

1 **Title:** Learning good therapeutic targets in ALS, neurodegeneration, using observational studies

2
3 Mohammadali Alidoost¹, Jeremy Y. Huang², Georgia Dermentzaki³, Anna S. Blazier,³ Giorgio Gaglia²,
4 Timothy R. Hammond³, Francesca Frau⁴, Mary Clare Mccorry⁵, Dimitry Ofengeim³, Jennifer L. Wilson^{1*}

5
6 **Affiliations:**

7 1. Department of Bioengineering, University of California Los Angeles, Los Angeles, CA, 90095, USA

8 2. Precision Medicine & Computational Biology, Sanofi Research US, Cambridge, MA, 01701, USA

9 3. Rare & Neurologic Diseases, Sanofi Research US, Cambridge, MA, 01701, USA

10 4. Evidence Generation & Decision Sciences, Sanofi Development, Frankfurt, 65929, Germany

11 5. Scientific Relations & Initiatives, Sanofi Research US, Cambridge, MA ,01701, USA

12 * To whom correspondence should be addressed: Jennifer L Wilson, jenniferwilson@g.ucla.edu, 420
13 Westwood Plaza, 5121 Eng V, Rm 4121D, Los Angeles CA, 90095, USA

14 **Funding**

15 This project was supported by a Sanofi iDEA-iTECH award.

16 **Competing interests**

17 The authors Jeremy Y. Huang, Georgia Dermentzaki, Anna S. Blazier, Giorgio Gaglia, Timothy R.

18 Hammond, Francesca Frau, Mary Clare Mccorry, and Dimitry Ofengeim are Sanofi employees and may

19 hold shares and/or stock options in the company.

20 **Keywords:** network biology, real-world data, systems modelling, neurodegeneration

21
22
23

24 **Abstract**

25 Analysis of real-world data (RWD) is attractive for its applicability to real-world scenarios but
26 RWD is typically used for drug repurposing and not therapeutic target discovery. Repurposing studies
27 have identified few effective options in neuroinflammatory diseases with relatively few patients such as
28 amyotrophic lateral sclerosis (ALS), which is characterized by progressive muscle weakness and death
29 with no disease-modifying treatments available. We previously reclassified drugs by their simulated
30 effects on proteins downstream of drug targets and observed class-level effects in the EHR, implicating
31 the downstream protein as the source of the effect. Here, we developed a novel ALS-focused pathways
32 model using data from patient samples, the public domain, and consortia. With this model, we simulated
33 drug effects on ALS and measured class effects on overall survival in retrospective EHR studies. We
34 observed an increased but non-significant risk of death for patients taking drugs associated with the
35 complement system downstream of their targets and experimentally validated drug effects on
36 complement activation. We repeated this for six protein classes, three of which, including multiple
37 chemokine receptors, were associated with a significant increased risk for death, suggesting that
38 targeting proteins such as chemokine receptors could be advantageous for these patients. We recovered
39 effects for drugs associated with complement activation and chemokine receptors in Parkinson's and
40 Myasthenia Gravis patients. We demonstrated the utility of network medicine for testing novel therapeutic
41 effects using RWD and believe this approach may accelerate target discovery in neuroinflammatory
42 diseases, addressing the critical need for new therapeutic options.

43
44

45 **Introduction**

46 We need novel treatment identification methods for rare, difficult-to-treat diseases, especially
47 those with insufficient experimental systems. Big data approaches and advanced algorithms have the
48 power to capture patient data and infer new disease drivers in ways previously not possible. Many
49 approaches suffer from limitations that were well-established a decade ago: animal models are valuable
50 for understanding mechanism but have limited relevance to human conditions, human epidemiology is
51 highly relevant to human disease but is insufficient for explaining causality, and natural conditions are the
52 most relevant and may explain causality but are rare¹. We hypothesized that a model-informed analysis of

53 patient data may overcome these hurdles, by predicting causality first, and later testing causal predictions
54 in relevant human disease contexts, thereby overcoming the limitations of traditional approaches.

55 Retrospective analyses of electronic health records (EHRs) are attractive for difficult-to-treat
56 conditions, including many neuroinflammatory conditions, where animal and model systems have failed to
57 deliver treatments with substantial impacts on survival. However, when inferring treatment impact from
58 real-world evidence, most studies are restricted to repurposing of known druggable targets. For instance,
59 the discovery of phosphodiesterase inhibitors as protective for Parkinson's² and amyotrophic lateral
60 sclerosis (ALS)³, peroxisome proliferator-activated receptor (PPAR) agonists as protective in Alzheimer's
61 Disease⁴, and the potential of multiple drugs, diazoxide, gefitinib, paliperidone, and dimethyltryptamine as
62 high-quality candidates for repurposing in ALS⁵ all leverage 'omics data, network medicine, and EHRs,
63 but they uncovered evidence for known, druggable targets. Indeed, reviews of the field emphasized that
64 real world data can support clinical trial emulation, sub phenotyping, or image analysis⁶, but the authors
65 failed to consider the historical record as sufficient for supporting *novel, untested* drug targets. Despite
66 the status quo, we had compelling prior evidence that it might be possible to more directly vet novel
67 targets in EHRs and sought to test that idea in difficult-to-treat neuroinflammatory populations.
68 Specifically, we learned that disease-specific network models improved prediction of drug effects⁷ and
69 that drug-effect predictions based on shared protein networks were detectable in EHRs⁸.

70 Debilitating neuroinflammatory conditions without substantially effective treatments have
71 motivated creative and compelling approaches to advance treatment options. Of the seven approved
72 treatments for ALS by the Federal Drug Administration (FDA), none are disease-modifying⁹ and none
73 prevent death. Multiple studies have identified distinct molecular mechanisms including the role of the
74 innate immune system (**Table S4**), among other genetic drivers^{9,10}. Targeting these disrupted pathways
75 or causal genetic drivers in the clinic has had limited efficacy and coverage due to the time to develop
76 new medicines and initiate new trials. Since ALS is a rare disease, clinical trial cohorts often struggle to
77 meet statistically significant thresholds, leading to increased interest in leveraging new insights from real-
78 world-data. Thus, we pursued our approach in neuroinflammatory conditions because of the potential for
79 accelerating discovery for these devastating conditions.

80 In this study, we tested our network inference platform on neuroinflammatory conditions and
81 corroborated our model predictions EHR investigation with experimental analysis. Intriguingly, our results
82 synergize with other investigations of dysregulated pathways in ALS, specifically, we recovered increased
83 risks for death associated with complement system activation. And despite the fact that ALS patients
84 were not prescribed any complement-targeted drugs, our retrospective analysis provided rationale for
85 directly inhibiting this pathway. We also asked whether these insights were generalizable across
86 additional neuromuscular and neurodegenerative diseases with larger patient populations, such as
87 Myasthenia gravis and Parkinson's disease. In total, we built and tested six network classes with different
88 observed significant impacts on overall patient survival in several modules. Encouragingly, our pipeline
89 independently identified therapeutic targets currently in clinical development and suggested additional
90 targets with stronger effects. By classifying approved therapies by their effects on downstream proteins,
91 rather than labeled primary targets, we can leverage large-scale cohort studies to identify novel targets
92 for ALS and other neurodegenerative diseases.

93

94 **Materials and Methods**

95 *Curating ALS-associated neurodegenerative disease pathways*

96 To customize our network analysis, we curated novel disease pathways from multiple sources. All
97 together, we included 195,252 gene-phenotype relationships across 1,096 unique disease pathways and
98 24,441 unique genes/transcripts. We curated this data from six sources listed below. Further, a list of all
99 derived pathway names, their sources, and the total number of genes is included in **Supplemental File 2**.
100 For a detailed list of pathway sources and processing for each source, and integration into our network
101 modeling platform please see *Supplemental Methods*.

