

Predictive model of response to tafamidis in hereditary ATTR polyneuropathy

Cecília Monteiro,¹ Jaleh S. Mesgazardeh,¹ João Anselmo,² Joana Fernandes,² Marta Novais,² Carla Rodrigues,² Gabriel J. Brighty,¹ David L. Powers,³ Evan T. Powers,¹ Teresa Coelho,^{2,4} and Jeffery W. Kelly¹

¹Departments of Chemistry and Molecular Medicine, The Scripps Research Institute, La Jolla, California, USA. ²Unidade Corino de Andrade, Centro Hospitalar do Porto, Porto, Portugal. ³Department of Mathematics, Clarkson University, Potsdam, New York, USA. ⁴Department of Neurophysiology, Centro Hospitalar do Porto, Porto, Portugal.

Conflict of interest: Conflict of interest statement: JWK discovered tafamidis at the Scripps Research Institute; is a shareholder of FoldRx (Pfizer), the companies that developed tafamidis into a drug; and he and ETP receive royalty payments from tafamidis sales. JWK is a paid consultant for the Pfizer Orphan and Rare Diseases organization. Pfizer supported JWK's expenses with travel and accommodation for scientific meetings. TC is a past and current investigator in clinical trials sponsored by FoldRx, Pfizer, Alnylam, Ionis, and Prothena; these trials were paid per protocol to Centro Hospitalar do Porto. TC is a consultant for Pfizer, Alnylam, Ionis, and Prothena. TC has spoken on behalf of Pfizer, Alnylam, Glaxo, and Prothena at scientific meetings and received financial compensation. Pfizer, Alnylam, Ionis, and Biogen supported TC's expenses with travel, accommodation, and registration for scientific meetings.

Copyright: © 2019, American Society for Clinical Investigation.

Submitted: November 30, 2018

Accepted: May 8, 2019

Published: June 20, 2019.

Reference information: *JCI Insight*. 2019;4(12):e126526. <https://doi.org/10.1172/jci.insight.126526>.

BACKGROUND. The hereditary transthyretin (TTR) amyloidoses are a group of diseases for which several disease-modifying treatments are now available. Long-term effectiveness of these therapies is not yet fully known. Moreover, the existence of alternative therapies has resulted in an urgent need to identify patient characteristics that predict response to each therapy.

METHODS. We carried out a retrospective cohort study of 210 patients with hereditary TTR amyloidosis treated with the kinetic stabilizer tafamidis (20 mg qd). These patients were followed for a period of 18–66 months, after which they were classified by an expert as responders, partial responders, or nonresponders. Correlations between baseline demographic and clinical characteristics, as well as plasma biomarkers and response to therapy, were investigated.

RESULTS. 34% of patients exhibited an almost complete arrest of disease progression (classified by an expert as responders); 36% had a partial to complete arrest in progression of some but not all disease components (partial responders); whereas the remaining 30% continued progressing despite therapy (nonresponders). We determined that disease severity, sex, and native TTR concentration at the outset of treatment were the most relevant predictors of response to tafamidis. Plasma tafamidis concentration after 12 months of therapy was also a predictor of response for male patients. Using these variables, we built a model to predict responsiveness to tafamidis.

CONCLUSION. Our study indicates long-term effectiveness for tafamidis, a kinetic stabilizer approved for the treatment of hereditary TTR amyloidosis. Moreover, we created a predictive model that can be potentially used in the clinical setting to inform patients and clinicians in their therapeutic decisions.

Introduction

Hereditary or familial transthyretin (TTR) amyloidoses (hATTRs) are a group of autosomal dominant degenerative diseases associated with one of the established 124 amyloidogenic mutations in the TTR gene (1, 2). The most common TTR mutation involves substitution for valine at position 30 by methionine (TTRVal30Met) (3), which is linked with familial amyloid polyneuropathy (FAP), a form of hATTR that presents with a predominantly neurological phenotype characterized by a progressive sensory, motor, and autonomic axonal neuropathy, with variable involvement of other organs (4, 5). If untreated, this disease is physically incapacitating in ≤ 5 years and invariably fatal in less than 2 decades (6, 7). This motivated the development of a fit-for-purpose small molecule kinetic stabilizer (tafamidis) (8) and one nonsteroidal antiinflammatory drug repurposed as a TTR kinetic stabilizer (diflunisal) (9), followed more recently by two approved oligonucleotide-based TTR mRNA-lowering drugs (10, 11). The availability of distinct therapeutic strategies for FAP raises the question of which drug to prioritize. Outcome predictors to specific therapies — i.e., baseline (or before treatment) demographic and clinical characteristics, as well as plasma biomarkers that help to determine which patients are more likely to respond to a certain drug — are essential in the emerging era of personalized precision medicine.

Tafamidis kinetically stabilizes the TTR tetramer, slowing its dissociation into monomers, which after misfolding enables aggregation that appears to be responsible for postmitotic cell dysfunction and death (12). Several observational open-label studies have been published after the registration trial showing that use of tafamidis halts or slows the progression of FAP, whereas others suggest that tafamidis is less effective in later-stage patients (7, 13–17). The North of Portugal has the highest number of patients worldwide exhibiting TTRVal30Met FAP (18). The FAP Reference Center in Porto (Unidade Corino de Andrade [UCA]) has ~300 patients treated with tafamidis (20 mg qd). Moreover, since the disease was first described in 1952 in Porto, UCA has followed more than 800 families with this mutation, including the longitudinal follow-up of presymptomatic mutation carriers.

Herein, we report a longitudinal follow-up study of 210 FAP patients treated with tafamidis (20 mg qd) for a maximum period of 66 months, with 2 main aims. First, we intended to scrutinize the long-term effectiveness of tafamidis in a large group of patients. Second, we sought baseline demographic and clinical characteristics, as well as plasma biomarkers that can be used as tafamidis outcome predictors. These factors were used to build a predictive model that is also described here.

Results

Study population. From July 2012 to January 2016, 306 patients were considered eligible to start tafamidis therapy at UCA (Figure 1A). Of those, 96 were excluded from this cohort study based on preestablished exclusion criteria (Supplemental Table 1; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.126526DS1>). All patients had at least 18 months of follow-up, and about half were followed for at least 48 months (Supplemental Figure 1).

Genetic, demographic, and clinical baseline characteristics for all patients included in this study are shown in Table 1. The majority of patients were heterozygous for the common TTRVal30Met mutation. The median age at disease onset was 34 years, reflecting a higher proportion of early-onset cases, as described for the Portuguese FAP patient population (18). Median disease duration at baseline was 2 years, also reflecting the situation within a clinical center that follows a large number of asymptomatic mutation carriers.

All patients presented with a predominantly neuropathic phenotype, characterized by length-dependent axonal sensory-motor and autonomic neuropathy. No asymptomatic mutation carriers were included in this study, as they are not eligible for therapy with tafamidis according to the current country guidelines. Outcome measures developed and validated for other neuropathies have been adapted and used in several clinical trials for FAP, including the neuropathy impairment score (NIS, an outcome measure based on objective evaluation of lower- and upper-limb sensory-motor function and reflexes by a trained neurologist; refs. 8, 19, 20); and the Norfolk Quality of Life Questionnaire–Diabetic Neuropathy (Norfolk QOL-DN [ref. 21]; a patient-reported outcome measure, sensitive to different features of this neuropathy, including autonomic involvement). More severely affected patients have higher NIS and/or Norfolk QOL-DN values. Additionally, we characterized sensory and motor neurological involvement by measuring sensory nerve action potentials (SNAPs) and compound motor action potentials (CMAPs) on routine nerve conduction tests. Our resulting neurophysiological score (S_{NF}) is defined as follows:

$$S_{NF} = SNAP_{ulnar\ nerve} + SNAP_{sural\ nerve} + CMAP_{ulnar\ nerve} + CMAP_{tibial\ nerve} + CMAP_{peroneal\ nerve}$$

(Equation 1)

More severely affected patients have low SNAPs and/or CMAPs, resulting in a lower S_{NF} . As expected based on the short median disease duration, most patients included in this study had mild to moderate disease, represented by a median baseline NIS of 8 points (Table 1).

Response classification. As expected in a complex disease such as FAP, no single available outcome measure reflects disease progression fully (20). For example, symptoms such as severe diarrhea, urinary retention, and erectile dysfunction, resulting from autonomic nervous system involvement, are frequently seen in these patients (4, 22). Changes in these disease manifestations are not captured by NIS or S_{NF} . Additionally, the impact of these symptoms in each patient's quality of life can be highly variable. For this reason, we based our response classification on an expert opinion. At the end of the observation period, a clinical expert in TTR-FAP on our research team individually reviewed each patient record, including neurology, cardiology and nephrology visits, as well as NIS, neurophysiology data, quality-of-life evaluations, and weight measurements.

