Tafamidis for Transthyretin Familial Polyneuropathy (TTR-FAP)

Evidence Review Group assessment of manufacturer submission

Produced by

CRD/CHE Technology Assessment Group (Centre for Reviews and Dissemination/Centre for Health Economics), University of York

Authors

Rita Faria, Research Fellow; Simon Walker, Research Fellow; Stephen Palmer, Professor

Centre for Health Economics

Mark Corbett, Research Fellow; Lisa Stirk, Information Specialist; Catriona McDaid, Senior Research Fellow.

Centre for Reviews & Dissemination

Acknowledgements

With thanks to Professor Philip Hawkins and Professor Mary Reilly who provided clinical advice to the ERG. With thanks to David Epstein for providing advice on the health economics.

Date completed **16th July 2012**

Declared competing interests of the authors

None

Table of Contents

1 SUMMARY 11
1.1 Scope of the Evidence Review Group assessment 11
1.2 Scope of the manufacturer submission 11
1.3 Summary of submitted clinical effectiveness evidence 11
1.4 Summary of submitted cost effectiveness evidence 12
1.5 Summary of additional work undertaken by the ERG 14
1.6 Commentary on the robustness of submitted evidence 15
2 BACKGROUND 20
2.1 Introduction 20
2.2 Description of health problem 20
2.2.1 Epidemiology 20
2.2.2 Clinical variability in TTR-FAP 21
2.2.3 V30M mutation 21
2.2.4 Stages of disease progression in V30M 21
2.2.5 Heterogeneity within V30M populations 22
2.2.6 Non-V30M mutations 23
2.2.7 Applicability of Coutinho classification of disease stages to NonV30M and non-endemic V30M populations 23
2.3 Current treatment 24
2.3.1 Treatment of symptoms 24
2.3.2 Liver transplantation 24
2.4 Tafamidis and treatments in development 26
3 THE DECISION PROBLEM 27
3.1 Population 27
3.2 Intervention 27
3.3 Comparators 27
3.4 Outcomes 28
3.5 Time frame 28
3.6 Treatment strategies 28
4 CLINICAL EFFECTIVENESS 29
4.1 Overview of the manufacturer's submission 29
4.2 ERG systematic review 30
4.3 Overview of included studies 31
4.4 Randomised Controlled Trial (Fx-005) 33

4.4.1 Objectives and methods 33
4.4.2 Analyses 33
4.4.3 Results 36
4.4.4 Assessment of study quality 43
4.4.5 FDA analyses of the RCT data 48
4.5 Open-label extension study (Fx-006) 49
4.5.1 Study details and main results 49
4.6 Non-randomised studies 52
4.6.1 Pre-post study (Fx1A-201) 53
4.6.2 Case series 55
4.7 Ongoing studies 56
4.8 Generalisability of the study results 56
4.9 Summary 57
5 ECONOMIC EVALUATION 59
5.1. Overview of manufacturer's economic evaluation 59
5.1.1. Review of existing literature 64
5.1.2. Reports of other Health Technology Assessment agencies on tafamidis 64
5.1.3. Manufacturer's 'de novo' economic evaluation 64
5.2. Patient population 68
5.3. Natural history 70
5.3.1 Disease severity and TQoL scores 70
5.3.2 Mortality 77
5.3.3 Liver transplantation 81
5.4 Treatment effectiveness within the submission 84
5.5 Health related quality of life 88
5.5.1 HRQoL associated with TTR-FAP 88
5.5.2 HRQoL associated with liver transplantation 92
5.5.3 HRQoL experienced by carers 93
5.5.4 HRQoL: Final issues 93
5.6 Resources and costs 94
5.6.1 Acquisition costs of tafamidis 94
5.6.2 Costs associated with healthcare resource use per disease stage 95
5.6.3 Costs associated with healthcare resource use from liver transplantation 96
5.6.4 Productivity costs accrued by patients and carers 96
5.7 Sensitivity analyses 97
5.8 Model validation 97

5.9 Results included in manufacturer's submission 98
5.9.1 Base-case 98
5.9.2 Sensitivity Analysis 99
5.9.3 Budget Impact considerations 100
5.10 Summary of uncertainties and issues 103
6 ADDITIONAL WORK UNDERTAKEN BY THE ERG 104
6.1 Overview 104
6.2 The ERG's base-case 105
6.2.1 Base-case results 108
6.3 Exploratory scenario analysis 111
6.4 Analysis of relative weighting of QALY benefits 121
6.5 Sensitivity analysis on budget impact 123
References 125

List of Appendices

Appendix 1: Systematic review methods 129
Appendix 2: Study Fx1A-OS-001 142
Appendix 3: Details of the Co-primary outcome assessment tools 145
Appendix 4: Points of clarification from manufacturer 150
Appendix 5: Quality Assessment using an economic modelling checklist [30] 175
Appendix 6 – Validation of ERG's model 176
Appendix 7 – Calculations for exploratory scenario analyses 177
Appendix 8 – Cost-effectiveness results for exploratory scenario analyses for different baseline TQoL scores 179
Appendix 9 – Calculations for sensitivity analysis to budget impact 187

List of Tables

Table 1: Characteristics of TTR-FAP patients who have had a liver transplant* 25
Table 2: Survival following transplant in patients with V30M and nonV30m TTR-FAP 26
Table 3: Studies of tafamidis included in the ERG's systematic review 32
Table 4: Patient demographic and baseline characteristics of the studies submitted by the manufacturer (Data were obtained from both the company submission (Tables 10, 17, and 28) and from the full study reports) 37

Table 5: Results of Co-primary endpoints in RCT Fx-005 39
Table 6: Results for secondary endpoints in the randomised controlled trial (ITT population) 42
Table 7: Risk of Bias assessment results for the RCT (Fx-005) 44
Table 8: ERG estimates of the rates of disease progression prior to randomisation 46
Table 9: Patient baseline characteristics in the trial open-label extension study Fx-006..... 50
Table 10: Results for the 'sustainability of the treatment effect' of tafamidis in study Fx-006, ITT population (source: full study report)..... 51
Table 11: Results for the 'superiority of the treatment effect' of tafamidis in study Fx-006, ITT population (source: full study report)..... 51
Table 12: Efficacy results for the pre-post study 54
Table 13: Checklist for the quality assessment for the pre-post study (Fx1A-201) 54
Table 14: Summary of the economic evaluation submitted by the manufacturer..... 60
Table 15: Age at model entry and baseline TQoL score (adapted from Table 39 p136 of MS) 68
Table 16: Conversion of the mPDS to the Coutinho et al (1980) disease stages (adapted from Table 3 of Points for Clarification and Coutinho et al (1980) [7] 71
Table 17: TQoL scores per stage recorded in the THAOS registry (adapted from Tables 41 p139 of MS and Appendix 12-15 of Points for Clarification) 72
Table 18: Threshold TQoL scores for cross-over between stages (adapted from Table 42 p139 of MS)..... 73
Table 19: Time spent in each disease stage according to the manufacturer's submission (adapted from Tables 43, p140 of MS) 75
Table 20: TQoL rate of change over 6-months applied in the economic model (adapted from Table 44, p140 of MS) 75
Table 21: Parameters for fitted Weibull without transplant survival 79
Table 22: Variables associated with reduced survival in a cohort of T60A patients [3]..... 80
Table 23: Liver transplant rates over 6-months used in the economic model (adapted from Table 47, p141) 81
Table 24: Parameters for fitted Weibull post-transplant survival..... 82
Table 25: Comparison of monthly rate of change observed in Fx-005 used in the economic model and those estimated from the results of the pre-specified primary end-points, and corresponding hazard ratio associated with tafamidis (Adapted from Tables 11 and 12, p48 of MS and adapted from Table 14, p108 of Clinical Study Report for Fx-005)..... 86
Table 26: HRQoL scores used in model 88
Table 27: Mean TQoL per stage and thresholds for moving disease stage and corresponding EQ-5D score using the manufacturer's mapping function 90
Table 28: EQ-5D utility values for mean TQoL per stage and stage threshold using the alternatives presented in the response to the points for clarification 91
Table 29: Costs by disease stage (adapted from Table 50, p148 of MS) 94

Table 30: Breakdown of healthcare costs per disease stage (adapted from Table 55,56,58, p154-168 of MS) 95
Table 31: Breakdown of productivity costs per disease stage (adapted from Table 55,56,58, p154-168 of MS) 96
Table 32: Cost-effectiveness results for the base-case population (combined V30M and non-V30M) and for the two subgroup populations, V30M and non-V30M patients (adapted from Tables 51-53 of MS) 99
Table 33: Sensitivity analysis to base-case population (combined V30M and non-V30M) ... 100
Table 34: Prevalence and incidence of ATTR-FAP in England 101
Table 35: Budget impact estimates (adapted from Table 33-5 p100-1 of MS) 102
Table 36: Comparison of model assumptions and key input parameters between the manufacturer's and the ERG's base-case 107
Table 37: Base-case results for tafamidis add-on therapy compared with conventional standard care (CST) alone and comparison with manufacturer's results 109
Table 38: Base-case results for different baseline TQoL scores 110
Table 39: Scenarios considered in the exploratory analysis, rationale and description 112
Table 40: Cost-effectiveness results for scenario analysis 118
Table 41: Analysis on the relative weighting of additional QALY benefits 121
Table 42: Scenarios for budget impact sensitivity analysis 123
Table 43: Sensitivity analysis to budget impact 124
Table 44: Summary of content of Norfolk QoL-DN Questionnaire 145

List of Figures

Figure 1: Structure of Markov model (adapted from Figure 25 p132 of MS) 66
Figure 2: Relationship between TQoL score and disease duration 76
Figure 3: Survival curves used in the economic model for the base-case population and patient subgroups for patient without liver transplantation (using parameters presented in Table 21) 80
Figure 4: Effect of tafamidis in TQoL score 85
Figure 5: Scatter plot of baseline EQ-5D scores by baseline TQoL scores (reproduced from Figure 29, p145 of MS) 89
Figure 6: Study flow chart 130

List of abbreviations

AGNSS	Advisory Group for National Specialised Services
AIC	Akaike information criterion
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
AWMSG	All-Wales Medicines Strategy Group
BIC	Bayesian information criterion
CEA	Cost-Effectiveness Analysis
CI	Confidence interval
CTR	Clinical trial report
EMA	European Medicines Agency
EQ-5D	EuroQol (5-Dimensions)
ERG	Evidence Review Group
FAPWTR	Familial Amyloidotic Polyneuropathy World Transplant Registry
FDA	Food and Drug Administration
HRQoL	Health-related Quality of Life
HTA	Health Technology Assessment
ICER	Incremental cost-effectiveness ratio
Infarmed	National Authority of Medicines and Health Products (Portugal)
IQWIG	Institute for Quality and Efficiency in Health Care (Germany)
ITT	Intention-to-treat
LOCF	Last observation carried forward
mBMI	Modified body mass index
mPDS	Modified polyneuropathy disability score
MD	Mean difference
MS	Manufacturer's submission
NAC	National Amyloidosis Centre
NHS	National Health Service
NHS-EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Clinical Excellence
NIS	Neuropathy Impairment Score
NIS-LL	Neuropathy Impairment Score – Lower Limb
NIS-UL	Neuropathy Impairment Score – Upper Limb
Norfolk QoL-DN	Norfolk Quality of Life – Diabetic Neuropathy
PSS	Personal Social Services
QALY	Quality-Adjusted Life Year
RCT	Randomised controlled trial
SMS	Scottish Medicines Consortium
SD	Standard deviation
SE	Standard error
SMC	Scottish Medicine Consortium
SPC	Summary of Product Characteristics

TEAE	Treatment-emergent adverse event
THAOS	Transthyretin Amyloidosis Outcomes Survey
TQoL	Total Quality of Life (See Norfolk QoL-DN)
TTR	Transthyretin
TTR-FAP	Transthyretin familial amyloid polyneuropathy
UK	United Kingdom

Glossary

Amyloids	Proteins with misfolded structures which interact with one another to form insoluble fibrils. Abnormal accumulation of amyloid fibrils in organs (amyloidosis) can play a role in various neurodegenerative disorders.
Autonomic neuropathy	Disease affecting autonomic nerves which control involuntary body functions, such as heart rate, blood pressure, perspiration and digestion.
Cachexia	Weight loss and deterioration in physical condition.
Cardiac amyloidosis	Disease caused by amyloid deposits in heart tissue, preventing the heart from functioning properly.
Efficacy-evaluable (EE) analysis	An analysis in which only patients who complete the clinical trial are included in the final results. Also called per-protocol analysis.
Hypertrophic cardiomyopathy	Changes in which the muscular heart wall thickens abnormally, preventing the heart muscle from relaxing normally during the filling phase, which can lead the flow of blood out of the heart to be blocked.
Intention-to-treat (ITT) analysis	An analysis in which all participants in a trial are analysed according to the intervention to which they were allocated, whether they received it or not.
Karnofsky performance status score	A measure of functional impairment. Scores range from 0 to 100. The lower the score, the worse the functional impairment.
Open-label trial	A type of clinical trial in which both the researchers and participants know which treatment is being administered.
Paraesthesia	A sensation of tingling, burning, pricking, or numbness of a person's skin.
Peripheral neuropathy	Disease of the peripheral nerves in which motor, sensory, or vasomotor nerve fibres may be affected and which is marked by muscle weakness, pain, and numbness.
Polyneuropathy	Neurological disorder occurring when multiple nerves throughout the body malfunction simultaneously.
Sensorimotor polyneuropathy	Neurological disorder affecting both motor and sensory types of nerves throughout the body.
Transthyretin (TTR)	A blood protein that binds thyroxine and retinol which is produced in the liver. Transthyretin is also called prealbumin.

1 SUMMARY

1.1 Scope of the Evidence Review Group assessment

This report presents the Evidence Review Group's (ERG) assessment of the manufacturer's submission (MS) to the Advisory Group for National Specialised Services (AGNSS) on the use of tafamidis (Vyndaqel) for the treatment of transthyretin familial amyloid polyneuropathy (TTR-FAP). The report includes an assessment of both the clinical evidence and economic analysis submitted by the company, supplemented with additional analyses undertaken by the ERG.

1.2 Scope of the manufacturer submission

The manufacturer evaluated the effectiveness and cost-effectiveness of the licensed dose of 20mg orally once daily of tafamidis. The population of interest in the assessment of clinical and cost-effectiveness was patients with confirmed stage one TTR-FAP. The key subgroups of interest were patients with the V30M variant and those with variants other than V30M, henceforth referred to as non-V30M. The manufacturer's definition of the decision problem was in line with the AGNSS scope.

1.3 Summary of submitted clinical effectiveness evidence

The evaluation of the clinical effectiveness of tafamidis for transthyretin familial amyloid polyneuropathy (TTR-FAP) was primarily based on a single randomised controlled trial (RCT) (Fx-005) of 18 months' duration. A 20mg daily dose taken orally was compared to placebo. This multi-centre trial included 128 patients from several countries with the V30M mutation. Over half the patients were from Portugal which is an endemic area for the V30M mutation. A one year open-label extension (Fx-006) study of the RCT was also undertaken during which all participants received tafamidis. Supplementary evidence was provided from a small single-arm pre-post study in 21 patients with a non-V30M population. Additionally, the ERG identified two small case series published recently as conference abstracts. No studies were identified that compared tafamidis to liver transplantation.

Participants in the RCT had fairly early stage disease (based on median duration of symptoms and baseline scores on the peripheral neuropathy assessment scale used), with the median age being around 35 years. The co-primary outcomes were the Norfolk Quality of Life-Diabetic Neuropathy Scale (TQoL) and the Neuropathy Impairment Score-Lower Limb (NIS-LL).

In the primary analysis of the intention-to-treat (ITT) population at 18 months, for TQoL the tafamidis group scores deteriorated by a mean of 2.0 points versus a placebo deterioration of 7.2 points. The difference was -5.2 points in favour of tafamidis with a 95% confidence interval (CI) from -11.8 to 1.3 points; this was not statistically significant (p=0.12). For NIS-LL the proportion of responders in the tafamidis group was 45%, versus 30% for placebo (difference =15%, 95% CI -1.15 to 32.0); which was also not statistically significant (p=0.07). There was evidence of benefit with tafamidis in secondary analyses of the co-primary endpoints (NIS-LL and TQoL), and for some of the secondary outcomes; statistically significant differences were found for both co-primary endpoints in the efficacy evaluable population, and for mBMI at all time points in the ITT population. The treatment was generally well-tolerated; although tafamidis was associated with more urinary tract infections than placebo, placebo was associated more with headache and neuralgia than tafamidis.

Seventy-one of the original 128 patients from the trial participated in the extension study to the RCT. Tafamidis patients in the trial received tafamidis for a further 12 months and the placebo group from the trial also started tafamidis and received it for 12 months. For several outcome measures the mean rate of change per month was broadly similar following an additional 12 months of tafamidis treatment compared to the rate of change with tafamidis treatment in the main trial. The investigators interpreted this as suggesting sustainability of the treatment effect of tafamidis over 30 months. In a similar analysis the rate of change for several outcomes was lower following 12 months of tafamidis treatment when compared with the previous 18 months of placebo treatment.

A small pre-post study was conducted in a population with non-V30M mutations, which was more similar to an English population that the RCT population. When pre-treatment and post-treatment periods were compared, a statistically significant slowing in disease progression was seen for NIS and TQoL, but the difference was not significant for large nerve fibre function.

1.4 Summary of submitted cost effectiveness evidence

The manufacturer conducted a systematic literature search to identify any published economic evaluations in TTR-FAP. No published economic evaluations were identified in TTR-FAP. In the absence of any previously published economic evaluation of tafamidis, the manufacturer's *de novo* economic evaluation formed the basis of the economic evidence submitted to AGNSS.

The *de novo* economic evaluation compared the lifetime costs and health outcomes of tafamidis as an add-on therapy to conventional support therapy with conventional support therapy alone in patients with TTR-FAP disease stage 1 using a decision analytic model. Tafamidis was discontinued once patients progress to disease stage 2. Patients in disease stage 1 in both arms of the model were eligible for liver transplantation, irrespective of genetic variant. The model evaluated costs from the perspective of the NHS and Personal Social Services (PSS) but also included productivity costs incurred by patients and carers. Costs were expressed in UK pound sterling at a 2010 price base. Outcomes in the model were expressed in terms of quality-adjusted life years (QALYs). Both costs and outcomes were discounted at 3.5% per annum. The base-case population consisted of a combined population including both V30M and non-V30M patients, and separate populations of patients with the V30M variant and patients with variants other than V30M (non-V30M) were also considered, which is in line with the scope defined by AGNSS.

The natural history of patients with TTR-FAP was modelled through the TQoL score for disease severity, whilst mortality was modelled separately. Greater TQoL scores indicate lower health related quality of life (HRQoL) and more severe disease. In order to use TQoL scores to model disease severity, a number of assumptions were made by the manufacturer: (i) the Coutinho disease stages, a classification system developed for an endemic V30M population, are a suitable classification of disease status for both V30M and non-V30M populations in England; (ii) TQoL captures disease progression for both V30M and non-V30M populations; (iii) the Coutinho disease stages can be defined using TQoL scores, with TQoL cut-offs between stages determined by TQoL scores observed in the THAOS registry and the use of arbitrary rules with no supporting evidence provided; (iv) the TQoL rate of change, and therefore disease progression over time, is the same for V30M and non-V30M patients; (vi) TQoL rate of change is dependent on disease stage; and (v) the TQoL rate of change over time is based on the relationship between TQoL and disease duration observed in a cross-sectional observational study conducted outside the UK and the use of cut-off TQoL scores for disease stages.

Other key assumptions made by the manufacturer included: (i) mortality is solely dependent on time in the model and is independent of disease severity; (ii) patients who undergo liver transplantation experience no further disease progression (i.e. it maintains their HRQoL at the time of transplant throughout the rest of their lifetime) and have improved survival; (iii) all patients in disease stage 1 are eligible for liver transplantation; and (iv) mortality post-liver transplantation depends on time from liver transplantation and on the patient's age.

Results were presented for the base-case (combined V30M and non-V30M) population and for the two subgroup populations, V30M and non-V30M patients. One-way sensitivity analyses were conducted for the base-case population. The incremental cost-effectiveness ratio (ICER) for the base-case population is £189,995 per QALY gained. The ICER for the V30M population is £174,634 and for the non-V30M population is £304,293 per QALY gained. The results of the sensitivity analyses indicate that the rate of liver transplantation is a key driver of cost-effectiveness: reducing the rate of liver transplant to zero resulted in the ICER more than doubling to £602,850 per QALY gained.

A separate budget impact analysis was also presented to assess the annual acquisition costs (and over a 5-year period) to the NHS in England of using tafamidis for a prevalent population of 17 patients. The annual cost of tafamidis starts at [redacted] in year 1 and increases progressively with the increase in patients treated to [redacted] in year 5. The cumulative budget impact in years 1 to 5 is [redacted].

1.5 Summary of additional work undertaken by the ERG

In order to reduce computation times and facilitate further analyses, the ERG reconstructed the manufacturer's decision analytic model using the same structure, data and assumptions but as a cohort model rather than as an individual patient simulation model. The ERG did, however, present an alternative base-case, informed by the critical appraisal of the manufacturer's submission and response to the points for clarification, together with expert clinical advice. In addition, the ERG undertook: (i) exploratory scenario analyses examining alternative assumptions for stopping rules, disease stage cut-offs, rules for liver transplantation, costs and HRQoL; (ii) an analysis on the relative weighting of QALY benefits to assess how much more the QALYs gained from tafamidis would need to be valued compared to QALYs for other treatments in order for tafamidis to be considered a cost-effective use of resources under various cost-effectiveness thresholds; and (iii) a set of sensitivity analyses on the budget impact.

The ERG's base-case compares tafamidis as an add-on therapy to conventional standard care alone from the perspective of the UK NHS and PSS. Results are presented separately for the V30M population and for the non-V30M population. Results for the combined V30M and non-V30M population correspond to the weighted average of the estimates for the V30M and non-V30M populations, weighted by the relative proportion of the prevalence of each variant in England. The key differences with the manufacturer's base-case are: (i) the perspective for both costs and outcomes is that of the NHS and PSS, hence productivity costs are not

included in the base-case; (ii) patients remain on tafamidis throughout stages 1 and 2 rather than discontinuing at progression to stage 2 given issues with use of the Coutinho stage classifications in the patient population in England; and (iii) the rate of liver transplantation is assumed to be zero given expert clinical advice that it is unlikely to be a therapeutic option for TTR-FAP patients in England. In addition, results are presented for different baseline TQoL scores in order to explore the impact of disease severity on the cost-effectiveness of tafamidis.

Based on the ERG's analyses the ICER for the V30M population is £1,074,450 per QALY gained, while the ICER for the non-V30M population is £1,138,813 per QALY gained. The ICER for the combined population is £1,126,565 per QALY gained. For both populations, the ICER increases as baseline TQoL increases. These results suggest that the cost-effectiveness of tafamidis may be more favourable for patients who are identified earlier, or those with less severe neuropathic impairment. The scenarios with the greatest impact on the expected costs and health benefits are those testing the assumptions regarding stopping rules, disease staging and TQoL rate of change and liver transplant. The threshold analysis on the relative weighting of additional QALY benefits indicated that, under the ERG's base-case, the QALY benefits associated with treatment would need to be valued at around thirty-five times the QALY benefits obtained in other diseases treated in the NHS which have been considered by the National Institute for Health and Clinical Excellence (NICE) for tafamidis to be considered cost-effective. The sensitivity analysis on the budget impact suggests that the cumulative cost of tafamidis to the NHS could reach up to [redacted] over 5 years.

1.6 Commentary on the robustness of submitted evidence

Strengths

The manufacturer performed a thorough search of the literature to identify all relevant studies of tafamidis and their systematic review process methods appeared robust and were well-documented. The main randomised trial on which the submission was based used appropriate methods to minimise the risk of bias affecting the study results. Safety data were made available for all studies.

The ERG acknowledges the difficulties of undertaking a valid evaluation of the efficacy and safety of tafamidis given the limited evidence available, particularly for the non-V30M population. The manufacturer's economic evaluation provides the main source of evidence on the cost-effectiveness of tafamidis for TTR-FAP. In general, the analysis follows the guidelines for economic evaluations to inform decisions in the NHS. In addition, all the potentially relevant costs and health outcomes have been included.

Weaknesses

There was no randomised trial evidence of tafamidis in patients with a non-V30M mutation. This was a significant limitation of the evidence since patients from England are much more likely to have a non-V30M mutation. Disease resulting from non-V30M mutations presents with varying symptoms and co-morbidities, such as early cardiac involvement, often not seen in V30M patients. Patients with non-V30M mutations can also have different rates of progression: they generally have a later disease onset and faster rates of overall disease progression, when compared to a V30M population. Although a small pre-post study was performed in a non-V30M population, the limitations associated with of this type of study design, and the size of the study, means that any observed benefit cannot be attributed to tafamidis with confidence. The submission did not discuss these limitations or consider their implications.

A key area of concern in relation to the main trial was the potentially clinically important differences between the tafamidis and placebo groups at baseline. The tafamidis group had lower baseline NIS-LL and TQoL scores than the placebo group, suggesting less severe disease. Importantly, baseline NIS-LL was found to be a significant covariate predictor (p=0.0112) of a co-primary outcome (NIS-LL response at 18 months). Furthermore, the tafamidis group also had a longer disease duration at baseline than the placebo group. The difference in NIS-LL scores at baseline, coupled with the difference in duration of symptoms, suggests the placebo group may have had an underlying faster rate of disease progression. Little consideration was given in the submission to possible impact on the results of the baseline imbalance between the treatment groups, based on the fact that the differences were not statistically significant. It is nevertheless possible that some of the effect estimates presented in the submission may be due to a combination of an underlying disease rate difference, and the effect of tafamidis treatment, rather than due to the effect of tafamidis treatment alone.

The manufacturer proposed that the similar reduced rates of progression with tafamidis in non-V30M patients (in the pre-post study) and V30M patients (in the trial) provided a rationale for extrapolating results of the trial to a non-V30M population. In addition to the limitations of the pre-post study, further analyses of the RCT data by an FDA statistician, indicated there was uncertainty about the relevance of the trial results to the whole V30M population, with variation in response appearing to be related to patients' origin of mutation (their endemic or non-endemic status). The plausibility of extrapolating results to a non-V30M population therefore appears questionable.

The ERG identified a number of weaknesses in the economic model. These include:

- The consideration of a combined population of V30M and non-V30M patients in the base case is inappropriate. Not only is this an identifiable source of heterogeneity but many of the parameters used to reflect the combined population (e.g. survival curves), are based on different proportions of V30M and non-V30M patients. Therefore, the ERG does not consider the results presented in the manufacturer's submission for the base case a reliable estimate. If it is considered appropriate to consider the results for the combined population, then the results should be calculated as the weighted average of the results for each population, weighted by their relative proportions in England.
- There are several issues with the use of TQoL to model natural history, namely it is uncertain: (i) whether TQoL captures disease progression and severity in both V30M and non-V30M populations; (ii) whether the Coutinho stages can be defined by TQoL scores; (iii) whether the TQoL cut-offs for disease stages used in the model are appropriate; (iv) whether TQoL rate of change estimated from a cross-sectional study appropriately captures disease progression; and (v) whether TQoL rate of change estimated in the aforementioned study is applicable to the progression of the disease in non-V30M patients.
- The absolute mean difference in TQOL between the treatment and control groups used in the model is considerably greater than observed in the FX-005 RCT. The model may therefore overestimate the benefits of the treatment.
- It is assumed that absolute gains in TQOL while on treatment are maintained after progression of disease and/or discontinuation. There is no evidence presented to confirm or refute this assertion.
- It remains unclear whether the Coutinho stages are applicable to non-V30M patients, given that disease progression generally involves cardiac amyloidosis and may not follow the pattern of progressive polyneuropathy to which the Coutinho stages refer to and that is typical of V30M cases.
- Mortality without liver transplant was modelled independently of disease severity and solely dependent on time from symptom onset. The manufacturer did not claim any benefit of the treatment on survival time. Given that survival time is likely to be correlated with disease severity, this assumption may understate the benefits of treatment.
- Clinical advice suggested liver transplantation is rarely a treatment option in the England. Further, the evidence used to estimate the benefits of liver transplantation does not appear to the ERG to be comparable with the population in England. One of the main benefits of treatment with tafamidis in the decision analytic model is that the

treatment delays the onset of stage 2 disease and increases the possibility of liver transplant for the patient

- Productivity costs should not be included when the perspective taken is that of the NHS & PSS. In addition, it is unclear if the estimates of productivity costs are appropriate to a patient population which enters the model at an age close to the UK retirement age.

Areas of uncertainty

The main areas of uncertainty with regard to the clinical effectiveness of tafamidis are:

While the 18 month RCT was of reasonable duration, and there is evidence on safety for a 30 month period for a small group of patients, the long-term safety and efficacy are still unknown. In clinical practice it is likely that patients would take tafamidis for longer than 30 months. Also, any durability of the effect of tafamidis in those patients who stop treatment is unknown.

Aside from the lack of a statistically significant difference between tafamidis and placebo in the primary analysis for NIL-LL responders and TQoL in the main RCT, it is unclear whether the magnitude of the effect for either outcome is clinically meaningful. Most of the evidence in the trial regarding the effectiveness of tafamidis relies on secondary outcomes and secondary analyses. In addition, some of the outcomes in these analyses may have been affected by baseline imbalances leading to an overestimation of the treatment effect. Therefore there is some uncertainty regarding the effect of tafamidis in a V30M population based on the single trial available.

Patients with non-V30M mutations can have different rates of progression, generally having later disease onset and faster rates of overall disease progression, when compared to a V30M population. Therefore, there is uncertainty as to whether the results from the trial are applicable to a non-V30M population.

In relation to the economic model, several key areas of uncertainty remain. It is unclear on the generalisability of the evidence on natural history of the disease from a V30M population to the population in England of predominantly non-V30M patients. The manufacturer based the model on TQoL, a measure of quality of life related to diabetic neuropathy, which has not been validated in the TTR-FAP patient population. There is considerable uncertainty as to whether TQoL is an appropriate marker of disease severity and progression in this disease, particularly in non-V30M patients, where peripheral neuropathy may not be the predominant symptom. In addition, given the heterogeneity observed in TQoL scores within the disease stages in V30M

patients, it is unclear whether TQoL scores can be used to define disease stages. It remains unclear whether the Coutinho stages are applicable to the English V30M and non-V30M population. If disease stages 1 and 2 are indistinguishable in the English population, the stopping rule of discontinuing tafamidis at progression to stage 2 may not be feasible in clinical practice. There is also uncertainty regarding the eligibility of these patients for liver transplantation and the rate of liver transplantation in this patient population. The clinicians contacted by the ERG considered that the most plausible rate is likely to be close to zero, whereas the manufacturer assumed that every 6 months approximately 1 in 20 patients in stage 1 receive a liver transplant. Finally, there is uncertainty as to whether the results from the trial are applicable to a non-V30M population and, therefore, whether tafamidis is an effective treatment for the patient population in England.

2 BACKGROUND

2.1 Introduction

The manufacturer's submission provides a fairly detailed account of the aetiology and symptoms associated with transthyretin familial amyloid polyneuropathy (TTR-FAP) as well as the physical and emotional impact on patients and their families. The intention is not to replicate that in this ERG assessment of the manufacturers' submission (MS). Our description of the health problem and current treatments focuses on those issues that we believe are most pertinent to assessing the validity and generalisability of the effectiveness and cost-effectiveness evidence presented in the submission.

2.2 Description of health problem

TTR-FAP is one of three types of familial amyloid polyneuropathy (FAP). It is caused by an inherited mutation which affects the structure of a protein called transthyretin (TTR), which is mainly made in the liver.[1] This causes the transthyretin protein to dissociate more easily which may result in deposits of amyloid (abnormal protein) in nerves and the heart, leading to nerve damage and heart failure. Symptoms include sensory motor neuropathy such as limb weakness and loss of sensation; autonomic dysfunction affecting the bladder, bowel and sexual function; the heart is frequently affected and other organs including the kidneys can also be involved.[2] Symptoms develop at any stage in adulthood and early-onset (third to fourth decade) and late-onset (sixth to eighth decade) presentations have been identified.[2-3] It is a progressive disease and on average people die within 10 years of developing symptoms.[2]

2.2.1 Epidemiology

TTR-FAP is a rare condition with an estimated prevalence of less than 0.1 people in 10,000, which is equivalent to less than 5,000 people in the European Union,[4-5] and 544 people in England and Wales.[6] This may not be an accurate estimate due to inherent limitations in prevalence studies of rare conditions;[5] it is most likely to be an overestimate as published prevalence studies for rare diseases tend to be undertaken in regions with higher prevalence. The manufacturer's submission presents prevalence data specific to England based on data from the National Amyloidosis Centre (NAC), which suggests a lower prevalence than the overall European prevalence rate and this data is discussed in more detail in the section on the economic model (section 5.3.3).

2.2.2 Clinical variability in TTR-FAP

Over 100 different variations (mutations) of TTR have been described and are associated with highly variable symptoms, course and prognosis.[2] Worldwide, the most common variant is the V30M mutation which results in a substitution of amino acids in the composition of transthyretin (methionine for valine at position 30). Patients have been identified all over the world, though there are several notable pockets (endemic areas) of the V30M mutation in Portugal, Sweden and Japan. Typical clinical presentation in patients with the V30M mutation is a progressive small fibre neuropathy affecting peripheral and autonomic nerves. Cardiac amyloidosis is relatively uncommon with this mutation in contrast with non-V30M mutations.[3] The most common variant in the USA is Val122Ile (V122I), detected in 3-4% of the African-American population. Patients mainly present with hypertrophic cardiomyopathy with mild or no neurological symptoms.[2] 26% of the current patients visiting the NAC in England have the V122I mutation with a similar presentation of cardiomyopathy without significant neurological features. The other common mutation in the English population is Thr60Ala (T60A) constituting 26% of NAC patients; almost half do not have peripheral neuropathy symptoms (Philip Hawkins, National Amyloidosis Centre, personal communication, 22/2/2012).

2.2.3 V30M mutation

Coutinho et al. undertook a retrospective analysis of the records of 483 V30M patients who had attended their clinic in Portugal between 1939 and 1979.[7] 42.2% of patients were seen only once during the course of their disease and the remaining patients had a follow-up of 1 to 17 years. The mean age of symptom onset was 31 years for men (range 13 to 68) and 33 years for women (range 19 to 60). Lower limb paresthesia (abnormal skin sensation such as pricking or tingling), pain and impaired temperature sensation were the most common initial symptoms (50.2%), followed by vomiting, constipation, diarrhoea, or alternating constipation (40.4%). The authors described an insidious onset and slow progression in the first years, followed by rapid deterioration. Approximately one third of patients only had symptoms of autonomic dysfunction and one third had a purely sensory neuropathy. The average duration from symptom onset to death was 10.8 years (range 3 to 26 years).

2.2.4 Stages of disease progression in V30M

Based on their review of the cases, Coutinho et al. described three progressive stages of TTR-FAP.[7] These were specified based on sensory motor neuropathy symptoms which the authors found to be consistent in how they presented. This was in contrast to autonomic

neuropathies, which Coutinho et al. reported were more variable in frequency and intensity. Stage 1, with an average duration of 5.6 years (SD 2.8), was defined as the period of time in which “the disease is limited to the lower limbs and the patient is still walking without help”. In stage 2, with an average duration of 4.8 years (SD 3.6), motor signs progress in the lower limbs and the patient needs help to move around. In terms of mobility the patient is “obviously handicapped but can still move around though needing help”. There is also wasting and weakness in hand muscles and temperature and pain sensory impairment appear in the upper limbs and trunk at this stage. In stage 3, which lasts for 4.8 years (SD 3.6), patients were described as being bedridden or confined to a wheelchair due to the severity of the symptoms eventually leading to death from cachexia (general physical wasting including loss of weight and muscle mass) or secondary infections.

2.2.5 Heterogeneity within V30M populations

Patient outcome can vary depending upon the type of mutation and even in patients with a specific mutation such as V30M, the clinicopathological features of the disease vary depending on age of onset and geographical location.[8] Koike et al. undertook a retrospective analysis of the natural history of late-onset TTR-FAP in 50 patients with the V30M mutation, from non-endemic areas of Japan, who had no relationship to the endemic foci in Japan for at least two generations.[8] They found that the natural history of these late-onset patients was different to the classical presentation of early onset V30M described by Coutinho et al. as sensory neuropathy, beginning in the lower limbs, progressing to upper limb neuropathy and motor symptoms presenting at a later stage.[7] The mean age at onset of symptoms in the non-endemic population was 64.5 years (SD 6.5, range 55 to77 years) and for diagnosis was 67.3 years (SD 5.9).[8]

Although lower limb sensory symptoms tended to present first (mean 65.1 years, SD 6.5), late-onset patients experienced sensory and motor symptoms in the upper and lower limbs within 18 months of first symptom onset and some patients had upper limb symptoms prior to lower limb symptoms. The authors noted that this was different to the classical presentation described by Coutinho et al. of several years between onset of lower and upper limb symptoms.[7] Cardiac symptoms tended to occur in the latest phase of the disease. The mean age at which patients began using a walking stick was 67.7 years (SD 6.6) and a wheelchair at 69.3 years (SD 6.2). The mean duration from disease onset to death was 7.3 years (range 2 to 15 years) and the most common cause of death was heart failure due to cardiomyopathy. The main causes of death in early onset TTR-FAP from endemic areas of Japan and Portugal include cachexia or secondary infection.[8]

2.2.6 Non-V30M mutations

Published information on the natural history of patients with non-V30M mutations is very sparse. A recent prospective study following 60 patients from the UK and Canada provided information on the clinical features, natural history and outcome of patients with the T60A mutation, the most common variant in the UK population.[3] Patients were followed for a median of 31 months from diagnosis (range 0.4 to 132 months). The authors highlighted a number of important differences between the natural histories of T60A and V30M: 1) peripheral neuropathy is not a predominant symptom in patients with T60A, where the heart and autonomic nerves are predominantly affected; 2) T60A patients are older when they present with symptoms; and 3) prognosis is poorer than for V30M. The median age at onset of symptoms was 63 years (range 45 to 78). The median delay from symptom onset to diagnosis was 24 months (range 2-132 months). 42% of patients presented with cardiac symptoms, with the majority of patients presenting with echocardiographic evidence of cardiac amyloidosis (56 of 58 patients). 75% of patients presented with autonomic neuropathy (e.g. postural hypotension, altered bowel habit, upper gastrointestinal tract symptoms, urinary retention, and impotence); 54% had peripheral neuropathy at diagnosis. 20% of patients had neither autonomic nor peripheral neuropathy at diagnosis. Median survival from onset of symptoms was 6.6 years (95% CI 0.2 to 14) and 3.4 years (95% CI: 2.7 to 5.3) from diagnosis.[3]

2.2.7 Applicability of Coutinho classification of disease stages to NonV30M and non-endemic V30M populations

Although the manufacturer's submission acknowledges the clinical variability of TTR-FAP and that "no single staging system captures the myriad of manifestations of TTR-FAP", the use of three stages based on the Coutinho et al. classification is central to the economic model.

However, the generalisability and clinical utility of staging TTR-FAP using the Coutinho descriptors in a non-V30M population and indeed in a V30M population from a non-endemic area is uncertain. Apart from one study provided as part of the manufacturer's submission (Fx1A-OS-001), we are not aware from our own literature searches or through discussion with our clinical advisors, of any other studies that have applied this approach or explored its utility and generalisability to describing the progression of TTR-FAP non-V30M populations or other V30M populations. As with the Coutinho et al. study, study Fx1A-OS-001 is based upon a Portuguese population, which is different clinically to a UK population (see Appendix 2 for further details of this study).

The Coutinho classification is based primarily on sensory motor neuropathy and the extent of assistance required walking. It does not incorporate symptoms which may be an important feature in some presentations of the disease, such as cardiac or autonomic symptoms without significant lower limb involvement. The implications of this could be that a patient with the presentation just described for T60A (Section 2.2.6) could apparently remain in stage 1 for a considerable period of time as the requirement for assistance walking may not be the most apparent clinical feature of their disease and the aspect that most impacts on quality of life. The appropriateness of the staging system is discussed in further detail in the evaluation of the economic model.

2.3 Current treatment

Treatment normally involves the management of symptoms, and liver transplantation (for those patients who are suitable for transplantation).

2.3.1 Treatment of symptoms

Supportive therapy involves management of symptoms (e.g. alleviation of neuropathic pain) and supporting the function of failing organs (e.g. surgical management of carpal tunnel syndrome, and implantation of a permanent pacemaker). Gastroparesis, orthostatic hypotension, and urinary retention are among the other conditions sometimes associated with TTR-FAP, for which a range of treatments can be given.[2]

2.3.2 Liver transplantation

Liver transplantation is the only treatment available which can be given with curative intent. It aims to stabilise the disease by replacing the main source of mutant TTR with a liver that makes normal transthyretin.[2] It is regarded as an effective treatment for patients with V30M mutation,[9-10] though controlled studies are not available. There is uncertainty about the effects of liver transplantation in patients with non-V30M forms of the disease with cardiac involvement.[3] If a combined heart and liver transplant is not possible severe cardiac amyloidosis precludes liver transplantation as there may be progression of the heart disease despite the liver transplant.[2] Disease duration prior to transplantation, initial presentation with autonomic rather than peripheral neuropathy, TTR mutation and poor nutritional status (mBMI <600) have been identified as significant factors influencing survival following transplant.[9] It is important that transplantation occurs early in the course of the disease before too much

damage to the nerves or heart has occurred. This is because the transplant stabilises neuropathy rather than reverses it.[2] The occurrence or worsening of cardiac dysfunction is the main factor affecting prognosis after liver transplantation.[2]

Several case series reporting survival rates following transplant have been published or reported at conferences,[11-20] though some of these were very early cohorts or were small cohorts from single centres. The most complete data set on survival following liver transplantation in TTR-FAP patients are available from the Familial Amyloidotic Polyneuropathy World Transplant Registry (FAPWTR). The manufacturer's submission refers to two sources of data on outcome following liver transplant: the 2009 report of the FAPWTR,[21] and a 2004 publication from FAPWTR including patients reported to the register by the end of 2000.[22] Neither publication reports beyond five-year survival. The ERG searches (see Appendix 1) identified a conference abstract reporting 10 year survival data from the FAPWTR;[9] when we contacted the authors they provided further details.

There are currently 1853 patients with TTR-FAP recorded in the Register who have had a transplant, 88% with the V30M mutation and 12% with a non-V30M mutation (Professor HE Wilczek, Karolinska University Hospital Huddinge, Sweden, personal communication 12/06/2012). As would be expected, given the different natural history of these mutations, the non-V30M donor recipients were older than V30M patients at the time of receiving the transplant and a higher proportion received a liver transplant in conjunction (simultaneous or sequential) with another organ (Table 1).

Table 1: Characteristics of TTR-FAP patients who have had a liver transplant*

Mutation	**% Male**	**Age, years** **Mean (SD)** **Median (range)**	**Disease duration, years** **Mean (SD)** **Median (range)**	**Type of transplant**
V30M (n=1625)	54.4%	39 (10.2) 36.5 (21.4-73)	3.8 years (2.7) 3 years (0.3-30)	98% liver; 1.8% liver+kidney; 0.2% liver+heart
Non-V30M (n=228)	64.9%	50.9 (11) 52.6 (22.6-70)	4.0 years (3.5) 3 years (0-20)	84.6% liver; 14.4% liver+heart; 0.5% liver+kidney; 0.5% liver+kidney+heart

*(FAPWTR, Professor HE Wilczek, Karolinska University Hospital Huddinge, Sweden, personal communication 12/06/2012)

One-year survival was slightly greater for V30M (88.1%) than for non-V30M (82.9%) and this difference widened considerably over time: at 10 year follow-up over 70% of V30M patients survived compared to just over 40% in the non-V30M group (Table 2). Survival data on the three most common non-V30M mutations suggest that there may also be some variability in outcome across patients with different non-V30M mutations (Table 2). Ten year survival for patients with T60A, which is the more common genetic mutation in the UK, was 38%.

Table 2: Survival following transplant in patients with V30M and nonV30m TTR-FAP

	Survival				
Mutation	**1 year**	**3 years**	**5 years**	**7 years**	**10 years**
V30M (n=1625)	88.1%	84.5%	81.9%	79.1%	73.5%
Non-V30M(n=228)	82.9%	70.6%	58.1%	52.9%	43.2%
Most common non-V30M mutations					
Thr60Ala (n=17)	82.4%	69.7%	50.7%	38%	38%
Ser77Tyr (n=33)	84.8%	68.5%	55.9%	49.6%	35.5%
Tyr114Cys (n=15)	93.3%	80%	80%	80%	60%

*(FAPWTR, Professor HE Wilczek, Karolinska University Hospital Huddinge, Sweden, personal communication 12/06/2012)

Clinical advice to the ERG indicates that the number of transplants amongst TTR-FAP patients resident in the UK is very low as most patients have a non-V30M mutation often with cardiac involvement, presenting with symptoms that mean they are not suitable for a transplant (see Section 5.3.3 for further details).

2.4 Tafamidis and treatments in development

Tafamidis is claimed to work by stabilising the transthyretin tetramer (its protein structure), and by doing this it inhibits tetramer dissociation, the rate limiting step in the formation of transthyretin amyloid. It is taken orally (20mg per day) and has the trade name Vyndaqel. In November 2011 tafamidis was granted marketing authorisation by the European Medicines Agency for the treatment of transthyretin amyloidosis "*in adult patients with stage 1 symptomatic polyneuropathy to delay peripheral neurologic impairment*".[23] The rationale for the licensing limitation is provided in the assessment report.[24] The Committee for Medicinal Products for Human Use (CHMP) state that *"126 out of 128 patients (in the trial) were of stage 1 of the disease showing that the study population was homogenous. In this context, the CHMP was of the view that no data are available for stages 2 and 3…".*[24]

A specific obligation of the authorisation was for the manufacturer to conduct a further study of tafamidis in non-V30M patients who are registered on the THAOS registry.[24] The THAOS registry is funded by the manufacturer, with the aim of studying the natural history of patients with transthyretin amyloidosis.

Another stabiliser, diflunisal (a non-steroidal anti-inflammatory) has also been undergoing clinical development for a TTR-FAP population; a randomised placebo-controlled trial in patients with TTR-FAP is currently ongoing.[25] It is also used off-license for some UK patients. A drug treatment with a different mode of action, doxycycline and tauroursodeoxycholic acid, is currently being studied in a phase II open-label study.[26] Studies are also planned, or are underway, for two different agents which target the transthyretin gene.[27]

3 THE DECISION PROBLEM

3.1 Population

The manufacturer's submission considers all patients in England with stage one TTR-FAP. They present the results for a base case population of V30M and non-V30M patients and for subgroups of V30M and non-V30M patients separately.

