Integrated Systems Pharmacology Analysis of Clinical Drug-Induced Peripheral Neuropathy

J Hur1, AY Guo2, WY Loh3, EL Feldman1 and JPF Bai4

Drug-induced peripheral neuropathy (DIPN) is a common but unintended effect of many lifesaving drugs and often leads to dose reduction or regimen modification that may compromise clinical outcomes. Various chemotherapeutic agents and anti-infective drugs are known to cause peripheral neuropathy (PN) during treatment; however, other drugs can also induce PN following prolonged use. For example, the statins, which are widely prescribed for lowering total cholesterol, induce PN following prolonged use. For example, the statins, which are widely prescribed for lowering total cholesterol, have been reported to cause PN with an incidence of 1 in 14,000 person-years of treatment.

DIPN can be broadly divided into three groups based on pathophysiology: (i) axonal degeneration, (ii) segmental demyelination, and (iii) neuron soma damage. Chemotherapeutic agents not only act on their intended therapeutic targets but can also affect various cellular compartments of neurons and their surrounding cells (i.e., glial cells and macrophages), resulting in neurotoxicity and functional changes. For example, bortezomib, paclitaxel, platinum, and vincristine all can cause mitochondrial toxicity, with bortezomib also acting on endoplasmic reticulum, microtubules, and myelin sheath of the dorsal root ganglia.

Genome-wide association studies have improved our understanding of the association between genes and differential clinical manifestations of DIPN. For example, genome-wide association studies have implicated regulatory factor X 2 (RFX2), ephrin type A receptor (EPHA4), and frizzled family receptor 3 (FZD3) in treatment-related PN of taxanes. However, it is generally not well understood how a wide spectrum of drug classes can cause treatment-related PN or how diverse these DIPN drugs are or whether there are any common biological factors connected to these particular compounds. Neither are the mechanisms of DIPN characterized across the drugs approved for treating different diseases.

To define the pharmacological space of DIPN drugs and to enable prediction and assessment of the likelihood of a new drug in causing DIPN, we adopted a hybrid approach of bioinformatics, systems pharmacology, and statistical regression tree analysis to link the clinical incidence and severity that are associated with individual DIPN drugs to their known and extended pharmacological targets. In this reported study, we mined the Drugs@FDA database, supplemented the findings by mining SIDER and NCBI DailyMed, curated the clinical incidence and severity of DIPN reported in published summaries of clinical trials, and then conducted statistical regression tree analyses to associate proteins in the extended pharmacological network with the clinical phenotypes of incidence.

RESULTS
Compilation and curation of the drug list
To identify the DIPN drugs, we referred to the terms listed in Common Terminology Criteria for Adverse Events (v4.03: 14 June 2010) published by the National Cancer Institute of the National Institutes of Health. PN, peripheral sensory neuropathy, and peripheral motor neuropathy were used as search terms. Medical Dictionary for Regulatory Activities (MedDRA; http://www.meddra.com/) terms of hyperesthesia, hypoesthesia, and paresthesia were also referenced. A total of 3,200 current drug label PDF files were downloaded from the Drugs@FDA database and were converted to text files using Xpdf (http://www.tulabs.com/xpdf/). Among a total of 532 drug labels containing neuropathy terms, manual curation confirmed 214 nonredundant DIPN drugs. Similarly, two additional resources, the NCBI DailyMed data and the SIDER2 were surveyed for the aforementioned terms to include 20 additional drugs. A final list of 234 drugs was
obtained (Supplementary Table S1), 183 (78%) of which were common in two or more resources (Figure 1a).

Indication-based drug categorization

DIPN drugs were categorized based on the indication for which each drug was approved by referencing DrugBank and Drugs@FDA. DrugBank included records for 224 out of the 234 drugs, with an average of three classes per drug. Referencing the US Food and Drug Administration (FDA)–approved labels allow the adoption of one indication for each drug. This categorization, though different from pharmacological target-based classification, was chosen to show the relation of DIPN to disease-specific treatment. The five largest categories were antineoplastic (n = 64; 27%), anti-infective (n = 38; 16%), anti-inflammatory (n = 16; 7%), anti-HIV (n = 16; 7%), and antidepressive (n = 10; 5%) (Figure 1b). Those categories with only one or two drugs were merged into the “miscellaneous” category (n = 34; 14%).

Construction of DIPN pharmacological networks

The known intended and unintended targets of DIPN drugs are defined hereafter as targets and were collected from DrugBank and Therapeutic Target Database (TTD; http://bidd.nus.edu.sg/group/cjttd/). Among the 234 DIPN drugs, 204 had at least one target, whereas 30 anti-infective agents only had nonhuman targets in the databases. We then generated a pharmacological network (Figure 2a) using the remaining 174 DIPN drugs and their 280 human targets, referred to as the base network. This network contained multiple subnetworks (Figure 2b), identified by the Fast Greedy community structure analysis algorithm11 based solely on the network topology, and the 10 largest subnetworks are indicated by nongray node colors. The largest subnetwork, highlighted in yellow (Figure 2b), includes 23 drugs (mostly antineoplastic) and 62 targets. The second largest subnetwork consisted of 24 drugs and 35 targets (Figure 2c) and included the drugs that are mainly used for treating neurological disorders such as Parkinson disease, bipolar disorder, unipolar depression and seizures, and the drugs that shared serotonin receptors as common therapeutic targets.

