



**HAL**  
open science

## **Prediction of Drug Approval After Phase I Clinical Trials in Oncology -RESOLVED2**

Guillaume Beinse, Virgile Tellier, Valentin Charvet, Eric Deutsch, Isabelle Borget, Christophe Massard, Antoine Hollebecque, Loic Verlingue

► **To cite this version:**

Guillaume Beinse, Virgile Tellier, Valentin Charvet, Eric Deutsch, Isabelle Borget, et al.. Prediction of Drug Approval After Phase I Clinical Trials in Oncology -RESOLVED2. JCO Clinical Cancer Informatics, 2019, 3, pp.1-10. 10.1200/CCI.19.00023 . hal-02894327

**HAL Id: hal-02894327**

**<https://hal.science/hal-02894327>**

Submitted on 8 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Title page

**Article type:** original article

### Prediction of Drug Approval After Phase I Clinical Trials in Oncology - RESOLVED2

Prediction of Drug Approval After Phase I in Oncology

#### Authors list:

Guillaume BEINSE, MD<sup>1</sup>; Virgile TELLIER<sup>1</sup>; Valentin CHARVET<sup>1</sup>; Prof. Eric DEUTSCH, MD PhD<sup>2,3,4</sup>; Prof. Isabelle BORGET, PharmD PhD<sup>5,6,7</sup>; Christophe MASSARD, MD PhD<sup>1</sup>; Antoine HOLLEBECQUE, MD PhD<sup>1</sup>; Loic VERLINGUE, MD<sup>1</sup>.

#### Affiliations:

- 1 - Department of Drug Development in Oncology DITEP, Gustave Roussy Cancer Campus, Villejuif, France;
- 2 - Department of Radiation Oncology, Gustave Roussy Cancer Campus, Villejuif, France;
- 3 - Université Paris-Saclay, Le Kremlin-Bicêtre, France;
- 4 - INSERM U1030, Molecular Radiotherapy, Université Paris-Saclay, Villejuif, France;
- 5 - Department of Biostatistics and Epidemiology, Gustave Roussy Cancer Campus, Villejuif, France;
- 6 - INSERM U1018, CESP, Université Versailles-Saint-Quentin-en-Yvelines, Villejuif, France;
- 7 - GRADES, Université Paris-Sud, France.

#### Corresponding author:

Christophe Massard, MD, PhD 114 Rue Edouard Vaillant, 94800 Villejuif, France.  
Phone: +33 142 114 385; email: [christophe.massard@gustaveroussy.fr](mailto:christophe.massard@gustaveroussy.fr)

#### Key words (MeSH): (/6)

Machine Learning; Clinical Trials, Phase I; Antineoplastic Agents; United States Food and Drug Administration; Databases, Pharmaceutical

**Research support for the study:** none. Training fees (VT, VC) were covered by the DITEP department.

Part of this study will be presented (poster communication) at the AACR 2019 Congress.

**Disclaimers:** none.

#### Key Points

**\*Key objective:** we aimed to evaluate the feasibility and utility of a machine-learning recommender system to predict drug development outcome in Oncology and therefore to support early go/no-go decision as soon as phase I trials completion.

**\*Knowledge generated:** RESOLVED2 is a Lasso-penalized Cox regression model. To train RESOLVED2, we developed a new metric, namely Food and Drug Administration approval free survival (FDA-aFS), defined by the time between publication of the first early clinical trial (ECT) reporting clinical effect of a drug, and FDA approval, censored by date of last news. From simple pharmacological data and ECT's PubMed abstract, RESOLVED2 can accurately predict the time to FDA approval of a new antineoplastic agent.

**\*Relevance:** our work demonstrates machine learning approaches can enhance drug development in Oncology by supporting early go/no-go decisions.

## **Abstract**

### **Purpose**

Drug development in Oncology is currently facing a conjunction of increasing number of anti-neoplastic agents (ANA) candidate for phase I clinical trials (P1CT) and an important attrition rate for final approval. We aimed to develop a machine learning algorithm (RESOLVED2) predicting drug development outcome, which could support early go/no-go decisions after P1CT, by better selection of drugs suitable for further development.