3.2 Intervention

The intervention considered is tafamidis (Vyndaqel). Tafamidis (20mg once daily) is given in addition to conventional support therapy to patients whilst they remain in stage one of the disease as per the licence and the scope. Once patients progress to stage two they discontinue treatment with tafamidis and receive conventional support therapy only. Patients on tafamidis may also undergo liver transplantation.

3.3 Comparators

Current treatment normally involves the management of symptoms. There is also the possibility of liver transplantation for those patients suitable for transplantations and with no cardiac involvement. Both comparators were included by the manufacturer in the statement of the decision problem, though it is noted that the rate of liver transplant in England is low as many patients have cardiac involvement.

3.4 Outcomes

The outcomes of interest specified in the scope were quality of life, progression of peripheral neuropathy, mortality, and cardiac outcomes. Adverse events were also investigated. The manufacturer's submission focused on the outcomes used in the Fx-005 study, namely the Norfolk Quality of Life-Diabetic Neuropathy Scale (TQoL) and the Neuropathy Impairement Score-Lower Limb (NIS-LL). For the economic evaluation, outcomes were measured in quality-adjusted life-years (QALYs).

3.5 Time frame

The decision model presented in the submission takes a life-time perspective which is appropriate.

3.6 Treatment strategies

The manufacturer assessed only one treatment strategy in its base-case, tafamidis treatment until the patient progresses to stage 2.

4 CLINICAL EFFECTIVENESS

This section presents the ERG's assessment of the clinical effectiveness evidence for tafamidis in patients with TTR-FAP. Section 4.1 presents an overview of the manufacturer's approach in the submission. Section 4.2 outlines the additional systematic searches undertaken by the ERG. The following sections provide a summary and critique of the submitted evidence for tafamidis and its applicability to the UK setting.

4.1 Overview of the manufacturer's submission

Details of the clinical trial programme for tafamidis were supplied by the manufacturer. They also undertook a systematic review; the MEDLINE, EMBASE and Cochrane Library databases were searched without date restrictions, with studies selected using pre-specified inclusion criteria (in line with the AGNSS scope). The search was conducted on 16th November 2011, so was as up to date as possible. Trial registries and numerous conference proceedings were also searched between 2009 and 2011. Non-English language papers were excluded, which was reasonable in this context. Overall, the search strategy and eligibility criteria were comprehensive and were well-documented in the submission appendices.

Two reviewers independently selected studies for inclusion, which minimised the risk of reviewer error and bias affecting the selection process. Seven relevant records were identified, all from conference proceedings, although the references were not provided. The manufacturer did not use the conference proceedings, noting that the full clinical study reports, which the manufacturer provided to the ERG, gave a more comprehensive and complete account of the studies. Based on the ERG literature searches, all the key evidence appears to have been included. There were no studies comparing the use of tafamidis with liver transplantation. The ERG did identify two additional conference abstracts of case series which do not appear to relate to any of the other studies, although it is likely that these were unavailable when the November 2011 searches were undertaken by the manufacturer. We have summarised these studies in section 4.6.2 for completeness.

The main limitation of the submission was the sparse evidence available for tafamidis applicable to a UK population, where non-V30M mutations are most common. The single RCT in the submission was undertaken in a V30M population. Only a small uncontrolled pre-post

study of tafamidis was submitted for a non-V30M population. In the absence of a control group it is uncertain whether any response observed was solely related to the treatment received.

In relation to the main trial, the key limitation of the evidence presented in the MS was the lack of consideration of the impact of potentially clinically-important baseline imbalances between the tafamidis and placebo group on the trial results. The submission noted that there were baseline imbalances but did little by way of considering them any further, stating that they were not statistically significant differences.

Furthermore, the submission relies on comparing change in outcome measures from baseline to follow-up in the (non-V30M) pre-post study, with outcomes in the tafamidis arm and placebo arm of the main RCT (MS Section 3.18.1). In the absence of a natural history group for a non-V30M population, this might be regarded as reasonable, however, it assumes that over a 12-18 month period the rate of natural disease progression would be the same in the V30M and non-V30M populations. As outlined in the ERG background section, the clinical presentation and progression in people with V30M and non-V30M mutations is different, which was reflected in the submitted data.

In the submission (MS Section 3.18.1) the manufacturer also proposes that the similar reduced rates of progression in non-V30M patients receiving tafamidis, and V30M patients in the trial receiving tafamidis provides a rationale for extrapolating the results of the trial to a non-V30M population. This rationale is based on the assumption that a change of two points on a scale in a more severely affected population (with a higher baseline score) has the same clinical meaning as a two point change in a less severely affected population (with a much lower baseline score). It is uncertain whether such an assumption is appropriate.

These issues are discussed further in the ERG's appraisal of the submitted evidence.

4.2 ERG systematic review

A systematic review was undertaken by the ERG to ensure that all relevant data on tafamidis for the treatment of TTR-FAP was included in the manufacturer's submission. A comprehensive search strategy was undertaken (databases were searched up to 6th March 2012) in order to identify records on tafamidis, as well as records on TTR-FAP. Following the main searches, current awareness searches were run on a weekly basis on MEDLINE and EMBASE (up until 6th June 2012) in order to keep as up to date as possible with new publications in the field. Further details about the search strategy and the other review methods and procedures used can be found in Appendix 1.

4.3 Overview of included studies

A summary of the studies included in the ERG's systematic review is presented in Table 3. The manufacturer's submission included two studies of the effectiveness of tafamidis: a randomised placebo-controlled trial (Fx-005) of 128 participants with the V30M mutation, and an observational study using a pre-post design (Fx1A-201) of 21 participants with varying non-V30M mutations. The RCT was extended to become an open-label (single treatment arm) study (Fx-006). A correlation study (Fx1A-OS-001) examining the relationship between disease stage (in V30M patients) and outcomes was also included in the submission; none of the patients received tafamidis, therefore this study was excluded from consideration in the review of clinical effectiveness. However, since this study was relevant to the economic model, an evaluation is presented in Appendix 2 and population characteristics are reported below in Table 4. The ERG also identified two small case series,[28-29] both recently published as conference abstracts. There were no studies comparing the use of tafamidis with liver transplantation.

The studies in the MS have not yet been published as full papers. In addition to the information provided in the MS, the ERG has had access to the clinical trial report (CTR) of the main RCT (Fx-005), the extension study (FX-006) and the pre-post study (Fx1A-201), though not all appendices from the CTRs. The MS also provided the study protocol and statistical analysis plan for the RCT and other relevant documents in response to a request by the ERG (see Appendix 4). In the following sections the ERG provides (i) a summary of the evidence on the clinical effectiveness from each of the studies based on the information provided in the MS, the clinical study reports and the ERGs' own searches and (ii) a critical appraisal of the evidence.

Table 3: Studies of tafamidis included in the ERG's systematic review

Study design Study name	Information available	Population (type of mutation)	Primary endpoints	No of participants	No of sites (no of countries)
RCT Fx-005	MS Unpublished CTR, protocol, and statistical analysis plan. Conference abstracts[30-38]	V30M	NIS-LL response at 18 months, TQoL change from baseline at 18 months	128	8 (7)
RCT open-label extension Fx-006	MS Unpublished clinical study report Conference abstracts[39-43]	V30M	Not stated. Primary objective was to provide safety data.	71	7 (6)
Pre-post Fx1A-201	MS Unpublished clinical study report, protocol, and analysis plan. Conference abstracts[44-45]	Non-V30M	Transthyretin stabilisation	21	4 (4)
Case series Lozeron (2011)	Conference abstract[28]	V30M	Unclear	9	Unclear (1)
Russo (2012)	Conference abstract[29]	Non-V30M	Unclear	7	Unclear (1)

CTR=Clinical trial report, MS=Manufacturer's submission

4.4 Randomised Controlled Trial (Fx-005)

4.4.1 Objectives and methods

The primary objectives of the RCT were to evaluate the effect of tafamidis (once daily, 20mg) on disease progression over 18 months, and to evaluate safety and tolerability. The co-primary endpoints, as stated in the MS and the trial protocol, were (i) categorical NIS-LL response at 18 months (NIS-LL responders were defined as having a less than two point increase at a given time point), and (ii) change from baseline TQoL score at 18 months, assessed using the Norfolk Quality of Life – Diabetic Neuropathy (QoL-DN) questionnaire.

The NIS-LL is a tool which evaluates motor (muscle strength), sensory and reflex activity in the lower limbs and was developed for use in diabetic neuropathy studies. The maximum possible score is 88. The Norfolk QoL-DN, also referred to as TQoL in the MS and this report, is a patient-reported outcome measure developed to be sensitive to the different features of diabetic neuropathy. There are 35 scored items and the possible total score ranges from -4 to 135. For both NIS-LL and TQoL higher scores represent worsening outcomes. Further details of these outcome measures including strengths and limitations, is presented in Appendix 3 (the complete tools are also presented on pages 303-307 of the MS). Their suitability for assessing patients with TTR-FAP is discussed in section 4.4.4.

The secondary objective of the RCT was to determine the TTR stabilisation effect. Secondary endpoints are listed in Table 6. Large nerve fibre function assesses predominantly motor function, and small nerve fibre function assesses sensory function (both by combining scores of several different measures). Higher scores represent worsening outcomes for large nerve fibre function, and small nerve fibre function. Modified body mass index (mBMI) differs from BMI by compensating for the oedema formation associated with TTR-FAP. Further details on these outcome measures can be found in the MS (MS Table 4).

The study was conducted at eight sites in seven countries. Eligible patients were aged between 18-75 years, had a documented V30M TTR mutation, a positive amyloid biopsy, and peripheral and/or autonomic neuropathy with a Karnofsky performance status score of ≥50. Therefore patients with a range of functional impairment states were eligible, ranging from patients with minor, or no signs of disease, who can carry on with normal daily activities (Karnofsky score of 90-100) to patients who are unable to work, and require considerable assistance and frequent medical care (Karnofsky score of 50). Many of the patients were awaiting a liver transplant.

4.4.2 Analyses

The protocol and CTR detailed numerous analyses. The focus here is on the analyses presented by the manufacturer in the submission, in particular the analysis of primary outcomes

and on any additional analyses that the ERG viewed as important for interpreting the evidence. The analyses have been categorised as follows: the co-primary endpoint (ITT population) analyses, the secondary analyses of co-primary endpoints (also referred to in the MS as supportive analyses), secondary endpoint analyses, and post-hoc analyses, which reflects the CTR categorisations.

Co-primary endpoints (ITT population)

The MS stated that superior treatment efficacy of tafamidis compared with placebo would be demonstrated if statistically significant treatment differences favouring tafamidis were demonstrated for each of the co-primary endpoints, evaluated in the intention to treat (ITT) population. The ITT population was defined as comprising all randomised patients who received at least one dose of study medication (tafamidis or placebo) and who had at least one post-baseline efficacy assessment for both NIS-LL and Norfolk QoL-DN, or who discontinued study treatment due to death or liver transplant.

An analysis of covariance (ANCOVA) with baseline as covariate was used to compare TQoL scores between treatment groups. For patients without post-baseline TQoL assessments (e.g. the small number of patients receiving a liver transplant prior to a follow up assessment), the mean change from baseline at month 18 for patients who had post-baseline assessments was used to impute the change from baseline within each treatment group. A chi-squared test for proportions was used to compare NIS-LL response rates between treatment groups. There was no adjustment for baseline values in the NIS-LL responder analysis. Patients who stopped treatment to undergo a liver transplant were treated as non-responders for NIS-LL for time points after the date of liver transplant. The last observation carried forward method was used in both the NIS-LL response and TQoL analyses for patients having missing data at 18 months and at least one post baseline measurement.

Secondary analysis of co-primary end points

Several secondary analyses of the co-primary endpoints were performed. These included analyses of the co-primary endpoints using the efficacy-evaluable (EE) population. This population was defined as comprising all patients with month 18 NIS-LL and TQoL scores, who took at least 80% of prescribed study medication and who had no major protocol violations.

Another supportive analysis was a sensitivity analysis to impute response for liver transplant patients for NIS-LL response at 18 months. This was undertaken because the assumption in the primary analysis that all liver transplant patients were non-responders was regarded by the trialists as highly conservative. An estimated group probability of NIS-LL response at 18 months was calculated using the median baseline NIS-LL score for patients who underwent liver transplant, derived from a logistic regression. This probability was used to impute NIS-LL response. In a separate analysis the NIS-LL categorical response at 18 months was modelled as a function of treatment, gender, age, duration of symptoms, study site, and baseline NIS-LL score using logistic regression. An analysis was also conducted comparing within-group TQoL between baseline and 18 months.

Secondary endpoints

Analyses of the secondary endpoints were performed using only the ITT population. Imputed values using LOCF were not used in any of the repeated measures ANOVA analyses.

Post-hoc analyses

A number of post-hoc efficacy analyses were performed (a full list can be found in the CTR, Section 9.7.1.8). The following analyses were referred to in the MS. Treatment group differences in baseline and change from baseline for NIS-LL, TQoL, large nerve fibre function, small nerve fibre function and mBMI were descriptively analysed by individual study site, gender, NIS-LL responder status, and age at symptom onset. The rate of disease progression per month (as measured by the slope of the change from baseline over 18 months) for TQoL was compared between groups using mixed-effects models. Results of rate of disease progression per month analyses for large nerve fibre function, small nerve fibre function and mBMI were not referred to in the MS.

Among other post-hoc analyses conducted, demographics and baseline disease characteristics were compared by treatment group, in patients who underwent liver transplant. Also, NIS-LL response rates in each treatment group were compared using various responder definitions (i.e. changes from baseline in NIS-LL at 18 months ranging from -4 to 10 points); the results are presented in Figure 6 (p102) of the CTR.

4.4.3 Results

Trial population

Of the 162 participants screened, 128 were randomised (65 to tafamidis and 63 to placebo). The majority of patients (74 of the 128) were recruited at one Portugese site. The ITT population consisted of 125 participants: 64 tafamidis and 61 placebo; of the three patients not included in the ITT analyses, one had a negative genotype, and two discontinued treatment (due to adverse events) before having a post-baseline assessment. The efficacy evaluable (EE) population consisted of 87 participants (45 tafamidis and 42 placebo). The main reason for exclusion from the EE analyses was liver transplantation (13 patients in each group); much smaller numbers were excluded for important protocol deviations (2 patients in each group) and adverse events, pregnancy, or withdrawal of consent (eight patients in total).

Patient baseline characteristics are presented in Table 4, which also presents characteristics from the other submitted studies for ease of comparison. Differences between the trial treatment groups in Fx-005 include duration of symptoms, NIS-LL scores, TQoL scores, and large nerve fibre function scores. The tafamidis group had lower NIS-LL, TQoL, and large nerve fibre scores than the placebo group indicating that the patients in the tafamidis group were less impaired at baseline than the group receiving placebo. However, the tafamidis group also had a longer disease duration than the placebo group. This difference, coupled with difference in duration of symptoms, suggests the placebo group may have an underlying faster rate of disease progression.

Based on a consensus report from the Peripheral Nerve Society, a between group difference of two points on this scale may be clinically meaningful (see Appendix 3 for further details),[46] however there is some uncertainty about this (see 4.4.4). The clinical significance of the 3 point TQoL difference is unclear; as far as the ERG has been able to determine, there is no consensus as to what would constitute a clinically meaningful difference between two groups. These baseline imbalances are discussed further in Section 4.4.4. No baseline data on Karnofsky performance status scores were provided for the RCT, despite a Karnofsky score of ≥50 being one of the study eligibility criteria; the ERG requested the data but were informed by the manufacturer that the data was not collected as part of the trial.

Table 4: Patient demographic and baseline characteristics of the studies submitted by the manufacturer (Data were obtained from both the company submission (Tables 10, 17, and 28) and from the full study reports)

Characteristic	Fx-005 RCT (ITT population)		Fx1A-201 Pre-post study	Fx1A-OS-001 Correlation study (of disease stage to outcomes)			
	Placebo	Tafamidis	Tafamidis	Control	Stage 1	Stage 2	Stage 3
Number of participants	61	64	21	16	29	16	16
Genotype (type of mutation)	V30M	V30M	Non-V30M (8 types)	No mutation	V30M	V30M	V30M
Mean Age in years (SD) Median Age	38.4 (12.9) 34.0	39.8 (12.7) 35.5	63.1 (9.9) 64.3	34.8 (10.0) 35.3	39.0 (11.6) 36.0	46.5 (12.8) 44.7	55.0 (10.0) 52.7
Age group % ≤ 65 years % >65 years	 95 5	 92 8	 52 48	NR	NR	NR	NR
% Female	57	50	38	50	59	44	38
Height (cm) Mean (SD) Median	 167 (11.2) 165.5	 167 (10.1) 167	 171 (9.4) 172	 170 (7.6) 170.5	 165 (11.2) 162	 166 (9.4) 165	 165 (10.1) 165.5
Weight (kg) Mean (SD) Median	 63.9 (13.4) 64.0	 64.1 (11.9) 62.0	 72.6 (16.5) 75.0	 73.8 (17.7) 73.5	 65.5 (13.1) 64.0	 62.8 (15.1) 60.0	 54.7 (23.4) 50.6
mBMI (at screening) Mean (SD) Median	 1012 (213) 984	 1005 (165) 975	 1053 (207) 1048	 1199 (241) NR	 1032 (214) NR	 886 (310) NR	 760 (316) NR
Karnofsky performance status score Mean (SD) Median	 NR but ≥50††	 NR but ≥50††	 74.8 (14.0) 70.0	 100 (0) NR	 86.2 (6.2) NR	 60.0 (5.2) NR	 45.6 (5.1) NR
TTR (mg/dL) Mean (SD) Median	NR	NR	 19.3 (4.7) 19.8	NR	NR	NR	NR
Duration of symptoms (months) Mean (SD) Median	 34.7 (32.9) 21.0	 47.0 (48.4) 28.0	 64.7 (60.8) 45.5	NA	 31.2 (33.6) 16.8	 92.4 (34.8) 94.8	 170.4 (52.8) 170.4
Age at symptom onset (years) Mean (SD) Median	 35.7 (11.5)* 32.2*	 36.6** 33.1**	 59.3 (9.2) 61.0	NA	 36.9 (10.7) 33.3	 39.3 (13.1) 35.0	 41.7 (12.6) 37.0
Age at diagnosis (years)							

Characteristic	Fx-005 RCT (ITT population)		Fx1A-201 Pre-post study	Fx1A-OS-001 Correlation study (of disease stage to outcomes)			
	Placebo	Tafamidis	Tafamidis	Control	Stage 1	Stage 2	Stage 3
Mean (SD) Median	NR	NR	61.6 (9.6) 61.0	NR	NR	NR	NR
NIS total score Mean (SD) Median	NR	NR	48.7 (44.3) 45	NR	NR	84.8 (16.1) 84.8	128.7 (35.8) 124.2
NIS-LL Mean (SD) Median	11.4 (13.5) 6 (range 0,57)	8.4 (11.4) 4 (range 0,54)	27.6 (24.7) 18 (range 0,70)	0.0 (0.0) 0	6.9 (6.3) 4 (range 0,28)	55.3 (6.7) 54.5 (range 43,69)	69.0 (9.8) 65 (range 52,86)
Norfolk QoL-DN (TQoL) Mean (SD) Median	30.8 (26.7) 22 (range 0,107)	27.3 (24.2) 19 (range -1, 110)	47.8 (35.1) 38 (range 5,104)	2.6 (5.0) 1 (range 0,20)	21.0 (14.5) 17 (range -1,61)	73.1 (27.5) 79 (range 13,123)	95.4 (21.7) 100.5 (range 32,121)
Large nerve fibre function (summated scores from 7 tests) Mean (SD) Median	7.8 (9.1) 7.4	8.7 (8.5) 9.7	6.1 (5.9) 6.7	NR	3.7 (8.4)	21.9 (2.3)	21.2 (3.0)

Abbreviations kg: kilograms, mBMI: modified body-mass index, NA: Not applicable, NR: Not reported, NIS: Neuropathy Impairment Score, NIS-LL: Neuropathy Impairment Score – Lower Limb, mg/dL: milligrams per decilitre, Norfolk QoL-DN: Norfolk Quality of Life – Diabetic Neuropathy, RCT: Randomised controlled trial, SD: standard deviation, TQoL: Total Quality of Life, TTR: transthyretin. * Reported only for n=63, ** Reported only for n=65 † A higher score represents worse nerve function †† A Karnofsky score of ≥50 was an inclusion criterion Large nerve fibre function for Fx1A-201 used summated scores from 5 tests

Co-primary endpoints (ITT population)

At 18 months in the ITT population the tafamidis group scores deteriorated by a mean of 2.0 points on the Norfolk TQoL versus a deterioration of 7.2 points in the placebo group (difference=-5.2, 95% CI: -11.8 to 1.3, p=0.12). There was no statistically significant difference between treatment groups in change from baseline TQoL (Table 5).

In the tafamidis group at 18 months, 45% of participants in the tafamidis group were classified as responders on the NIS-LL compared to 30% in the placebo group. There was also no statistically significant difference between groups on this outcome measure (Table 5)

Table 5: Results of Co-primary endpoints in RCT Fx-005

Primary Endpoint (at 18 months)	**Tafamidis N=64**	**Placebo N=61**	**Difference**
Change from baseline in TQoL Mean (SD) Least squares Mean (SE)	 2.4 (14.6) 2.0 (2.3)	 6.9 (22.9) 7.2 (2.4)	 -5.2 (95% CI: -11.8 to 1.3), p=0.12
NIS-LL response to treatment	29/64 45.3%	18/61 29.5%	15% (95% CI: -1.15 to 32[†]) p=0.07

[†]Confidence intervals calculated by the ERG

Secondary analyses of co-primary endpoints

At 18 months in the EE population both the difference in change from baseline TQoL (tafamidis: 0.1 versus placebo: 8.9; difference=-8.8, 95% CI: -17.4 to -0.2, p=0.045) and the difference in the proportion of NIS-LL responders (60% tafamidis versus 38% placebo; difference=22%, 95% CI: 1 to 40, p=0.04) were statistically significantly different, favouring treatment with tafamidis.

A sensitivity analysis for NIS-LL response was also reported which used an imputed response based on median baseline NIS-LL score for liver transplant patients, rather than classifying them as non-responders as in the primary analysis; there was a statistically significant difference between groups favouring treatment with tafamidis (55% tafamidis versus 36% placebo, p=0.04).

When the NIS-LL categorical response at 18 months was modelled as a function of treatment, gender, age, duration of symptoms, study site, and baseline NIS-LL score using logistic regression, gender (p=0.0033) and baseline NIS-LL (p=0.0112) were the only significant covariate predictors of NIS-LL response at Month 18 in the ITT population; these results were

not reported in the MS, but were obtained from the CTR (CTR Section 11.4.2.1 and Table 14.2.2.1).

Secondary endpoints

Table 6 summarises the results for the main secondary endpoints (all for the ITT population).

The results in Table 6 show statistically significant results favouring treatment with tafamidis at the 12 and 18 month time points for mean change from baseline in NIS-LL, and small nerve fibre function.

At 6, 12, and 18 months there was an improvement in modified BMI in the tafamidis group compared to baseline and a deterioration in the placebo group and the between group differences were statistically significant at all time-points. At 18 months, 98% of tafamidis patients had achieved TTR stabilisation compared to none of the placebo group. For large nerve fibre function the placebo group deteriorated more than the tafamidis group and the difference was statistically significant at 6 and 12 months, but not at 18 months (p=0.066).

In relation to the endpoints relating most to the study's primary endpoints, the change from baseline TQoL differences at 6 months and 12 months were not statistically significant; the results favoured placebo treatment at six months and treatment with tafamidis at 12 months. For the NIS-LL responder analyses a higher proportion of the tafamidis group were responders than the placebo group and this difference was statistically significant at 12 months.

The NIS-LL subscales analyses indicated a statistically significant difference between groups in favour of tafamidis for the muscle weakness subscale (p=0.013) but not for the reflexes (p=0.26) or sensation (p=0.64) subscales at 18 months.

The Norfolk QoL-DN change from baseline individual domain score results showed no statistically significant differences between the treatment groups for any of the domains at 18 months: physical function/large fibre, p=0.12; activities of daily living, p=0.76; symptoms, p=0.31; small fibre function, p=0.38; and autonomic function, p=0.61.

Post-hoc analyses

The average rate of change on the NIS-LL was an [redacted]NIS-LL units per month in the tafamidis group compared to an [redacted] NIS-LL units per month in the placebo. The difference ([redacted] units per month) was [redacted] (p= [redacted]). The average rate of change for TQoL showed an [redacted] of [redacted] units per month for tafamidis compared to [redacted] units per month for placebo. The difference ([redacted] units per month) was [redacted] (p=[redacted]).

Participants who left the trial to receive a liver transplant had a longer disease duration than the remaining participants, although the baseline characteristic between the two liver transplant groups appeared balanced.

Table 6: Results for secondary endpoints in the randomised controlled trial (ITT population)

Secondary Endpoint	Results (95% CI)		Difference between groups (95% CI)	p-value
	Tafamidis	Placebo		
TQoL at 6 months*	1.2 (-2.8 to 5.1)	0.2 (-3.8 to 4.3)	0.9 (-4.7 to 6.6)	0.74
TQoL at 12 months*	1.5 (-3.0 to 6.1)	4.5 (-0.1 to 9.1)	-3.0 (-9.4 to 3.5)	0.37
% NIS-LL responders at 6 months	60.9 (49 to 72.9)	54.1 (41.6 to 66.6)	6.8	0.44
% NIS-LL responders at 12 months	54.7 (42.5 to 66.9)	32.8 (21 to 44.6)	21.9	0.01
NIS-LL at 6 months*	1.26 (-0.01 to 2.53)	2.08 (0.77 to 3.38)	-0.81 (-2.63 to 1.00)	0.38
NIS-LL at 12 months*	1.35 (-0.23 to 2.93)	4.72 (3.14 to 6.31)	-3.37 (-5.61 to -1.13)	0.004
NIS-LL at 18 months*	2.81 (0.93 to 4.69)	5.83 (3.93 to 7.73)	-3.024 (-5.70 to -0.35)	0.03
mBMI at 6 months*	17.1 (-0.6 to 34.7)	-30.5 (-48.6 to -12.4)	47.6 (22.3 to 72.9)	0.0003
mBMI at 12 months*	19.4 (-0.8 to 39.7)	-30.7 (-51.0 to -10.5)	50.2 (21.5 to 78.8)	0.0007
mBMI at 18 months*	39.3 (16.5 to 62.2)	-33.8 (-57.2 to -10.3)	73.1 (40.4 to 105.8)	<0.0001
Large nerve fibre function at 6 months*	0.58 (-0.36 to 1.53)	1.93 (0.97 to 2.90)	-1.35 (-2.71 to 0.00)	0.050
Large nerve fibre function at 12 months*	1.11 (0.04 to 2.18)	3.07 (2.0 to 4.13)	-1.96 (-3.47 to -0.45)	0.012
Large nerve fibre function at 18 months*	1.52 (0.29 to 2.76)	3.17 (1.92 to 4.43)	-1.65 (-3.41 to 0.11)	0.066
Small nerve fibre function at 6 months*	0.24 (-0.26 to 0.74)	0.72 (0.20 to 1.23)	-0.48 (-1.19 to 0.24)	0.19
Small nerve fibre function at 12 months*	0.39 (-0.15 to 0.94)	1.32 (0.78 to 1.87)	-0.93 (-1.70 to -0.16)	0.018
Small nerve fibre function at 18 months*	0.34 (-0.28 to 0.95)	1.62 (0.99 to 2.24)	-1.28 (-2.16 to -0.40)	0.005
% achieving TTR stabilisation at 8 weeks	98 (95 to 100)	7 (0.4 to 13.0)	91	<0.0001
% achieving TTR stabilisation at 6 months	100 (100 to 100)	5 (0 to 11)	95	<0.0001
% achieving TTR stabilisation at 12 months	98 (94 to 100)	2 (0 to 6)	96	<0.0001
% achieving TTR stabilisation at 18 months	98 (94 to 100)	0 (0 to 0)	98	<0.0001

* mean change from baseline (least squares) Table data obtained from the clinical trial report

Safety results

Adverse events were presented by MedDRA (v.10) system organ class and by preferred term. Nearly all patients reported at least one treatment-emergent adverse event (TEAE). A TEAE was defined as either an event starting inclusively in the period after the start of study treatment and up to 30 days after the last study treatment administration, or a specific event which started prior to starting study treatment but worsened after dosing.

The MS provides a summary of adverse events which reflects the data provided in the CTR (MS Tables 25, 26, and 27). A brief summary of these is provided below.

Only 5 tafamidis patients and 2 placebo patients did not report a TEAE. The TEAEs reported more frequently in the tafamidis group compared with the placebo group included diarrhoea, urinary tract infection, pain in extremity, upper-abdominal pain, myalgia, and vaginal infections. The TEAEs reported more frequently in the placebo group included constipation, atrioventricular first degree block, fatigue, muscle spasm, paraesthesia, neuralgia, and hypoaesthesia.

Of the TEAEs considered to be due to study treatment by the clinical investigators, there was greater incidence of urinary tract infections (10.8% versus 0%), upper abdominal pain (7.7% versus 3.2%) and pain in extremity (7.7% versus 4.8%) in the tafamidis group, and greater incidence of headache (15.9% versus 7.7%), neuralgia (11.1% versus 1.5%) muscle spasms (7.9% versus 1.5%), peripheral oedema (7.9% versus 1.5%), fatigue (7.9% versus 0%) and paraesthesia (9.5% versus. 0%) in the placebo group.

The incidence of moderate and severe TEAEs, and of withdrawals due to TEAEs (7 withdrawals in total), was similar between the groups. Four patients died (three randomised to placebo and one to tafamidis), all following liver transplantation.

4.4.4 Assessment of study quality

The risk of bias (internal validity) is discussed in this section along with other aspects relating to study quality. Issues relating to generalisability (external validity) are discussed in Section 4.8.

Risk of bias assessment

The ERG undertook an independent assessment of the risk of bias in the RCT, using the Cochrane Risk of Bias Tool (Table 7). The randomisation process was centralised, utilising an

interactive voice response system which produced participant treatment allocations and study identification numbers. Both participants and investigating staff were blinded to treatment, with the two treatments being identical in appearance. Across the two groups, similar numbers of participants withdrew from the trial, and for similar reasons. All the outcomes pre-specified in the trial protocol were reported in the trial report. The ERG therefore concluded, based on the methodology described in the study protocol and CTR, that the trial had a low risk of bias for all domains of the Cochrane Risk of Bias tool.

Table 7: Risk of Bias assessment results for the RCT (Fx-005)

Bias	ERG judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Centralised randomisation was used (an interactive voice response system)
Allocation concealment (selection bias)	Low risk	It appeared that the centralised randomisation system produced study identification numbers and treatment allocations by participant (rather than producing an allocation list). A statistician who was independent of the study conduct was the only individual unblinded to treatment identity.
Blinding of participants and researchers (performance bias)	Low risk	The two treatments were identical in appearance. A statistician who was independent of the study conduct was the only individual unblinded to treatment identity.
Blinding of outcome assessment (detection bias)	Low risk	A statistician who was independent of the study conduct was the only individual unblinded to treatment identity.
Incomplete outcome data (attrition bias)	Low risk	128 patients were randomised and 125 patients were included in the modified intention-to-treat analysis (one patient from the tafamidis group and two from the placebo group were excluded from the ITT analyses). Balance across groups was also seen for the EE population.
Selective reporting (reporting bias)	Low risk	Results for all pre-specified endpoints in the protocol and statistical analysis plan were reported in the CTR which the ERG had access to.
Other bias	None identified	-

Baseline differences and their implications

However, notwithstanding the Cochrane Risk of Bias tool assessment results, it was nevertheless apparent from the baseline characteristics (Table 4) that there were potentially clinically meaningful differences between the groups at baseline. Bearing in mind the risk of bias assessment results, it is likely that these differences arose by chance (the use of 'minimisation' in the randomisation process was not reported). The differences suggested that the tafamidis group was less impaired than the placebo group at baseline, and also implied

the possibility of an underlying difference in disease progression rates between the treatment groups.

In the CTR it was noted that a 2 point difference in NIS-LL was clinically meaningful. This was based on a consensus report from the Peripheral Nerve Society.[46] This difference existed between the tafamidis and placebo groups at baseline, with the tafamidis group having lower scores (i.e. less neurological impairment) than the placebo group. Furthermore, a baseline difference in duration of symptoms was evident which contrasted with this NIS-LL difference. Despite having less (NIS-LL) impairment, the duration of TTR-FAP disease symptoms was longer in the tafamidis group than the placebo group; a difference of 7 months in the median duration of symptoms (difference in means: 12.3 months).

The manufacturer's submission merely stated that there were no statistically significant differences at baseline between the treatment groups. There was no further consideration in the MS of the possible implications of the baseline differences outlined in Section 4.4.3 for the interpretation of the results of the trial. Although baseline TQoL scores were incorporated into the primary analysis of change from baseline in TQoL, no such analyses were presented in the MS which incorporated baseline NIS-LL scores. This has implications for the reliability of many of the results presented in the MS.

A basic estimate can be made of the difference between groups in disease progression at baseline by dividing the change in median NIS-LL score (from time of first symptoms, to baseline) by the median duration of symptoms for each group, and then subtracting one from the other. The ERG undertook this calculation based on an assumption (based on clinical advice) that i) patients are likely to have a very low NIS-LL score when they notice symptoms for the first time, though not zero and ii) that, whichever low value is used (we have used a value of 2) it would be unlikely to differ much between the treatment groups at such an early stage of disease.

This estimate suggests that, prior to randomisation, the underlying rate of disease progression was more than twice as fast in the placebo group (Table 8). Even if it were assumed that the placebo group (having faster disease progression) had a median NIS-LL score of 3 at the time of first symptoms, and the tafamidis group had a score of 2, the rate of disease progression to baseline is still estimated to be twice as fast in the placebo group. These calculations are not intended to provide accurate stand-alone estimates of disease progression for each treatment group. They are simply used here to highlight the likely impact of the baseline imbalances on the *difference* in disease progression rates between the tafamidis and placebo groups.

It therefore appears plausible that some of the effect estimates presented in the submission may be due to a combination of an underlying disease rate difference, and the effect of tafamidis treatment, rather than due to the effect of tafamidis treatment alone. This issue was not explored in the MS, although in the CTR baseline NIS-LL was reported to be a significant covariate predictor of NIS-LL response at 18 months. The ERG did not have access to the trial data set to explore the rate difference issue further, but the FDA (having access to the dataset) did conduct exploratory analyses. Further details of the FDA analyses are provided in section 4.4.5.

Table 8: ERG estimates of the rates of disease progression prior to randomisation

Treatment Group	Median Baseline NIS-LL score	Median duration of Symptoms (months)	Estimated rate of disease progression (NIS-LL units/month)	
			Scenario 1†	Scenario 2††
Placebo	6.0	21.0	0.190	0.143
Tafamidis	4.0	28.0	0.071	0.071
Difference	2.0	-7.0	0.119	0.072

†Assuming both treatment groups have a median NIS-LL score of 2 when symptoms are first noticed ††Assuming the placebo group has a median NIS-LL score of 3, and the tafamidis group a median score of 2, when symptoms are first noticed.

Co-primary outcomes

Given the rarity of the condition, it is unsurprising that outcome assessment tools tailored specifically for the TTR-FAP population do not exist. The co-primary endpoints used in the trial were the NIS-LL and Norfolk QoL-DN tools which have mainly been used amongst patients with diabetes. This seems a reasonable approach in the absence of experience with other measures in people with TTR-FAP, since these are validated measures, and there is experience of their use in diabetes trials. However, there are nevertheless limitations to consider. A key limitation of the NIS-LL is that it only evaluates lower limb impairment. The use of NIS, which evaluates both upper and lower limb impairment, may have provided a more complete assessment of neuropathy, particularly since it has been reported that V30M patients from non-endemic areas appear to have rapid progression of symptoms, with some patients reporting upper limb symptoms before lower limb symptoms.[8]

The Peripheral Nerve Society suggests that a mean change of two points between an intervention and placebo on the NIS is clinically meaningful and this was used by the manufacturer in their submission.[46] The consensus is based on the rationale that the least degree of neurological abnormality that a physician can recognise is equivalent to two points on the NIS (one point for each side of body). The robustness of this assumption is unclear,

particularly as applied to a TTR-FAP population. It is also unclear whether from a patient perspective this is a difference that patients would perceive as beneficial.

The Norfolk QoL-DN tool is considered to have reasonable reliability and validity in a diabetic population, but it does not capture the emotional/psychological components of TTR-FAP. Also, it may not capture the impact of all important aspects of the condition on quality of life where peripheral neuropathy is not the predominant symptom; although the measure does assess autonomic nerve function, there are only three items on this domain (related to vomiting, diarrhoea and dizziness). These are common symptoms in TTR-FAP, but other relevant symptoms that may impact on quality of life are not assessed, such as renal and urinary symptoms and erectile or sexual dysfunction. As far as the ERG is aware from the information presented in the MS and the ERG's own searches, a minimum clinically meaningful difference has not been established for this scale. See Appendix 3 for further details of the strengths and limitations of both these measures when applied to a population with TTR-FAP.

Other analyses

Although balance was seen across the groups in the EE population with respect to the loss of patients to liver transplant and losses due to other reasons, this is a subset of the original randomised population. In addition, it is likely to have been subject to the baseline imbalances discussed earlier. Also there were many secondary analyses performed, with no adjustment for multiple statistical tests.

It should also be noted that one analysis (TQoL rate of change per month) was classified in the MS as a supportive (secondary) analysis. It was clear from the CTR (Section 9.7.1.8), and the statistical analysis plan, that the analyses of rates of change per month were not pre-planned supportive analysis, but were performed post-hoc.

Handling of drop-outs or missing data

The last observation carried forward (LOCF) approach is commonly used to handle the issue of missing data, especially for continuous outcomes, and can provide a conservative estimate of treatment effects. However, such conservative estimates are only likely when treatments improve a symptom. Conversely, where treatment goals are to slow the rate of disease progression, or maintain symptoms, as is the case with tafamidis, LOCF can sometimes result in overestimation of the treatment effect.

For liver transplant patients the manufacturer used a conservative approach for the categorical NIS-LL response outcome: patients stopping treatment due to liver transplantation

(13 in each treatment group) were deemed to be NIS-LL non-responders. For the remaining patients (four randomised to tafamidis and two to placebo) a less conservative approach was used: LOCF was used to impute missing data for month 18. Similarly, LOCF was used for the TQoL for the 18 months analysis (for 12 tafamidis and 10 placebo patients). The use of LOCF for these two primary outcomes could have slightly biased the analyses in favour of tafamidis, since slightly more tafamidis patients were assigned a LOCF value (resulting in the recording of a favourable response, despite the data being missing). However, the effect of this remains unclear when also considering the likely baseline differences in disease progression between the treatment groups i.e. it could possibly favour placebo if there was a faster rate of disease progression in the placebo group.

Length of treatment and follow up

The 18 month treatment duration used in the trial seems reasonable, and there is evidence on safety for a 30 month period for a small group of patients. However, the long-term safety and efficacy are unknown. In practice, it is likely that patients would take tafamidis for longer than 30 months. Also, any durability of the effect of tafamidis in those patients who stop treatment is unknown.

4.4.5 FDA analyses of the RCT data

The US FDA produced a report on tafamidis with the purpose of informing the Peripheral and Central Nervous System Drugs Advisory Committee for their Meeting of May 24, 2012.[47] Having access to the full trial data set, an FDA statistician undertook further analyses, with results presented within the report. Some of these are summarised below.

In light of the baseline imbalances between treatment groups an analysis was performed to explore whether the NIS-LL responder status was associated with baseline NIS-LL. A logistic regression with NIS-LL response at month 18 as dependent variable and treatment and baseline NIS-LL score as independent variables was undertaken. The FDA report states that this analysis produced a p-value of 0.161, compared to the p-value for the protocol-specified primary analysis for NIS-LL of 0.068. The report went on to suggest that NIS-LL score at baseline was therefore associated with responder status – patients with lower scores at baseline were more likely to be NIS-LL responders. Given that the tafamidis group had lower scores at baseline there is therefore a possibility that the treatment effect could have been overestimated for NIS-LL.

When an analysis of NIS-LL response (at 18 months) by site was performed, a higher proportion of responders was seen in the tafamidis group at the Porto (Portugal) site (61%

versus 28%) than at all the other sites combined (25% versus 32%). A possible explanation for this mentioned in the FDA report was that there may have been bias at site 1 (the report mentions the feasibility of effectively unblinding patients by knowledge of baseline and follow up TTR concentrations, although it should be noted there is no evidence to suggest this occurred). It was also noted that the explanation could be due population differences. The Porto site recruited a largely homogenous (endemic) population, and the group of combined other sites was likely to be more variable.

A comparison of baseline characteristics was made between the largest recruiting site ('site 1', Porto, which recruited 74 of the trial's 128 participants) and the remaining sites. The Porto participants were younger, had lower median NIS-LL and TQoL scores, and had a shorter duration of disease at baseline (by 24 months) than patients at the remaining sites. Further analysis by FDA showed that, at baseline, the Porto patients were evenly balanced across treatment groups for median age, NIS-LL, and duration of symptoms, although placebo patients had higher median TQoL scores (tafamidis: 14 versus placebo: 19). For all the other sites combined (54 patients in total), the placebo patients had larger median NIS-LL scores (tafamidis: 3.8 versus placebo: 10.3) but shorter duration of symptoms (tafamidis: 51.8 months versus placebo: 38.6 months), suggesting a 'much faster disease course' (see p69 of FDA report) in the placebo group.[47]

These FDA analyses suggest that baseline NIS-LL differences in the trial may have resulted in an overestimation of treatment effect on the NIS-LL. They also suggest that patients from both endemic and non-endemic areas were recruited into the RCT, and that tafamidis might be effective in an endemic V30M population. However, it should be borne in mind that these were exploratory analyses.

4.5 Open-label extension study (Fx-006)

4.5.1 Study details and main results

This study aimed to evaluate the long-term safety, tolerability, and efficacy of tafamidis and to determine TTR stabilisation. The open-label extension was for 12 months, following on from the 18 month RCT treatment. Patients randomised to placebo in the RCT received tafamidis (referred to as the placebo-tafamidis group) and patients randomised to tafamidis continued to receive the drug (the tafamidis-tafamidis group). The efficacy outcomes were NIS-LL, TQoL, large nerve fibre function, small nerve fibre function, and mBMI. There were 38 patients in the tafamidis-tafamidis group and 33 patients in the placebo-tafamidis group. The patient baseline

characteristics are outlined in Table 9 (further detail is reported in Table 22 of the MS). Data were analysed using a mixed model analysis of variance, with no imputations used.

Table 9: Patient baseline characteristics in the trial open-label extension study Fx-006

Characteristic	Tafamidis-tafamidis	Placebo-tafamidis
Number of participants	38	33
Age (years)		
Mean (SD)	42.0 (14.1)	40.7 (13.7)
Median	37.5 (range 26 to 76)	36.0 (range 24 to 73)
NIS-LL		
Mean (SD)	8.4 (13.2)	17.5 (20.8)
Median	5.3	10.0
TQoL		
Mean (SD)	21.1 (21.9)	38.1 (31.9)
Median	11.0	28.0
Large fibre function		
Mean (SD)	6.7 (8.5)	10.1 (10.7)
Median	5.0	10.8
Small fibre function		
Mean (SD)	4.8 (4.3)	7.1 (4.4)
Median	4.2	7.4
mBMI		
Mean (SD)	1068 (142)	990 (265)
Median	1038	946

*Medians

Although the full study report notes that this study was not powered to detect statistically significant differences, it nevertheless aimed to test three main hypotheses: 1) the sustainability of the effect of tafamidis in delaying disease progression in patients treated with tafamidis for 30 months across studies Fx-005 and Fx-006, 2) the superiority of tafamidis compared with placebo in delaying disease progression in patients treated with tafamidis for 12 months (placebo-tafamidis group) and, 3) an early start treatment effect in delaying disease progression. The results relating to these hypotheses are summarised below:

Table 10 presents the mean monthly rates of change for the outcome measures for the group that received tafamidis in the main trial and continued to receive tafamidis in Fx006. The submission reported that the treatment effect of tafamidis was sustained over 30 months, with no statistically significant differences (between patients receiving tafamidis in Fx-005 for 18 months and the tafamidis-tafamidis group in Fx-006) found for NIS-LL, large nerve fibre function, small nerve fibre function, and TQoL. However, the increase in mBMI seen over the first 18 months was not observed over the last 12 months (rate of change: Fx-005, 1.85/month versus Fx-006, -2.0/month, p=0.0006).

Table 10: Results for the 'sustainability of the treatment effect' of tafamidis in study Fx-006, ITT population (source: full study report)

Endpoint	Mean rate of change per month (SE)				
	Tafamidis-Tafamidis (n=38)				Placebo (n=61)
	Tafamidis 18 months (Fx-005)	Tafamidis-12 months (Fx-006)	p-value	Tafamidis-Tafamidis 30 months (Fx-006)	Placebo 18 months (Fx-005)
NIS-LL	0.08 (0.06)	0.11 (0.07)	0.600	0.10 (0.04)	0.34 (0.05)
TQoL	-0.03 (0.15)	0.25 (0.20)	0.163	0.04 (0.07)	0.46 (0.15)
Large Nerve Fibre Function	0.06 (0.03)	0.05 (0.05)	0.930	0.06 (0.02)	0.18 (0.03)
Small Nerve Fibre Function	0.03 (0.02)	0.05 (0.02)	0.335	0.03 (0.01)	0.09 (0.02)
Modified BMI	1.85 (0.73)	-2.0 (1.04)	0.0006	0.37 (0.44)	-1.62 (0.62)

Table 11 presents the mean monthly rates of change for the outcome measures for the group that received placebo in the main trial and then moved to tafamidis in Fx006. The rates of change of NIS-LL (p=0.01) and TQoL (p=0.0003) were statistically significantly lower (possibly indicating a slowing of disease progression) following 12 months of tafamidis treatment in the placebo-tafamidis group, compared with the previous 18 months of placebo treatment (Table 11).