Extension of DIPN pharmacological networks

The base network was extended by referencing the database of BioGRID (http://thebiogrid.org)12 to add intermediators that had either genetic or physical protein–protein interactions with the targets. A total of 2,807 intermediators had interactions with any targets of at least one drug. For each interleaver, drug-degree was defined as the number of interacting drugs that the interleaver indirectly interacted through their known targets. With drug-degree as the cutoff, we generated 20 different extended networks, denoted as Min01 through Min20. The Min01 network included all 2,807 intermediators, whereas the Min20 network included 12 intermediators with at least 20 interacting drugs (Supplementary Table S2). The number of intermediators substantially decreased as the minimum drug-degree threshold increased. As an example, the Min05 extended network including 274 intermediators, ~10% of all intermediators, is depicted in Supplementary Figure S1. Compared with the base pharmacological network, the drugs and their targets in this extended network are highly interconnected via intermediators.

Gene set enrichment analysis on drug targets and intermediators

To identify significantly enriched biological functions among the targets and intermediators, gene set enrichment
Figure 2 Base pharmacological network. (a) The base pharmacological network including 174 drug-induced peripheral neuropathy drugs and 280 human drug targets. The layout of the network was generated using edge-embedded spring layout available in Cytoscape, then was manually adjusted to make sure that all node names are visible. Node color represents the community structures identified by the Greedy algorithm implemented in GLayout, where gray is used as a default color and nongray colors were used for the 10 biggest clusters. (b) The biggest subnetwork with 23 drugs and 62 targets, in yellow. Drugs mostly from the antineoplastic class. (c) The second biggest subnetwork with 24 drugs and 35 targets. Drugs mostly from antidepressive, antiparkinson, and anticonvulsant agents. Shape of notes represents types of entity: circle for drug target and rectangle for drug. The size of node corresponds to the number of interacting drugs.

Pharmacological Network and Clinical Drug Safety

Identification of significant DIPN-associated intermediators
To identify the intermediators significantly associated with DIPN, we generated 1,000 random drug sets from DrugBank and constructed extended networks for each random network. The drug-degree of each intermediary from the DIPN extended networks were compared with those from the extended networks of the 1,000 random networks using a Z-test. Among the total 2,807 intermediators, 230 (Supplementary Table S3) met our significance criteria: having at least five interacting drugs via their targets, a BH-corrected P value < 0.05, and a minimum of 1.5-fold enrichment of drug-degree in the DIPN networks compared with the average drug-degrees of the random networks. The 20 significant intermediators with the highest number of drug-degrees are listed in Table 1, which includes three transcription factors and nine transcription cofactors according to MatBase (Genomatix, Munich, Germany; http://www.genomatix.de/). Tumor protein 53 (TP53) was the most highly connected significant intermediary,
interacting with 49 DIPN drugs in the Min01 network, which is 1.7 times more than the average 37.1 drugs in the 1,000 random networks.

These significant intermediators were further examined for common features using gene set enrichment analysis. The 10 most significantly enriched biological functions in KEGG

![Heatmap of enriched biological functions](image)

**Figure 3** Functional enrichment of drug targets and intermediators. A heat-map was generated using top 10 most enriched biological functions in each gene set (the human target set and 20 sets of intermediators at different level). Benjamini–Hochberg (BH)–corrected $P$ values of the top 10 most significant functional terms are represented in a heat-map with $-\log_{10}$ (BH-corrected $P$ value) as color gradient; the darker the red, the more significant. Any blank cell indicates that the corresponding biological function in the gene set is not statistically significant (BH-corrected $P$ value $\geq 0.05$).
pathway and three GO categories are listed in Table 2. The ubiquitin-mediated proteolysis pathway was the most significantly enriched pathway, comprising 10% of the significant intermediators. Apoptosis-related terms were also highly enriched.

Comparison of known targets and intermediators with the genes implicated in neuropathy

We compiled 72 human genes whose mutations result in neurological disorders with neuropathy,15 and 36 genes for which single-nucleotide polymorphisms or differential expression was reportedly to be associated with chemotherapy-related neuropathy.4,6,16–20 When these two lists of genes were compared with the known targets and significant intermediators, little overlap was observed (Supplementary Table S4), suggesting that the DIPN pharmacological networks were not directly linked to the genes that are predisposed to neuropathy.

Integration of clinical information with pharmacological network

We extracted the clinical data (neuropathy incidence, grade, and the size of clinical trials) for 90 DIPN drugs from drug labels and published new drug application reviews in Drugs@FDA, and 43 drugs from published clinical studies, resulting in a nonredundant set of 97 drugs. Over 30 drugs included data from multiple clinical trials. The trial with the highest incidence was chosen first followed by highest severity if multiple trials of similar sizes were found for a drug. The majority of these 97 drugs had only the incidence of all grades (grade 0–4 combined), whereas only 20 drugs had a detailed incidence profile of grade 1–2 and/or grade 3–4 (Table 3).