### **Methods**

PubMed abstracts of P1CT reporting ANA were used together with pharmacological data from DrugBank5.0 database to model time to US Food & Drug Administration (FDA) approval (FDA approval-free-survival, FDA-aFS) since the first P1CT publication. RESOLVED2 model was trained with machine learning methods. Its performances were evaluated on an independent test set with weighted concordance index (IPCW).

### **Results**

We identified 462 ANA from PubMed matching with DrugBank5.0 database (P1CT publication dates: 1972-2017). Among 1411 variables, 28 were used by RESOLVED2 to model the FDA-aFS with an IPCW of 0.89 on the independent test set. RESOLVED2 outperformed a model based on efficacy/toxicity (IPCW=0.69). In the test set, at 6 years of follow-up, 73% (95%CI[49%;86%]) of drugs predicted approved were approved while 92% (95%CI[87%;98%]) of drugs predicted non-approved were still not approved (log-rank  $p$ -value $<10^{-10}$ ). A predicted approved drug was 16 times more likely to be approved compared to a predicted non-approved drug (Hazard ratio=16.4; 95%CI [8.40; 32.2]).

### **Conclusion**

As soon as P1CT completion, RESOLVED2 can accurately predict the time to FDA approval. We provide the proof of concept that drug development outcome can be predicted by machine learning strategies.

## INTRODUCTION

Drug development in Oncology is a fast-evolving field with numerous challenges<sup>1</sup>: more than 1,000 antineoplastic agents (ANA) are under investigation in 2018<sup>2</sup>. Oncology had the highest overall attrition rate for Food and Drug Administration (FDA) approval from phase I trial (95% between 2006 and 2015), phase II (92%) and phase III trials (67%)<sup>3, 4</sup>. The community aims at limiting the recruitment of patients in phase II and/or large phase III studies evaluating treatment that won't be approved for various reasons: it impairs recruitment of patients in other studies, it slows down the whole drug development process, and obviously results in substantial financial loss for pharmaceutical industry and academic institutions<sup>5</sup>. Exposure of patients to ineffective treatments and financial loss currently urge the pharmaceutical industry and academic investigators to develop new tools to enhance drug development strategies<sup>6</sup>, such as computer-assisted decisions.

Phase I trials in oncology are usually dedicated to safety analysis and meanwhile can provide early signals of efficacy of the compounds<sup>7</sup>. Classical strategies to improve research & development<sup>5</sup> are the use of surrogate markers of efficacy (overall response rate as a surrogate of overall survival)<sup>8, 9</sup> or predictive biomarkers of efficacy (molecular alterations from the tumor or liquid biopsy)<sup>10, 11</sup>. Biomarker-based strategy used in phase I can indeed significantly increase response rate and the likelihood of FDA approval<sup>12</sup>.

A tool to individually predict FDA approval for new compounds based on early clinical data is still lacking. Pharmacological data may be a cornerstone to perform such predictions for compounds with original targets or new mechanisms of action. High-volume pharmacological data are currently available in open-source databases such as DrugBank5.0<sup>13</sup>.

In the present study, we aimed to demonstrate the feasibility and utility of a recommender system based on machine-learning that could enhance drug development (RESOLVED2) in Oncology by supporting early go/no-go decision as soon as phase I trial completion.

## **MATERIALS & METHODS**

### ***Antineoplastic agents: identification and selection***

We extracted all PubMed abstracts in English and related to phase I trials in Oncology evaluating ANA for adult patients (package *RISMed* on R v3.3.3), without limitation for date of publication (**Supplementary Method 1.A**). Drug names were extracted from the titles of PubMed articles by regular expressions using the R package *stringr*. Drug names were identified by their suffix: for example, all monoclonal antibody names were queried from the “-mab” suffix (suffix list: **Supplementary Method 1.B**). For encoded drug names, we extracted drug code using the following regular expression: get any word with 2 to 4 numerical characters, optionally followed by 1 to 8 digital character (“`\\b[:alpha:]{2,4}-?[:digit:]{1,8}\\b`”).

All article titles where no drug or code was identified were manually checked. PubChem database<sup>14</sup> and Chemical Identifier Resolver<sup>15</sup> were download to annotate all identified drugs with all available alias. Individual drug names/codes were manually confirmed with search on NCI drug dictionary<sup>16</sup>, PubChem<sup>14</sup>, and Google, and to get alias for drug code, when available.

