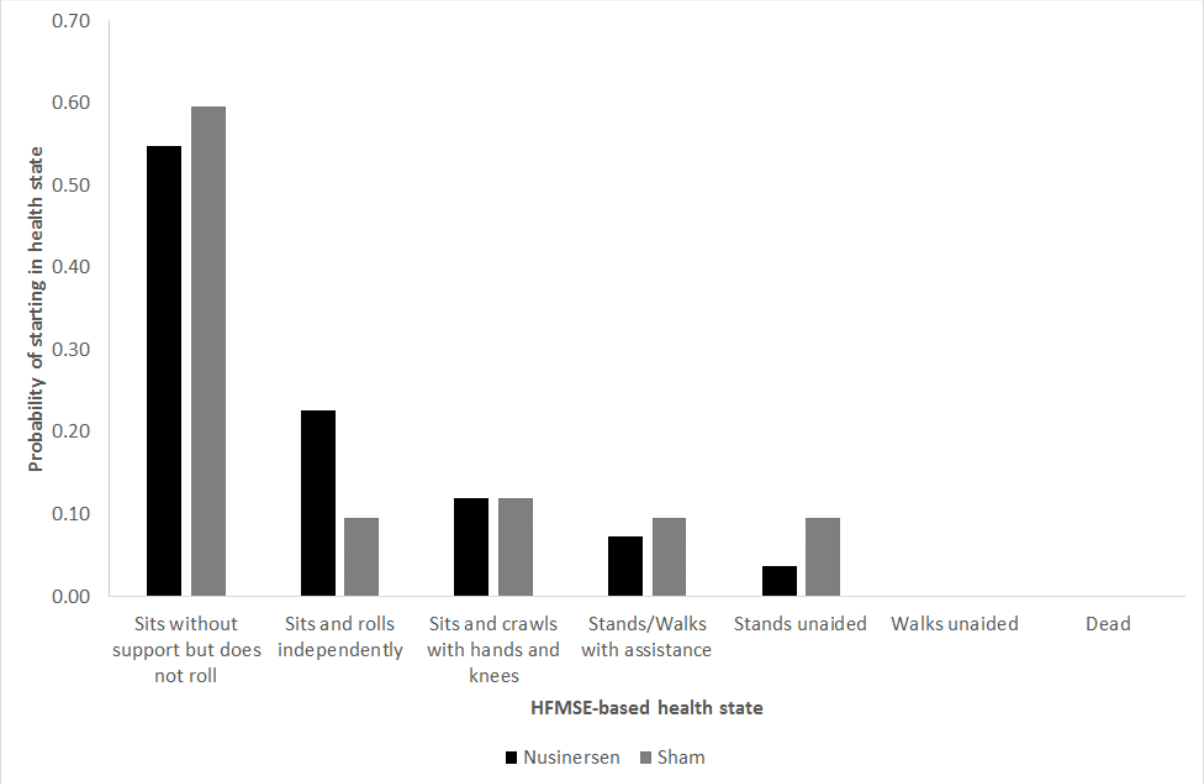


Figure 14: Initial HFMSE health state distribution of patients in the company's later onset model (based on CHERISH)



Correspondence between the written submission and the model

Overall, the implemented model structure and inputs correspond to the description in the CS.[1] However, the ERG notes that the CS is unclear with respect to: (a) how patients' trajectories are modelled after discontinuing due to scoliosis surgery (including the use of tunnel states which are not described in the CS); (b) when scoliosis surgery is applied (once only or during each cycle).

The ERG was able to generate probabilistic ICERs using the company's models which are similar to those reported within the CS. The ERG was also able to replicate the results of the company's deterministic base case analyses, DSAs and scenario analyses. As noted in Section 5.3, the results of the subgroup analyses for the early onset population presented in the CS are incorrect; corrected results were provided by the company following the clarification process (shown in Table 53).[3]

Correspondence of the model inputs and the original sources of parameter values

The ERG attempted to reproduce the transition matrices beyond the end of ENDEAR and CHERISH using the data reported in the CS (see Table 35 and Table 57); the resulting matrices were slightly different to those used in the company's models. It is likely, but not definite, that this is a consequence of rounding errors.

The source of the assumed cost of end-of-life care is not mentioned in the CS, but is cited in the model. The ERG was unable to locate the cost estimate within the NICE Guideline 61 resource use template.[40]

The documentation relating to the UK SMA advisory board meeting[45] (the source of the Type II and Type III mortality adjustment assumptions) does not report the actual adjustment factors applied to the better health states (early onset model, mortality adjustment factor = 0.90; later onset model, mortality adjustment factor = 0.50).

The ERG attempted to replicate IPD from Gregoretti *et al*,[31] and to adjust the data as described by the company. The resulting Kaplan-Meier estimates showed some deviation, as illustrated in Figure 21. This is likely to reflect expected uncertainty in the replication process rather than an error.

All other inputs applied in the base case analysis appear to reflect the original source material.

(3) Concerns regarding model structures which focus only on motor milestones

The ERG has some concerns regarding the structures of the early and later onset models. Both models focus exclusively on the achievement/loss of motor milestones (and death). Clinical advisors to the ERG agreed that the achievement/loss of motor milestones is important in SMA and that the company's model structures are broadly reasonable in terms of functional symptoms of SMA. The clinical advisors also commented that HINE-2 and HFMSE are appropriate instruments through which to classify motor milestones in SMA. They also noted that CHOP INTEND, which is used to inform the long-term extrapolation of motor function in the early onset model, is an appropriate functional scale for infants with Type I SMA, but may be less relevant for older or fragile children or for those with the ability to sit. The clinical advisors further commented that other symptoms and outcomes besides motor function may also be important - in particular, aspects of SMA relating to respiratory function, the explicit use of ventilation and the possibility of infections; these factors are not explicitly captured in either of the company's model structures. The clinical advisors also stated that motor function is not the sole determinant of HRQoL and that the ability to participate in activities and a lack of negative symptoms (e.g. pain and infection) may be more important than motor function. Despite these concerns, the ERG considers that both models are consistent with key outcomes measured in the ENDEAR and CHERISH trials[14, 15] and that alternative characterisations of the disease would likely be hindered by a lack of evidence.

(4) Highly favourable assumptions regarding the expected trajectory of nusinersen-treated patients through modelled motor milestone health states

Within the early onset model, transition probabilities beyond the end of follow-up in ENDEAR[14] are based on the rate of change in CHOP INTEND score over the trial duration, and mean CHOP INTEND

scores conditional on HINE-2 model health state within ENDEAR (supplemented with additional data from Study CS3A[34] for the best two health states). These mean CHOP INTEND scores are treated as thresholds that define whether the patient is in the current state or the next best/worst health state. A similar approach is used within the later onset model, whereby transition probabilities are derived using the rate of change in HFMSE score over the course of the CHERISH trial,[15] together with mean HFMSE scores for each HFMSE model health state within CHERISH (supplemented using data from Study CS2 and CS12[25] for the best health state). In both models, patients receiving nusinersen are assumed either to improve or stay in the same state (deterioration is not permitted), whilst patients in the usual care group are assumed either to worsen or stay in the same state (improvement is not permitted).

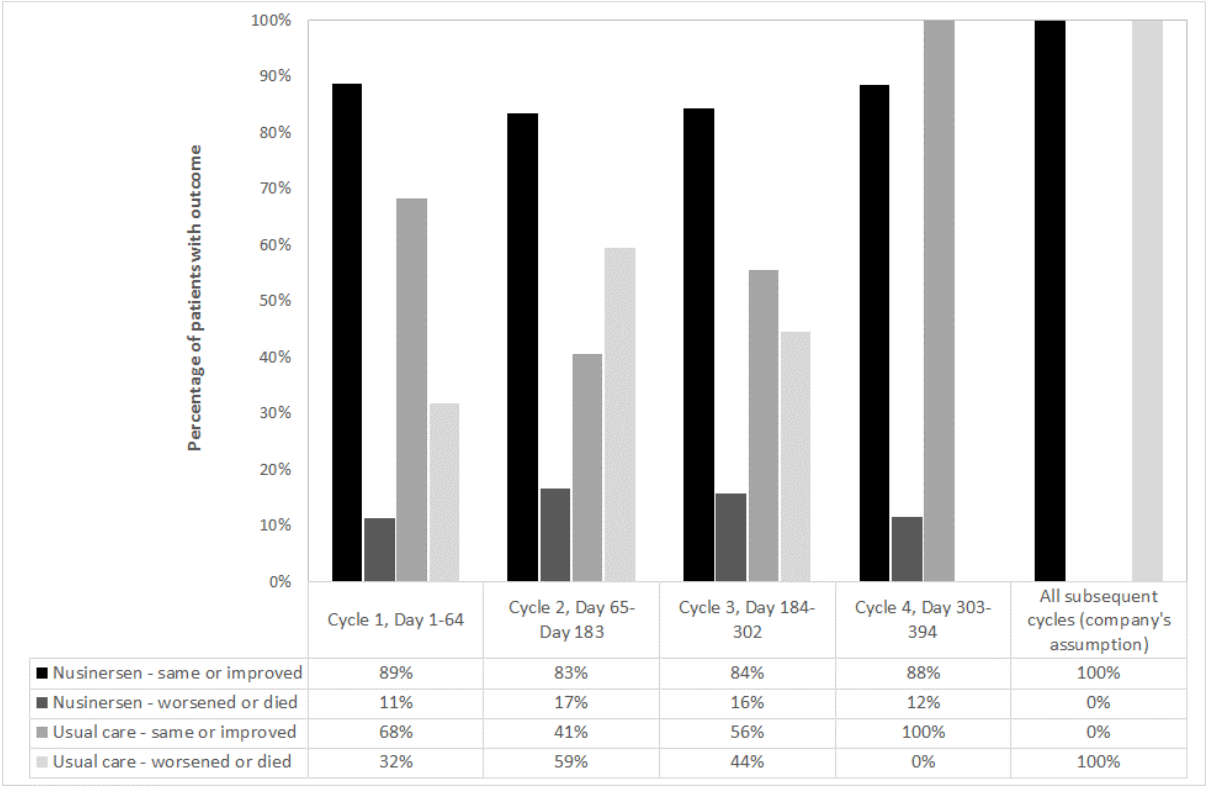
The ERG has several concerns regarding the company's approach for estimating transition probabilities; these concerns are detailed below.

(a) Highly favourable assumptions regarding improvements for nusinersen-treated patients beyond the end of the ENDEAR and CHERISH trials

Clinical advisors to the ERG considered that the company's assumption that patients receiving usual care would not experience improvements in motor milestones beyond the observed follow-up periods of ENDEAR and CHERISH may be broadly reasonable, although they noted that Type III patients in CHERISH may develop some further motor skills. However, the advisors noted that there is considerable uncertainty surrounding the long-term benefits of nusinersen on motor function and that it is possible that patients may lose milestones despite treatment with nusinersen. They considered this to represent a key uncertainty in the clinical evidence base and noted that improving motor milestones increases the burden on the respiratory system.

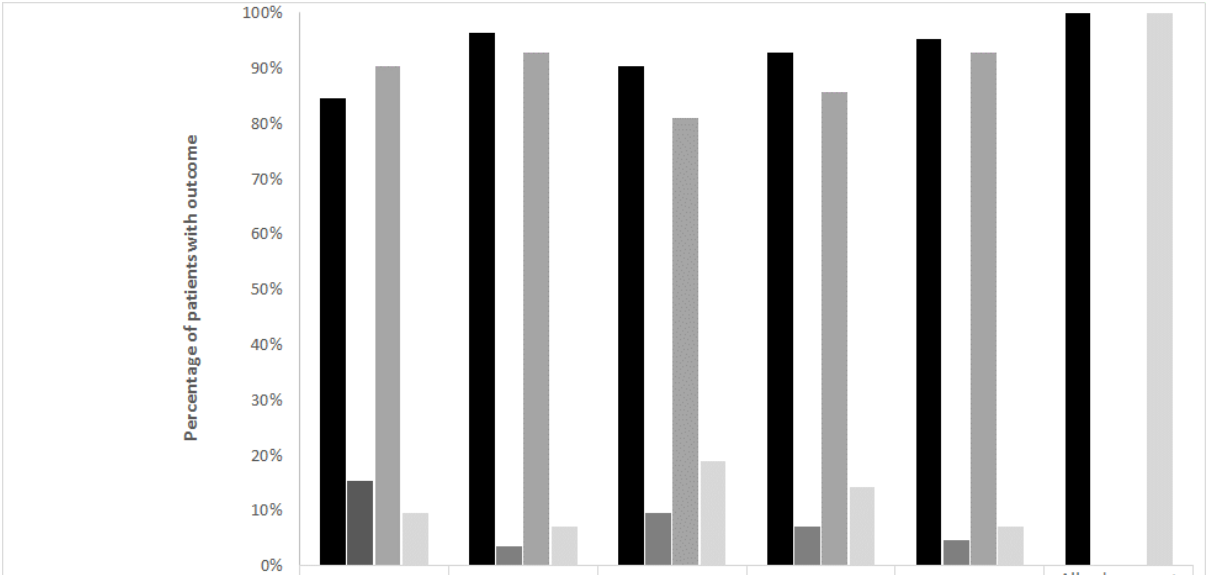
The ERG also notes that the company's assumptions of no deterioration for nusinersen and no improvement for usual care do not reflect the observed clinical trial data. Figure 15 presents observed data from ENDEAR[14] relating to the probability that a patient who is alive and at risk either: (a) stays in the same health state or improves or (b) worsens or dies. Figure 16 presents the equivalent data from CHERISH.[15] As shown in both figures, during every time interval, a proportion of surviving patients receiving nusinersen transited to a worse health state. In addition, during all cycles except for cycle 4 (the interval between days 303 and 394) in ENDEAR,[14] a proportion of surviving patients receiving the sham procedure transited to an improved health state, whilst in CHERISH,[15] a proportion of patients receiving sham transited to an improved state during every cycle. As such, the observed data do not support the assumptions employed in the extrapolated periods of the company's model.

Figure 15: Observed and assumed transitions between HINE-2 health states over time, early onset model



	Cycle 1, Day 1-64	Cycle 2, Day 65-Day 183	Cycle 3, Day 184-302	Cycle 4, Day 303-394	All subsequent cycles (company's assumption)
Nusinersen - same or improved	89%	83%	84%	88%	100%
Nusinersen - worsened or died	11%	17%	16%	12%	0%
Usual care - same or improved	68%	41%	56%	100%	0%
Usual care - worsened or died	32%	59%	44%	0%	100%

Figure 16: Observed and assumed transitions between HFMSE health states over time, later onset model



	Cycle 1, Day 1-92	Cycle 2, Day 93-169	Cycle 3, Day 170-274	Cycle 4, Day 275-365	Cycle 5, Day 366-456	All subsequent cycles (company's assumption)
Nusinersen - same or improved	85%	96%	90%	93%	95%	100%
Nusinersen - worsened or died	15%	4%	10%	7%	5%	0%
Usual care - same or improved	90%	93%	81%	86%	93%	0%
Usual care - worsened or died	10%	7%	19%	14%	7%	100%

In response to a request for clarification from the ERG,[2] the company provided additional information regarding the assumption of continued improvement for patients receiving nusinersen beyond the end of the trials:

[redacted] *the mean rate of improvement in CHOP INTEND was continued beyond the end of ENDEAR trial follow-up although, given the uncertainty around this parameter, a proportion of nusinersen patients can be assumed to reach a plateau or deteriorate. In the RWC arm, the base case assumes that the deterioration in CHOP INTEND observed in the ENDEAR trial continues beyond the end of trial follow-up. A lower rate of deterioration can be applied in the model by selecting the natural history study by Finkel et al. (2014) in infantile onset SMA or Kaufman et al. (2012) in later onset SMA."* (Company's clarification response,[2] question B16a).