Regression tree analysis was applied to the 97 drugs, of which treatment-related neuropathy incidences were reported, with their neuropathy incidence as the dependent variable. Multiple sets of factors, including intermediators from the pharmacological network (minimum level of 5), known molecular targets, and the drug categories and substructures obtained from DrugBank, were used as explanatory variables. Regression tree models for neuropathy incidence were obtained for all grades. A model of potentially important intermediators is shown in Figure 4a. Among the 17 drugs interacting with v-myc avian myelocytomatisis viral oncogene homolog (MYC), 8 drugs interacting also with PAF15 had an average neuropathy incidence of 38.1%, whereas 9 drugs not interacting with proliferating cell nuclear antigen-associated factor (PAF15) had an average incidence of 2.9% (Figure 4b; see Supplementary Table S5 for additional clinical and demographic data). Figure 4c illustrates a subnetwork centered on MYC, proliferating cell nuclear antigen-associated factor (PAF15), and all of their interacting drugs in the base DIPN pharmacological network. The subnetwork in the middle consists of 14 drugs and 11 drug targets interacting with both MYC and PAF15. Among these 14 drugs, 6 drugs highlighted in yellow had no known neuropathy incidence available from clinical trial data.

We also performed linear regression analyses to determine the effect of drug targets and intermediators on neuropathy incidence. Using univariate analysis, three targets and 18 intermediators were identified with false discovery rate < 0.05 as the cutoff (Supplementary Table S6).

### Table 1: Top 20 highly connected significant intermediators in DIPN

| Gene ID | Symbol | Description | No. of interacting drugs (drug-degree) | Enrichment fold | TF role |
|---------|--------|-------------|----------------------------------------|-----------------|---------|
| 7157    | TP53   | Tumor protein p53 | 49 31.8 3.45E-03 1.6 TF | --- | --- |
| 6613    | SUMO2  | Small ubiquitin-like modifier 2 | 46 25.9 1.73E-04 1.8 --- | --- | --- |
| 10987   | COP5   | COP9 signalosome subunit 5 | 44 29.1 1.05E-02 1.5 Cofactor | --- | --- |
| 7341    | SUMO1  | Small ubiquitin-like modifier 1 | 42 27.1 7.33E-03 1.5 --- | --- | --- |
| 4738    | NEDD8  | Neural precursor cell expressed, developmentally downregulated 8 | 34 11.1 4.75E-11 3.1 Cofactor | --- | --- |
| 4609    | MYC    | V-myc avian myelocytomatisis viral oncogene homolog | 33 16.3 3.40E-02 2.0 TF | --- | --- |
| 5371    | PML    | Promyelocytic leukemia | 32 15.0 4.96E-05 2.1 Cofactor | --- | --- |
| 867     | CBL    | Cbl proto-oncogene, E3 ubiquitin protein ligase | 32 18.3 2.67E-03 1.7 --- | --- | --- |
| 672     | BRCA1  | Breast cancer 1, early onset | 32 19.2 6.43E-03 1.7 Cofactor | --- | --- |
| 5925    | RB1    | Retinoblastoma 1 | 31 10.2 9.80E-03 3.0 Cofactor | --- | --- |
| 3164    | NRA1   | Nuclear receptor subfamily 4, group A, member 1 | 28 11.4 1.15E-05 2.5 TF | --- | --- |
| 387082  | SUMO4  | Small ubiquitin-like modifier 4 | 28 13.5 2.17E-04 2.1 --- | --- | --- |
| 142     | PARP1  | Poly (ADP-ribose) polymerase 1 | 27 14.1 1.98E-03 1.9 Cofactor | --- | --- |
| 3717    | JAK2   | Janus kinase 2 | 27 16.4 1.55E-02 1.6 --- | --- | --- |
| 6605    | SMARCE1| SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily e, member 1 | 26 17.2 4.14E-02 1.5 Cofactor | --- | --- |
| 8065    | CUL5   | Culin 5 | 25 10.0 1.46E-05 2.5 --- | --- | --- |
| 6421    | SFPO   | Splicing factor proline/glutamine-rich | 25 12.4 8.48E-04 2.0 Cofactor | --- | --- |
| 8453    | CUL2   | Culin 2 | 25 14.9 1.95E-02 1.7 --- | --- | --- |
| 9219    | MTA2   | Metastasis associated 1 family, member 2 | 25 15.0 2.08E-02 1.7 Cofactor | --- | --- |
| 6595    | SMARCA2| SWI/SNF-related, matrix-associated, actin-dependent regulator of chromatin, subfamily a, member 2 | 25 16.6 3.91E-02 1.5 Cofactor | --- | --- |

DIPN, drug-induced peripheral neuropathy; FDR, false discovery rate; enrichment fold, the ratio of drug-degree between DIPN network and 1,000 random networks (drug-degree\_DIPN/drug-degree\_1000 networks); TF, transcription factor.
Based on this result, we performed three multivariate analyses: (i) using all 21 significant factors; (ii) using 12 factors, each with minimum three incidence data points; and (iii) using 7 factors, each with minimum five incidence data points. One drug target (tubulin beta-3 chain (TUBB3)) and three intermediators (protein phosphatase 1 alpha catalytic subunit (PPP1CA), proteasome 26s subunit ATPase 2 (PSMC2), and F-box protein 25 (FBXO25)) were found to be significant.