### ***Antineoplastic agents' annotations: pharmacological data***

To use pharmacological characteristics in modeling, we matched drugs identified from the PubMed abstract corpus to DrugBank5.0 database identifiers<sup>13</sup>. From DrugBank5.0 we extracted drug pharmacological category and drug molecular

target<sup>17</sup>. Briefly, \*.xml database file was download and processed using a Python-based program (ElementTree module). Each drug being annotated with features lists of various lengths, each feature was transformed as binary-encoded dummy variables (Supplementary Data 1-2). Scripts are available at <https://github.com/DITEP/RESOLVED2>.

### **Results of early clinical trial data**

For each ANA, the earliest phase I trial (without ANA combination) was selected as follow: first clinical study reporting observations of tolerability in human, or first phase I trial, or *first-in-human* trial, regardless of the inclusion criteria. Drugs for which the earliest phase I trial abstract was not available were excluded. We selected only drugs initially developed as ANAs (i.e. antibiotics or anti-rheumatoid agents subsequently developed as ANAs were not retained). Additionally, when only phase II trial publications were available on PubMed, congress abstracts reporting dose-escalation results from phase I were considered (Abstract identifier list: **Supplementary Method 2**).

The following variables were manually extracted from abstracts: primary tumor sites; study enrichment with a specific tumor site; mention of: drug target, dose expansion cohort, molecular, pathological, or circulating biomarkers, dose limiting toxicity (DLT) or maximum tolerated dose (MTD), anti-tumor clinical activity, complete tumor response, objective response rate, occurrence of common terminology criteria for adverse events (CTCAE) 4.0 grade V treatment-related adverse event, treatment-related cardiac or neuro-psychological adverse events (Variable definitions: **Supplementary Method 3**).

**Primary outcome of interest: Food and Drug Administration approval-free-survival**

We considered time to FDA approval as a right-censored variable in order to consider the unknown probability of future approval for drugs under follow-up. FDA approval database (Drugs@FDA) was extracted from the FDA portal on 2018, July 30<sup>th</sup>, thus considered as the date of point for the non-approved drugs<sup>18</sup>. FDA approval was considered as an event, whereas drugs without FDA approval were censored at time of date-point. We defined FDA approval-free-survival (FDA-aFS) as the time between the first publication date of the earliest phase I trial to the date of first FDA approval, censored by date of last news.

### ***Statistical analysis: machine learning model and performances evaluation***

Descriptive statistics were used to describe earliest phase I trial data (absolute value and percent for binary variables; median and interquartile range [Q1-Q3] for continuous variables). FDA-aFS follow-up was described using the 1-Kaplan-Meier method, with median, range, and [Q1-Q3].

Statistical analyses were performed on *R* v3.3.3 (Scripts: <https://github.com/DITEP/RESOLVED2>). The dataset was randomly split using the package *caret* in a training set (70% - for model training) and a test set (30% - for model performances evaluation) [with similar time-distribution](#). The test set remained unused during all the training. All DrugBank [categorical](#) features were encoded as binary variables. Finally, all features were binary variables. [In order to maximize the number of drugs with complete annotation, variables were rejected rather than drugs when missing data were >2.5%.](#)

[A multivariable](#) Cox model with Lasso penalization was trained to predict the FDA-aFS for each ANA<sup>19</sup>. To avoid overfitting and to allow feature selection, inverse performance of the Lasso (L1 normalization) penalization parameter ( $\lambda$  value) was

minimized on a 100-fold cross validation set derived from the training set with the package *glmnet*<sup>20</sup>.

Performances of the RESOLVED2 predictions for FDA-aFS were estimated by the concordance index (C-index) using both non-weighted (*survcomp* package), and weighted (the inverse of the probability of censoring weighted estimate (IPCW), *pec* package) methods. The area under curve (AUC) of Receiver Operating Characteristic (ROC) curve (AUROC) was computed using the predicted probabilities and censored survival data at cut-off of *t*-years (*survivalROC* package).

We compared RESOLVED2 predictions to predictions based on variables frequently used to estimate the success of a phase I trial: clinical activity detected (yes/no), are complete response(s) reported, and identification of DLT or MTD reached. The so-called EffTox model was trained using the same method and split rules than RESOLVED2, without penalization.