Table 11: Results for the 'superiority of the treatment effect' of tafamidis in study Fx-006, ITT population (source: full study report)

Endpoint	Mean rate of change per month (SE)			
	Placebo-Tafamidis (n=33)			Tafamidis (n=64)
	Placebo 18 months (Fx-005)	Tafamidis-12 months (Fx-006)	p-value	Tafamidis 18 months (Fx-005)
NIS-LL	0.34 (0.06)	0.16 (0.08)	0.0103	0.16 (0.05)
TQoL	0.61 (0.16)	-0.16 (0.21)	0.0003	0.12 (0.15)
Large Nerve Fibre Function	0.18 (0.04)	0.11 (0.05)	0.2133	0.09 (0.03)
Small Nerve Fibre Function	0.09 (0.02)	0.04 (0.03)	0.055	0.02 (0.02)
Modified BMI	-1.77 (0.78)	5.19 (1.13)	<0.0001	2.05 (0.61)

At 30 months the tafamidis-tafamidis group had statistically significantly lower NIS-LL scores (p=0.04), and preserved large nerve fibre function (p=0.007) than the placebo-tafamidis

group. However the differences were not significantly different for small nerve fibre function (p=0.09), mBMI (p=0.44) and TQoL (p=0.30).

The main safety results can be summarised as follows: 37 of 44 (84%) patients in the tafamidis-tafamidis group and 40 of 41 (98%) patients in the placebo-tafamidis group reported at least one TAEA. No particular trends were observed between treatment sequence groups. 14 of 44 patients (31.8%) in the tafamidis-tafamidis group, and 18 of 41 patients (43.9%) in the placebo-tafamidis group experienced at least one TEAE that was considered at least possibly related to study medication. Six patients, five in the tafamidis-tafamidis group and one in the placebo-tafamidis group, discontinued to undergo liver transplantation. There were no discontinuations due to a TAEA.

Critique of the study design

The manufacturer's submission did not present any consideration of the limitations of the design used in this study. Participants were a subset of the original trial population (eligible patients had to have completed the month 18 visit in the RCT) and it is unclear whether they fully represent all the patients originally randomised. For example, patients who left the trial for a liver transplant had longer disease duration than those who remained.

For comparisons between tafamidis-tafamidis and placebo-tafamidis, the results may have been influenced by baseline imbalances. For the within-group comparisons it is possible that natural changes in disease progression rates across time could have affected the outcomes (rather than just an effect from tafamidis).

Another limitation relates to blinding: both patients and outcome assessors knew that tafamidis was being taken, which may have influenced the results for several of the assessed outcomes. However, the study does provide a further year of data on safety, and increases the sample of patients for which safety data has been recorded after taking tafamidis.

4.6 Non-randomised studies

There were three non-randomised studies of tafamidis: one pre-post study, which was included in the submission (Fx1A-201), and two small case series which were recently published as conference abstracts and identified by the ERG systematic review (see Table 3 in section 4.3)

4.6.1 Pre-post study (Fx1A-201)

This study investigated the effects of tafamidis in a group of patients with a nonV30M mutation. The primary objective of the pre-post study was to determine TTR stabilization at steady-state as measured by a validated immunoturbidimetric assay. The secondary objectives were to evaluate safety and efficacy using a similar set of outcome measures to the main trial. Additional outcome measures were NIS, NIS-Upper Limb (UL), NIS-Lower limb (LL), SF36 (a quality of life measure) and Karnofsky Performance Status score. Monthly rate of change on several of the outcome measures was also calculated.

Pre-treatment rate of change of disease progression was defined as the ratio of the baseline scores for each efficacy outcome and duration of symptom onset. Duration of symptom onset was estimated from the time of first symptom as reported by the patient. In the full study report the manufacturer stated that, due to the uncertainty in assessing the time of symptom onset and the small number of patients, the results for such analyses were exploratory, and should be interpreted with caution.

There were 21 patients in the study from four countries, There were eight different mutations including T60A (4 patients). Mean patient age was 63 years, mean duration of symptoms was 65 months and the median NIS-LL (18.0) and TQoL (38.0) scores were higher than for the RCT groups (see Table 4). Three patients discontinued treatment early, one each due to: transient ischaemic attack, liver transplant, and combined liver and heart transplant.

Treatment with tafamidis for 12 months resulted in stabilisation in 95% of patients by week 6 and stabilisation in all patients at months 6 and 12. Changes from baseline at 6 and 12 months for the efficacy endpoints are presented in Table 12. There was a mean deterioration in the NIS-LL of 2.7 points at 12 months compared to baseline and similarly there was a small deterioration on the NIS and NIS-UL. The MS states that this indicates minimal progression of neuropathy at 12 months. Modified BMI showed a decrease at month 6 (suggesting worsening symptoms) but an increase at month 12 (suggesting improvement). At 12 months there was minimal change in quality of life and large nerve fibre function.

When comparing the pre-treatment and post-treatment periods a statistically significant slowing in disease progression was seen for NIS (p=[redacted]) and [redacted] ([redacted]), but the difference was [redacted] for large nerve fibre function (p=[redacted]); further details were presented in the MS. Results for NIS-LL and NIS-UL were not presented (in neither the MS nor the clinical study report).

Table 12: Efficacy results for the pre-post study

Endpoint	Mean change from baseline at 6 months (95% CI)	Mean change from baseline at 12 months (95% CI)
TQoL	-4.3 (-10.7 to 2.1)	0.1 (-8.9 to 9.0)
NIS (range 0 to 244)	2.0 (-2.9 to 6.9)	5.3 (-1.0 to 11.5)
NIS-LL (range 0 to 88)	-0.5 (-3.2 to 2.3)	2.7 (-0.4 to 5.8)
NIS-UL (range 0 to 156)	2.0 (-0.8 to 4.9)	2.5 (-1.2 to 6.2)
mBMI	-22.4 (-62.0 to 17.2)	16.6 (-31.0 to 64.2)
Large nerve fibre function*	0.6 (-0.7 to 1.9)	0.2 (-1.5 to 1.8)

*based summated scores of five tests, SD=standard deviation

Safety

Most patients (81%) experienced a TEAE. The most common were falls (24%), diarrhoea (24%), pain in extremity (19%), dizziness (14%), dysponea (14%), vomiting (14%), and constipation (14%). Four SAEs, in three patients (14%), were considered to possibly be related to study treatment by the investigator: ankle fracture, malaise, urinary retention, and transient ischaemic attack.

Study quality assessment

Although this study was conducted in a non-V30M population, which is more relevant to an English setting, the manufacturer's submission did not present any consideration of the limitations of the design used (in which patients acted as their own control). Study quality was therefore assessed. The results are presented in Table 13. Within the limits of this design, the study fulfilled most of the checklist criteria, although fewer than 90% of patients were followed up for 12 months, and it was unclear whether patients were recruited consecutively.

Table 13: Checklist for the quality assessment for the pre-post study (Fx1A-201)

Criterion	ERG judgement
Were selection/eligibility criteria adequately reported?	Yes
Was the selected population representative of that seen in normal practice?	Yes
Was an appropriate measure of variability reported?	Yes
Was loss to follow-up reported or explained?	Yes
Were at least 90% of those included at baseline followed up?	No
Were patients recruited prospectively?	Yes
Were patients recruited consecutively?	Unclear
Did the study report relevant prognostic factors?	Yes

The use of this design was understandable, considering the difficulties in recruiting sufficient numbers of patients for a RCT, but it nevertheless means it is not possible to confidently ascribe any benefits directly to tafamidis treatment. The primary endpoint for this study was TTR stabilisation, which is not a clinically relevant endpoint for evaluating effectiveness: in the main trial (Fx-005) 98% of tafamidis patients had TTR stabilisation at both 12 and 18 months i.e. even patients who were classified as non-responders on the NIS-LL. Also (as for the RCT extension study) comparisons were made at different time points within the single treatment group, meaning the impact of temporal effects unrelated to treatment cannot be ruled out.

4.6.2 Case series

Two small case series, both recently published as conference abstracts, were identified by the ERG (see Table 3). One was a study of nine V30M French patients with moderately severe disease (mean NIS-LL score of 42.5, mean Karnofsky score of 70%); the mean age was 71 years. It seems likely that there was no overlap of patients between this study and the studies reported in the submission, since patients in this study were receiving tafamidis as part of the French Early Access programme. The ERG did make attempts to contact the first author to confirm this, but were unable to obtain a response. Brief results were reported after 6 months of follow up: four patients had deteriorated in terms of NIS-LL (mean loss of 12.5 points) and three had remained stable (< 2 point worsening). Two patients stopped treatment due to severe urinary tract infections. The authors concluded that tafamidis had limited use in patients with moderately severe V30M TTR-FAP.[28]

The other study reported preliminary data at three months of follow up for seven non-V30M Italian patients (six of which had the GL89 mutation). From the mutations studied it was clear that there was no overlap of patients between this study and the studies reported in the submission. Although Norfolk QoL and NIS were assessed, no results for these outcomes were reported. The authors reported that tafamidis was well tolerated and no adverse effects had been reported. Two patients were reported to have experienced a moderate increase in weight with a substantial improvement in gastrointestinal improvement.[29]

Study quality assessment

These studies were only available as abstracts and it was not possible to undertake a formal quality assessment. Overall, they are of limited usefulness in this assessment due to the very small sample sizes, the absence of control groups, and the short treatment durations.

4.7 Ongoing studies

The MS listed two ongoing studies. One study (B3461029), which was set up as a requirement of the EMA marketing authorisation, aims to compare disease progression in symptomatic TTR-FAP patients with non-V30M mutations over at least 12 months of pre-treatment with current standard of care with a following 12 month period of tafamidis treatment. Eligible patients have to have enrolled on the THAOS registry, and yearly updates are to be provided to the EMA. The THAOS registry is funded by Pfizer, and is a multi-centre, longitudinal, observational survey which collects data on the natural history and treatment of transthyretin amyloidosis. It was established in 2007 and comprises data on patients from 14 countries. Data are currently available for 827 patients, of which 541 have symptomatic mutations (patients with wild-type transthyretin are also included in the survey).[48]

A second study (Fx1A-303) aims to provide long-term safety and efficacy data for patients who have successfully completed the RCT extension study (Fx-006) or the pre-post study (Fx1A-201). Patients will take tafamidis and will be assessed at 6-monthly follow up appointments.

The ERG identified a further ongoing single-group study in Japan which primarily aims to assess the effect of tafamidis on TTR stabilisation (the primary outcome) in patients with V30M or non-V30M mutations. Secondary outcomes relate to measures of efficacy. It was first received on clinicaltrials.gov on 13th September 2011, and has the following identifier: NCT01435655.

4.8 Generalisability of the study results

The key evidence for tafamidis was from an RCT conducted in a TTR-FAP population with the V30M mutation. Based on the baseline characteristics from the trial, on average the patients had early stage disease: the median duration of symptoms was 28 months in the tafamidis group and 21 months in the placebo group; and the mean score of both groups on the NIS-LL was on the lower end of severity on the possible score range of 0 to 88 (see Table 4). Most patients were 65 years or under, and the mean age of onset was mid-thirties. Therefore the

results are likely to be most applicable to patients with early-onset TTR-FAP related to a V30M mutation. There may be clinical variability within the trial population. 58% of patients were from a single centre in Portugal, an endemic area for V30M. Additional analyses undertaken by FDA (Section 4.4.6) suggested that the majority of responders to treatment (based on NIS-LL) were from this centre and suggest the possibility that people with V30M from endemic and non-endemic areas may respond differently to treatment.[47] However, this can only be considered an exploratory analysis.

The key issue in the UK context is that the best available evidence is for a V30M population which is not a common TTR-FAP mutation in the UK population. Disease resulting from non-V30M mutations presents with varying symptoms and co-morbidities, such as early cardiac involvement, often not seen in V30M patients. Patients with non-V30M mutations can have different rates of progression: they generally have a later disease onset and faster rates of overall disease progression, when compared to a V30M population. Therefore, there is uncertainty as to whether the results from the trial are applicable to a non-V30M population.

The MS included a single small pre-post study of non-V30 patients who received tafamidis. The patients in this study were more likely to be generalisable to the UK population than the trial population: they were older than those in the trial, had a much older age of onset and had a range of different genetic mutations including T60A (see Table 4). However, this was a pre-post study and the limitations of this study design mean the findings cannot be considered as robust.

4.9 Summary

The evaluation of the clinical effectiveness of tafamidis for TTR-FAP was primarily based on a single RCT (Fx-005) of 18 months' duration, with a one year open-label extension (Fx-006) in patients with the V30M mutation. Supplementary evidence was provided from a small single-arm pre-post study of 12 months duration in a non-V30M population. Additionally, the ERG identified two small case series published recently as conference abstracts.

The population in the RCT was not representative of an English population, in which the V30M mutation is rare. In the primary analysis of the intention-to-treat (ITT) population at 18 months, for TQoL the tafamidis group scores deteriorated by a mean of 2.0 points versus a placebo deterioration of 7.2 points. The difference was -5.2 points in favour of tafamidis with a 95% confidence interval (CI) from -11.8 to 1.3 points; this was not statistically significant (p=0.12). For NIS-LL the proportion of tafamidis responders was 45%, versus 30% for placebo (difference =15%, 95% CI -1.15 to 32.0); this was not statistically significant (p=0.07). The

magnitude and direction of the effect in favour of tafamidis was similar for TQoL and NIS-LL response in the efficacy evaluable population (patients with co-primary endpoint scores at 18 months who took 80% of their medication) and the differences were statistically significant in this analyses. Tafamidis appeared to be generally well-tolerated: tafamidis was associated with a greater incidence of urinary tract infections, and placebo was associated with a greater incidence of headache and neuralgia.

Although the risk of bias affecting the RCT results was assessed as being low, baseline differences (in TQoL, NIS-LL, and duration of symptoms) between the placebo and tafamidis groups may have contributed to the difference observed between groups at 18 months for NIS-LL outcomes, and for some analyses for TQoL where baseline score was not incorporated into the analysis as a covariate. The MS describes the baseline imbalances as not statistically significant and does not consider them any further. The ERG identified an analysis in the main clinical trial report where NIS-LL categorical response at 18 months was modelled as a function of treatment group, gender, age, duration of symptoms, and baseline NIS-LL score using logistic regression; gender (p=0.0033) and baseline NIS-LL (p=0.0112) were significant covariate predictors of NIS-LL response at Month 18 in the ITT population.

An extension study to the RCT appeared to suggest sustainability of the treatment effect of tafamidis over 30 months, a superiority of tafamidis treatment when comparing the results of placebo patients with the same patients who later received tafamidis, but a short-term and attenuating effect of tafamidis on mBMI. However, comparisons were made across time points within a randomised group and the use of this type of study design means the possibility that results may have been influenced by natural disease progression (temporal) effects cannot be ruled out. A small pre-post study was conducted in a population with a non-V30M population (which is more applicable to an English population) but the absence of a control group, and the small sample size, means the outcomes observed cannot be confidently attributed to tafamidis.

The manufacturer proposed that the similar reduced rates of progression with tafamidis in non-V30M patients (in the pre-post study) and V30M patients (in the trial) provides a rationale for extrapolating results of the trial to a non-V30M population. Notwithstanding the limitations of the pre-post study, further analyses of the RCT data by an FDA statistician, indicated there was uncertainty about the relevance of the trial results to the whole V30M population, with variation in response possibly related to patients' origin of mutation (their endemic or non-endemic status). The plausibility of extrapolating results to a non-V30M population therefore appears questionable.

5 ECONOMIC EVALUATION

5.1. Overview of manufacturer's economic evaluation

This section of the ERG report focuses on the economic evaluation submitted by the manufacturer as well as their responses to the ERGs points of clarification relating to the economic evaluation. The economic evaluation is subject to a critical review on the basis of the MS, the responses to points of clarification and by examination of the electronic version of the model. The critical review is conducted with the aid of a checklist [49] to assess the quality of the economic evaluations and a narrative highlighting the key assumptions and possible limitations. Section 6 presents a description of the additional work undertaken by the ERG to address any remaining uncertainties.

A summary of the manufacturer's economic evaluation and signposts to the relevant section in their submission are reported in Table 14.

Table 14: Summary of the economic evaluation submitted by the manufacturer

	Details	Source/justification	Signpost
Perspective	NHS & PSS and productivity costs accrued by patients and their carers.	The productivity costs represent the income forgone by patients and carers as a result of the disease.	Section 5.8.1, p103-4. Appendix B, Section 8.3, p148
Model	Cost-utility analysis using a decision analytic model and individual patient simulation.	The model extrapolates TQoL scores, health-related quality of life (HRQoL) and costs over the patient's lifetime.	Section 5.8.1, p103-4. Appendix B, Section 1.1, p131-132.
States and events	The model includes one event, liver transplant, and 5 health states: • Disease stage 1 – patient does not require assistance with ambulation. • Disease stage 2 – patient requires caregiver assistance or a walking aid. • Disease stage 3 – patient is bed-bound. • Post-liver transplant disease stage 1 • Death.	As there are no quantitative measures of TTR-FAP disease progression, the manufacturer used the disease stages defined by Coutinho et al (1980), which classify the progression of the disease according to the extent of motor neuropathy in V30M population.[7] A patients stage and their progression between A patients stage and their progression between stages depends on the patient's TQoL score. The marketing authorisation specifies that tafamidis is indicated for the treatment of TTR-FAP in adult patients with stage 1 symptomatic polyneuropathy.[23]	Section 5.8.1., p103. Appendix B, Section 1.2, p132-3.
Intervention and comparators	Tafamidis (20mg once daily) in addition to conventional support therapy is compared with conventional support therapy alone in patients with TTR-FAP stage 1 and symptomatic neuropathy. Tafamidis is discontinued once patients progress to stage 2.	In line with the marketing authorisation for tafamidis and the scope issued by AGNSS. Tafamidis is the only licensed treatment for TTR-FAP. Conventional support therapy includes liver transplant.	Section 5.8, p103. Appendix B, Section 1.5, p133 and Section 3 p136.
Patient	The base-case population consists of a combined cohort of male and	In line with the marketing authorisation for tafamidis and the	Appendix B,

population and Subgroups	female adult patients with symptomatic TTR-FAP caused by any type of TTR mutation, both V30M and non-V30M. Each patient is assigned a specific TQoL score and age at baseline, based on the distribution of ages and TQoL scores of patients in the THAOS registry. Two subgroups are considered: • Patient with V30M mutation. • Patients with mutations other than V30M. The two subgroups differ in: • Age – the average age at diagnosis is the same (61.6 years) but the standard deviation differs. • Baseline TQoL score (TQoL score at treatment initiation) • Mortality risk pre-liver transplantation • Mortality risk post-transplantation • Probability of liver transplantation.	scope issued by AGNSS.	Section 2, p135.
Natural history	The model tracks the progression of TQoL scores over time. Greater TQoL scores indicate lower HRQoL and more severe disease. The rate of TQoL progression differs by disease stage. The disease stages were defined according to TQoL cut-offs. The TQoL score at which a patient transitions from stage n to n+1 was calculated to be the midpoint of the 90th percentile for patients classified in stage N and the 10th percentile for patients classified in Stage N+1. Patients in stage 1 are eligible for liver transplant, regardless of disease genotype. Liver transplant is assumed to halt disease progression and improve survival. Mortality pre-liver transplant depends on disease duration from time at symptom onset. Mortality post-liver transplant depends on time	TQoL, via Norfolk QoL-DN, was collected during Fx-005, a multicentre double-blind RCT, and on the observational cross-sectional study Fx1A-OS-001. The TQoL cut-offs were based on data from the THAOS registry and the definitions proposed by Coutinho et al (1980).[7] The rate of change in TQoL score is based on the relationship between TQoL, disease duration and disease stage, using data collected during Fx1A-OS-001. The probability of liver transplant is based on expert advice from NAC and differs according to mutation (V30M and non-V30M). The mortality risk in the model was informed by fitting parametric survival curves to published Kaplan-Meier curves. Mortality pre-	Section 5.8, p103. Appendix B, Section 1.3-1.4, p134-144

	from liver transplantation and age.	liver transplant is obtained from the study by Sattianayagam et al (2012), which presents Kaplan-Meiers curves for T60A and V30M patients.[3] Mortality post-liver transplant is based on the study by Herlenius et al (2004), which presents Kaplan-Meier curves for V30M and non-V30M patients.[22] Mortality for the base-case population is obtained from combining the Kaplan-Meier curves for V30M and non-V30M populations.	
Treatment effectiveness	Tafamidis reduces the rate of change in TQoL and therefore slows down disease progression. Tafamidis is only given during stage 1.	The hazard ratio between tafamidis and placebo is the ratio between the rate of change in TQoL in the tafamidis group and the placebo group observed between baseline and month 18 in Fx-005.	Appendix B, Section 6.1.1, p140.
Health-related quality of life	Quality-adjusted life years (QALYs) were calculated using EQ-5D scores. EQ-5D scores were derived from the TQoL using a mapping function. As the patient progresses through the model, TQoL changes and so does its HRQoL. Liver transplant is associated with one-off QALY decrement. A QALY loss of 0.01 is included in stage 3 to reflect the impact on carers.	The relationship between TQoL and EQ-5D scores was obtained by mapping TQoL to EQ-5D using data from the THAOS registry. QALY decrement as a result of liver transplant based on the study by Ratcliffe et al (2002).[50] HRQoL loss due to impact on caregivers based on the NICE final appraisal for the treatment of Alzheimer patients.[51]	Appendix B, Section 7, p145-146.
Adverse events	Not considered.	The manufacturer justified the exclusion of adverse effects from the model with their mild to moderate impact on health and therefore negligible impact on HRQoL.	Appendix B, Section 7.6, p146.
Resource utilisation and costs	The costs include: • Acquisition cost of tafamidis, excluding VAT. • One-off costs as a result of healthcare resource use as a result of transition to stage 2 and stage 3. • Costs of liver transplantation • Recurrent costs as a result of healthcare resource use during stage 1, 2, 3, and medical management post-liver	Tafamidis 20mg once daily costs [redacted] per patient per year. The acquisition cost excludes VAT due to the home delivery arrangements proposed by the manufacturer. Costs are estimated by applying UK unit costs to resource use estimates obtained from Swedish physicians.	Appendix B, Section 8, p147-148. Appendix B, Section 13, p154-169.

	transplantation. • Productivity costs, which are also stage dependent.		
Discount rates	A 3.5% discount rate was applied to both costs and health consequences.	As defined in the NICE reference case and UK Treasury guidance .[52]	Appendix B, Section 5, p137.
Sensitivity analysis	Sensitivity analyses on: • Baseline TQoL • Baseline age • Rate of liver transplantation • Reduction in disease progression due to tafamidis • Costs	No justification was presented for the parameters and alternative ranges tested in the sensitivity analyses.	Appendix B, Section10.1, p150-151.

5.1.1. Review of existing literature

The manufacturer conducted a systematic literature search to identify any published economic evaluations in TTR-FAP. The search strategy was described in Appendix A of the manufacturer's submission. The search strategy was checked by the ERG and found to be appropriate. No published economic evaluations were identified in TTR-FAP.

5.1.2. Reports of other Health Technology Assessment agencies on tafamidis

Tafamidis has been the subject of appraisals by other Health Technology Assessment agencies, namely the All Wales Medicine Strategy Group (AWMSG, for Wales), the Institute for Quality and Efficiency in Health Care (IQWIG, for Germany) and the National Authority of Medicines and Health Products (Infarmed, for Portugal). The AWMSG was unable to endorse tafamidis for use within the NHS Wales due to the absence of submission from the manufacturer.[53] The IGWIG concluded that the clinical evidence suggested some positive effect of tafamidis on neurological degeneration and therefore recommended tafamidis for use in Germany.[54] The Infarmed assessed the clinical effectiveness, cost-effectiveness and budget impact of tafamidis for the Portuguese National Health Service.[55] The assessment concluded that despite the limited evidence on effectiveness, tafamidis fulfils an unmet need and delays disease progression. The cost-effectiveness results and budget impact estimates were considered acceptable given the characteristics of tafamidis and of TTR-FAP. However, the price agreed with the manufacturer was not disclosed.

5.1.3. Manufacturer's 'de novo' economic evaluation

In the absence of any previously published economic evaluation of tafamidis, the manufacturer's de novo economic evaluation forms the basis of the economic evidence submitted to AGNSS. The *de novo* economic evaluation compared the lifetime costs and health outcomes of tafamidis as an add-on therapy to conventional support therapy with conventional support therapy alone in patients with TTR-FAP stage 1. Patients in both arms could receive liver transplant, although the economic evaluation assumes that only stage 1 patients are eligible. The model evaluated costs from a societal perspective, expressed in UK pound sterling at a 2010 price base. Outcomes in the model were expressed in terms of quality-adjusted life years (QALYs). A QALY is a generic measure of health which combines the length of life with the health-related quality of life (HRQoL) experienced during the period of time considered. Both costs and health outcomes were discounted at a rate of 3.5% per annum, as recommended by UK guidance).[52]

Results were presented in terms of mean costs and QALYs, incremental costs and QALYs of tafamidis compared with conventional support therapy and incremental cost-effectiveness ratios. The base-case population consists of a combined population including both V30M and non-V30M patients, however the exact proportion of each was not explicit. Subgroup analyses were presented for two subgroup populations, patients with the V30M mutation and patients with mutations other than V30M. This is in line with the scope defined by AGNSS.

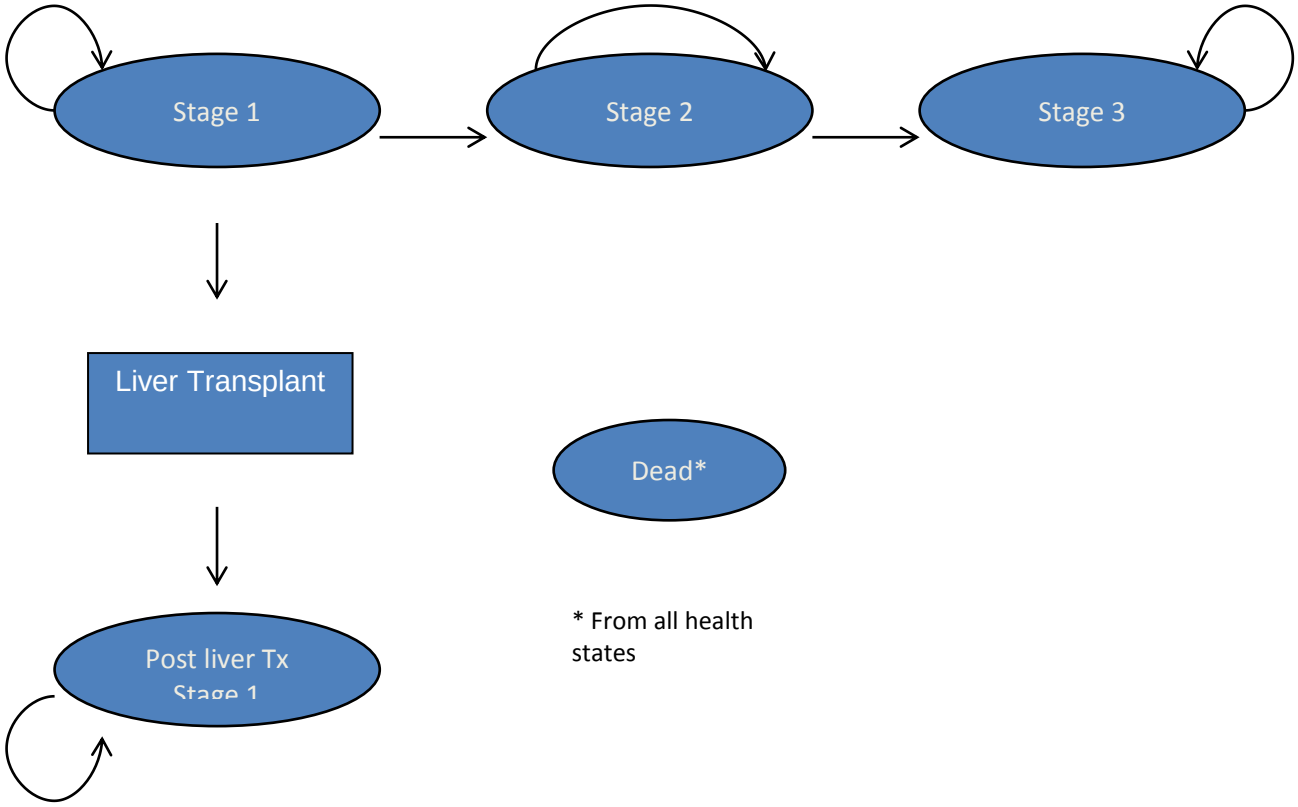
The critical review is complemented with a checklist designed to assess the quality of economic evaluations, presented in Appendix 5.

Model structure

The economic model is a decision analytic model using patient level simulation of 20,000 hypothetical patients, in which each simulation represents an individual patient. At model entry, each patient is assigned a TQoL score and age, drawn from the distribution of TQoL scores and ages for the specific population (combined V30M and non-V30M in the base case and V30M and non-V30M separately for the subgroup analyses). The cycle length is 6 months. As time progresses, the patient is at risk of transitioning to another health state or remaining in their current state.

The model includes one event, liver transplantation, and 5 health states: (i) disease stage 1, where the patient does not require assistance with ambulation, (ii) disease stage 2, where the patient requires carer assistance or a walking aid, (iii) disease stage 3, where the patient is bed-bound, (iv) post-liver transplantation disease stage 1 and (v) death. Figure 1 depicts the structure employed in the model. The circles represent the health states. The box indicates the event (liver transplantation), which leads to the post-liver transplantation health state. The arrows represent the possible movements (transitions) between health states in any given cycle.

Figure 1: Structure of Markov model (adapted from Figure 25 p132 of MS)



The model tracks the patient's TQoL score, which relates to disease severity. Greater TQoL scores indicate lower HRQoL and more severe disease. The patient's TQoL score increases at a rate conditional on disease stage. The TQoL rate of change is greater in stage 1 than stage 2, and greater in stage 2 than stage 3. Patients progress to higher disease stages once their TQoL scores reaches defined cut-off values, at which point their TQoL rate of change decreases to the higher stage's TQoL rate of change. Patients in stage 1 are eligible for liver transplantation, regardless of disease genotype. Liver transplantation is assumed to halt disease progression and improve survival.

The model captures costs and QALYs over the patient's lifetime. Each health state is associated with recurrent costs, which are a result of healthcare resource use and productivity losses due to the disease. Transitions to stage 2 and to stage 3 are associated with one-off costs as a result of healthcare resource use. Liver transplantation is also associated with a one-off cost. HRQoL was captured via EQ-5D, which was calculated from the patient's TQoL score using a mapping function. The mapping function between EQ-5D and TQoL scores was obtained using data from the THAOS registry.

Key assumptions

A brief overview of the key assumptions used in the manufacturer's economic evaluation is reported below. This is followed by a more detailed critique of the economic evaluation and its assumptions:

- The natural history of patients with TTR-FAP is modelled through the TQoL score for disease severity, whilst mortality is modelled separately. Therefore, TQoL is assumed to be an appropriate measure of disease severity and progression for both V30M and non-V30M populations.
- The patient's TQoL score defines their disease stage (with the patient transitioning to the next stage of their disease once they reach particular TQoL cut-off scores) and the patient's HRQoL, whilst the TQoL rate of change is dependent on their disease stage.
- Coutinho et al (1980)[7] disease stages are a suitable classification of disease status for both V30M and non-V30M populations.
- The cut-off TQoL scores for each disease stage were based on data on TQoL scores and disability level from the THAOS registry. Patients were classified into disease stages by converting their score in the modified Polyneuropathy Disability Scale (mPDS) to Coutinho disease stages, although no evidence was provided to support the mapping of mPDS onto disease stages. The TQoL score at which a patient transitions from stage n to n+1 was calculated to be the midpoint of the 90th percentile for patients classified in stage N and the 10th percentile for patients classified in Stage N+1.
- The TQoL rate of change over time is based on the relationship between TQoL and disease duration as reported in Fx-01-OS-001, a cross-sectional observational single centre study outside the UK. It was assumed that the TQoL rate of change was constant for each disease stage, this was calculated using the function relating TQoL to disease duration and the cut-off scores for disease stages.
- Mortality is solely dependent on time in the model and is independent of disease severity.
- Patients who undergo liver transplantation are assumed to experience no further disease progression (i.e. it maintains their HRQoL at the time of transplant throughout their lifetime) and to have improved survival. In addition, liver transplantation is associated with a one-off QALY decrement in the first cycle following the procedure.

- All patients in disease stage 1 are eligible for liver transplantation. The rate of liver transplantation was informed by expert clinical advice at NAC and differs by subgroup.
- Mortality post-liver transplantation depends on time from liver transplantation and on the patient's age.
- All patients in disease stage 1 are eligible for tafamidis treatment in accordance with its license. Tafamidis is discontinued once patients reach stage 2.

5.2. Patient population

At model entry, the patient is assigned a baseline TQoL score and age drawn from a distribution of baseline TQoL score and age specific to the base-case (combined V30M and non-V30M), V30M and non-V30M populations. The relative proportion of V30M and non-V30M patients in the base-case population was not explicitly presented in the manufacturer's submission. However, based on data from the MS, the ERG calculated that one sixth of patients have the V30M variant and the remaining 83% have a non-V30M variant.

Table 15 presents the data used to inform the distribution of age and baseline TQoL scores used in the model.

Table 15: Age at model entry and baseline TQoL score (adapted from Table 39 p136 of MS)

Population	Age (SD)	Baseline TQoL score (SD)
Base case (combined V30M and non-V30M)	63 (16.31)	48.97 (24.91)
Source	Lane et al. (2011).[56]	THAOS registry
Subgroup 1 (V30M)	61.6 (6.89)	49.64 (25.62)
Source	THAOS registry and Sattianayagam et al (2012) [3]	THAOS registry
Subgroup 2 (non-V30M)	61.6 (9.63)	44.89 (20.87)
Source	Fx1A-201	THAOS registry

For the base-case (combined V30M and non-V30M) population, the baseline age was obtained from the study by Lane et al [56] and the baseline TQoL was obtained from the distribution of TQoL scores in the THAOS registry in those classified as stage 1 (the classification of patients into stages is discussed in more detail later) . Lane et al investigated the characteristics of the patients with the diagnosis of ATTR registered in the NAC database.[56] Of the patients registered in NAC, 292 patients were found to have hereditary ATTR caused by variant TTR, of which 78 (26%) had the V30M variant, 75 (26%) the T60A variant, 76 (26%) the V122I variant and 63 (22%) one of 33 other mutations. The

median age at presentation was 63 years, inter-quartile range 49 to 71. Mean age at presentation or standard deviation were not reported. Therefore, the ERG has not been able to ascertain how the manufacturer obtained a mean age of 63 years with a standard deviation 16.31. For the TQoL score at baseline, the manufacturer has used the mean and standard deviation of TQoL scores of all patients in stage 1 recorded in the THAOS registry. The use of two sources to define the baseline characteristics in terms of age and TQoL of patients entering the model does not appear appropriate to the ERG as it is unclear if these sources are comparable. Whilst the information on age relates to an incident population, as it is based on the age at presentation, which the ERG has treated as akin to age at diagnosis, the information on TQoL score relates to a prevalent population, where it is possible patients may have been diagnosed many years previously.

For the V30M population, age was obtained from the distribution of ages in late onset V30M patients in the THAOS registry, whilst the baseline TQoL score refers to the distribution of TQoL from all V30M patients in stage 1 in the THAOS registry. Patients with late onset V30M disease were selected from the wider V30M population in order to better reflect the characteristics of the V30M patients in the UK. Since the THAOS registry records age at symptom onset rather than age at diagnosis, the 2 years median delay from symptom onset to diagnosis reported in Sattianayagam et al (2012) was added to estimate age at diagnosis.[3] The ERG was unable to confirm the data referring to late-onset V30M patients obtained from the THAOS registry. As with the base-case population, the ERG has concerns about the comparability of the sources of the two baseline parameters. The ERG considers it reasonable to use the age distribution of patients with the late onset V30M variant for the UK population, since it reflects an incident population. In contrast, TQoL score is based on all V30M patients in stage 1 in the THAOS registry, which not only represents a prevalent population but also includes non-late onset patients.

For the non-V30M population, the age was based on age at diagnosis of patients in Fx1A-201. As discussed in Appendix 2, Fx1A-201 was a phase II clinical trial in 21 patients. Although baseline TQoL scores were recorded in Fx1A-201 (mean=47.8, standard deviation=35.14), the manufacturer used TQoL scores recorded in the THAOS registry for all non-V30M patients in Stage 1. As with the base case and V30M population, the ERG has concerns about the comparability of the two sources used, not least because one reflects an incident population and the other a prevalent population.

In the ERG's view, it is not appropriate to consider a combined population given the difference in the characteristics of V30 and non-V30M. As discussed in Section 2.2.2, patients with the V30M variant typically present with small fibre neuropathy affecting

peripheral and autonomic nerves. Cardiac amyloidosis is relatively uncommon with this mutation in contrast with non-V30M mutations.[3] Therefore, the cost-effectiveness analysis should have been conducted separately for V30M and non-V30M patients. Results for the combined population should be calculated as the weighted average of the results for each population, weighted by their relative proportions in the UK. Furthermore, age and TQoL are also identifiable sources of heterogeneity in patients. Therefore, rather than considering the distributions around age and TQoL in their analyses, the manufacturer could have treated these parameters deterministically as different subgroups. Section 6 explores these issues in more detail.

5.3. Natural history

The natural history of patients with TTR-FAP is modelled through the TQoL score for disease severity, whilst mortality is modelled separately. The patient's TQoL score defines their disease stage (with the patient transitioning to the next stage of their disease once they reach particular TQoL cut-offs), and the patient's HRQoL, whilst their TQoL rate of change is dependent on their disease stage (i.e. once a patient reaches a TQoL cut-off, they transition to the next stage of the disease, and their TQoL rate of change decreases to that for the new stage). Mortality is solely dependent on time in the model and is independent of disease severity.

As mentioned previously, patients in stage 1 are considered eligible for a liver transplant. Patients who undergo liver transplantation are assumed to experience no further disease progression and to have improved survival. In this section we will discuss the modelling of disease severity and mortality separately for non-transplanted patients. We will then briefly describe the natural history of patients following liver transplantation.

5.3.1 Disease severity and TQoL scores

As previously stated, the patient's TQoL score defines their disease stage (with the patient transitioning to the next stage of their disease once they reach particular TQoL cut-offs), and the patient's HRQoL, whilst their TQoL rate of change is dependent on their disease stage. Below, we consider in turn, the derivation of the TQoL stage cut-off values, the TQoL rate of change and its relationship with disease stages.

TQoL cut-off for disease stages

Cut-offs TQoL scores for disease stages were defined using data from the THAOS registry. Since the THAOS registry does not record information on disease stage, the patients' scores on the modified Polyneuropathy Disability scale (mPDS) were used to calculate the Coutinho disease stages. The mPDS assesses walking ability. The conversion from mPDS score to the Coutinho disease stages assumed that score 1 and 2 corresponds to stage 1, score 3a and 3b corresponds to stage 2, and score 4 to stage 3 following the algorithm outlined in Table 16. No justification was provided for the method of mapping mPDS to Coutinho disease stage and the ERG has been unable to verify its validity. The use of Coutinho et al disease stages also assumes these are an appropriate method for characterising these patients, however, the ERG is uncertain of the generalisability and clinical utility of staging TTR-FAP using the Coutinho descriptors in a non-V30M population and indeed in a V30M population from a non endemic area (see Section 2.2.6)

Table 16: Conversion of the mPDS to the Coutinho et al (1980) disease stages (adapted from Table 3 of Points for Clarification and Coutinho et al (1980) [7]

Modified Polyneuropathy Disability Scale	Coutinho et al (1980) disease stages
0 = Normal.	**Note**. The Coutinho stages apply only when the patient is symptomatic.
1 = Sensory disturbances in feet but able to walk without difficulty.	**Stage 1** "… the disease is limited to the lower limbs and the patient is still walking without any help. On examination, the more common findings are a slight weakness of the extensors of the big toes, absent ankle jerks (sometimes with very brisk knee jerks) and some difficulty in standing on the heels. Pain and temperature sensations are extensively impaired, while light touch and joint position senses are still spared."
2 = Some difficulties in walking but can walk without aid.	
3a = Able to walk with one cane or crutch.	**Stage 2** "In stage 2, motor signs progress in the lower limbs with steppage and distal amyotrophies, while the muscles of the hands begin to be wasted and weak. Temperature and pain sensory impairment appear in the upper limbs and in the trunk and light touch loss begins to be evident in the feet and legs with a stocking distribution… The patient is by then obviously handicapped but can still move around, although needing help."
3b = Able to walk with 2 canes or crutches.	
4 = Not ambulatory; confined to wheelchair or bedridden.	**Stage 3** "Patient is bedridden or confined to a wheelchair and has generalised weakness, atrophies and areflexia. Temperature and pain are not felt all over the body except for the head and neck. The touch is diminished in a glove and stocking distribution."

Following conversion from mPDS to Coutinho stages, the manufacturer used the distribution of TQoL scores by disease stage of the patients in the THAOS registry to define TQoL cut-offs between stages. The TQoL score at which the patient transitions from stage N to stage N+1 was taken to be the midpoint between the 90th percentile of TQoL scores for stage N (i.e. the TQoL score at which 90% of patients classified as stage N have a TQoL score below this) and the 10th percentile for stage N+1 (i.e. the TQoL score at which 10% of patients classified as stage N+1 have a TQoL score below this value). The manufacturer's submission mistakenly states it is the midpoint between the upper 95% confidence interval value of stage N and the lower 95% confidence interval value of stage N+1.

Table 17 presents the TQoL scores per disease stage for the combined V30M and non-V30M population recorded in the THAOS registry. For this population, the average TQoL score in stage 1 is 48.97, in stage 2 is 72.68 and in stage 3 is 94.83. The width of the 10th percentile to 90th percentile range demonstrates the heterogeneity in TQoL scores within disease stages. In stage 1, the 10th percentile is 21 and the 90th percentile is 87; for stage 2, the 10th percentile is 21 and the 90th percentile is 103; and for stage 3, the 10th percentile is 79 and the 90th percentile is 107. Therefore, there are some patients classified as stage 1 that have markedly greater TQoL scores than some patients classified in stage 2, and similarly there are some patients in stage 2 that have markedly greater scores than some patients in stage 3. This raises issues about the appropriateness of using TQoL scores to define disease stages. Given the heterogeneity in TQoL scores over the disease stages, it is questionable whether TQoL is a valid marker of disease progression or if the Coutinho disease stages are a useful classification of disease severity.

Table 17: TQoL scores per stage recorded in the THAOS registry (adapted from Tables 41 p139 of MS and Appendix 12-15 of Points for Clarification)

Disease stage	Mean TQoL	P10	P90	Number of patients
Base-case: Combined Val30Met and non-Val30Met				
1	48.97	21.0	87.0	64 (86% V30M)
2	72.68	21.0	103.0	38 (79% V30M)
3	94.83	79.0	107.0	23 (91% V30M)

Table 18 presents the cut-offs TQoL scores between stages. The cut-off for crossing over to stage 2 from stage 1 is 54, just over 5 points above the average TQoL score of stage 1 patients (which is the baseline TQoL score used in the mode for the base case population). The cut-off for crossing over to stage 3 from stage 2 is 91, less than 4 points below the average TQoL score for stage 3 patients. No rationale was provided for this choice of calculating the cut-offs for each disease stage. As part of the additional work undertaken by the ERG alternative assumptions about TQoL cut-offs were considered (see Section 6).

Table 18: Threshold TQoL scores for cross-over between stages (adapted from Table 42 p139 of MS)

Disease stage	TQoL scores for cross-over between stages
Baseline	2.6
Stage 1 to Stage 2	54
Stage 2 to Stage 3	91
Maximum TQoL	135

There are a number of issues in the use of TQoL to define disease stages. Firstly, and as discussed in Appendix 3, there is uncertainty as to whether TQoL captures all aspects of HRQoL associated with TTR-FAP, namely the emotional/psychological components of the disease and other components where peripheral neuropathy is not the predominant symptom, such as the cardiomyopathy, which is a frequent symptom in non-V30M patients. Secondly, the statistics presented in Table 17 suggest that there is considerable heterogeneity in TQoL scores by disease stage. Thirdly, the clinicians contacted by the ERG considered that the Coutinho stages are not applicable to non-V30M patients, given that disease progression generally involves cardiac amyloidosis and does not follow the pattern of progressive polyneuropathy to which the Coutinho stages refer to and typical of V30M cases. Therefore, it is unclear whether TQoL is a valid marker of disease progression or if the Coutinho disease stages are a useful classification of disease severity. In addition, the definition of stage cut-offs as the midpoint between percentile 10 and 90 has not been justified and appears somewhat arbitrary to the ERG.

TQoL rate of change by disease stage

In the manufacturer's model, TQoL scores increase over time although at a decreasing rate as the patient progresses through the disease stages (with a different rate of change for each disease stage). The rate of change of TQoL score was estimated from the data collected in Fx1A-OS-001, which considered the relationship between disease duration and TQoL. The function was then combined with the TQoL cut-offs for each stage to calculate constant rates of change for each disease stage. It is unclear to the ERG why evidence on the rate of change of TQoL from the Fx-005 trial was not considered.

As discussed in Appendix 2, the Fx1A-OS-001 was an observational cross-sectional single centre study (in Oporto, Portugal) in V30M patients with stage 1 to 3 of TTR-FAP and healthy volunteers. The aim of Fx1A-OS-001 was to evaluate the relationship between the clinical measures assessed by NIS-LL and TQoL scores and the Coutinho disease stages. Fx1A-OS-001 recorded current TQoL score and disease duration, from date at symptom onset as recalled by patients to date of enrolment into the study. This information was used

to capture the relationship between symptom severity, as measured by TQoL, and disease duration.

TQoL score and disease duration were related using the polynomial regression, which is presented in Equation 1. The dependent variable is disease duration and TQoL score is the explanatory variable. However, in the ERG's view this is an inappropriate way to model the relationship, as it would be more appropriate to have TQoL as the dependent variable (as causality would suggest that it is disease duration which determines severity, and not severity which determines disease duration). In addition, the ERG does not consider it appropriate to use cross-sectional data set to determine the relationship between TQoL and disease duration, as strong assumptions need to be made about the homogeneity of the patients included in the study, and given the heterogeneity observed in TTR-FAP patients with regards to disease progression, it is unlikely such assumptions would hold. However, the ERG understands that no other data may be available to inform the relationship between disease severity and duration in this disease, although evidence was available at least for stage one patients from the Fx-005 trial. Section 6 explores this issue in more detail.