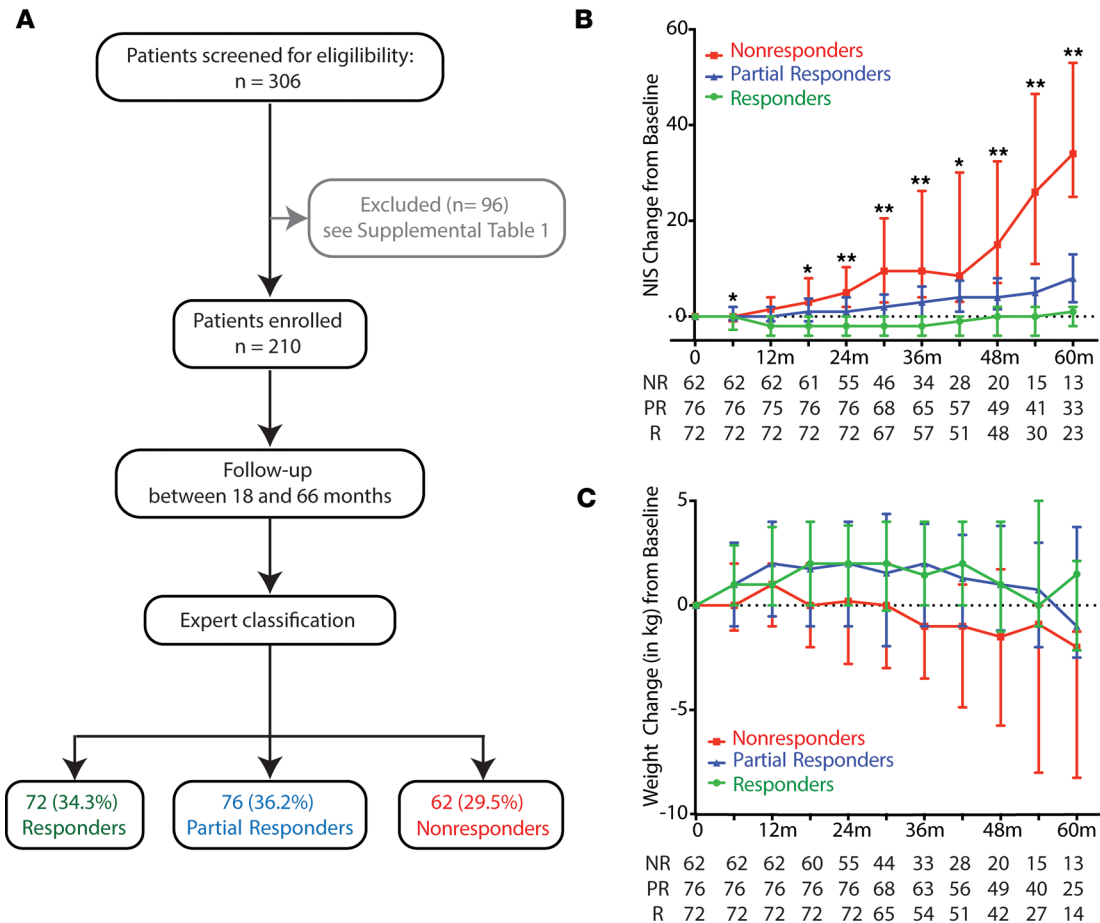


Figure 1. Patients can be classified in 3 groups according to expert opinion and selected outcome measures. (A) Study flowchart. **(B)** NIS change from baseline and **(C)** change in weight (in kg) from baseline according to expert opinion response classification in 3 groups. x axis in **B** and **C** represent number of follow-up months (m); numbers below **B** and **C** represent number of patients evaluated at each time point (NR, nonresponders; PR, partial responders; R, responders). Data are shown as median, with error bars representing interquartile range; *P* values were calculated using Kruskal-Wallis test with Dunn’s multiple-comparisons correction. **P* < 0.05, ***P* < 0.01; *P* values are only shown when differences exist between the 3 groups using multiple-comparisons correction.

Based on this complete evaluation, patients were classified as responders, partial responders, and nonresponders (Figure 1A). Seventy-two patients (34%) were considered to have no discernible disease progression and were therefore classified as responders. Change in NIS, calculated as the difference between NIS at each time point and NIS at baseline, showed that responders’ disease was mostly stable or improved (i.e., NIS change from baseline ≤ 0 for almost the entire study period; Figure 1B, green line). Continuous worsening of sensory, motor, and/or autonomic neuropathy, not different from the expected progression without therapy, was seen in 62 patients (30%), who were classified as nonresponders. This group showed an NIS increase from baseline of 3 points or more, after 18 months of treatment with tafamidis (Figure 1B, red line). An additional group of 76 patients (36%) were classified as partial responders, including patients whose autonomic nervous system manifestations improved dramatically while their sensory and/or motor neuropathy still progressed, or patients who continued progressing overall, albeit more slowly than nonresponders. Accordingly, this group showed an NIS increase from baseline of 1 point or more, also after 18 months of treatment with tafamidis (Figure 1B, blue line). Overall, when the rate of NIS change was analyzed, responders remained stable (0 points/yr), while partial responders increased 1.8 points/yr, and nonresponders increased 5.9 points/yr (Supplemental Figure 2). As explained above, the overlap of the rates of change in NIS seen among the 3 groups reflects the existence of other disease manifestations that are not measured by NIS, but which were considered in our expert classification. For example, responders and partial responders can be separated from nonresponders, as the first 2 groups gained weight or remained stable significantly more than nonresponders for the first 4 years of follow-up (Figure 1C). The same results were observed with modified BMI (Supplemental Figure 3).

Table 1. Genetic, demographic, and clinical characteristics of the study population at baseline

	Study population (n = 210)
Genotype	Val30Met/WT: 206 (98.1%), Val30Met/Val30Met: 2 (1.0%), Val28Met/WT: 1 (0.5%), Val30Met/Thr119Met: 1 (0.5%)
Sex F:M (%M)	93:117 (55.7%)
Age of disease onset (yr)	34.1 (29.0–43.1; <i>19.6–81.0</i>)
Age at baseline (yr)	36.5 (32.0–45.7; <i>22.5–84.3</i>)
Disease duration at baseline (yr)	2.1 (1.3–3.6; <i>0.3–16.5</i>)
NIS	8 (4–16; <i>0–116</i>)
Norfolk QOL-DN (n = 201)	22 (10–38; <i>0–108</i>)
S_{NF} (n = 207)	57.9 (32.8–79.0; <i>0.1–139.7</i>)

F, female; M, male. Median values are shown for all continuous variables; interval between the first quartile (Q1) and the third quartile (Q3) is shown in parentheses; minimum and maximum are also included between parenthesis in italic. Nine patients did not complete the Norfolk QOL-DN questionnaire at baseline and 3 patients did not perform baseline neurophysiological measurements. NIS, neuropathy impairment score; Norfolk QOL-DN, Norfolk Quality of Life Questionnaire–Diabetic Neuropathy; S_{NF} , neurophysiological score.

Norfolk QOL-DN data were collected at baseline and yearly after the start of tafamidis treatment. Interestingly, no differences in changes to the Norfolk QOL-DN data existed among patients who were classified as responders, partial responders, and nonresponders (Supplemental Figure 4).

In summary, our data suggest that in the long-term commercial setting, tafamidis fully stabilizes neuropathy and weight loss in about one-third of FAP patients. Additionally, another third exhibited slower overall disease progression, or continued progressing, while having clear improvements from autonomic neuropathy and weight loss. Finally, these data also suggest that despite therapy with tafamidis, about a third of patients continue progressing as rapidly as untreated FAP patients.

Clinical and demographic baseline outcome predictors. We next sought to investigate baseline characteristics that could help predict response to tafamidis. In a univariate analysis, sex was a strong predictor of response, with women more likely to become responders than men (Table 2).

Disease severity was the next sex-independent predictor variable. Patients with mild neurological impairment at baseline, as reflected by a median NIS of 6 and S_{NF} of 70.5, were significantly more likely to become responders. In addition, moderate disease at baseline was significantly linked with partial response, while more advanced patients were less likely to respond well to tafamidis (Table 2). Baseline patient quality-of-life self-assessment measured by Norfolk QOL-DN was also able to distinguish responders from nonresponders, with patients reporting worse quality of life more likely to become nonresponders (Table 2).

Age at baseline, age at disease onset, and disease duration were not predictors of response to tafamidis (Supplemental Table 2). Routine laboratory tests — including renal, liver, and thyroid function, lipid profile, complete blood count, general inflammatory markers, uric acid, and iron — did not show significant differences among the 3 groups at baseline (Supplemental Table 2).

Baseline plasma biomarkers predicting outcome. Plasma TTR levels were previously reported to be lower in patients with TTR amyloidoses (23). Moreover, low TTR levels have been associated with a decrease in survival in patients with WT-TTR cardiomyopathy (24). In order to determine whether there were significant differences among the 3 groups in tetrameric TTR levels at baseline, we measured TTR concentration (C_{TTR}) employing a fluorogenic small molecule (A2) that binds and reacts with natively folded TTR, rendering the conjugate fluorescent, which when combined with ultra-performance liquid chromatography (UPLC) separation enables quantification of native TTR tetramer (25, 26). We found that patients who become responders had higher levels of tetrameric TTR (3.0 μ M) than partial responders (2.7 μ M) and nonresponders (2.3 μ M) at baseline (Table 2 and Figure 2A). This difference was not explained by sex differences in plasma TTR levels (Figure 2B). Likewise, C_{TTR} showed weak or no correlation with disease severity, as measured by S_{NF} (Figure 2C) and NIS (Figure 2D). This suggests that tetrameric TTR levels can be used as an independent predictor of response to therapy, i.e., patients who have higher levels of tetrameric TTR are more likely to become responders.

Troponin T, a measure of cardiac muscle damage, was significantly lower at baseline in the responders compared with the other 2 groups (Table 2). This difference was independent of sex ($P = 0.136$, Mann-Whit-

Table 2. Baseline characteristics by response classification group

	Responders (R)	Partial responders (PR)	Nonresponders (NR)	R vs. NR	R vs. PR	PR vs. NR
Demographic characteristics						
Sex (%F)	68.1%	35.5%	27.4%	<0.0001	<0.0001	NS
Disease severity						
NIS	6 (4–10)	10 (6–5)	14 (8–31)	<0.0001	0.004	0.014
S_{NF}	70.5 (56.2–96.1)	58.0 (35.7–77.3)	30.5 (16.5–55.5)	<0.0001	0.0002	0.003
Norfolk QOL-DN	16 (9–29)	20.5 (11–36)	28.5 (14–52)	0.002	NS	NS
Plasma biomarkers						
Troponin T (ng/mL)	0.007 (0.004–0.011)	0.013 (0.007–0.019)	0.013 (0.010–0.028)	0.015	0.033	NS
C_{TTR} (μM)	3.0 (2.3–3.8)	2.7 (2.1–3.3)	2.3 (1.5–3.1)	0.002	NS	NS

Median values are shown for all continuous variables; interval between Q1 and Q3 is shown in parentheses. P values for continuous variables were calculated using a Kruskal-Wallis test with Dunn's multiple-comparisons correction; P values for sex were calculated using a χ^2 test. Troponin T values were only available from 51 patients (21 responders, 18 partial responders, 12 nonresponders); C_{TTR} is the concentration of TTR in plasma, also shown in Figure 2A. NIS, neuropathy impairment score; S_{NF} , neurophysiological score; Norfolk QOL-DN, Norfolk Quality of Life Questionnaire–Diabetic Neuropathy.

ney U test); however, there was a correlation between troponin T values and disease severity as measured by S_{NF} (Spearman's correlation coefficient = -0.59 , $P < 0.0001$) and NIS (Spearman's correlation coefficient = 0.49 , $P = 0.0003$). Although the number of patients for whom troponin T values were available was low ($n = 51$), this suggests that having less cardiac involvement might be also predictive of a better response to tafamidis, albeit not independent from the other measures of clinical severity.