However, in their response to clarification question (B16b), the company noted that the clinicians attending the advisory board meeting believed that [redacted] The clinical advisors to the ERG stated that it is more likely that outcomes for patients receiving nusinersen would follow a distribution whereby some would improve, whilst others would worsen. Therefore, the assumptions employed within the company's models regarding long-term improvements in motor function for patients receiving nusinersen do not fully reflect clinical advice received by the company or the ERG. Rather, the company's approach to extrapolating transition probabilities for the nusinersen group within both models appears to be unrealistically optimistic. Within the later onset model, the company's approach to extrapolating transition probabilities for the usual care group may be unduly pessimistic, at least for some Type III SMA patients.

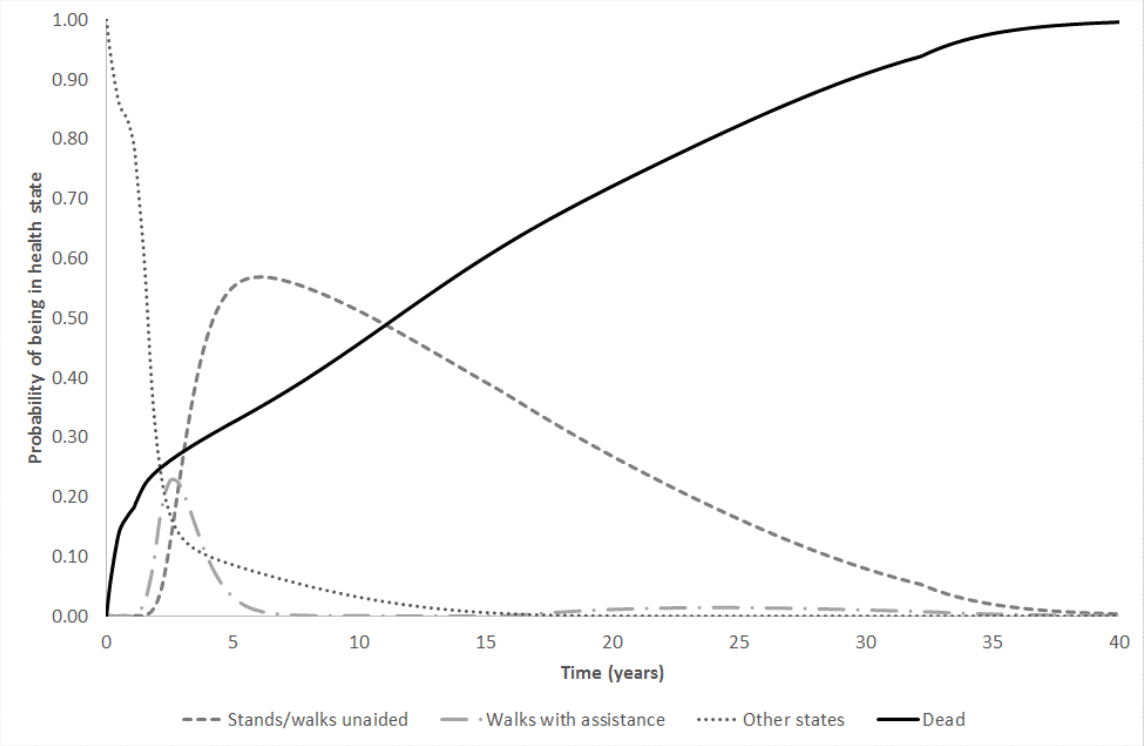
(b) Model predictions that all surviving patients will reach the best health states were not observed within the ENDEAR and CHERISH trials

Figure 17 shows the model-predicted health state occupancy within the nusinersen group over the time horizon within the early onset model; Figure 18 presents the equivalent estimates for the later onset model. As shown in Figure 17, within the nusinersen group, the vast majority of surviving patients reach the best health state (HINE-2 State [vii] Stands/walks unaided) within the first five years of the model time horizon. As a consequence of the assumption regarding no deterioration within the nusinersen group and the very low probability of undergoing scoliosis surgery and discontinuing treatment, almost all patients remain in this state until death. However, within ENDEAR, no patients achieved milestones which would locate them in the best two health states, and by the end of trial follow-up, only one patient

had reached State (v) Stands with assistance at any timepoint. The health state projections predicted by the early onset model therefore appear highly favourable given the observed data. The ERG notes that these favourable projections are driven by the company's combined use of CHOP INTEND data in the unobserved period and the assumption that motor function cannot deteriorate for patients receiving nusinersen.

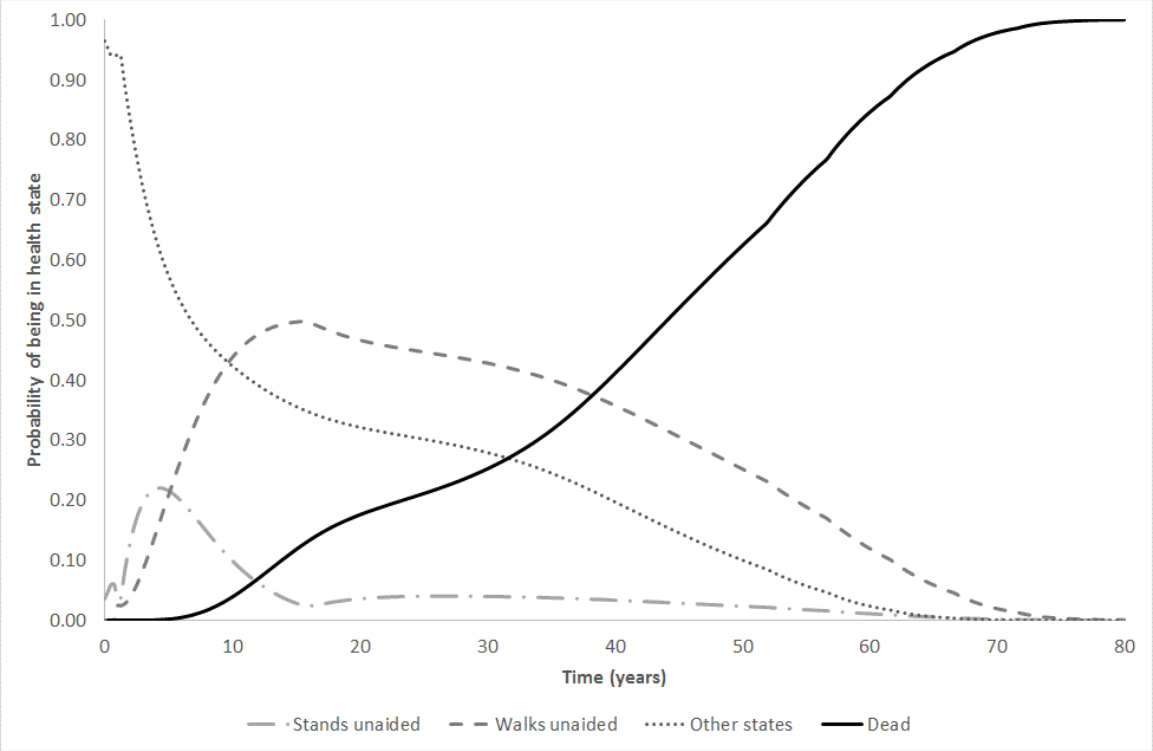
Similarly, within the later onset model, approximately 49.8% of patients reach the best health state (State [vi] Walks unaided) by around 15 years (see Figure 18). However, only two patients reached this milestone within the nusinersen group of CHERISH. The ERG notes that the company's later onset model predictions are driven by the assumption that the motor function for nusinersen-treated patients cannot deteriorate.

Figure 17: Health state occupancy over time, early onset model, nusinersen group



Note: Stands/walks unaided is the best state; Walks with assistance is the second best state

Figure 18: Health state occupancy over time, later onset model, nusinersen group



Note: Walks unaided is the best state; Stands unaided is the second best state

(c) Concerns regarding the company's approach to calculating transition probabilities

Within the early onset model, the company's approach for deriving transition probabilities for the unobserved period relies on an assumption of perfect correlation between CHOP INTEND and HINE-2 health state. The CHOP INTEND scores represent a threshold for being in the current state or the next best/worst state. However, the company's approach applies a different set of thresholds depending on treatment group (the mean CHOP INTEND scores for each health state are different for the nusinersen and usual care groups, see Table 35). The ERG considers the joint interpretation of these two assumptions to be unclear.

In response to a request for clarification from the ERG,[2] the company stated that: *"There won't necessarily be a direct relationship between the changes on one measure and the changes in the other both because they are measuring different aspects of motor ability and because of the different properties of the two measurement scales. For example, considering patients' absolute scores, patients are closer to zero on the HINE-2 scale than on the CHOP INTEND scale at baseline, thus limiting the scope for further reductions in score over time with HINE-2."* (Company's clarification response,[2] question B17b). The company's response calls to question the appropriateness of assuming a perfect correlation between the HINE-2 and CHOP INTEND instruments. A similar issue regarding the

definition of health states with treatment-specific HFMSE thresholds also applies to the company's later onset model (see Table 57).

The rate of improvement/worsening in CHOP INTEND and HFMSE are assumed to be constant with respect to time and are applied monotonically to each permitted transition. Figure 14 of the company's clarification response[2] and Figure 21 of the CS[1] suggest that the mean change in CHOP INTEND score in each group in ENDEAR and the mean change in HFMSE in CHERISH are not constant.

In addition, as shown in Equations [iv] to [vii], the company's calculation approach involves applying a constraint which prevents the estimated transition probabilities from exceeding 1.0. This constraint is necessary in the usual care group of the early onset model for the transition from State (v) Stands with assistance to State (iv) Sits without support. Based on the company's calculation, the unconstrained transition probability is [redacted] ([redacted]*4]/[52.7-48.0]); the ERG has concerns with the appropriateness of the calculation, given that this value exceeds 1.0. A further issue applies to the transition between State (iv) Stands/walks with assistance and State (iii) Sits and crawls with hands and knees within the company's later onset model, whereby the threshold between states is the same hence the denominator is zero; this calculation returns a #DIV/0! error unless a constraint is applied (see Table 57). These issues raise further questions regarding the appropriateness of the approach used to calculate transition probabilities within both models.

On the basis of the above issues, the ERG considers the company's extrapolation to be highly optimistic, mathematically unsound and inconsistent with the available evidence from ENDEAR and CHERISH.

(5) Highly favourable assumptions regarding the expected survival of nusinersen-treated patients

The company use a complex multi-stage approach for extrapolation using external data. As described by the company, it is widely recommended that longer-term data should be used to inform the extrapolation of clinical trial data with limited follow up.[55, 56] However, the ERG has concerns regarding how this has been implemented by the company and considers that a simpler approach would have greater plausibility and would provide more transparent survival predictions. The main points are summarised below; these are discussed in further detail in the following sections.

(i) Complexity of modelling approach
- Not clearly described. Assumptions not clearly stated or justified
- Some standard parametric models fitted to the observed data provided plausible predictions

(ii) Use of external data from Gregoretti *et al*[31] to inform early onset model

- Assumption that after adjustment for age, mortality is the same in Gregoretti *et al*[31] and ENDEAR[14] is not plausible
- Uncertainty due to reconstruction of IPD from published Kaplan-Meier curve

(iii) Use of external data from Zerres *et al*[33] to inform later onset model

- Assumption that mortality is the same as in CHERISH is not justified

(iv) Use of general population mortality

- Assumption that long-term mortality is systematically different between the studies and the general population (by assuming a constant HR) is not plausible

(v) Assumptions regarding treatment effect

- Description that a conservative HR of 1.0 is applied is misleading due to the implementation of the Type II adjustment

(vi) Concerns regarding SMA Type II adjustment

- No observed data to justify the use of Zerres *et al*[33] data or the adjustment factors used.

(i) Complexity of modelling approach

Jackson *et al*[49] present a framework for survival extrapolation using external data which is referenced by the company in justifying their approach (see clarification response,[2] question B9). If the external population has the same mortality at all times (or in the long-term) as that of the external population, then survival estimates from the external population can be used directly without adjustment. This assumption permits the direct use of data from Gregoretti *et al*[31] and Zerres *et al*[33] in the early onset and late onset models, respectively. Alternatively, OS may be assumed to be different, but systematically similar in such a way that the external data can be adjusted to estimate OS in the target population. This assumption permits the application of the adjusted general population mortality data. The validity of these assumptions is paramount to the reliability of the survival predictions; however, no clear justification for either assumption was presented by the company. The ERG considers that the plausibility of these assumptions is questionable and considers each case in further detail below.

Given the concerns regarding the use of external data, the ERG considers that a simpler approach based on extrapolating parametric models fitted to observed trial data may have been both more informative and more transparent than the approach adopted by the company. Consideration of appropriate external data is important; however, it could be used more simply to judge the plausibility of models fitted to observed data, or to inform certain parameters.[56] In their response to clarification questions from the ERG[2] (question B9), the company states that some parametric models provided plausible extrapolations and so the ERG considers that using these would be a reasonable approach. Details of which models provided plausible predictions were not provided by the company.

As summarised in Section 5.3.3, the company provides a detailed account of model fitting to each observed data source; however, the long-term fitted survival probabilities are of limited relevance given that composite functions are applied in the model. The survival functions as applied in the model are shown in Figure 19 and Figure 20 for the early onset and late onset models, respectively.

With respect to the early onset usual care group, one clinician believed that the survival curve was reasonable. The second advisor believed that the curve was optimistic compared with the patients seen in her clinical practice and commented that in routine care, many families do not have the resources to manage NIV or aggressive management and instead 'opt' for a palliative approach. The advisor also noted that in some areas, resources and experience in supporting small infants with SMA are limited. With respect to the early onset nusinersen group, one clinician stated that the survival curve reflected a "big assumption" whilst the other believed it was optimistic as she would not expect any patients to survive to 35 years. One of the advisors had particular concerns regarding the plausibility of the company's mortality adjustment in the better states of the early onset model, and noted that longer-term evidence from the SHINE and NURTURE studies may provide useful information.