102

103 *Running ALS-PathFX on all approved drugs*

104 Like previous versions of PathFX, the ALS version required drug binding proteins as inputs. We
105 used drug target information from DrugBank version 5.1.6. as inputs for ALS-PathFX analysis. We
106 generated networks and association tables for all drugs and then searched the PathFX-generated
107 association tables for phenotypes of interest. The association tables are standard outputs from PathFX

108 and included all network-associated phenotypes, the multiple-hypothesis-corrected p-values for the
109 associations, and which disease pathway proteins were discovered in the networks. All ALS-PathFX
110 predictions, created by merging all associations tables for drugs analyzed, are included in **Supplemental**
111 **File 2**.

112
113 *Generating network classes*

114 Using the merged table of all ALS-PathFX drug-phenotype associations, we counted how often a
115 network protein was used to make a drug-phenotype prediction (e.g., CXCL13 was used to predict a
116 network association from gabapentin to the “MS CSF” and “Bulk RNAseq ALS vs Control” pathways, and
117 C3 was used to predict a network association from gabapentin to the “Bulk RNAseq ALS vs Control”
118 pathway). After identifying network proteins occurring in the most drug-ALS PathFX associations, we
119 developed new drug classes by grouping drugs that shared network proteins (e.g., CXCL13 or C3). We
120 completed this analysis using the pandas module in python. Because we wanted to study patient data, we
121 also filtered out experimental or unapproved drugs and counted the number of approved active
122 ingredients with each network protein.

123 In a few instances, we observed several related network proteins and made multi-protein classes
124 based on qualitative and quantitative assessment of the ranked proteins, such as in the case for the
125 complement system proteins and chemokine proteins. Grouping drugs based on network proteins yielded
126 groups for the “target” cohorts (“network” class). To generate comparable “comparator” cohorts, we
127 assessed all remaining approved drugs that had network associations to the same ALS phenotypes, but
128 lacked a network protein in the association (“non-network” class). For instance, all drugs with an
129 association to the “MS CSF” or “Bulk RNAseq ALS vs Control” pathways would be considered in the non-
130 CXCL13 comparator class. All network proteins and their presence in drug networks are contained in
131 **Supplemental File 3**). Code for this analysis is included in our GitHub.

132
133 *Clinical data – Optum Market Clarity*

134 We used the Optum’s Market Clarity data: Optum® Market Clarity Data is an innovative dataset
135 of integrated multi-source medical and pharmacy claims and electronic health record (EHR) data that

136 enables detailed evaluation and assessment of the patient journey. The Market Clarity Data Asset
137 combines robust transactional pharmacy and medical claims data with best-of-breed EHR data. It is a
138 fully HIPAA-compliant, statistician-certified, de-identified precision data set that uncovers hidden
139 intelligence to turn insights into quick action. The Market Clarity Dataset links EHR data with historical,
140 linked administrative claim data, pharmacy claims, physician claims, facility claims (with clinical
141 information) and is inclusive of medications prescribed and administered. Clinically rich and specific data
142 elements sourced from the EHR include lab results, vital signs and measurements, diagnoses,
143 procedures and information derived from unstructured clinical notes using natural language processing.

144 We specifically assessed patient records available through February 2023 and accessed the
145 following data tables – the membership info for collecting date of birth, sex, race, ethnicity, and gender,
146 when available; the RX claims table to assessing filled prescriptions as a proxy for drug exposures; the
147 diagnosis table for identifying ALS patients and other comorbid conditions; and the date of death table for
148 collecting the month/year of death, when available.

149

150 *Observational studies in the electronic health record*

151 To measure effects for our network-derived drug classes, we used a standard target-comparator
152 approach using best practices in propensity score estimation and weighted treatment effects. We've used
153 a similar approach in our own work⁸ and describe the approach in greater detail in *Supplemental*
154 *Methods*.

155

156 *Stem Cell Culture*

157 Human iPSC astrocytes (iCell astrocytes, Fujifilm, Cat# C1037) were cultured in Matrigel-coated
158 (Corning 354277) 96-well plates with DMEM/F12 (Life Technologies 11330057) media supplemented with
159 heat-inactivated FBS (Gibco A38400-01) and N-2 supplement (Gibco 17502-048).

160 *Stimulation and drug treatment*

161 Three days post-plating, iCell astrocytes were treated with selective compounds (**Table S2**) at
162 100 nM, 1 μ M, and 10 μ M concentrations. For astrocyte stimulation (A1 subtype) we used the A1 cocktail:
163 IL-1a (R&D Systems, Cat# 200-LA, 0.3 ng/mL), TNF alpha (R&D Systems, Cat# 210-TA, 3 ng/mL) and

164 C1q (MyBiosource, Cat# MBS147305, 40 ng/mL) factors. Both the compounds, at multiple doses (100
165 nm, 1 uM, and 10 uM), and the A1 cocktail factors were added simultaneously on the cultures for
166 overnight treatment (24h).

167 168 *ELISA for complement system activation*

169 Media were collected 24h post-treatment and complement levels were assessed with the Milliplex
170 Human Complement Panel 1 (Millipore, HCMP1MAG-19K / Factors: C1q, C3, C3b/iC3b, C4, Factor B,
171 Factor H) and 2 (Millipore, HCMP2MAG-19K / Factors: C2, C4b, C5, C5a, Factor D, Mannose-binding
172 lectin (MBL), Factor I) according to the kit's instructions. The assay was run in a Bio-Plex 2000 instrument
173 (Biorad).

174 *Data availability*

175 The Market Clarity data used were licensed from Optum and are not publicly available.
176 To increase the rigor and reproducibility of our work, we have released code and data for the network
177 analysis, and anonymized code for accessing clinical data through a GitHub repository:
178 https://github.com/jenwilson521/network_drug_classes_als/ All supplemental Excel files are included in
179 their original format in the GitHub repository to increase readability and reusability of the data.

180 181 **Results**

182 *ALS-PathFX predicted drug network associations to ALS through distinct proteins*

183 Our overall approach used pathway modeling to identify novel drug classes and observational
184 studies in the health record to test the utility of these drug classes (**Figure 1**). To build a custom network
185 medicine platform, we curated several high-quality datasets from the TargetALS data portal and several
186 patient-derived datasets¹¹. The TargetALS datasets included differential genes measured by RNAseq
187 from multiple tissue types including cerebellum, motor cortex, and spine. We derived multiple gene sets
188 from this data using different fold-change thresholds (see *methods*). We also leveraged multiple novel
189 sequencing and proteomics data derived from patient samples, including snRNAseq in multiple tissue
190 types, proteomics from patient cerebral spinal fluid (CSF), and RNAseq data from reactive astrocytes and
191 microglia (for a full list see *methods* and¹¹). We specifically modified our highly-flexible PathFX platform
192 (released with ¹²) to emphasize these gene lists (referred to as “pathway phenotypes”) into PathFX¹²⁻¹⁴

193 using the original bias reduction methods published in ¹³ and described in methods. We then used our
194 custom network algorithm, ALS-PathFX, to model drug effects and derive new drug classes based on
195 predicted effects on nearby, but untargeted (not bound by drug) proteins (**Figure 1B**). Taken together, our
196 ALS-PathFX network model leverages several high-quality disease-relevant datasets to make predictions
197 about drug effects.