Equation 1: Relationship between disease duration and TQoL scores

[redacted]

Equation 1 allowed the calculation of disease duration at each TQoL cut-off for diseases stage. This in turn allows for calculation of the time spent in each disease stage according to equation 1. Table 19 presents the time spent in each stage calculated from Equation 1. The time spent in stage 1 was estimated at 3.58 years, because the time to move to a TQoL score of 54 from the minimum TQoL score of 2.6 (the cut-off to move from stage 1 to stage 2) is 3.58 years. The time spent in stage 2 was estimated at 4.92 years, because that is the time to move from a TQoL score of 54 to a TQoL score of 91 (the threshold to move to stage 3). The manufacturer did not report the time spent in stage 3, however, using the same method and the maximum TQoL score of 135, they have estimate the time until they reach the maximum TQoL score of 135, and used this to calculate the rate of TQoL progression in stage 3. The clinicians contacted by the ERG felt that patients would experience more time in stage 1 than suggested by the manufacturer and around the same amount of time in stage 2.

Table 19: Time spent in each disease stage according to the manufacturer's submission (adapted from Tables 43, p140 of MS)

Stage	Time spent in stage (years)
1	3.58
2	4.92
3	[redacted]

*Estimated from equation relating TQoL score and disease duration presented in the manufacturer's submission

The manufacturer then assumed that the TQoL rate of change was constant for each stage. Table 20 presents the 6-monthly rate of change in TQoL score used in the model and Figure 2 shows how the manufacturer estimated TQoL rates of change. In Figure 2, the bold line represents the relationship between disease duration and TQoL estimated from Equation 1, while the dashed line shows the TQoL rates of change for each disease stage. The TQoL rate of change for each stage corresponds to the linear slope of the line between disease duration at each stage cut-off, i.e. the ratio of the difference between the two stages TQoL cut-off scores and the time spent in each stage. For example, for stage 2 the difference between the upper and lower bounds of the TQoL cut-offs which define stage 2 is 37 (91-54) and the time spent in stage 2 is 4.92 years. Therefore, the TQoL rate of change for stage 2 is 3.761 TQoL units per 6-months.

Table 20: TQoL rate of change over 6-months applied in the economic model (adapted from Table 44, p140 of MS)

Stage	6-monthly rate of change in TQoL score
1	7.170
2	3.761
3	1.022

Figure 2: Relationship between TQoL score and disease duration

It is worth noting that the time spent in stage 1 as calculated by Equation 1 is not the time that patients spend in stage 1 in the economic model. Patients enter the model with a baseline TQoL score of 48.97, just 5 points less than the cut-off to stage 1. Applying Equation 1 to the baseline TQoL score at model entry indicates that patients enter the model are have already had disease duration of 3.3 years. Since the TQoL rate of change is 7.17 units per 6 months and the TQoL cut-off score to crossover to stage 2 is 54, the economic model estimates that on average patients diagnosed with the condition spend less than 5 months in stage 1. This reflects the confusion between incident and prevalent population discussed earlier (Section 5.2). The baseline TQoL score at model entry is taken from a prevalent population, who may have suffered from the disease for a number of years. Therefore, the cost-effectiveness results, ignoring age for simplicity, refer to a prevalent population who initiate tafamidis with a disease severity close to stage 2.

It also should be noted that the TQoL rate of change observed in Fx-005 was little more than a third of the rate of change applied in the economic model for stage 1. For the placebo group, the monthly rate of change of 0.4618 corresponds to a 6-monthly rate of change of 2.77, while the economic model uses a 6-monthly rate of change of 7.17. If the 6-monthly

rate of change in TQoL from Fx-005 is a more accurate reflection of the disease progression, untreated patients spend a longer time in stage 1 than predicted by the model, whilst the impact of tafamidis treatment is also biased, with a larger absolute reduction in the rate of change of TQoL in the model than would be observed in practice. Therefore, the economic model may overestimate the benefits of tafamidis. It is unclear to the ERG why evidence on the rate of change of TQoL from the Fx-005 trial was not considered, although the manufacturer states that this is because the Fx-005 trial was in patients with early onset which is not comparable with the late onset population of patients in England.

Since both the time spent in each stage (conditional on baseline TQoL score) and the TQoL cut-off scores between stages are the same for V30M and non-V30M patients, the model implicitly assumes that disease progression, in terms of TQoL, is identical for V30M and non-V30M patients. However, the published literature [3] and clinical advice contacted by the ERG strongly suggest that, although the progression of polyneuropathy is probably similar, prognosis is poorer than for V30M. Furthermore, the typical UK patient has the T60A variant, in which peripheral neuropathy is a minor component of the disease. Therefore, it is unclear to the ERG how accurately TQoL scores in general and the rate of change in TQoL scores in particular are able to reflect disease progression and severity in the patient population in England.

5.3.2 Mortality

Mortality in the economic model is assumed independent of disease severity measured by TQoL. Instead mortality is only dependent on the amount of time spent in the model (with the exception of those who undergo liver transplantation, which is discussed later).

Mortality without liver transplant is based on the study by Sattianayagam et al (2012).[3] Sattianayagam et al (2012) report the clinical presentation, histological findings, cardiac status, and clinical outcomes of all 60 patients with TTR T60A who were diagnosed and prospectively followed up at the UNAC and at the University of Western Ontario, Canada, from 1992 to 2009. Kaplan-Meier curves are presented for a cohort of the 52 non-transplanted T60A patients, with a median age of 63 at symptom onset, and for a V30M patient cohort of 26 Swedish non-transplanted patients on whom no information is available on age. The manufacturer digitalised the Kaplan-Meier curves for T60A and for V30M patients reported by the Sattianayagam et al (2011) to estimate the parameters for Weibull survival functions. The data from each Kaplan-Meier curve were then combined to obtain an overall Kaplan-Meier curve for a mixed population of V30M and T60A patients.

Whilst the Sattianayagam et al paper presents survival curves separately based on time since symptom onset and on time since diagnosis, the manufacturer has based their analyses on the survival curves on time from symptom onset. Given the age of patients entering the model is based on the age at diagnosis (see Section 5.2), the ERG considers that it would have been more appropriate to base their survival analyses on the survival curves based on time since diagnosis, particularly given that patients entering the model are eligible for tafamidis immediately, which would clearly not be appropriate before diagnosis. For the T60A patient cohort, median (95% CI) survival from onset of symptoms was 6.6 years (0.2-14.0) and from diagnosis was 3.4 years (2.7-5.3). For the Swedish V30M cohort, median survival from symptom onset was 12.0 years (5.9-20.2) and from diagnostic was 8.2 years (1.2-15.4). The impact on the comparative results of the choice of survival from symptom onset will bias the results in favour of tafamidis as it will result in greater survival. Therefore patients will benefit from the treatment effect in terms of lower TQoL for longer as survival is over estimated (treatment effect is discussed in more detail in Section 5.3).

The survival curve for the base-case (V30M and non-V30M population) refers to the combined survival of the 52 T60A patients and of the 26 V30M patients. Therefore the survival for the base-case population refers to a population of T60A and V30M patients at a ratio of 2 to 1. However, the base-case population in the model is composed of 16.66% of V30M and 83.33% of non-V30M patients (based on the manufacturer's estimate of the eligible population in England). Therefore, the survival for the base-case population refers to a considerably different population. Not only is the proportion of V30M to non-V30M markedly different from the UK population, but there are other causes for concern as well. It is unclear if the T60A patients are representative of all non V30M patients in England given the heterogeneous nature of the disease, although the age of patients in the Sattianayagam et al study is at least similar to that in the model. It is also unclear to the ERG whether the V30M patients in the study are representative of the V30M patients in England given Sweden has an endemic V30M population, whilst England has a non endemic V30M population. Unfortunately, no information on the characteristics of the V30M population in the Sattianayagam study was provided, therefore the ERG has been unable to assess the comparability between those patients and those in the model. As stated previously, the ERG does not consider it appropriate to consider a combined population of V30M and non V30M given the heterogeneity in disease progression between the two groups.

For the subgroup analyses, the survival curves were used separately for V30M and non V30M. However, as with the base case, the manufacturer has, in the ERG's view, incorrectly used the survival curves based on time from symptom onset rather than time from diagnosis.

As stated previously it is also unclear to the ERG whether the V30M and non-V30M patients from the study are representative of those in England.

Table 21 presents the parameters for the Weibull survival function used in the model. A shape parameter greater than 1 indicates that the mortality rate increases over time.

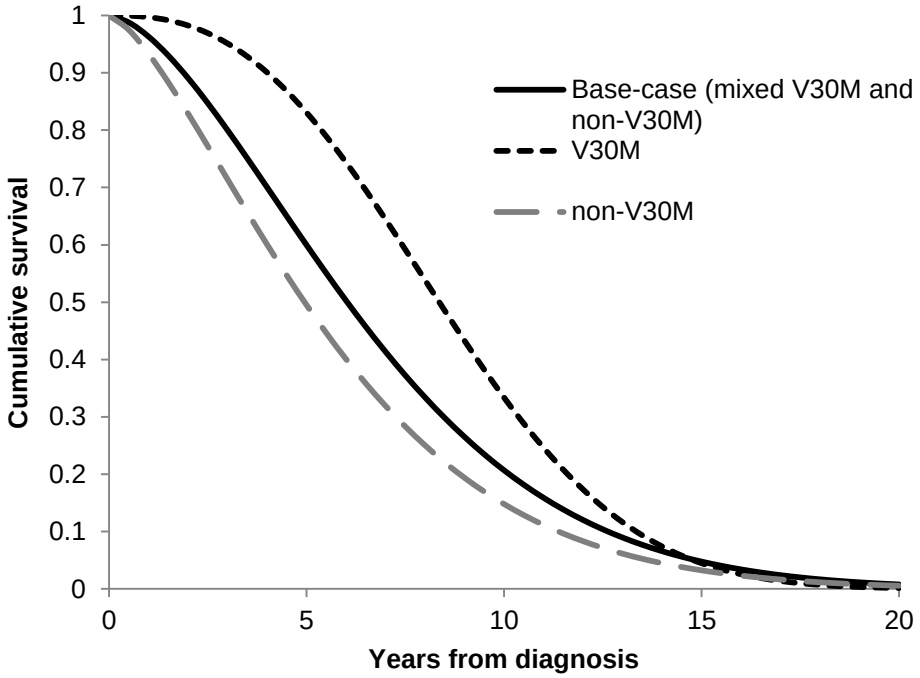
Table 21: Parameters for fitted Weibull without transplant survival

Population	Scale	Shape
Base-case (combined)	0.03744	1.624
V30M	0.00298	2.566
Non-V30M	0.06888	1.443

The Kaplan-Meier curves and the fitted Weibull survival curves used to estimated mortality risk in the model are presented in Figures 3 to 5, page 15 of the manufacturer's response to the points for clarification (Appendix 4). On visual examination, the Weibull function for the V30M cohort appears to fit reasonably well. However, this does not appear to the ERG to be the case for the non-V30M (T60A) cohort.

Figure 3 presents the Weibull survival curves used to estimate the probability of death in the manufacturer's model. Survival is greater for V30M patients compared to non-V30M patients. However, at 15 years post-diagnosis, survival is similar for V30M and non-V30M patients. The Sattianyagam study reports a median (95% CI) survival from diagnosis for non-V30M patients from diagnosis of 3.4 years (2.7-5.3).[3] However, according to the Weibull curve, at 3.4 years, 75% of the patient cohort is still alive and only at 6 years does cumulative survival reach 50%. Therefore, the Weibull curve chosen by the manufacturer may not reflect the survival of non-V30M patients (again with the same caveats about whether T60A is representative of all non V30M patients). For the Swedish V30M cohort, median survival from diagnosis was 8.2 years (1.2-15.4). This is in line with the Weibull curve, where at 8 years around 53% of the cohort is still alive.

Figure 3: Survival curves used in the economic model for the base-case population and patient subgroups for patient without liver transplantation (using parameters presented in Table 21)



Although the manufacturer modelled mortality risk to be solely dependent on time from diagnosis, the results in Sattianayagam et al (2012) indicate that mortality risk may depend on other variables.[3] Table 22 presents the hazard ratios for each of the variables analysed in Sattianayagam et al (2012). The data in Table 22 suggest that variables other than time from symptom onset have an impact on mortality risk, such as age. Therefore, the approach taken by the manufacturer to incorporate mortality in the economic model may not be appropriate. It is unclear to the ERG how the approach taken by the manufacturer impacts on the cost-effectiveness results.

Table 22: Variables associated with reduced survival in a cohort of T60A patients [3]

	Hazard ratio	95% confidence interval
Univariate analysis		
Age for each 10 years older	2.49	1.28-4.85
IVS thickness for 17 mm vs. ≥17 mm	0.31	0.14–0.72
NT-proBNP for ,400 pmol/L vs. ≥400 pmol/L	0.39	0.16 – 0.96
Diastolic dysfunction for grade 0–1 vs. grade 2–4	0.33	0.12–0.91
LVPW thickness for 17 mm vs. ≥17 mm	0.42	0.18 – 0.95
Weight loss at diagnosis for weight loss vs. no weight loss	2.85	1.08–7.54
Multivariate analysis		

NT–proBNP for 400 pmol/L vs. ≥400 pmol/L	0.17	0.03–0.92
LVPW thickness 17 mm vs. ≥17 mm	0.17	0.03–0.97
Left atrial area for .20 mm2vs.≤20 mm2	9.24	1.27–67.40
IVS – Intravertricular septal NT-proBNP – N-terminal prohormone brain natriuretic peptide LVPW – left ventricular posterioir wall thickness		

As stated previously, the ERG has a number of concerns with the approach taken to mortality in the manufacturer's model. These include: (i) the use of survival curves from time of symptom onset, rather than time of diagnosis; (ii) in the base case a use of combined survival curve where the population is not in keeping with that of England (it had a greater proportion of V30M patients than in the population in England); and (iii) whether the V30M and non V30M patients from Sattianayagam et al are comparable to those in England. A further, and perhaps more significant, concern is that the manufacturer has modelled disease severity independently of mortality.

5.3.3 Liver transplantation

In the model, all patients in stage 1 are assumed eligible for liver transplantation, although only a fraction of patients will receive one, with the rate based on expert advice from the NAC. Liver transplantation is assumed to halt the progression of the disease (leaving TQoL constant), resulting in constant HRQoL. It is also assumed to improve survival. Liver transplantation is associated with a one-off HRQoL loss as a result of the procedure, one-off costs and some recurrent costs. Below, we consider the rate of liver transplantation, the impact on HRQoL and the impact on survival in turn.

Table 23 presents the 6-monthly rate of liver transplant applied in the economic model. In order to obtain the rate of liver transplant, the number of liver transplants in the past year for ATTR patients followed by NAC was divided by the number of ATTR patients followed by NAC. The ATTR population includes foreign nationals and patient living outside England.

Table 23: Liver transplant rates over 6-months used in the economic model (adapted from Table 47, p141)

	Transplant rate	Calculation
Base case (combined V30M and non-V30M)	5.72%	2 transplants in 18 patients
Subgroup 1 (V30M)	6.50%	1 transplant in 8 patients
Subgroup 2 (non-V30M)	5.13%	1 transplant in 10 patients

The rate of liver transplantation in TTR-FAP patients in England is likely to be an overestimate. Expert clinical advice from NAC informed the ERG that in the last 5 years, there were very few liver transplants in TTR-FAP patients residing in England and that it is not generally considered a treatment option in the UK. Therefore, the rate of liver transplantation in patients in England with TTR-FAP is likely to be close to zero.

The manufacturer's model assumes that once a patient receives a liver transplant further disease progression is prevented and the patient receives a constant HRQoL (based on their TQoL score at the time of liver transplant) until their death. The reasonableness of this assumption is unclear to the ERG. For example, in non-V30M patients, where there is cardiac involvement, it appears particularly unreasonable as there may be progression of the heart disease despite the liver transplant (see Section 2.3.2).

Mortality post-liver transplant is based on the study by Herlenius et al (2004).[22] Herlenius et al (2004) examined the patient characteristics and outcomes post-liver transplant of the patients registered up to December 2000 and with a minimum 1 year of follow-up in the Familial Amyloidotic Polyneuropathy World Transplant Registry (FAPWTR). A total of 539 patients underwent 575 liver transplantations, of which 17 were combined liver and heart transplantation and 1 was a combined liver, heart and kidney transplantation.

Survival curves were available for both V30M and non-V30M patients. The manufacturer digitalised the Kaplan-Meier curves to estimate patient numbers at each time point. Weibull survival curves were fitted to the survival data for each population (449 V30M patients and 62 non V30M patients) and then considered the combined population for the base-case analysis. Since the patients reported in Herlenius et al (2004)[22] are younger than the patient cohort in the model, and the fitted survival functions showed decreasing probability of survival over time, mortality rates according to the survival functions were below those of the general population for the age of patients in the model. Therefore, the manufacturer used the age-specific all-cause mortality rate for the UK general population if this was greater than the mortality suggest by the Weibull survival curves. Therefore, the risk of death in each cycle following liver transplant is the maximum between mortality estimated from Herlenius et al (2004) and age-specific all-cause death for the UK general population. Table 24 presents the parameters for the Weibull curves.

Table 24: Parameters for fitted Weibull post-transplant survival

Population	Scale	Shape
Base-case (combined)	0.1283	0.489
V30M	0.1120	0.476
Non-V30M	0.2598	0.565

The Herlenius et al (2004) study followed patients up to 5 years after transplantation.[22] The Weibull functions fitted to the survival data allow for the extrapolation of survival over longer time horizons. According to the Weibull functions, at 10 years after transplant, 79% of V30M patients and 52% of non-V30M patients are still alive. These estimates can be compared with the 10-year survival estimates reported by Wilczec et al (2012) based on FAPWTR data (Table 2).[57] At 10 years, 73.5% of V30M patients and 43.2% of non-V30M patients are alive. Therefore, the Weibull survival functions used in the economic model may overestimate survival post-liver transplant, in particular for the non-V30M population.

The use of data from the FAPWTR patients to inform mortality post liver transplant in the economic model raises several issues. Firstly, for the base case population involving both V30M and non-V30M patients, the relative proportions of each group used to inform the survival analysis is not in keeping with the relative proportions in the patient population in England (V30M patients made up roughly 80% of all patients in the Herlenius et al study but only 16.66% of all patients in the economic model), and given the V30M patients have better survival, this will overestimate the benefits of liver transplant in the base case. However, as previously stated, the ERG does not consider it appropriate to consider a combined population in any case. Secondly, the patients ages differ markedly from the population in England (the mean age at transplant in the Herlenius et al study was 40.6 years of age, yet the average age of patients entering the model was 63 years of age in the base case). The ERG would therefore expect that this would over estimate the benefit of liver transplant. This is reflected by the estimates of mortality from the Weibull survival functions falling below those from the UK general population for the age of patients entering the model. In an attempt to address this issue, the manufacturer used the maximum of the mortality rate from the survival functions or the UK general population, however, if the UK general population estimate is used the implicit assumption is that following liver transplant patients will have the same mortality as the general population, an assumption which appears unreasonable and overly optimistic to the ERG.

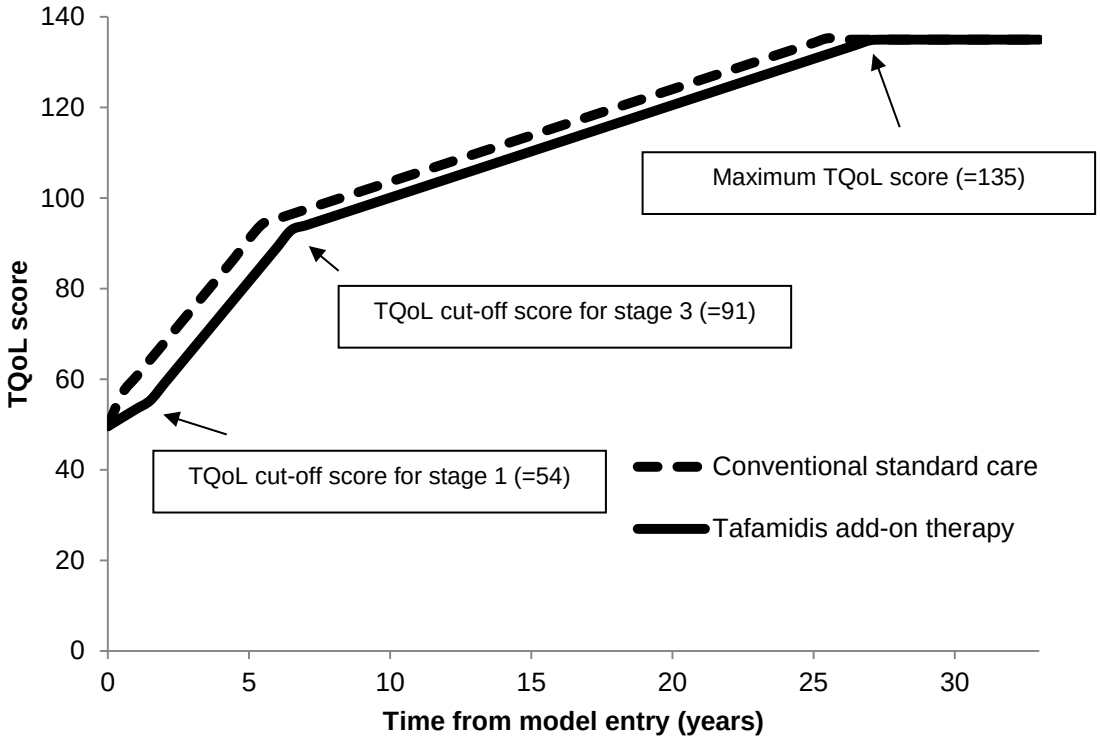
As stated previously, the ERG has a number of concerns with the approach taken with liver transplantation in the economic model. Firstly, it is unclear if the rate of liver transplantation for TTR-FAP patients in the UK would be as great as suggested in the economic model. Secondly, the assumption that liver transplantation halts disease progression is unproven, and appears particularly unjustified in non V30M patients where there may be cardiac involvement as well. Finally, the study used to model mortality following liver transplantation does not appear representative of the patients with TTR-FAP in England and may over estimate survival.

5.4 Treatment effectiveness within the submission

In the economic model, tafamidis reduces the TQoL rate of change whilst a patient is on treatment. The reduction in TQoL rate of change implies that disease progression is slowed, although mortality risk remains unchanged. This results in tafamidis improving health outcomes through a number of ways in the economic model. Firstly, health outcomes are improved because TQoL scores are lower for patients on tafamidis than for patients on conventional standard care, and TQoL scores determine HRQoL (Section 5.5 discusses the mapping of TQoL scores into the HRQoL measure in more detail). Therefore, patients on tafamidis have improved HRQoL scores compared to those on conventional standard care. This impact will remain until death in the model because their TQoL score will remain lower than those on conventional standard care for all cycle.

Figure 4 shows how tafamidis has a perpetual effect on TQoL scores even though it is only given for a short period of time, whilst the patient is in stage 1. The dashed line shows the change in TQoL score for conventional standard care, whereas the full line shows the change in TQoL score for tafamidis add-on therapy. Although patients are only on tafamidis during stage 1, which in the economic model is one year due to the baseline TQoL score at model entry of 49.64, the beneficial effects of treatment are modelled to last throughout the patient's lifetime. Patients on tafamidis not only reach stage 2 at a later time compared to conventional standard care (approximately 18 months for patients on tafamidis versus 5 months for patients on conventional standard care), but also cross-over to stage 3 at later in their lifetimes (approximately 6.5 years for patients on tafamidis versus 5 years for patients on tafamidis). There is no evidence to suggest that the benefit from treatment is sustained after treatment discontinuation, although there is also no evidence of a rebound effect (i.e. patients go back to natural history after discontinuation). Whether this is a reasonable assumption remains unclear to the ERG.

Figure 4: Effect of tafamidis in TQoL score



The second, and most important, benefit of tafamidis treatment with regards to the cost-effectiveness results is that tafamidis increases the likelihood that patients receive a liver transplantation by keeping them in stage 1, where they are eligible for liver transplant, for longer. The impact of liver transplantation on survival and HRQoL are discussed in detail in Section 5.3. Thirdly, tafamidis also has an impact on costs associated with the management of the disease. Since tafamidis is modelled to reduce the TQoL rate of change, patients remain in stage 1, a less costly stage in terms of resource use, for longer. In addition, the one-off cost incurred in the transition to stage 2 and subsequently to stage 3 occurs later, and as a result of discounting this will reduce the net present cost of the transitions.

The manufacturer's model assumes that tafamidis treatment has no impact on mortality, other than through increasing the chance of a patient receiving a liver transplant. There is no evidence to suggest that tafamidis improve survival, however, the clinicians contacted by the ERG considered that it would be reasonable to expect that, if tafamidis does have a beneficial impact on disease progression, survival may also be improved. Therefore, the assumption of no impact on mortality may be considered to be conservative.

Table 25 compares the different estimates of monthly TQoL rates of change and corresponding hazard ratios derived using three alternative methods and data collected during Fx-005. The economic model uses the hazard ratio of 0.265, which corresponds to the ratio of the monthly rate of change in the tafamidis group by that of the placebo group, estimated using slopes of graphs (post-hoc analysis). However, the pre-specific primary analysis of Fx-005 indicates different estimates of monthly TQoL rate of change. The ERG calculated the hazard ratio corresponding to each. The hazard ratio using the TQoL rate of change calculated using means is 0.348, whereas the hazard ratio calculated using the TQoL rate of change estimated from least square means and adjusting for baseline TQoL scores is 0.278. Since a lower hazard ratio indicate greater treatment effect, the hazard ratio chosen for the economic model favours the cost-effectiveness results towards tafamidis. The manufacturer provided no justification on the choice of the hazard ratio employed in the model.

Table 25: Comparison of monthly rate of change observed in Fx-005 used in the economic model and those estimated from the results of the pre-specified primary end-points, and corresponding hazard ratio associated with tafamidis (Adapted from Tables 11 and 12, p48 of MS and adapted from Table 14, p108 of Clinical Study Report for Fx-005)

	TQoL change over 18 months		Monthly TQoL rate of change		Hazard ratio
	Placebo	Tafamidis	Placebo	Tafamidis	
Post-hoc analysis used in the economic model					
TQoL rate of change calculated using slopes	[redacted]	[redacted]	0.4618	0.1225	0.265
Pre-specified primary analysis					
TQoL rate of change calculated using means	6.9	2.4	0.3833†	0.1333†	0.348‡
TQoL rate of change calculated using LSM and adjusting for baseline TQoL	7.2	2.0	0.4000†	0.1111†	0.278‡

†Calculated by the ERG assuming constant TQoL rate of change over the 18 months duration of Fx-005;
‡Calculated by the ERG based on the monthly TQoL rate of change for tafamidis and placebo.
LSM – Least Square Means.

The manufacturer assumes that treatment effect is the same across patient populations, independently of genetic variant (V30M or non-V30M). However, the key evidence for tafamidis is from Fx-005, a RCT conducted in a TTR-FAP population with the V30M mutation. As discussed in Section 4.8, the results of Fx-005 are likely to be most applicable to patients with early-onset TTR-FAP related to a V30M mutation. However, patients in

England have predominantly non-V30M mutations, and even those with the V30M mutation are unlikely to be characterised as early onset. Therefore there is uncertainty as to whether the results from the trial are applicable to the UK patient population.

It should be noted that the TQoL rate of change in Fx-005 was little more than a third of the rate of change applied in the economic model for stage 1 (the calculation of which was described and critiqued in detail previously). For the placebo group, the monthly rate of change of 0.4618 corresponds to a 6-monthly rate of change of 2.77, while the economic model uses a 6-monthly rate of change of 7.17. If the 6-monthly rate of change in TQoL from Fx-005 is a more accurate reflection of the disease progression, untreated patients spend a longer time in stage 1 than predicted by the model and would be more likely to receive a liver transplant. The benefits from tafamidis treatment would also be exaggerated, with the model producing a greater absolute reduction in the rate of increase of TQoL than would be observed in practice. Therefore, the economic model may overestimate the benefits of tafamidis.

Patients on tafamidis treatment can discontinue treatment due to adverse effects up to month 18. The discontinuation rate observed during Fx-005 of [redacted] ([redacted] over the [redacted] months of the trial) was applied in the economic model to the costs but not on the health outcomes. This implies that patients who discontinue tafamidis continue to experience the benefits from treatment whilst no acquisition costs are incurred. Since there is no evidence suggesting that treatment effect is sustained after discontinuation, nor does the manufacturer makes a claim of such effect, the ERG concluded that discontinuation rate was applied erroneously in the manufacturer's model. Since health benefits are sustained without tafamidis acquisition costs, this error in the incorporation of discontinuation from treatment favours the cost-effectiveness results towards tafamidis.

In summary, the effectiveness of tafamidis is modelled through the reduction in TQoL rate of change, which in turn improves HRQoL, reduces costs and increases the likelihood of the patient receiving a liver transplantation. Assuming that tafamidis has no effect on survival may have underestimated treatment benefits in the model, however there is no evidence to suggest the existence of such survival benefit. On the other hand, there are three key issues which may have resulted in a overestimation of the benefits from treatment: (i) the assumption that benefits from treatment are sustained throughout lifetime (i.e. there is no catch up following discontinuation of treatment); (ii) the hazard ratio chosen to reflect the effect of tafamidis in TQoL rate of change; and (iii) the generalisability of the results from a trial conducted in the V30M population to the patients in England, who typically present with a non-V30M variant.

5.5 Health related quality of life

Health outcomes were expressed in terms of QALYs by quality-adjusting the cycles the patient was alive within the model with the appropriate HRQoL scores. The manufacturer considered three elements of HRQoL: (i) HRQoL associated with TTR-FAP, (ii) HRQoL associated with liver transplantation, and (iii) HRQoL losses by the patient's carer. Table 26 summarises the HRQoL scores used in the model. The following sections discuss each of the elements of HRQoL considered in the manufacturer's submission in turn.

Table 26: HRQoL scores used in model

Health state or event	HRQoL	
Disease stages	TQoL	EQ-5D
Stage 1	48.97 to 54	0.607 to 0.636
Stage 2	54 to 91	0.397 to 0.636
Stage 3	91 to 135	0.147 to 0.397
Liver transplantation		
One-off QALY loss	-0.2 QALYs applied in the first 6-months post-liver transplantation	
Post-liver transplantation	Assumed constant and equal to EQ-5D score at transplant	
Effect on carers		
QALY loss	-0.01 per cycle once patient enters stage 3	

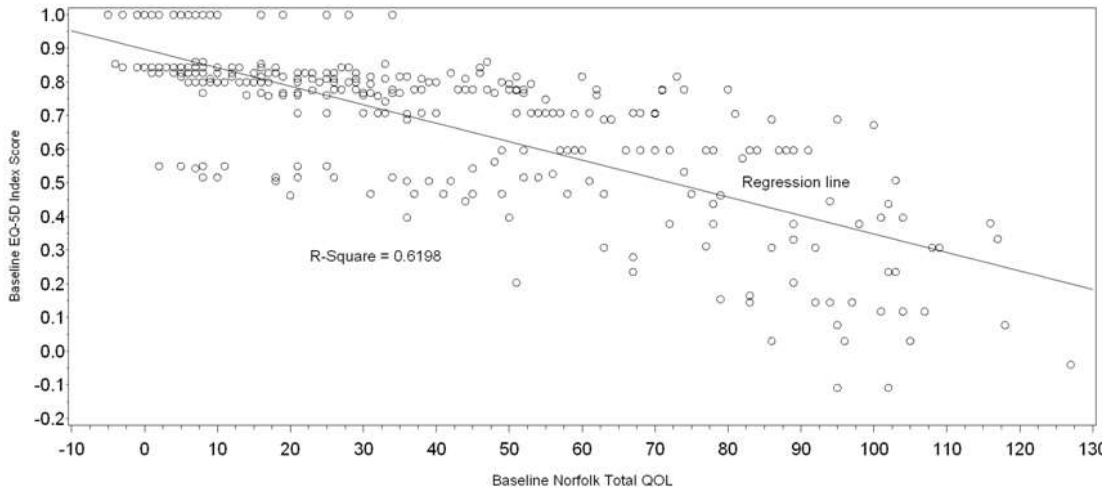
5.5.1 HRQoL associated with TTR-FAP

The HRQoL score experience by the patient in any cycle was determined by the patient's TQoL score via a mapping function of TQoL scores onto EQ-5D scores, a generic measure of HRQoL. Therefore, HRQoL scores depend on the patient's TQoL score. As TQoL scores increase, marking more severe disease, the EQ-5D scores decrease accordingly. Lower EQ-5D scores indicate lower HRQoL.

The EQ-5D questionnaire allows patients to describe how they experience their HRQoL and comprises five dimensions of health: mobility, ability to self-care, ability to undertake usual activities, pain and discomfort, and anxiety and depression. A set of preference values, which was elicited from a sample of the UK population, can be applied to the patients' self-reported descriptions of their HRQoL to generate HRQoL scores.[52] An EQ-5D score of 1 represents perfect health, whilst 0 represents death.

The relationship between EQ-5D and TQoL was obtained from a cross-sectional analysis of baseline data in the THAOS registry. Figure 5 presents the scatter plot of baseline EQ-5D scores by baseline TQoL scores from the data in the THAOS registry. Equation 2 below shows the manufacturer's mapping function, which corresponds to the regression line in Figure 5. The analysis of the scatter plot suggests a concentration of EQ-5D scores at approximately the 0.7-0.8 for a range of TQoL scores from 0 to 65 points. In addition, there is another set of TQoL scores, from 0 to 100, in the region of EQ-5D =0.5-0.6. From TQoL score 60 onwards, the EQ-5D ranges from 0 to 0.7. The manufacturer provided no evidence on the validity of the mapping function in their submission. Given that TQoL is a condition-specific measure of HRQoL, i.e. a measure of quality of life related with diabetic neuropathy, it may not be appropriate to map it to a generic measure of HRQoL such as EQ-5D. If TQoL does not capture a spectrum of HRQoL as wide as EQ-5D, mapping TQoL scores to EQ-5D may not appropriately reflect the HRQoL experienced by patients.

Figure 5: Scatter plot of baseline EQ-5D scores by baseline TQoL scores (reproduced from Figure 29, p145 of MS)



Equation 2: Manufacturer's mapping function used for the base-case analysis

$$EQ - 5D = 0.913991 - 0.005682 * TQoL$$

Table 27 shows the mean TQoL per stage, the thresholds for moving disease stage and the corresponding EQ-5D score, using the manufacturer's mapping function. The maximum EQ-

5D score attainable is 0.914, which corresponds to TQoL score of zero. The minimum EQ-5D score attainable is 0.147, corresponding to the maximum TQoL score of 135 points. The mean TQoL per stage was sourced from the manufacturer's responses to the points for clarification, which presented summary statistics per stage from THAOS registry data (Appendix 13 of manufacturer's responses to the points for clarification). The mean TQoL score for stage 1, and also the baseline TQoL score at model entry, is 48.97 points, which corresponds to 0.636 in the EQ-5D measure. For stage 2, mean EQ-5D is estimated at 0.501. For stage 3, mean EQ-5D is 0.375. These values can be compared with the EQ-5D reported by the general population and by patients with other chronic conditions.[58] For example, the average EQ-5D score for patients in stage 1 of 0.636 is similar to that reported by the individuals of the general population between 80 and 89 years of age. The average EQ-5D score for patients in stage 2 of 0.501 is similar to that of patients with rheumatoid arthritis and related disorders of 0.510. The average EQ-5D score for patients in stage 3 of 0.375 is similar to that of patients with paralysis of 0.350.

Table 27: Mean TQoL per stage and thresholds for moving disease stage and corresponding EQ-5D score using the manufacturer's mapping function

Mapping function: $EQ-5D = 0.913991 - 0.005682 * TQoL$	**TQoL**	**EQ-5D**
Minimum and maximum EQ-5D scores attainable		
Minimum HRQoL	135	0.147
Maximum HRQoL	0	0.914
Mean TQoL score per stage		
Disease stage 1	48.97	0.636
Disease stage 2	72.68	0.501
Disease stage 3	94.83	0.375
Thresholds for moving disease stage		
Stage 1 to Stage 2	54	0.607
Stage 2 to Stage 3	91	0.397

Given the issues discussed above, the ERG requested from the manufacturer alternative functional forms, including squared and cubic functions, and separate models by disease stage. Table 28 compares the EQ-5D scores corresponding to the average TQoL per disease stage and for each stage threshold, using the original mapping function and using the functions presented in the response to the points for clarification. The minimum EQ-5D score attainable is -0.145 for the quadratic function, -0.222 for the cubic function and -0.107

using the linear function by disease stage. The maximum EQ-5D score attainable is 0.89 for the quadratic function, 0.91 for the cubic function and 0.930 using the linear function by disease stage. All these three options map the maximum TQoL of 135 points, corresponding to the worst HRQoL, to negative EQ-5D values. The mean TQoL per disease stage is fairly consistent across mapping functions, which the exception of the mean HRQoL score for stage 3. Using the linear function by stage, the mean QoL score at stage 3 of 94.83 is converted into EQ-5D=0.17, which is less than half of that obtained using the manufacturer's original mapping function (EQ-5D=0.375). This difference can also be observed for the conversion of TQoL scores into EQ-5D for the cut-off scores between stages. The cut-off score between stage 2 and 3 is 91, which, using the manufacturer's original mapping function, corresponds to 0.397 while using the linear function by stage corresponds to 0.196. Without access to the original patient level data, the ERG is unable to determine which is the most appropriate function and whether other functions could better model the relationship between TQoL and EQ-5D.

Table 28: EQ-5D utility values for mean TQoL per stage and stage threshold using the alternatives presented in the response to the points for clarification

	TQoL	**Quadratic***	**Cubic****	**By stage*****
Minimum and maximum attainable				
Maximum HRQoL	0	0.89	0.91	0.930
Minimum HRQoL	135	-0.015	-0.222	-0.107
Mean TQoL per disease stage				
Disease stage 1	48.97	0.646	0.662	0.705
Disease stage 2	72.68	0.494	0.539	0.551
Disease stage 3	94.83	0.331	0.366	0.17
Thresholds for moving disease stage				
Stage 1 to Stage 2	54	0.616	0.639	0.631
Stage 2 to Stage 3	91	0.36	0.402	0.196

* Quadratic: $EQ\text{-}5D=0.89-0.004*TQoL-0.00002*TQoL^2$

** Cubic: $EQ\text{-}5D=0.90979-0.00712*TQoL+0.00007123*TQoL^2-0.000000596927*TQoL^3$

*** By stage: Stage 1: EQ-5D=0.930807-0.004613*TQoL; Stage 2: EQ-5D=0.861597-0.004278*TQoL; Stage 3: EQ-5D=0.822396-0.006884*TQoL.

EQ-5D scores directly elicited from patients could have been used in the model rather than mapping from TQoL to EQ-5D. The THAOS registry collects EQ-5D directly from patients and the manufacturer classified the patients in the THAOS registry according to disease stage using mPDS (see Section 5.2 for more details). This alternative approach is in line with the recommendations of the NICE methods guide [52]and would avoid the uncertainty associated with mapping from a HRQoL measure focussed on neuropathy symptoms (TQoL) to a generic measure of HRQoL such as EQ-5D. Furthermore, the approach taken by the manufacturer makes HRQoL, through EQ-5D scores, dependent on TQoL rate of change. Therefore, the issues with manufacturer's approach to estimate TQoL rate of change, discussed in Section 5.2, are compounded further in the model due to the link to HRQoL.

5.5.2 HRQoL associated with liver transplantation

In the model, liver transplantation impacts HRQoL in two aspects. Firstly, liver transplantation results in a one-off loss of 0.2 QALYs during the first 6-months following transplantation. Secondly, and most important for the cost-effectiveness of tafamidis, liver transplantation is assumed to halt the progression of TQoL scores. Therefore, liver transplantation maintains EQ-5D scores unchanged throughout the patient's lifetime.

One-off QALY loss following liver transplantation

The one-off loss of 0.2 QALYs during the first 6-months following transplantation aims to account for the impact of the procedure on HRQoL. The QALY loss of 0.2 was based on the study by Ratcliffe et al (2002).[50] The ERG has doubts on whether the study by Ratcliffe et al (2002) provides evidence on the HRQoL loss associated with the liver transplant procedure, given that the study appears to show a short term (at 3 months) gain from liver transplant compared to baseline. However, the ERG considers it reasonable to assume that there is a one-off HRQoL loss as a result of liver transplantation, although the exact value is unclear.

Effect of liver transplantation on HRQoL

As discussed in Section 5.3.3, liver transplantation is assumed to maintain HRQoL constant through halting TQoL progression. Therefore, in the model, patients following liver transplantation maintain the same EQ-5D score throughout their lifetime. This assumption may overestimate the benefits of liver transplantation in two ways. First, evidence on HRQoL

in the general population indicates that individuals experience a natural decline in HRQoL as they get older. Therefore, it is unrealistic to assume patients to maintain HRQoL constant throughout their lifetime. Second, and as discussed in Section 2.3.2 and 5.3.3, there is uncertainty about the effects of liver transplantation in patients with non-V30M forms of the disease with cardiac involvement, which is the most frequent presentation in England. Therefore, the benefits of liver transplantation in HRQoL may have been overestimated.

Since a key component of the benefits from tafamidis lies in increasing the likelihood of receiving a liver transplant, by maintaining patients in stage 1 for longer, the possible overestimation of benefits from liver transplantation results in an overestimation in the benefits associated with tafamidis. Therefore, the cost-effectiveness results are favoured towards tafamidis.

5.5.3 HRQoL experienced by carers

A QALY loss of 0.01 was applied to stage 3 patients to account for the impact on carers. The value of 0.01 was based on the NICE final appraisal determination for the treatment of Alzheimer patients.[51] Although Alzheimer disease may impose a different level of burden to carers compared with TTR-FAP, the ERG considers that it may be reasonable to use the value of 0.01 to account for the impact on carers.

5.5.4 HRQoL: Final issues

In summary, ERG considers that the manufacturer generally followed standard economic evaluation methods to incorporate HRQoL in the model, namely: (i) the estimation of health outcomes in terms of QALYs and the use of EQ-5D to estimate QALYs, following the guidelines for economic evaluations for the NHS [52], and (ii) inclusion of all health effects relevant to the disease. Nonetheless, the manufacturer's approach raises a number of issues. Firstly, EQ-5D scores directly elicited from patients, recorded in the THAOS registry, could have been used, rather than mapping TQoL to EQ-5D. Secondly, the manufacturer provided no evidence on the validity of the mapping function in their submission. Given that TQoL is a condition-specific measure of HRQoL, it is unclear whether it is appropriate to map it to a generic measure such as EQ-5D. Thirdly, liver transplantation is assumed to maintain HRQoL unchanged throughout the patients' lifetime, although evidence from HRQoL in the general population indicates that HRQoL naturally worsens at older ages. Fourthly, liver transplantation is assumed to have the same HRQoL benefits for V30M and non-V30M patients, despite the considerable uncertainty about the effects of liver

transplantation in non-V30M patients. As a result, the cost-effectiveness estimates may be biased in favour of tafamidis.

5.6 Resources and costs

The model included the acquisition costs of tafamidis, the costs associated with healthcare resource use related to both disease stage and liver transplantation, and the costs associated with productivity losses. Table 29 summarises the costs used in the manufacturer's submission. Each element of the costs is discussed in turn.

Table 29: Costs by disease stage (adapted from Table 50, p148 of MS)

	Six-month cost	One off cost
Drug acquisition costs		
Tafamidis cost	[redacted]	-
Costs associated with healthcare resource use per disease stage		
Stage 1	£2,272	-
Stage 2	£7,548	£1,803
Stage 3	£10,076	£4,021
Costs associated with healthcare resource use due to liver transplantation		
Transplant procedure	£20,501	-
6-month cost of transplant in 1st year	£22,232	-
6-month cost of transplant after 1st year	£814	-
Productivity costs accrued by patients and carers		
Stage 1	£2,514	-
Stage 2	£8,238	-
Stage 3	£8,238	-

5.6.1 Acquisition costs of tafamidis

The drug acquisition cost of tafamidis is [redacted] per patient per year, based on a provisional pack cost of [redacted], awaiting confirmation from the Department of Health. The manufacturer proposes to deliver tafamidis to patients' addresses, making the acquisition of tafamidis not subject to VAT. Following the request for details by the ERG on the points for clarification, the manufacturer clarified that the arrangements have not been finalised.

The arrangements for home delivery may have an impact on the cost-effectiveness of tafamidis. Treatment discontinuation may result in unused tablets in the patient's possession which cannot be used for other patients. Given the acquisition cost of tafamidis, any medication wastage will make tafamidis less favourable in terms of cost-effectiveness.

5.6.2 Costs associated with healthcare resource use per disease stage

Data on resource use was obtained from a group of clinicians based in Sweden since the manufacturer stated they received no responses from the UK-based specialists consulted. The UK clinicians contacted by the ERG considered the data on resource use provided by the Swedish clinicians as generally applicable to the UK clinical setting. Unit costs were sourced from national sources. [59-61] Table 30 summarises the costs associated with recurrent healthcare resource use by stage. Homecare represents the greatest proportion of healthcare costs for each of the 3 stages. For home care, patients in stage 1 are assumed to require 6 hours of home care worker service per month. Patients in stage 2 are assumed to require 36 hours of home care worker service per month. Patients in stage 3 are assumed to require 36 hours of home care service per month and 1 day of special housing (in a residential or nursing care home unit for adults with physical disabilities) per month.

Table 30: Breakdown of healthcare costs per disease stage (adapted from Table 55,56,58, p154-168 of MS)

Type of resource	Stage 1		Stage 2		Stage 3	
	6-month cost	%	6-month cost	%	6-month cost	%
Polyneuropathy	£133.30	6%	£718.83	10%	£555.55	6%
Gastrointestinal disorders	£397.09	17%	£230.01	3%	£296.70	3%
Cardiac arrhythmias	£306.19	13%	£505.55	7%	£305.75	3%
Bladder dysfunction	£9.13	0%	£12.78	0%	£140.38	1%
Ocular problems	£14.33	1%	£49.31	1%	£31.45	0%
Other	£467.96	21%	£694.15	9%	£2,956.31	29%
Primary care	£139.60	6%	£602.86	8%	£285.46	3%
Aids	£3.23	0%	£9.41	0%	£0.00	0%
Homecare	£800.80	35%	£4,724.72	63%	£5,504.29	55%
Total	£2,271.63		£7,547.62		£10,075.89	

In addition to recurrent costs described in Table 30, the model applies a one-off cost at the progression to stage 2 and subsequently at progression to stage 3. The one-off cost associated with progression to stage 2 is £1,803 and the one-off cost associated with progression to stage 3 is £4,021. The clinicians contacted by the ERG considered the resource use that these one-off costs refer to to be reasonable and applicable to the UK setting.