To facilitate interpretability and applicability of RESOLVED2, a binary classification model was computed from the previously computed scores. The main objective was to identify and therefore prevent the development of predicted non-approved drugs, to improve the current important attrition rate in drug development (i.e. drugs that will fail to be approved after phase III trials). The training set was used to identify the cut-off that maximized the difference in observed FDA-aFS between predicted approved drugs versus predicted non-approved drugs (corresponding to minimizing the log-rank-derived *p*-value).

FDA-aFS was described using Kaplan-Meier curves and *t*-year event rate estimations. Significance was defined as *p*-value<0.05.

## RESULTS

### ***FDA approval-free-survival of antineoplastic agents***

On the basis of MESH-term search, 2606 PubMed entries were identified as early clinical trials that assessed ANAs (**Figure S1**). Among these, 2415 publication titles were found to quote one drug or more (**Figure 1**). There were 619 compounds matching to 551 DrugBank entries (**Supplementary List 1**). Sixty-eight compounds derived from a parent compound were not registered in DrugBank (i.e. liposomal-encapsulated drugs, pegylated drugs, modified galenic forms etc...), or were prodrugs/compounds not used as therapeutic agents (Floxuridine for fluorouracil; exisulind for sulindac; adenosine 5'-triphosphate), and 314 drugs did not match any DrugBank entry (**Supplementary List 2**). Among the 551 DrugBank entries, 486 (88%) were developed initially as ANAs. For 24 drugs, earliest phase I trial publications were identified but abstract were not available (**Figure 1**: flow chart).

Finally, 462 drugs were selected, for which earliest phase I trial dates of publication ranged from June 1972, to October 2017, with 368 (80%) trials published after 2000. Based on abstract text, most phase I trials included all cancer types (68%), mentioned a drug target (80%), at least one DLT observed or MTD reached (77%) and a clinical activity of efficacy (69%). Few abstracts mentioned dose expansion cohort (16%) or a molecular biomarker (9%) (**Table 1**).

The median follow-up was 134 months (range: [1-425]; [Q1-Q3]: [77-203]). At 3 and 6 years of follow-up, 13% (95% confidence interval (95%CI) [10%;16%]) and 20% (95%CI [16%;24%]) of drugs have been approved, respectively. Overall, we observed that 131 out of the 462 drugs obtained FDA approval (28%; 95%CI [24%;32%]). The non-approved drug with the shortest follow-up was depatuxizumab mafodotin (10 months) while the longest was observed for treosulfan (495 months).

### ***Training of the models to predict FDA approval***

PubMed abstract data and DrugBank annotations resulted in 1411 binary variables. Ten drugs were removed because of unavailable data for clinical activity or DLT identified/MTD reached. The overall response rate was not considered because of too much unavailable data (**Table 1**).

The multivariable Cox model was penalized with the Lasso procedure, thus facilitating features selection by filtering on non-zero learned coefficients (**Figure S2, S3**). Twenty-eight features were finally selected (**Figure 2, Table S1**). Relevant treatment targets (*PDGFR $\alpha$ 1*, *PD-L1*, *HDAC1*), pharmacokinetic properties (*CYP450 substrates/modulators*; *P-glycoprotein ABCB1 substrates/modulators*), and drug categories (*Kinase Inhibitors*, *Purine Analogues*, *Antibodies*, *Proteins*) were among best predictors of FDA-aFS. Phase I trial results with *complete response reported* and trial designs such as *tumor type enrichment or selection*, *dose expansion cohort*, and *molecular biomarker* (**Figure 2**) were also identified as important features (DrugBank5.0 features definitions: **Table S2**; drugs annotations with model features: **Table S3**).

The predictive score for FDA-aFS was used for classification. The cut-off found in the training set that maximized FDA-aFS difference between predicted approved drugs and predicted non-approved drugs (**Figure S4; Figure 3A**) was further applied and evaluated in the independent test set.

#### ***Generalization of predicted FDA approval on the test set***

The predictions of RESOLVED2 were highly related with the observed FDA-aFS of ANAs included in the previously unseen test set. For the time-dependent scores, the non-weighted C-index was 0.90 and the weighted C-index (IPCW) was 0.89. Moreover, the AUROC based on Kaplan-Meier curves were 0.97 and 0.94 at 3 and 6 years, respectively (**Figure S5.A, S5.C**).