198 We generated ALS-PathFX networks for 7,013 approved and experimental drugs using targets
199 from DrugBank^{15,16}. Briefly, ALS-PathFX first identified high-confidence proteins downstream of druggable
200 targets, using interaction specificity analysis, a process that prioritizes protein interactions while reducing
201 connections through hub proteins. After generating a protein network for each drug, ALS-PathFX
202 identified network-associated phenotypes, by assessing which disease pathway proteins were enriched in
203 the drug network relative to the entire interactome, and using multiple hypothesis correction and
204 controlling for biases in the number of pathway-associated genes/proteins (as described in ¹³ and in
205 *methods*). In total, we discovered 13,967 drug-phenotype associations containing 2,984 unique drugs
206 and 90 unique disease pathways. For example, PathFX generated a network of 235 proteins and 238
207 edges for the antipsychotic, acetophenazine, and connected the network to 138 phenotypes, including 4
208 neurodegenerative phenotypes (subnetwork associated with four phenotypes shown in **Figure 2**).
209 Acetophenazine's network contained the complement system proteins Complement C3 (C3),
210 Complement C5 (C5), Complement C5a Receptor 1 (C5AR1), and Complement C3a Receptor 1
211 (C3AR1). Network proteins may be associated with more than one phenotype, and for instance, C3 and
212 C3AR1 were associated with "Multiple sclerosis" and "sequencing data 1". We assigned acetophenazine
213 to the "complement-associated drug class" or "CS-class" for brevity. We considered any drugs predicted
214 to affect the same phenotypes (e.g., "Multiple sclerosis" or "sequencing data 1") but without complement
215 proteins in their networks, as sufficient comparator drugs and considered them in the "non-CS-class" of
216 drugs. We provide more detail about ALS-PathFX results and assessment of network-protein drug
217 classes in the *Supplemental Results*.

218

219 *Complement-system-associated drugs are associated with changes in survival*

220 We hypothesized that drugs with complement system proteins in their networks would have
221 distinct clinical effects. This hypothesis is further supported by multiple recent developments of
222 complement system inhibitors for neurodegenerative disease, specifically myasthenia gravis (**Table S4**).
223 Aside from recent discoveries, complement system inhibitors have been approved since 2007 for several
224 other non-neurodegenerative diseases including autoimmune and blood disorders (**Table S4**). Despite
225 several approved compounds, none were prescribed in our ALS study patient population (*described*
226 *below and depicted in Figure S6A*, all compounds listed in **Supplemental File 4**), further emphasizing
227 the need for a way to anticipate their effects on overall survival.

228 To test whether ALS-PathFX-predicted drugs affected overall survival, we conducted a novel
229 observational study using a standard treatment-comparator study design and large-scale propensity
230 scoring to adjust for possible confounding variables. More specifically, we accessed the de-identified
231 Optum Market Clarity data. Optum Market Clarity deterministically links medical and pharmacy claims
232 with EHR data from providers across the continuum of care. This data is available for researchers without
233 IRB approval (*longer description in Methods*). The novelty in our approach was deriving network-protein-
234 based drug classes instead of using traditional drug classifications such as anatomical therapeutic
235 category (ATC) classification systems. We provide a more intuitive description of this process in the
236 *Supplemental Results*.

237 After controlling for possible confounding (**Figure S6**), we estimated the differences in overall
238 survival and observed a marginal, decreased, but not significant survival time for patients with an
239 exposure to CS-class drugs (HR = 1.031 95% confidence interval (0.937, 1.134), $p = 0.534$) (**Figure 3A**,
240 **Table 1**). The study discovered a detrimental effect early, where CS-class drugs are separated from non-
241 CS-class drugs before ~2,000 days (~5.5 years). At later timepoints, ~10,000 days (~27.4 years), overall
242 survival was minimally distinguishable. While the result is modest and insufficient to directly change
243 clinical practice, this result affirmed that network-derived drug classes could predict differential effects on
244 ALS survival in historical patient cohorts and suggested the possibility of testing other network drug
245 classes. Yet, as mentioned previously, the results do not confirm whether drugs activate or inhibit the
246 complement system.

247
248 *Network derived drug classes activate complement system proteins*

249 We next sought additional data sources and experiments to better understand possible drug-
250 induced effects on the complement system. We first sought drug-induced gene expression changes from
251 two public data sources, PharmOmics²¹ and LINCS²² (*Supplemental methods*) and saw little or
252 ambiguous effects in these data sources (*supplemental results*). We next hypothesized that protein-level
253 changes may better inform the observed effect on overall survival. We investigated complement levels
254 using ELISA, specifically measuring C3, Complement 4 (C4), and factor B levels released from healthy
255 stem cells following exposure to drugs from our network and non-network classes (**Figure 4**). Healthy
256 stem cells were a sufficient system because we started following patients before their ALS diagnosis and
257 this could indicate whether study drugs induced or suppressed complement system proteins. We
258 prioritized compounds for experimental testing based on those that had the highest number of unique
259 patients and total exposure days per patient (reflecting repeat prescription claims). The CS-class drugs,
260 mirtazapine and methylprednisolone increased activation of both C3 and C4, and the CS-class drug,
261 gabapentin, appeared to have no effect on either C3 and C4. Of the non-CS-class drugs tested,
262 cephalexin, amoxicillin, cefuroxime, nitroglycerin, and solatol had no effect on C3 or C4 levels, however,
263 the non-CS-class drugs, famotidine and erythromycin, increased C3 and C4 levels (**Figure 4**). Drugs from
264 both classes seemed to induce Factor B, but to a lesser effect. Some of these effects are supported by
265 additional literature such as different methylprednisolone regimens in myasthenia gravis patients led to
266 an increase in C3 and C4 levels²³. This increase in C3, C4 and Factor B for the above compounds
267 indicate that these compounds induce further complement activation once complement activation is
268 initiated, since treatment of human iPSC astrocytes alone w/o the A1 cocktail factors did not elicit any
269 complement response (undetectable levels).

270 In our network model, methylprednisolone, gabapentin and mirtazapine have distinct network
271 connections to the complement system proteins, C3, C5, complement C3a receptor (C3AR1), and
272 complement C5a receptor (C5AR1). Specifically, methylprednisolone, gabapentin, and mirtazapine
273 connected through annexin A1 (ANXA1), the adenosine A1 (ADORA1) receptor, and the opioid receptor
274 kappa 1 (OPRK1), respectively (all pathway connections in **Supplemental File 2**). This suggests that
275 alternative pathways could have differential effects on the complement system, but that the observed
276 trends in overall survival might be associated with increased activation of the complement system.

277

278 *Several network classes alter overall survival in ALS-patients*

279 We repeated our observational study analysis pipeline for an additional 5 drug classes and
280 ranked their overall effects on survival (**Figure 3B-H, Table 1**). We observed a large number of
281 prescription-days per study drug as observed with the complement system study, though, the exact brand
282 names changed per study (**Table S6, Figures S9, S11, S13, S15, S17**). Again, we saw low predictive
283 performance of the logistic regression, suggesting good covariate balance (**Table S7**), and similar
284 propensity score plots, suggesting sufficient class balance (**Figures S10, S12, S14, S16, S18**). Across all
285 drug classes, we saw an increased risk for death in the network drug class compared to the non-network
286 classes (**Figure 3G,H, Table 1**). The CXCR5 class had the greatest increased risk of death (HR = 1.151,
287 95% confidence interval (1.021-1.298), $p = 0.0214$) (**Figure 3B**), and the chemokine class had the next
288 highest risk (HR = 1.123, 95% confidence interval (1.018-1.239), $p = 0.0202$) (**Figure 3C**). The CNR2
289 class had a marginally decreased, but not significant risk (HR = 0.994, 95% confidence interval (0.890-
290 1.097), $p = 0.897$), most similar to the value observed for the complement system drug class. All other
291 survival curves provided in **Figures S19-S21**. All of the retained brand names, the number of unique
292 patients with an RX claim, the prescription-days per drug, and patient demographics are included in the
293 extended supplement for each network class comparison: chemokine (**Supplemental File 6**), CXCR3
294 (**Supplemental File 7**), CXCR5 (**Supplemental File 8**), CNR2 (**Supplemental File 9**), and NPY
295 (**Supplemental File 10**). All hazard ratios also reported in **Supplemental File 16**. We also investigated
296 the extent to which network drug classes shared brand names and observed shared, but distinct brand
297 names in each drug class comparison as measured by Jaccard similarity (**Figure S22C, D, Supplemental**
298 *methods and results*).