Tafamidis and extent of TTR stabilization in plasma correlates with clinical outcome. Next, we hypothesized that the tafamidis plasma concentration, which correlates with the extent of TTR kinetic stabilization, could explain why some patients respond to tafamidis better than others (26). TTR tetramer dissociation into monomers is the rate-limiting step of the TTR amyloidogenesis cascade. The process of TTR amyloidogenesis is thought to be the cause of tissue degeneration and disease progression (27). Tafamidis acts by kinetically stabilizing the TTR tetramer, slowing down tetramer dissociation into monomers, and consequently slowing aggregation and disease progression (12) (Figure 3A). We developed biochemical methods to measure total tafamidis plasma concentration (C_{Taf}) (26), as well as a TTR subunit exchange method to quantify the rate of tetramer dissociation to monomers in plasma (25). When analyzing C_{Taf} by response classification, we found that nonresponders had significantly lower tafamidis levels (Figure 3B; $6.8 \mu\text{M}$) than partial responders ($8.4 \mu\text{M}$) and responders ($8.7 \mu\text{M}$). The same difference was seen at 24 months of therapy, and the same tendency was found in the samples for which a last time point (more than 24 months) was analyzed (Supplemental Figure 5, A and B). These patients were all taking tafamidis meglumine 20 mg once daily (corresponding to 12.5 mg tafamidis). Blood samples were collected at the time of the visit, and information regarding the last tafamidis dose was not collected. However, tafamidis levels at 12 months correlated well with tafamidis levels at 24 months within the same individual, and the overall differences between 12 and 24 months were not significant, suggesting that temporal variations in drug levels were minor (Supplemental Figure 6, A and B). This was not unexpected based on the reported long half-life of tafamidis (~ 50 hours). Interestingly, when analyzing men and women separately, we found that the differences in C_{Taf} between different response groups were driven entirely by male patients (Figure 3C, $C_{Taf/NR-women} = 8.4 \mu\text{M}$, $C_{Taf/PR-women} = 9.2 \mu\text{M}$, $C_{Taf/R-women} = 8.4 \mu\text{M}$; and Figure 3D, $C_{Taf/NR-men} = 6.3 \mu\text{M}$, $C_{Taf/PR-men} = 7.6 \mu\text{M}$, $C_{Taf/R-men} = 9.7 \mu\text{M}$).

C_{Taf} should correlate with the extent of TTR tetramer kinetic stabilization (defined as the difference between the rate of tetramer dissociation at 12 months or 24 months and the rate of tetramer dissociation at baseline), and it did (Figure 3E and Supplemental Figure 7A), reflecting target engagement. Consistent with the differences in C_{Taf} reported above, the extent of stabilization was also different according to response classification in male patients (Figure 3F).

According to these observations, C_{Taf} and consequently, the extent of TTR tetramer stabilization, is a predictor of response only in male patients. This suggests that additional mechanisms for nonresponse may play an important role in women.

Factors that influence C_{Taf} . Interestingly, our data show that C_{Taf} has a wide range of concentrations, i.e., from $\sim 20 \mu\text{M}$ to $\sim 0 \mu\text{M}$ (Figure 3B). We next sought to investigate factors that could help explain this

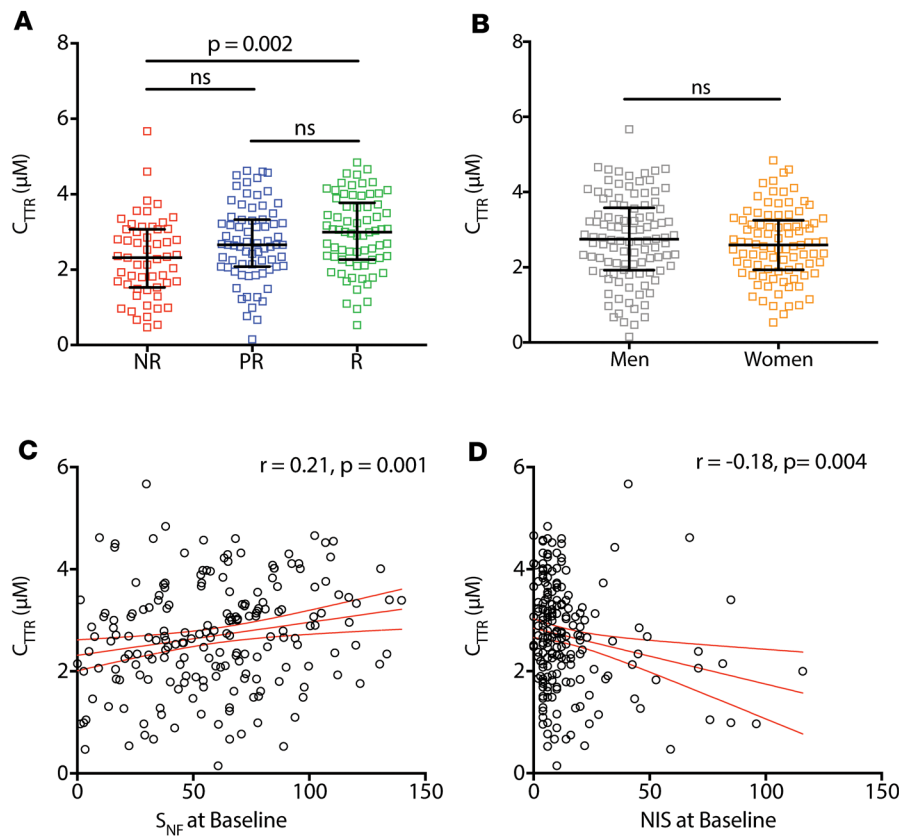


Figure 2. Concentration of tetrameric plasma TTR at baseline is significantly different between responders and nonresponders. (A) Concentration of tetrameric plasma TTR (C_{TTR}) at baseline is significantly higher in responders than in nonresponders; P values were calculated using Kruskal-Wallis test with Dunn's multiple-comparisons correction; $n_{NR} = 55$, $n_{PR} = 72$, $n_R = 70$. (B) No differences were seen between men and women in C_{TTR} at baseline. P values were calculated with Mann-Whitney U test; $n_{men} = 106$, $n_{women} = 91$; for A and B, horizontal bars represent median and error bars represent interquartile range. (C) Weak correlation or no correlation is seen between C_{TTR} and S_{NF} ($n = 194$), and (D) C_{TTR} and NIS at baseline ($n = 197$); for C and D, r represents Spearman's correlation coefficient. Red line represents best fit; curved red lines represent 95% confidence intervals of the best fit.

variation in C_{Taf} . In general, drug levels are influenced by sex, age, drug absorption and metabolism, body weight and size, and renal and/or hepatic impairment (28). Women had overall higher levels of tafamidis in plasma at 12 months than men (8.6 μM versus 7.6 μM , Figure 4A), but this difference was not statistically significant. Moreover, at 24 months, both men and women had the same median C_{Taf} in plasma (7.6 μM ; Figure 4B). When adjusting for BMI, there were no significant differences in C_{Taf} between men and women at 12 months ($C_{Taf,men} = 0.34 \mu\text{M}/[\text{kg}/\text{m}^2]$, $C_{Taf,women} = 0.35 \mu\text{M}/[\text{kg}/\text{m}^2]$) and 24 months ($C_{Taf,men} = 0.36 \mu\text{M}/[\text{kg}/\text{m}^2]$, $C_{Taf,women} = 0.33 \mu\text{M}/[\text{kg}/\text{m}^2]$).

Differences in metabolism could also explain some of the C_{Taf} variation. Approximately 20% of tafamidis is reported to be metabolized by glucuronidation to acylglucuronide (herein referred to as tafamidis-gluc.). If, in fact, patients with low C_{Taf} are faster metabolizers, one would expect to find higher tafamidis-gluc. levels in the plasma of patients with low unmetabolized tafamidis. We were able to detect the peak corresponding to tafamidis-gluc. in plasma of patients taking tafamidis by comparison to synthesized tafamidis-gluc. (Supplemental Figure 8, A–D). By comparing the area of the metabolite peak (elution volume 9 mL) with the area of the unmetabolized drug (elution volume 13 mL), we found a strong positive correlation between the two, suggesting that increased glucuronidation is unlikely to explain low tafamidis levels (Figure 4C). In the subset of patients analyzed, tafamidis-gluc. represented 1%–15% of the total C_{Taf} .

Finally, we did not find significant correlation between tafamidis plasma levels and age, weight, BMI and modified BMI (mBMI), estimated glomerular filtration rate (eGFR; measure of renal function) and aspartate aminotransferase (AST; measure of hepatic function) at 12 or 24 months (Table 3). Only alanine aminotransferase (ALT; measure of hepatic function) at 24 months had a weak correlation with tafamidis levels ($r < -0.20$, Table 3).

In summary, we were able to exclude common factors that play a role in drug pharmacokinetics as being responsible for the lower tafamidis concentrations seen preferentially in nonresponders; however, it is worth noting that our analysis did not fully exclude the hypothesis of faster tafamidis excretion or poor absorption in patients with low tafamidis levels.