Figure 19: Fitted survival curves, early onset model

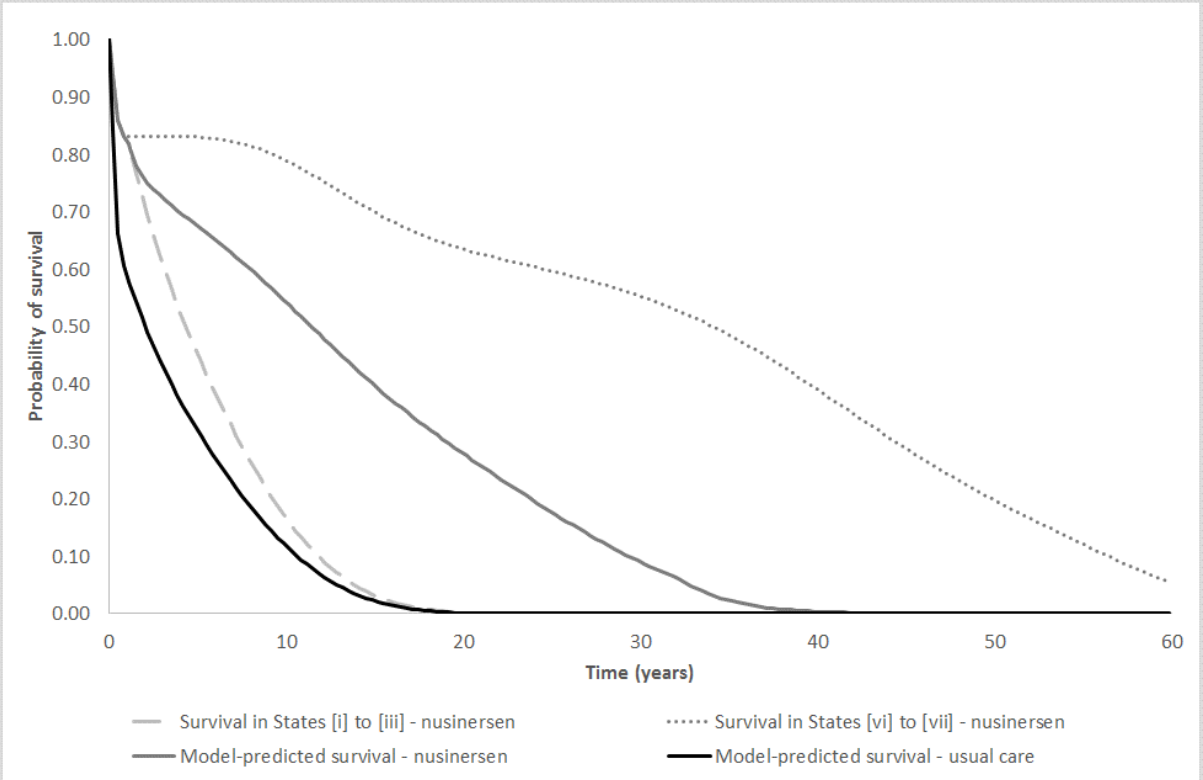
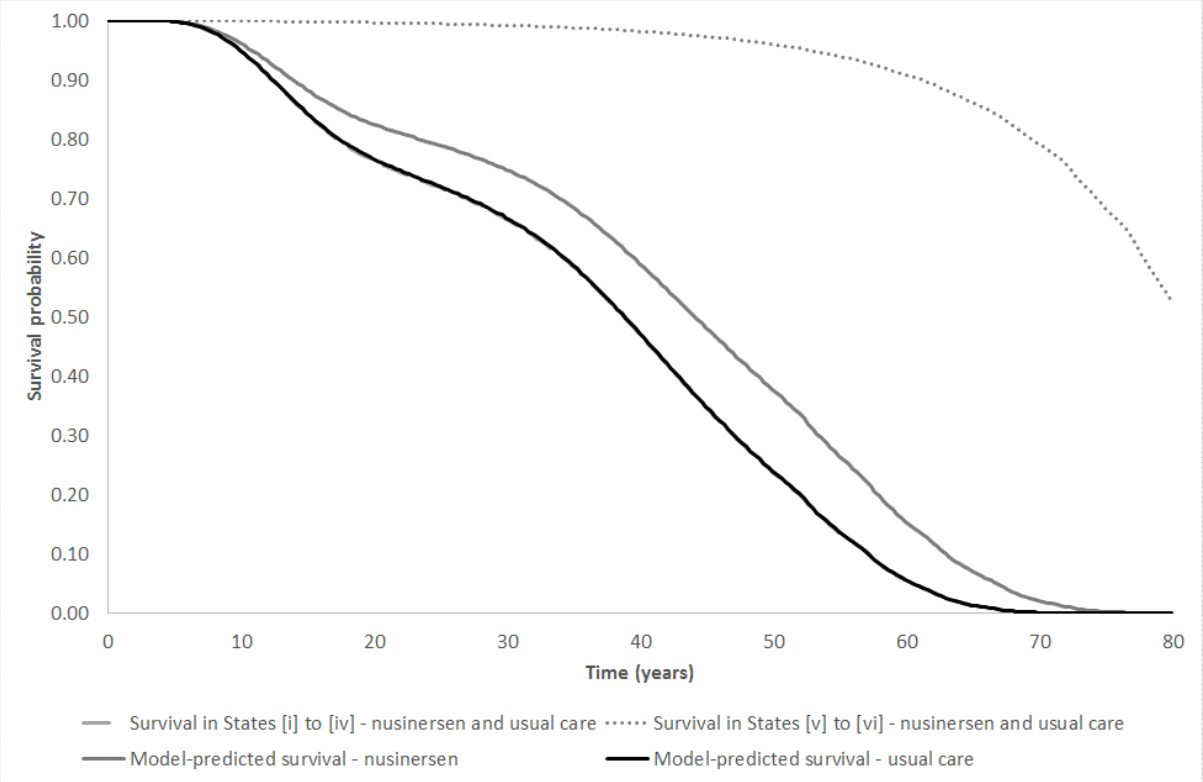


Figure 20: Fitted survival curves, later onset model



Note: the mortality adjustment has almost no effect in the usual care group due to the small proportion of patients in the best two states by the end of month 15

(ii) Use of external data from Gregoretti et al[31]

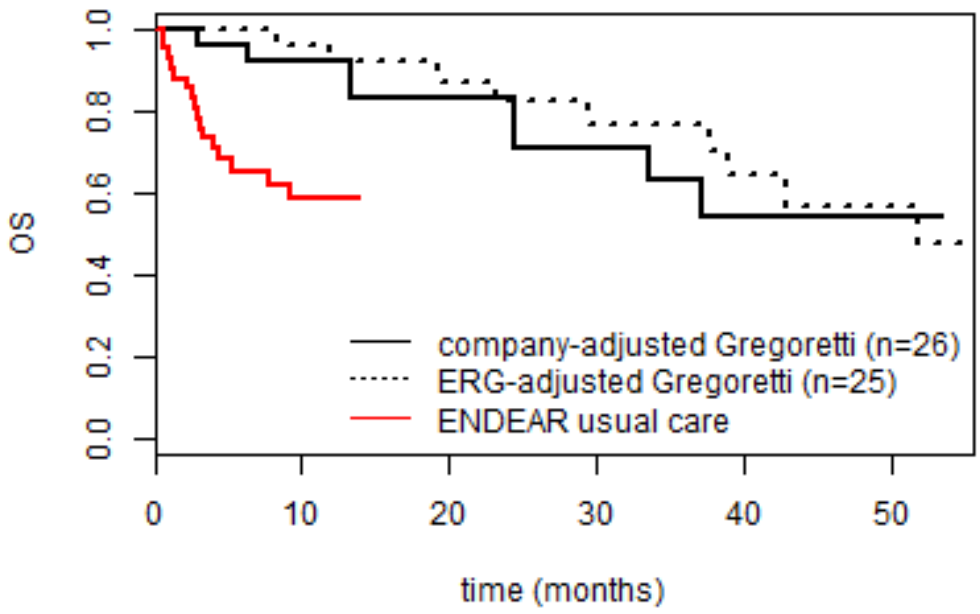
The ERG has concerns regarding the use of data from Gregoretti *et al*[31] to represent the usual care arm in ENDEAR. Of the 31 patients receiving NRA, 21 patients (67.8%) were over 3 months (~13 weeks) old at onset of symptoms, whereas in ENDEAR, patients tended to be younger at symptom onset (mean age of symptom onset of 9.6 weeks, range 1-20 weeks). As discussed in the original publication, mortality in the NRA cohort was higher (45.2%) than reported elsewhere.[57-59] The study authors comment that NIV and mechanically assisted coughing were used differently over the years of the study; the clinical advisors to the ERG noted that the reported outcomes from the study are poorer than would be expected in current clinical practice.

In order to fit parametric survival models, IPD were reconstructed by the company using the algorithm reported by Guyot *et al*.[44] The accuracy of the reconstruction depends on the amount of information provided in the original publication. In the case of Gregoretti *et al*,[31] the authors provide the total number of events (14 out of 31 patients died) but a number at risk table was not provided which results in a reconstruction with a higher degree of uncertainty. This is highlighted in Figure 21 by the difference between the ERG's reconstruction and that reported in Figure 34 of the CS.[1] A further limitation of the reconstructed IPD is the lack of information about important individual-level covariates. The company

adjusted the data to account for differences in the mean age of the populations, resulting in a reduction in the sample size from 31 to 26. However, there is potential for other confounding factors to remain.

The observed OS from ENDEAR and the adjusted Gregoretti *et al* NRA data are shown in Figure 21. There is a marked difference in OS between the two populations which indicates that the age-correction performed by the company was not sufficient to account for differences in baseline characteristics between the two groups. The company's clarification response[2] states that survival was "*greater than that expected from the clinical advice we received for UK patients and the sample size was small. From this paper there is insufficient information to draw conclusions on why survival was higher in the Italian patient population compared to the UK patient population.*" (Company's clarification response,[2] question B12).

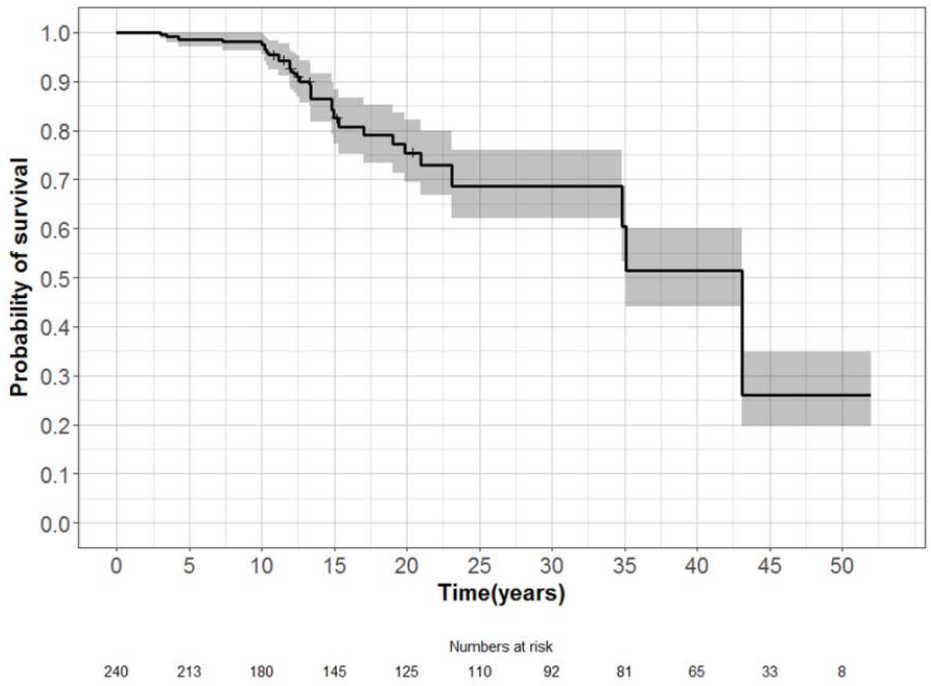
Figure 21: Kaplan-Meier OS estimates from adjusted Gregoretti *et al* and ENDEAR



(iii) Use of external data from Zerres et al[33]

The company's Kaplan-Meir curve of reconstructed IPD from Zerres *et al*[33] is shown in Figure 22. At 15 months, OS is 100%, as was observed in CHERISH. However, insufficient information was presented in Zerres *et al*[33] and the CS to allow the ERG to determine whether key characteristics of the two populations were similar.

Figure 22: Kaplan-Meier curve based on reconstructed IPD from Zerres *et al* (reproduced from CS, Figure 44)



The clinical advisors to the ERG noted that it is unclear whether any respiratory support was provided in the Zerres *et al* cohort. As this study predates publications on effective NIV use in paediatric cohorts, this is unlikely to reflect patients treated in current clinical practice.

(iv) Use of general population mortality

Beyond the trial data, OS is informed by general population mortality life tables.[32] OS is assumed to be systematically different between Gregoretti *et al* and the general population, or between Zerres *et al* and the general population, as characterised by a constant HR. The company acknowledge that the assumption of proportional hazards is not expected to hold; however, they state that "*this is a conservative approach since we would expect hazard rates to get closer to those from the general population with time. However, the available data did not provide information on how the hazard ratio may change over time*" (Company's clarification response,[2] question B13).

(v) Assumptions regarding treatment effect

The treatment effect in the first 13 months of the early onset model is derived from observed data in ENDEAR.[14] The company's preferred model (1-knot spline) provides a constant HR. Beyond the observed trial data, the company state that a conservative HR of 1.0 is applied in the base case; however, this is misleading as survival in the nusinersen treatment group is largely driven by an assumed switch to the Type II SMA mortality curve (proportion = 0.90).

(vi) Concerns regarding SMA type assumption

Survival for nusinersen-treated patients reaching model health states (iv) to (vii) was assumed to lie close to that observed in a Type II SMA population, with a weight of 0.90 applied to the survival prediction from Zerres *et al*[33] and a weight of 0.10 applied to the prediction from Gregoretti *et al.*[31] Justification of these weightings was provided by referencing an advisory board meeting on SMA.[45] However, no further details on how this figure was agreed or elicited were provided, despite a request from the ERG in clarification question B8.[2] Clinical advisors to the ERG considered that this was a large and optimistic assumption. The clinical advisors noted that there is a trade-off between gaining motor ability and placing a greater burden on the respiratory system, the impact of which is not clear.

(6) Issues relating to estimated patient utilities

(a) Poor face validity of patient utilities

The ERG considers that the mapped utility values used in the company's early and later onset models (see Table 43) have poor face validity. The worst states (early onset model - State [i] No milestones; later onset model - State [i] Sits without support but does not roll) are associated with a utility of [redacted], whilst the best states (early onset model - State [vii] Stand or walks without assistance; later onset model - State [vi] Walks without assistance) are associated with a utility of [redacted] The ERG considers that it is implausible that over the course of 10 years, a notional patient with SMA who never develops any motor milestones would accrue [redacted] undiscounted QALYs.

The ERG's clinical advisors also did not consider the company's patient utility values to be plausible, and noted in particular the high valuations for the worse states and the limited range of utility gain between the valuations for the best and worst states. They stated that although the utility of [redacted] for an infantile onset type I SMA patient who has achieved no milestones may be reasonable during the first few months of life (before motor function develops in healthy children), this would not be valid as the child gets older. One clinical advisor also commented that whilst mobility may have some influence on HRQoL, the ability of patients to participate in usual activities and a lack of negative symptoms (such as pain and infections) are likely to be key determinants of HRQoL.

(b) Issues relating to using mapped PedsQL data to represent utilities for patients with SMA

The algorithm used by the company (Khan *et al*[36]) mapped the PedsQL to the EQ-5D-Y (valued using the adult EQ-5D tariff). There are two main limitations associated with using the mapped values to generate utility values for patients with SMA.

Firstly, the study in which the mapping algorithm was developed was based on healthy schoolchildren aged 11-15 years. This population is very different to the populations represented within the company's models. The ERG believes that a healthy population completing both the PedsQL and EQ-5D-Y would

likely have very different responses to patients with early onset SMA or later onset SMA. Most of the children recruited into the mapping study had no problems in any dimensions of the EQ-5D-Y (percentage in Khan *et al*[36] with no problems in each domain - 95% mobility, 98% self-care, 95% usual activities, 76% pain/discomfort and 83% anxiety depression) and high PedsQL scores (scores of ≥80 in physical, emotional, social and school functioning, where the maximum score is 100). CHERISH PedsQL scores are not reported in the CS or the appendices, but they are unlikely to be as high as this in a population where severe motor function problems are characteristic of the disease. It is therefore unlikely that a mapping function developed in such a different population would be appropriate for the patient population under consideration. Khan *et al*[36] note that *"the performance of these algorithms in childhood populations, which differ according to age or clinical characteristics to our own, remains to be evaluated."*

Secondly, Khan *et al*[36] comment that they had few responses at the more severe end of the EQ-5D; this will have impacted upon the accuracy of the derived mapping functions. Mapping may overestimate the utility values for those at the severe end, primarily due to lack of data to accurately fit a regression model. The high utility values reported in the CS may well be a reflection of this problem.