As a comparison, performances of the EffTox model in the test set (**Figure S6**) were lower in term of concordance (C-index = 0.79; IPCW = 0.69) and in term of sensitivity and specificity (survival AUROC were 0.76 and 0.84 at 3 and 6 years, respectively) (**Figure S5.B, S5.D**).

The classifier version of RESOLVED2 predicted that 81% of ANA from the test set would be non-approved. Predictions were strongly related with observed FDA-aFS (**Figure 2.B**). For example, at 3 years of follow-up, 95% (95%CI [91%;99%]) of predicted non-approved drugs were not approved, while 50% (95%CI [27%;66%]) of predicted approved drugs were indeed approved. For a later follow-up of 6 years, we found that 92% (95%CI [87%; 98%]) of predicted non-approved drugs were still not approved, while 73% (95%CI [49%; 86%]) of predicted approved drugs were indeed approved. A predicted approved drug was 16 times more likely to be approved compared to a predicted non-approved drug (Hazard ratio=16.4; 95%CI [8.40; 32.2];  $p < 10^{-10}$ ). [Proportional hazard assumption was verified \(cox.zph R function\)](#).

### ***Applications of RESOLVED2***

We applied RESOLVED2 on recent examples of early drug development (**Table 2, Table S4**). Rovalpituzumab Tesirine (Roval-T) is a bi-specific antibody developed in a hard-to-treat cancer type, namely small cell lung cancer. Epacadostat is an IDO inhibitor developed in combination with immune checkpoint blockers in melanoma. Based on available data from single-agent phase I publication (Epacadostat being not yet indexed in DrugBank5.0), RESOLVED2 found that Roval-T had a probability of FDA approval within 6 years of 73% while epacadostat had a 92% risk of failure. Only longer follow-up will confirm or not these predictions. We also applied RESOLVED2 predictions for treatments with relatively complex development and finally approved (**Table 2**). Interestingly, 4 out of 5 treatments were indeed predicted

approved and the one missed by RESOLVED2 had nevertheless a relatively high score. Several false-positive of the classifier should be mentioned (N=14 drugs predicted non-approved but finally approved, among 109 predicted non-approved drugs in the test set), however, scores of these 14 drugs were also relatively high, close to the classifier threshold (**Table S5**).

## DISCUSSION

Given the limited success rate of recent ANA development in oncology, improving early go/no-go decision after phase I clinical trial is a timely challenge. RESOLVED2 used the earliest phase I PubMed abstracts and simple pharmacological characteristics to predict the likelihood of FDA approval for individual ANA. When RESOLVED2 was used for classification, it was highly correlated with time-to-approval in the independent evaluation test set: within the first 6 years of follow-up, RESOLVED2 predictions were right for 73% of approved drugs and 92% of non-approved drugs. RESOLVED2 could potentially reduce by 81% the number of ANA undergoing further development and that would fail to achieve FDA approval.

Features included in the RESOLVED2 model supported its external validity. For instance, we found that targeting the immune checkpoint PD-L1 was associated with successful developments<sup>21 22</sup>. The model also included drug characteristics related to antibodies, such as *Complement C1q subcomponent subunit A* target, *Antibodies*, and *Proteins*, or related to targeted therapies, such as *Kinase inhibitors* and known targets (*Platelet derived growth factor receptor alpha 1*, *Histone deacetylase 1*). Trial designs also played an important role in RESOLVED predictions, such as the use of molecular biomarker, cohort enrichment with specific tumor types, and a dose expansion cohort. These findings are consistent with the positive impact of

biomarker-based strategies in oncology<sup>10</sup>, along with the rise of precision medicine, both in clinics<sup>23, 24</sup>, and in the drug development landscape<sup>1, 25</sup>. Features describing cytotoxic chemotherapy targets or categories were also found as predictive of drug approval, such as *Purine Analogues* and *DNA interacting agents*. Overall, RESOLVED2 used features describing therapeutic breakthrough, which could suggest that the development of innovative drugs should be preferred to development of “me-too” drugs”.