5.6.3 Costs associated with healthcare resource use from liver transplantation

The costs from liver transplant consist of the costs as a result of the procedure itself, the costs in the first year of transplant, and recurrent costs after the first year. The costs associated with the transplant procedure are £20,501 and the subsequent costs in the first 6-month cycle after the transplant are £22,232. The manufacturer obtained both values from the National Specialist Commissioning Team. The 6-month cost post-liver transplant is £814 and was based on the systematic review and economic evaluation of treatments for chronic hepatitis C commissioned by NICE by Hartwell et al (2011).[62] Following liver transplantation, patients continue to incur costs associated with stage 1, namely £2,272 of healthcare resources and £2,514 for productivity costs.

5.6.4 Productivity costs accrued by patients and carers

The manufacturer included the productivity costs incurred by patients and their carers in the base-case. Productivity costs refer to the income forgone by patients and carers as a consequence of the disease. Table 31 shows the breakdown of productivity costs used in the model. The productivity costs for stage 1 are £2,514 per patient over a 6-month period. These costs assume that relatives lose 4.3 days of work due to the patient's disease, 7.5% of patients are on sick leave and 10% have taken early retirement. The productivity costs for stage 2 and stage 3 are £8,238 per patient over a 6-month period. These costs assume that relatives lose 4.3 days of work due to the patient's disease, 57% of patients are on sick leave and 22% have taken early retirement. The sources for the days of work lost by relatives, and the proportion of patients in early retirement and in sick leave were not provided.

Table 31: Breakdown of productivity costs per disease stage (adapted from Table 55,56,58, p154-168 of MS)

Cost item	Unit cost	Stage 1		Stage 2 = Stage 3	
		Productivity loss‡	6-monthly costs	Productivity loss	6-monthly costs
Days of productivity loss, relatives over 6-months	£103.56	4.3	£444	4.3	£444
Early retirement	£103.56	10%	£1,310	22%	£2,882
Sick leave 100%	£103.56	3.76%	£493	13%	£1,703
Sick leave 75%	£103.56	1.28%	£126	17%	£1,670
Sick leave 50%	£103.56	1.86%	£122	20%	£1,310

Sick leave 25%	£103.56	0.62%	£20	7%	£229
		Total	£2,514	Total	£8,238
†Unit cost obtained from the ONS Annual Survey of Hours and Earnings 2011 (ref)and corresponds to the median annual UK salary for full-time employees per working day (£26,200/253 working days = £103.56 gross). Productivity loss in terms of number of work days lost by relatives per year or proportion of patients in each of the sick leave or early retirement categories.					

The inclusion of productivity raises two issues. Firstly, it is unclear what the perspective of the analysis should be. Productivity costs should not be included when the perspective taken is that of the NHS & PSS.[52, 63] The NICE methods guide, which informs economic evaluations for the NHS, allows for the inclusion of costs falling outside the NHS budget only 'in exception circumstances' and 'if this has been specifically agreed with the Department of Health, usually before referral of the topic'.([52]) A further issue is that this patient population is approaching retirement. Therefore it is unclear if productivity costs should be included at all (the age at model entry is between 61 and 63 years of age depending on the genetic variant whilst the State Pension age in the UK is 65 years of age for men and between 60 and 65 years of age for women).([64]) Given that productivity costs are approximately three times greater for stages 2 and 3 compared to stage 1, and that tafamidis keeps patients in stage 1 for longer, the inclusion of productivity costs favours the cost-effectiveness results towards tafamidis.

5.7 Sensitivity analyses

The manufacturer conducted a number of one-way sensitivity analyses on the base-case (combined V30M and non-V30M) analysis. No rationale was provided for the choice of variables and ranges tested in the sensitivity analysis.

5.8 Model validation

The manufacturer's submission did not discuss the validation process or whether any model validation took place. The ERG conducted a detailed examination of the electronic model. The model was generally consistent with that described in the manufacturer's report although a number of consistencies were identified. Firstly, the costs and benefits of the first 6 months of the model are excluded from the overall results for costs and QALYs. This biased the results in favour of the tafamidis strategy, as patients receiving tafamidis experience the benefits of reduced TQoL in the first cycle at no cost. Secondly, patients who die in the cycle in which they receive their liver transplants incur no costs or QALY

decrements for the liver transplant. Thirdly, the manufacturer's model considered patients discontinuing from tafamidis treatment, however, in the model patients who discontinued still received the benefit of the treatment, in terms of a lower rate of TQoL progression in stage 1, but without incurring the cost. This biased the results in favour of tafamidis but as the rate of discontinuation is low, the bias was minor. Finally, the model did not draw from the distributions of baseline TQoL and age when the patient level simulation was run over 20,000 patients. This meant that the ERG was not initially able to replicate the results shown in the manufacturer's submission. However, the ERG was able to correct for this issue and was able to replicate the results.

Whilst the issues noted above were considered relatively minor by the ERG, with the exception of exclusion of first period costs and QALYs, the ERG considered the way in which the model was constructed overly complex and laboursome. The use of individual patient simulation in the model, and the need to run it for 20,000 simulations to calculate mean costs and QALYs made re-analysis challenging and time consuming. As a result of this the ERG reconstructed the model as a cohort model using the same assumptions and parameters as the manufacturer's model. Section 6 discusses the ERG model in more detail.

5.9 Results included in manufacturer's submission

Results were presented for the base-case (combined V30M and non-V30M) population, and for the two subgroup populations, V30M and non-V30M patients. One-way sensitivity analysis was conducted for the base-case populations.

5.9.1 Base-case

Table 32 presents the cost-effectiveness results for the base-case population (combined V30M and non-V30M) and for the two subgroup populations, V30M and non-V30M patients. The ICER for the base-case population is £189,995 per QALY gained. The ICER for the V30M population is £174,634 and for the non-V30M population is £304,293 per QALY gained. The ICER for the non-V30M patient population is greater than the ICER for V30M and for the combined population as a result of their greater mortality risk both without and post-liver transplantation, thereby reducing the time over which the gains from tafamidis (i.e. the lower level of TQoL for every period than in the conventional standard care arm) are captured.

Table 32: Cost-effectiveness results for the base-case population (combined V30M and non-V30M) and for the two subgroup populations, V30M and non-V30M patients (adapted from Tables 51-53 of MS)

Intervention	Mean Costs	Mean QALYs	Incremental Costs	Incremental QALYs	ICER
Base-case (V30M and non-V30M)					
Conventional supportive therapy	£175,789	2.92	-	-	-
Tafamidis + Conventional support therapy	£356,007	3.86	£180,218	0.95	£189,995
V30M					
Conventional supportive therapy	£221,909	3.47	-	-	-
Tafamidis + Conventional support therapy	£410,694	4.55	£188,785	1.08	£174,634
Non-V30M					
Conventional supportive therapy	£144,595	2.47	-	-	-
Tafamidis + Conventional support therapy	£328,906	3.07	£184,312	0.61	£304,293

5.9.2 Sensitivity Analysis

Table 33 presents the results of the one-way sensitivity analysis on the base-case population. No sensitivity analysis was conducted for the subgroup populations. The ICER changed little to variations in baseline TQoL scores and to alternative assumptions about costs associated with the management of the disease. In contrast, rate of liver transplantation, the effect of tafamidis on TQoL progression and age at model entry affected the ICER to a greater degree. Reducing the rate of liver transplant to zero resulted in the ICER more than doubling to £602,850 per QALY gained. A threefold increase in the rate of liver transplant to 15.02% caused the ICER to fall to £98,112 per QALY gained. Varying the hazard ratio effect of tafamidis in TQoL progression had a marked impact on the ICER. The ICER increased to £273,481 when the effectiveness of tafamidis was reduced (the hazard increasing from 0.265 to 0.531). Improving the effectiveness of tafamidis (from 0.265 to 0.133) reduced the ICER to £164,814 per QALY gained. Raising the age at model entry to 73 years resulted in the ICER increasing to £244,690 and reducing to 53 years resulted in the ICER decreasing to £163,328.

Table 33: Sensitivity analysis to base-case population (combined V30M and non-V30M)

Analysis description	ICER (£/QALY)	% change from base-case
Base-case	£189,995	-
Baseline TQoL (=48.97)		
+ 2 points = 50.97	£184,884	-2.69%
- 2 points = 46.97	£192,349	1.24%
- 10 points = 38.97	£209,177	10.10%
Rate of liver transplant (=5.72%)		
0%	£602,850	217.30%
15.02%	£98,112	-48.36%
Tafamidis effect (Rate ratio=0.265)		
Rate ratio=0.133	£164,814	-13.25%
Rate ratio=0.531	£273,481	43.94%
Patient's age at model entry (age=63)		
+10 years = 73	£244,690	28.79%
- 10 years = 53	£163,328	-14.04%
Costs associated with conventional support therapy (£44,712)		
+ 25% (=£55,888)	£183,821	-3.25%
-25% (=£33,533)	£202,746	6.71%
Minus patient productivity costs (=£27,052)	£203,105	6.90%
Minus patient and carer productivity costs (=£25,964)	£204,043	7.39%

The results of the sensitivity analysis indicate that rate of liver transplantation is a key driver of cost-effectiveness. However, and as discussed in Section 2.3.2 and 5.3.3, liver transplantation is unlikely to be an option for UK patients. Therefore, the ICER for tafamidis of £602,850 per QALY gained, which assumes that the rate of liver transplantation is zero, may be a more appropriate estimate for the UK.

5.9.3 Budget Impact considerations

The manufacturer estimated a prevalence of 17 patients with symptomatic TTR-FAP in stage 1 and an incidence of 10 patients a year, based on expert opinion at the NAC, epidemiology data from the NAC database,[56] and a study on T60A patients .[3] Table 34 summarises the data used to estimate the prevalence and the incidence of symptomatic TTR-FAP at stage 1. The estimates for prevalence and incidence were confirmed by expert opinion at NAC by the ERG.

Table 34: Prevalence and incidence of ATTR-FAP in England

Prevalence	
Number of patients followed by NAC with hereditary ATTR	100
V122I patients, which presents with severe, restricted cardiomyopathy without significant neuropathic features.	Exclude 26
T60A patients without peripheral neuropathy	Exclude 12
Patients not residing in England	Exclude 7
Patients with V30M mutations who are foreign nationals living abroad	Exclude 20
Total number of patients with hereditary ATTR and peripheral neuropathy	35
Patients in stage 1	50%
Total number of patients with hereditary ATTR-FAP in stage 1	**17**
Incidence	
Number of patients with hereditary ATTR presenting between April 2010 and April 2011	42
V122I patients, which presents with severe, restricted cardiomyopathy without significant neuropathic features.	Exclude 11
T60A patients without peripheral neuropathy	Exclude 5
Patients not residing in England	Exclude 8
Patients with V30M mutations who are foreign nationals living abroad	Exclude 3
Total number of patients with hereditary ATTR and peripheral neuropathy	15
Patients in stage 1	70%
Total number of patients with hereditary ATTR-FAP in stage 1	10

A number of assumptions were made to estimate the budget impact to the NHS in England:

- Patients remain in stage 1 for an average of 3 years, based on the economic model estimate of 3.58 years in stage 1.
- The current prevalent population have remained in stage 1 for one year. Consequently, the prevalent population in year 1 transitions to stage 2 in year 3 and discontinues tafamidis. The incident population in year 1 transitions to stage 2 in year 4 and discontinues tafamidis.
- Treatment uptake by incident patients is 25% in the first year, increasing to 60% by year 5. No justification was given to the predicted treatment uptake.
- Tafamidis will be delivered to patients' addresses, thereby avoiding the VAT impact on costs.
- Patients only discontinue tafamidis due to transition to stage 2. Discontinuation due to adverse effects or non-compliance was not considered.

- No medication wastage due to discontinuation. Patients are delivered the medication frequently enough to prevent wastage due to discontinuation.

Table 35 presents the manufacturer estimates of the budget impact on the NHS in England of the introduction of tafamidis. The annual cost of tafamidis starts at [redacted] in year 1 and increases progressively with the increase in patients treated to [redacted] in year 5. The cumulative budget impact in year 1 to 5 is [redacted].

Table 35: Budget impact estimates (adapted from Table 33-5 p100-1 of MS)

	Year 1	Year 2	Year 3	Year 4	Year 5
Prevalence	17	27	37	30	30
Incidence	10	10	10	10	10
Patients who have moved to stage 2	-	-	17	10	10
Total eligible patients	27	37	30	30	30
Treatment uptake	25%	30%	40%	50%	60%
Prevalent treated patients	4	7	10	10	12
Incident treated patients	3	3	4	5	6
Patients who discontinue (due to progression to stage 2)	0	0	4	3	3
Total patients treated	7	10	10	12	15
Tafamidis budget impact					
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

The assumptions used by the manufacturer may underestimate the budget impact to the NHS for a number of reasons. Firstly, patients are assumed to take tafamidis for 3 years, which corresponds to duration of patients in stage 1 according to the economic model if a patient enters at symptom onset. However, some patients may remain in stage 1 for longer. Alternatively, given that the clinicians contacted by the ERG considered that it may not be possible to distinguish between stage 1 and 2 in the UK patient population, given the Coutinho staging may not be appropriate in the UK patient population, patients may not discontinue tafamidis upon progression to stage 2. Therefore, more patients may be treated with tafamidis than those estimated by the manufacturer. Secondly, treatment uptake is only 25% at year 1 and progressively increases to 60% at year 5. Since tafamidis is the only licensed treatment for TTR-FAP, it is plausible that all eligible patients will initiate treatment. Therefore, treatment uptake may be closer to 100%, which would have a considerable effect on budget impact. Thirdly, the manufacturer assumes that the acquisition of tafamidis will not incur VAT due to home delivery. However, home delivery arrangements are yet to be finalised and, as far as the ERG are aware, has not been guaranteed by the manufacturer.

As a result, the acquisition of tafamidis may be subject to VAT, which would increase budget impact by 20%. Lastly, the manufacturer did not take into account medication wastage as a result of discontinuation. If patients may discontinue tafamidis in the middle of the packet, or if delivery arrangements are such that patients may accumulate medication at home, the unused medication cannot be re-used by other patients and would still represent a cost to the NHS.

5.10 Summary of uncertainties and issues

The ERG considered that the manufacturer's economic evaluation were subject to a number of important uncertainties and issues which raise questions about the validity of the results presented. The main issues were:

- The consideration of a combined population of V30M and non-V30M patients in the base case. Not only is this an identifiable source of heterogeneity but many of the parameters used to reflect the combined population (e.g. survival curves), are based on different proportions of V30M and non-V30M patients. Therefore, the ERG does not consider the results presented in the manufacturer's submission for the base case a reliable estimate.

- The ERG is uncertain on the validity of the assumption of independence between disease severity and mortality in the manufacturer's model.

- It is unclear to the ERG whether TQoL is an appropriate measure for modelling disease severity in the patient population in England.

- The appropriateness of Coutinho disease stages for the patient population in England is unclear given Coutinho is based on early onset V30M patients, whilst most patients in England are non-V30M, and even those who are V30M are late onset.

- No evidence was provided on the appropriateness of using mPDS to map to Coutinho disease stages.

- No justification was provided by the manufacturer for the method used to create TQoL cut-off values for the Coutinho disease stages.

- The use of cross sectional data to model the relationship between disease duration and TQoL appears inappropriate to the ERG. The analysis would only be valid in a homogenous patient population, which is not the case with TTR-FAP. Further, the

use of disease duration, rather than TQoL, as the dependent variable is inappropriate.

- It is unclear on the comparability of the different data sources used, in particular the use of incident and prevalent populations for different parameters is a concern. This raises issues about the reliability of the results.
- The use of mortality data based on survival from symptom onset is inappropriate.
- The rate and benefit of liver transplantation in the model does not appear appropriate. Clinical advice suggested liver transplantation may not be a treatment option in the UK. Further, the evidence used to estimate the benefits of liver transplantation does not appear to the ERG to be comparable with the population in England. It is unclear to the ERG if the evidence on effectiveness from the Fx-005 trial is generalisable to the patient population in England.

6 ADDITIONAL WORK UNDERTAKEN BY THE ERG

6.1 Overview

As discussed in Section 5.8, in order to reduce computation times and facilitate the exploratory analysis, the ERG reconstructed the manufacturer's model using the same structure but as a cohort model rather than as an individual patient simulation model. The major issues noted with the manufacturer's model were corrected in the ERG's model, namely (i) the exclusion of the costs and benefits of the first 6 months, and (ii) the exclusion of costs and QALY losses due to liver transplantation for those patients who died in the same cycle as the procedure. The ERG's model was comprehensively checked by two researchers and compared with the manufacturer's model to ensure comparability. Details of the validation process can be found in Appendix 6.

In the following sections, the ERG presents an alternative base-case, informed by the critical appraisal of the manufacturer's submission and response to the points for clarification, together with input from expert clinical advice (Section 6.2). In addition, the ERG has undertaken further exploratory work to address several of the issues and uncertainties identified during the review of the manufacturer's submission. This additional work undertaken by the ERG includes:

- Exploratory scenario analyses examining alternative assumptions for stopping rules, stage cut-offs, rules for liver transplantation, costs and HRQoL (Section 6.3)
- An analysis on the relative weighting of QALY benefits to assess how much more the QALYs gained from tafamidis would need to be valued compared to QALYs for other treatments for tafamidis to be considered a cost-effective use of resources under various cost-effectiveness thresholds (based on the upper bound of the conventional NICE cost-effectiveness threshold, the upper bound of empirical estimates of society's willingness to pay for a QALY and the highest indicative ICER based on preliminary estimates previously reported by NICE associated with existing ultra-orphan drugs that are currently provided within the NHS) (Section 6.4).
- A set of sensitivity analyses on the budget impact based on different assumptions on treatment uptake (Section 6.5).

6.2 The ERG's base-case

The ERG's base-case compares tafamidis as an add-on therapy to conventional standard care alone from the perspective of the UK NHS and PSS. Consequently, only the costs falling on the UK NHS budget, namely health care resource use, and costs associated with PSS, are included (i.e. productivity costs have been excluded).

Results are presented separately for the V30M population and for the non-V30M population. As discussed in Section 2.2, ATTR is a very heterogeneous disease with patients experiencing different symptoms and rates of disease progression. Since the heterogeneity between V30M and non-V30M patients is observable and relates to the different mutation, the cost-effectiveness of tafamidis for each patient population can and should be analysed separately. Nonetheless, results are presented for a combined population of 16.7% of V30M patients and 83.3% of non-V30M patients in order to facilitate comparison with the manufacturer's results based on weighted average of the results for the V30M and the non-V30M populations. This weighted average avoids the use of inputs referring to combined populations with different proportions of V30M and non-V30M patients used for the manufacturer's combined base-case and discussed in Section 5.2.

The model structure is identical to that employed by the manufacturer in their submission (see **Error! Reference source not found.**). However, some of the input parameters and ssumptions differ, namely tafamidis stopping rules and the rate of liver transplantation. In contrast with the manufacturer's submission but following expert clinical advice, patients

remain on tafamidis until transition to stage 3 rather than discontinuing treatment at stage 2. The clinicians contacted by the ERG advised that the Coutinho et al (1980) stages may not be appropriate to classify the UK patient population, and are not currently used to classify UK patients. The Coutinho et al (1980) stages are based on the severity of the neuropathy and do not consider cardiac and autonomic involvement. Although disease progression in the Portuguese (endemic) V30M variant is evident from the degree of neuropathy, in the UK patients there may be considerable overlap in the neuropathy at stage 1 and stage 2 and disease progression is mainly evident in the autonomic and cardiac symptoms. Therefore, it may not be possible to distinguish between stage 1 and 2 in the UK patient population. As a result, the ERG's base-case assumes that patients remain on tafamidis throughout stage 1 and 2 and that tafamidis continues to reduce the TQoL rate of change. The clinicians also hypothesised that, given the heterogeneity in patients presenting with the disease, the variability in disease progression, and the assessment of response being unfeasible, treatment discontinuation may not take place upon transition to stage 3 in clinical practice. Sensitivity analyses explores the impact of continuing treatment tafamidis throughout the patient's lifetime, as well as discontinuing at progression to stage 2, as proposed by the manufacturer.

The rate of liver transplantation is assumed to be zero. The clinicians contacted by the ERG considered that liver transplantation is unlikely to be a therapeutic option for UK TTR-FAP patients, see section 5.3.3 for more discussion of this issue. Nonetheless, the ERG acknowledges the high degree of uncertainty on the issue. Therefore, sensitivity analyses explores the impact of varying the rate of liver transplantation on the cost-effectiveness results.

Table 36 compares the input parameters and assumptions in the manufacturer's base-case with the ERG's. The key differences with the manufacturer's base-case are: (i) the perspective for both costs and outcomes is that of the NHS & PSS, hence productivity costs are not included in the base-case; (ii) patients remain on tafamidis throughout stage 1 and 2 rather than discontinuing at progression to stage 2; and (iii) the rate of liver transplantation is assumed to be zero. In addition, results are presented for different baseline TQoL scores in order to explore the impact of disease severity as measured by TQoL scores in the cost-effectiveness of tafamidis. This was considered more appropriate than an average of all TQoLs as this is explainable heterogeneity in patients and should not be ignored.

Table 36: Comparison of model assumptions and key input parameters between the manufacturer's and the ERG's base-case.

Parameter/Assumption	Manufacturer's base-case	ERG's base-case
Perspective	Outcomes: NHS Costs: NHS & PSS and productivity costs to society.	Costs and outcomes: NHS & PSS. Productivity costs included in a scenario analysis.
Decision rules		
Starting rules	All patients at stage 1 are eligible for tafamidis.	Same.
Stopping rules	Tafamidis is discontinued once patients reach stage 2.	Tafamidis is discontinued once patients reach stage 3.
Patient population	Base-case: combined V30M and non-V30M. Subgroups: V30M and non-V30M separately.	Base-case: V30M and non-V30M as separate populations. Combined V30M and non-V30M presented for comparison.
Natural history		
Baseline TQoL	Combined V30M and non-V30M: 48.97 V30M: 49.68 Non-V30M: 44.89	Same for V30M and non-V30M. Combined population: 45.68 (weighted average) Results presented for different baseline TQoL scores (10, 20, 30 and 40).
Disease stages	Transition from stage 1 to 2 occurs when TQoL reaches 54 points. Transition from stage 2 to 3 occurs when TQoL reaches 91 points.	Same. Tested in sensitivity analysis.
TQoL rate of change	Stage 1: 7.17 points in 6 months. Stage 2: 3.761 points in 6 months. Stage 3: 1.022 points in 6 months.	Same. Tested in sensitivity analysis.
Mortality (without liver transplantation)	Weibull functions fitted to data presented in Sattianayagam et al (2012)[3]	Same. Tested in sensitivity analysis.
Liver transplantation	Patients in stage 1 are eligible for liver transplant. Rate of liver transplantation over 6 months is: Combined V30M and non-V30M: 5.72% V30M: 6.50% Non-V30M: 5.13%	No patients are eligible for liver transplant. Rate of liver transplantation is zero. Tested in sensitivity analysis.
Mortality post-liver transplantation	Weibull functions fitted to data presented by Herlenius et al (2004).[22]	Not applicable for base-case. Used in the sensitivity analysis.

Parameter/Assumption	Manufacturer's base-case	ERG's base-case
Benefits of liver transplantation	Liver transplantation halts disease progression, maintains HRQoL and increases survival.	Not applicable for base-case. Used in the sensitivity analysis.
Tafamidis effect		
Treatment effectiveness	Tafamidis reduces the rate of change in TQoL and therefore slows down disease progression while on treatment.	Same.
Adverse effects	Not considered.	Same.
Discontinuation due to non-compliance or adverse effects.	Considered only in terms of costs. Patients who discontinue tafamidis continue to receive the benefits from treatment.	Not considered.
HRQoL		
Without liver transplantation	EQ-5D values depend on TQoL scores using mapping function.	Same. Alternative mapping functions tested in sensitivity analysis.
Liver transplantation	Liver transplantation is associated with one-off QALY decrement of 0.2.	Not applicable for base-case. Used in the sensitivity analyses.
Carer's HRQoL loss	A HRQoL loss of 0.01 is included in stage 3 to reflect the impact on carers.	Same.
Resource use and costs		
Associated with disease stages	Costs are estimated by applying UK unit costs to resource use obtained from Swedish clinicians.	Same.
Liver transplantation.	Costs provided by the National Specialist Commissioning Team.	Not applicable for base-case. Used in sensitivity analysis.
Productivity costs	Base-case includes productivity costs.	Not included in base-case. Included in sensitivity analysis.
Tafamidis cost	Tafamidis costs [redacted] per patient per year.	Same. Impact of wastage tested in sensitivity analyses.

6.2.1 Base-case results

Table 37 presents the cost-effectiveness results for the base-case populations. For both populations, tafamidis is more costly but also more effective than conventional standard therapy alone. The ICER for the V30M population is £1,074,450 per QALY gained, while the

ICER for the non-V30M population is £1,138,813 per QALY gained. The ICER for the combined population is £1,126,565 per QALY gained.

Table 37: Base-case results for tafamidis add-on therapy compared with conventional standard care (CST) alone and comparison with manufacturer's results

Intervention	Mean costs (£)	Mean QALYs	ICER (£/QALY)
V30M population			
CST	126,159	3.38	
Tafamidis	1,075,441	4.27	£1,074,450
Non-V30M population			
CST	79,466	2.58	
Tafamidis	743,561	3.16	£1,138,813
Combined population			
CST	87,248	2.71	
Tafamidis	798,874	3.35	1,126,565

Table 38 presents the base-case results for different baseline TQoL scores. For both populations, the ICER increases as baseline TQoL increases. For the V30M population, the base-case ICER (baseline TQoL = 49.64 points) is £1,074,450 per QALY gained, while the ICER for a baseline TQoL of 10 points is £663,229 per QALY gained. For the non-V30M population, the base-case ICER (baseline TQoL = 44.89 points) is £1,138,183 per QALY gained whereas the ICER for a baseline TQoL of 10 points is £778,141 per QALY gained. The TQoL rate of change is greatest in stage 1. Given treatment effectiveness is modelled through a hazard ratio on the rate of change of TQoL, there is a larger absolute gain in terms of reduced progression of TQoL the longer a patient remains in stage 1. Since patients entering the model with lower baseline remain in stage 1 for longer, the benefits from tafamidis are greater (see Section 5.4 for more discussion of the impact of treatment in the economic model). As a result, the cost-effectiveness of tafamidis appears more favourable for patients with lower baseline TQoL. These results suggest that the cost-effectiveness of tafamidis may be more favourable for patients who are identified earlier, or those with less severe neuropathic impairment. Note that in the single RCT evaluating tafamidis, which included exclusively V30M patients, the baseline TQoL score in the placebo group was 30.8 (SD=26.7) points and in the tafamidis group was 27.3 (SD=24.2) points. In Fx1A-201, a before-and-after study evaluating tafamidis in a cohort of non-V30M patients, the baseline TQoL score was 47.8 (SD=35.1) points. Although the prevalent patient population in England consists mainly of patients with a non-V30M variant of the disease, where the baseline TQoL score used in the base-case of 44.89 may be appropriate, the TQoL scores

for incident (new) patients may be lower. Therefore, the cost-effectiveness profile of tafamidis may be more favourable for incident patients, with milder disease, than for prevalent patients, with more advanced progression.

Table 38: Base-case results for different baseline TQoL scores

	Conventional standard care		Tafamidis		ICER (£/QALY)
TQoL	Costs (£)	QALYs	Costs (£)	QALYs	
V30M population					
Base-case: 49.64	126,159	3.38	1,075,441	4.27	1,074,450
10	88,571	4.32	1,013,203	5.72	663,229
20	99,260	4.09	1,020,647	5.31	756,057
30	107,234	3.81	1,033,172	4.91	836,106
40	119,096	3.63	1,051,967	4.56	1,010,156
Non-V30M population					
Base-case: 44.89	79,466	2.58	743,561	3.16	1,138,813
10	54,085	3.26	712,593	4.11	778,141
20	62,121	3.07	715,902	3.82	874,490
30	67,956	2.85	721,963	3.53	957,860
40	78,357	2.71	733,221	3.27	1,162,822
Combined population					
Base-case: 45.68	87,248	2.71	798,874	3.35	1,128,086
10	59,833	3.44	762,695	4.38	758,989
20	68,311	3.24	766,693	4.07	854,751
30	74,502	3.01	773,831	3.76	937,568
40	85,147	2.86	786,345	3.49	1,137,378

The cost-effectiveness results for the ERG's base-case are different from the manufacturer's results. The manufacturer's ICER for the V30M population is £174,634 per QALY gained, markedly lower than the ERG's ICER. Similarly, the manufacturer's ICER for the non-V30M population is £304,293 per QALY gained. In order to understand the reasons for the differences in results between the manufacturer's and the ERG's results, the section below (Section 6.3) uses a series of alternative scenarios to compare and contrast the different assumptions and parameter inputs used in both models. In addition, sensitivity analyses over a range of alternative parameter values are used to explore any remaining areas of uncertainty.

6.3 Exploratory scenario analysis

The ERG tested a number of alternative scenarios in order to i) identify the key drivers of the cost-effectiveness results and ii) determine the main areas of uncertainty. Table 39 summarizes the alternative scenarios considered. For each element, the rationale for the change and the description of the methodology applied are provided.

Table 39: Scenarios considered in the exploratory analysis, rationale and description

Topic	Scenario	Rationale	Description
Perspective	1. Inclusion of productivity costs.	Productivity costs accrued by patients and carers may represent a significant burden.	The productivity costs presented by the manufacturer are included in the analysis. • Stage 1: £2,514 per 6 months; • Stage 2: £8,238 per 6 months; • Stage 3: £8,238 per 6 months.
Duration of treatment	2. Lifetime treatment duration.	The clinicians contacted by the ERG felt that treatment discontinuation may be unfeasible even at progression to stage 3.	Tafamidis treatment is continued throughout lifetime.
	3. Treatment only during stage 1.	The marketing authorisation specifies that tafamidis is indicated for stage 1 of the disease.	Tafamidis treatment is discontinued at progression to stage 2.
Disease staging	4A. Cut-off TQoL score between stages defined as half-way between the mean TQoL score of stage N and stage N+1	The manufacturer did not provide a justification for the definition of cut-off TQoL scores between stages. Alternative definitions may be appropriate.	Cut-off TQoL scores between stages: • Stage 1 → Stage 2: 59 for V30M; 66 for non-V30M. • Stage 2 → Stage 3: 82 for V30M; 90 for non-V30M.
	4B. Cut-off TQoL score between stages defined as half-way between the mean TQoL score of stage N and stage N+1 AND TQoL rate of change recalculated	The TQoL rate of change depends on the cut-off TQoL scores between stages.	TQoL rate of change for V30M (over 6 months): • Stage 1: 6.3355 • Stage 2: 4.3771 • Stage 3: 1.1308 TQoL rate of change for non-V30M (over 6 months): • Stage 1: 6.3412 • Stage 2: 3.2440

			• Stage 3: 1.0334
	5A. Cut-off TQoL score between stages defined as the mean of stage N+1	See Scenario 4A.	Cut-off TQoL scores between stages: • Stage 1 → Stage 2: 69 for V30M; 86 for non-V30M. • Stage 2 → Stage 3: 95 for V30M; 93 for non-V30M.
	5B. Cut-off TQoL score between stages defined as the mean of stage N+1 AND TQoL rate of change recalculated	See Scenario 4B.	TQoL rate of change for V30M (over 6 months): • Stage 1: 6.2720 • Stage 2: 2.7764 • Stage 3: 0.9768 TQoL rate of change for non-V30M (over 6 months): • Stage 1: 5.2715 • Stage 2: 2.1668 • Stage 3: 0.9991
TQoL rate of change	6. TQoL rate of change for stage 1 is that observed in the placebo group of Fx-005	Fx-005 enrolled patients with V30M variant at stage 1, which may be similar to some of the patients in England.	TQoL rate of change for both populations: Stage 1: 2.7708
	7. TQoL rate of change independent of stage	See Scenario 4A and 4B.	TQoL rate of change for both populations across all stages =2.084 points/6 months.
Mortality (without liver transplant)	8. Doubled mortality risk	The parametric curve used to model the Kaplan-Meier data from Sattianayagam et al (2011) appears not to fit well the non-V30M curve.	The scale parameter of the Weibull function was doubled for both populations.
Liver transplantation	9. Patients are eligible for liver transplantation during stage 1 at the rate used in the manufacturer's submission.	There is considerable uncertainty regarding the eligibility of patients in England to liver transplantation.	As described in the Scenario column.

	10. Patients are eligible for liver transplant during stages 1 and 2 at the rate used in the manufacturer's submission.	See Scenario 9.	As described in the Scenario column.
	11. Patients are eligible for liver transplant during lifetime at the rate used in the manufacturer's submission.	See Scenario 9.	As described in the Scenario column.
Tafamidis costs	12. Acquisition costs of the drug increased by 20%.	The manufacturer proposed to deliver tafamidis to patients' addresses, which would avoid VAT. However, if the acquisition of tafamidis is subject to VAT, its acquisition costs would increase by 20%. Conversely, if patients discontinue tafamidis with some medicine still at home, the wasted medication will represent a loss to the NHS.	As described in the Scenario column.
HRQoL	13. Quadratic mapping function from TQoL to EQ-5D	EQ-5D scores are calculated from TQoL scores using a mapping function. However, alternative mapping functions may provide a better fit.	As described in the Scenario column.
	14. Cubic mapping function from TQoL to EQ-5D	See Scenario 13.	As described in the Scenario column.
Manufacturer's Assumptions	15. Using the manufacturer's assumptions regarding liver transplantation and stopping rules for tafamidis.	There is uncertainty regarding eligibility for liver transplantation and stopping rules for tafamidis.	Tafamidis is discontinued upon progression to stage 2. Patients are eligible for liver transplant at the rate proposed by the manufacturer: 6.5% over 6-months for V30M and 5.13% over 6-months for non-V30M.

Table 40 summarizes the results of the sensitivity analyses on incremental costs and health benefits, and compares them with ERG's base-case. Appendix 7 provides details on the calculations required for the analyses and Appendix 8 presents the full results of the sensitivity analysis by different baseline TQoL scores. The scenarios with the greatest impact on the expected costs and health benefits are those testing the assumptions regarding stopping rules, disease staging and TQoL rate of change and liver transplant. Scenario 3, where tafamidis is discontinued once patients progress to stage 2, reduces the ICER from £1,074,450 to £635,218 per QALY gained for the V30M population and from £1,138,813 to £834,830 per QALY gained for the non-V30M population. There is a reduction in the expected benefits because tafamidis is discontinued at progression to stage 2, and therefore disease progression, as measured by TQoL, is not reduced during this stage and returns to the natural rate of progression. Also for this reason, the acquisition costs of tafamidis are accrued for a shorter period of time and expected costs decrease. In addition, since TQoL rate of change is greatest in stage 1, and although the relative effect of tafamidis on TQoL progression is assumed constant throughout stages when a patient is on treatment, the greatest absolute effect on TQoL scores is felt at stage 1, making stage 1 the most beneficial period to use tafamidis. Therefore, assuming that tafamidis is used only during stage 1 reduces the ICER and favours the cost-effectiveness results towards tafamidis.

The impact of disease stages and TQoL rate of change in the cost-effectiveness results is also evident in Scenario 7, where TQoL rate of change was assumed constant and independent from disease stages. The ICER increased considerably from £1,074,450 to £1,608,195 per QALY gained for the V30M population and from £1,138,813 to £1,915,439 per QALY gained for the non-V30M population. The expected costs remained approximately the same because the acquisition cost of tafamidis is much greater compared with the costs associated with the disease itself and as tafamidis is given for a similar period of time, i.e. stages 1 and 2. The reduction in expected benefits is the result of the lower TQoL rate of change during stages 1 and 2, which translates into a lower absolute benefit derived from tafamidis.

As discussed in Section 5.2, the estimation of the TQoL rate of change is fraught with assumptions and methodological issues, namely: (i) the data used to model the relationship between disease duration and TQoL was from a cross-sectional sample of Portuguese V30M patients, which may not be representative of V30M or non-V30M patients in the UK; (ii) disease duration was modelled as the dependent variable and TQoL as the explanatory variable despite causality suggesting that disease duration determines TQoL and not the opposite; and (iii) the cut-off TQoL scores for crossover between stages were used to determine time in each stage, although the definition of TQoL cut-offs was not justified by the

manufacturer. As demonstrated in Scenarios 4 to 7, alternative assumptions to estimate the TQoL rate of change have an impact on the cost-effectiveness of tafamidis. Given the absence of justification for the choice of cut-off TQoL scores between stages and the methodological issues mentioned above, the ERG is unable to determine which set of assumptions best represents the natural history of TRR-FAP, and therefore which of the corresponding cost-effectiveness results are the most appropriate reflection of the cost-effectiveness of tafamidis. However, as discussed previously, TQoL might not be an appropriate measure for capturing disease severity in these patients, and none of the scenarios described above have been able to examine this issue.

Liver transplantation is another key driver of cost-effectiveness in the model. Scenarios 9 and 10 explore the impact of including the option of liver transplantation at the different disease stages. Across the scenarios, the ICERs decreased considerably by 34% to 64% depending on the patient population and assumption tested. Including the option of liver transplantation at the rate proposed by the manufacturer has a dual effect. The expected costs reduce because, once liver transplantation occurs, the patient discontinues tafamidis and disease progression is halted. The expected health benefits increase, firstly as a result of the reduced mortality risk experienced by patients post-liver transplantation in comparison to the risk pre-liver transplantation, and secondly because liver transplantation is assumed to halt TQoL progression, hence maintaining EQ-5D constant over the patient's remaining lifetime. The largest reduction occurred for Scenario 10, in which patients are assumed eligible for liver transplantation during stages 1 and 2. The ICER reduced from £1,074,450 to £383,233 per QALY gained for the V30M population and from £1,138,813 to £657,439 per QALY gained for the non-V30M population. Scenario 10 presents the largest reduction in the ICER as a result of the combined effect of tafamidis and liver transplantation. Since tafamidis reduces the TQoL rate of change, patients remain in the earlier stages of the disease for longer. Therefore, patients are eligible for liver transplantation for a longer period of time than compared with CST, which results in a more favourable cost-effectiveness profile for tafamidis.

The lowest ICER is obtained in Scenario 15, which uses the manufacturer's assumptions regarding liver transplantation and stopping rules for tafamidis. Assuming that tafamidis is discontinued upon progression to stage 2, which assumes that stage 2 is distinguishable in these patients, and that patients are eligible for liver transplant during stage 1 at a 6-monthly rate of 6.5% for V30M patients and 5.13% for non-V30M patients reduces the expected costs and increases the expected health benefits considerably. For the V30M population, the ICER is reduced from £1,074,450 to £214,197 per QALY gained and for the non-V30M population from £1,138,813 to £427,561 per QALY gained. If patients are eligible for liver

transplant during stage 1 and 2, the ICER for the V30M population increases to £397,314 per QALY gained and for the non-V30M population for £615,193 (not shown). Therefore, these results emphasize the importance of these assumptions on the cost-effectiveness of tafamidis.

Table 40: Cost-effectiveness results for scenario analysis

	Scenario	Patient Population	Inc. Costs (£)	% Change	Inc. QALYs	% Change	ICER (£/QALY)	% Change
Base-case	Not applicable (NA)	V30M	949,282	NA	0.88	NA	1,074,450	NA
		Non-V30M	664,095	NA	0.58	NA	1,138,813	NA
		Combined	£711,626	NA	0.63	NA	£1,129,565	NA
Perspective	1.Inclusion of productivity costs	V30M	936,190	-1.38%	0.84	-4.55%	1,108,075	3.13%
		Non-V30M	649,697	-2.17%	0.58	0.00%	1,114,123	-2.17%
		Combined	697,446	-1.99%	0.62	-1.06%	1,118,897	-0.94%
Duration of treatment	2. Lifetime treatment duration	V30M	949,491	0.02%	0.88	0.00%	1,074,673	0.02%
		Non-V30M	664,712	0.09%	0.58	0.00%	1,139,713	0.08%
		Combined	712,175	0.08%	0.63	0.00%	1,130,437	0.08%
	3. Treatment only during stage 1	V30M	177,563	-81.30%	0.28	-68.18%	635,218	-40.88%
		Non-V30M	273,449	-58.82%	0.33	-43.10%	834,830	-26.69%
		Combined	257,468	-63.82%	0.32	-48.94%	800,419	-29.14%
Disease staging	4A. Cut-off TQoL score between stages defined as half-way between the mean TQoL score of stage N and stage N+1	V30M	925,427	-2.51%	0.77	-12.50%	1,206,863	12.32%
		Non-V30M	645,475	-2.80%	0.57	-1.72%	1,130,235	-0.75%
		Combined	692,134	-2.74%	0.60	-4.23%	1,147,183	1.56%
	4B. Cut-off TQoL score between stages defined as half-way between the mean TQoL score of stage N and stage N+1 AND TQoL rate of change recalculated	V30M	946,492	-0.29%	0.95	7.95%	994,714	-7.42%
		Non-V30M	662,764	-0.20%	0.53	-8.62%	1,257,552	10.43%
		Combined	710,052	-0.22%	0.60	-4.76%	1,183,420	4.77%
	5A. Cut-off TQoL score between stages defined as the mean of stage N+1	V30M	921,518	-2.92%	0.92	4.55%	1,000,815	-6.85%
		Non-V30M	638,962	-3.78%	0.67	15.52%	951,363	-16.46%
		Combined	686,055	-3.59%	0.71	12.96%	964,011	-14.66%
	5B. Cut-off TQoL score between stages defined as the mean of stage N+1	V30M	954,657	0.57%	0.74	-15.91%	1,284,583	19.56%
		Non-V30M	663,238	-0.13%	0.39	-32.76%	1,693,232	48.68%

	Scenario	Patient Population	Inc. Costs (£)	% Change	Inc. QALYs	% Change	ICER (£/QALY)	% Change
Base-case	Not applicable (NA)	V30M	949,282	NA	0.88	NA	1,074,450	NA
		Non-V30M	664,095	NA	0.58	NA	1,138,813	NA
		Combined	£711,626	NA	0.63	NA	£1,129,565	NA
	AND TQoL rate of change recalculated	Combined	711,808	0.03%	0.45	-28.84%	1,587,675	40.56%
TQoL rate of change	6. TQoL rate of change for stage 1 is that observed in the placebo group of Fx-005	V30M	941,873	-0.78%	0.86	-2.27%	1,097,917	2.18%
		Non-V30M	654,876	-1.39%	0.51	-12.07%	1,289,131	13.20%
		Combined	702,709	-1.25%	0.57	-9.79%	1,236,438	9.46%
	7. TQoL rate of change independent of stage (rate=2.084 points/6 months)	V30M	948,319	-0.10%	0.59	-32.95%	1,608,195	49.68%
		Non-V30M	658,149	-0.90%	0.34	-41.38%	1,915,439	68.20%
		Combined	706,511	-0.72%	0.38	-39.42%	1,851,120	63.88%
Mortality (pre-liver transplant)	8. Doubled mortality risk	V30M	761,048	-19.83%	0.61	-30.68%	1,231,949	14.66%
		Non-V30M	444,204	-33.11%	0.30	-48.28%	1,501,709	31.87%
		Combined	497,011	-30.16%	0.35	-44.18%	1,413,302	25.12%
Liver transplantation	9. Rate of liver transplantation as manufacturer's submission	V30M	840,586	-0.01379	1.30	47.73%	645,281	-39.94%
		Non-V30M	556,626	-16.18%	0.79	36.21%	715,273	-37.19%
		Combined	603,953	-15.13%	0.88	38.89%	690,232	-38.89%
	10. Patients are eligible for liver transplantation in stage 1 and 2	V30M	578,491	-39.06%	1.51	71.59%	382,233	-64.43%
		Non-V30M	475,731	-28.36%	0.72	24.14%	657,439	-42.27%
		Combined	492,858	-30.74%	0.85	35.19%	578,698	-48.77%
	11. Patients are eligible for liver transplantation throughout lifetime	V30M	572,597	-39.68%	1.12	27.27%	511,076	-52.43%
		Non-V30M	474,665	-28.52%	0.63	8.62%	752,880	-33.89%
		Combined	490,987	-31.00%	0.71	12.96%	689,911	-38.92%

	Scenario	Patient Population	Inc. Costs (£)	% Change	Inc. QALYs	% Change	ICER (£/QALY)	% Change
Base-case	Not applicable (NA)	V30M	949,282	NA	0.88	NA	1,074,450	NA
		Non-V30M	664,095	NA	0.58	NA	1,138,813	NA
		Combined	£711,626	NA	0.63	NA	£1,129,565	NA
Tafamidis costs	12. Acquisition costs of the drug increased by 20% to account for potential wastage or VAT effect	V30M	1,144,467	20.56%	0.88	0.00%	1,295,370	20.56%
		Non-V30M	801,317	20.66%	0.58	0.00%	1,374,126	20.66%
		Combined	858,509	20.64%	0.63	0.00%	1,362,712	20.64%
HRQoL	13. Quadratic mapping function from TQoL to EQ-5D	V30M	949,282	0.00%	1.10	25.00%	865,908	-19.41%
		Non-V30M	664,095	0.00%	0.71	22.41%	941,271	-17.35%
		Combined	711,626	0.00%	0.78	23.02%	918,227	-18.71%
	14. Cubic mapping function from TQoL to EQ-5D	V30M	949,282	0.00%	1.10	25.00%	865,945	-19.41%
		Non-V30M	664,095	0.00%	0.67	15.52%	991,874	-12.90%
		Combined	711,626	0.00%	0.74	17.72%	959,496	-15.06%
Manufacturer's assumptions	15. Using the manufacturer's assumptions regarding liver transplantation and stopping rules for tafamidis.	V30M	165,929	-82.52%	0.77	-12.50%	214,197	-80.06%
		Non-V30M	250,181	-62.33%	0.59	1.72%	427,561	-62.46%
		Combined	236,139	-66.82%	0.62	-1.59%	380,869	-66.28%

6.4 Analysis of relative weighting of QALY benefits

In the absence of a pre-defined cost-effectiveness threshold to which to compare the ICER for tafamidis with, the ERG estimated how much more would the incremental QALY benefits associated with tafamidis need to be valued relative to QALY benefits for other treatments for tafamidis to be considered cost-effective according to three alternative cost-effectiveness thresholds, £30,000/QALY, £70,000/QALY and £391,000/QALY. Each threshold represents the willingness to pay for an additional QALY. The threshold of £30,000 per QALY corresponds to the willingness to pay for an additional QALY in the NHS and corresponds to the upper bound of the range used by NICE.[52] The threshold of £70,000 per QALY gained was estimated as the upper bound to the society's willingness to pay for an additional QALY.[65]The threshold of £391,000 per QALY gained corresponds to the highest ICER associated with a product currently in use for ultra-orphan diseases. [66] The threshold of £391,000 per QALY gained implies that the extra QALY obtained for that particular ultra-orphan disease is valued thirteen times greater than a QALY obtained in the rest of the NHS (£391,000/£30,000=13).