In response to a request for clarification regarding to appropriateness of the mapping algorithm (see clarification response,[2] question B25), the company acknowledged that the Khan *et al*[36] mapping algorithm is *"not ideal"*, but noted that as the PedsQL was the only HRQoL questionnaire administered in either clinical trial (ENDEAR or CHERISH), mapping should be undertaken. The ERG disagrees and notes that two alternative sources could have been used: Bastida *et al*[37] and Lloyd *et al*[46] (previously described in Section 5.3.3). Whilst these studies used parents/clinicians as a proxy for SMA patients, both studies include valuations for health states associated with SMA.

Within Bastida *et al*,[37] mean values from UK respondents were reported to be [redacted] for Type I, [redacted] for Type II and [redacted] for SMA Type III (see Table 74). Health state valuations were highly variable between respondents from each country. Within Lloyd *et al*,[46] clinicians' valuations of health states for Type I SMA health states ranged from -0.33 to 0.71, whilst valuations for Type II SMA health states ranged from -0.13 to 0.72. The clinical advisors to the ERG commented that whilst Bastida *et al*[37] and Lloyd *et al*[46] are not subject to the same methodological problems as the mapping analysis, they also appear to have limited face validity, in particular, due to the very low (negative) valuations for patients in the worst health states which undermines the HRQoL of non-ambulant patients. The clinical advisors further commented that the valuations from these studies may not reflect those of other clinicians and families of SMA patients.

Table 74: EQ-5D utilities (parent proxy) reported by Bastida *et al*

SMA type	UK	Spain	France	Germany
All SMA types	[redacted]	[redacted]	[redacted]	[redacted]
Type I SMA	[redacted]	[redacted]	[redacted]	[redacted]
Type II SMA	[redacted]	[redacted]	[redacted]	[redacted]
Type III SMA	[redacted]	[redacted]	[redacted]	[redacted]

SMA - spinal muscular atrophy

Table 75: Elicited utilities from Lloyd *et al* vignette study

SMA Type I health states and associated HRQoL scores	
Health state	**Utility value**
Baseline	-0.12
Worsened	-0.24
Improvement	-0.17
Reclassified as SMA Type II	-0.04
Stands with assistance	0.04
Walks with assistance*	0.52
Reclassified as SMA Type III*	0.71
SMA after scoliosis surgery	-0.22
Gastric/nasogastric tube	-0.17
Requires ventilation	-0.33
SMA Type II health states and associated HRQoL scores	
HFMSE health state	**Utility value**
Baseline	0.04
Worsened	-0.13
Mild improvement	0.04
Moderate improvement	0.10
Stands/walks with assistance*	0.39
Stands/walks unaided	0.72
Loss of ambulation with/without assistance*	-0.12

SMA - spinal muscular atrophy; HRQoL – health-related quality of life
** Denotes health states where 2 index scores were calculated for one of the participants*

Overall, the ERG considers that none of the sources are ideal, but prefers the vignette study[46] as this broadly aligns with the final models' health states and is based on EQ-5D assessments of clinical experts in SMA. The ERG also notes that owing to the company's extrapolation assumptions regarding no deterioration in motor function for nusinersen-treated patients and no motor function improvement for patients receiving usual care, the utility values for the best and worst states have the greatest influence on the ICER in both the early and later onset models.

(7) Arbitrary calculations underpinning caregiver disutilities

Carer health utility values are based on self-reported EQ-5D-5L values of carers of patients with SMA (Bastida *et al*[37]). No detail is provided on the scoring of the EQ-5D-5L. Caregiver health utility values are adjusted by patient disutility between different states; the difference between this adjusted utility value and general population utility is used to calculate the caregiver disutility. The reasons for adjusting

are not clear from the CS.[1] The ERG has three main issues relating to the company's approach to estimating caregiver disutilities:

(i) Caregiver utilities are estimated based on differences in patient utility between HINE-2/HFMSE health states. However, it is unclear whether the impact of achieving a particular milestone for a patient would be equal to that for a carer, and the assumption that the ordering of health states for patients is the same as that for impacts on caregivers health is not adequately justified in the CS. One clinical advisor to the ERG considered that some degree of correlation might be expected, but noted that caregiver burden would be driven by other factors besides restricted motor function e.g. the incidence of recurrent infections and pain, educational development, availability of support and emotional burden. The other clinical advisor stated that impacts on carers are *"very individual and impossible to tease out."*

(ii) The calculations used in the company's model are arbitrary and most are informed by utilities for other states than the one being valued.

(iii) The ERG and its clinical advisors do not consider the patient utilities obtained from the mapping study to have face validity. This has a direct impact on the face validity of the company's estimated caregiver disutilities.

Given that Bastida *et al*[37] reports EQ-5D utilities from caregivers according to SMA type (see Table 75), it is unclear why these estimates were not used directly for health states defined by milestones associated with SMA type (as is assumed for the health state costs).

Table 76: Caregiver utilities reported by Bastida *et al*[37]

SMA type	Caregiver utility value			
	UK	Spain	France	Germany
All SMA types	[redacted]	[redacted]	[redacted]	[redacted]
Type I SMA	[redacted]	[redacted]	[redacted]	[redacted]
Type II SMA	[redacted]	[redacted]	[redacted]	[redacted]
Type III SMA	[redacted]	[redacted]	[redacted]	[redacted]

SMA - spinal muscular atrophy

(8) Issues relating to health state costs

The ERG notes the following issues relating to the costs included in the company's models:

(i) End-of-life costs are included in the early onset model, but not the later onset model. This is inconsistent.

(ii) The model does not include a cost associated with scoliosis surgery. The inclusion of scoliosis surgery costs is, however, unlikely to have a significant impact on the ICER for nusinersen.

(iii) Health state costs are taken from the cross-sectional study reported by Bastida *et al.*[37] Clinical advisors to the ERG noted that the estimated costs for Type I SMA and Type II SMA milestones appeared to be low, given the high degree of dependency associated with these

patients and the resources required to manage their condition. Both clinical advisors noted that the costs of managing SMA are likely to be dependent on age. This is not captured in the company's models.

(9) Representation of uncertainty

As highlighted in Table 49 and Table 66, the company's PSA in both the early and later onset models is subject to limitations, specifically:

(i) Several uncertain model parameters (for example, the initial distributions and the mortality adjustment factors) are held fixed at their mean values. These values are uncertain and should be characterised using probability distributions.

(ii) Health utilities are sampled using a single random number. This leads to over-correlation between each individual health state utility value;[47] as such, the uncertainty surrounding these parameters will be underestimated.

(iii) Priors are included for some but not all unobserved transitions (the ERG presumes that this is to ensure that the assumptions concerning improvement/deterioration of motor function are maintained in the PSA).

However, the ERG notes that correcting these issues is likely to have a negligible impact on the probabilistic ICER for nusinersen.

More generally, the post-trial transition probabilities, the patient health utilities and the mortality risks applied in both the early and later onset models are all highly uncertain. The ERG does not consider the company's exploration of the impact of this uncertainty to be sufficient.

5.6 Exploratory analyses undertaken by the ERG

5.6.1 ERG's exploratory analyses - methods

The ERG undertook eight sets of exploratory analyses; the same analyses were applied to both the early and later onset models. The ERG's preferred analysis includes: (i) the use of a common initial distribution across health states for both treatment groups; (ii) the inclusion of end-of-life costs for the later onset population; (iii) the use of patient utilities from Lloyd *et al*[46] and (iv) the application of caregiver utilities by SMA type (from Bastida *et al*[37]) to states relating to SMA milestones. Additional sensitivity analyses were undertaken using the ERG's analysis to explore: (i) the use of alternative HRQoL estimates for patients; (ii) the exclusion of the mortality adjustment factor applied to the better health states and (iii) alternative assumptions regarding long-term transition probabilities. The methods used to implement these analyses are described below; technical details for implementing the analyses in the company's models are presented in Appendix 4.

Exploratory analysis 1: Use of the average initial distribution for both treatment groups

Within this analysis, the initial distributions were set equal to the weighted average probability of being in each state in both groups at baseline in ENDEAR[14] and CHERISH.[15] The correction of this error is applied to all subsequent exploratory analyses.

Exploratory analysis 2: Inclusion of end-of-life costs for the later onset model

In order to maintain consistency between the early and later onset models, end-of-life costs were included in the later onset model. No amendment was made to the early onset model as these costs were already included.

Exploratory analysis 3: Use of patient utilities from the vignette study

As discussed in Section 5.5, the ERG has concerns regarding the validity of the utilities based on mapping the PedsQL data to the EQ-5D. Within this analysis, the data reported in the abstract by Lloyd *et al*[46] for Type I SMA are applied to the early onset model and the values for Type II SMA are applied to the later onset model (see Table 77). The ERG recognises that, based on clinical advice, the values for the worse health states also appear to be subject to face validity issues.

Table 77: Health utilities from vignette study applied in ERG's exploratory analyses

Early onset model		
HINE-2 health state	**Mapped PedsQL utility (company's base case[1])**	**Utilities elicited within vignette study[46]**
No milestones achieved	[redacted]	-0.24
Mild milestones	[redacted]	-0.12
Moderate milestones	[redacted]	-0.17
Sits without support	[redacted]	-0.04
Stands with assistance	[redacted]	0.04
Walks with assistance	[redacted]	0.52
Stands/Walks unaided	[redacted]	0.71
Later onset model		
HFMSE health state	**Mapped PedsQL utility (company's base case[1])**	**Utilities elicited within vignette study[46]**
Sits without support but does not roll	[redacted]	0.04*
Sits and rolls independently	[redacted]	0.04†
Sits and crawls with hands and knees	[redacted]	0.10‡
Stands/Walks with assistance	[redacted]	0.39
Stands unaided	[redacted]	0.72
Walks unaided	[redacted]	0.72

HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; HFMSE - Hammersmith Functional Motor Scale-Expanded; PedsQL - Paediatric Quality of Life Inventory

** Assumed to reflect "Baseline" state in vignette study; † Assumed to reflect "Mild improvement" state in vignette study; ‡ Assumed to reflect "Moderate improvement" state in vignette study*

Exploratory analysis 4: Application of caregiver utilities by SMA type from Bastida et al[37] to states relating to SMA milestones in both the early and later onset models

As discussed in Section 5.5 (critical appraisal point 7), the ERG considers that the company's approach to incorporating health state-dependent caregiver disutilities is arbitrary and lacks adequate justification. Given that caregiver EQ-5D values are reported by SMA type within Bastida *et al*,[37] the ERG considers the use of these data directly to be more appropriate than the values applied within the company's models.

Table 78: Caregiver utilities applied in the ERG's exploratory analyses

Early onset model		
HINE-2 health state	**Caregiver utility applied in company's base case[1])**	**Caregiver utilities from Bastida *et al*[37] applied in ERG exploratory analysis**
No milestones achieved	[redacted]	[redacted]
Mild milestones		
Moderate milestones		
Sits without support		
Stands with assistance		
Walks with assistance		
Stands/Walks unaided		
Later onset model		
HFMSE health state	**Caregiver utility applied in company's base case[1])**	**Caregiver utilities from Bastida *et al*[37] applied in ERG exploratory analysis**
Sits without support but does not roll	[redacted]	[redacted]
Sits and rolls independently		
Sits and crawls with hands and knees		
Stands/Walks with assistance		
Stands unaided		
Walks unaided		

HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; HFMSE - Hammersmith Functional Motor Scale-Expanded; ERG - Evidence Review Group
** Based on SMA Type I ; † Based on SMA Type II; ‡ Based on SMA Type III*

Exploratory analysis 5: ERG-preferred analysis

The ERG's preferred analysis combines exploratory analyses (i) to (iv). It should be noted that this analysis does not address the ERG's concerns regarding the company's modelled survival and motor function trajectories. As such, the ERG's "preferred" ICERs are very likely to be underestimated in both SMA populations.

Exploratory analysis 6: Use of alternative patient HRQoL estimates (Bastida et al[37] and expert clinical judgement)

Two alternative analyses were undertaken to explore the impact of using different HRQoL estimates for patients with SMA:

(6a) Analysis using utilities reported by Bastida *et al.*[37] Within this analysis, the UK patient utilities by SMA type reported by Bastida *et al*[37] (Type I utility=[redacted]; Type II utility=[redacted]; Type III utility=[redacted]) are applied to the model health states defined by milestones consistent with these SMA types.

(6b) Analysis using HRQoL estimates obtained from ERG clinical advisors. Within this analysis, the clinical advisors to the ERG were asked to provide plausible estimates of HRQoL for the health states included in the company's early and later onset models (see Table 79). It should be noted that these HRQoL estimates should be interpreted with caution as they are not preference-based.

Table 79: Clinical advisors' estimates of HRQoL associated with model health states

Early onset model	
HINE-2 health state	**HRQoL estimate**
No milestones achieved	0.20
Mild milestones	0.25
Moderate milestones	0.35
Sits without support	0.60
Stands with assistance	0.65
Walks with assistance	0.75
Stands/Walks unaided	0.85
Later onset model	
HFMSE health state	**HRQoL estimate**
Sits without support but does not roll	0.60
Sits and rolls independently	0.60
Sits and crawls with hands and knees	0.60
Stands/Walks with assistance	0.75
Stands unaided	0.85
Walks unaided	0.85

HRQoL - health-related quality of life; HINE-2 - Module 2 of the Hammersmith Infant Neurological Examination; HFMSE - Hammersmith Functional Motor Scale-Expanded

Exploratory analysis 7: Exclusion of mortality adjustment factors for better health states

Within this analysis, the mortality adjustment factors applied to the better health states were set equal to zero. The ERG also believes that there would be value in exploring alternative simpler parametric models for OS rather than applying complex piecewise methods using multiple external data sources as the company noted that some of these were plausible; however, these models were not reported in the CS.[1]

Exploratory analysis 8: Alternative assumptions regarding long-term transition probabilities

Five alternative scenario analyses were undertaken using the ERG's preferred models to explore the impact of long-term transition probabilities on the cost-effectiveness of nusinersen versus usual care:

(8a) 5% nusinersen patients lose milestones during each cycle (subtracted proportionally from those improving and those remaining in their current state)

(8b) 10% nusinersen patients lose milestones during each cycle (subtracted proportionally from those improving and those remaining in their current state)

(8c) 20% nusinersen patients lose milestones during each cycle (subtracted proportionally from those improving and those remaining in their current state)

(8d) All patients remain their final health state after the end of follow-up in ENDEAR[14]/CHERISH[15] (applied to both treatment groups)

(8e) All patients lose all milestones previously achieved immediately after the end of follow-up in ENDEAR[14]/CHERISH[15] (applied to both treatment groups). The ERG notes that this latter analysis is particularly pessimistic.

5.6.2 Results of the ERG's exploratory analyses – early onset SMA

The results of the ERG's preferred analysis are presented in Table 80. Additional exploratory analyses undertaken using the ERG's preferred model are presented in Table 81. All exploratory analyses were undertaken using the deterministic version of the company's model; the ERG expects that the probabilistic ICERs would be slightly higher.