**DISCUSSION**
To better understand the underlying mechanisms of DIPN and to facilitate future development of drugs with a minimal potential of inducing DIPN, we conducted a pharmacological network-based analysis. We first compiled the list of DIPN drugs and defined the pharmacological space of DIPN as broadly as possible. A substantial number of drugs with PN identified as an adverse reaction from the postmarketing surveillance were also included, as long as such information is available.
| Drug name       | No. of subjects in clinical trial | Neuropathy incidence | Source          |
|-----------------|----------------------------------|----------------------|-----------------|
| Aldesleukin     | 525                              | <1.0%                | -- Drugs@FDA    |
| Almotriptan     | 1,313                            | 1.0%                 | -- Drugs@FDA    |
| Amiodarone      | 241                              | 3.7%                 | -- Drugs@FDA    |
| Amlodipine      | 1,730                            | 0.1–1%               | -- Drugs@FDA    |
| Anastrozole     | 60                               | 0.1%                 | -- Drugs@FDA    |
| Aprepitant      | 261                              | >0.5%                | -- Drugs@FDA    |
| Atorvastatin    | 5158                             | 0.2%                 | -- Drugs@FDA    |
| Betaxolol       | 509                              | <2.0%                | -- Drugs@FDA    |
| Bexarotene      | 156                              | <10.0%               | -- Drugs@FDA    |
| Bicalutamide    | 401                              | 8.0%                 | -- Drugs@FDA    |
| Bortezomib      | 1,163                            | 39.0% 27.0% 12.0%   | -- Drugs@FDA    |
| Cabazitaxel     | 371                              | 13.0% 12.0% 1.0%    | -- Drugs@FDA    |
| Capcetinibine   | 596                              | 10.0%                | -- Drugs@FDA    |
| Carboplatin     | 1,893                            | 4.0%                 | -- Drugs@FDA    |
| Carfilzomib     | 526                              | 14.0%                | <2% Drugs@FDA   |
| Cevimeline      | 533                              | <1.0%                | -- Drugs@FDA    |
| Ciprofloxacin   | 1,085                            | 0.1%                 | -- Drugs@FDA    |
| Clasilatin      | 79                               | 38.0%                | PMID:2541260    |
| Clomipramine    | 3,525                            | 0.1%                 | -- Drugs@FDA    |
| Crizotinib      | 255                              | 23.0%                | <1% Drugs@FDA   |
| Cyclophosphamide| 7,807                            | <1.0%                | -- Drugs@FDA    |
| Cyclosporine    | 705                              | 1.0%                 | -- Drugs@FDA    |
| Cytarabine      | 257                              | 4.0%                 | -- Drugs@FDA    |
| Dapptomycin     | 989                              | 0.7%                 | -- Drugs@FDA    |
| Dasatinib       | 911                              | 13.0% 12.0% 1.0%    | -- Drugs@FDA    |
| Dexamethasone   | 332                              | 9.0%                 | 1.0% PMID:2641260|
| Didanosine      | 395                              | 19.0%                | -- Drugs@FDA    |
| Diltiazem       | 800                              | 1.1%                 | -- Drugs@FDA    |
| Docetaxel       | 2,045                            | 49.0%                | 4.0% Drugs@FDA  |
| Doxorubicin     | 163                              | 6.0%                 | 0.0% PMID:10561296|
| Eletriptan      | 4,597                            | 3.1%                 | -- Drugs@FDA    |
| Enalapril       | 2,341                            | 0.5–1%               | -- Drugs@FDA    |
| Eribulin        | 503                              | 35.0%                | 8.0% Drugs@FDA  |
| Eszopiclone     | 287                              | 0.1–1%               | -- Drugs@FDA    |
| Etoposide       | 2,081                            | 1–2%                 | -- Drugs@FDA    |
| Etravirine      | 599                              | 1.7%                 | -- Drugs@FDA    |
| Exemestane      | 2,252                            | 6.0%                 | -- Drugs@FDA    |
| Fludarabine     | 101                              | 4.0%                 | -- Drugs@FDA    |
| Fluoxetine      | 1,145                            | 0.1%                 | -- Drugs@FDA    |
| Fosaprepitant   | 438                              | 0.5–3%               | -- Drugs@FDA    |
| Foscarnet       | 189                              | 5.0%                 | -- Drugs@FDA    |
| Ganciclovir     | 326                              | 8.0%                 | -- Drugs@FDA    |
| Gatiloxacin     | 5,000                            | 0.1–3%               | -- Drugs@FDA    |
| Ifosfamide      | 1,317                            | 0.4%                 | -- Drugs@FDA    |
| Imatinib        | 337                              | 5.9%                 | 0.0% Drugs@FDA  |
| Indomethacin    | 175                              | <1.0%                | -- Drugs@FDA    |
| Ipromide        | 1,142                            | <1.0%                | -- Drugs@FDA    |
| Itraconazole    | 107                              | 17.0%                | -- PMID:21685202|
| Ixabepiline     | 126                              | 70.0%                | 14.0% PMID:19171553|