299

300 *Network-drug classes affect overall survival in multiple neurodegenerative diseases*

301 An intriguing utility for our analysis could be in identification of key nodes of pathophysiology
302 across multiple diseases on the nervous system. We first demonstrated that neurodegenerative diseases
303 shared pathways information using network models (**Figure S23, Supplemental methods and results**).
304 Intriguingly, neurodegenerative diseases share more genes near druggable targets than if comparing all

305 disease genes. We thus repeated measurements for some network classes in multiple disease
306 indications. We specifically considered effects of the complement system in myasthenia gravis and
307 Parkinson's disease and the CXCR3, and CXCR5 classes in myasthenia gravis. We again assessed
308 prescription days per drug for all studies (**Table S6, Figures S24, S26, S28, S30**), measured low
309 predictive performance of a logistic regression analysis (**Table S7**, and saw similar propensity score plots
310 between target and comparator cohorts (**Figures S25, S27, S29, S31**). Interestingly, we observed a
311 similar trend for myasthenia gravis patients taking the complement system drugs where early survival
312 (before ~2,000 days) showed separation between the CS-class and non-CS-class drugs, however, the
313 curves crossed-over at later timepoints (~10,000 days), suggesting a protective effect of complement-
314 associated drugs at early time points. Unlike our analysis in ALS patients, this analysis yielded a slight,
315 but significant increased risk for death (HR 1.098, 95% confidence interval (1.018-1.183), $p = 0.015$)
316 (**Figure 3D**). This finding combined with our experimental results and the recent approvals of complement
317 inhibitors for myasthenia gravis further suggest that activated complement is detrimental for overall
318 survival in neurodegeneration, but only at early timepoints. The effect of CXCR5 drugs was similar in
319 myasthenia gravis patients relative to ALS patients (HR 1.143, 95% confidence interval (1.052-1.241), $p =$
320 0.00155) (**Figure 3E**) and the effect of CXCR3 drugs was greater than in ALS patients (HR 1.321, 95%
321 confidence interval (1.223-1.427), $p = 1.47 \times 10^{-12}$) (**Figure 3F**). The effect of complement-system drugs
322 in Parkinson's patients showed, similar, but starker cross-over effects like we observed in ALS and
323 myasthenia gravis patients. We observed no overall survival benefit or risk (HR 0.992, 95% confidence
324 interval (0.965-1.020), $p = 0.579$) (**Figure S32**). Like before, all of the retained brand names, the number
325 of unique patients with an RX claim, the prescription-days per drug, and patient demographics are
326 included in the extended supplement for each network class comparison: CXCR5 in myasthenia gravis
327 (**Supplemental File 12**), CXCR3 in myasthenia gravis (**Supplemental File 13**), complement system in
328 myasthenia gravis (**Supplemental File 14**), and complement system in Parkinson's (**Supplemental File**
329 **15**).

330

331 **Discussion**

332 We presented an approach for using proteins downstream of druggable targets to define new
333 drug classes and tested the effects of these drug classes on overall survival in ALS and two additional
334 neurodegenerative diseases. We used observational studies as exploratory analyses, and while they
335 would require further, dedicated clinical investigations to confirm their effects, they are compelling for their
336 ability to recover known effects – specifically the effect of complement activation as detrimental to survival
337 in ALS, myasthenia gravis, and Parkinson's disease. Our preliminary experimental data combined with
338 recent approvals of complement inhibitors for myasthenia gravis patients (**Table S4**) further supports the
339 potential of complement inhibitors in ALS. Overall, this suggests a novel and impactful approach for
340 identifying disease-drivers in rare and difficult-to-treat diseases and may provide valuable insights to
341 therapeutic target selection.

342 Our results are further compelling for their utility in difficult-to-treat populations. ALS is
343 insufficiently explained by genetics, with less than 10% of familial and sporadic ALS cases explained by
344 genetic drivers^{9,10}. Without sufficient molecular targets, novel approaches are desperately needed to
345 identify treatment strategies for the disease. Further, the ALS population is relatively small and is typically
346 insufficient for observational studies. Indeed, through a related project, the Veterans Affairs dataset,
347 which contained nearly 20,000 ALS patients, is the largest repository of historical ALS patient data and is
348 nearly double the size of the cohort in this study²⁴. Yet, we still resolved statistically-significant clinical
349 effects by emphasizing proteins downstream of druggable targets, which allowed the inclusion of more
350 patients.

351 While we believe our paradigm can greatly enhance model-informed drug development efforts,
352 further work could overcome limitations in our current study. For instance, we had insufficient network
353 connections to all possible druggable targets and were limited to deriving network classes for proteins
354 that were “close” to approved targets. Our network drug classes also shared many approved drugs
355 indicating that our current interaction network was only sufficiently connected to some pathways.
356 However, the differences in network phenotypes effectively separated drugs to discern relatively stronger
357 effects, such as the increased risks associated with CXCR5 networks. This emphasizes the importance of
358 deriving high-quality, disease pathway phenotypes. We also tested these network classes in a single
359 patient population and further testing using data from additional health systems could elevate the impact

360 of these findings. Indeed, previous meta-analyses across healthcare systems have successfully
361 discovered drug-class effects on cardiovascular disease outcomes²⁵. Further, an intriguing opportunity, or
362 possible challenge, of historical data analysis is how to interpret the outcomes in the context of EHR
363 timescales and understand their relevance to dedicated clinical investigations. We followed patients for
364 nearly two decades, when most clinical trials are months to years in duration. Of the four FDA-approved
365 therapies for ALS, the riluzole trials measured outcomes at 12 months²⁶ or 18 months follow-up²⁷ and the
366 trials for edavarone, phenylbutyrate-taurursodiol (PB-TUDCA), and tofersen measured changes in the
367 revised ALS functional rating scale²⁸ at shorter durations – ranging from 24-28 weeks^{9,29-32}. Our
368 measured differences in proportions of surviving patients at short durations (~2 years) are similar to effect
369 sizes in the riluzole trials, suggesting that future observational studies, which leverage drug network
370 classes, could shorten observation times to better anticipate effect sizes measurable in the duration of a
371 clinical study.

372 Our discovery of disease drivers complements an evolving landscape of shared molecular
373 underpinnings of clinically-distinct diseases. Specifically, we predicted that activated complement would
374 increase risk for death across multiple disease indications. Our pathways analysis of shared
375 dysregulation is supported by other findings, specifically, Arneson et al³³ noted that when disease genes
376 and proteins weren't shared across neurodegenerative diseases, gene and protein functional information
377 was shared, further emphasizing shared dysregulation across neurodegenerative diseases. Separate
378 investigations of post-mortem brain samples further corroborated shared gene-level changes across
379 patients with four neurodegenerative diseases (Alzheimer's Disease, Parkinson's disease, Huntington's
380 disease, and ALS) and discovered that shared genes were involved in functional processes such as
381 inflammation, mitochondrial dysfunction, and oxidative stress^{33,34}. Additionally, multiple canonical
382 inflammatory reactivity inducers, including complement subcomponent C1q, interleukin 1 alpha (IL-1 α),
383 and tumor necrosis factor (TNF), convert astrocytes into A1 reactive astrocytes³⁵, and these
384 proinflammatory, reactive astrocytes secrete multiple complement factors, reduce synaptic and neuronal
385 connectivity, and are linked extensively to neurodegenerative disease³⁶. Taken together, this further
386 supports the possibility of shared molecular changes in these diseases, though, further experimental
387 validation would be required to confirm this general finding.