As shown in Table 80, the application of a common initial distribution has only a minor impact on the ICER for nusinersen. The use of utilities from the vignette study[46] increases the ICER when only patient health gains are considered, but decreases the ICER when caregiver QALY losses are also included. The use of caregiver utilities by SMA type reported by Bastida *et al*[37] increases the ICER considerably. When these amendments are combined within the ERG's preferred analysis, nusinersen is expected to produce 5.2 incremental QALYs at an additional cost of £2,192,722 per patient compared with usual care. The inclusion of caregiver QALY losses reduces the incremental health gain to 3.47 QALYs. The ICERs for nusinersen versus usual care are estimated to be £421,303 per QALY gained (including patient health gains only) and £631,583 per QALY gained (including patient health gains and caregiver QALY losses).

It should be noted that the ERG's preferred analysis does not include any modification to the company's optimistic assumptions regarding survival and motor function trajectories; as such, it is very likely that the true ICERs for nusinersen will be higher. The additional exploratory analyses presented in Table 81 indicate that the use of alternative patient utilities from Bastida *et al*,[37] the exclusion of the mortality adjustment factor and the inclusion of assumptions regarding nusinersen-treated patients losing milestones have the propensity to considerably increase the ICER for nusinersen.

Table 80: ERG preferred analysis, early onset

Option	QALYs (patient)	QALYs (patient+ caregiver)	Cost	Inc. QALYs (patient)	Inc. QALYs (patient+ caregiver)	Inc. cost	ICER (patient)	ICER (patient+ caregiver)
Company's base case								
Nusinersen	7.86	7.61	£2,258,852	5.37	5.44	£2,187,311	**£407,605**	**£402,361**
Usual care	2.49	2.17	£71,540	-	-	-	-	-
ERG exploratory analysis 1 – mean initial distribution applied to both treatment group								
Nusinersen	7.87	7.63	£2,264,226	5.38	5.45	£2,192,722	**£407,417**	**£402,159**
Usual care	2.49	2.18	£71,504	-	-	-	-	-
ERG exploratory analysis 2 - include end-of-life cost								
Nusinersen	7.87	7.63	£2,264,226	5.38	5.45	£2,192,722	**£407,417**	**£402,159**
Usual care	2.49	2.18	£71,504	-	-	-	-	-
ERG exploratory analysis 3 – patient utilities based on vignette study (Lloyd *et al*[46])								
Nusinersen	4.42	4.15	£2,264,226	5.20	5.56	£2,192,722	**£421,303**	**£394,023**
Usual care	-0.78	-1.42	£71,504	-	-	-	-	-
ERG exploratory analysis 4 - caregiver utilities based on Bastida *et al*[37]								
Nusinersen	7.87	5.88	£2,264,226	5.38	3.65	£2,192,722	**£407,417**	**£600,882**
Usual care	2.49	2.23	£71,504	-	-	-	-	-
ERG exploratory analysis 5 - ERG preferred analysis (including ERG analyses 1, 2, 3 and 4)								
Nusinersen	4.42	2.43	£2,264,226	5.20	3.47	£2,192,722	**£421,303**	**£631,583**
Usual care	-0.78	-1.04	£71,504	-	-	-	-	-

ERG - Evidence Review Group; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; Inc. - incremental

Table 81: Additional exploratory analyses undertaken using the ERG preferred model, early onset

Option	QALYs (patient)	QALYs (patient+ caregiver)	Cost	Inc. QALYs (patient)	Inc. QALYs (patient+ caregiver)	Inc. cost	ICER (patient)	ICER (patient+ caregiver)
ERG exploratory analysis 6a - patient utilities based on Bastida *et al*[37]								
Nusinersen	3.87	1.88	£2,264,226	3.23	1.49	£2,192,722	**£679,469**	**£1,467,413**
Usual care	0.64	0.38	£71,504	-	-	-	-	-
ERG exploratory analysis 6b - patient HRQoL estimates based on clinical judgement								
Nusinersen	6.69	4.70	£2,264,226	5.99	4.25	£2,192,722	**£366,289**	**£515,511**
Usual care	0.70	0.44	£71,504	-	-	-	-	
ERG exploratory analysis 7 - no mortality adjustment								
Nusinersen	1.16	0.45	£1,188,262	1.95	1.49	£1,116,759	**£573,922**	**£750,195**
Usual care	-0.78	-1.04	£71,504	-	-	-	-	-
ERG exploratory analysis 8a- 5% nusinersen patients lose milestones each cycle								
Nusinersen	4.00	2.27	£2,229,247	4.79	3.31	£2,157,744	**£450,926**	**£652,213**
Usual care	-0.78	-1.04	£71,504	-	-	-	-	-
ERG exploratory analysis 8b- 10% nusinersen patients lose milestones each cycle								
Nusinersen	3.45	1.98	£2,175,120	4.23	3.02	£2,103,616	**£496,787**	**£696,405**
Usual care	-0.78	-1.04	£71,504	-	-	-	-	-
ERG exploratory analysis 8c- 20% nusinersen patients lose milestones each cycle								
Nusinersen	2.01	1.04	£1,957,022	2.79	2.09	£1,885,518	**£674,945**	**£904,003**
Usual care	-0.78	-1.04	£71,504	-	-	-	-	-
ERG exploratory analysis 8d – all patients stay in final state indefinitely after end of ENDEAR								
Nusinersen	-0.66	-1.03	£1,660,017	0.09	-0.01	£1,588,513	**£16,788,055**	**Dominated**
Usual care	-0.76	-1.02	£71,504	-	-	-	-	-
ERG exploratory analysis 8e – all patients lose all milestones after end of ENDEAR								
Nusinersen	-1.03	-1.37	£567,615	-0.25	-0.33	£496,111	**Dominated**	**Dominated**
Usual care	-0.78	-1.04	£71,504	-	-	-	-	-

ERG - Evidence Review Group; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; Inc. - incremental

5.6.3 Results of the ERG's exploratory analyses – later onset SMA

The results of the ERG's preferred analysis are presented in Table 82. Additional exploratory analyses undertaken using the ERG's preferred model are presented in Table 83.

As shown in Table 82, the application of a common initial distribution and the inclusion of end-of-life care costs slightly reduces the ICER nusinersen in the later onset population. The use of utilities for later onset SMA from the vignette study[46] significantly reduces the ICER for nusinersen. In contrast, the inclusion of caregiver utilities by SMA type from Bastida *et al*[37] drastically reduces the net QALY gains accrued when caregiver QALY losses are included in the analysis. When these amendments are combined within the ERG's preferred analysis, nusinersen is expected to produce 7.37 incremental QALYs at an additional cost of £3,014,078 per patient compared with usual care. The inclusion of caregiver QALY losses reduces the net incremental health gain to 4.76 QALYs. The ICERs for nusinersen versus usual care are estimated to be £408,769 per QALY gained (including patient health gains only) and £632,850 per QALY gained (including patient health gains and caregiver QALY losses).

Again, the ERG's preferred analysis for the later onset population does not include any modification to the company's optimistic assumptions regarding survival and motor function trajectories; as such, it is very likely that the true ICERs for nusinersen will be higher. The additional exploratory analyses presented in Table 83 indicate that the use of alternative patient utilities from Bastida *et al*,[37] the use of HRQoL estimates from the ERG's clinical advisors, and the inclusion of assumptions regarding nusinersen-treated patients losing milestones have the propensity to result in considerably higher ICERs for nusinersen. The ERG notes that the exclusion of the mortality adjustment factor results in less favourable ICERs for nusinersen, however the impact is less marked than that for the early onset model.

Table 82: ERG preferred analysis, later onset

Option	QALYs (patient)	QALYs (patient+ caregiver)	Cost	Inc. QALYs (patient)	Inc. QALYs (patient+ caregiver)	Inc. cost	ICER (patient)	ICER (patient+ caregiver)
Company's base case								
Nusinersen	16.88	15.66	£3,148,754	2.37	3.30	£2,964,442	**£1,252,991**	**£898,164**
Usual care	14.52	12.36	£184,312	-	-	-	-	-
ERG exploratory analysis 1 – mean initial distribution applied to both treatment group								
Nusinersen	16.95	15.76	£3,200,341	2.47	3.47	£3,014,655	**£1,221,051**	**£869,639**
Usual care	14.48	12.29	£185,686	-	-	-	-	-
ERG exploratory analysis 2 - include end-of-life cost								
Nusinersen	16.95	15.76	£3,203,766	2.47	3.47	£3,014,078	**£1,220,817**	**£869,472**
Usual care	14.48	12.29	£189,688	-	-	-	-	-
ERG exploratory analysis 3 – patient utilities based on vignette study (Lloyd *et al*[46])								
Nusinersen	8.53	7.34	£3,200,341	7.37	8.37	£3,014,655	**£408,847**	**£360,122**
Usual care	1.15	-1.03	£185,686	-	-	-	-	-
ERG exploratory analysis 4 - caregiver utilities based on Bastida *et al*[37]								
Nusinersen	16.95	13.54	£3,200,341	2.47	-0.14	£3,014,655	**£1,221,051**	**Dominated**
Usual care	14.48	13.68	£185,686	-	-	-	-	-
ERG exploratory analysis 5 - ERG preferred analysis (including ERG analyses 1, 2, 3 and 4)								
Nusinersen	8.53	5.12	£3,203,766	7.37	4.76	£3,014,078	**£408,769**	**£632,850**
Usual care	1.15	0.36	£189,688	-	-	-	-	-

ERG - Evidence Review Group; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; Inc. - incremental

Table 83: Additional exploratory analyses undertaken using the ERG preferred model, later onset

Option	QALYs (patient)	QALYs (patient+ caregiver)	Cost	Inc. QALYs (patient)	Inc. QALYs (patient+ caregiver)	Inc. cost	ICER (patient)	ICER (patient+ caregiver)
ERG exploratory analysis 6a - patient utilities based on Bastida *et al*[37]								
Nusinersen	6.97	3.56	£3,203,766	4.80	2.19	£3,014,078	**£627,612**	**£1,375,278**
Usual care	2.16	1.37	£189,688	-	-	-	-	-
ERG exploratory analysis 6b - patient HRQoL estimates based on clinical judgement								
Nusinersen	15.44	12.03	£3,203,766	3.54	0.93	£3,014,078	**£850,597**	**£3,231,764**
Usual care	11.89	11.10	£189,688	-	-	-	-	
ERG exploratory analysis 7 - no mortality adjustment								
Nusinersen	7.49	4.42	£2,929,515	6.34	4.07	£2,739,998	**£432,191**	**£673,128**
Usual care	1.15	0.35	£189,517	-	-	-	-	-
ERG exploratory analysis 8a- 5% nusinersen patients lose milestones each cycle								
Nusinersen	6.78	4.03	£2,756,403	5.63	3.67	£2,566,715	**£455,934**	**£699,062**
Usual care	1.15	0.36	£189,688					
ERG exploratory analysis 8b- 10% nusinersen patients lose milestones each cycle								
Nusinersen	4.97	2.88	£2,296,390	3.81	2.52	£2,106,702	**£552,283**	**£834,754**
Usual care	1.15	0.36	£189,688					
ERG exploratory analysis 8c- 20% nusinersen patients lose milestones each cycle								
Nusinersen	2.49	1.28	£1,539,734	1.34	0.92	£1,350,046	**£1,011,268**	**£1,459,562**
Usual care	1.15	0.36	£189,688					
ERG exploratory analysis 8d – all patients stay in final state indefinitely after end of CHERISH								
Nusinersen	2.85	1.72	£2,993,988	0.81	0.73	£2,809,679	**£3,465,629**	**£3,831,118**
Usual care	2.04	0.98	£184,309	-	-	-	-	-
ERG exploratory analysis 8e – all patients lose all milestones after end of CHERISH								
Nusinersen	0.91	0.20	£721,228	0.04	0.03	£529,189	**£14,994,339**	**£18,436,952**
Usual care	0.88	0.17	£192,038	-	-	-	-	-

ERG - Evidence Review Group; QALY - quality-adjusted life year; ICER - incremental cost-effectiveness ratio; Inc. - incremental

5.7 Discussion

The CS[1] includes a systematic review of published economic evaluations of treatments for SMA together with two *de novo* health economic evaluations of nusinersen for the treatment of early onset and later onset SMA. The company's review did not identify any economic evaluations of treatments for SMA.

The company's early onset model assesses the cost-effectiveness of nusinersen versus usual care for the treatment of patients with early onset SMA (initial age = 5.58 months), based on the ENDEAR trial. The incremental health gains, costs and cost-effectiveness of nusinersen are evaluated over a 60-year time horizon from the perspective of the NHS and PSS. The company's early onset model adopts a state transition approach, with health states defined by motor function milestones based on the HINE-2 instrument. The model parameters were largely informed by: HINE-2 and CHOP INTEND outcomes collected within ENDEAR;[14] mortality outcomes from ENDEAR[14] and other observational data (Gregoretti *et al*,[31] Zerres *et al*[33] and general population life tables[32]); a mapping exercise to translate PedsQL outcomes collected in the CHERISH trial to the EQ-5D;[1, 36] a cross-sectional study of the costs and caregiver HRQoL impacts of SMA[37] and standard costing sources.[39] The model assumes that treatment using nusinersen will be discontinued for patients who do not achieve any milestones (or lose previously achieved milestones) after 13 months, and for patients undergoing scoliosis surgery who cannot subsequently receive nusinersen administration via lumbar puncture. The company's early onset model employs two key assumptions: (i) after month 13, nusinersen-treated patients who reach health states consistent with Type II/III SMA milestones gain an additional survival advantage, and (ii) after month 13, the motor function of nusinersen-treated patients cannot deteriorate, whilst the motor function of patients receiving usual care cannot improve.

Based on a re-run of the probabilistic version of the company's early onset model by the ERG, nusinersen is expected to generate an additional 5.29 QALYs at an additional cost of £2,160,048 per patient; the corresponding ICER for nusinersen versus usual care is £408,712 per QALY gained. The inclusion of caregiver QALY losses leads to a slightly lower probabilistic ICER of £404,270 per QALY gained. The probability that nusinersen produces more net benefit than usual care at WTP thresholds below £337,000 per QALY gained is approximately zero. The company's subgroup analyses suggest that the cost-effectiveness profile for nusinersen may be improved in early onset SMA patients with shorter disease duration (≤12 weeks subgroup ICER≈£375,000 per QALY gained, ICER includes patient health gains only).

The company's later onset model assesses the cost-effectiveness of nusinersen versus usual care for the treatment of patients with later onset SMA (initial age = 43.71 months), based on the CHERISH trial. The incremental health gains, costs and cost-effectiveness of nusinersen are evaluated over an 80-year

time horizon from the perspective of the NHS and PSS. The company's later onset model adopts a state transition approach, with health states defined by motor function milestones based on the HFMSE instrument. The model parameters were largely informed by: HFMSE outcomes collected within CHERISH;[15] mortality outcomes from CHERISH[15] and other observational data (Zerres *et al*[33] and general population life tables[32]), and the same cost and HRQoL sources as those used in the early onset model[37, 39] (see above). The company's model assumes that treatment using nusinersen will be discontinued for patients who do not achieve milestones beyond the Sits without support but does not roll state after 15 months, and for patients undergoing scoliosis surgery who cannot subsequently receive nusinersen administration via lumbar puncture. The later onset model includes two key assumptions: (i) after month 15, patients in either treatment group who reach health states consistent with Type III SMA milestones 15 gain an additional survival advantage, and (ii) after month 15, the motor function of nusinersen-treated patients cannot deteriorate, whilst the motor function of patients receiving usual care cannot improve.