DIPN, drug-induced peripheral neuropathy; PMID, PubMed unique identifier.
in their FDA-approved labels. The incidence for these drugs is typically less than 0.1% (Table 3). The DIPN causality for those identified from postmarketing surveillance cannot be established, nor can those from epidemiology studies. On the contrary, the identification of PN as a treatment-related problem in clinical trials is considered to have been judged by physicians in charge of the trials. In light of our goal to project an inclusive pharmacological network, it is not considered crucial whether the causality is established for each of the 234 drugs. Furthermore, the pharmacological networks in the current study do not distinguish severity and incidence. As a result, the DIPN list covered the drugs with a wide range of pharmacological characteristics. Because the 234 drugs were approved for treating various diseases (Figure 1), their pharmacological and extended pharmacological networks could conceivably highlight, at least, part of the network connecting those diseases. On the other hand, these drugs do share a common characteristic of clinical DIPN, and one could argue that these pharmacological networks would biologically reflect the DIPN network and, therefore, would be useful for understanding DIPN.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN.21 In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence. Because the 234 drugs were approved for treating various diseases (Figure 1), their pharmacological and extended pharmacological networks could conceivably highlight, at least, part of the network connecting those diseases. On the other hand, these drugs do share a common characteristic of clinical DIPN, and one could argue that these pharmacological networks would biologically reflect the DIPN network and, therefore, would be useful for understanding DIPN.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence. Because the 234 drugs were approved for treating various diseases (Figure 1), their pharmacological and extended pharmacological networks could conceivably highlight, at least, part of the network connecting those diseases.

As a result, the DIPN list covered the drugs with a wide range of pharmacological characteristics. Because the 234 drugs were approved for treating various diseases (Figure 1), their pharmacological and extended pharmacological networks could conceivably highlight, at least, part of the network connecting those diseases. On the other hand, these drugs do share a common characteristic of clinical DIPN, and one could argue that these pharmacological networks would biologically reflect the DIPN network and, therefore, would be useful for understanding DIPN.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence.

As shown in Table 3, clinical DIPN incidence varied from drug to drug. It should be noted that the underlying diseases could have affected the DIPN incidence; for instance, multiple myeloma itself can cause PN. In the present study, the impact of disease was, however, not taken into account due to the lack of a comprehensive understanding of how individual diseases at different stages had affected the incidence and severity of PN in individual clinical trials. When trials were conducted for a chemotherapeutic agent as compared with a “standard-of-care” treatment, we only adopted those from the trials that had a head-to-head comparison between two single-agent regimens. When multiple trial results for a drug for either the same indication or different indications were available, we adopted a subjective criterion of adopting the result from the trial with the highest incidence first followed by the trial with the second highest incidence. When neither the first nor the second highest incidences were available, the incidence was assigned to the trials with available incidence. When the incidence was available only in postmarketing surveillance, it was assigned as the corresponding drug’s clinical incidence. Postmarketing surveillance can be widely used in the guidance of drug development by regulatory agencies, and studies that show the clinical outcomes of drug development can be more reliable than those of postmarketing surveillance. In the present study, all the incidences from postmarketing surveillance were used in the calculations of clinical DIPN incidence.
by highest severity if multiple trials of similar sizes were found for a drug. As a result, we may have inflated the severity and/or incidence of DIPN for some drugs. Our goal was not to be exclusive by imposing strict rules but to employ consistent rules for adopting the published results.

To identify the pharmacological network with a broad coverage of biological pathways for future development of a useful predictive tool, we extended the base pharmacological network to include intermediators that interacted with the known targets of DIPN drugs. As shown in Figure 3, the most enriched biological functions of these intermediators, very similar across different sets of intermediators, included “transcription cofactor binding,” “cell cycle,” and “regulation of apoptosis.” We speculated that this was likely due to the fact that 27% of the DIPN drugs were antineoplastic agents; however, the significant intermediators, with the antineoplastic or anti-infective drugs excluded, were still highly enriched with these functions (Supplementary Figure S2). It was concluded that these enriched functions were not limited to antineoplastic or anti-infective agents. There seemed to be a distinct pattern in the enriched biological functions between the drug–target group and the various groups of intermediators (Figure 3). Drug targets are mainly receptors or specific kinases, whereas intermediators are cellular proteins that interact with drug targets. It seems reasonable that the spectrum of biological functions covered by intermediators would be different, including regulation and transcription.

Though all the 97 drugs listed in Table 3 were used for our regression tree analyses, many classes of drugs were pruned off the regression tree by the Generalized, Unbiased Interaction Detection and Estimation tree algorithm due to low predictive power. As a result, higher clinical DIPN incidences of antineoplastic drugs as compared with other classes of drugs could have led to the dominance of these drugs identified in regression tree modeling. Our regression tree modeling of neuropathy incidence suggests that MYC and PAF15 may play an important role in DIPN (Figure 4).