388 Our approach motivates extending the integration of clinical health data into Model Informed Drug
389 Development (MIDD). Already, the FDA has provided guidance about the use of real-world data (RWD)
390 for generating real world evidence (RWE) from sources other than randomization, controlled trials
391 (RCTs)(**Table S9**). Indeed, many have generated RWE to support regulatory decision making including
392 accelerated approvals, learning intrinsic factors, and optimizing dosing (**Table S9**). Yet these studies did
393 not leverage RWD for novel target support.

394 Other sources have identified novel targets from RWD, although, they have focused directly on
395 target-driven effects instead of downstream effects. Genetic data repositories have helped identify lead
396 candidates from patient data, such as in the case of Proprotein convertase subtilisin/kexin type 9
397 (PCSK9) inhibitors where inactivating and gain-of-function mutants were associated with elevated or
398 reduced cholesterol, respectively^{37,38}. Mendelian randomization is another method of inferring drug-target
399 effects in new populations³⁹, but not every suitable target has sufficient genetic connections to support
400 later therapeutic development. Additionally, analysis of PheWAS and expression quantitative trait loci
401 (eQTLs) have also coupled gene variation with observable clinical effects, especially side-effects^{40,41}. In
402 contrast to our approach, these approaches overlook downstream effects and require sufficient genetic
403 variation in patient populations to advance therapeutic discovery.

404 For the most part, prior work combining molecular models of drug effects with patient data have
405 been limited to drug repurposing studies, and we believe our approach greatly expands the possibilities
406 for learning promising therapeutic targets from real-world patient data. Knowing that electronic health
407 records will be increasingly useful to support new target identification, we anticipate that network
408 medicine will extend capabilities beyond target-centric approaches.

409

410 **Study Highlights**

411 Currently, the impact of real-world data (RWD) on druggable targets with translational impact is limited to
412 drug-repurposing. However, we hypothesized that network models of drug-induced signaling could
413 elevate the impact of RWD for identifying novel drug targets. This study addressed the feasibility of that
414 approach by simulating the effects of novel targets and testing these effects in neuroinflammatory
415 conditions. We demonstrated that drugs that alter complement activation have deleterious effects on

416 overall survival and these results are consistent with effects of recently FDA-approved drugs, although
417 none of our patients were taking them. We believe this paradigm will change translational science in how
418 RWD may be used to identify novel, druggable targets.

419 **Acknowledgments**

420 The authors would like to acknowledge the Sanofi real world data team, and especially Thomas Verrier,
421 the RWD-Platform Support team for their invaluable support navigating data access, without which, the
422 study could not be completed.

423 **Contribution statement**

424 J.L.W., M.A., J.Y.H., G.D., A.S.B., F.F., D.O. wrote the manuscript. J.L.W., J.Y.H, G.D., A.S.B., D.O.
425 designed the research. J.L.W., J.Y.H., M.A., G.D. performed the research. J.L.W., M.A., G.D. analyzed
426 the data. A.S., G.G., T.H., F.F., MC. M. contributed new reagents/analytical tools.

427

428

429 **References**

430

- 431 1. Plenge, R. M., Scolnick, E. M. & Altshuler, D. Validating therapeutic targets through human genetics.
432 *Nat. Rev. Drug Discov.* **12**, 581–594 (2013).
- 433 2. Adesuyan, M. *et al.* Phosphodiesterase Type 5 Inhibitors in Men With Erectile Dysfunction and the Risk
434 of Alzheimer Disease. *Neurology* **102**, e209131 (2024).
- 435 3. Fang, J. *et al.* Endophenotype-based in silico network medicine discovery combined with insurance
436 record data mining identifies sildenafil as a candidate drug for Alzheimer’s disease. *Nat Aging* **1**, 1175–
437 1188 (2021).
- 438 4. Fang, J. *et al.* Artificial intelligence framework identifies candidate targets for drug repurposing in
439 Alzheimer’s disease. *Alzheimer’s Res Ther* **14**, 7 (2022).
- 440 5. Cheng, F. *et al.* Network medicine informed multi-omics integration identifies drug targets and
441 repurposable medicines for Amyotrophic Lateral Sclerosis. (2024).doi:10.21203/rs.3.rs-4271755/v1
- 442 6. Cheng, F. *et al.* Artificial intelligence and open science in discovery of disease-modifying medicines for
443 Alzheimer’s disease. *Cell Rep. Med.* **5**, 101379 (2024).

- 444 7. Wilson, J. L., Gravina, A. & Grimes, K. From random to predictive: a context-specific interaction
445 framework improves selection of drug protein–protein interactions for unknown drug pathways. *Integr Biol*
446 (2022).doi:10.1093/intbio/zyac002
- 447 8. Wilson, J. L. *et al.* A network paradigm predicts drug synergistic effects using downstream protein–
448 protein interactions. *Cpt Pharmacometrics Syst Pharmacol* (2022).doi:10.1002/psp4.12861
- 449 9. Ilieva, H., Vullaganti, M. & Kwan, J. Advances in molecular pathology, diagnosis, and treatment of
450 amyotrophic lateral sclerosis. *BMJ* **383**, e075037 (2023).
- 451 10. Ghasemi, M. & Brown, R. H. Genetics of Amyotrophic Lateral Sclerosis. *Cold Spring Harb. Perspect.*
452 *Med.* **8**, a024125 (2018).
- 453 11. Zelic, M. *et al.* Glial state changes and neuroinflammatory RIPK1 signaling are a key feature of ALS
454 pathogenesis. *bioRxiv* 2024.04.12.589201 (2024).doi:10.1101/2024.04.12.589201
- 455 12. Wilson, J. L., Wong, M., Stepanov, N., Petkovic, D. P. PhenClust, a standalone tool for identifying
456 trends within sets of biological phenotypes using semantic similarity and the Unified Medical Language
457 System. *JAMIA Open* **4**, ooab079 (2021).
- 458 13. Wilson, J. L. *et al.* PathFX provides mechanistic insights into drug efficacy and safety for regulatory
459 review and therapeutic development. *PLoS Comput Biol* **14**, e1006614 (2018).
- 460 14. Wilson, J. L. *et al.* PathFXweb: a web application for identifying drug safety and efficacy phenotypes.
461 *Bioinformatics* **11**, 265–4506 (2019).
- 462 15. Wishart, D. S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration.
463 *Nucleic Acids Research* **34**, D668-72 (2006).
- 464 16. Law, V. *et al.* DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Research* **42**,
465 D1091-7 (2014).
- 466 17. Tian, Y., of, M. S. I. journal & 2018 Evaluating large-scale propensity score performance through real-
467 world and synthetic data experiments. *academic.oup.com* at <<https://academic.oup.com/ije/article-abstract/47/6/2005/5043131>>
468
- 469 18. Glynn, R. J., Schneeweiss, S. & Stürmer, T. Indications for propensity scores and review of their use
470 in pharmacoepidemiology. *Basic Clin. Pharmacol. Toxicol.* **98**, 253–259 (2006).
- 471 19. Sekhon, J. S. Multivariate and Propensity Score Matching Software with Automated Balance
472 Optimization: The Matching Package for R. at
473 <<https://papers.ssrn.com/sol3/Delivery.cfm?abstractid=1009044>>
- 474 20. Austin, P. C. The use of propensity score methods with survival or time-to-event outcomes: reporting
475 measures of effect similar to those used in randomized experiments. *Stat. Med.* **33**, 1242–1258 (2014).
- 476 21. Chen, Y.-W. *et al.* PharmOmics: A species- and tissue-specific drug signature database and gene-
477 network-based drug repositioning tool. *iScience* **25**, 104052 (2022).
- 478 22. Keenan, A. B. *et al.* The Library of Integrated Network-Based Cellular Signatures NIH Program:
479 System-Level Cataloging of Human Cells Response to Perturbations. *Cell Syst.* **6**, 13–24 (2018).