Some limitations should be stressed. Regarding the data used, among 619 ANAs retrieved in PubMed, 314 were not recorded in DrugBank5.0. The approval rate in our cohort was 28%, which was higher than the 5% of likelihood of approval reported elsewhere<sup>3</sup>. This could be explained by the selection of compounds reported both in PubMed and DrugBank databases.. A publication bias could have also influenced our model (i.e delayed phase I publication reporting promising drug; restriction to English-written abstract). To date, RESOLVED2 provides a probability of approval for compounds given in mono-therapy. Predicting approval for treatment combinations is a current challenge that would require dedicated data and/or a flexible modeling approach that can be derived from RESOLVED2.

As reported in 2018, FDA approval regulations evolve quickly<sup>28</sup> to improve the balance between enhanced drug access and patient safety<sup>29, 30</sup>. The Cox model assumes that the strength of predictors is constant over time. Because it has been trained on data from a long time-interval (1972-2017), RESOLVED2 predictions may have been influenced by approval rules evolution. Despite the good performances of RESOLVED2 were confirmed on an independent test set with a similar time-distribution, it would benefit from future independent prospective validation. Another assumption of the Cox model is that events are independent: here, “drug A approval”

could have influence “drug B approval”. Nevertheless, the strength of such dependencies is arguable. For example, “me-too” drugs are equivalently approved for the treatment of tumor types such as: kidney cancers ( $N = 4$  anti-angiogenic ANA), *BRAF*-mutated melanomas ( $N = 2$  *BRAF* inhibitors) and *EGFR*-mutated lung adenocarcinomas ( $N = 6$  small molecules) ([https://www.nccn.org/professionals/physician\\_gls/default.aspx](https://www.nccn.org/professionals/physician_gls/default.aspx), <sup>31</sup>). On the other hand, me-too drugs can also fail to obtain approval because of poor activity or safety: for example, panitumumab has failed approval in Head and Neck cancers while cetuximab is a long-standing standard of care<sup>32</sup>. To evaluate the added value of RESOLVED2 compared to the current approach, we designed EffTox as a proxy based on variables most frequently used to estimate the success of a phase 1, when limited information on a drug is available. Despite the final approval decision is highly contextual and uses more information, standardized decisions in this context are becoming more frequent<sup>33, 34</sup>. A basic approach as EffTox and an advanced approach as RESOLVED2 could be applied to support decision in this context.

Using a model for censored data to predict drug approval is new, accounts for heterogeneity in follow-up and maximizes the amount of data used, including recent examples. Lasso-penalized Cox model offers the advantage of only one hyperparameter to train, which is valuable in the context of a limited number of examples. Moreover, it facilitates the interpretation of the model by automated features selection and generation of hazard ratios per feature, which is valuable compared to ensemble modeling technics or neural networks, for example<sup>35</sup>.

Regarding the choice of data, numerous drug databases with various topic and aims are available<sup>26,27</sup>: DrugBank5.0 had the advantage to provide qualitative, accurate, and manually curated annotations on pharmacological properties and drug targets<sup>13</sup>.

In conclusion, the seminal RESOLVED2 experience showed that machine-learning models could efficiently support early go/no-go decision, before phase II/III trials<sup>36,37</sup>. Such models could improve the current landscape of drug development for patients, academic centers and pharmaceutical industry.