Interacting with either MYC or PAF15 alone was associated with relatively low neuropathy incidence of 2.94% or 2.90%, respectively, whereas interaction with both genes had a high rate of 38.14%. The drugs linked to MYC were dominated by floxacins, and then by tyrosine kinase inhibitors, whereas those linked to PAF15 seemed to include drugs from diverse categories. Those drugs linked to both MYC and PAF15 were mostly nucleoside analogues or metabolic inhibitors, followed by taxanes. When we examined the 10 drugs with the highest neuropathy incidences, 4 (paclitaxel, docetaxel, bortezomib, and vinorelbine) were connected to both MYC and PAF15 (Table 3 and Figure 4c), suggesting that these two genes might contribute to DIPN in some drugs. The remaining six drugs with top 10 incidences included four antineoplastic drugs (oxaliplatin, ixabepilone, eribulin, and pemetrexed), one anti-infective drug (metronidazole), and one anti-HIV drug ( stavudine). In the pharmacological networks, four of these drugs (metronidazole, oxaliplatin, eribulin, and stavudine) did not have any known human targets; thus, unknown off-targets may play a role. It should be noted that our computational exercise is limited by what pharmacological targets and intermediators are available in the databases. Eribulin with the eighth highest rate (36%), for example, is known to directly bind tubulin, thus disrupting microtubule structures, which was not properly annotated in neither DrugBank nor TTD.

To demonstrate the usefulness of our theoretical exercises, we further explored the pharmacological networks focusing two antineoplastic agents. Both bortezomib and sorafenib are widely used anticancer drugs, where the former is a proteasome inhibitor and the latter is a tyrosine kinase inhibitor. They both interact with MYC through their targets: RARF1 (v-raf-1 murine leukemia viral oncogene homolog 1) for sorafenib and PSMB5 (proteasome subunit beta type-5) or PSMD2 (26S proteasome non-ATPase regulatory subunit 2) for bortezomib (Figure 4c). Interestingly, they had different incidences of DIPN: 39% for bortezomib and 6.6% for sorafenib. In the extended subnetwork centered on bortezomib and sorafenib, we observed that only the targets of bortezomib interacted with tubulins (Supplementary Figure S3). This finding is in line with other studies that related perturbation of tubulin or microtubules in neural cell lines and in rat dorsal root ganglion neurons to the neurotoxicity of bortezomib. In fact, five drugs (ixabepilone, paclitaxel, docetaxel, bortezomib, and vinorelbine) belonging to the top 10 list were connected to tubulins (Table 3 and Supplementary Figure S4). These results, though not new, are supported by other studies demonstrating that dysregulated microtubule polymerization by microtubule-targeting agents such as taxanes are strongly related to DIPN. Furthermore, these findings confirmed the quality of the databases and knowledge bases that were used for this study. The three intermediators (PPP1CA, PSMC2, and FBXO25) identified in the multivariate linear regression analysis are preliminary but worthy of further investigation.

The current study demonstrates the usefulness of leveraging the knowledge of the FDA-approved drugs and accumulated knowledge generated from pharmacological as well as medical research for constructing the extended pharmacological network of DIPN. Such extended pharmacological network can potentially be useful for predicting the neurotoxicity potential of new drugs. Based on the modeling result, the usefulness of targeting the genes regulated by MYC and PAF15 could be worthy of further investigation for the discovery of potential drugs for alleviating the symptoms of DIPN. Our next step will be to refine our network analysis by incorporating high-throughput omics data such as gene expression profiles of DIPN drugs in neuronal cells. Our ultimate goal is to develop our theoretical analysis into a predictive tool of DIPN with an acceptable specificity and sensitivity, based on the pharmacological networks and multiplex measurements including transcriptomic profiling. Such tool will be valuable for minimizing the severity and incidence of DIPN.

METHODS
Compilation and curation of DIPN drugs

Text-mining and curation of DIPN drugs

Pharmacological Network and Clinical Drug Safety

Hur et al.
The flanking texts surrounding the identified neuropathy terms in the drug labels were manually examined for their accuracy, and a nonredundant set of DIPN drugs was compiled. Two other drug resources, DailyMed (http://dailymed.nlm.nih.gov/) and SIDER2 (http://sideeffects.embl.de/), were to identify additional DIPN drugs. The drugs collected only from DailyMed or SIDER2 were subjected to confirmation for inclusion by referencing the literature (PubMed).

Drug name normalization, integration, and manual curation. The drug names collected were normalized to be consistent and comparable. The official names in the DrugBank database were used as the representative names, if available. Drugs without a DrugBank record were processed by a Perl script to automatically identify the core ingredient name. For example, “sumatriptan succinate” was normalized to “sumatriptan.” The lists of drugs collected above were normalized and combined and then thoroughly reviewed by a pharmacology expert to include only the single active ingredients with evidence of DIPN.

Pharmacological network analysis

Construction of pharmacological network. A base pharmacological network consisting of DIPN drugs and their known intended and unintended targets was constructed by referencing the DrugBank version 3.0 (downloaded on 21 October 2012; http://www.drugbank.ca/) and Therapeutic Target Database version 4.3.02 released on 25 August 2011 (TTD; http://bidd.nus.edu.sg/group/cjttd/). This base network was extended by adding the proteins or genes, denoted as intermediators, that have protein–protein interactions or genetic interactions with the known targets of DIPN drugs. The protein–protein interactions and genetic interactions were derived from the BioGRID release 3.2.97 (http://thebiogrid.org/). Drug-degree was defined for each intermediary as the number of drugs indirectly interacting via the known targets. Twenty extended networks were generated using the criterion of drug-degrees ranging from 1 to 20 as the threshold for intermediators to be included in a network.