- 480 23. Lin, X. & Qi, G. Observation on the efficacy of different methylprednisolone regimens in the treatment
481 of myasthenia gravis. *Pak. J. Méd. Sci.* **38**, 910–915 (2022).
- 482 24. Reimer, R. J. *et al.* An electronic health record cohort of Veterans with amyotrophic lateral sclerosis.
483 *Amyotroph. Lateral Scler. Front. Degener.* **24**, 695–701 (2023).
- 484 25. Suchard, M. A. *et al.* Comprehensive comparative effectiveness and safety of first-line
485 antihypertensive drug classes: a systematic, multinational, large-scale analysis. *The Lancet* **394**, 1816–
486 1826 (2019).
- 487 26. Bensimon, G., Lacomblez, L. & Meininger, V. A Controlled Trial of Riluzole in Amyotrophic Lateral
488 Sclerosis. *N. Engl. J. Med.* **330**, 585–591 (1994).
- 489 27. Il, A. L. S. S. G. *et al.* Dose-ranging study of riluzole in amyotrophic lateral sclerosis. *Lancet* **347**,
490 1425–1431 (1996).
- 491 28. Cedarbaum, J. M. *et al.* The ALSFRS-R: a revised ALS functional rating scale that incorporates
492 assessments of respiratory function. *J. Neurol. Sci.* **169**, 13–21 (1999).
- 493 29. Abe, K. *et al.* Safety and efficacy of edaravone in well defined patients with amyotrophic lateral
494 sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* **16**, 505–512 (2017).
- 495 30. Paganoni, S. *et al.* Trial of Sodium Phenylbutyrate–Taurursodiol for Amyotrophic Lateral Sclerosis. *N.*
496 *Engl. J. Med.* **383**, 919–930 (2020).
- 497 31. Miller, T. M. *et al.* Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS. *N. Engl. J. Med.* **387**,
498 1099–1110 (2022).
- 499 32. Brooks, B. R. *et al.* Edaravone efficacy in amyotrophic lateral sclerosis with reduced forced vital
500 capacity: Post-hoc analysis of Study 19 (MC1186-19) [clinical trial NCT01492686]. *PLoS ONE* **17**,
501 e0258614 (2022).
- 502 33. Arneson, D., Zhang, Y., Yang, X. & Narayanan, M. Shared mechanisms among neurodegenerative
503 diseases: from genetic factors to gene networks. *J. Genet.* **97**, 795–806 (2018).
- 504 34. Li, M. D., Burns, T. C., Morgan, A. A. & Khatri, P. Integrated multi-cohort transcriptional meta-analysis
505 of neurodegenerative diseases. *Acta Neuropathol. Commun.* **2**, 93 (2014).
- 506 35. Liddelow, S. A. *et al.* Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**,
507 481–487 (2017).
- 508 36. Lawrence, J. M., Schardien, K., Wigdahl, B. & Nonnemacher, M. R. Roles of neuropathology-
509 associated reactive astrocytes: a systematic review. *Acta Neuropathol. Commun.* **11**, 42 (2023).
- 510 37. Cohen, J. C., Boerwinkle, E., Mosley, T. H. & Hobbs, H. H. Sequence Variations in PCSK9, Low LDL,
511 and Protection against Coronary Heart Disease. *N. Engl. J. Med.* **354**, 1264–1272 (2006).
- 512 38. Horton, J. D., Cohen, J. C. & Hobbs, H. H. Molecular biology of PCSK9: its role in LDL metabolism.
513 *Trends Biochem. Sci.* **32**, 71–77 (2007).
- 514 39. Bowden, J. & Holmes, M. V. Meta-analysis and Mendelian randomization: A review. *Res. Synth.*
515 *Methods* **10**, 486–496 (2019).

516 40. Duffy, Á. *et al.* Tissue-specific genetic features inform prediction of drug side effects in clinical trials.
517 *Sci. Adv.* **6**, (2020).

518 41. Nguyen, P. A., Born, D. A., Deaton, A. M., Nioi, P. & Ward, L. D. Phenotypes associated with genes
519 encoding drug targets are predictive of clinical trial side effects. *Nat. Commun.* **10**, 1579 (2019).

520

521

522 **Tables**

<i>Pathway Name</i>	ALS - HR	ALS - CI low	ALS- CI high	MG - HR	MG - CI low	MG - CI high	PD - HR	PD - CI low	PD - CI high
CXCR5	1.151	1.021	1.298	1.143	1.052	1.241			
CXCR3	1.124	0.995	1.269	1.321	1.223	1.427			
Chemokine Pathways	1.123	1.018	1.239						
Neuropeptide Y	1.088	0.985	1.203						
CNR2	1.0.99	1.890	1.097						
Complement System	1.031	0.937	1.134	1.098	1.018	1.183	0.992	0.965	1.020

523

524 **Table 1. Hazard ratios and confidence intervals for 6 drug classes across 3**

525 **neurodegenerative indications.** ALS = Amyotrophic lateral sclerosis, MG = myasthenia gravis,
 526 and PD = Parkinson’s disease. CI = 95% confidence interval.

527

528

529 **Figure Legends**

530 **Figure 1. A combined pathways and observational study approach for identifying pathways with**
531 **protective or harmful effects on overall survival in ALS. (A)** Our overall workflow consisted of ‘omics
532 data integration using several high-quality data sets, network modeling, identification of downstream
533 proteins, identifying patient cohorts, and effects estimation **(B)** For network inference, we modeled all
534 approved, active ingredients using ALS-PathFX and then created novel drug classes based on
535 downstream proteins included in drug networks. In this example, we assessed drugs with network
536 associations to multiple ALS sequencing datasets and included them in the “complement system” (“CS”)
537 drug class if they had any complement-system proteins in their networks. We highlighted five drugs that
538 are associated with any one of two disease pathways (i.e., “ALS sequencing data 1”, or “ALS sequencing
539 data 2”). In our schematic, we also highlighted that drugs “1”, “2”, and “3” contain downstream
540 associations to complement-system proteins and drugs “4” and “5” are connected to the same pathway
541 phenotypes, but through different proteins. Drugs, network proteins, and disease phenotypes are
542 represented by triangles, circles, or yellow boxes, respectively. **(C)** Later, we measured overall survival for
543 patients exposed to at least one of the “complement class” or “non-complement class” drugs using
544 observational studies in the Optum Market Clarity dataset. In this example, we grouped drugs “1”, “2”, and
545 “3” into the “target” class, and drugs “4” and “5” into the “comparator” class. We repeated this analysis for
546 6 drug classes and ranked classes by their effects on overall survival. Three complement proteins are
547 shown in the diagram, but a total of nine were considered when assigning drugs to the complement
548 system class: “Complement component 1, R subcomponent (C1R)”; “Complement component 1, S
549 subcomponent (C1S)”; “Complement component 1, Q subcomponent, alpha polypeptide (C1QA)”;
550 “Complement component 1, Q subcomponent, beta polypeptide (C1QB)” “Complement component 1, Q
551 subcomponent, alpha polypeptide (C1QA)”; “Complement component 1, Q subcomponent, gamma
552 polypeptide (C1QC)”; “Complement C3 (C3)”; “Complement C5 (C5)”; “Complement C3a receptor 1
553 (C3AR1)”; & “Complement C5a receptor 1 (C5AR1)”.