Table 41 presents the results for the analysis of relative weighting of additional QALY benefits for the base-case and scenarios. Under the base-case, for tafamidis to be considered cost-effective at a threshold of £30,000 per QALY gained (corresponding upper bound of the ICER used by NICE to the NHS), the QALY benefits associated with treatment would need to be valued at around thirty-five times over QALY benefits obtained in other diseases treated in the NHS. Applying the threshold of £70,000 per QALY gained would require valuing the QALY benefits associated with tafamidis at around fifteen to sixteen times the societal value for a QALY. Finally, for tafamidis to be considered cost-effective under the threshold of £391,000, the QALY benefits would need to be valued almost three times more than the benefits obtained from the ultra-orphan drug with the ICER of £391,000, on which the threshold is based.

Table 41: Analysis on the relative weighting of additional QALY benefits

	Scenario	Patient Population	ICER (£/QALY)	Relative weight for threshold		
				£30,000	£70,000	£391,000
Base-case		V30M	1,074,450	35.96	15.41	2.76
		Non-V30M	1,138,813	38.17	16.36	2.93
		Combined	£1,129,565	37.65	16.14	2.89
Perspective	1	V30M	1,108,075	37.15	15.92	2.85
		Non-V30M	1,114,123	37.34	16.00	2.86
		Combined	1,118,897	37.50	16.07	2.88

	Scenario	Patient Population	ICER (£/QALY)	Relative weight for threshold		
				£30,000	£70,000	£391,000
Duration of treatment	2	V30M	1,074,673	35.97	15.41	2.76
		Non-V30M	1,139,713	38.20	16.37	2.93
		Combined	1,130,437	37.68	16.15	2.89
	3	V30M	635,218	21.14	9.06	1.62
		Non-V30M	834,830	27.62	11.84	2.12
		Combined	800,419	26.82	11.49	2.06
Disease staging	4A	V30M	1,206,863	40.06	17.17	3.07
		Non-V30M	1,130,235	37.75	16.18	2.90
		Combined	1,147,183	38.45	16.48	2.95
	4B	V30M	994,714	33.21	14.23	2.55
		Non-V30M	1,257,552	41.68	17.86	3.20
		Combined	1,183,420	39.45	16.91	3.03
	5A	V30M	1,000,815	33.39	14.31	2.56
		Non-V30M	951,363	31.79	13.62	2.44
		Combined	964,011	32.21	13.80	2.47
	5B	V30M	1,284,583	43.00	18.43	3.30
		Non-V30M	1,693,232	56.69	24.29	4.35
		Combined	1,587,675	52.73	22.60	4.05
TQoL rate of change	6	V30M	1,097,917	36.51	15.65	2.80
		Non-V30M	1,289,131	42.80	18.34	3.28
		Combined	1,236,438	41.09	17.61	3.15
	7	V30M	1,608,195	53.58	22.96	4.11
		Non-V30M	1,915,439	64.52	27.65	4.95
		Combined	1,851,120	61.97	26.56	4.76
Mortality (pre-liver transplant)	8	V30M	1,231,949	41.59	17.82	3.19
		Non-V30M	1,501,709	49.36	21.15	3.79
		Combined	1,413,302	47.33	20.29	3.63
Liver transplantation	9	V30M	645,281	21.55	9.24	1.65
		Non-V30M	715,273	23.49	10.07	1.80
		Combined	690,232	22.88	9.80	1.76
	10	V30M	382,233	12.77	5.47	0.98
		Non-V30M	657,439	22.02	9.44	1.69
		Combined	578,698	19.33	8.28	1.48
	11	V30M	511,076	17.04	7.30	1.31
		Non-V30M	752,880	25.11	10.76	1.93
		Combined	689,911	23.05	9.88	1.77
Tafamidis costs	12	V30M	1,295,370	43.35	18.58	3.33
		Non-V30M	1,374,126	46.05	19.74	3.53

	Scenario	Patient Population	ICER (£/QALY)	Relative weight for threshold		
				£30,000	£70,000	£391,000
		Combined	1,362,712	45.42	19.47	3.49
HRQoL	13	V30M	865,908	28.77	12.33	2.21
		Non-V30M	941,271	31.18	13.36	2.39
		Combined	918,227	30.41	13.03	2.33
	14	V30M	865,945	28.77	12.33	2.21
		Non-V30M	991,874	33.04	14.16	2.54
		Combined	959,496	32.06	13.74	2.46
Manufacturer's assumptions	15	V30M	214,197	7.18	3.08	0.55
		Non-V30M	427,561	14.13	6.06	1.08
		Combined	380,869	12.70	5.44	0.97

6.5 Sensitivity analysis on budget impact

As discussed in Section 5.9, the manufacturer made a number of assumptions to estimate the budget impact to the NHS. In order to determine the main drivers,

Table 42 summarises the alternative scenarios considered. For each element, the assumption tested and the rationale for the change are provided.

Table 42: Scenarios for budget impact sensitivity analysis

No.	Assumption	Scenario	Rationale
1	Patients discontinue tafamidis upon transition to stage 2.	Patients remain on tafamidis during stage 1 and 2 (=8.5 years)	Corresponds to ERG's base-case for cost-effectiveness analysis.
2	Treatment uptake by incident patients is 25% in the first year, increasing to 60% by year 5.	All patients initiate treatment.	Since tafamidis is the only licensed treatment for TTR-FAP, it is plausible that all eligible patients initiate treatment.
3	Tafamidis will be delivered to patients' addresses, thereby avoiding the VAT impact on costs.	The acquisition of tafamidis incurs VAT at 20%.	Simulates two alternative situations: Tafamidis incurs VAT. Tafamidis delivered to patients' addresses but at a rate that allows for accumulation and considerable wastage due to discontinuation.
4	Scenarios 1, 2 and 3 combined		Worst-case scenario for budget impact.

Table 43 presents the results of the sensitivity analysis to the budget impact and Appendix 9 provides full details on the calculations. Under the manufacturer's assumptions, the budget

impact to the NHS is [redacted] in the first year, increasing to [redacted] in year 5, at a cumulative impact of [redacted] over five years. Changing each assumption individually resulted in an increase in the cumulative impact over 5 years from [redacted] under Scenario 1 (in which patients remained on tafamidis throughout stages 1 and 2), to [redacted] under Scenario 2 (in which all prevalent and incident patients initiate tafamidis) under Scenario 2, and to [redacted] under Scenario 3 (in which the cost of tafamidis is increased by 20%). Scenario 4 is the result of combining the Scenarios 1 to 3 in a single analysis. The cumulative budget impact for Scenario 4 is [redacted], which is 335% greater than the manufacturer's estimate. However, any of these Scenarios are likely to be overestimates, since none accounts for discontinuation from treatment due to side-effects or patient's preferences. Nevertheless, these figures suggest that the budget impact to the NHS may be greater than that estimated in the manufacturer's submission.

Table 43: Sensitivity analysis to budget impact

Scenario	**Annual impact (£,000)**					**Cumulative impact (£,000)**	
	Year 1	Year 2	Year 3	Year 4	Year 5	Over 5 years	**% change from base-case**
Manufacturer's estimate	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	Not applicable
Scenario 1	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	39%
Scenario 2	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	185%
Scenario 3	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	20%
Scenario 4	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	335%

References

1. National Amyloidosis Centre. *Amyloidosis overview*. National Amyloidosis Centre; [cited 2012 January 24]. Available from: http://www.ucl.ac.uk/silva/medicine/amyloidosis/nac/overview.
2. Plante-Bordeneuve V, Said G. Familial amyloid polyneuropathy. *Lancet Neurol* 2011;**10**:1086-97.
3. Sattianayagam PT, Hahn AF, Whelan CJ, Gibbs SDJ, Pinney JH, Stangou AJ, et al. Cardiac phenotype and clinical outcome of familial amyloid polyneuropathy associated with transthyretin alanine 60 variant. *Eur Heart J* 2012;**33**:1120-7.
4. European Medicines Agency. *Recommendation for maintenance of orphan designation at the time of marketing authorisation. Vyndaqel (N-methyl D-(2,3,4,5,6-pentahydroxy-hexyl)-ammonium; 2-(3,5-dichloro-phenyl)-benzoxazole-6-carboxylate) for treatment of familial amyloid polyneuropathy*. European Medicines Agency; 2011. [cited 2012 January 23]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Committee_meeting_report/2011/12/WC500118738.pdf.
5. *Prevalence of rare diseases: bibliographic data. Orphanet Report Series, Rare Diseases collection, Number 1*; 2012. Available from: http://www.orpha.net/orphacom/cahiers/docs/GB/Prevalence_of_rare_diseases_by_alphabetical_list.pdf
6. National Horizon Scanning Centre. *Tafamidis for transthyretin amyloid polyneuropathy*. Birmingham: National Horizon Scanning Centre (NHSC); 2010.
7. Coutinho P, da Silva AM, Lima JL, Barbosa AR. Forty years of experience with type 1 amyloid neuropathy: review of 483 cases. In: Glenner GG, e Costa PP, de Freitas AF, editors. *Amyloid and amyloidosis: proceedings of the Third International Symposium on Amyloidosis, Povoa de Varzim, Portugal, 23-28 September 1979*. Amsterdam: Excerpta Medica; 1980. p. 88-98.
8. Koike H, Tanaka F, Hashimoto R, Tomita M, Kawagashira Y, Iijima M, et al. Natural history of transthyretin Val30Met familial amyloid polyneuropathy: analysis of late-onset cases from non-endemic areas. *J Neurol Neurosurg Psychiatry* 2012;**83**:152-8.
9. Wilczek HE, Stangou AJ, Suhr O, Larsson ME, Ericzon BG. Liver transplantation for familial amyloidosis: the 20-year report from the familial amyloidotic polyneuropathy world transplant registry (FAPWTR). *Hepatology* 2011;**54**:413A-4A.
10. Stangou AJ, Hawkins PN. Liver transplantation in transthyretin-related familial amyloid polyneuropathy. *Curr Opin Neurol* 2004;**17**:615-20.
11. Adams D, Antonini T, Lozeron P, Azoulay D, Mincheva Z, Lacroix C, et al. Long-term effect of liver transplantation on the neuropathy and survival in TTR (familial amyloid polyneuropathy) FAP: a prospective monocentric study in 167 patients. *Eur J Neurol* 2011;**18**:28.
12. Adams D, Lozeron PP, Antonini T, Kreib AM, Azoulay D, Ducot B, et al. Long term effect of liver transplantation (LT) on the neuropathy and survival in met30 transthyretin familial amyloid polyneuropathy (met30TTR-FAP). *Neurology* 2007;**68**:A406.
13. Singer R, Mehrabi A, Schemmer P, Kashfi A, Hegenbart U, Goldschmidt H, et al. Indications for liver transplantation in patients with amyloidosis: a single-center experience with 11 cases. *Transplantation* 2005;**80**:S156-9.
14. Stangou AJ, Hawkins PN, Booth DR, O'Grady J, Jewitt D, Rela M, et al. Liver transplantation for non-Met30 TTR associated familial amyloid polyneuropathy: the UK experience. *Hepatology* 1999;**30**:304A-A.
15. Stangou AJ, Heaton ND, Rela M, Jewitt D, Mathias CJ, O'Grady J, et al. Orthotopic liver transplantation for familial amyloid polyneuropathy: the UK experience. In: *VIIIth International Symposium on Amyloidosis*; 1998 7-11 August 1998; Rochester, MN. 1998. p. 238-40.

16. Steen L, Holmgren G, Suhr O, Wikstrom L, Groth CG, Ericzon BG. Worldwide survey of liver-transplantation in patients with familial amyloidotic polyneuropathy. *Amyloid* 1994;**1**:138-42.
17. Suhr OB, Ando Y, Holmgren G, Wikstrom L, Friman S, Herlenius G, et al. Liver transplantation in familial amyloidotic polyneuropathy (FAP). A comparative study of transplanted and non-transplanted patient's survival. *Transpl Int* 1998;**11** S160-3.
18. Yamashita T, Ando Y, Okamoto S, Yohei M, Hitahara T, Ueda M, et al. Effect of liver transplantation on the survival of patients with ordinary onset FAP in Japan. *Amyloid* 2010;**17**:131.
19. Peripheral neuropathies: liver transplant improves long-term survival rates in individuals with familial amyloid polyneuropathy. *Nat Rev Neurol* 2012;**8**:180.
20. Zeldenrust S, Pem R, Brandhagen D, Rosen C, Poterucha JJ, Gertz M, et al. Liver transplantation for familial amyloidotic polyneuropathy does not prevent disease progression in a majority of patients. *Amyloid* 2004:351-53.
21. *Familial Amyloidotic Polyneuropathy World Transplant Register and the Domino Transplant Register: update on behalf of reporting members of the FAPWTR*; 2009.
22. Herlenius G, Wilczek HE, Larsson M, Ericzon B. Ten years of international experience with liver transplantation for familial amyloidotic polyneuropathy: results from the familial amyloidotic polyneuropathy world transplant registry. *Transplantation* 2004;**77**:64-71.
23. European Medicines Agency. *Vyndaqel. Annex I: Summary of product characteristics*. [cited 2012 July 3]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/002294/WC500117862.pdf.
24. European Medicines Agency. *Assessment report: Vyndaqel - tafamidis meglumine*. European Medicines Agency; 2011. [cited 2012 January 23]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/002294/WC500117838.pdf.
25. Berk JL, Dyck PJ, Obici L, Zeldenrust SR, Sekijima Y, Yamashita T, et al. The diflunisal trial: update on study drug tolerance and disease progression. *Amyloid* 2011;**18**:196-7.
26. Obici L, Cortese A, Lozza A, Lucchetti J, Gobbi M, Palladini G, et al. Doxycycline plus tauroursodeoxycholic acid for transthyretin amyloidosis: a phase II study. *Amyloid* 2012;**19** 34-6.
27. Said G, Grippon S, Kirkpatrick P. Tafamidis. *Nat Rev Drug Discov* 2012;**11**:185-6.
28. Lozeron P, Theaudin M, Lacroix C, Adams D. Evaluation of tafamidis efficiency in moderately severe Met 30 TTR-familial amyloid polyneuropathies (FAP). *Eur J Neurol* 2011;**18**:323.
29. Russo M, Stancanelli C, Gentile L, Toscano A, Vita G, Mazzeo A. Safety and tolerability of orally administered tafamidis meglumine in TTR FAP: preliminary data at 3-month follow-up. *J Peripher Nerv Syst* 2012;**17**:S48-S9.
30. Coelho T, Maia L, Martins da Silva A, Waddington Cruz M, Plante-Bordeneuve V, Lozeron P, et al. Tafamadis (Fx-1006A): a first-in-class disease-modifying therapy for transthyretin familial amyloid polyneuropathy. *Neurology* 2010;**74**:A286.
31. *Safety and Efficacy Study of Fx-1006A in Patients With Familial Amyloidosis*. Available from: http://ClinicalTrials.gov/show/NCT00409175
32. Coelho T, Maia L, Martins Da Silva A, Waddington-Cruz M, Plante-Bordeneuve V, Suhr OB, et al. A landmark clinical trial of a novel small molecule transthyretin stabilizer, Fx-1006A, in patients with TTR amyloid polyneuropathy: A phase II/III randomized, double-blind, placebo-controlled study. *J Peripher Nerv Syst* 2009;**14**:35-6.
33. Waddington-Cruz M, Coelho T, Maia L, da Silva AM, Plante-Bordeneuve V, Suhr O, et al. A landmark clinical trial of a novel small molecule transthyretin (TTR) stabiliser, Fx-1006A, in patients with TTR amyloid polyneuropathy: a phase II/III, randomised, double-blind, placebo-controlled study. *J Neurol* 2008;**255**:107.
34. Coelho T, Waddington-Cruz M, Plante-Bordeneuve V, Cros D, Grogan DR, Packman J. Correlation of clinical outcomes and disease burden in patients with transthyretin (TTR)

amyloid polyneuropathy: study Fx-005, a landmark clinical trial of Fx-1006A, a novel small molecule TTR stabiliser. *J Neurol* 2008;**255**:78.
35. Coelho T, Maia L, da Silva AM, Waddingtion-Cruz M, Plante-Bordeneuve V, Lozeron P, et al. Tafamidis (Fx-1006A): a first-in-class disease-modifying therapy for transthyretin familial amyloid. *Amyloid* 2010;**17**:75-6.
36. Coelho T, Maia L, da Silva AM, Cruz MW, Plante-Bordeneuve V, Lozeron P, et al. A comprehensive evaluation of the disease-modifying effects of tafamidis in patients with transthyretin type familial amyloid polyneuropathy. *Neurology* 2011;**76**:A111.
37. Coelho T. New pharmacological treatment. *J Neurol* 2011;**258**:S4.
38. Cruz MW, Coelho T, Maia L, da Silva AM, Plante-Bordeneuve V, Lozeron P, et al. Tafamidis and nerve fiber function in transthyretin-type familial amyloid polyneuropathy. *Muscle Nerve* 2011;**44**:672-.
39. Coelho T, Maia L, da Silva AM, Cruz MW, Plante-Bordeneuve V, Lozeron P, et al. The long-term effects of Tafamidis for the treatment of transthyretin type familial amyloid polyneuropathy. *Neurology* 2011;**76**:A183-4.
40. *An Extension of Study Fx-005 Evaluating Long-Term Safety And Clinical Outcomes Of Fx-1006A In Patients With Transthyretin Amyloid Polyneuropathy*. Available from: http://ClinicalTrials.gov/show/NCT00791492
41. Coelho T, Maia L, Martins Da Silva A, Waddington Cruz M, Plante-Bordeneuve V, Suhr OB, et al. Long-term effects of tafamidis - a new therapeutic option for patients with transthyretin familial amyloid polyneuropathy (TTR-FAP). *J Hepatol* 2012;**56**:S543.
42. Coelho T, Maia L, Silva A, Waddington Cruz M, Plante-Bordeneuve V, Suhr OB, et al. Sustainability of response to tafamidis, a disease-modifying therapy for transthyretin type familial amyloid polyneuropathy. *Eur J Neurol* 2010;**17**:35.
43. Coelho T, Maia L, Da Silva AM, Cruz MW, Plante-Bordeneuve V, Suhr OB, et al. Early-treatment effects of tafamidis in transthyretin type familial amyloid polyneuropathy. *J Peripher Nerv Syst* 2011;**16**:S24-S5.
44. Plante-Bordeneuve V, Schmidt H, Merlini G, Obici L, Judge DP, Packman J, et al. The effects of tafamidis on transthyretin stabilization and clinical outcomes in patients with non-V30M transthyretin amyloidosis. *Eur J Neurol* 2011;**18**:29.
45. *The Effects of Fx-1006A on Transthyretin Stabilization and Clinical Outcome Measures in Patients With Non-V30M Transthyretin Amyloidosis*. In. Available from: http://ClinicalTrials.gov/show/NCT00630864
46. Peripheral Nerve Society. Diabetic polyneuropathy in controlled clinical trials: consensus report of the Peripheral Nerve Society. *Ann Neurol* 1995;**38**:478-82.
47. Food and Drug Administration, Center for Drug Evaluation and Research, Division of Neurology Products. *Vyndaqel (tafamidis meglumine). NDA 202737. Proposed Indication: Treatment of transthyretin (TTR) familial amyloid polyneuropathy. Peripheral and Central Nervous System Drugs Advisory Committee Background Package*. 2012. [cited 2012 May 23]. Available from: http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/UCM304830.pdf.
48. Coelho T, Maurer M, Plante-Bordeneuve V, Rapezzi C. Baseline demographics in THAOS - The transthyretin amyloidosis outcomes survey. *Neurology* 2012;**78**:Abstract: P03149.
49. Drummond M, Sculpher M, Torrance G, O'Brien BJ, Stoddart G. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press; 2005.
50. Ratcliffe J, Longworth L, Young T, Bryan S, Burroughs A, Buston M. Assessing Health-Related Quality of Life Pre– and Post–Liver Transplantation: A Prospective Multicenter Study. *Liver Transplantation* 2002;**8**:263-70.
51. National Institute for Health and Clinical Excellence. *Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease*. London: National Institute for Health and Clinical Excellence; 2009.

52. National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisal*. London: National Institute for Health and Clinical Excellence; 2008.
53. *Tafamidis (Vyndaqel) for the treatment of transthyretin amyloidosis in adult patients with stage 1 symptomatic polyneuropathy to delay neurologic impairment*. All Wales Medicines Strategy Group; 2012. [cited 2012 May]. Available from: http://www.wales.nhs.uk/sites3/Documents/371/Statement%20of%20Advice%20Vyndaqel.pdf.
54. IQWiG - Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. *Tafamidis: approval denotes proven added benefit*. [cited 2012 May]. Available from: https://www.iqwig.de/sidaua9d28tc8kkrrm52gogol17j2lvqe2k/tafamidis-approval-denotes-proven-added-benefit.1437.en.html.
55. *Relatorio de avaliacao previa do medicamento para uso humano em meio hospitalar - Tafamidis*. Lisboa, Portugal: Infarmed - Autoridade National do Medicamento e Produtos de Saude I.P.; 2012. Available from: http://www.infarmed.pt/portal/page/portal/INFARMED/MEDICAMENTOS_USO_HUMANO/AVALIACAO_ECONOMICA_E_COMPARTICIPACAO/MEDICAMENTOS_USO_HOSPITAL/DL_N_195_2006_3_OUT/RELATORIOS_AVALIACAO_PREVIA/Relat%F3rios%20de%20avalia%E7%E3o%20pr%E9via%20%E0%20utiliza%E7%E3o%20em%20meio%20hospita1/151_ParecerNet_tafamidis%5BVyndaqel%5D_Deferimento_Mai2012_convertido.pdf
56. Lane T, Cavalla F, Pinney J, Gibbs S, Dung J, Banypersad S, et al. 21 year experience of transthyretin amyloidosis in a single national centre. In: *VIIth International Symposium on Familial Amyloidotic Polyneuropathy*; 2011 November 20th-23rd, 2011; Kumamoto University, Japan. 2011.
57. Wilczek HE, Larsson M, Ericzon BG. Long-term data from the Familial Amyloidotic Polyneuropathy World Transplant Registry (FAPWTR). *Amyloid* 2011;**18**:188-90.
58. Sullivan PW, Slejko JF, Sculpher M, Ghushchyan V. Catalogue of EQ-5D Scores for the United Kingdom. *Med Decis Making [serial online]* 2011. [cited April 16, 2012].
59. *British National Formulary no. 62*. London: BMJ Publishing Group Ltd and Royal Pharmaceutical Society of Great Britain; 2011.
60. Curtis L. *Unit costs of health and social care 2011*. Canterbury: Personal Social Services Research Unit (PSSRU); 2011.
61. Department of Health. *NHS reference costs 2010-2011*. London: Department of Health; 2012.
62. Hartwell D, Jones J, Baxter L, Shepherd J. *Peginterferon alfa and ribavirin for chronic hepatitis C in patients eligible for shortened treatment, re-treatment or in HCV/HIV co-infection: a systematic review and economic evaluation*: Technology Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Clinical Excellence; 2009.
63. Claxton K, Walker S, Palmer S, Sculpher M. *Appropriate perspectives for health care decisions*. York: Centre for Health Economics, University of York; 2010.
64. Directgov. *Calculating your state pension age*. [cited 2012 May]. Available from: http://www.direct.gov.uk/en/pensionsandretirementplanning/statepension/dg_4017919.
65. Towse A. Should NICE's threshold range for cost per QALY be raised? Yes. *BMJ* 2009;**338**:b181.
66. National Institute for Health and Clinical Excellence. *Appraising Orphan Drugs*. 2006. [cited 2011 May]. Available from: http://www.nice.org.uk/niceMedia/pdf/smt/120705item4.pdf.
67. Vinik EJ, Hayes RP, Oglesby A, Bastyr E, Barlow P, Ford-Molvik SL, et al. The development and validation of the Norfolk QOL-DN, a new measure of patients' perception of the effects of diabetes and diabetic neuropathy. *Diabetes Technol Ther* 2005;**7**:497-508.
68. Smith SC, Lamping DL, Maclaine GD. Measuring health-related quality of life in diabetic peripheral neuropathy: a systematic review. *Diabetes Res Clin Pract* 2011;**96**:261-70.
69. Dyck PJ, Davies JL, Litchy WJ, O'Brien PC. Longitudinal assessment of diabetic polyneuropathy using a composite score in the Rochester Diabetic Neuropathy Study cohort. *Neurology* 1997;**49**:229-39.

70. Bril V. NIS-LL: the primary measurement scale for clinical trial endpoints in diabetic peripheral neuropathy. *Eur Neurol* 1999;**41**:8-13.
71. Apfel SC, Schwartz S, Adornato BT, Freeman R, Biton V, Rendell M, et al. Efficacy and safety of recombinant human nerve growth factor in patients with diabetic polyneuropathy: a randomized controlled trial. *JAMA* 2000;**284**:2215-21.
72. Perkins BA, Bril V. Diabetic neuropathy: a review emphasizing diagnostic methods. *Clinical Neurophysiology* 2003;**114**:1167-75.

Appendix 1: Systematic review methods

The following sections describe the methods used to perform the ERG's systematic review.

Inclusion criteria

Studies meeting the following criteria were included:

Population

Adults with stage 1 TTR-FAP, V30M or non-V30M, were included.

Intervention

Tafamidis

Comparators

Supportive therapy including liver transplantation. Placebo.

Outcomes

The outcomes of interest included quality of life, progression of peripheral neuropathy, mortality, cardiac outcomes and adverse events.

Study design

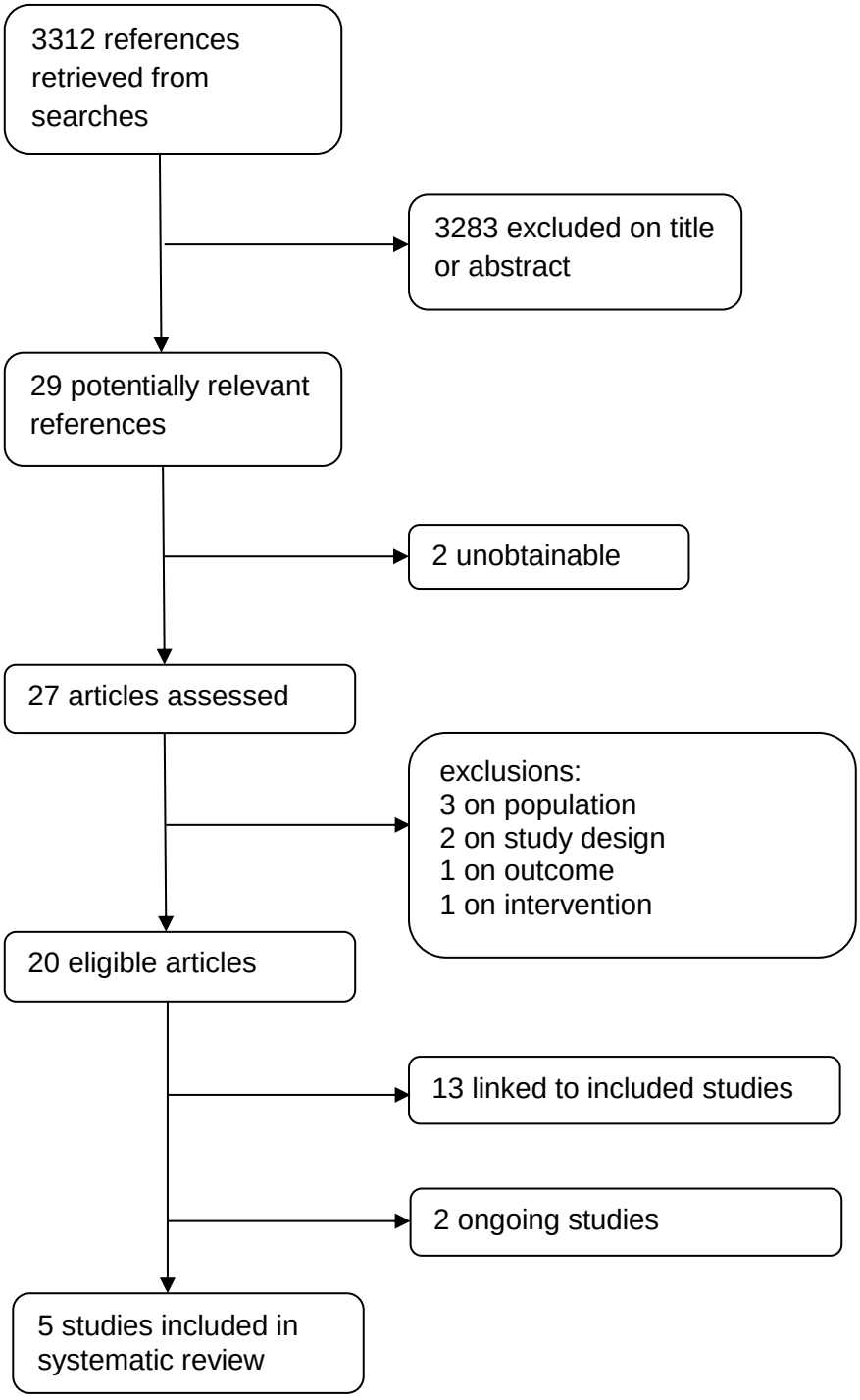
Randomised controlled trials (RCTs) were included. Where RCTs were not available for the comparisons of interest non-randomised controlled studies and observational study designs (including cohort studies, case-control studies and case series) were eligible for inclusion. These study designs were also eligible to inform the assessment of adverse events.

Published and unpublished data were eligible for inclusion provided that there were sufficient methodological details reported to allow assessment of risk of bias. Investigators were contacted to obtain further methodological details where necessary. Pre-clinical and biological studies were excluded.

Abstracts of identified studies were independently assessed for inclusion by two reviewers using the criteria outlined above. Disagreements were resolved through discussion and, where necessary, by consultation with a third reviewer. For studies identified as potentially

relevant, full papers were assessed independently by two reviewers with disagreements resolved by the same procedure. A flow chart summarising study selection is presented below in Figure 6.

Figure 6: Study flow chart



Data extraction

Data on key study characteristics (population, intervention, comparator, study methods) and results were extracted by one reviewer and checked by a second reviewer. Disagreement were resolved by consensus or by a third reviewer where necessary. Attempts were made where possible to contact investigators for clarification of data.

Assessment of risk of bias

The Cochrane risk of bias tool was used to assess RCTs. For other study designs, a checklist used in previous reviews by the TAR group were used to inform an assessment of study quality. The assessments were undertaken by one researcher and checked by a second. Discrepancies were resolved by discussion or by a third reviewer where necessary.

Methods of analysis and synthesis

It was anticipated that there would be insufficient data available to undertake a meta-analysis. Data were reported in tables and discussed in a narrative. Where possible, data related to patients V30M and non-V30M mutations were reported and analysed separately.

Literature searches

Date searches were conducted: 5-6 March 2012

Limits:	English language publications only
	No publication date limit
Records found:	After deduplication: 3312
	Before deduplication: 7477

The aim of the literature searches was to systematically identify studies on the effectiveness of Tafamidis for the treatment of transthyretin familial amyloid polyneuropathy. The base search strategy was constructed using MEDLINE and then adapted to the other resources searched. The search included the following components:

1. transthyretin familial amyloid polyneuropathy terms
OR
2. tafamidis terms

The two sets of terms were combined with the Boolean operator OR, in order to identify all records on Tafamidis, as well as all records on transthyretin familial amyloid polyneuropathy. No language or study design filters were applied. Results were limited to English language only publications.

Search terms were identified by scanning key papers identified at the beginning of the project, through discussion with the review team and the use of database thesauri. The creation of the search strategy was an iterative process originally using the MEDLINE database and then adapted as appropriate to the other sources searched.

Databases

Sources of information were identified by an Information Specialist with input from the project team. The following databases were searched from date of inception to the most recent date available for relevant studies:

Biosis (via Dialog 1926 to March 2012)
CINAHL (via EBSCO 1982 to March 2012)
Cochrane Library (Issue 2 of 12 February 2012):

- Cochrane Database of Systematic Reviews
- Database of Abstracts of Reviews of Effects (DARE)
- Cochrane Central Register of Controlled Trials
- HTA Database
- NHS Economic Evaluation Database

Conference Proceedings Citation Index- Science (via Web of Knowledge 1990 to March 2012)
Dissertation Abstracts (via Dialog 1861 to March 2012)
EconLit (via OvidSP 1961 to February 2012)
EMBASE (via OvidSP 1974 to 2 March 2012)
Inside Conferences (via Dialog October 1993 to March 2012)
MEDLINE and MEDLINE In Process and Other Non-Indexed Citations (via OvidSP 1946 to February Week 4 2012)
Science Citation Index Expanded (via Web of Knowledge 1899 to March 2012)

Ongoing studies were identified from the following databases:

Clinical Trials.gov (via website at http://www.clinicaltrials.gov/ to March 2012)

Full search strategies for each database searched are provided below.

Current awareness searches for Tafamidis and Diflunisal, a relevant comparator, were run on a weekly basis on MEDLINE and EMBASE in order to keep as up to date as possible with new publications in the field. This was done until June 6, 2012.

SEARCH STRATEGIES

MEDLINE and MEDLINE In Process and Other Non-Indexed Citations (Ovid)

Date range: 1946 – February Week 4 2012

Date searched: 5 March 2012

Records found: 1490

1 Amyloid Neuropathies, Familial/ (401)

2 amyloidosis, familial/ (230)

3 familial amyloid$ polyneuropath$.ti,ab. (1167)

4 transthyretin amyloidosis.ti,ab. (98)

5 (transthyretin related hereditary amyloidosis or transthyretin related familial amyloidosis).ti,ab. (4)

6 (transthyretin type hereditary amyloidosis or transthyretin type familial amyloidosis).ti,ab. (0)

7 TTR amyloid polyneuropathy.ti,ab. (2)

8 (ttr-fap or attr).ti,ab. (320)

9 corino de andrade$ disease.ti,ab. (2)

10 (neuropath$ adj2 amyloid$ adj2 (familial or hereditary)).ti,ab. (64)

11 or/1-10 (1746)

Line 11 captures transthyretin familial amyloid polyneuropathy terms

12 exp animals/ not humans/ (3667503)

13 11 not 12 (1716)

Line 13 excludes animal-only studies

14 tafamidis.af. (5)

15 Vyndaqel.af. (1)

16 Fx-1006A.af. (2)

17 14 or 15 or 16 (6)

Line 17 captures Tafamidis terms

18 13 or 17 (1718)

Line 18 groups transthyretin familial amyloid polyneuropathy and Tafamidis terms into one set

19 limit 18 to english language (1490)

Line 19 limits the results to English language studies only

Key

/ = indexing term (MeSH heading)

exp = exploded MeSH heading

$ = truncation

.ti,ab. = terms in either title or abstract fields

adj = terms adjacent to each other (same order)

adj2 = terms within two words of each other (any order)

Biosis (Dialog)

Date range: 1926 - date

Date searched: 6 March 2012

Records found: 1661

S (familial(w)amyloid?(w)polyneuropath?)/ti,ab,de

S (transthyretin(w)amyloidosis)/ti,ab,de

S ((transthyretin(w)related(w)hereditary(w)amyloidosis) or (transthyretin(w)related(w)familial(w)amyloidosis))/ti,ab,de

S ((transthyretin(w)type(w)hereditary(w)amyloidosis) or (transthyretin(w)type(w)familial(w)amyloidosis))/ti,ab,de

S (ttr(w)amyloid(w)polyneuropathy)/ti,ab,de

S (ttr(w)fap or ttr-fap)/ti,ab,de

S attr/ti,ab,de

S (corino(w)de(w)andrade?(w)disease)/ti,ab,de

S (neuropath?(2n)amyloid?(2n)(familial or hereditary))/ti,ab,de

S (tafamidis or vyndaqel or fx(w)1006a or fx-1006a)

S s1:s10

Key

/ti,ab,de = searches title, abstract and descriptor fields

(w) = terms adjacent to each other (same order)

(2n) = terms within two words of each other (any order)

? = truncation

CINAHL (EBSCO)

Date range: 1982 - date

Date searched: 5 March 2012

Records found: 42

NB – ‘fx-1006a’ removed from strategy as it was not searched correctly on CINAHL

S1 (MH "Amyloid Neuropathies, Familial") (12)

S2 (MH "Amyloidosis, Familial") (6)

S3 "familial amyloid* polyneuropath*" (28)

S4 "transthyretin amyloidosis" (5)

S5 "transthyretin related hereditary amyloidosis" or "transthyretin related familial amyloidosis" (0)

S6 "transthyretin type hereditary amyloidosis" or "transthyretin type familial amyloidosis" (0)

S7 "ttr amyloid polyneuropathy" (0)

S8 "ttr fap" or attr (9)

S9 "corino de andrade* disease" (0)

S10 neuropath* n2 amyloid* n2 familial (12)

S11 neuropath* n2 amyloid* n2 hereditary (1)

S12 tafamidis or vyndaqel or "fx-1006a" (0)

S13 S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 (44)

S14 tafamidis or vyndaqel (0)

S15 S13 or s14

S16 (ZL "english") (2600120)

S17 S15 and S16 (42)

Key

MH = indexing term (CINAHL heading)

* = truncation

" " = phrase search

n2 = terms within one word of each other (any order)

ZL = language field

Clinical Trials.gov

http://www.clinicaltrials.gov/

Date range: all to date

Date searched: 6 March 2012

Records found: 63

Search 1 "familial amyloid polyneuropathy" OR "transthyretin amyloidosis" OR "transthyretin related hereditary amyloidosis" OR "transthyretin related familial amyloidosis" (12 records)

Search 2 "transthyretin type hereditary amyloidosis" OR "transthyretin type familial amyloidosis" OR "ttr amyloid polyneuropathy" OR "ttr fap" OR attr OR "corino de andrade disease" (61 records)

Search 3 (neuropathy AND amyloid AND familial) OR (neuropathy AND amyloid AND hereditary) OR tafamidis OR vyndaqel OR "fx-1006a" (10 records)

Combining searches 1-3 = 63 unique records

Key

" " = phrase search

The Cochrane Library

http://www.thecochranelibrary.com/

Issue searched: Issue 1 of 12, Feb 2012

Date searched: 5 March 2012

Records found:

Cochrane Database of Systematic Reviews (1)
Database of Abstracts of Reviews of Effects (DARE) (0)
Cochrane Central Register of Controlled Trials (6)
HTA Database (2)
NHS Economic Evaluation Database (0)

#1 MeSH descriptor Amyloid Neuropathies, Familial, this term only 4

#2 MeSH descriptor Amyloidosis, Familial, this term only 1

#3 "familial amyloid* polyneuropath*" 0

#4 "transthyretin amyloidosis" 0

#5 "transthyretin related hereditary amyloidosis" or "transthyretin related familial amyloidosis" 0
#6 "transthyretin type hereditary amyloidosis" or "transthyretin type familial amyloidosis" 0
#7 "ttr amyloid polyneuropathy" 1
#8 "ttr-fap" or attr 4
#9 "corino de andrade* disease" 0
#10 neuropath* near/2 amyloid* near/2 (familial or hereditary) 4
#11 tafamidis or vyndaqel or "fx-1006a" 3
#12 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11) 9

Key
MeSH descriptor = indexing term (MeSH heading)
* = truncation
" " = phrase search
near/2 = terms within one word of each other (any order)

Conference Proceedings Citation Index- Science (Web of Science)
Date range: 1990 - date
Date searched: 5 March 2012
Records found: 282

1 TS=("familial amyloid* polyneuropath*")
2 TS=("transthyretin amyloidosis")
3 TS=("transthyretin related hereditary amyloidosis" or "transthyretin related familial amyloidosis")
4 TS=("transthyretin type hereditary amyloidosis" or "transthyretin type familial amyloidosis")
5 TS=("ttr amyloid polyneuropathy")
6 TS=("ttr-fap" or attr)
7 TS=("corino de andrade* disease")
8 TS=(neuropath* near/2 amyloid* near/2 (familial or hereditary))
9 TS=(tafamidis or vyndaqel or "fx-1006a")
#10 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9

Limits: English language only; Lemmatization - OFF

Key

Topic = terms in Title, Abstract, Author Keywords and Keywords Plus fields

* = truncation

" " = phrase search

near/2 = terms within one word of each other (any order)

Dissertation Abstracts (Dialog)

Date range: 1861 to date

Date searched: 6 March 2012

Records found: 37

S (familial(w)amyloid?(w)polyneuropath?)/ti,ab,de

S (transthyretin(w)amyloidosis)/ti,ab,de

S ((transthyretin(w)related(w)hereditary(w)amyloidosis) or
(transthyretin(w)related(w)familial(w)amyloidosis))/ti,ab,de

S ((transthyretin(w)type(w)hereditary(w)amyloidosis) or
(transthyretin(w)type(w)familial(w)amyloidosis))/ti,ab,de

S (ttr(w)amyloid(w)polyneuropathy)/ti,ab,de

S ((ttr(w)fap) or ttr-fap or attr)/ti,ab,de

S (corino(w)de(w)andrade?(w)disease)/ti,ab,de

S (neuropath?(2n)amyloid?(2n)(familial or hereditary))/ti,ab,de

S (tafamidis or vyndaqel or fx(w)1006a or fx-1006a)

S s1:s9

Key:

/ti,ab,de = searches title, abstract and descriptor fields

(w) = terms adjacent to each other (same order)

(2n) = terms within two words of each other (any order)

? = truncation

EconLit (Ovid)

Date range: 1961 – February 2012

Date searched: 5 March 2012

Records found: 0

1 familial amyloid$ polyneuropath$.ti,ab. (0)

2 transthyretin amyloidosis.ti,ab. (0)

3 (transthyretin related hereditary amyloidosis or transthyretin related hereditary amyloidosis).ti,ab. (0)

4 (transthyretin type hereditary amyloidosis or transthyretin type hereditary amyloidosis).ti,ab. (0)

5 TTR amyloid polyneuropathy.ti,ab. (0)

6 (ttr-fap or attr).ti,ab. (0)

7 corino de andrade$ disease.ti,ab. (0)

8 (neuropath$ adj2 amyloid$ adj2 (familial or hereditary)).ti,ab. (0)

9 (tafamidis or vyndaqel or fx-1006a).af. (0)

10 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 (0)

Key:

$ = truncation

.ti,ab. = terms in either title or abstract fields

adj2 = terms within two words of each other (any order)

EMBASE (Ovid)

Date range: 1974 – 02 March 2012

Date searched: 5 March 2012

Records found: 1736

1 familial amyloid polyneuropathy/ (726)

2 familial amyloidosis/ (334)

3 familial amyloid$ polyneuropath$.ti,ab. (1398)

4 transthyretin amyloidosis.ti,ab. (138)

5 (transthyretin related hereditary amyloidosis or transthyretin related hereditary amyloidosis).ti,ab. (4)

6 (transthyretin type hereditary amyloidosis or transthyretin type hereditary amyloidosis).ti,ab. (0)

7 TTR amyloid polyneuropathy.ti,ab. (3)

8 (ttr-fap or attr).ti,ab. (427)

9 corino de andrade$ disease.ti,ab. (2)

10 (neuropath$ adj2 amyloid$ adj2 (familial or hereditary)).ti,ab. (91)

11 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 (2224)

12 exp animal/ (1680441)
13 exp nonhuman/ (3803372)
14 12 or 13 (5467464)
15 exp human/ (13062869)
16 14 not (14 and 15) (4360019)
17 11 not 16 (2052)
18 (tafamidis or vyndaqel or fx-1006a).af. (33)
19 17 or 18 (2061)
20 limit 19 to english language (1736)

Key:
/ = indexing term (EMTREE heading)
exp = exploded EMTREE heading
$ = truncation
.ti,ab. = terms in either title or abstract fields
adj2 = terms within two words of each other (any order)

Inside Conferences (Dialog)
Date range: October 1993 to date
Date searched: 6 March 2012
Records found: 108

S (familial(w)amyloid?(w)polyneuropath?)/ti,ab,de
S (transthyretin(w)amyloidosis)/ti,ab,de
S ((transthyretin(w)related(w)hereditary(w)amyloidosis) or
(transthyretin(w)related(w)familial(w)amyloidosis))/ti,ab,de
S ((transthyretin(w)type(w)hereditary(w)amyloidosis) or
(transthyretin(w)type(w)familial(w)amyloidosis))/ti,ab,de
S (ttr(w)amyloid(w)polyneuropathy)/ti,ab,de
S ((ttr(w)fap) or ttr-fap or attr)/ti,ab,de
S (corino(w)de(w)andrade?(w)disease)/ti,ab,de
S (neuropath?(2n)amyloid?(2n)(familial or hereditary))/ti,ab,de
S (tafamidis or vyndaqel or fx(w)1006a or fx-1006a)
S s1:s9
S s10/eng

Key:

/ti,ab,de = searches title, abstract and descriptor fields

(w) = terms adjacent to each other (same order)

(2n) = terms within two words of each other (any order)

? = truncation

Science Citation Index Expanded (Web of Science)

Date range: 1899 - date

Date searched: 5 March 2012

Records found: 1791

1 TS=("familial amyloid* polyneuropath*")

2 TS=("transthyretin amyloidosis")

3 TS=("transthyretin related hereditary amyloidosis" or "transthyretin related familial amyloidosis")

4 TS=("transthyretin type hereditary amyloidosis" or "transthyretin type familial amyloidosis")

5 TS=("ttr amyloid polyneuropathy")

6 TS=("ttr-fap" or attr)

7 TS=("corino de andrade* disease")

8 TS=(neuropath* near/2 amyloid* near/2 (familial or hereditary))

9 TS=(tafamidis or vyndaqel or "fx-1006a")

#10 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9

Limits: English language only; Lemmatization – OFF

Key

Topic = terms in Title, Abstract, Author Keywords and Keywords Plus fields

* = truncation

" " = phrase search

near/2 = terms within one word of each other (any order)

Appendix 2: Study Fx1A-OS-001

Context

Fx1A-OS-001 is referred to in the report as demonstrating the correlation between disease progression and quality of life (1.3). The manufacturer's submission concludes that results from this observational study demonstrated that these scales (NIS-LL and Norfolk QoL) correlated well with disease progression and were relevant measures to be used in TTR-FAP studies (MS Section 3.15.3.4). Of particular relevance to this assessment is the quantile regression analysis with restricted cubic spline function, to estimate the relationship between disease duration and Norfolk TQoL. The results of this analysis were used in the manufacturer's submission to estimate length of time spent in each disease stage for the decision model and this is discussed in further detail in the economics section.