A predictive model of response to tafamidis. The observations in the preceding sections show that several variables that can be measured before or soon after tafamidis therapy is initiated vary significantly among

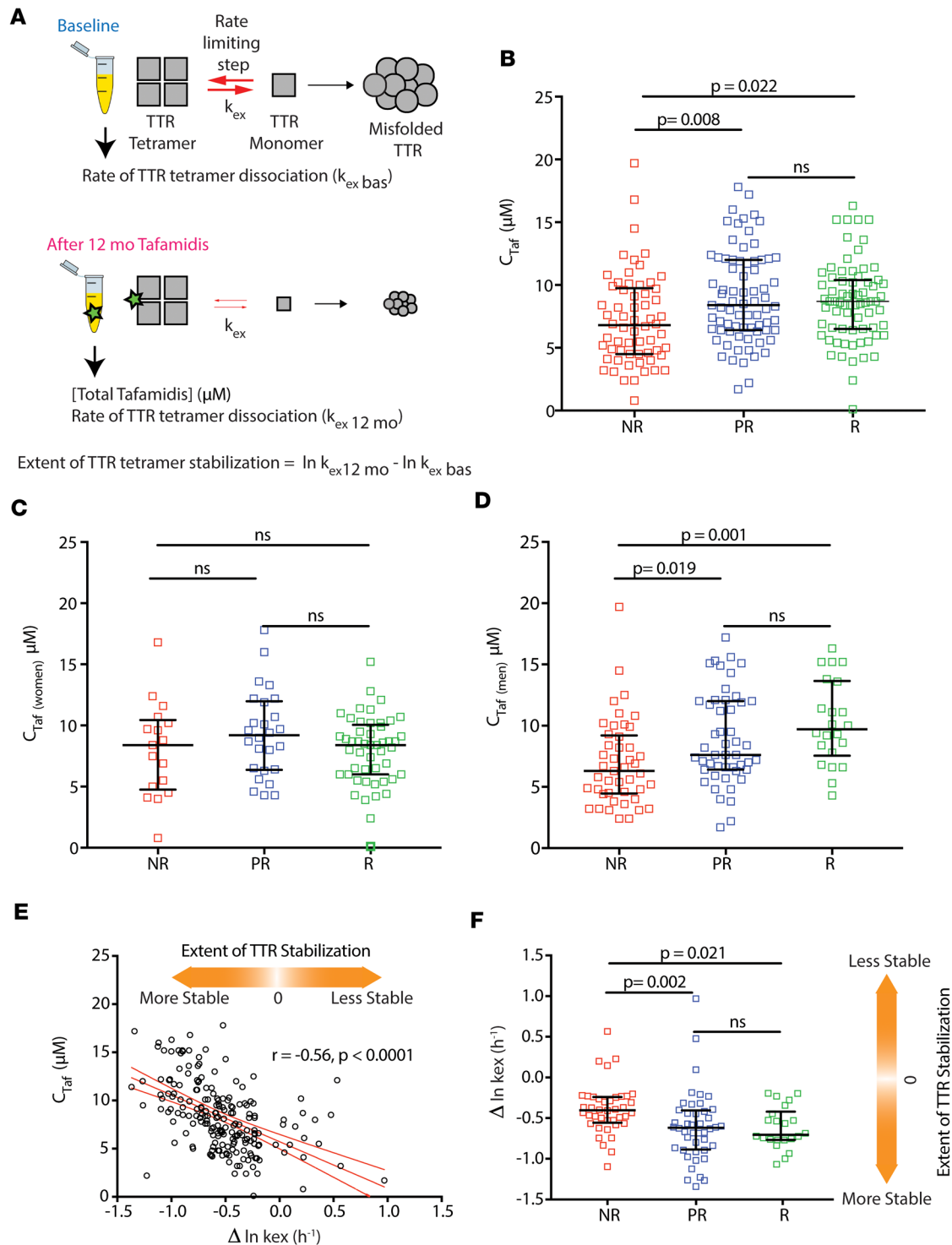


Figure 3. Concentration of tafamidis in plasma and extent of tetramer stabilization are different according to response classification. (A) Diagram representing the TTR amyloidogenesis cascade hypothesis and the experimental approach for this experiment. Using samples collected at baseline (bas) and after 12 months of therapy, the rate of dissociation of the TTR tetramer was determined at baseline and after 12 months of tafamidis therapy; the extent ($\Delta \ln k_{ex}$) of TTR tetramer stabilization was calculated as shown in the figure; additionally, total concentration of tafamidis (C_{Taf}) in plasma was measured in the 12-months samples. (B) Total C_{Taf} in plasma is significantly lower in nonresponders ($n = 62$) when compared with partial responders ($n = 75$) and responders ($n = 71$). (C) Total C_{Taf} in women is not significantly different according to response classification (NR: $n = 17$, PR: $n = 26$, R: $n = 49$). (D) C_{Taf} in men is significantly lower in nonresponders (NR: $n = 45$, PR: $n = 49$, R: $n = 22$). (E) Correlation between plasma C_{Taf} and the extent of TTR stabilization ($n = 194$), showing that a higher C_{Taf} is positively correlated with a more stable TTR tetramer. r , Spearman's correlation coefficient; red line represents best fit ($R^2 = 0.26$); curved red lines represent 95% confidence intervals of the best fit. (F) TTR in plasma of male nonresponders ($n = 40$) is significantly less stable when compared with male partial responders ($n = 44$) and male responders ($n = 22$); for B–D and F, horizontal bars represent median and errors bar represent interquartile range. P values were calculated using Kruskal-Wallis test with Dunn's multiple-comparisons correction.

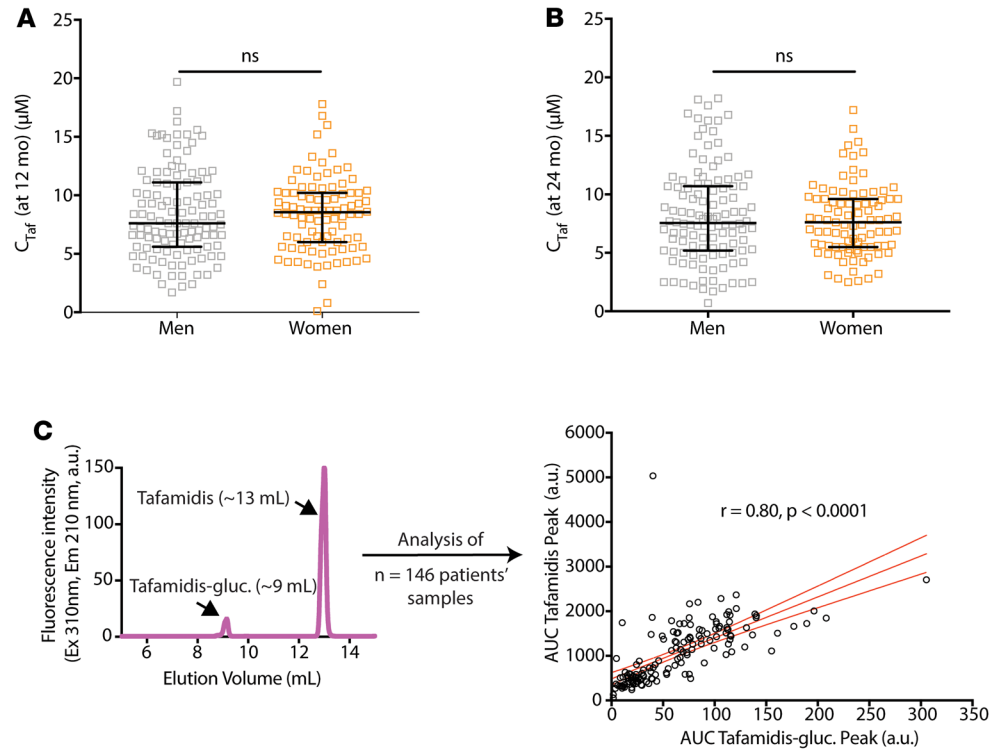


Figure 4. Variation in C_{Taf} is not explained by sex or by tafamidis glucuronidation. (A) C_{Taf} is not different between women and men at 12 months ($P = 0.70$; $n_{men} = 116$, $n_{women} = 92$) and (B) 24 months of therapy ($P = 0.77$; $n_{men} = 116$, $n_{women} = 93$); for A and B, horizontal bars represent median and errors bar represent interquartile range. P values were calculated using a Mann-Whitney U test. (C) Left panel: Chromatogram representing fluorescence detection of tafamidis-gluc. and unmetabolized tafamidis in the plasma of a patient taking oral tafamidis (for details, see Supplemental Methods and Supplemental Figure 8). Right panel: Correlation between the AUC for the tafamidis peak and the metabolite peak (tafamidis-gluc.). r , Spearman's correlation coefficient; red line represents best fit ($R^2 = 0.44$); curved red lines represent 95% confidence intervals of the best fit.

nonresponders, partial responders, and responders. Therefore, these variables could in principle be used to predict a patient's responsiveness to tafamidis. To explore this possibility, we converted the categorical outcome variable (i.e., the response classification), which has the values "nonresponder", "partial responder", and "responder", to a numerical variable with corresponding values of 0, 0.5, and 1, respectively. We then fit the linear model shown below to the numerized responses for the 210 patients in our database:

$$S_R = a_0 + b_{NF} \times S_{NF} + b_M \times M + b_{M \times Taf} \times M \times C_{Taf} + b_{TTR} \times C_{TTR}, \quad (\text{Equation 2})$$

where S_R is the numerized response score; a_0 is the constant of regression; S_{NF} is the neurophysiology score (a sum of several SNAPs and CMAPs); M is the sex of the patient, coded as 1 for males and 0 for females; and C_{Taf} and C_{TTR} are the micromolar plasma concentrations of tafamidis and tetrameric TTR, respectively.

Note that the fourth term has a product of 2 variables, M and C_{Taf} . Given the coding for M , $M \times C_{Taf} = 0$ for females and $M \times C_{Taf} = C_{Taf}$ for males. This interaction variable therefore accounts for our observation that plasma C_{Taf} differed significantly among the response classifications only for male patients.