554

555 **Figure 2. ALS-PathFX connected acetophenazine to multiple neurodegenerative phenotypes.**

556 PathFX connected the drug, acetophenazine (orange triangle) to four relevant neurodegenerative

557 phenotypes (green boxes): Alzheimer disease, Huntington's chorea, seq_data_1, and Multiple Sclerosis.
558 PathFX used three drug target proteins (red circles): androgen receptor (AR), and the dopamine
559 receptors D1 and D2 (DRD1, DRD2), and 99 adjacent proteins (gray ellipses) when predicting
560 associations to these phenotypes.

561
562 **Figure 3. Patient cohorts in network drug classes have sufficient covariate balance (A)** We plotted
563 the total exposure days (sum of all prescriptions) per brand name per patient. We repeated this process
564 for the target, CS-class (left) and comparator, non-CS-class (right) drugs. Each dot is a single patient. **(B)**
565 The number of patients is plotted against their propensity score for patients on non-CS-class drugs (left)
566 or CS-class drugs (right). The propensity score is the predicted probability using LogisticRegression on
567 patient demographic, diagnostics, and medical prescription claims features. Patient demographics
568 information for race **(C)** and gender **(D)** concepts.

569
570 **Figure 3. Multiple network drug classes have effects on overall survival in ALS, myasthenia gravis**
571 We plotted Kaplan Meier survival curves for the complement system class **(A)**, the chemokine receptor 5
572 (CXCR5) class, a relatively high-risk class **(B)**, and the chemokine pathway class, also a relatively-high-
573 risk class **(C)** in ALS patients. We also plotted Kaplan Meier plots for three network classes in myasthenia
574 gravis patients: the complement system **(D)**, chemokine receptor 5 (CXCR5) **(E)**, and the chemokine
575 receptor 3 (CXCR3) **(F)** classes. In all plots, the proportion of the surviving population plotted against
576 time in days. Patients exposed to a “network class” or “Non-network class” drug are represented with a
577 dotted or solid line, respectively. Hazard ratio values with the 95% confidence interval for six network-
578 drug classes **(G)** and their $-\log(-p\text{-values})$ **(H)** as measured with weighted Cox proportional hazards
579 (coxph from the R survival package).

580
581 **Figure 4. Complement-associated drugs activate multiple components of the complement**
582 **pathway in ELISA assays.** Normalized values (to CTL: A1 cocktail factors) of fold change for the
583 complement factors C3, C4, and Factor B following treatment of human iPSC astrocytes with the A1
584 cocktail factors and selective classes of compounds (100 nM, 1 μ M and 10 μ M). Addition of

585 methylprednisolone, mirtazapine and albuterol showed the strongest effect in increasing the complement
586 levels (especially C3).

587

588 **Supplementary Material**

589 Supplemental File 1 – extended methods, results, Tables S1-8, Figures S1-32.

590 Supplemental File 2 – ALS-PathFX pathways, all modeling results.

591 Supplemental File 3 – All network-protein derived drug classes.

592 Supplemental File 4 – Drug brand names, prescription days, patient demographic data, and covariates for
593 the study of ALS patients and the CS-class.

594 Supplemental File 5 – All diagnosis codes used to select patients.

595 Supplemental File 6 – Drug brand names, prescription days, patient demographic data, and covariates for
596 the study of ALS patients and the CXC-class.

597 Supplemental File 7 – Drug brand names, prescription days, patient demographic data, and covariates for
598 the study of ALS patients and the CXCR3-class.

599 Supplemental File 8 – Drug brand names, prescription days, patient demographic data, and covariates for
600 the study of ALS patients and the CXCR5-class.

601 Supplemental File 9 – Drug brand names, prescription days, patient demographic data, and covariates for
602 the study of ALS patients and the CNR2-class.

603 Supplemental File 10 – Drug brand names, prescription days, patient demographic data, and covariates
604 for the study of ALS patients and the NPY-class.

605 Supplemental File 11 – PathFX predictions for multiple neurodegenerative diseases.

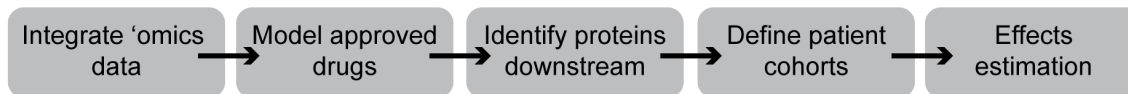
606 Supplemental File 12 – Drug brand names, prescription days, patient demographic data, and covariates
607 for the study of MG patients and the CXCR5-class.

608 Supplemental File 13 – Drug brand names, prescription days, patient demographic data, and covariates
609 for the study of MG patients and the CXCR3-class.

610 Supplemental File 14 – Drug brand names, prescription days, patient demographic data, and covariates
611 for the study of MG patients and the CS-class.

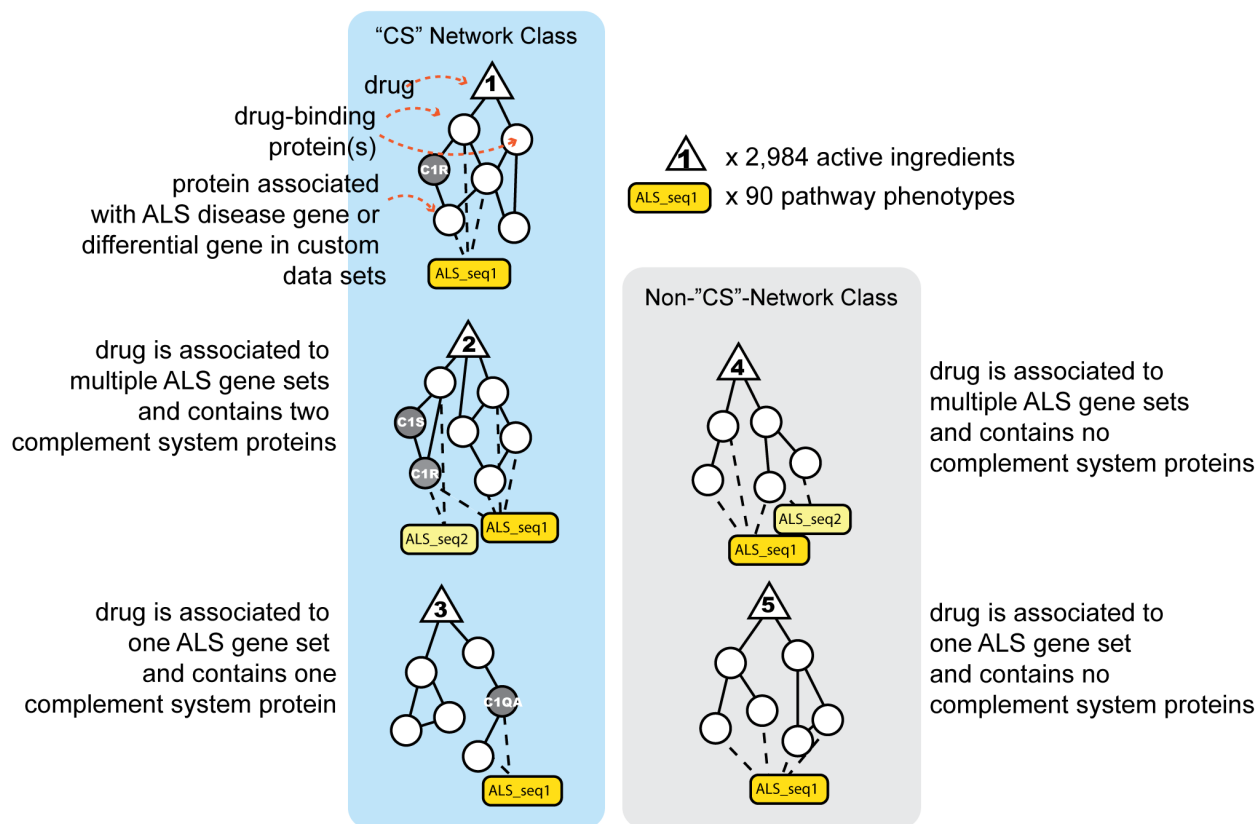
- 612 Supplemental File 15 – Drug brand names, prescription days, patient demographic data, and covariates
613 for the study of PD patients and the CS-class.
614 Supplemental File 16 – Table of hazard ratios and confidence intervals for all studies.