Functional enrichment of the pharmacological networks. To identify the enriched biological functions of the known targets of DIPN drugs and their intermediators, gene set enrichment analysis was employed using a locally implemented version of DAVID. GO terms and KEGG pathway terms were used as biological functions. Terms with a BH-corrected P value < 0.05 were deemed significant. A heat-map with −log10 (BH-corrected P) was used as biological functions. Terms with a BH-corrected P < 0.05 were deemed significant intermediators: (i) BH-corrected P value < 0.05 in the Min01 network, (ii) minimum drug-degree of 5, and (iii) the drug-degree in DIPN network is at least 1.5 times higher than that in random networks. These DIPN-specific intermediators were subjected to further functional enrichment analysis.

Integration of clinical information with pharmacological networks

Clinical information from drug labels and literature. The data of PN incidence, severity, and trial size from clinical trials were collected from drug labels, published new drug application reviews in Drugs@FDA, and published literature in PubMed. Following automatic text-mining, manual curation extracted the clinical phenotypes of incidence and severity of DIPN for individual DIPN drugs from Drugs@FDA. PubMed was also surveyed to compile relevant clinical information of DIPN drugs. Over 21 million abstracts in PubMed were surveyed using a text-mining script to identify publications containing any of DIPN drugs and keywords, “clinical trial” and “neuropathy.” Then, manual curation was performed to extract relevant clinical information. In the case of drugs with multiple clinical trial data available, the data from the source with the largest trial size first, and the highest incidence followed by the highest severity if the trial sizes were similar, were included for further analyses.

Modeling using regression tree algorithm. The collected neuropathy incidence data were used to build a statistical model with the Generalized, Unbiased, Interaction Detection and Estimation algorithm, a multipurpose machine-learning algorithm for constructing regression trees. Generalized, Unbiased, Interaction Detection and Estimation algorithm uses recursive partitioning to select the splits of the tree nodes. After an initial large tree is constructed, the algorithm uses 10-fold cross-validation to prune the nodes and estimate prediction accuracy. Only those models with estimated prediction accuracy within half a standard error of the minimum are reported.

Univariate and multivariate linear regression analysis. Linear regression analyses were performed to examine the effect of drug targets and intermediators on neuropathy incidence. Significant factors, defined at false discovery rate < 0.05 in univariate regression, were further examined in a multivariate regression analysis.

Acknowledgments. The authors thank Dr. Darrell Abernethy at the US Food and Drug Administration for his support and comments on the manuscript. The authors also thank Dr. Phillippe O’Brien at the University of Michigan for his editorial support. This work was supported by the US Food and Drug Administration, the Program for Neurology Research and Discovery at the University of Michigan, the A. Alfred Taubman Medical Research Institute, and a Juvenile Diabetes Research Foundation Postdoctoral Fellowship (to J.H.). W.-Y.L.'s work was partially supported by the US Army Research Office grant W911NF-09-1-0205 and National Science Foundation grant DMS-1305725.
The views expressed in this article are those of the authors and do not necessarily reflect the official views of the US Food and Drug Administration.

Author Contributions. J.H. wrote the manuscript. J.P.F.B. and E.L.F. designed the research. J.H., J.P.F.B., A.Y.G., and W.-Y.L. performed the research. J.H., A.Y.G., and W.-Y.L. analyzed the data.

Conflict of interest. The authors declared no conflict of interest.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?
Many antineoplastic and anti-inflammatory agents cause DIPN, resulting in dose reduction or treatment modification that may compromise clinical outcomes.

WHAT QUESTION DID THIS STUDY ADDRESS?
This study attempts to define the pharmacological space of DIPN, to identify the commonly associated genes and proteins, and to build statistical models based on clinical incidence and severity using an integrated approach of bioinformatics, systems pharmacology, and statistical regression analysis.

WHAT THIS STUDY ADDS TO OUR KNOWLEDGE
This study expands the current list of DIPN drugs, typically limited to antineoplastic and anti-inflammatory drugs, to include the drugs for treating many other diseases. This study suggests that two high-level gene regulators, MYC and PAF15, might be involved in DIPN.

HOW THIS MIGHT CHANGE CLINICAL PHARMACOLOGY AND THERAPEUTICS
The integrated pharmacological networks of DIPN and the statistical model may be informative for future development of drugs with a potential of inducing peripheral neuropathy.