Based on a re-run of the probabilistic version of the company's later onset model by the ERG, nusinersen is expected to generate an additional 2.28 QALYs at an additional cost of £2,938,441 per patient: the corresponding ICER for nusinersen versus usual care is £1,286,149 per QALY gained. The inclusion of caregiver QALY losses leads to a markedly lower probabilistic ICER of £933,088 per QALY gained. The probability that nusinersen produces more net benefit than usual care is approximately zero even at WTP thresholds of £500,000 per QALY gained. The company's subgroup analyses are inconclusive with respect to whether the cost-effectiveness profile for nusinersen is improved for later onset SMA patients with shorter disease duration (<25 months).

The ERG's critical appraisal identified a number of issues relating to the company's economic analyses and the evidence used to inform them. The most pertinent of these include: (i) the absence of economic evidence relating to Type 0 and Type IV SMA; (ii) the unnecessary complexity of the company's implemented models; (iii) highly favourable assumptions regarding the expected trajectory of nusinersen-treated patients through modelled motor milestone health states; (iv) highly favourable assumptions regarding the expected survival of nusinersen-treated patients; (v) poor face validity of patient utilities used in the models, and (vi) arbitrary calculations underpinning the caregiver disutilities used in the models.

The ERG undertook eight sets of exploratory analyses using the deterministic version of the company's models. The ERG's preferred scenario includes: (i) the use of a common initial distribution across health states for both treatment groups; (ii) the inclusion of end-of-life costs for the later onset population; (iii) the use of patient utilities from the vignette study[46] and (iv) the application of caregiver utilities by SMA type from Bastida *et al*[37] to states relating to SMA milestones in both the early and later onset models.

Importantly, the preferred analyses do not address the ERG's concerns regarding the optimistic assumptions underpinning the company's modelled survival and motor function trajectories; as such, it is very likely that the true ICERs for nusinersen will be higher. In order to address this uncertainty, additional sensitivity analyses were undertaken to explore the use of alternative patient utilities, the exclusion of the mortality adjustment factors and alternative long-term transition probabilities.

The ERG's preferred analyses within the early onset population results in ICERs for nusinersen versus usual care of £421,303 per QALY gained (including patient health gains only) and £631,583 per QALY gained (including patient health gains and caregiver QALY losses). The ERG's additional exploratory analyses lead to ICERs ranging from £366,289 per QALY gained to dominated.

The ERG's preferred ICER for nusinersen versus usual care in the later onset population is estimated to be £408,769 per QALY gained (including patient health gains only). The inclusion of caregiver QALY losses increases the ICER to £632,850 per QALY gained. The ERG's additional exploratory analyses lead to ICERs ranging from £432,191 per QALY gained to in excess of £18.4million per QALY gained.

6. END OF LIFE

NICE end of life supplementary advice should be applied in the following circumstances and when all the criteria referred to below are satisfied:

- The treatment is indicated for patients with a short life expectancy, normally less than 24 months and;
- There is sufficient evidence to indicate that the treatment offers an extension to life, normally of at least an additional 3 months, compared to current NHS treatment.

The CS[1] makes the case that NICE's end of life criteria apply to the infantile onset SMA population, but not for the later onset population. The ERG agrees that the later onset population does not meet the end of life criteria. The evidence presented in this chapter therefore relates only to the early onset (Type I) SMA population. Table 84 presents the main evidence for nusinersen relating to NICE's end of life criteria; additional evidence from natural history studies is presented in Table 32 of the CS.[1]

Table 84: Evidence supporting the application of end of life criteria presented in the CS (adapted from CS, Table 31)

Criterion	Evidence available
Nusinersen is indicated for patients with a short life expectancy, normally less than 24 months	Survival is highly dependent upon the nature and extent of supportive care, which may vary by country, institution and physician and patient preference. The median age for death or permanent respiratory support (a composite endpoint used in clinical trials and natural history studies in this population) is approximately 9–13 months.[60] [redacted]
There is sufficient evidence to indicate that nusinersen offers an extension to life, normally of at least an additional 3 months, compared with current NHS treatment	Infants in the ENDEAR study who received nusinersen had a significantly higher likelihood of EFS (final analysis: HR for death or the use of permanent assisted ventilation, 0.53; *p*=0.005) and OS (HR for death, 0.37; *p*=0.004) than infants who underwent a sham procedure, despite the fact that more infants in the nusinersen group than in the control group were receiving ventilatory support at baseline. The median time to death or the use of permanent assisted ventilation was 22.6 weeks in the control group and was not reached in the nusinersen group; the median time to death was not reached in either group (ITT population at end of study). In addition, at the latest data cut-off, all pre-symptomatic children in NURTURE (including those with 2 *SMN2* copy number) are still alive.

ITT- intention-to-treat; SMA - spinal muscular atrophy; SMN – survival motor neuron

With respect to the criterion relating to short life expectancy, the CS[1] makes the following points:

- There are no published studies on the natural history of SMA in English or UK populations.

- Survival of Type I SMA patients is highly dependent upon the nature and extent of supportive care received. This may vary between countries, institutions and according to physician and patient preferences.
- "Proactive" supportive care can prolong survival (for example, due to nutritional support using gastrostomy tubes and NIV or tracheostomy/ventilator support). [redacted]
- Changes in standard of care over time and the variable use of tracheostomy and invasive mechanical ventilation lead to variations in reported survival rates.
- The CS[1] makes the case that "survival free of permanent ventilation", which is generally accepted as intubation or tracheostomy with mechanical ventilation or >16 hours/day NIV support for >14 consecutive days (16+/14+) in the absence of an acute reversible illness or following surgery, may be a more relevant endpoint, as permanent ventilation may not be provided in England.
- On the basis of natural history studies included in the CS,[1] the median time to death or permanent respiratory support is reported to be 9-13 months. [redacted]

  Further details of the studies used to inform these estimates are provided in CS[1] Table 32.

The clinical advisors to the ERG did not share the same view regarding the expected survival of Type I SMA patients. One clinical advisor considered that the low survival rates for Type I SMA patients cited in the CS[1] are outdated and reflect an era before the use of ventilation, and noted that some less severe Type Ic SMA patients diagnosed between 3 and 6 months may survive to school age. The shift is seen in a greater proportion receiving ventilation. In contrast, the second clinical advisor considered that the mean survival for Type I SMA patients is likely to be less than 2 years and noted that she did not have any SMA patients who were older than 2 years of age (almost all of these patients had or have Type Ib SMA). This advisor noted that practice has changed and that the availability of improved expertise and equipment with NIV to support younger children will lead to longer survival. The clinical advisors commented that survival free of permanent ventilation is a useful surrogate outcome for severe impairment and weakness which allows for comparisons between studies. However, the advisors considered that ventilation for >16 hours a day is arguably better than death and that parents of infants with SMA may also share this view.

The ERG notes that the mean predicted survival for the usual care group of the company's early onset model is 3.87 years; on the basis of the company's model, the short life expectancy criterion is not met. As discussed in Section 5.5.3, one of the ERG's clinical advisors considered this predicted survival trajectory to be overly optimistic and expected the function to be steeper. Despite the differences in clinical opinion received by the ERG, it should be noted that the company's statement that "patients rarely survive to their second birthday" is inconsistent with the company's own model predictions.

With respect to the criterion relating to a life extension of 3 months or greater, the CS[1] notes the following:

- In ENDEAR, the median time to death or the use of permanent assisted ventilation was 22.6 weeks in the control group and was not reached in the nusinersen group. Overall, the risk of death or the use of permanent assisted ventilation was 47% lower in the nusinersen group than in the control group (hazard ratio, 0.53; 95% CI, 0.32–0.89; *p*=0.005)
- Despite a poorer prognosis in the nusinersen group of ENDEAR at baseline, the overall risk of death was 63% lower in the nusinersen group compared with the sham group (HR=0.37 [95% CI: 0.18, 0.77])
- Despite a poorer prognosis in the nusinersen group of ENDEAR at baseline, a lower proportion of infants receiving nusinersen received permanent assisted ventilation compared with those receiving sham (23% versus 32%, HR=0.66, *p*=0.13)
- All pre-symptomatic infants in NURTURE were alive and none had required respiratory intervention (invasive or NIV for ≥6 hours/day, continuously for ≥7 days or tracheostomy).

The company's early onset model suggests that nusinersen extends mean survival by 9.12 years compared with usual practice.

Both clinical advisors noted that there was considerable uncertainty regarding the expected survival duration for Type I patients receiving nusinersen and considered the model-predicted survival trajectory for the nusinersen group to be overly optimistic (see Section 5.5.3). However, they did believe that it was plausible that nusinersen would extend survival by at least 3 months. The clinical advisors also noted that there were infants treated with nusinersen who had gains in motor function but progressive deterioration in respiratory function; this has implications for long-term survival, especially as it is not yet clear whether these motor milestones will be maintained.

7. OVERALL CONCLUSIONS

Clinical effectiveness conclusions

The CS[1] did not contain a systematic review as would be expected in a submission to the NICE STA process. As such, it is not entirely certain that all nusinersen studies have been identified, although the ERG is confident that all relevant studies of nusinersen for SMA have been included in the CS. No information was provided for the BSC comparator listed in the NICE scope.[12] Two key RCTs were presented in the CS: ENDEAR, in early (infantile) onset SMA patients and CHERISH, in later onset SMA patients.

Nusinersen appears to provide significant clinical benefit to patients and the safety profile reported in the studies was acceptable and generally more favourable than that for the sham control group. The patient groups in the study arms for the ENDEAR and CHERISH studies were broadly similar although the nusinersen groups had more severe symptoms and longer duration of treatment.

Cost-effectiveness conclusions

With respect to the early onset model, the ERG's preferred assumptions increase the ICER for nusinersen versus usual care (including patient health gains only) from £407,605 per QALY gained (the company's base case) to £421,303 per QALY gained. When caregiver QALY losses are included in the analysis, the ERG's preferred assumptions increase the ICER from £402,361 per QALY gained (the company's base case) to £631,583 per QALY gained.

With respect to the later onset model, the ERG's preferred assumptions decrease the ICER for nusinersen versus usual care (including patient health gains only) from £1,252,991 per QALY gained (the company's base case) to £408,769 per QALY gained. When caregiver QALY losses are included in the analysis, the ERG's preferred assumptions reduces the ICER from £898,164 per QALY gained (the company's base case) to £632,850 per QALY gained. The main driver of these differences between the ICERs generated by the company and the ERG relates to the HRQoL impact on patients and caregivers.

The ERG's preferred analyses do not include any modification to the optimistic assumptions underpinning the company's modelled survival and motor function trajectories. The ERG's additional exploratory analyses show that the use of less optimistic assumptions has the propensity to markedly increase the ICERs for nusinersen in both populations.

The long-term probabilities of achieving, maintaining and losing motor function for nusinersen-treated patients, the long-term survival advantage of nusinersen and the relationship between motor function

and HRQoL in patients with SMA are all highly uncertain. However, the ERG also notes that given the acquisition cost of nusinersen, the level of decision uncertainty with respect to NICE's usual thresholds for cost-effectiveness is low.

7.1 Implications for research

Longer-term studies are required to determine the full impact of nusinersen on survival and motor function outcomes and AEs for patients with SMA; SHINE may provide useful information on these outcomes. Future clinical studies of nusinersen for the treatment of SMA should include a preference-based measure of HRQoL for patients (if applicable) and/or caregivers. Future research studies may also be worthwhile to determine whether nusinersen offers benefits to patients with Type 0 SMA and patients with Type IV SMA.

8. REFERENCES

1. Biogen Idec Ltd. Nusinersen for treating spinal muscular atrophy [ID1069] - company's evidence submission to NICE. Biogen; Berkshire, 2018.
2. Biogen Idec. Nusinersen for treating spinal muscular atrophy [ID1069] - company's response to clarification questions from the ERG. Biogen; Berkshire, 2018.
3. Biogen Idec. Nusinersen for treating spinal muscular atrophy [ID1069] - additional clarification responses provided by the company. Biogen; Berkshire, 2018.
4. European Medicines Agency. Nusinersen - European Public Assessment Report and Summary of Product Characteristics (annex). EMA: London; 2017.
5. Darras BT. Spinal Muscular Atrophies. *Pediatric Clinics of North America* 2015;62(3):743-66.
6. European Medicines Agency. First medicine for spinal muscular atrophy. EMA:London; 2017.
7. Arnold WD, Kassar D, Kissel JT. Spinal muscular atrophy: Diagnosis and management in a new therapeutic era. *Muscle & Nerve* 2015;51(2):157-67.
8. Farrar MA, Park SB, Vucic S, Carey KA, Turner BJ, Gillingwater TH, *et al*. Emerging therapies and challenges in spinal muscular atrophy. *Annals of Neurology* 2017;81(3):355-68.
9. Ogino S, Wilson RB, Gold B. New insights on the evolution of the SMN1 and SMN2 region; Simulation and meta-analysis for allele and haplotype frequency calculations. European *Journal of Human Genetics* 2004;12(12):1015-23.
10. Mercuri E, Finkel RS, Muntoni F, Wirth B, Montes J, Main M, *et al*. Diagnosis and management of spinal muscular atrophy: part 1: recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscular Disorders* 2018;28(2):103-15.
11. Finkel RS, Mercuri E, Meyer OH, Simonds A, Schroth M, Graham RJ, *et al*. Diagnosis and management of Spinal Muscular Atrophy: Part 2: pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. *Neuromuscular Disorders* 2018;28(3):197-207.
12. National Institute for Health and Care Excellence (NICE). Nusinersen for treating spinal muscular atrophy – final scope. NICE; London, 2018. Available from: https://www.nice.org.uk/guidance/gid-ta10281/documents/final-scope (Accessed 01/04/2018).
13. BMJ Group and the Royal Pharmaceutical Society of Great Britain. British National Formulary. London; 2018.
14. Ionis Pharmaceuticals. Clinical study report - A Phase 3, randomized, double-blind, sham-procedure controlled study to assess the clinical efficacy and safety of ISIS 396443 administered intrathecally in patients with infantile-onset spinal muscular atrophy [ENDEAR]. Ionis Pharmaceuticals, Inc; Carlsbad, CA, 2017.
15. Ionis Pharmaceuticals. Clinical study report - Full - Final: A Phase 3, randomized, double-blind, sham-procedure controlled study to assess the clinical efficacy and safety of ISIS 396443 administered intrathecally in patients with later-onset spinal muscular atrophy [CHERISH]. Ionis Pharmaceuticals, Inc; Carlsbad, CA, 2017.
16. Centre for Reviews and Dissemination (CRD). Systematic review: CRD's guidance for undertaking reviews in health care. University of York; York, 2009.
17. National Institute for Health and Care Excellence. Single technology appraisal: User guide for company evidence submission template. 2015. NICE; London. Available from: https://www.nice.org.uk/process/pmg24/chapter/instructions-for-companies (Accessed 11/05/2018).
18. National Institute for Health and Care Excellence. Methods for the development of NICE public health guidance. 2012. NICE; London. Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-f-quality-appraisal-checklist-quantitative-intervention-studies (Accessed 11/05/2018).
19. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, *et al*. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. The Ottowa Hospital; Ottowa, 2011. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp (Accessed 11/05/2018).