The quantities denoted by b with subscripts are the regression coefficients for the corresponding variables. The best-fit values for the parameters are as follows: $a_0 = 0.195 \pm 0.076$, $b_{NF} = 0.0046 \pm 0.0007$, $b_M = -0.447 \pm 0.083$, $b_{M \times Taf} = 0.027 \pm 0.008 \mu\text{M}^{-1}$, and $b_{TTR} = 0.068 \pm 0.022 \mu\text{M}^{-1}$. The model captures a moderate amount — about 35% ($R^2 = 0.35$) — of the variation in the responses of individual patients to tafamidis therapy. However, given the discrete nature of the response classification variable, this is not the best way to judge the utility of the model. Instead, we divided the model outputs into quintiles and calculated the proportions of nonresponders, partial responders, and responders in each quintile (Figure 5A).

Table 3. Relation between plasma C_{Taf} and patient characteristics that could influence tafamidis pharmacokinetics

	C_{Taf} (at 12 mo)		C_{Taf} (at 24 mo)	
	<i>r</i>	<i>P</i> value	<i>r</i>	<i>P</i> value
Age	-0.0485	0.4868	-0.0721	0.2993
Weight	-0.0129	0.8528	-0.0878	0.2127
BMI	-0.0013	0.9846	-0.0818	0.2460
mBMI	0.1067	0.1260	0.0208	0.7687
eGFR	-0.0482	0.4903	-0.0065	0.9261
AST	-0.0403	0.5634	-0.0606	0.3847
ALT	-0.1029	0.1390	-0.1690	0.0147

BMI is shown in kg/m²; mBMI (modified BMI) in kg/m² × g/L of serum albumin; and eGFR in mL/min/1.73 m². AST, aspartate transaminase; ALT, alanine transaminase; eGFR, estimated glomerular filtration rate. Only ALT at 24 months is not independent from C_{Taf} ($P < 0.05$ highlighted in bold); however, the Spearman's correlation coefficient (*r*) for ALT suggests negligible correlation.

The proportions of each response category in a given quintile can then be taken as estimates of the probabilities that a new patient who also falls into that quintile (based on their calculated S_R , obtained from evaluating Equation 2 with that patient's sex, S_{NF} , C_{Taf} , and C_{TTR} values) will have a given response to tafamidis. For example, a hypothetical male patient ($M = 1$) with $S_{NF} = 32.9$, $C_{Taf} = 14.9$ μM, and $C_{TTR} = 3.6$ μM has a calculated response score value of $S_R = 0.195 + (0.0046 \times 32.9) + (-0.447 \times 1) + 0.027 \times 1 \times 14.9 + 0.071 \pm 0.068 \times 3.6 = 0.55$. An S_R of 0.55 would place this hypothetical patient in the third quintile (Figure 5A), where their probability of being a nonresponder, partial responder, or responder is $19\% \pm 6\%$, $43\% \pm 8\%$, or $38\% \pm 7\%$, respectively (uncertainties represent the standard errors for proportions).

The robustness of this model is difficult to assess using traditional methodology employed to evaluate binary classifier systems (for example, receiver operating characteristic [ROC] curves) because it has 3, rather than 2, classifications. Therefore, to give a sense of the ability of our model to discriminate between the response categories, we show 2 ROC curves for our model (Figure 5B): one showing discrimination between the combined nonresponder and partial responder categories versus the responder category (NR+PR vs. R; purple curve) and the other discrimination between the nonresponder category versus the combined partial responder and responder categories (NR vs. PR+R; cyan curve). The former curve emphasizes identification of responders, whereas the latter emphasizes identification of nonresponders. The AUCs of the 2 ROC curves were 0.82 and 0.81, respectively, indicating that our model is effective at discriminating both responders and nonresponders.

The relative importance of the variables in our model can be assessed by leaving each one out of the model in turn and assessing the performance of the abbreviated models. This exercise revealed that S_{NF} was the most important variable in our model, since leaving it out diminished the AUCs of the ROC curves to 0.75 for the NR+PR vs. R curve and 0.72 for the NR vs. PR+R curve. The importance of the other variables decreased in the following order: sex (M) > $M \times C_{Taf}$ > C_{TTR} . The best fit parameters and AUCs for the abbreviated models are presented in Supplemental Table 3.

NIS was left out of our model because S_{NF} and NIS provide somewhat redundant information about the extent of disease progression in terms of nerve function, but S_{NF} better discriminated responders, partial responders, and nonresponders (Table 2). In fact, S_{NF} decreased exponentially as NIS increased, so that S_{NF} decreased precipitously over a narrow range of small NIS values (Supplemental Figure 9). Thus, S_{NF} seems to be more sensitive than NIS to early disease progression in FAP. However, NIS is more widely available and a fully validated clinical score. For this reason, we built an alternative model (Model-NIS), in which S_{NF} was replaced by NIS (Supplemental Figure 10). Finally, since the biochemical methods to measure C_{Taf} and C_{TTR} might not be available in all clinical centers following these patients, we built a simplified version of our model (Model-S), in which plasma measurements of tafamidis and TTR concentrations were left out (Supplemental Figure 11). In both cases (Model-NIS and Model-S), discrimination between the 3 response groups was still possible (Supplemental Figures 10 and 11), although not as clear as with the original model (Figure 5A).

The usefulness of our model for estimating an individual patient's chances of being a nonresponder, partial responder, or responder to tafamidis therapy (20 mg qd) can only truly be tested by prospectively applying it to a cohort of patients as they begin treatment and then comparing the proportions of each response category in each quintile to the expected proportions from our model. However, in the interim,

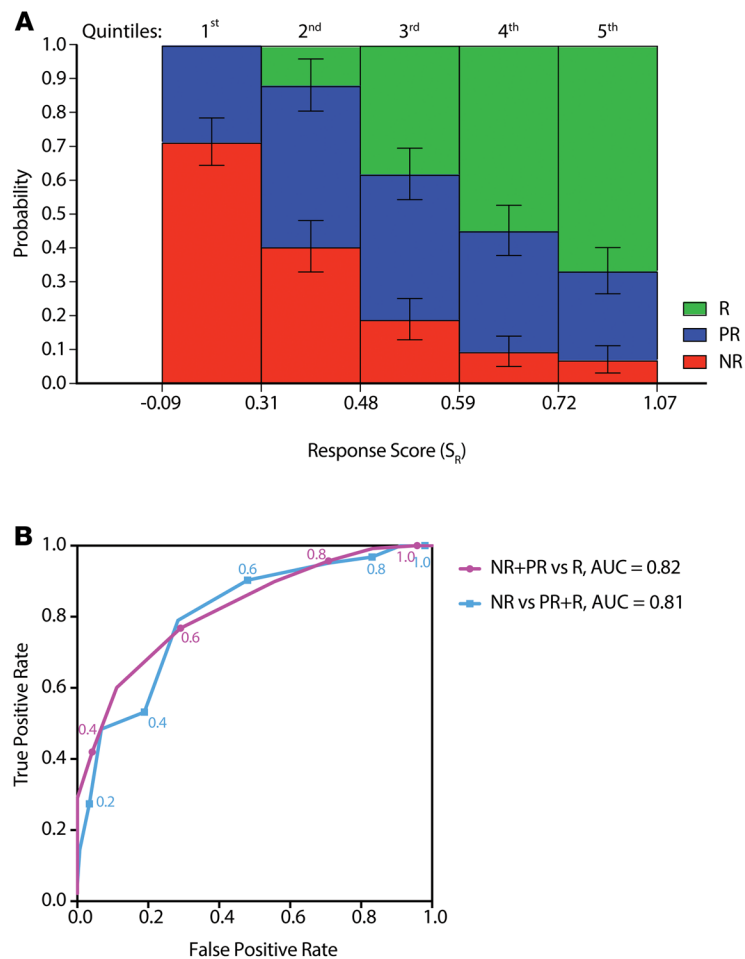


Figure 5. Predictive model of response to tafamidis. (A) The response score (S_R) interval of each quintile is shown in the x axis, and the probability for each response classification within each quintile is shown in the y axis. Error bars represent 95% confidence limits. (B) Receiver operating characteristic (ROC) curves for the predictive model. Purple curve: Discrimination between the combined nonresponder and partial responder categories versus the responder category (NR+PR vs. R). True positives in this case are patients who are nonresponders or partial responders with S_R below a given cutoff; false positives are responders with S_R below a given cutoff (S_R cutoffs are shown as purple numbers). Cyan curve: Discrimination between the nonresponder category vs. the combined partial responder and responder categories (NR vs. PR+R). True positives in this case are patients who are nonresponders with S_R below a given cutoff; false positives are partial responders or responders with S_R below a given cutoff (S_R cutoffs are shown as cyan numbers).

we were able to test the validity of our approach by randomly dividing the patient data set into “training” and “test” sets. The training set was used for fitting the model, setting the boundaries for the quintiles, and calculating the expected proportions of each response category in each quintile. The fitted model was used to calculate response values for the test set, which were then sorted into the quintiles defined by the training group, and the proportion of each response category in the quintiles from the test set was compared with the expected proportions from the training set. We repeated this procedure 100 times and found that the proportions of the responder categories in the quintiles from the training set predicted the proportions in the test set with a root mean square deviation of $\pm 14\%$. This error was consistent with the standard errors shown in Figure 5 for the proportions in the quintiles calculated from the full data set — which are typically on the order of 7%–8% — given that the training and test sets were each half the size of the full data set. These results suggest that our model can indeed be useful for predicting the likelihood that a given patient will respond to tafamidis therapy.