Methods

Table 23 of the manufacturer's submission provides an outline of the study methods. In summary, people with stage 1, 2 and 3 V30M TTR-FAP and healthy volunteers were recruited into the study at a single centre in Portugal. The baseline data for 29 patients from the Portugal centre of the main trial (FX005) with stage 1 disease were utilised as the stage 1 group. A single clinician classified the disease stage of each participant with TTR-FAP using Coutinho's classification. This was determined based on assistance required to walk: stage 1, fully ambulatory; stage 2, required assistance with ambulation; stage 3, wheelchair bound. It also appears likely that the same clinician administered the primary outcome measures that have been used in the main trial for tafamidis as well as other secondary outcome measures.

The primary analysis, which was pre-specified in the statistical analysis plan, was to compare scores on the Norfolk TQoL and NIS-LL by disease stage and healthy volunteer group using ANOVA and pair-wise comparisons. Several additional analyses were undertaken, several of which were not pre-specified in the analysis plan.

Results

The mean age at symptom onset was progressively older in each stage group: stage 1 39.0 years, stage 2 46.5 years and stage 3 55 years. The mean duration of symptoms was 31.2 months in the Stage 1 group, 92.4 in Stage 2 and 170.4 in stage 2. Further participant characteristics are reported in Table 4 in the main body of the ERG report. There was a statistically significant difference in NIS-LL scores across the four groups with score

increasing (more severe neuropathy) by disease stage, though scores for stage 2 and 3 patients were fairly close (Manufacturer's submission Figure 16). There was also a statistically significant difference in Norfolk TQoL scores across the four groups with score increasing (worsening quality of life) by disease stage (Manufacturer's submission Figure 17). Scores on the outcome measures for each disease stage group are reported in Table 4.

There was a [redacted] correlation (r^2 = [redacted], p[redacted]) between disease duration and TQol scores: quality of life worsens with longer disease duration (Manufacturer's submission, Figure 26). The overall estimated rate of change on Norfolk TQoL was [redacted] points per year.

Critique

This study has a number of limitations that impact on the robustness of the results:

- Patient selection

There was lack of clarity in the method of selecting patients for this study. The unpublished clinical study report states that baseline data from the FX005 trial for 29 stage 1 patients from the centre were used for the stage 1 group. It is unclear how these patients were selected from the 64 patients from Porto who participated in the FX005 trial, for example whether they were randomly selected or whether they were chosen based on specific clinical criteria. This introduces the possibility of bias in the stage 1 sample. The process of selection of stage 2 and 3 patients is also unclear.

- Patient population

There are two sources of uncertainty regarding the generalisability of the results of this study. First, the patients all had the V30Met mutation whereas this mutation is unusual in England where non-Val30Met mutations are more common. There is limited evidence of the validity and reliability of using the Coutinho's classification in a non-Val30Met population and clinical advice suggests that this staging system does not capture the severe autonomic and cardiac involvement that patients with some non-Val30Met mutations present with, particularly in the England population.

Second there is uncertainty regarding generalisability to other Val30Met populations. Examination of the Norfolk TQoL scores and NIS-LL scores for stage 1 patients from the cross-sectional study (Fx1A-OS-001) compared to the placebo and tafamidis population in the trial FX005 indicate that the patients from Fx1A-OS-001 have lower (better) quality of life scores and slightly lower (better) NIS-LL scores and shorter disease duration. This raises

uncertainty about how representative these patients are of Val30Met patients in early stage disease and therefore the applicability of the findings of the study to other V30Met populations.

- Measurement of the variables used

A single clinician undertook the staging of patients using Coutinho's classification. While this will have avoided introduction of variability between different assessors, the inter-rater reliability of this staging tool is unknown and it is unclear how ambulatory status was actually assessed. It seems likely that the same clinician undertook the NIS-LL assessment. This introduces the risk that the knowledge of stage may have influenced, even unknowingly, the assessment of neuropathy and vice versa. All the assessments appear to have been undertaken unblinded to other clinical information such as disease duration as well as Karnofsky score and other clinical measures. Therefore there is a real possibility that the correlation between disease duration and NIS-LL has been overestimated. The Norfolk TQoL was completed by patients therefore the risk is less with this outcome measure, though knowledge of the score could have contaminated staging.

Appendix 3: Details of the Co-primary outcome assessment tools

Norfolk Quality of Life – Diabetic Neuropathy (QoL-DN) Questionnaire

This is a patient-reported outcome measure. It was developed to be sensitive to the different features of diabetic neuropathy: small fibre, large fibre and autonomic nerve function. [67] The scale has five domains: (i) physical functioning/large nerve fibre function; (ii) activities of daily living; (iii) symptoms; (iv) small nerve fibre function; (v) autonomic nerve function. The questionnaire is completed in relation to symptoms and problems in the previous four weeks.

Seven items on the scale related to symptoms are scored as a 1 or 0 to indicate presence or absence of the symptom; 26 of the items are scored on a 5-point Likert scale ranging from 0 (no problem) to 4 (severe problem), including one symptom item; two items are scored on a scale of 0 (excellent) to -4 (poor). Table 44 provides a summary of the content of each of the domains.

Table 44: Summary of content of Norfolk QoL-DN Questionnaire

Domain	Content
Symptoms	*Presence of symptoms in feet, legs, hands, arms* -numbness -tingling, pins & needles -electric shocks -unusual sensations -superficial pain -deep pain -weakness - touch of sheets, clothes shoes bothersome
Large nerve fibre	-pain at night -symptoms preventing usual daytime activities -unsteady on feet when walking -problems getting out of chair without using hands -problems going down stairs -difficulty walking -needing to reduce time spent on work or other activities -accomplishing less than would like -limited in kind of activities can do -extra effort needed to do work or other activities -general health now -general health compared to 3 months ago -physical health interfering with normal social activities -pain interfering with work or housework -weakness/shakiness interfering with work or housework

Small nerve fibre	*Unable to* - feel a burn or injury - feel feet when walking -tell hot/cold water with hand -tell hot/cold water with feet
Activities of daily living	*Difficulty with* -fine movements -bathing/showering -dressing -getting on/off toilet -using eating utensils
Autonomic nerve function	*Problems with* -vomiting -Diarrhoea/bowel control -fainting dizziness when standing

There are 35 scored items and the possible total score, which is the sum of all 35 items (without any weighting) ranges from -4 to 135. In the development study for the questionnaire the mean total score (SE) was 41.8 (3.1) for a sample of diabetic patients with neuropathy, 13 (1.5) for diabetic patients without neuropathy and 3.8 (0.5) for healthy controls.[67]

A recent systematic review appraising the psychometric evidence for health related quality of life measures concluded that there was acceptable evidence for the Norfolk QoL-DN Questionnaire meeting some of the psychometric criteria used in relation to a diabetic population.[68] There was at least some acceptable evidence for internal consistency, test-retest reliability, content validity and some dimensions of construct validity in a diabetic population. Limitations identified were the absence of a psychological/emotional domain, limited data on the responsiveness of the questionnaire and limited experience of its implementation in clinical trials.[68]

There is no existing validated disease-specific quality of life outcome measure for TTR-FAP that could have been used for the trial or supporting studies in the submission. The Norfolk QoL-DN measure was chosen to allow assessment of the impact of the intervention on quality of life related to the progression of peripheral neuropathy (Manufacturer's Submission 3.1). While the measure is considered to have reasonable reliability and validity in a diabetic population, there is some uncertainty as to whether it captures all aspects of quality of life associated with TTR-FAP. It does not capture the emotional/psychological components of the condition and may not capture the impact of all important aspects of the condition on quality of life where peripheral neuropathy is not the predominant symptom. Although the

measure does include autonomic nerve function there are only three items on this subscale related to vomiting, diarrhoea and dizziness. These are common symptoms in TTR-FAP, but other relevant symptoms that may impact on quality of life are missing such as renal and urinary symptoms and erectile or sexual dysfunction. It is unclear whether the instrument would sufficiently capture cardiac symptoms, though this is more relevant to the study including non-V30M (Fx1A-201) patients than the main trial (Fx-005).

Neuropathy Impairment Score-Lower Limb (NIS-LL)

The NIS-LL is a tool for neurological examination that evaluates motor (muscle strength), sensory and reflex activity in the lower limbs.[69] It was developed for use in diabetic neuropathy studies. For muscle strength the clinician grades eight lower limb muscle groups (hip, knee, ankle, toe) on an 8-point scale from 0 (normal) to 4 (paralysis) (there are partial points on the scale). For the sensory subscale, the dorsal surface of the great toe at the base of the nail at the terminal phalanx is assessed using four modalities (touch pressure, pinprick vibration using a tuning fork and joint position) as being normal (scored 0), decreased (1) or absent (2) for each modality. Reflexes are assessed in the same way for the quadriceps and ankle for the reflex subscale. All assessments are undertaken on both sides of the body. Total score on the scale ranges from 0 (normal, no neuropathy) to 88 (no lower limb motor, sensory or reflex activity). The maximum possible score on the three sub-scales are 64 for muscle strength, 16 for sensory testing and 8 for reflexes.[70] The Peripheral Nerve Society suggests that a mean change of two points between an intervention and placebo on the NIS is clinically meaningful and this was used by the manufacturer in their submission.[46] The consensus is based on the rationale that the least degree of neurological abnormality that a physician can recognise is equivalent to two points on the NIS (one point for each side of body). The robustness of this assumption is unclear, particularly as applied to a TTR-FAP population. It is also unclear whether from a patient perspective this is a difference that patients would perceive as beneficial.

The NIS-LL was derived from the Neuropathy Impairment Score (NIS) which assesses whole body function. The NIS, which was used in Fx1A-201, provides a single score of total body neuropathic deficits and total body scores for cranial nerves, muscle weakness, reflexes and sensations. The NIS-LL+7 tests which was used as a secondary outcome measure in the main trial, is a composite score of the NIS-LL plus scores for nerve conduction studies, quantitative sensory testing and heart rate response to deep breathing.[69]

NIS-LL was chosen as an outcome measure for the trial and supporting studies based on the rationale that lower limbs are most affected in early TTR-FAP and are therefore most appropriate to evaluate clinical progression over time (Manufacturer's Submission 3.1). There is some evidence of reasonable sensitivity and specificity in patients with diabetic neuropathy [69] and it has been used in some trials of interventions for diabetic neuropathy.[71] Scoring of muscle strength is subjective and may be open to bias.[70] In the trial (FX005), the same neurologist undertook the NIS-LL assessment where possible, to minimise any problems with inter-rater reliability, and the mean of two successive NIS-LL at least 24 hours apart within one week were used for each assessment point to reduce any problems with

intra-rater reliability(CTR, p49). These measures address variability in measurement within centres, though not across centres. To address variability between centres all site neurologists received a NIS-LL training video and were required to submit five sample test cases to an independent neurologist for certification prior to performing study patient assessments (CTR, p68).

One limitation of the NIS-LL is that it is heavily weighted by the motor function/muscle strength items,[72] which contribute a 64 points to the 88 point scale in comparison to sensory activity which contributes a maximum of 16 points. Clinical advice suggests that this limitation is of relevance in TTR-FAP where sensory symptoms are an important feature of the disease, particularly in the population included in the trial. At baseline in the RCT (Fx005) patient scores ranged from 0 to 14 in the tafamidis group and 0 to 16 in the placebo group out of a maximum score of 16 on the sensory subscale. This suggested that there was a ceiling effect i.e. for some patients further deterioration from baseline could not have been captured on this part of the scale. In addition, the focus on lower limb only in this outcome measure results in sensory impairment of upper limbs not being captured. Although descriptions of disease progression suggest that in early disease lower limb involvement precedes upper limb involvement in patients with V30M from endemic areas,[7] in other populations these upper and lower symptoms present more closely in time at the early stages of the disease.[8]

It has also been noted that the NIS-LL does not capture autonomic functioning.[70] Symptoms of autonomic dysfunction such as sexual impotence, nausea, vomiting and diarrhoea, and urinary incontinence and retention are common in TTR-FAP and these aspects of the condition are not captured by the NIS-LL. For example, participants in the main trial (FX-005) had a range of symptoms related to their TTR-FAP in addition to sensory motor neuropathy including gastrointestinal disorders (63%), metabolism and nutritional disorders (18%), renal and urinary disorders (29%) and erectile or sexual dysfunction (16%) Clinical Study Report, Table 14.1.4.2). Only 20% of participants had a peripheral neuropathy only. These limitations are also very pertinent in relation to use of the NIS-LL in a non-V30M population as in study Fx1A-os-001 in the manufacturer's submission, though that study also included the NIS.

Appendix 4: Points of clarification from manufacturer

Tafamidis for the treatment of TTR-FAP: Key points for clarification

Pfizer Ltd Response 8th May 2012

We would like to begin by thanking the ERG for their review of our submission and AGNSS for the opportunity to provide further clarification.

In order to fully engage in the AGNSS process, we have responded to each of the ERG's technical questions. However, our level of response is hampered to a great extent by the paucity of data available to support the estimate of a cost per QALY available for an ultra orphan condition.

Furthermore, we would like to highlight the following points for consideration in relation to our response:

- We are confident that tafamidis is an innovative new treatment meeting a high unmet need for patients with Transthyretin Familial Amyloid Polyneuropathy (TTR-FAP), a progressive and fatal ultra-orphan disease. Today there are no pharmacological treatments beyond symptomatic therapy to treat these patients. Tafamidis is the first, licensed treatment with demonstrated efficacy in terms of delaying peripheral neurologic impairment, such that patients may retain their independence and remain in stage 1 for longer compared to conventional supportive therapy (CST). Tafamidis could be provided to patients in England with a reasonable budget impact and minimal service implications.
- We have adopted a pragmatic approach to the economic modelling that is sufficiently robust to aid decision making. Specifically, we utilised the best available evidence for V30M and nonV30M to pool the data to minimise uncertainty, increase sample size and produce the most credible ICER given the inherent challenges of modelling for ultra orphan conditions.
- It is also important to highlight that the level of technical scrutiny that has been applied by the ERG, whilst appropriate for traditional NICE HTA appraisals, is less feasible in this setting given the paucity of data. The request for further statistical analyses does not help to aid the precision in the ICER estimate for tafamidis. That said, in our submission, we have attempted to address the uncertainty in our economic analysis using one-way sensitivity analyses and demonstrated that these results fall within the range of other available ICERs for ultra-orphan drugs. Many additional, sophisticated, analytical techniques, such as probabilistic sensitivity analysis (PSA) could have been undertaken, but their impact on the certainty or the acceptability of the cost/QALY estimate would be limited given the underlying uncertainty associated with the data for ultra orphan conditions.
- Moreover, the single dimension QALY measure doesn't capture the broader value of tafamidis. For example, the QALY measure does not account for the high unmet need, the severity and burden of the disease or the innovation

required to introduce the first pharmacological treatment for this orphan disease.

As a consequence, we would recommend that the ICER for tafamidis, whilst informative, should not be given priority in terms of the value of tafamidis over other equally important factors within the AGNSS multi-criteria decision making framework, such as societal value, severity and ability of patients to benefit and innovation. In line with the decision-making framework outlined by the AGNSS process, the ICER should be considered separately, alongside the other criteria when reaching a decision regarding the broader value of tafamidis.

Responses to technical queries

References to tables and figures relate to the Detailed Appraisal Information document, unless otherwise stated.

1.1 Study Fx-005 (RCT):

1.1.1 Please provide the study protocol (appendix 16.1.1 of clinical study report) and full details of the analysis plan (appendix 16.1.9 of clinical study report).

Please see appendices 1 and 2 attached for the study Fx-005 protocol (appendix 1) and full details of the analysis plan for study Fx-005(appendix 2).

1.1.2 Baseline characteristics (Table 10): For each treatment arm, please provide data for the following: age at symptom onset, age at diagnosis, IENF density, and baseline Karnofsky score.

Please note that data was not collected on the following parameters in study Fx-005: IENF density, age at diagnosis, age at symptom onset and baseline Karnofsky score.

However, we are able to provide estimates of mean age at symptom onset based on the available data for age and duration of symptoms, both at study entry.

The results of these new analyses are presented in Table 1.

Table 1: Age at Onset of Symptoms (Study Fx-005)

Safety Population	Tafamidis 20 mg N=65	Placebo N=63
Age, years		
Mean (SD)	36.6 (10.9)	35.7 (11.5)
Median	33.1	32.2
Range	23, 65	22, 63
ITT Population	**Tafamidis 20 mg N=64**	**Placebo N=61**
Age, years		
Mean (SD)	36.3 (10.8)	36.0 (11.5)
Median	33.0	32.9
Range	23, 65	22, 63

1.1.3 There are some discrepancies between Table 29 and elsewhere in the submission and the clinical trial report. The mean scores at 18 months for TQoL from Fx-005 in this table do not seem to tally with the LSM or mean reported elsewhere for this outcome. Could you please clarify this?

We would like to clarify that in our submission table 12 reports TQoL change from baseline to 18 months (co-primary outcome), which is different to that reported in Table 29 and figure 5 showing the TQoL change over time (supportive analyses).

Please see below a detailed explanation of the statistical approach taken to calculate these outcomes:

- For Table 12 in the dossier (co-primary endpoint analysis), an ANCOVA model, with baseline as a covariate, was used to obtain the LS mean at Month 18. For patients with post-baseline TQoL assessments but no assessment at Month 18, LOCF was used to impute missing data at Month 18. For patients without post-baseline TQoL assessments, the mean change from baseline at Month 18 for patients who had post-baseline assessments was used to impute the change from baseline within each treatment group.

- For Table 29 and figure 5 in the submission dossier (supportive analysis), a repeated measures mixed model was used to obtain the LS means at each time point based on observed cases and did not utilize data imputation methods. For the reference for this data, please see figure 10 in the CSR and Table 7, page 11 in the Q100 EMA response.

We haven't been able to check whether this applies to the other time points and NIS-LL as the data for other time points are reported in graphs only.

The ANCOVA model was only applied for the outcome TQoL change from baseline to 18 months (Table 12). For TQoL at the other time points (Figure 5; Table 29) and the NIS-LL (Figure 6; Table 29), the mixed model was used.

The number of patients at 6 month follow-up in the FX-005 tafamidis group (n=70) appears to be a misprint – can you please confirm.

We are able to confirm that there is a misprint in Table 29 in our submission for the number of patients in the tafamidis group at 6 month follow up. This should read 60 patients and not the n=70 that has been previously stated.

1.1.4 In each treatment arm, for how many patients were data i) imputed (i.e. no post-baseline assessment) and ii) carried forward (LOCF) for TQol and NIS-LL at the 6, 12, and 18 month timepoints?

We are able to clarify the number of patients with no post-baseline assessment. Table 2 below presents the number of patients with missing post-baseline assessment, by visit, for the NIS-LL and TQOL outcomes.

Table 2: The number of patients with missing post-baseline assessment, by visit for the NIS-LL and TQoL outcomes in study Fx-005

Visit	Tafamidis	Placebo	Total
Month 6	4	4	8
Month 12	15	11	26
Month 18	16	14	30

Eight patients (four in each treatment group) had no post-baseline assessment.

In terms of the number of patients with data carried forward, we would like to start by clarifying that a LOCF approach was utilised for the categorical NIS-LL responder analysis only. However those patients who discontinued due to liver transplant were treated as non-responders for NIS-LL for timepoints after the date of liver transplant as pre-specified in the analysis plan.

There were 12 patients in each treatment group who discontinued due to liver transplant and they were treated as non-responders. For the remaining six patients (four in Tafamidis group and two in Placebo group), LOCF was used to impute the missing Month 18 value for the responder analysis (co-primary endpoint analysis).

The NIS-LL continuous analysis of the change from baseline was carried out using a repeated measures analysis of variance (ANOVA) model. This method utilises all available data without the need for imputing missing data.

We are able to clarify that for the TQoL change from baseline to 18 months (co-primary endpoint analysis), LOCF was used to impute missing data at Month 18 for patients with post-baseline TQoL assessments but no assessment at Month 18. There were 12 patients in the Tafamidis group and 10 patients in the placebo group with missing Month 18 value and LOCF was used to impute the missing data.

For patients without post-baseline TQoL assessments, the mean change from baseline at Month 18 for patients who had post-baseline assessments was used to impute the change from baseline within each treatment group. There were eight patients (four in each treatment group) with no post-baseline assessment.

1.2 Study Fx1A-201 (before-after study):

1.2.1 Please provide the study protocol (Appendix 16.1.1) and analysis plan (Appendix 16.1.9).

As requested, please see appendixes 3 and 4 attached for the Fx1A201 study protocol (appendix 3) and analysis plan (appendix 4).

1.2.2 Can you provide Norfolk QoL-DN results for each of the five domains (mentioned in Methodology details, Table 16)?

As requested, these data are included in appendix 5, which includes a table showing the change from baseline for each of the five Norfolk QoL-DN domain scores.

1.2.3 Could we please have further details of exactly how rate of change in efficacy endpoints was estimated from pre-study symptom duration, and full results from Tables 14.4.2, 14.4.3 and 14.4.4 of the clinical study report.

An explanation of the methodology employed to estimate the rate of change in efficacy endpoints from the pre-study symptom duration in study Fx1A-201 is presented below:

- Please note that Fx1A-201 was an open label study and that the monthly rate of change prior to the initiation of tafamidis treatment was compared with the monthly rate of change during tafamidis treatment for key efficacy outcomes in patients completing 12 months of treatment (n=18). Thus, each patient served as his/her own control for this analysis.
- The pre-study monthly rate of change prior to treatment was defined as the ratio of the baseline scores for each efficacy outcome and duration of symptom onset (defined as the period from the first reported ATTR-related symptom to the start of the study).

- The on-study monthly rate of change for tafamidis treatment was estimated using the slope of the linear regression analyses using the available data gained during the study.
- The differences between the pre-study and on-treatment rates for each patient were analyzed using the signed rank test.

The requested data tables from the clinical study report showing the results of these analyses exploring the impact of tafamidis on disease progression using comparisons between on-study and pre-study rates of change in key outcomes are attached in appendices 6, 7 and 8.

1.3 Study Fx1A-OS-001:

1.3.1 Could we please have details of the process used to classify patients into each of the three stages. Was this undertaken by a single or multiple clinicians, how were the Coutinho descriptions applied and was any assessment of the reliability of the classifications undertaken?

Please find below an explanation of the process adopted in study Fx1A-OS-001 for classifying patients into each of the three stages:

- At the screening visit (Days -7 to -1), the ATTR-PN stage and participant eligibility were determined.

- At the evaluation visit (day 1 of the study) all participants ATTR-PN staging was re-evaluated according to the staging criteria published by Coutinho (1980 insert reference). In particular, the ATTR-PN stage was determined primarily by ambulatory status (Stage 1 – fully ambulatory; Stage 2 – required assistance with ambulation; Stage 3 – wheelchair bound).

  In addition, we would like to clarify that in study Fx1A-OS-001, there was only one site involving one principle investigator: Dr Teresa Coelho in Porto, Portugal was involved in this study. Dr Coelho independently determined the staging for all patients in the study according to the Coutinho criteria described above. We can confirm that no retest reliability was carried out regarding the staging.

1.3.2 How was time from symptom onset established, i.e. the date of earliest ATTR-related symptom)?

We can confirm that the time from symptom onset was established in study Fx1A-OS-001 based on the time between first symptom experienced and the date of enrolment into the study based on patient responses.

1.3.3 Could we please have the full statistical analysis plan (Appendix 16.1.9 of the clinical study report).

Please see appendix 9 attached for the full statistical analysis plan for study Fx1A-OS-001.

1.4 Patient population estimates applied in the health economic model

1.4.1 Could further clarification be provided on the exact number (or proportions) of V30M patients and non V30M patients used to estimate the base case results (combined analysis) in the health economic model, and further details on how these have been calculated. In Table 47 it appears to suggest that there are 8 prevalent V30M patients and 10 prevalent non V30M patients but this does not appear to tally with our estimates based on the information provided in Section 1.1.5.

We would like to clarify that in our submission, the most relevant estimate from the perspective of AGNSS in terms of informing the budget impact of tafamidis is the estimate of the number of patients reported in Section 1.1.5 as this only includes patients living in England.

The patient numbers reported in Appendix B: Pharmaco-economic model, Table 47 also includes foreign nationals who could receive a transplant in England and these data are only used to inform the rate of transplant included in our economic model. A more detailed explanation of these calculations is provided in appendix 10.

1.4.2 For the survival analyses presented in tables 48 and 49, how have the base case (combined) survival estimates been calculated? Are they based on the proportions of V30M patients and non V30 patients reported in the two separate studies or are they based on the estimate used to derive the base case cost-effectiveness results (combined analysis) in the health economic model? Please provide further details.

We can confirm that the V30M and non-V30M survival analyses presented in Tables 48 and 49 within our submission are based on the combined pseudo-patient level datasets (See 1.7.1 below) from the associated studies (Disease related survival from Sattianayagam et al and post-liver transplant survival from Helenius et al, 2004). Therefore the proportion of V30 and non-V30M patients reflects the studies.

1.5 THAOS registry

1.5.1 Could more details be provided on the THAOS registry, in particular with regard to:

- How was staging classified and recorded in the THAOS registry?

Please note that there is limited observational data regarding TTR-FAP and that the THAOS registry, whilst limited in the number of observations and lack of follow-up data, does provide the best available evidence to reflect TTR-FAP patients in England. Moreover, the THAOS registry includes patients with V30M and nonV30M mutations from countries where TTR-FAP is endemic and non-endemic in order to reflect the heterogeneity of the TTR-FAP population.

We would like to clarify that staging, as described by Coutinho (1980), was not recorded within the THAOS registry. Rather, the registry collected data on the modified Polyneuropathy Disability Scale (mPDS) to assess walking ability, which was mapped to the Coutinho stages as shown in the table below. The mapping algorithm in SAS was developed by FoldRx, who held the rights of tafamidis prior to Pfizer. This algorithm is applied each time data analyses regarding Coutinho staging are carried out using the THAOS registry.

It should be acknowledged that the mapping process relies solely on mobility whereas the Coutinho staging also assesses other aspects of the illness, such as sensory and autonomic symptoms. To date, the mapping process has only been applied to baseline data, because there are very few follow-up data points.

Table 3: Mapping of the mPDS to the Coutinho Disease Stages in THAOS

mPDS	Coutinho Stage
0 = Normal	None (Coutinho stages only apply when symptomatic)
1= Sensory disturbances in feet but able to walk without difficulty 2 = Some difficulties walking but can	Stage 1

walk without aid	
3a = Able to walk with 1 cane or crutch 3b = Able to walk with 2 canes or crutches	Stage 2
4 = Not ambulatory; confined to wheelchair or bedridden	Stage 3

- Please report both the number of patients and the proportion of V30M and non-V30M patients for each stage used to estimate Table 41.

  We would like to clarify that the data used to populate Table 41 in our submission was based on an analysis including a total of 125 patients. Of these, there were 106 V30M patients and 19 non-V30M patients. Please see table 4 below for a description of the number of patients in each stage.

  Table 4: Patients with baseline TQoL assessments in THAOS, presented by mutation type and disease stage

V30M			Non-V30M		
Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
N=55	N=30	N=21	N=9	N=8	N=2

- Please report the frequency of different mutations within the non-V30M group.

  Please see Appendix 11 attached including a table showing the frequency of different mutations for non-V30M patients with baseline TQoL assessments in THAOS.

- For each stage (1, 2 and 3) could more information be provided on the breakdown of TQoL scores, in particular could the mean, standard deviation, median, interquartile range and minimum and maximum values for TQoL be provided. Could demographic characteristics by stage also be provided.

  For more information concerning the summary statistics for TQoL scores by Coutinho stage, please see Appendix 12 attached.

  For more information regarding the demographic and baseline characteristics of patients with baseline TQoL assessments in THAOS by Coutinho stage, please see Appendix 13 attached.

- Could the data requested above also be provided separately for V30M and non V30M patients.

  For more information on TQoL scores for each stage reported separately for V30M and nonV30M patients, please see appendices 14 and 15 attached.

  For demographic characteristics by stage reported separately for V30M and non-V30M patients, please see appendix 16 attached.

1.5.2 It would be useful to know if the THAOS registry data are collected longitudinally. If so, could justification be given for why the data available has not been used to estimate rates of change in TQoL.

We can confirm that the THAOS registry data are collected longitudinally, but there is currently insufficient follow-up data to estimate rates of change in TQoL using the THAOS registry:

- Although the THAOS registry was established in 2007, most patients were enrolled in the past year and at the time of the AGNSS submission, less than a quarter of the TTR-FAP patients with baseline TQoL assessments (27/125) had data from follow-up visits.

1.6 Relationship between disease duration and TQoL scores (Figure 26)

1.6.1 Could more clarification be provided on what the equation used in Figure 26 represents (i.e. what is the dependent variable and what is the explanatory variable).

We would like to clarify that the equation used in Figure 26 in our submission is based on a multinomial line of best fit plotted against the correlation curve between time since symptom onset and TQoL Score, the original correlation curve was extracted from study FX1A-OS-001 and transposed to switch the variables. The dependant variable is TQOL and the explanatory variable is disease duration.

1.6.2 Could justification also be provided for the choice of functional form, in comparison with other possible forms/models. In doing so could reference be made to relevant goodness of fit statistics including the AIC and BIC statistics.

We are unable to provide further justification for the choice of functional form as the original correlation curve from study FX1A-OS-001 was developed originally by FoldRX (who held the rights to tafamidis prior to Pfizer) and no information is available to us.

For this reason, we have undertaken an additional one-way sensitivity analysis in the tafamidis economic model to explore the impact on the ICER of varying the 6-monthly rate of change in TQoL by ± 50%.

The results below demonstrate that even if the rate of change is reduced by 50% which is an extremely conservative estimate in stage 1, the cost/QALY is still less than £300,000.

Table 5: Results of one-way sensitivity analysis varying six-month rate of change by ±50%

Variable	CE with high value (+ 50%)	CE with low value (- 50%)
6 month rate of change in TQOL - Stage 1 (10.755 to 3.585; base case 7.170)	£158,246	£289,281
6 month rate of change in TQOL - Stage 2 (5.642 to 1.881; base case 3.761)	£178,019	£210,707
6 month rate of change in TQOL - Stage 3 (1.533 to 0.511; base case 1.022)	£186,364	£191,871

1.7 Survival analyses

1.7.1 Could clarification be provided on the survival analyses conducted for the disease related survival and the post liver transplant survival, in particular with regard to:

- If the calculations were based on individual patient data:

We can clarify that the survival analyses conducted for the disease related survival and the post liver transplant survival described in Section 6.1.5 of our submission were based on two separate published data sources. For patients who had not received a liver transplant in the model survival estimates were taken from data reported by Sattianayagam et al, 2011, who reported survival from diagnosis for V30M and T60A (non-V30M) patients. Whereas, Post-liver transplant survival specific to TTR-FAP patients was obtained from the Familial Amyloidotic Polyneuropathy World Transplant Registry (FAPWTR) reported by Helenius et al, 2004.

The Kaplan-Meier curves reported in both papers were then digitised using GetData Graph Digitizer (Version 2.24 – http://getdata-graph-digitizer.com) to extract the survival data from the graphs and the incremental number of events between each data point was estimated using total patient numbers.

The resulting aggregate level data set was used to create a pseudo-patient level dataset using the expandcl command in Stata version 12.0. Note that this method did not account for censoring. In the absence of detailed information on censoring we applied the same approach as detailed in Craig C et al. An assessment of Methods to Combine Published Survival Curves. Medical Decision Making. 20: 104. 2000.

- Were other functional forms considered? If so, could goodness of fit statistics including AIC and BIC statistics be provided to support the use of Weibull functions instead of other functional forms.

A number of different functional forms were considered within the model development phase only and these included the exponential, gompertz, lognormal, loglogistic and weibull distributions. The generalised gamma distribution was not considered because we were unable to generate estimates for some subgroups.

In the majority of the subgroups considered the loglogistic and lognormal distributions provided the lowest AIC score.

Early modelling revealed that the use of loglogistic and lognormal distributions resulted in unrealistic life expectancy i.e., this appeared to be caused as a result of the 'fat-tails' present in these distributions. Therefore, these functional forms were discounted and no further statistical analyses using these forms were undertaken. The decision to fit a weibull distribution was therefore preferred as a more clinically and biologically plausible distribution and because in most instances the estimated AIC statistics were broadly comparable across models. The model summary statistics for all functional models considered for both pre-liver transplant survival and post-liver transplant survival are presented in Tables 6-11 below.

Model summary statistics:

Pre-liver transplant survival

Table 6: Model summary statistics for pre-liver transplant survival curve analysis for combined cohort

Model	Obs	ll(null)	ll(model)	Df	AIC	BIC
exponential	79	-98.73771	-98.73771	1	199.4754	201.8449
gompertz	79	.	-92.64088	2	189.2818	194.0207
lognormal	79	.	-89.096	2	182.192	186.9309
loglogistic	79	.	-88.91095	2	181.8219	186.5608
Weibull	79	-89.33229	-89.33229	2	182.6646	187.4035

Table 7: Model summary statistics for pre-liver transplant survival curve analysis in V30M patients

Model	Obs	ll(null)	ll(model)	Df	AIC	BIC
exponential	27	-30.23417	-30.23417	1	62.46834	63.76418
gompertz	27	.	-19.8112	2	43.62239	46.21407
lognormal	27	.	-19.46951	2	42.93903	45.5307
loglogistic	27	.	-19.19451	2	42.38902	44.9807
Weibull	27	-18.30796	-18.30796	2	40.61591	43.20759

Table 8: Model summary statistics for pre-liver transplant survival curve analysis in non-V30M patients

Model	Obs	ll(null)	ll(model)	Df	AIC	BIC
exponential	52	-67.74285	-67.74285	1	137.4857	139.4369
gompertz	52	.	-67.27876	2	138.5575	142.46
lognormal	52	.	-61.22267	2	126.4453	130.3478
loglogistic	52	.	-60.6276	2	125.2552	129.1577
Weibull	52	-64.45404	-64.45404	2	132.9081	136.8106

Post-liver transplant survival

Table 9: Model summary statistics for post-liver transplant survival curve analysis for combined cohort

Model	Obs	ll(null)	ll(model)	Df	AIC	BIC
exponential	511	-548.0073	-548.0073	1	1098.015	1102.251
gompertz	511	.	-501.5799	2	1007.16	1015.633
lognormal	511	.	-496.6609	2	997.3218	1005.795
loglogistic	511	.	-501.7503	2	1007.501	1015.973
Weibull	511	-503.4395	-503.4395	2	1010.879	1019.352

Table 10: Model summary statistics for post-liver transplant survival curve analysis for V30M patients

Model	Obs	ll(null)	ll(model)	Df	AIC	BIC
exponential	449	-439.1019	-439.1019	1	880.2038	884.3108
gompertz	449	.	-399.4912	2	802.9824	811.1964
lognormal	449	.	-396.9947	2	797.9893	806.2034
loglogistic	449	.	-401.0769	2	806.1537	814.3678
Weibull	449	-402.2806	-402.2806	2	808.5612	816.7753

Table 11: Model summary statistics for post-liver transplant survival curve analysis for non-V30M patients

Model	Obs	ll(null)	ll(model)	Df	AIC	BIC
exponential	62	-99.22075	-99.22075	1	200.4415	202.5686
gompertz	62	.	-92.99567	2	189.9913	194.2456
lognormal	62	.	-90.91073	2	185.8215	190.0757
loglogistic	62	.	-91.80162	2	187.6032	191.8575
Weibull	62	-92.31939	-92.31939	2	188.6388	192.893

- Could the variance covariance matrices for the functions be provided.

Please see the results below presented in Tables 12-17 for the Covariance matrix of coefficients for the weibull model.

Pre-liver transplant survival

Table 12: Covariance matrix of coefficients of weibull model: pre-liver transplant for combined cohort

		_t	ln_p
	e(V)	_cons	_cons
_t	_cons	.13411504	
ln_p	_cons	-.03451339	.01001293

Table 13: Covariance matrix of coefficients of weibull model: pre-liver transplant for V30M patients

		_t	ln_p
	e(V)	_cons	_cons
_t	_cons	.9741731	
ln_p	_cons	-.15164591	.02457646

Table 14: Covariance matrix of coefficients of weibull model: Pre-liver transplant for non-V30M patients

		_t	ln_p
	e(V)	_cons	_cons
_t	_cons	.15094242	
ln_p	_cons	-.04668343	.01730427

Post-liver transplant survival

Table 15: Covariance matrix of coefficients of weibull model: Post-liver transplant for combined cohort

		_t	ln_p
	e(V)	_cons	_cons
_t	_cons	.01179214	
ln_p	_cons	-.00526531	.0074373

Table 16: Covariance matrix of coefficients of weibull model: Post-liver transplant for V30M patients

		_t	ln_p
	e(V)	_cons	_cons
_t	_cons	.01529348	
ln_p	_cons	-.0068386	.00981011

Table 17: Covariance matrix of coefficients of weibull model: Post-liver transplant for non-V30M patients

		_t	ln_p
	e(V)	_cons	_cons
_t	_cons	.05035774	
ln_p	_cons	-.02156044	.02928209

- If the calculations were based on aggregate data could more information be provided on how these were fitted and if other functional forms were considered.

We can confirm that no survival calculations were based on aggregate data; however the pseudo-patient level data sets were derived from the presented aggregate level Kaplan-Meier curves from the published papers by Sattianayagam et al, 2011 and Helenius et al, 2004.

1.7.2 Can a figure similar to Figure 27 showing the survivor function and fitted Weibull survival function be provided for the V30M and non V30M subgroups.

Please see the requested disease-related survival from Sattianayagam et al, with Weibull fitted survival distribution for V30M and non V30M subgroups.

Figure 1: Disease-related survival from diagnosis from Sattianayagam et al, with Weibull fitted survival (V30M)

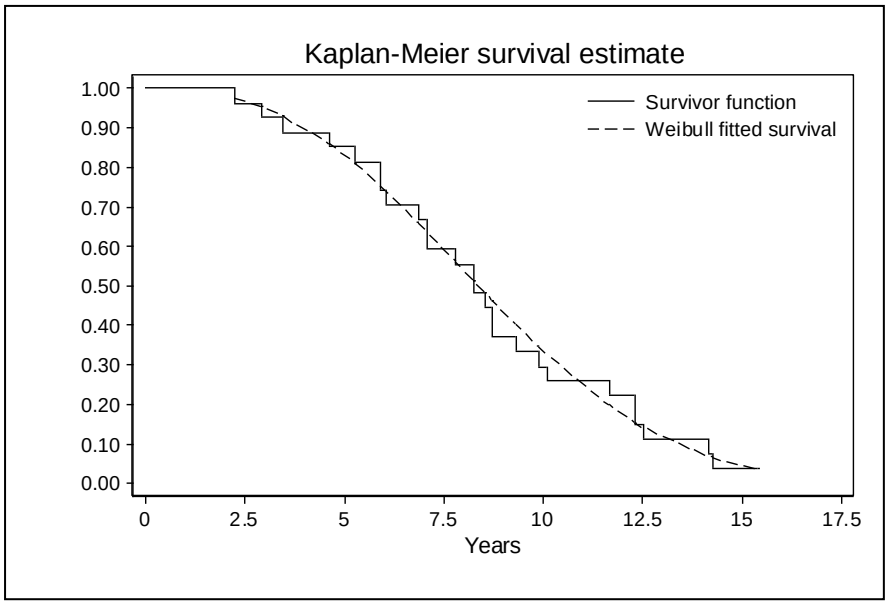
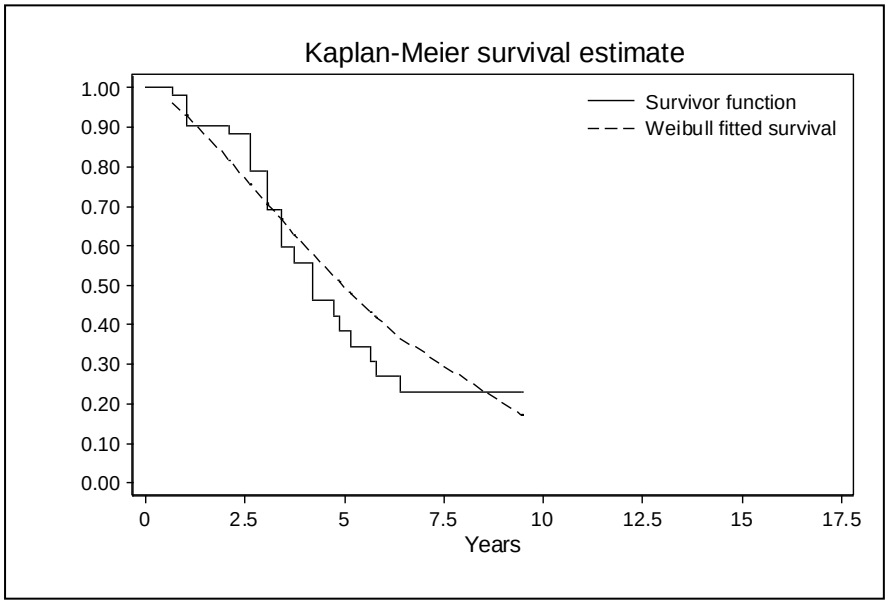


Figure 2: Disease-related survival from diagnosis from Sattianayagam et al, with Weibull fitted survival (non-V30M)



1.7.3 Can Figure 28 be provided, with the actual survival data shown alongside the fitted function. Similarly, can the same figure be provided for the two subgroups.

Please see the requested graphs below showing the actual survival data alongside the fitted weibull function for all patients and for the V30M and non V30M subgroups respectively.

Figure 3: Post-liver transplant survival rates from Helenius et al, with Weibull fitted survival (V30M and non-V30M)

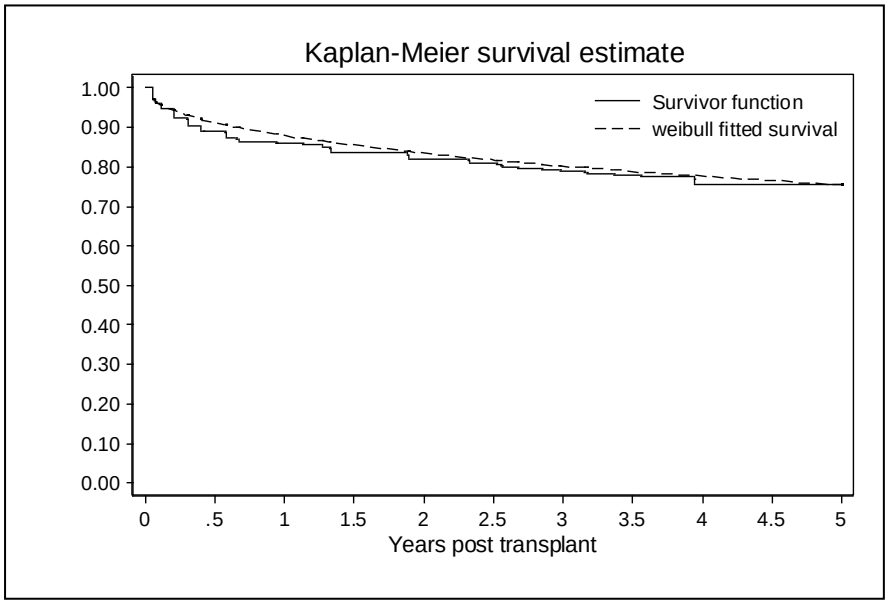


Figure 4: Post-liver transplant survival rates from Helenius et al, with Weibull fitted survival (V30M)

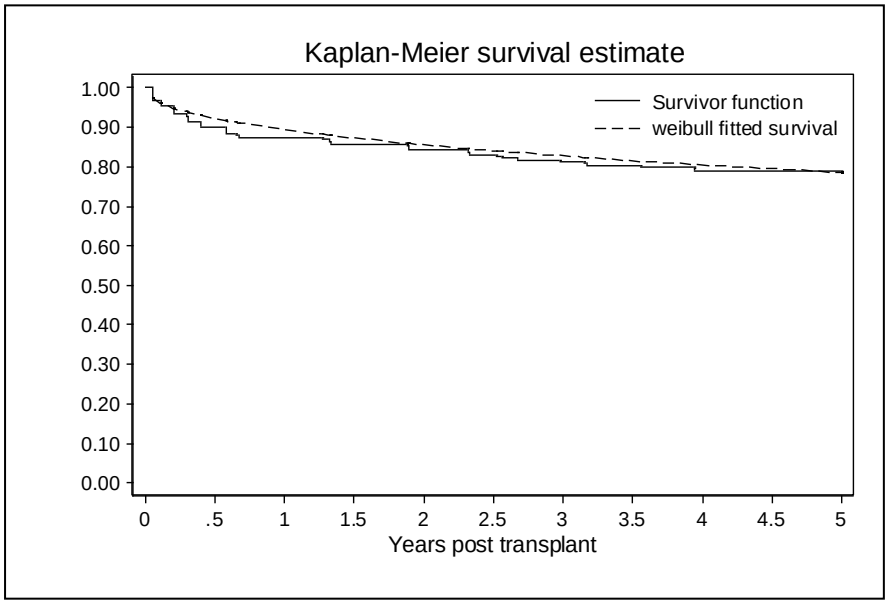
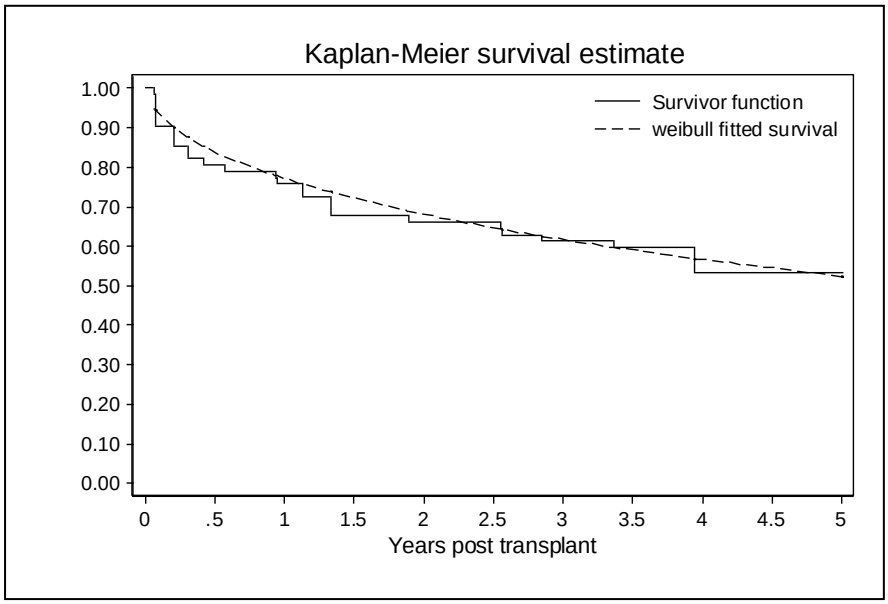


Figure 5: Post-liver transplant survival rates from Helenius et al, with Weibull fitted survival (non-V30M)



1.8 EQ-5D

1.8.1 Could the individual patient level data used to calculate the relationship between EQ-5D and TQoL be provided to the ERG? If this is not possible could the following clarifications and additional analyses be conducted:

- Could justification be provided for the functional form chosen, in particular with regard to other functional forms considered. Your choice of functional form should be supported with provision of relevant goodness of fit statistics for all analyses, including AIC and BIC statistics.
- Could additional analyses be conducted including squared and cubed terms of the explanatory variable (TQoL) and report the regression coefficients.
- Could the variance covariance matrices for all analyses be provided.