## Disclosures

**Guillaume BEINSE:** Travel and accommodation expenses for congress from Novartis, Pharmamar.

**Isabelle BORGET:** Consultant/Advisory fees from Novartis, Janssen, Merck, Takeda.

**Eric DEUTSCH:** Consultant/Advisory role for Boehringer, Medimmune, Amgen, Merck KGA, Accuray. Courses and training BMS, MSD, Amgen. Invitations to national or international congresses from Roche, Astrazeneca, Merck KGA. Research Funding from: AstraZeneca, Bayer, Biotrachea, Boehringer, BMS, Cleavel, EDF, Lilly, GSK, Merck, MSD, Medimmune, Nanobiotix, OSEO, Pierre Fabre, RaySearch Laboratories, Roche, IPSEN, Servier, Takeda. Principal/sub-Investigator of Clinical Trials for Abbvie, Agios Pharmaceuticals, Amgen, Argen-X Bvba, Arno Therapeutics, Astex Pharmaceuticals, Astra Zeneca, Aveo, Bayer Healthcare Ag, Bbb Technologies Bv, Blueprint Medicines, Boehringer Ingelheim, Bristol Myers Squibb, Celgene Corporation, Chugai Pharmaceutical Co., Clovis Oncology, Daiichi Sankyo, Debiopharm S.A., Eisai, Eli Lilly, Exelixis, Forma, Gamamabs, Genentech, Inc., Glaxosmithkline, H3 Biomedicine, Inc, Hoffmann La Roche Ag, Innate Pharma, Iris Servier, Janssen Cilag, Kyowa Kirin Pharm. Dev., Inc., Loxo Oncology, Lytix Biopharma As, Medimmune, Menarini Ricerche, Merck Sharp & Dohme Chibret, Merrimack Pharmaceuticals, Merus, Millennium Pharmaceuticals, Nanobiotix, Nektar Therapeutics, Novartis Pharma, Octimet Oncology Nv, Oncoethix, Onyx Therapeutics, Orion Pharma, Oryzon Genomics, Pfizer, Pharma Mar, Pierre Fabre, Roche, Sanofi Aventis, Taiho Pharma, Tesaro, Inc, Xencor. Research Grants from Astrazeneca, BMS, Boehringer Ingelheim, Janssen Cilag, Merck, Novartis, Pfizer, Roche, Sanofi. Non-financial support (drug supplied) from Astrazeneca, Bayer, BMS, Boringher Ingelheim, Johnson & Johnson, Lilly, Medimmune, Merck, NH TherAGuiX, Pfizer, Roche.

**Antoine HOLLEBECQUE:** Consultant/Advisory role for Amgen, Spectrum Pharmaceuticals, Lilly. Travel and accommodation expenses from Servier, Amgen, Lilly. Courses, trainings for Bayer. Principal/sub-Investigator of Clinical Trials for Abbvie, Agios Pharmaceuticals, Amgen, Argen-X Bvba, Arno Therapeutics, Astex Pharmaceuticals, Astra Zeneca, Aveo, Bayer Healthcare Ag, Bbb Technologies Bv, Blueprint Medicines, Boehringer Ingelheim, Bristol Myers Squibb, Celgene Corporation, Chugai Pharmaceutical Co., Clovis Oncology, Daiichi Sankyo, Debiopharm S.A., Eisai, Eli Lilly, Exelixis, Forma, Gamamabs, Genentech, Inc., Glaxosmithkline, H3 Biomedicine, Inc, Hoffmann La Roche Ag, Innate Pharma, Iris Servier, Janssen Cilag, Kyowa Kirin Pharm. Dev., Inc., Loxo Oncology, Lytix Biopharma As, Medimmune, Menarini Ricerche, Merck Sharp & Dohme Chibret, Merrimack Pharmaceuticals, Merus, Millennium Pharmaceuticals, Nanobiotix, Nektar Therapeutics, Novartis Pharma, Octimet Oncology Nv, Oncoethix, Onyx Therapeutics, Orion Pharma, Oryzon Genomics, Pfizer, Pharma Mar, Pierre Fabre, Roche, Sanofi Aventis, Taiho Pharma, Tesaro, Inc, Xencor. Research Grants from Astrazeneca, BMS, Boehringer Ingelheim, Janssen Cilag, Merck, Novartis, Pfizer, Roche, Sanofi. Non-financial support (drug supplied) from Astrazeneca, Bayer, BMS, Boringher Ingelheim, Johnson & Johnson, Lilly, Medimmune, Merck, NH TherAGuiX, Pfizer, Roche

**Christophe MASSARD:** Consultant/Advisory fees from Amgen, Astellas, Astra Zeneca, Bayer, BeiGene, BMS, Celgene, Debiopharm, Genentech, Ipsen, Janssen, Lilly, MedImmune, Novartis, Pfizer, Roche, Sanofi, Orion. Principal/sub-Investigator of Clinical Trials for Abbvie, Agios Pharmaceuticals, Amgen, Argen-X Bvba, Arno Therapeutics, Astex Pharmaceuticals, Astra Zeneca, Aveo, Bayer Healthcare Ag, Bbb