Discussion

Herein, we have shown that over a maximum follow-up period of 66 months, at least one-third of patients taking tafamidis respond to the drug with almost complete cessation of disease progression. Our criteria to define responders (resulting in a median NIS score change from baseline ≤ 0 for 54 months) are more stringent than the criteria used in previous studies (change in NIS in the lower limbs [NIS-LL] < 2 points in 18 months) (8, 13). We identified an additional one-third of patients with partial remission of important disease manifestations while continuing progressing from other disease aspects (defined as partial responders). The lack of an untreated control group was a limitation of our study that could

not be avoided, given that tafamidis is the standard of care for patients in Portugal. However, one can compare our results with previously reported natural history studies or placebo-controlled clinical trials for FAP. It is clear that patients in the responder and partial responder groups progressed slower (median NIS progression rate: 0 points/yr for responders and 1.8 points/yr for partial responders) than what was described in the pivotal tafamidis study (8, 13) for the placebo-treated cohort (NIS-LL progression rate of approximately 5 points in 12 months), which includes a patient population comparable to ours. Additionally, a combined analysis of 252 matched untreated FAP patients had an NIS-LL progression rate of 2.8 points/yr (17). A multinational natural history study including more advanced patients (median baseline NIS of 32) estimated a rate of NIS progression of 14.3 points/yr (29). These data confirm that untreated patients progress at a rate of more than 2 points per year in their NIS-LL. The fact that NIS-LL is a subpart of NIS will lead to an identical or higher progression rate when considering the total score as we do in the current study. Additionally, we report a third of patients using our more stringent analysis as nonresponders with a median NIS progression rate of 5.9 points/yr, which is within the range reported for the rate of progression of untreated FAP patients in the above-mentioned placebo-controlled studies or observational and natural history studies (8, 13). Moreover, our data support the tafamidis registration trial (8), in which 60% of the efficacy-evaluable population was considered to be “NIS-LL responders.” The response classification by one single expert is a weakness in our study; however, the clinical scores and the opinions of the neurologists observing each patient were taken into consideration by the expert. A consensus of multiple experts on response classification should be the aim of future studies. It is important to note that progression of Norfolk QOL-DN scores was not a useful measure of tafamidis response in this group of patients. Patients who were unequivocally progressing from their motor and sensory neuropathy showed minimal changes in this patient-reported outcome measure, possibly reflecting patients’ expectations of a treatment effect and the open-label nature of this study.

One limitation of our 3-group approach is that, while clinically rigorous and supported by different scores, it might have led us to lose power in our prediction model. Despite this limitation, we believe that this classification-by-response model can be used to rigorously compare mechanistically distinct therapies in the commercial setting, where observation times are longer and more meaningful. Additionally, the 3-group approach can be used to evaluate patients with other FAP genotypes treated with tafamidis at the same or higher doses. Ideally, data from multiple clinical centers should be combined in order to acquire information on larger patient data sets.

An important aim of this study was to identify baseline demographic and clinical characteristics and identify plasma biomarkers that can be used as tafamidis outcome (or response-to-therapy) predictors. Our data show that female sex is a clear predictor of a positive response to tafamidis. These data raise the possibility that sex differences influence underlying disease mechanisms, a hypothesis that is also supported by the peripheral blood cell transcriptional profiling differences between men and women reported previously (30). Future studies of disease-specific and response-to-therapy biomarkers should take sex in consideration. Our data also confirm that early-stage patients are more likely to respond positively to tafamidis. This raises the hypothesis that TTR dissociation, misfolding, and aggregation initiate a downstream cascade of abnormal events that can auto-perpetuate itself even when the proteinopathy trigger to pathology is removed by tafamidis treatment. A better understanding of early and late disease events, as well as the role of neuroinflammation, is essential, as it may lead to complementary treatment approaches. Additionally, we have identified tafamidis concentration (and consequently TTR kinetic stabilization) as a predictor of response in a subset of patients, predominantly men. This finding raises the possibility that some FAP patients might benefit from a higher dose of tafamidis, and this should be further investigated.

These data also highlight the need for early diagnostic biomarkers in FAP to allow early initiation of therapy. Using a method that detects only natively folded tetrameric TTR, we found that patients who had higher levels of native TTR at baseline were more likely to become responders than nonresponders. Future studies should further investigate the differences between our A2-UPLC tetrameric TTR quantification method and the routinely used immunoturbidimetric methods to define TTR levels as a prognostic indicator of response to kinetic stabilizers. The reasons why FAP patients have low levels of TTR are not understood. Interestingly, reduced CSF levels of A β ₄₂ were established as diagnostic biomarkers after longitudinal studies following genetic Alzheimer’s disease cases (31). Analogously, longitudinal follow-up of presymptomatic FAP mutation carriers should help define the role of native TTR levels in detecting early symptomatic disease stages.

Using data from this retrospective patient study, we were able to build a predictive model of response to tafamidis (20 mg qd), which can potentially be used in the clinical setting to help clinicians and patients predict response, and eventually to prioritize therapy. To build the model, we used plasma tafamidis concentration at 12 months (C_{Taf}); however, based on the reported pharmacokinetic characteristics of tafamidis and also on our comparisons between 12 and 24 months, we consider that concentration after 1 month of therapy should be used in future studies to further validate this model. Moreover, additional response-to-therapy biomarkers that might strengthen this model are currently under investigation using unbiased plasma proteomic approaches. It is important to keep in mind that our model was built using patients with a predominantly neuropathic phenotype, and most of our patients were heterozygous for the common Val30Met mutation (98.1%). Future studies with other mutations and mixed cardiac/neuropathy phenotypes should indicate whether a similar model can be applied more broadly.

Finally, our work comprises detailed clinical and biochemical human data supporting the amyloid hypothesis applied to the TTR amyloidoses. The amyloid hypothesis posits that the misfolding and aggregation of a given protein into a spectrum of non-native structures, including cross- β -sheet amyloid fibrils, is linked to tissue degeneration (32). Overall, disease progression is slowed when the dissociation rate of plasma TTR tetramers is reduced by correspondingly higher plasma concentration of the kinetic stabilizer tafamidis. This provides a direct link between neuropathy progression and the inhibition of TTR aggregation via TTR native tetramer stabilization.

Methods

Study design

This was a nonrandomized, longitudinal observational study of all patients eligible to start therapy with tafamidis in a single clinical center (Unidade Corino de Andrade), with at least 24 months follow-up from baseline (or a minimum of 18 months for patients who did not respond during that period of time), up to 66 months. Patients were considered eligible to start therapy with tafamidis when they met the following criteria: (i) confirmed TTR genetic mutation, (ii) tissue biopsy with confirmed Congo red–positive amyloid deposition, and (iii) symptoms or signs confirming involvement of the peripheral nervous system (sensory, autonomic, and/or motor). For this study, the following exclusion criteria were considered: (i) patients included in the clinical trial Fx005 (ClinicalTrials.gov, NCT00409175) or Fx006 (NCT00791492), (ii) inclusion in the open-label add-on trial with the siRNA drug patisiran (NCT01961921), (iii) comorbidities with peripheral and/or central nervous system clinical manifestations, (iv) therapy suspension before 2 years for reasons other than nonresponse, (v) absent clinical follow-up after 12 months of therapy, (vi) no baseline or subsequent plasma samples collected, and (vii) refusal to participate in the study.

Study population

All patients eligible for this study with baseline assessment between July 2012 and January 2016 were included. Patients were assessed at treatment start (baseline visit) and subsequently every 6 months up to 5.5 years. Patients were observed by one of 6 trained neurologists working or previously working at the FAP clinic. When considered necessary, they were also observed by a cardiologist, ophthalmologist, and/or nephrologist. A standardized protocol included the outcome measures detailed below: NIS, Norfolk QOL-DN, weight, routine neurophysiologic tests including sensory and motor nerve amplitude measurements, and routine blood tests. Information regarding TTR mutation was available for all patients. Disease duration was based on judgment by the treating neurologist regarding the information provided by each patient for the onset of symptoms and signs associated with TTR-FAP.

Outcome measures

NIS, Norfolk QOL-DN, weight, BMI, and mBMI. Patients had their NIS score determined at baseline and every 6 months. For most patients, the neurologist scoring was the same at baseline and follow-up, reducing the chance for interobserver variability. Concomitantly, patients had height measurement at baseline or at an asymptomatic stage and weight measurements at all visits of the study. Serum albumin was determined as part of the routine blood testing at the hospital clinical laboratory.

S_{NF}. Routine nerve conduction studies were performed concurrently with each neurology observation. The SNAPs of ulnar and sural nerves were measured (peak to peak). Average SNAP was used when bilat-

eral measurements were available for the same nerve. Most patients had bilateral sural nerve and unilateral ulnar nerve SNAP measurements. CMAPs of the tibial, peroneal, and ulnar nerves were also determined, in most cases unilaterally. In patients for whom bilateral measurements were available, average CMAPs were calculated. The neurophysiology score corresponds to: sural SNAP + ulnar SNAP + tibial CMAP + peroneal CMAP + ulnar CMAP. No adjustments for age, sex, or height were made. In cases where sensory or motor potentials were not present in a specific nerve, repeated measurements were not attempted in the same nerve in the subsequent visit, and the corresponding SNAP or CMAP was considered 0.

Routine blood tests

Routine blood tests were performed on all patients at the hospital clinical laboratory, including renal function (creatinine, blood urea nitrogen [BUN], and uric acid), liver function (total bilirubin, aspartate and alanine transaminases, alkaline phosphatase, gamma-glutamyl transaminase, serum albumin), cardiac markers (troponin T [Roche Elecsys high-sensitivity], NT-Pro-BNP [Roche Elecsys]), total cholesterol, triglycerides, serum albumin, total protein, free thyroxine, thyroid-stimulating hormone, complete blood count, sedimentation rate, and C-reactive protein. Urine samples were assessed for albuminuria and proteinuria. eGFR was calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation (33).