We would like to start by clarifying that the relationship between the EQ-5D and TQoL included in the economic model in our submission is based on a cross-sectional analysis of baseline data in the THAOS registry due to the lack of follow-up data in the registry (as described in the answer to question 1.5.2 above).

Therefore, it would not be appropriate to use AIC/BIC statistics and the variance/covariance matrix to assess goodness of fit for different models to map this relationship as these analyses require more than one observation per individual.

The R squared (R^2) and F statistic are more appropriate statistics for cross-sectional data. Accordingly, we have undertaken these analyses for the cubic, quadratic and linear models and these results are presented in appendix 17. The results show that the R-squared for each model are similar.

The results from the additional quadratic and cubed models were then applied to the tafamidis economic model and the results show that these different models have little impact on the ICER for the base case: £180,637 (quadratic model), £184,826 (cubed model) compared to the base case ICER of £189,995 using the linear model in our submission.

Please see tables 18 and 19 below for further detail of these results.

Table 18: Results of tafamidis economic model incorporating the quadratic formula (0.89-0.004*TQoL-0.00002*TQoL2)

	[redacted]					
	[redacted]		[redacted]		[redacted]	
	[redacted]				[redacted]	
[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]

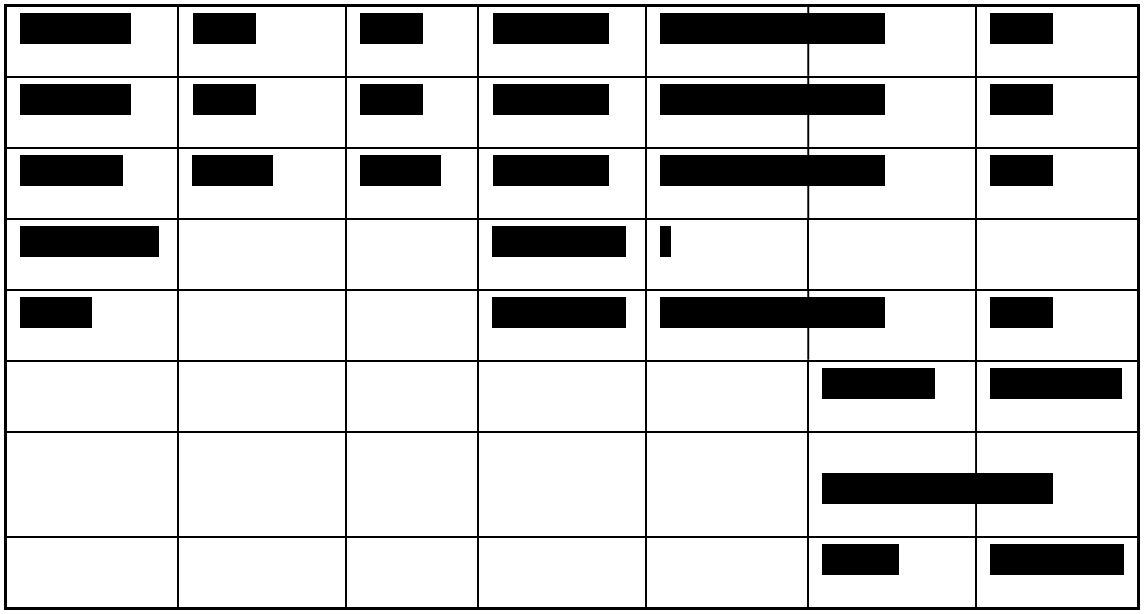
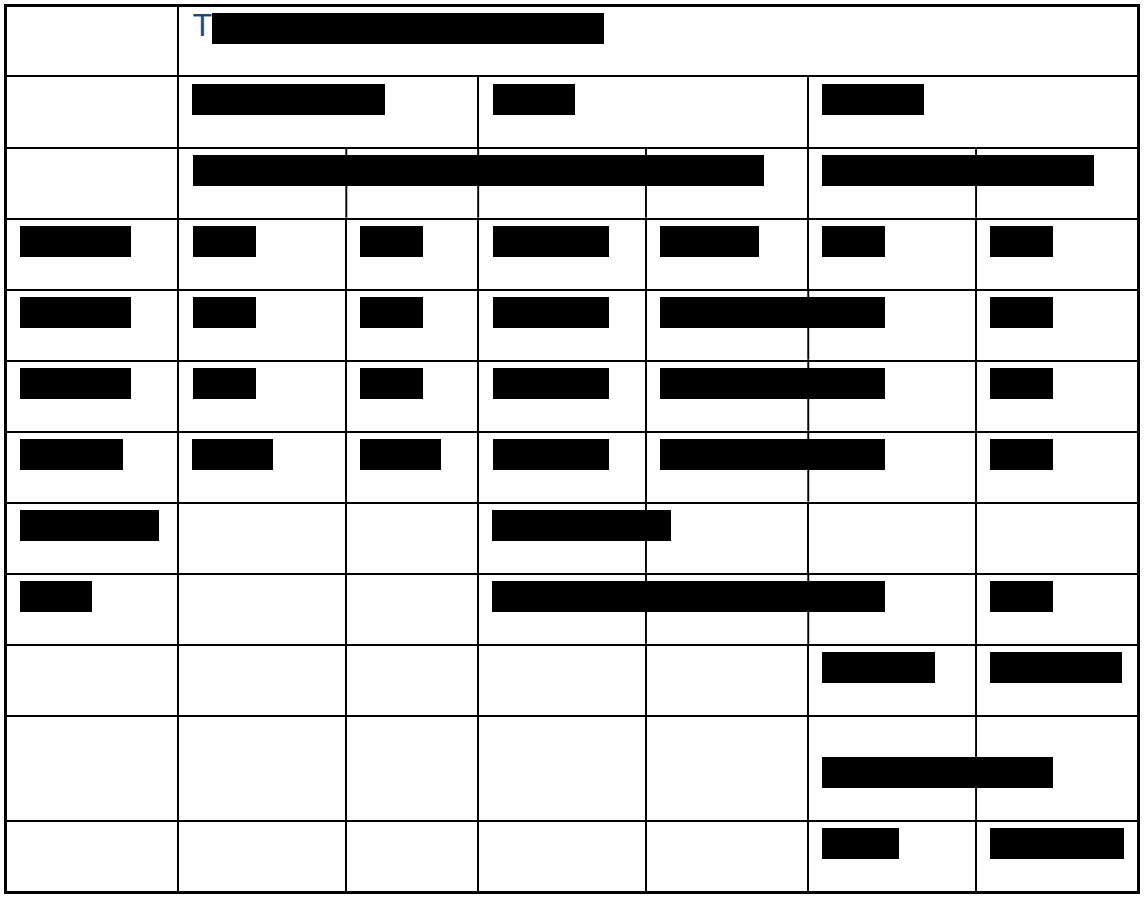


Table 19: Results of tafamidis economic model incorporating the cubic formula (0.90979-0.00712*TQoL+0.00007123*TQoL2-0.000000596927*TQoL3)



- Could additional analyses be conducted where the data is separated into the 3 disease stages based on the cut-offs used in the model, and linear

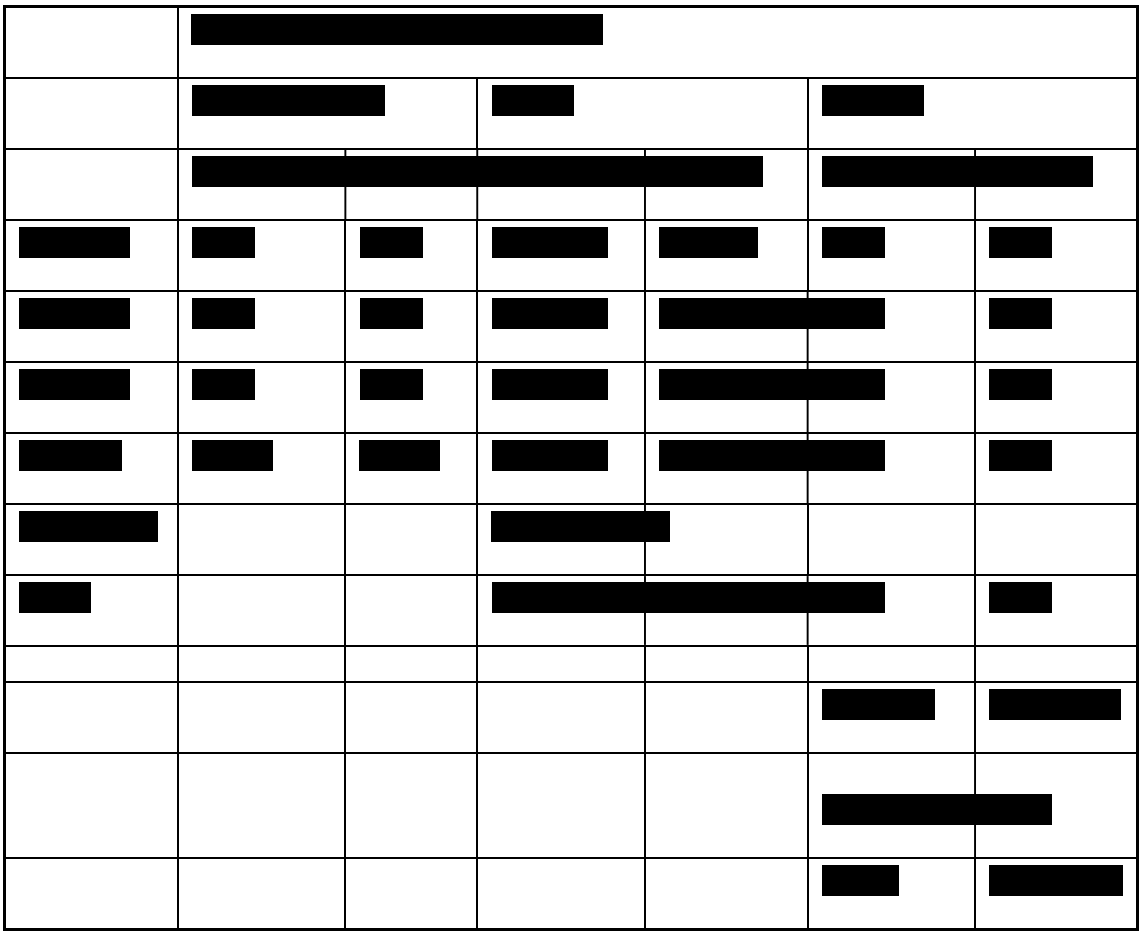
estimates of the relationship between TQoL and EQ-5D be calculated separately for each stage. Please report the regression coefficients.

As described above, the R squared (R^2) and F statistic are more appropriate statistics for cross-sectional data. Accordingly, we have undertaken these analyses for the linear model for each stage and these results are presented in appendix 18.

The data for each stage has then been included in the tafamidis economic model and the results show a slight reduction in the ICER to £152,169 from the original base case ICER of £189,995.

Please see table 19 below for further detail of the model results incorporating linear models for each disease stage.

Table 19: Results of the tafamidis economic model incorporating linear models by disease stage (Stage 1: 0.930807-0.004613*TQoL, Stage 2: 0.861597-0.004278*TQoL, Stage 3: 0.822396-0.006884*TQoL)



1.9 Conditional response rates

1.9.1 Could information on the rate of change in TQoL in the trial be provided separately for Tafamidis patients who were classed as responders and non-responders based on the NIS-LL measure. Furthermore:

- Could this be provided separately for both those who were classed as responders and non-responders at 12 and 18 months.

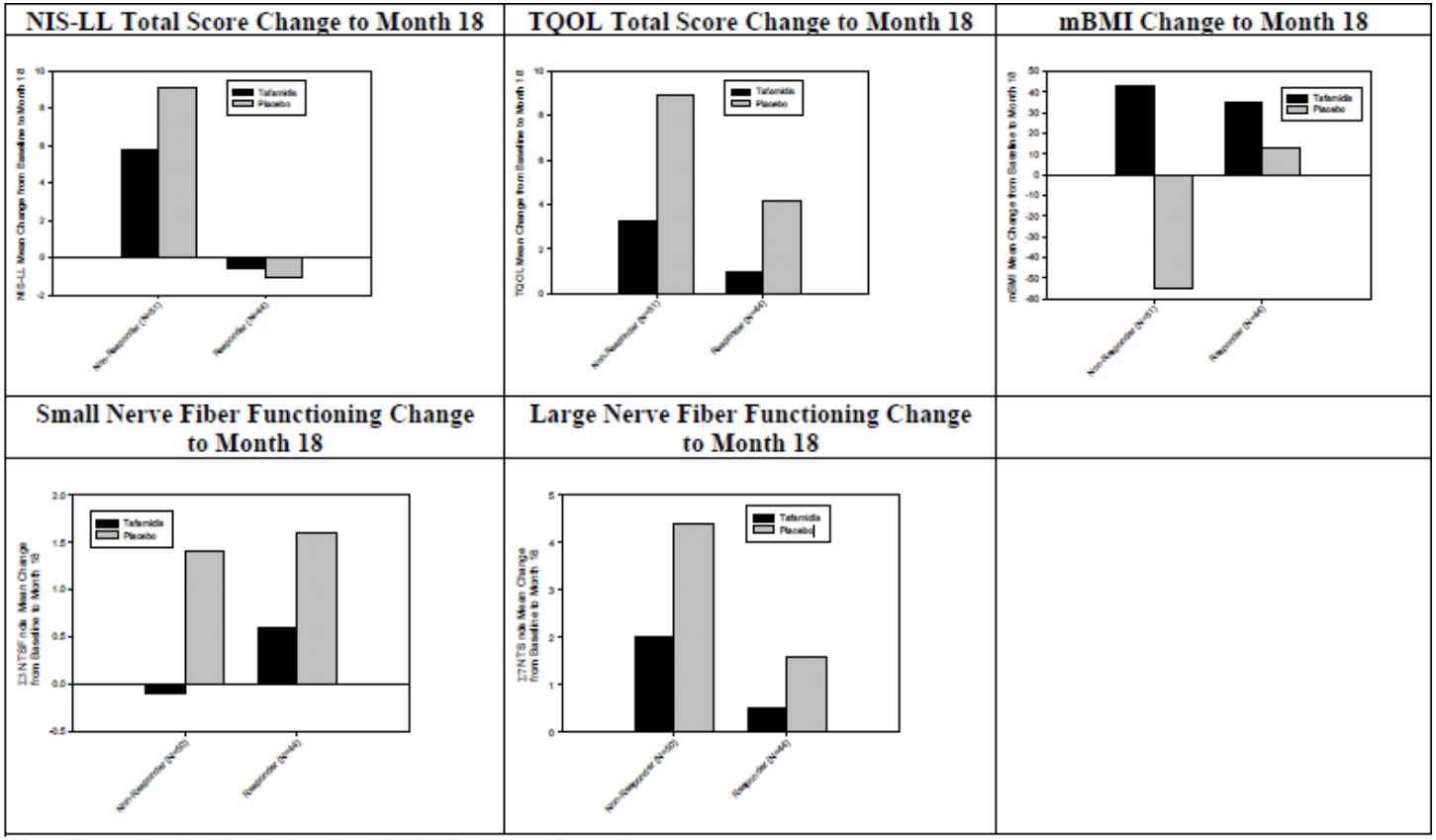
We are able to provide further information concerning the rate of change in TQoL in study Fx-005 by NIS-LL responder status at 18 months only:

- An analysis was carried out to determine whether there were differences in outcomes between tafamidis and placebo in patients categorized as responders (had no disease progression [change in NIS-LL of <2 points]) versus non-responders (had disease progression [change in NIS-LL of ≥2 points]).
- As shown in figure 8, patients receiving tafamidis had better outcomes than patients on placebo (including TQoL), irrespective of responder status.
- These results suggest that no single outcome captures the multi-faceted aspects of overall disease progression in placebo patients, or the full effect of tafamidis treatment.

No data is available for patients who were classified as responders and non-responders at 12 months.

Please see appendix 19 attached giving more information concerning the change of Norfolk QOL-DN from baseline by NIS-LL responder status.

Figure 6: Change from Baseline to Month 18 in Key Efficacy Endpoints by NIS-LL Responder Status (NIS-LL<2 Responder Vs. NIS-LL ≥2 Non-Responder) – ITT Population, Observed Case (Study Fx-005)



Note: Responder classification based on NIS-LL Total Score Responder assessment.

1.10 Electronic model

1.10.1 The ERG is unclear if the model is programmed to examine parameter uncertainty through a probabilistic sensitivity analysis as well as variability in patients' baseline characteristics. The model does not appear to allow for the inclusion of parameter uncertainty as none of the parameters appear to be characterised with uncertainty in the "Abacus_data" worksheet (with the exception of patient baseline characteristics). In the sheet "Abacus_Data" it refers to the "Number of patients" in cell F1 and the "Number of simulations" in cell F2, however, it is unclear to us whether/how the results characterise the uncertainty based on the different simulations. Further clarification is required.

We would like to clarify that in our economic model we have undertaken simple one way sensitivity analyses exploring the impact of changing the baseline characteristics. The results of these analyses are presented in Table 54, page 151 of our submission. We did not consider that additional, sophisticated, analytical techniques, such as probabilistic sensitivity analysis (PSA) were warranted given the underlying uncertainty associated with the data for ultra orphan conditions.

1.11 Home Care provision of Tafamidis

1.12 Home Care provision of Tafamidis

1.12.1 Could more details be provided on the arrangements for home provision. In particular:

- Whether the arrangements have been finalised?

  We would like to clarify that the home care provision for tafamidis have not been finalised. It is likely that the service provided by Home Care will be limited to a delivery service and only include the delivery of tafamidis packs, as prescribed.

  We estimate that these arrangements will be finalised in September/October 2012.

  If so, what are the anticipated frequency of visits and number of tablets provided at each visit?

  It is anticipated that Home care will deliver a supply of tafamidis which is determined by the TTR-FAP specialist prescribing the treatment.

- If home care provision is not provided, is the anticipated frequency of visits to clinic/hospital and number of tablets provided at each visit?

Home Care provides the most efficient way of delivering tafamidis capsules to patients. It is likely that most patients will only see a specialist at the NAC once or twice a year. It is envisaged that the prescription of tafamidis are likely to be issued more frequently than this, yet it is not efficient for patients in other areas of the country to travel to the NAC in London solely to receive a prescription. Retail pharmacists will not stock or obtain tafamidis. Home Care provides a solution to ensure that patients receive tafamidis tablets at regular intervals as advised by the TTR-FAP specialist.

- Finally, could any tablets that have not been used because of withdrawal or other reasons be provided to another patient, and if so have arrangements for collection/re-use been made?

  Whilst we appreciate the potential cost benefits of this type of arrangement, we are concerned that the capsules may not be stored appropriately or could become contaminated. Therefore, this type of arrangement would not be considered appropriate by Pfizer due to safety concerns.

Appendix 5: Quality Assessment using an economic modelling checklist [49]

Item	Critical Appraisal	Reviewer Comment
Was a well-defined question posed in answerable form?	Yes	The purpose of the economic evaluation was to determine the cost-effectiveness of tafamidis over the lifetime of patients with TTR-FAP in England compared with conventional support therapy, including liver transplant for eligible patients.
Was a comprehensive description of the competing alternatives given (i.e. can you tell who did what to whom, where, and how often)?	Yes	Tafamidis 20mg once daily is licensed for the treatment of TTR-FAP in patients in stage 1 of the disease. The comparator is conventional support therapy, including liver transplant for eligible patients.
Was the effectiveness of the programme or services established?	No	There is considerable uncertainty regarding the effectiveness of tafamidis in the patient population in England.
Were all the important and relevant costs and consequences for each alternative identified?	Yes	Costs were identified from the perspective of the NHS & PSS, as well as patients and their carers. Consequences were identified in terms of health for the patient and their carer.
Were costs and consequences measured accurately in appropriate physical units (e.g. hours of nursing time, number of physician visits, lost work-days, gained life years)?	Yes	Resource use estimates were obtained from Swedish physicians. The UK clinicians contacted by the ERG confirmed the resource use estimates as reasonable for UK clinical practice. Productivity costs borne by the patient and their carer were also included. Consequences were measured in quality-adjusted life years (QALYs).
Were the cost and consequences valued credibly?	Yes	Costs were estimated by applying UK unit costs to resource use data. QALYs were estimated using EQ-5D as the measure of HRQoL. EQ-5D scores were derived from the TQoL using a mapping function.
Were costs and consequences adjusted for differential timing?	Yes	Costs and consequences were discounted at 3.5% per annum.
Was an incremental analysis of costs and consequences of alternatives performed?	Yes	The results were presented as incremental cost-effectiveness ratios.
Was allowance made for uncertainty in the estimates of costs and consequences?	?	Some one way sensitivity analyses have been conducted. However, many areas of uncertainty have been ignored. A probabilistic sensitivity analysis was also not conducted.
Did the presentation and discussion of study results include all issues of	No	Many of the ERGs concerns were not discussed.

concern to users?		

Appendix 6 – Validation of ERG's model

Table 45 presents the results of the comparison of the manufacturer's model with the ERG's model without correction for the methodological issues identified by the ERG and discussed in Section 5.10. The estimates of mean costs and mean QALYs are not coincidental due to the formal of each model. While the manufacturer's individual patient level simulation samples from the distribution of ages and baseline TQoL scores, the ERG's cohort model uses the mean age and baseline TQoL for the population. In light of these results, the ERG is confident that the ERG's cohort model accurately reflects the manufacturer's individual patient simulation.

Table 45: Comparison of results between the manufacturer's model and the ERG's model without correction for the methodological issues identified by the ERG

Intervention	Mean costs (£)	Mean QALYs
Combined V30M and non-V30M patient population: Manufacturer's model		
CST	£187,000	2.57
Tafamidis	£278,000	3.16
Combined V30M and non-V30M patient population: ERG's model		
CST	£188,000	2.52
Tafamidis	£271,000	3.16

Appendix 7 – Calculations for exploratory scenario analyses

For scenario 4B: Cut-off TQoL score between stages defined as half-way between the mean TQoL score of stage N and stage N+1
AND
TQoL rate of change recalculated

V30M				
Stage 1	TQoL	10	58	points
up to 59	Disease duration	0.0519	3.84006	years
	rate	6.3355		
Stage 2	TQoL	59	81	points
from 59 to 82	Disease duration	3.90967	6.42273	years
	rate	4.37713		
Stage 3	TQoL	82	135	points
from 82 onwards	Disease duration	6.59814	30.03315	years
	rate	1.1308		
Non-V30M				
Stage 1	TQoL	10	65	points
until 66	Disease duration	0.0519	4.38865	years
	rate	6.3412		
Stage 2	TQoL	66	89	points
from 66 to 90	Disease duration	4.48038	8.02537	years
	rate	3.2440		
Stage 3	TQoL	90	135	points
from 90 onwards	Disease duration	8.2599	30.03315	years
	rate	1.0334		

For scenario 5B: Mean of TQoL score of stage N+1
AND
TQoL rate of change recalculated

V30M				
Stage 1	TQoL	10	68	points
up to 69	Disease duration	0.0519	4.67566	years
	rate	6.2720		
Stage 2	TQoL	69	94	points
from 69 to 95	Disease duration	4.77957	9.28182	years
	rate	2.7764		
Stage 3	TQoL	95	135	points
from 95	Disease duration	9.55915	30.03315	years
	rate	0.9768		
Non-V30M				
Stage 1	TQoL	10	85	points
up to 86	Disease duration	0.0519	7.16565	years
	rate	5.2715		
Stage 2	TQoL	86	92	points

from 86 to 93	Disease duration	7.36918	8.75374	years
	rate	2.1668		
Stage 3	TQoL	93	135	points
from 93	Disease duration	9.01341	30.03315	years
	rate	0.9991		

For scenario 7: TQoL rate of change is independent of stage

TQoL	10	135	points
Disease duration	0.0519	30.03315	years
rate	2.084636		

Appendix 8 – Cost-effectiveness results for exploratory scenario analyses for different baseline TQoL scores

Scenario 1 Inclusion of productivity costs

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£244,148	3.38	£1,182,293	4.27	£1,061,844
10	£174,780	4.32	£1,053,935	5.72	£630,609
20	£195,471	4.09	£1,069,058	5.31	£716,835
30	£208,695	3.81	£1,094,706	4.91	£800,053
40	£231,462	3.63	£1,133,464	4.56	£976,730
Non-V30M population					
44.89	£155,291	2.58	£804,988	3.16	£1,114,123
10	£107,766	3.26	£741,075	4.11	£748,364
20	£123,548	3.07	£747,828	3.82	£835,029
30	£133,814	2.85	£760,255	3.53	£917,486
40	£154,182	2.71	£783,566	3.27	£1,117,578
Combined population					
45.68	£170,101	2.71	£867,872	3.35	£1,105,410
10.00	£118,935	3.44	£793,218	4.38	£728,738
20.00	£135,535	3.24	£801,366	4.07	£815,330
30.00	£146,294	3.01	£815,997	3.76	£897,914
40.00	£167,062	2.86	£841,882	3.49	£1,094,103

Scenario 2 Patients remain on tafamidis throughout their lifetime

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.38	£1,075,650	4.27	£1,074,673
10	£88,571	4.32	£1,013,203	5.72	£663,229
20	£99,260	4.09	£1,020,647	5.31	£756,057
30	£107,234	3.81	£1,033,172	4.91	£836,107
40	£119,096	3.63	£1,051,986	4.56	£1,010,175
Non-V30M population					
44.89	£79,466	2.58	£744,178	3.16	£1,139,713
10	£54,085	3.26	£712,602	4.11	£778,151
20	£62,121	3.07	£715,934	3.82	£874,528
30	£67,956	2.85	£722,067	3.53	£957,994
40	£78,357	2.71	£733,619	3.27	£1,163,424
Combined population					
45.68	£87,248	2.71	£799,423	3.35	£1,128,873
10.00	£59,833	3.44	£762,702	4.38	£758,997
20.00	£68,311	3.24	£766,720	4.07	£854,783
30.00	£74,502	3.01	£773,918	3.76	£937,680
40.00	£85,147	2.86	£786,680	3.49	£1,137,883

Scenario 3 Patients remains on tafamidis only during stage 1

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.38	£303,722	3.66	£635,218
10	£88,571	4.32	£979,470	5.7	£644,698
20	£99,260	4.09	£900,258	5.25	£688,809

30	£107,234	3.81	£765,359	4.76	£687,850
40	£119,096	3.63	£561,238	4.23	£744,412
Non-V30M population					
44.89	£79,466	2.58	£352,916	2.9	£834,830
10	£54,085	3.26	£690,945	4.1	£762,648
20	£62,121	3.07	£655,532	3.78	£830,629
30	£67,956	2.85	£590,121	3.46	£861,819
40	£78,357	2.71	£466,448	3.1	£978,515
Combined population					
45.68	£87,248	2.71	£344,717	3.03	£801,561
10.00	£59,833	3.44	£739,033	4.37	£742,990
20.00	£68,311	3.24	£696,320	4.03	£806,992
30.00	£74,502	3.01	£619,327	3.68	£832,824
40.00	£85,147	2.86	£482,246	3.29	£939,498

Scenario 4A Cut-off TQoL score between stages defined as half-way between the mean TQoL score of stage N and stage N+1

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£129,339	3.44	£1,054,766	4.21	£1,206,863
10	£93,006	4.39	£1,012,030	5.71	£692,375
20	£99,324	4.07	£1,016,035	5.3	£749,521
30	£107,873	3.81	£1,024,580	4.89	£848,257
40	£119,832	3.63	£1,042,258	4.52	£1,035,570
Non-V30M population					
44.89	£75,840	2.52	£721,316	3.09	£1,130,235
10	£51,574	3.23	£711,310	4.11	£754,263
20	£56,386	2.99	£712,141	3.81	£805,820
30	£61,992	2.77	£714,675	3.51	£877,357
40	£71,295	2.6	£719,114	3.23	£1,031,707
Combined population					
45.68	£84,757	2.67	£776,891	3.28	£1,143,006
10.00	£58,479	3.42	£761,430	4.38	£743,948
20.00	£63,542	3.17	£762,790	4.06	£796,437
30.00	£69,639	2.94	£766,326	3.74	£872,507
40.00	£79,385	2.77	£772,971	3.45	£1,032,351

Scenario 4B Cut-offs halfway between means and TQoL rate of change recalculated

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£128,114	3.29	£1,074,606	4.25	£994,714
10	£88,571	4.43	£1,011,609	5.8	£673,395
20	£95,991	4.1	£1,016,052	5.38	£718,788
30	£107,234	3.81	£1,027,379	4.96	£795,635
40	£115,820	3.52	£1,047,643	4.58	£880,210
Non-V30M population					
44.89	£77,373	2.66	£740,137	3.19	£1,257,552
10	£52,554	3.4	£711,816	4.16	£867,244
20	£57,013	3.15	£713,882	3.86	£920,461
30	£66,324	2.96	£719,082	3.57	£1,066,267
40	£72,502	2.74	£730,658	3.3	£1,163,582
Combined population					
45.68	£85,830	2.77	£795,882	3.37	£1,213,746
10.00	£58,557	3.57	£761,782	4.43	£834,936
20.00	£63,509	3.31	£764,244	4.11	£886,849
30.00	£73,142	3.10	£770,465	3.80	£1,021,162
40.00	£79,722	2.87	£783,489	3.51	£1,116,353

Scenario 5A Cut-off TQoL score between stages defined as the mean of stage N+1

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£117,775	3.18	£1,039,293	4.1	£1,000,815
10	£81,293	4.2	£1,010,542	5.71	£614,361
20	£91,401	3.94	£1,012,031	5.29	£680,653
30	£99,324	3.65	£1,016,049	4.87	£749,531
40	£105,918	3.34	£1,024,806	4.46	£820,920
Non-V30M population					
44.89	£66,175	2.39	£705,137	3.06	£951,363
10	£45,490	3.16	£710,658	4.1	£705,265
20	£49,870	2.92	£710,376	3.8	£745,613
30	£56,545	2.7	£709,663	3.5	£814,871
40	£62,663	2.49	£706,678	3.21	£900,539
Combined population					
45.68	£74,775	2.52	£760,830	3.23	£959,605
10.00	£51,457	3.33	£760,639	4.37	£690,114
20.00	£56,792	3.09	£760,652	4.05	£734,786
30.00	£63,675	2.86	£760,727	3.73	£803,981
40.00	£69,872	2.63	£759,699	3.42	£887,269

Scenario 5B Cut-offs mean of stage N+1 and TQoL rate of change recalculated

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£120,981	3.62	£1,075,638	4.36	£1,284,583
10	£80,013	4.54	£1,011,609	5.8	£738,339
20	£89,869	4.28	£1,016,052	5.39	£838,564
30	£101,576	4.06	£1,027,385	4.99	£1,001,672
40	£109,162	3.77	£1,047,768	4.63	£1,085,558
Non-V30M population	Non-V30M population				
44.89	£73,541	2.85	£736,779	3.24	£1,693,232
10	£44,905	3.55	£711,150	4.22	£994,379
20	£51,327	3.32	£712,295	3.92	£1,097,547
30	£59,215	3.11	£715,933	3.63	£1,265,053
40	£68,290	2.92	£725,883	3.35	£1,531,900
Combined population					
45.68	£81,448	2.98	£793,256	3.43	£1,625,124
10.00	£50,756	3.72	£761,227	4.48	£951,706
20.00	£57,751	3.48	£762,921	4.17	£1,054,383
30.00	£66,275	3.27	£767,842	3.86	£1,221,156
40.00	£75,102	3.06	£779,531	3.56	£1,457,510

Scenario 6 TQoL rate of change for stage 1 is that observed in the placebo arm of Fx-005

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£119,096	3.56	£1,060,969	4.42	£1,097,917
10	£49,842	5.35	£1,010,250	6.16	£1,196,336
20	£59,277	4.91	£1,010,250	5.73	£1,158,129
30	£76,651	4.45	£1,010,456	5.3	£1,094,079
40	£93,161	4	£1,017,367	4.87	£1,054,613

Non-V30M population					
44.89	£67,087	2.8	£721,963	3.31	£1,289,131
10	£31,916	3.9	£710,732	4.36	£1,455,270
20	£36,609	3.59	£710,749	4.06	£1,423,700
30	£46,510	3.28	£711,075	3.76	£1,362,145
40	£57,597	2.96	£714,449	3.46	£1,310,517
Combined population					
45.68	£75,755	2.93	£778,464	3.50	£1,257,262
10.00	£34,904	4.14	£760,652	4.66	£1,412,114
20.00	£40,387	3.81	£760,666	4.34	£1,379,438
30.00	£51,534	3.48	£760,972	4.02	£1,317,467
40.00	£63,524	3.13	£764,935	3.70	£1,267,866

Scenario 7 Uniform rate of change for all stages equivalent to the average TQoL rate of change

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£120,941	3.24	£1,065,611	4.34	£853,668
10	£63,741	4.92	£1,010,255	6.03	£845,897
20	£77,623	4.49	£1,010,456	5.61	£832,338
30	£88,571	4.06	£1,013,989	5.18	£826,326
40	£107,234	3.64	£1,030,147	4.75	£826,019
Non-V30M population					
44.89	£73,485	2.6	£730,357	3.24	£1,023,777
10	£38,922	3.64	£710,785	4.29	£1,030,317
20	£46,950	3.34	£711,074	3.99	£1,020,033
30	£54,085	3.04	£712,949	3.69	£1,015,978
40	£67,956	2.75	£720,361	3.39	£1,011,653
Combined population					
45.68	£81,394	2.71	£786,233	3.42	£995,426
10.00	£43,059	3.85	£760,697	4.58	£999,580
20.00	£52,062	3.53	£760,971	4.26	£988,751
30.00	£59,833	3.21	£763,122	3.94	£984,369
40.00	£74,502	2.90	£771,992	3.62	£980,714

Scenario 7 TQoL rate of change independent of stage

	CST		Tafamidis		ICER
TQoL	Costs	QALYs	Costs	QALYs	(/QALY)
V30M population					
49.64	£103,666	3.95	£1,051,985	4.53	£1,608,195
10	£38,815	5.64	£1,010,250	6.23	£1,665,221
20	£46,574	5.22	£1,010,250	5.8	£1,651,726
30	£60,679	4.79	£1,010,252	5.37	£1,626,435
40	£81,666	4.36	£1,012,031	4.95	£1,589,343
Non-V30M population					
44.89	£58,694	3.01	£716,843	3.36	£1,915,439
10	£26,720	4.06	£710,731	4.4	£2,000,914
20	£30,250	3.76	£710,732	4.1	£1,989,954
30	£37,484	3.46	£710,766	3.8	£1,967,035
40	£50,949	3.16	£712,036	3.5	£1,926,990
Combined population					
45.68	£66,189	3.17	£772,700	3.56	£1,864,232

	10.00	£28,736	4.32	£760,651	4.71	£1,944,965
	20.00	£32,971	4.00	£760,652	4.38	£1,933,583
	30.00	£41,350	3.68	£760,680	4.06	£1,910,268
	40.00	£56,069	3.36	£762,035	3.74	£1,870,716

Scenario 8 Doubled mortality risk

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£96,442	2.87	£857,490	3.48	£1,238,949
10	£61,145	3.72	£808,294	4.7	£762,626
20	£71,064	3.5	£810,890	4.36	£858,952
30	£78,160	3.25	£818,629	4.03	£941,777
40	£89,656	3.09	£834,512	3.73	£1,155,391
Non-V30M population					
44.89	£47,127	1.88	£491,331	2.18	£1,501,709
10	£27,398	2.43	£475,133	2.85	£1,055,949
20	£33,228	2.27	£475,651	2.65	£1,157,131
30	£37,412	2.1	£477,525	2.45	£1,245,946
40	£46,572	1.98	£483,574	2.26	£1,535,448
Combined population					
45.68	£55,346	2.05	£552,358	2.40	£1,457,916
10.00	£33,023	2.65	£530,660	3.16	£1,007,062
20.00	£39,534	2.48	£531,524	2.94	£1,107,435
30.00	£44,203	2.29	£534,376	2.71	£1,195,251
40.00	£53,753	2.17	£542,064	2.51	£1,472,105

Scenario 9 Patients are eligible for liver transplantation during stage 1 at the rate used in the manufacturer's submission

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.38	£966,745	4.68	£645,281
10	£97,123	5.99	£674,720	8.76	£208,137
20	£102,991	5.22	£683,475	7.99	£209,645
30	£108,704	4.63	£711,045	7.08	£246,235
40	£118,853	3.91	£787,450	5.95	£327,206
Non-V30M population					
44.89	£79,790	2.68	£646,416	3.48	£715,273
10	£60,541	3.9	£540,178	5.08	£409,276
20	£65,764	3.52	£545,646	4.68	£412,581
30	£70,144	3.18	£560,070	4.26	£454,822
40	£78,738	2.82	£598,833	3.77	£544,785
Combined population					
45.68	£87,518	2.80	£699,804	3.68	£703,608
10.00	£66,638	4.25	£562,602	5.69	£375,753
20.00	£71,969	3.80	£568,618	5.23	£378,758
30.00	£76,571	3.42	£585,233	4.73	£420,058
40.00	£85,424	3.00	£630,269	4.13	£508,522

Scenario 10 Liver transplant stage 1 and 2

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£129,170	4.97	£707,660	6.48	£382,233
10	£104,439	6.83	£673,608	8.8	£289,636
20	£111,030	6.37	£675,590	8.16	£314,847
30	£114,971	5.79	£680,335	7.54	£323,519
40	£123,903	5.42	£690,262	6.96	£368,307
44.89	£81,921	3.15	£557,652	3.88	£657,439
10	£62,631	4.14	£538,794	5.09	£502,040
20	£68,470	3.86	£539,972	4.72	£548,820
30	£72,564	3.55	£542,838	4.36	£577,905
40	£81,392	3.34	£549,833	4.02	£681,463
Combined population					
45.68	£89,796	3.45	£582,653	4.31	£611,571
10.00	£69,599	4.59	£561,263	5.71	£466,639
20.00	£75,563	4.28	£562,575	5.29	£509,825
30.00	£79,632	3.92	£565,754	4.89	£535,507
40.00	£88,477	3.69	£573,238	4.51	£629,270

Scenario 11 Liver transplant throughout lifetime

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£135,065	5.36	£707,662	6.48	£511,076
10	£107,235	6.99	£673,608	8.8	£313,366
20	£114,363	6.56	£675,590	8.16	£351,183
30	£119,529	6.08	£680,335	7.54	£384,381
40	£129,124	5.77	£690,262	6.96	£471,154
44.89	£82,988	3.25	£557,653	3.88	£752,880
10	£63,140	4.17	£538,794	5.09	£521,113
20	£69,070	3.91	£539,972	4.72	£577,606
30	£73,378	3.62	£542,839	4.36	£627,172
40	£82,329	3.41	£549,834	4.02	£768,651
Combined population					
45.68	£91,668	3.60	£582,655	4.31	£712,579
10.00	£70,489	4.64	£561,263	5.71	£486,489
20.00	£76,619	4.35	£562,575	5.29	£539,869
30.00	£81,070	4.03	£565,755	4.89	£586,707
40.00	£90,128	3.80	£573,239	4.51	£719,068

Scenario12. Acquisition costs of the drug increased by 20%.

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.38	£1,270,626	4.27	£1,295,370
10	£88,571	4.32	£1,208,429	5.72	£803,262
20	£99,260	4.09	£1,215,873	5.31	£916,253
40	£119,096	3.63	£1,247,190	4.56	£1,221,552
50	£128,114	3.4	£1,270,626	4.25	£1,352,277
Non-V30M population					

44.89	£79,466	2.58	£880,783	3.16	£1,374,126
10	£54,085	3.26	£849,937	4.11	£940,436
20	£62,121	3.07	£853,242	3.82	£1,058,193
30	£67,956	2.85	£859,288	3.53	£1,158,987
40	£78,357	2.71	£870,486	3.27	£1,406,560
Combined population					
45.68	£87,248	2.71	£945,757	3.35	£1,361,000
10.00	£59,833	3.44	£909,686	4.38	£917,574
20.00	£68,311	3.24	£913,681	4.07	£1,034,536
30.00	£74,502	3.01	£920,806	3.76	£1,134,555
40.00	£85,147	2.86	£933,270	3.49	£1,375,725

Scenario 13 Quadratic mapping function

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.19	£1,075,441	4.29	£865,908
10	£88,571	4.24	£1,013,203	5.74	£614,780
20	£99,260	4	£1,020,647	5.36	£674,733
30	£107,234	3.69	£1,033,172	4.98	£717,275
40	£119,096	3.5	£1,051,967	4.61	£840,116
44.89	£79,466	2.49	£743,561	3.19	£941,271
10	£54,085	3.22	£712,593	4.12	£736,649
20	£62,121	3.03	£715,902	3.85	£795,261
30	£67,956	2.8	£721,963	3.58	£834,790
40	£78,357	2.64	£733,221	3.31	£979,351
Combined population					
45.68	£87,248	2.61	£798,874	3.37	£928,711
10.00	£59,833	3.39	£762,695	4.39	£716,338
20.00	£68,311	3.19	£766,693	4.10	£775,173
30.00	£74,502	2.95	£773,831	3.81	£815,204
40.00	£85,147	2.78	£786,345	3.53	£956,145

Scenario 14 Cubic mapping function

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.43	£1,075,441	4.53	£865,945
10	£88,571	4.4	£1,013,203	5.72	£701,012
20	£99,260	4.18	£1,020,647	5.38	£770,128
30	£107,234	3.9	£1,033,172	5.07	£794,244
40	£119,096	3.74	£1,051,967	4.78	£893,848
44.89	£79,466	2.66	£743,561	3.32	£991,874
10	£54,085	3.31	£712,593	4.09	£838,479
20	£62,121	3.14	£715,902	3.85	£918,150
30	£67,956	2.93	£721,963	3.62	£942,907
40	£78,357	2.8	£733,221	3.41	£1,064,938
Combined population					
45.68	£87,248	2.79	£798,874	3.52	£970,886
10.00	£59,833	3.49	£762,695	4.36	£815,568
20.00	£68,311	3.31	£766,693	4.11	£893,480
30.00	£74,502	3.09	£773,831	3.86	£918,130
40.00	£85,147	2.96	£786,345	3.64	£1,036,423

Scenario 15 Using the manufacturer's assumptions regarding liver transplantation and stopping rules for tafamidis.

	CST		Tafamidis		
TQoL	Costs	QALYs	Costs	QALYs	ICER (/QALY)
V30M population					
49.64	£126,159	3.38	£292,088	4.16	£214,197
10	£97,123	5.99	£667,530	8.76	£205,739
20	£102,991	5.22	£645,070	7.97	£197,041
30	£108,704	4.63	£591,490	7.01	£202,940
40	£118,853	3.91	£480,884	5.75	£197,026
44.89	£79,790	2.68	£329,971	3.27	£427,561
10	£60,541	3.90	£533,730	5.07	£404,926
20	£65,764	3.52	£520,985	4.67	£395,996
30	£70,144	3.18	£489,989	4.22	£405,134
40	£78,738	2.82	£414,313	3.66	£399,743
Combined population					
45.68	£87,518	2.80	£323,657	3.42	£392,001
10.00	£66,638	4.25	£556,030	5.69	£371,729
20.00	£71,968	3.80	£541,666	5.22	£362,837
30.00	£76,571	3.42	£506,905	4.68	£371,435
40.00	£85,424	3.00	£425,408	4.01	£365,957

Appendix 9 – Calculations for sensitivity analysis to budget impact

	Year 1	Year 2	Year 3	Year 4	Year 5
Scenario 1: Patients remain on tafamidis for stage 1 and 2					
Prevalence	17	27	37	47	57
Incidence	10	10	10	10	10
Patients who have moved to stage 2	0	0			
Total eligible patients	27	37	47	57	67
Treatment uptake	25%	30%	40%	50%	60%
Prevalent treated patients	4	7	10	14	19
Incident treated patients	3	3	4	5	6
Patients who discontinue (due to progression to stage 2)	0	0	0	0	0
Total patients treated	7	10	14	19	25
Scenario 2: 100% uptake					
Prevalence	17	27	37	30	30
Incidence	10	10	10	10	10
Patients who have moved to stage 2	0	0	17	10	10
Total eligible patients	27	37	30	30	30
Treatment uptake	100%	100%	100%	100%	100%
Prevalent treated patients	17	27	37	30	30
Incident treated patients	10	10	10	10	10
Patients who discontinue (due to progression to stage 2)	0	0	17	10	10
Total patients treated	27	37	30	30	30
Scenario 3: 20% increase in acquisition costs					
Prevalence	17	27	37	30	30
Incidence	10	10	10	10	10
Patients who have moved to stage 2	0	0	17	10	10
Total eligible patients	27	37	30	30	30
Treatment uptake	25%	30%	40%	50%	60%
Prevalent treated patients	4	7	10	10	12
Incident treated patients	3	3	4	5	6
Scenario 4:Worst case scenario					
Prevalence	17	27	37	47	57
Incidence	10	10	10	10	10
Patients who have moved to stage 2	0	0	0	0	0
Total eligible patients	27	37	47	57	67
Treatment uptake	100%	100%	100%	100%	100%
Prevalent treated patients	17	27	37	47	57
Incident treated patients	10	10	10	10	10