A



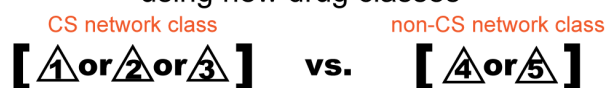
B

Model drugs with PathFX, assess ALS-associated network proteins, define new drug classes

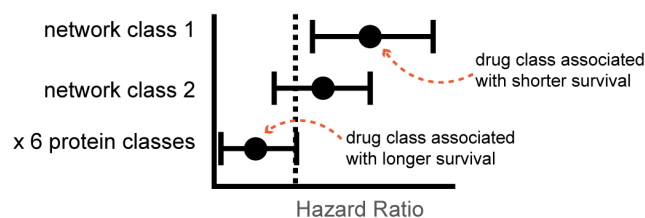


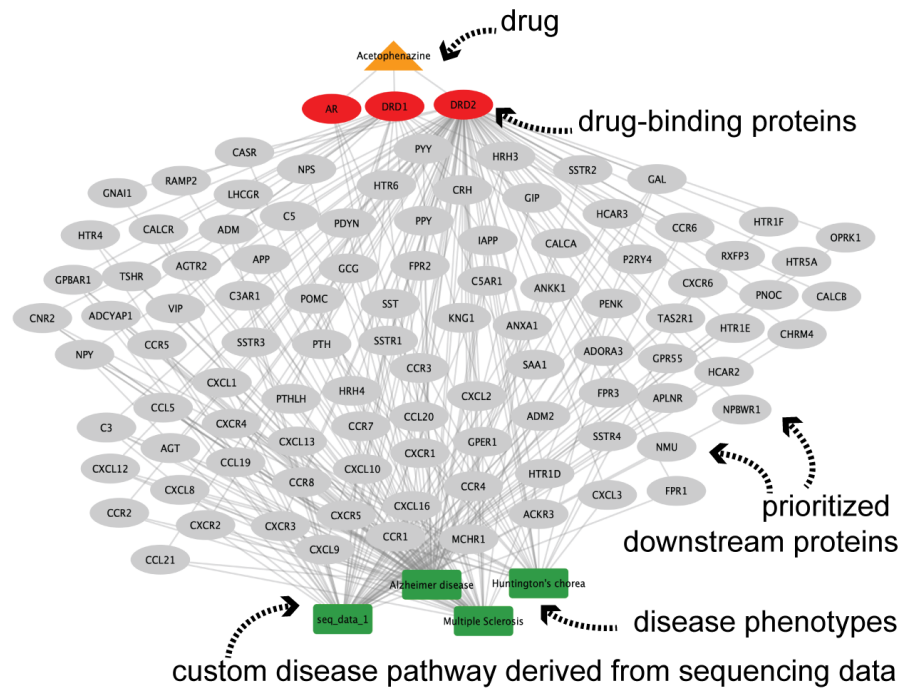
C

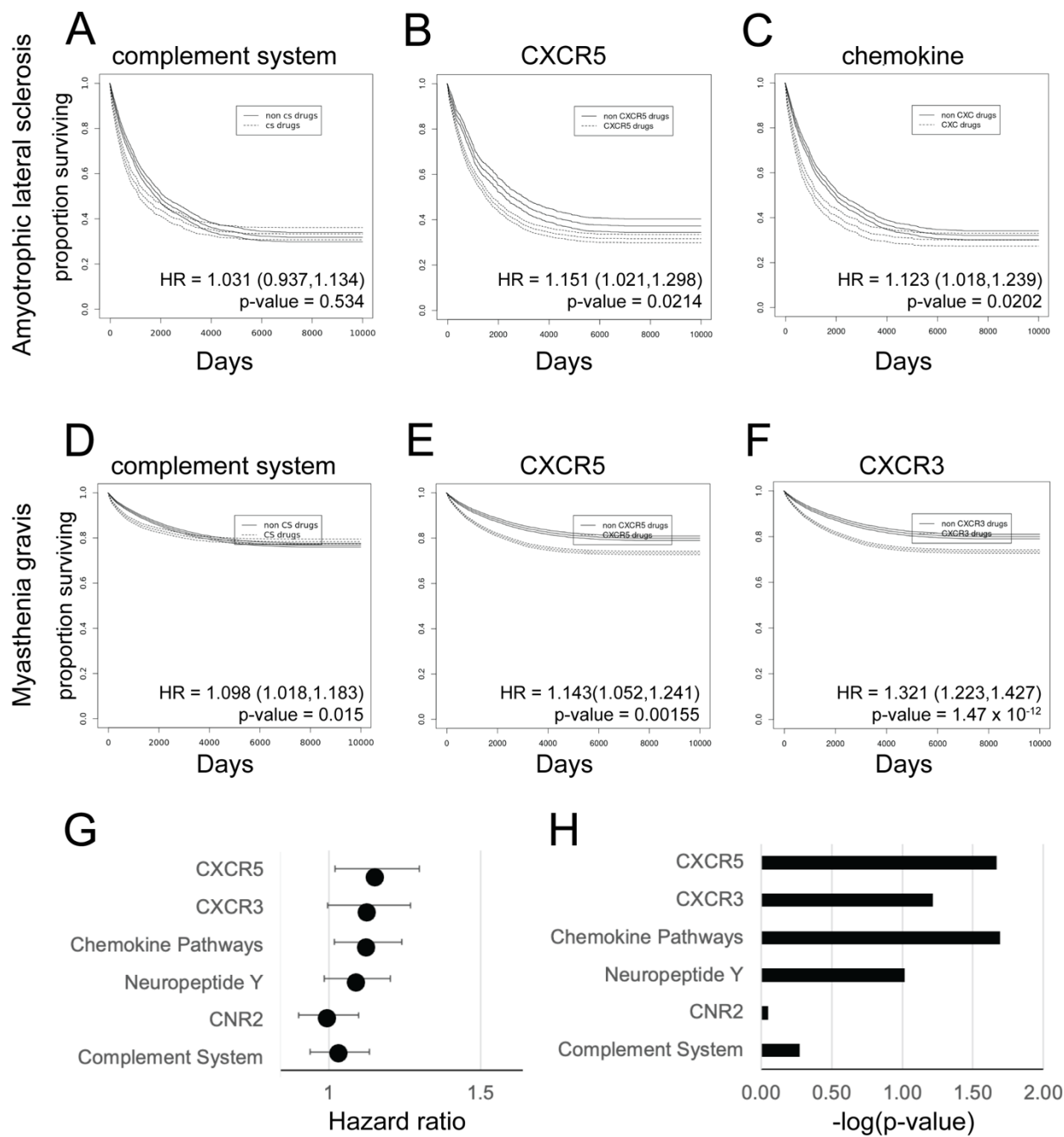
Assess differences in overall survival using new drug classes



Ranked all classes by HR effect size







Pathway Name	ALS - HR	ALS - CI low	ALS - CI high	MG - HR	MG - CI low	MG - CI high	PD - HR	PD - CI low	PD - CI high
CXCR5	1.151	1.021	1.298	1.143	1.052	1.241			
CXCR3	1.124	0.995	1.269	1.321	1.223	1.427			
Chemokine Pathways	1.123	1.018	1.239						
Neuropeptide Y	1.088	0.985	1.203						
CNR2	1.099	1.890	1.097						
Complement System	1.031	0.937	1.134	1.098	1.018	1.183	0.992	0.965	1.020