Responder classification

In January 2018, i.e., 66 months after the study began, an expert in TTR-FAP (TC) classified all patients as responders, partial responders, and nonresponders. Patient clinical records were scrutinized individually. Additionally, NIS, Norfolk QOL-DN, weight, and neurophysiology score were also considered. Responders were patients who showed no progression of sensory, autonomic, and/or motor neuropathic symptoms and signs. The usual criterion of progression of less than 2 points in the NIS score was not applied (29). Patients were classified as nonresponders when they showed a continuous and rapid worsening of sensory, motor, and/or autonomic neuropathy not different from the expected progression without therapy. A group of partial responders included patients who showed slower than expected progression of sensory and/or motor neuropathy with stability of autonomic nervous system manifestations, or patients who continued progressing from sensory and/or motor neuropathy while showing a manifest improvement of the autonomic neuropathy and weight gain. Autonomic nervous system manifestations can be devastating in these diseases; therefore, we decided that improvement of objective symptoms such as diarrhea, orthostatic hypotension and erectile dysfunction, even in the presence of continuous progression of sensory neuropathy, should be considered at least as a partial response. None of the biochemical measurements regarding TTR concentration, tafamidis concentration, or TTR stabilization played any role in the classification of patients as responders, partial responders, or nonresponders.

Plasma samples

Blood was collected in tubes with EDTA, followed by centrifugation at 1500 *g* for 20 minutes. The resulting supernatant (plasma) was carefully removed and centrifuged for an additional 20 minutes to remove any remaining cells. Plasma was transferred to 1.5 mL cryovials with a cap and stored at -75°C until it was shipped. After shipment in dry ice, frozen patient plasma samples were kept at -80°C . All plasma samples were thawed at ambient temperature before being aliquoted and used for the different assays. No more than 2 cycles of freeze-thaw were applied. Most patients had blood collection at baseline and every 6 months. For most assays, only baseline and annual samples were analyzed (12, 24, 36, 48, and 60 months). In the cases where a 12-month sample was not available, a 6-months or 18-month sample was used; similarly, when a 24-month sample was not available, an 18-month sample was used. Blood from healthy volunteers was obtained from the Scripps Research Institute Normal Blood Donor Services Center.

Recombinant protein expression and purification

Recombinant WT TTR (WT-TTR) and dual-FLAG-tagged WT-TTR (FT₂-WT-TTR) were expressed and purified from *Escherichia coli* as described previously. The molar absorptivity (ϵ) of WT-TTR ($73,800\text{ M}^{-1}\text{ cm}^{-1}$) and FT₂-WT-TTR ($85,720\text{ M}^{-1}\text{ cm}^{-1}$) tetramers in standard phosphate buffer was used to prepare TTR solutions of known concentration. After purification, aliquots were stored at -80°C in 50 mM phosphate buffer pH 7.6 (standard phosphate buffer).

Measurement of tetrameric TTR in plasma

Tetrameric TTR levels in plasma were quantified using a Waters ACQUITY H-Class Bio-UPLC instrument employing a Waters Protein-Pak Hi Res Q ion exchange column (strong anion exchanger, 5- μ m particle size, 4.6 \times 100 mm column). A standard curve was prepared using recombinant WT-TTR at concentrations of 10 μ M, 5 μ M, and 2.5 μ M in 50 mM phosphate buffer, pH 7.6. Subjects' samples and 2 healthy control samples were thawed and filtered using a 0.45- μ m syringe filter. Subjects' plasma samples, healthy control plasma, and standard curve samples were aliquoted (9 μ L) in duplicate into a 96-well plate with 1 μ L of the fluorescent small molecule A2 (500 μ M) in each well (34). The same 2 healthy controls from the Normal Blood Donor Services Center were used in duplicate in each plate to determine intra-assay variability. After incubation, the samples were injected into the ion exchange column and separated using a linear 24%–29% buffer B gradient over 10 minutes (flow 0.5 mL/min, buffer A: 25 mM Tris pH 8; buffer B: same as buffer A, but with 1 M NaCl added). The TTR-(A2)₂ fluorescent conjugate peak (excitation 328 nm, emission 430 nm; elution time, 6 minutes) was integrated, and the concentration of TTR in patient samples was quantified using the standard curve. Intraplate variability was \leq 15%, and interplate variability was \leq 10%.

Subunit exchange assay to assess kinetic stability

Subunit exchange rates were determined as described previously (25) with the following minor modifications. FT₂-WT-TTR (2 μ L of a 40- μ M solution) was added to plasma (40 μ L) to afford a final FT₂-WT-TTR concentration of 2 μ M. The samples were incubated at 25°C for 48 hours to allow subunit exchange to occur. At 48 hours, the reaction was stopped by the addition of the fluorogenic small molecule A2 at a final concentration of 500 μ M. The samples were incubated with A2 for at least 3 hours to allow complete covalent labeling of TTR, before being injected into the ion exchange column and separated using a linear 24%–39% buffer B gradient over 29 minutes using the same buffers as described above. The rate of exchange for all subunit exchange experiments was calculated using peak 1, as previously described. All patient samples were analyzed in duplicate (technical replicates), and most samples corresponding to the same patient were run in the same plate. Coefficient of variability within the samples was \leq 15%. The same 2 healthy control plasmas were used in duplicate in all subunit exchange plates. The inter-plate variability was \leq 17%.

HPLC analysis of tafamidis concentration

Tafamidis levels were quantified in the plasma samples employing the HPLC method previously described (26), with the following minor modifications. Four sets of standard curves were generated simultaneously by adding tafamidis (using a 50 \times stock solution in DMSO) to a healthy donor plasma and incubating overnight at 37°C. The final C_{Taf} employed for the standard curve equaled 0, 1, 3, 6, or 12 μ M. Protein extraction was performed by mixing 40 μ L of each standard curve/healthy plasma sample with 200 μ L acetonitrile containing 1% (w/v) trichloroacetic acid. Samples were shaken at 1500 rpm for 30 minutes at 25°C and then centrifuged for 10 minutes at maximal speed in a tabletop microcentrifuge. The supernatant was separated into 4 tubes (50 μ L each), and each tube was kept at -80° C before use. One experiment was done to confirm that frozen standard curve tubes and freshly prepared standard curve samples gave exactly the same results. Each day, 10–15 aliquots from patients taking tafamidis were thawed. 20 μ L plasma was mixed with 100 μ L acetonitrile containing 1% (w/v) trichloroacetic acid. Shaking and centrifugation were done as with the standard curve samples. After protein extraction, standard curve and patient sample supernatants (10 μ L) were injected into a Thermo Fisher Scientific BetaBasic 18 (50 \times 4.6 mm) column using an Agilent 1260 Infinity HPLC with fluorescence detection (excitation at 310 nm, emission at 370 nm). A gradient from 90% A (water 95%, 4.9% acetonitrile, 0.1% trifluoroacetic acid)/10% B (acetonitrile 95%, 4.9% water, 0.1% trifluoroacetic acid) to 0% A/100% B was run for 20 minutes at 1 mL/min. A peak at \sim 13 mL was seen both in plasma samples with tafamidis added ex vivo and in patient plasma samples. Pure tafamidis in DMSO also produces a peak at the same retention time, and we have previously shown by mass spectrometry that this peak corresponds to tafamidis in patient samples. Under the conditions used, the detection limit for tafamidis was 6.3 nM. All samples (including patient and standard curves) were prepared and injected in duplicate. The coefficient of variability was \leq 10%. To confirm that no tafamidis was lost in the protein precipitate using our acetonitrile method, we performed one experiment in which tafamidis was added ex vivo to healthy control plasma (5, 10, 15 μ M), protein was extracted as detailed above, and concentration was measured using the same HPLC method and standard curve of pure tafamidis in DMSO. Calculated concentrations were 5.34 ± 0.05 μ M, 11.15 ± 0.07 μ M, and 15.21 ± 0.07 μ M, showing that our method measures total C_{Taf} in plasma.

Tafamidis-gluc. was synthesized as described in Supplemental Figure 8 and Supplemental Methods. Fluorescence absorption and emission scans were taken to confirm that tafamidis-gluc. has the same fluorescence properties as unmodified tafamidis (i.e., maximum excitation 310 nm and maximum emission 370 nm). Tafamidis-gluc. was then injected in the HPLC using the same method as above, producing a unique peak at ~9 mL. We noticed that in our analysis of patient samples, the same peak at 9 mL was seen, and that this peak was not present when tafamidis was added ex vivo to healthy plasma. We performed the same experiment as above to confirm that all tafamidis-gluc. was extracted using our acetonitrile method. We then retrospectively analyzed 146 chromatograms and compared the AUC of the tafamidis peak (13 mL) to the tafamidis-gluc. peak (9 mL). While doing these experiments, we noticed that synthesized tafamidis-gluc. was hydrolyzed in healthy plasma when incubated overnight at 37°C. Therefore, we cannot exclude that some endogenous tafamidis-gluc. was being hydrolyzed to tafamidis, leading to an overestimation of plasma tafamidis levels.

Prediction model

The linear model represented by Equation 2 was fit to the numericized Response values using linear regression as implemented in Mathematica 11.3 (Wolfram) or Excel (Microsoft). In the few cases where data were missing (3 of 210 SNF scores; 9 of 210 TTR concentrations; and 2 of 210 tafamidis concentrations), we used mean imputation to replace them; that is, the missing values were replaced with the average value for patients of the same sex and response category. The patients were then divided into quintiles based on their calculated response values and the proportions of each response category in each quintile were calculated. The standard errors of the proportions were calculated as usual for probabilities, i.e., standard error = $(P_i \times (1 - P_i)/n)^{1/2}$, where P_i is the probability of event i and n is the number of patients in the sample (for our quintiles, $n = 210/5 = 42$) (35). Note that this standard error only considers the probabilities of the proportions individually. The confidence intervals of the proportions considered jointly are larger by about 60% (35). The alternative models, Model-NIS and Model-S, were created in an exactly analogous way.

The validation of the model by comparing the results from a training subset of the patient database with a test subset was also performed using Mathematica 11.3 to randomly split the patients into groups of equal size, fit the data to the training set as described above, split the patients into quintiles, and calculate the proportions of nonresponders, partial responders, and responders in each quintile. The fitted model was then applied to the test data set; the patients in the test dataset were sorted into the quintiles obtained from the training dataset; and the proportions of nonresponders, partial responders, and responders in each quintile were calculated. This process was iterated 100 times, and then the proportions of each response category were compared in the training and test datasets for each iteration. Finally, the root mean square of the differences was calculated.