20. Finkel RS, Mercuri E, Darras BT, Connolly AM, Kuntz NL, Kirschner J, *et al.* Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *New England Journal of Medicine* 2017;377(18):1723-32.
21. Biogen. Clinical study report - Full - Interim 2 - An open-label study to assess the efficacy, safety, tolerability, and pharmacokinetics of multiple doses of ISIS 396443 delivered intrathecally to subjects with genetically diagnosed and presymptomatic spinal muscular atrophy. Biogen Idec Research Limited; Berkshire, 2017.
22. ClinicalTrials.gov. A study to assess the efficacy and safety of nusinersen (ISIS 396443) in infants with spinal muscular atrophy (ENDEAR). 2017. https://clinicaltrials.gov/ct2/show/results/NCT02193074?term=NCT02193074&rank=1&sect=Xg0156&view=results (Accessed 08/05/2018).
23. Mercuri E, Darras BT, Chiriboga CA, Day JW, Campbell C, Connolly AM, *et al.* Nusinersen versus sham control in later-onset spinal muscular atrophy. *New England Journal of Medicine* 2018;378(7):625-35.
24. ClinicalTrials.gov. A study to assess the efficacy and safety of nusinersen (ISIS 396443) in participants with later-onset spinal muscular atrophy (SMA) (CHERISH). 2018. https://clinicaltrials.gov/ct2/show/results/NCT02292537?term=NCT02292537&rank=1&sect=Xf0156&view=results#outcome9 (Accessed 08/05/2018).
25. Darras BT, Chiriboga CA, Montes J, Swoboda KJ, Johnson N, Iannaccone ST, *et al.* Nusinersen in treatment-naïve patients with later-onset spinal muscular atrophy (SMA): efficacy results from a phase 1b/2a multicentre study (CS2) and its open-label extension (CS12). Paper presented at: 21st International Congress World Muscle Society; Granada, Spain.
26. Chiriboga CA, Darras BT, Montes J, Swoboda KJ, Johnson N, Iannaccone ST, *et al.* Nusinersen in treatment-naive children with later-onset spinal muscular atrophy (SMA): Efficacy results from a Phase 1b/2a multicenter study (CS2) and its open-label extension (CS12). Paper presented at: Annual Spinal Muscular Atrophy Conference; Orlando, Florida.
27. Crawford TO, De Vivo DC, Betini E, Hwu W-L, Foster R, Gheuens S, *et al.* Efficacy and safety of nusinersen in genetically diagnosed infants with presymptomatic spinal muscular atrophy (SMA): Results from the second interim analysis of the ongoing, Phase 2 NURTURE study. Paper presented at: 2017 Annual Spinal Muscular Atrophy Conference; Orlando, FL.
28. Mercuri E, Finkel R, Farrar M, Richman S, Foster R, Hughes S, *et al.* Infants and children with SMA treated with nusinersen in clinical trials: An integrated safety analysis. Paper presented at: 2017 Annual Spinal Muscular Atrophy Conference; Orlando, FL.
29. Center for Drug Evaluation and Research. Safety Team Leader Review - Spinraza (nusinersen). 2017. Available from: https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/209531Orig1s000MedR.pdf (Accessed 24/05/2018).
30. De Sanctis R, Coratti G, Pasternak A, Montes J, Pane M, Mazzone ES. Developmental milestones in type I spinal muscular atrophy. *Neuromuscular Disorders* 2016;26(11):754-9.
31. Gregoretti C, Ottonello G, Chiarini Testa MB, Mastella C, Rava L, Bignamini E, *et al.* Survival of patients with spinal muscular atrophy type 1. *Pediatrics* 2013;131(5):e1509-14.
32. Office for National Statistics. National life tables: England and Wales. ONS; London, 2017. Available from: https://www.ons.gov.uk (Accessed 01/05/2018).
33. Zerres K, Rudnik-Schöneborn S, Forrest E, Lusakowska A, Borkowska J, Hausmanowa-Petrusewicz I. A collaborative study on the natural history of childhood and juvenile onset proximal spinal muscular atrophy (type II and III SMA): 569 patients. *Journal of Neurological Science* 1997;146(1):67-72.
34. Finkel RS, Chiriboga CA, Vajsar J, Day JW, Montes J, De Vivo DC, *et al.* Treatment of infantile-onset spinal muscular atrophy with nusinersen: a phase 2, open-label, dose-escalation study. *The Lancet* 2016;388(10063):3017-26.
35. Haaker G, Fujak A. Proximal spinal muscular atrophy: current orthopedic perspective. *The Application of Clinical Genetics* 2013;6(11):113-20.
36. Khan KA, Petrou S, Rivero-Arias O, Walters SJ, Boyle SE. Mapping EQ-5D utility scores from the PedsQL™ generic core scales. *Pharmacoeconomics* 2014;32(7):693-706.

37. Bastida JL and Research Team. Social economic costs and health-related quality of life in patients with spinal muscular atrophy in Europe. University of Castilla-La Mancha (Ciudad Real, Spain); Toledo, Spain; 2016.
38. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value in Health* 2010;13(5):509-18.
39. Department of Health. NHS Reference Costs 2015/16. DH; London: DH, 2016.
40. National Institute for Health and Care Excellence. NICE guideline: End of life care for infants, children and young people with life-limiting conditions: planning and management. NICE; London, 2016.
41. Foundation for Statistical Computing. R. In. Vienna, Austria.
42. Graf Erika, Schmoor Claudia, Sauerbrei Willi, Schumacher Martin. Assessment and comparison of prognostic classification schemes for survival data. *Statistics in Medicine* 1999;18(17-18):2529-45.
43. Zerres K, Rudnik-Schöneborn S. Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications. *Archives of Neurology* 1995;52(5):518-23.
44. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Medical Research Methodology* 2012;12.
45. Biogen Idec. Spinal Muscular Atrophy (SMA) – UK Advisory Board Meeting, 11th September 2017, London, UK. Meeting summary report. Biogen; Berkshire, 2017.
46. Lloyd A, Gallop K, Thompson R, Vaidya S, Teynor M. Estimation of the health-related quality of life benefits of treatment for spinal muscular atropy (SMA). *Value in Health* 2017;20(9):A559.
47. Ren S, Minton J, Whyte S, Latimer NR, Stevenson M. A new approach for sampling ordered parameters in probabilistic sensitivity analysis. *PharmacoEconomics* 2018;36(3):341-7.
48. O'Hagen JM, Glanzman AM, McDermott MP, Ryan PA, Flickinger J, Quigley J, *et al.* An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. *Neuromuscular Disorders* 2007;17(9-10):693-7.
49. Bladen CL, Thompson R, Jackson JM, Garland C, Wegel C, Ambrosini A, *et al.* Mapping the differences in care for 5,000 Spinal Muscular Atrophy patients, a survey of 24 national registries in North America, Australasia and Europe. *Journal of Neurology* 2014;261(1):152-63.
50. Kaufmann P, McDermott MP, Darras BT, Finkel RS, Sproule DM, Kang PB. Prospective cohort study of spinal muscular atrophy types 2 and 3. *Neurology* 2012;79(18):1889-97.
51. Klug C, Schreiber-Katz O, Thiele S, Schorling A, Zowe J, Reilich P, *et al.* Disease burden of spinal muscular atrophy in Germany. *Orphanet Journal of Rare Diseases* 2016;11:58-67.
52. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes. 3rd edn. Oxford University Press; New York, 2005.
53. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, *et al.* Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value in Health* 2003;6(1):9-17.
54. National Institute for Health and Care Excellence Guide to the methods of technology appraisal 2013. NICE; London, 2013. Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781 (Accessed 01/04/2018).
55. Guyot P, Ades AE, Beasley M, Lueza B, Pignon J-P, Welton NJ. Extrapolation of survival curves from cancer trials using external information. *Medical Decision Making* 2017;37(4):353-66.
56. Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, *et al.* Extrapolating survival from randomized trials using external data: A review of methods. *Medical Decision-Making* 2017;37(4):377-90.

57. Bach JR. There are other ways to manage spinal muscular atrophy type 1. *Chest* 2005;127:1463; author reply -4.
58. Bach JR. Medical considerations of long-term survival of Werdnig-Hoffmann disease. *American Journal of Physical Medicine and Rehabilitation* 2007;86(5):349-55.
59. Ottonello G, Mastella C, Franceschi A, Bosticco D, Wolfler A, Pedemonte M, *et al.* Spinal muscular atrophy type 1: avoidance of hospitalization by respiratory muscle support. *American Journal of Physical Medicine and Rehabilitation* 2011;90(11):895-900.
60. Chiriboga CA. Nusinersen for the treatment of spinal muscular atrophy. *Expert Review of Neurotherapeutics* 2017;17(10):955-62.
61. Biogen Idec. Data on file - survey inputs. Biogen; Berkshire, 2017.
62. Glanzman AM, Mazzone E, Main M, Pelliccioni M, Wood J, Swoboda KJ, *et al.* The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND): test development and reliability. *Neuromuscular Disorders* 2010;20(3):155-61.
63. Pera MC, Coratti G, Forcina N, Mazzone ES, Scoto M, Montes J, *et al.* Content validity and clinical meaningfulness of the HFMSE in spinal muscular atrophy. *BMC Neurology* 2017;17(39).
64. WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: windows of achievement for six gross motor development milestones. *Acta Paediatrica Supplement* 2006;suppl 450:86-95.

9. APPENDICES

Appendix 1: Exclusion criteria for ENDEAR and CHERISH studies

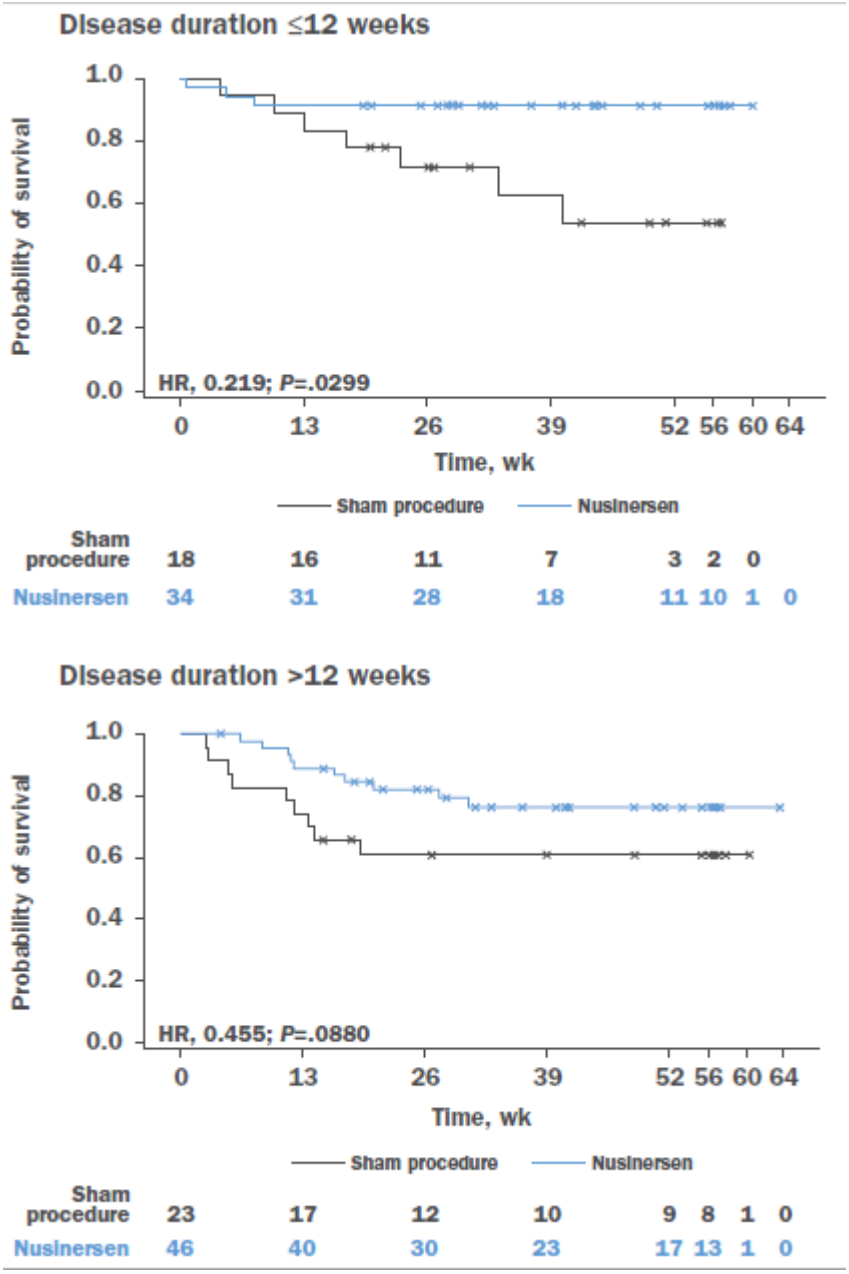
Table 85: Exclusion criteria for ENDEAR and CHERISH (adapted from Table 8, CS, page 34)

Exclusion criteria	ENDEAR	CHERISH
	• Peripheral oxygen desaturation (oxygen saturation below 96% without ventilation support) during screening • SMA symptoms within the first week of birth • Presence of an active infection requiring systemic antiviral or antibacterial treatment during screening • History of brain or spinal cord disease that would interfere with lumbar puncture, CSF circulation, or safety assessments • Presence of an implanted CSF drainage shunt or central nervous system catheter; abnormalities in haematology or clinical chemistry parameters at screening that would prevent inclusion as assessed by the site investigator • Treatment of SMA with an investigational drug, biological agent, or device within 30 days of screening • History of gene therapy, prior ASO therapy, or cell transplantation • The parent/guardian is unable to understand a basic description of the study or does not agree to comply with the schedule of assessments as defined by the protocol • The infant's caregiver does not adhere to the standard-of-care guidelines • Presence of a medical condition that would interfere with the infant's ability to participate in the study as assessed by the site investigator.	• Respiratory insufficiency at screening (defined by the medical necessity for invasive or non-invasive ventilation for >6 hours during a 24-hour period) • Medical necessity for a gastric feeding tube, where most feeds are given by this route; severe contractures (any contracture that, according to the investigator, could interfere with HFMSE) or severe scoliosis (Cobb Angle >40 degrees) evident on X-ray examination at screening • Hospitalisation for surgery (i.e., scoliosis surgery, other surgery), pulmonary event, or nutritional support within 2 months of screening or planned during the duration of the study • Presence of an untreated or inadequately treated active infection

ASO – antisense oligonucleotide, CSF - cerebrospinal fluid; SMA - spinal muscular atrophy