1. Paice, J.A. Clinical challenges: chemotherapy-induced peripheral neuropathy. Semin. Oncol. Nurs. 23, S8–S19 (2009).
2. Gaist, D., García Rodríguez, L.A., Huerta, C., Hallas, J., & Sindrup, S.H. Are users of lipid-lowering drugs at increased risk of peripheral neuropathy? Eur. J. Clin. Pharmacol. 56, 931–933 (2001).
3. Ferrier, J., Pereira, V., Busserolles, J., Authier, N., & Balayssac, D. Emerging trends in understanding chemotherapy-induced peripheral neuropathy. Curr. Pain Headache Rep. 17, 364 (2013).
4. Wheeler, H.E. et al.; Cancer and Leukemia Group B. Integration of cell line and clinical trial genome-wide analyses supports a polygenic architecture of Paclitaxel-induced sensory peripheral neuropathy. Clin. Cancer Res. 19, 491–499 (2013).
5. Leandro-Garcia, L.J. et al. Genome-wide association study identifies ephrin type A receptors implicated in paclitaxel-induced peripheral sensory neuropathy. J. Med. Genet. 50, 599–605 (2013).
6. Baldwin, R.M. et al. A genome-wide association study identifies novel loci for paclitaxel-induced sensory peripheral neuropathy in CALGB 40101. Clin. Cancer Res. 18, 5099–5109 (2012).
7. National Cancer Institute. Common Terminology Criteria for Adverse Events v4.03 (CTCAE). <http://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf>
8. Drugs@FDA. <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/). Accessed 25 November 2012.
9. The Gene Ontology. <http://www.geneontology.org/).
10. Kuhn, M., Campillos, M., Letunic, I., Jensen, L.J., & Bork, P. A side effect resource to capture phenotypic effects of drugs. Mol. Syst. Biol. 6, 343 (2010).
11. Su, G., Kuchinsky, A., Morris, J.H., States, D.J., & Meng, F. GLay: community structure analysis of biological networks. Bioinformatics 25, 3135–3137 (2010).
12. Chatr-Aryamontri, A. The BioGRID interaction database: 2013 update. Nucleic Acids Res. 41, D816–D823 (2013).
13. Kuhn, M., Campillos, M., Letunic, I., & Jensen, L.J. A side effect resource to capture phenotypic effects of drugs. Mol. Syst. Biol. 6, 343 (2010).
14. Huang da, W., Sherman, B.T., & Lempicki, R.A. Bioinformatics enrichment tools: paths toward understanding chemotherapy-induced peripheral neuropathy. Clin. Cancer Res 19, 599–605 (2013).
15. NIH GeneReviews book. <http://www.ncbi.nlm.nih.gov/books/NBK1116/>. Accessed 25 November 2012.
16. Leandro-Garcia, L.J. et al. Genome-wide association study identifies ephrin type A receptors implicated in paclitaxel-induced peripheral sensory neuropathy. J. Med. Genet. 50, 599–605 (2013).
17. Ogun, T. et al. Genetic polymorphisms associated with oxaliplatin-induced peripheral neurotoxicity in Japanese patients with colorectal cancer. Int. J. Clin. Pharmacol. Ther. 51, 475–481 (2013).
18. Leandro-Garcia, L.J. et al. Regulatory polymorphisms in J−tubulin IIIa are associated with paclitaxel-induced peripheral neuropathy. Clin. Cancer Res. 18, 4441–4448 (2012).
19. Won, H.K. et al. Polymorphic markers associated with severe oxaliplatin-induced, chronic peripheral neuropathy in colon cancer patients. Cancer 118, 2829–2836 (2012).
20. Brojoi, A. et al. Mechanisms of peripheral neuropathy associated with bortezomib and vinristine in patients with newly diagnosed multiple myeloma: a prospective analysis of data from the HOVON-65/GMMG-HD4 trial. Lancet Oncol. 11, 1057–1065 (2010).
21. Richardson, F.G. et al. Management of treatment-emergent peripheral neuropathy in multiple myeloma, Leukemia 26, 595–608 (2012).
22. LaPointe, N.E., Morfini, G., Brady, S.T., Feinstein, S.C., Wilson, L. & Jordan, M.A. Effects of eribulin, vincristine, paclitaxel and ixabepilone on fast axonal transport and kinesin-1 driven microtubule gliding: implications for chemotherapy-induced peripheral neuropathy. Neurotoxicology 37, 231–240 (2013).
23. Smith, J.A. et al. Eribulin binds at microtubule ends to suppress dynamic instability. Biochem. Biophys. Res. Commun. 41, 1331–1337 (2010).
24. Poruchynsky, M.S., Sackett, D.L., Robey, R.W., Ward, Y., Annunziata, C. & Foy, T. Proteasome inhibitors increase tubulin polymerization and stabilization in tissue culture cells: a possible mechanism contributing to peripheral neuropathy and cellular toxicity following proteasome inhibition. Cell Cycle 9, 940–949 (2008).
25. Staff, N.P. et al. Bortezomib alters microtubule polymerization and axonal transport in rat dorsal root ganglion neurons. Neurotoxicology 39, 124–131 (2013).
26. Carlson, K. & Oancea, A.J. Peripheral neuropathy with microtubule-targeting agents: occurrence and management approach. Clin. Breast Cancer 11, 73–81 (2011).
27. Knox, C. et al. DrugBank 3.0: a comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1044 (2011).
28. Zhu, F. et al. Therapeutic target database update 2012: a resource for facilitating target-oriented drug discovery. Nucleic Acids Res. 40, D1128–D1136 (2012).
29. Loh, W.Y. Regression trees with unbiased variable selection and interaction detection. Stat. Sinica 12, 361–386 (2002).
30. Loh, W.Y. Variable selection for classification and regression in large p, small n problems. In Probability Approximations and Beyond (eds Barbour, A et al.)35–159 (Springer, New York, 2012).

Supplementary information accompanies this paper on the CPT: Pharmacometrics & Systems Pharmacology website (http://www.nature.com/psp)