Blinding

All plasma samples were analyzed for $C_{T_{49}}$ tetrameric TTR concentration, and TTR kinetic stability by a researcher blinded to response to therapy classification. Treatment status was not blinded. All patients were classified as responders, partial responders, or nonresponders by a researcher blinded to the biochemical analysis.

Statistics

All graphs and statistics were prepared in Prism 7 (GraphPad Software). For comparisons among 3 or more groups, we used a Kruskal-Wallis test with Dunn's multiple comparison correction. For comparisons between only 2 groups, we used a 2-tailed Mann-Whitney U test. Correlations were assessed by calculating the Spearman's correlation coefficient. All P values are indicated in the legends and figures. Scatter plots include a horizontal bar for the median value and error bars that represent the first quartile (Q1) and the third quartile (Q3). For analyses in which variables were compared among response categories at baseline, missing values were few (in the worst case, there were 13 missing values of 210 for baseline subunit exchange rate constants) and occurred at random because of technical problems (insufficient sample for the assay, high sample viscosity, etc.). Moreover, these missing values were spread across the response categories (5 nonresponders, 6 partial responders, and 2 responders). Therefore, these values were simply omitted from the analysis. For our longitudinal analysis of the change in NIS and weight as a function of time, we had many more data points at early than later time points. This circumstance arose simply because the patients in our population started tafamidis treatment on a rolling basis between July 2012 and January 2016. Data for the later time points were

only available for patients who began tafamidis therapy at the beginning of our window. This observation suggested that the missing values were missing at random and could be omitted.

Study approval

The study was approved by the ethical and institutional review boards at Centro Hospitalar do Porto and The Scripps Research Institute prior to subject enrollment. Written informed consent was received from participants prior to inclusion in the study.

Author contributions

CM, TC, and JWK designed the study; CM and JSM analyzed C_{Taf} , C_{TTR} , and TTR kinetic stability. CM, JA, JF, MN, and CR collected and introduced genetic, demographic, clinical, analytical, and neurophysiologic data into the study database. GJB synthesized tafamidis-gluc. CM, DLP, ETP, TC, and JWK analyzed the data. DLP and ETP built the predictive model. TC classified patients as responders, partial responders, and nonresponders. CM and ETP wrote the manuscript.

Acknowledgments

We would like to acknowledge the team of neurologists, cardiologists, nephrologists and ophthalmologists who have carefully observed and followed the patients at the Unidade Corino de Andrade, in particular: Ana Martins da Silva, Maria Cristina Alves, Márcio Cardoso, and Luis Maia. We thank also the nurses who collected the blood samples, in particular Susana Ferreira, and study coordinators Katia Valdez, Inês Cardoso, and Alexandra Sousa. We thank Joel Buxbaum for critically reading the final version of the manuscript. Finally, we thank all the patients and their families for their willingness to participate and support this study. This work was supported by an American Heart Association Predoctoral Fellowship (16PRE31130009) and the Ellen Browning Scripps Foundation fellowship awarded to CM and critical NIH support (DK 46335 to JWK and TC).

Address correspondence to: Teresa Coelho, Unidade Corino de Andrade, Centro Hospitalar do Porto, Largo do Prof. Abel Salazar, 4099-001 Porto, Portugal. Phone: 351.22.207.7500; Email: tcoelho@netcabo.pt. Or to: Jeffery W. Kelly, Departments of Chemistry and Molecular Medicine, The Scripps Research Institute, 10550 N Torrey Pines Rd, La Jolla, California 92037, USA. Phone: 858.784.9880; Email: jkelly@scripps.edu.

1. Rowczenio D, Ashutosh M. Mutations in Hereditary Amyloidosis. <http://www.amyloidosismutations.com/mut-attrib.php>. Created 2015. Accessed May 22, 2019.
2. Sipe JD, et al. Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 nomenclature guidelines. *Amyloid*. 2016;23(4):209–213.
3. Saraiva MJ, Birken S, Costa PP, Goodman DS. Family studies of the genetic abnormality in transthyretin (prealbumin) in Portuguese patients with familial amyloidotic polyneuropathy. *Ann N Y Acad Sci*. 1984;435:86–100.
4. Conceicao I, et al. “Red-flag” symptom clusters in transthyretin familial amyloid polyneuropathy. *J Peripher Nerv Syst*. 2016;21(1):5–9.
5. Andrade C. A peculiar form of peripheral neuropathy; familial atypical generalized amyloidosis with special involvement of the peripheral nerves. *Brain*. 1952;75(3):408–427.
6. Coutinho P, eds. *Amyloid and Amyloidosis*. Amsterdam, The Netherlands: Excerpta Medica; 1980:88–98.
7. Coelho T, Ines M, Conceicao I, Soares M, de Carvalho M, Costa J. Natural history and survival in stage 1 Val30Met transthyretin familial amyloid polyneuropathy. *Neurology*. 2018;91(21):e1999–e2009.
8. Coelho T, et al. Tafamidis for transthyretin familial amyloid polyneuropathy: a randomized, controlled trial. *Neurology*. 2012;79(8):785–792.
9. Berk JL, et al. Repurposing diflunisal for familial amyloid polyneuropathy: a randomized clinical trial. *JAMA*. 2013;310(24):2658–2667.
10. Benson MD, et al. Inotersen treatment for patients with hereditary transthyretin amyloidosis. *N Engl J Med*. 2018;379(1):22–31.
11. Adams D, et al. Patisiran, an RNAi therapeutic, for hereditary transthyretin amyloidosis. *N Engl J Med*. 2018;379(1):11–21.
12. Bulawa CE, et al. Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proc Natl Acad Sci USA*. 2012;109(24):9629–9634.
13. Coelho T, et al. Long-term effects of tafamidis for the treatment of transthyretin familial amyloid polyneuropathy. *J Neurol*. 2013;260(11):2802–2814.
14. Cortese A, et al. Monitoring effectiveness and safety of Tafamidis in transthyretin amyloidosis in Italy: a longitudinal multi-center study in a non-endemic area. *J Neurol*. 2016;263(5):916–924.
15. Lozeron P, et al. Effect on disability and safety of Tafamidis in late onset of Met30 transthyretin familial amyloid polyneuropathy. *Eur J Neurol*. 2013;20(12):1539–1545.
16. Planté-Bordeneuve V, et al. Long-term treatment of transthyretin familial amyloid polyneuropathy with tafamidis: a clinical and

- neurophysiological study. *J Neurol*. 2017;264(2):268–276.
17. Mundayat R, et al. Positive effectiveness of tafamidis in delaying disease progression in transthyretin familial amyloid polyneuropathy up to 2 years: an analysis from the transthyretin amyloidosis outcomes survey (THAOS). *Neurol Ther*. 2018;7(1):87–101.
 18. Sousa A, Coelho T, Barros J, Sequeiros J. Genetic epidemiology of familial amyloidotic polyneuropathy (FAP)-type I in Póvoa do Varzim and Vila do Conde (north of Portugal). *Am J Med Genet*. 1995;60(6):512–521.
 19. 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(1):229–239.
 20. Coelho T, Vinik A, Vinik EJ, Tripp T, Packman J, Grogan DR. Clinical measures in transthyretin familial amyloid polyneuropathy. *Muscle Nerve*. 2017;55(3):323–332.
 21. Vinik EJ, 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(3):497–508.
 22. Sekijima Y, Ueda M, Koike H, Misawa S, Ishii T, Ando Y. Diagnosis and management of transthyretin familial amyloid polyneuropathy in Japan: red-flag symptom clusters and treatment algorithm. *Orphanet J Rare Dis*. 2018;13(1):6.
 23. Buxbaum J, Anan I, Suhr O. Serum transthyretin levels in Swedish TTR V30M carriers. *Amyloid*. 2010;17(2):83–85.
 24. Hanson JLS, et al. Use of serum transthyretin as a prognostic indicator and predictor of outcome in cardiac amyloid disease associated with wild-type transthyretin. *Circ Heart Fail*. 2018;11(2):e004000.
 25. Rappley I, et al. Quantification of transthyretin kinetic stability in human plasma using subunit exchange. *Biochemistry*. 2014;53(12):1993–2006.
 26. Cho Y, et al. Personalized medicine approach for optimizing the dose of tafamidis to potentially ameliorate wild-type transthyretin amyloidosis (cardiomyopathy). *Amyloid*. 2015;22(3):175–180.
 27. Eisele YS, et al. Targeting protein aggregation for the treatment of degenerative diseases. *Nat Rev Drug Discov*. 2015;14(11):759–780.
 28. Rostami-Hodjegan A, Tucker GT. Simulation and prediction of in vivo drug metabolism in human populations from in vitro data. *Nat Rev Drug Discov*. 2007;6(2):140–148.
 29. Adams D, et al. Rapid progression of familial amyloidotic polyneuropathy: a multinational natural history study. *Neurology*. 2015;85(8):675–682.
 30. Kurian SM, et al. Peripheral blood cell gene expression diagnostic for identifying symptomatic transthyretin amyloidosis patients: male and female specific signatures. *Theranostics*. 2016;6(11):1792–1809.
 31. Khan TK. An algorithm for preclinical diagnosis of Alzheimer's disease. *Front Neurosci*. 2018;12:275.
 32. Hardy J. The amyloid hypothesis for Alzheimer's disease: a critical reappraisal. *J Neurochem*. 2009;110(4):1129–1134.
 33. Levey AS, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150(9):604–612.
 34. Choi S, Ong DS, Kelly JW. A stilbene that binds selectively to transthyretin in cells and remains dark until it undergoes a chemoselective reaction to create a bright blue fluorescent conjugate. *J Am Chem Soc*. 2010;132(45):16043–16051.
 35. May WL, Johnson WD. A SAS macro for constructing simultaneous confidence intervals for multinomial proportions. *Comput Methods Programs Biomed*. 1997;53(3):153–162.