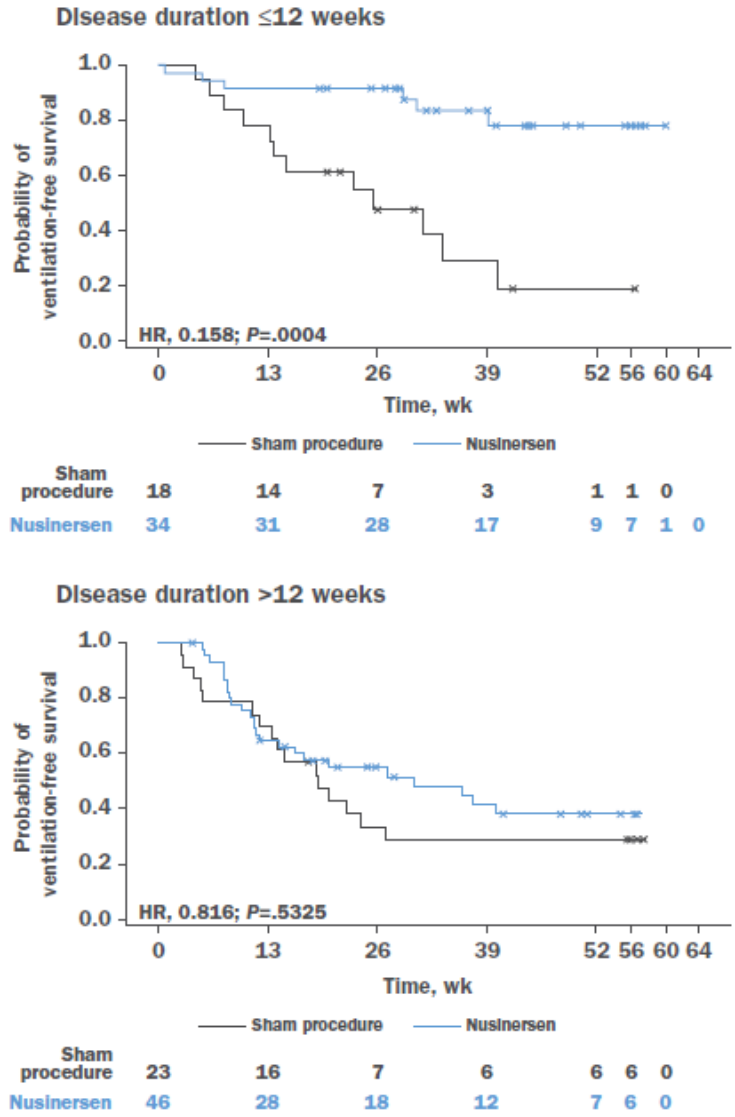
Appendix 2: Overall survival and event free survival by disease duration subgroup

Figure 23: Overall survival by disease duration (reproduced from CS, Appendix E, Figure 8)



HR- hazard ratio

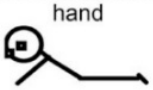
Figure 24: Event free survival by disease duration subgroup (reproduced from CS, Appendix E, Figure 9)



HR- Hazard ratio

Appendix 3: Instruments used to inform transition probabilities within the company's models

Table 86: HINE-2 classification (reproduced from CS Figure 5)

Head control	Unable to maintain head upright normal up to 3m	Wobbles normal up to 4m	Maintained upright all the time normal from 5m		
Sitting	Cannot sit	With support at hips  normal at 4m	Props  normal at 6m	Stable sit  normal at 7-8m	Pivots (rotates)  normal at 9m
Voluntary grasp – note side	No grasp	Uses whole hand	Index finger and thumb but immature grasp	Pincer grasp	
Ability to kick in supine	No kicking	Kicks horizontally but legs do not lift	Upward (vertically)  normal at 3m	Touches leg  normal at 4-5m	Touches toes  normal at 5-6m
Rolling	No rolling	Rolling to side (normal at 4m)	Prone to supine (normal at 6 m)	Supine to prone (normal at 6 m)	
Crawling or bottom shuffling	Does not lift head	On elbow  (normal at 3 m)	On outstretched hand  (normal at 4m)	Crawling flat on abdomen  (normal at 8m)	Crawling on hands and knees  (normal at 10m)
Standing	Does not support weight	Supports weight (normal at 4m)	Stands with support (normal at 7m)	Stands unaided (normal at 12m)	
Walking		Bouncing (normal at 6m)	Cruising (walks holding on) (normal at 12m)	Walking independently (normal by 15m)	

Source: De Sanctis et al[30] and CS[1]

Table 87: CHOP INTEND domains and scoring

CHOP INTEND item	CHOP INTEND activities	Scoring within CHOP INTEND item
1	Spontaneous movement (upper extremity)	0-4
2	Spontaneous movement (lower extremity)	0-4
3	Hand grip	0-4
4	Head in midline with visual stimulation	0-4
5	Hipadductors	0, 2 or 4
6	Rolling elicited from legs	0-4
7	Rolling elicited from arms	0-4
8	Shoulder and elbow flexion and horizontal abduction	0-4
9	Shoulder flexion and elbow flexion	0-4
10	Unnamed – relates to knee extension	0-4
11	Hip flexion and foot dorsiflexion	0, 2, 3 or 4
12	Head control	0-4
13	Elbow flexion	0, 2 or 4 (score with item 14)
14	Neck flexion	0, 2 or 4 (score with item 13)
15	Head/neck extension (Landau)	0, 2 or 4
16	Spinal incurvation (Galant)	0, 2 or 4
Total score, best score on each side for each item (maximum 64 points)		

Source: Glanzman et al[62]

Table 88: HFMSE domains and scoring

HMFSE item	HMFSE activities	Scoring within HMFSE item
1	Able to sit on chair or with legs off bed with or without hand support	0, 1 or 2
2	Able to sit on floor cross legged or legs stretched in front	0, 1 or 2
3	Able to bring hands to face at eye level	0, 1 or 2
4	Able to bring hands to head	0, 1 or 2
5	Roll to side	0, 1 or 2
6-9	Roll	0, 1 or 2
10	Able to lie down from sitting	0, 1 or 2
11	Able to raise head when lying prone	0, 1 or 2
12-13	Able to prop on forearms or extend arms	0, 1 or 2
14	Able to sit up from lying	0, 1 or 2
15	Able to four-point kneel	0, 1 or 2
16	Able to crawl	0, 1 or 2
17	Lift head from supine	0, 1 or 2
18	Stand with support	0, 1 or 2
19	Stand without support	0, 1 or 2
20	Able to walk	0, 1 or 2
21-22	Able to flex hip from supine	0, 1 or 2
23-26	Able to half knee	0, 1 or 2
27	Able to go from standing to sitting	0, 1 or 2
28	Able to squat	0, 1 or 2
29	Able to jump	0, 1 or 2
30-33	Go up and down stairs	0, 1 or 2

Source: Pera et al[63]

Table 89: WHO motor milestones

Source: WHO Multicentre Growth Reference Study Group[64]

Appendix 4: Methods for implementing the ERG's exploratory analyses

(a) Infant onset model

Exploratory analysis 1

Replace the values in worksheet "Markov Nusinersen T1" cells F335:N335 and worksheet "Markov RWC T1" cells F335 to N335 with the values presented in Table 90.

Table 90: ERG analysis 1 - baseline distribution for early onset model

Health state	Baseline proportion
No milestones	0.59
Mild milestones	0.31
Moderate milestones	0.10
Sits without support	0.01
Stands with assistance	0.00
Walks with assistance	0.00
Stands/walks unaided	0.00
Loss	0.00
Dead	0.00

Exploratory analysis 2

No amendment is required for the early onset model.

Exploratory analysis 3

For the early onset model, go to worksheet "Utility T1" drop-down box in row 11, select "Clinical experts – EQ-5D-Y vignette study"

Exploratory analyses 4

Go to worksheet "Utility T1" cells I18 to I25. Replace with the values shown in Table 91.

Table 91: ERG exploratory analysis 4 - caregiver utilities for early onset model (Bastida)

Health state	Caregiver utility
No milestones	[redacted]
Mild milestones	
Moderate milestones	
Sits without support	
Stands with assistance	
Walks with assistance	
Stands/walks unaided	
Loss of later onset motor function	0.00

Exploratory analysis 5

Apply all changes from ERG exploratory analyses 1-4, as described above. Analyses 6-8 should start from this version of the model.

Exploratory analyses 6a

Go to worksheet "Utility T1" cells F18:F25. Replace with values shown in Table 92.

Table 92: ERG exploratory analysis 6a – patient utilities for early onset model (Bastida)

Health state	Patient utility
No milestones	[redacted]
Mild milestones	
Moderate milestones	
Sits without support	
Stands with assistance	
Walks with assistance	
Stands/walks unaided	
Loss of later onset motor function	0.00

Exploratory analysis 6b

Go to worksheet "Utility T1" cells F18:F25. Replace with values shown in Table 93

Table 93: ERG exploratory analysis 6b – patient utilities for early onset model (ERG's clinical advisors)

Health state	HRQoL estimate
No milestones	0.20
Mild milestones	0.25
Moderate milestones	0.35
Sits without support	0.60
Stands with assistance	0.65
Walks with assistance	0.75
Stands/walks unaided	0.85
Loss of later onset motor function	0.00

Exploratory analysis 7

Go to worksheet "Efficacy T1" cell I104. Set value equal to zero.

Exploratory analysis 8

For exploratory analysis 8a, 8b and 8c, replace the values in worksheet "Markov Nusinersen T1" cells F419:L425 with the values presented in Table 94, Table 95 and Table 96, respectively. For exploratory analysis 8d and 8e, replace the values in worksheet "Markov Nusinersen T1" cells F419:L425 and worksheet "Markov RWC T1" cells F419:L425 with the values presented in Table 97 and Table 98, respectively.

Table 94: Transition matrix for ERG exploratory analysis 8a - 5% of nusinersen-treated patients deteriorate to next worst state

From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]						
Mild milestones	[redacted]	[redacted]	[redacted]				
Moderate milestones		[redacted]	[redacted]	[redacted]			
Sits without support			[redacted]	[redacted]	[redacted]		
Stands with assistance				[redacted]	[redacted]	[redacted]	
Walks with assistance					[redacted]	[redacted]	[redacted]
Stands/walks unaided						[redacted]	[redacted]

Table 95: Transition matrix for ERG exploratory analysis 8b - 10% of nusinersen-treated patients deteriorate to next worst state

From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mild milestones	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Moderate milestones	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits without support	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Walks with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands/walks unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 96: Transition matrix for ERG exploratory analysis 8c - 20% of nusinersen-treated patients deteriorate to next worst state

From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Mild milestones	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Moderate milestones	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits without support	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Walks with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands/walks unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 97: Transition matrix for ERG exploratory analysis 8d – all patients remain in the state achieved at the end of ENDEAR follow-up

From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]						
Mild milestones		[redacted]					
Moderate milestones			[redacted]				
Sits without support				[redacted]			
Stands with assistance					[redacted]		
Walks with assistance						[redacted]	
Stands/walks unaided							[redacted]

Table 98: Transition matrix for ERG exploratory analysis 8e – all patients revert to no milestones state at the end of ENDEAR follow-up

From\To state	No milestones	Mild milestones	Moderate milestones	Sits without support	Stands with assistance	Walks with assistance	Stands/walks unaided
No milestones	[redacted]						
Mild milestones	[redacted]						
Moderate milestones	[redacted]						
Sits without support	[redacted]						
Stands with assistance	[redacted]						
Walks with assistance	[redacted]						
Stands/walks unaided	[redacted]						

(b) Later onset model

Exploratory analysis 1

Replace the values in worksheet "Markov Nusinersen T1" cells F335:N335 and worksheet "Markov RWC T1" cells F335 to N335 with the values presented in Table 99.

Table 99: ERG analysis 1 - baseline distribution for later onset model

Health state	Baseline proportion
Sits without support but does not roll	0.56
No improvement	0.00
Sits and rolls independently	0.18
Sits and crawls with hands and knees	0.12
Stands/Walks with assistance	0.08
Stands unaided	0.06
Walks unaided	0.00
Loss	0.00
Dead	0.00

Exploratory analysis 2

Go to worksheet "Cost T2" drop-down box in row 170. Select "Apply".

Exploratory analysis 3

Go to worksheet "Utility T2" cells F18:F25. Replace values with those presented in Table 100.

Table 100: ERG exploratory analysis 3 – patient utilities for later onset model (vignette)

Health state	Patient utility
Sits without support but does not roll	0.04
No improvement	0.00
Sits and rolls independently	0.04
Sits and crawls with hands and knees	0.10
Stands/Walks with assistance	0.39
Stands unaided	0.72
Walks unaided	0.72
Loss	0.00

Exploratory analyses 4

Go to worksheet "Utility T2" cells I18: I25. Replace values with those presented in Table 101.

Table 101: ERG exploratory analysis 4 - caregiver utilities for later onset model (Bastida)

Health state	Caregiver utility
Sits without support but does not roll	[redacted]
No improvement	
Sits and rolls independently	
Sits and crawls with hands and knees	
Stands/Walks with assistance	
Stands unaided	
Walks unaided	
Loss	

Exploratory analysis 5

Apply all changes from ERG exploratory analyses 1-4, as described above. Analyses 6-8 should start from this version of the model.

Exploratory analyses 6a

Go to worksheet "Utility T2" cells F18:F25. Replace values with those presented in Table 102.

Table 102: ERG exploratory analysis 6a – patient utilities for later onset model (Bastida)

Health state	Patient utility
Sits without support but does not roll	[redacted]
No improvement	
Sits and rolls independently	
Sits and crawls with hands and knees	
Stands/Walks with assistance	
Stands unaided	
Walks unaided	
Loss	

Exploratory analyses 6a

Go to worksheet "Utility T2" cells F18:F25. Replace values with those presented in Table 103.

Table 103: ERG exploratory analysis 6b – patient utilities for later onset model (ERG's clinical advisors)

Health state	HRQoL estimate
Sits without support but does not roll	0.60
No improvement	0.00
Sits and rolls independently	0.60
Sits and crawls with hands and knees	0.60
Stands/Walks with assistance	0.75
Stands unaided	0.85
Walks unaided	0.85
Loss	0.00

Exploratory analysis 7

Go to worksheet "Efficacy T2" cell I185. Set value equal to zero.

Exploratory analysis 8

For exploratory analysis 8a, 8b and 8c, replace the values in worksheet "Markov Nusinersen T2" cells F416:M423 with the values presented Table 104, Table 105 and Table 106, respectively. For exploratory analysis 8d and 8e, replace the values in worksheet "Markov Nusinersen T2" cells F416:M423 and worksheet "Markov RWC T2" cells F416:M423 with the values presented in Table 107 and Table 108, respectively.

Table 104: Transition matrix for ERG exploratory analysis 8a - 5% of nusinersen-treated patients deteriorate to next worst state

From\To state	Sits without support but does not roll	No improvement	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No improvement	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits and rolls independently	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits and crawls with hands and knees	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands/walks with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Walks unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 105: Transition matrix for ERG exploratory analysis 8b - 10% of nusinersen-treated patients deteriorate to next worst state

From\To state	Sits without support but does not roll	No improvement	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No improvement	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits and rolls independently	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits and crawls with hands and knees	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands/walks with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Walks unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 106: Transition matrix for ERG exploratory analysis 8c - 20% of nusinersen-treated patients deteriorate to next worst state

From\To state	Sits without support but does not roll	No improvement	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
No improvement	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits and rolls independently	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Sits and crawls with hands and knees	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands/walks with assistance	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Stands unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Walks unaided	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

Table 107: Transition matrix for ERG exploratory analysis 8d – all patients remain in the state achieved at the end of CHERISH follow-up

From\To state	Sits without support but does not roll	No improvement	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll	1.000						
No improvement							
Sits and rolls independently			1.000				
Sits and crawls with hands and knees				1.000			
Stands/walks with assistance					1.000		
Stands unaided						1.000	
Walks unaided							1.000

Table 108: Transition matrix for ERG exploratory analysis 8e – all patients revert to no milestones state at the end of CHERISH follow-up

From\To state	Sits without support but does not roll	No improvement	Sits and rolls independently	Sits and crawls with hands and knees	Stands/walks with assistance	Stands unaided	Walks unaided
Sits without support but does not roll	1.000						
No improvement							
Sits and rolls independently	1.000						
Sits and crawls with hands and knees	1.000						
Stands/walks with assistance	1.000						
Stands unaided	1.000						
Walks unaided	1.000						