Technologies Bv, Blueprint Medicines, Boehringer Ingelheim, Bristol Myers Squibb, Celgene Corporation, Chugai Pharmaceutical Co., Clovis Oncology, Daiichi Sankyo, Debiopharm S.A., Eisai, Eli Lilly, Exelixis, Forma, Gamamabs, Genentech, Inc., Glaxosmithkline, H3 Biomedicine, Inc, Hoffmann La Roche Ag, Innate Pharma, Iris Servier, Janssen Cilag, Kyowa Kirin Pharm. Dev., Inc., Loxo Oncology, Lytix Biopharma As, Medimmune, Menarini Ricerche, Merck Sharp & Dohme Chibret, Merrimack Pharmaceuticals, Merus, Millennium Pharmaceuticals, Nanobiotix, Nektar Therapeutics, Novartis Pharma, Octimet Oncology Nv, Oncoethix, Onyx Therapeutics, Orion Pharma, Oryzon Genomics, Pfizer, Pharma Mar, Pierre Fabre, Roche, Sanofi Aventis, Taiho Pharma, Tesaro, Inc, Xencor. Research Grants from Astrazeneca, BMS, Boehringer Ingelheim, Janssen Cilag, Merck, Novartis, Pfizer, Roche, Sanofi. Non-financial support (drug supplied) from Astrazeneca, Bayer, BMS, Boringher Ingelheim, Johnson & Johnson, Lilly, Medimmune, Merck, NH TherAGuiX, Pfizer, Roche. **Loic VERLINGUE**: receives consulting fees from ADAPTherapy and Pierre Fabre, not related to the current work.

All other authors state they have no conflict of interest to disclose.

### **Acknowledgements**

The authors are thankful to Dr Vincent Runge for educational support.

## References

1. Wong KM, Capasso A, Eckhardt SG: The changing landscape of phase I trials in oncology. *Nat Rev Clin Oncol* 13:106–117, 2016
2. PhRMA Cancer Medicines in Development 2018 [Internet]. Phrma [cited 2018 Oct 11] Available from: <http://www.phrma.org/report/list-of-2018-medicines-in-development-for-cancer>
3. BIO, Biomedtracker: Clinical Development Success Rates 2006-2015 [Internet]. 2016 Available from: <https://www.bio.org/sites/default/files/Clinical%20Development%20Success%20Rates%202006-2015%20-%20BIO,%20Biomedtracker,%20Amplion%202016.pdf>
4. Jonsson B, Bergh J: Hurdles in anticancer drug development from a regulatory perspective. *Nat Rev Clin Oncol* 9:236–243, 2012
5. Paul SM, Mytelka DS, Dunwiddie CT, et al: How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* 9:203–214, 2010
6. Nass SJ, Rothenberg ML, Pentz R, et al: Accelerating anticancer drug development - opportunities and trade-offs. *Nat Rev Clin Oncol* , 2018
7. Hamid O, Robert C, Daud A, et al: Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N Engl J Med* 369:134–144, 2013
8. Chakiba C, Grellety T, Bellera C, et al: Encouraging Trends in Modern Phase 1 Oncology Trials. *N Engl J Med* 378:2242–2243, 2018
9. Jain RK, Lee JJ, Ng C, et al: Change in tumor size by RECIST correlates linearly with overall survival in phase I oncology studies. *J Clin Oncol Off J Am Soc Clin Oncol* 30:2684–2690, 2012

- 10.** Jardim DL, Fontes Jardim DL, Schwaederle M, et al: Impact of a Biomarker-Based Strategy on Oncology Drug Development: A Meta-analysis of Clinical Trials Leading to FDA Approval. *J Natl Cancer Inst* 107, 2015
- 11.** Stahel R, Bogaerts J, Ciardiello F, et al: Optimising translational oncology in clinical practice: strategies to accelerate progress in drug development. *Cancer Treat Rev* 41:129–135, 2015
- 12.** Schwaederle M, Zhao M, Lee JJ, et al: Association of Biomarker-Based Treatment Strategies With Response Rates and Progression-Free Survival in Refractory Malignant Neoplasms: A Meta-analysis. *JAMA Oncol* 2:1452–1459, 2016
- 13.** Wishart DS, Feunang YD, Guo AC, et al: DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 46:D1074–D1082, 2018
- 14.** <https://pubchem.ncbi.nlm.nih.gov/>
- 15.** <https://cactus.nci.nih.gov/chemical/structure>
- 16.** <https://www.cancer.gov/publications/dictionaries/cancer-drug>
- 17.** <https://www.drugbank.ca/>
- 18.** Research C for DE and: Drug Approvals and Databases - Drugs@FDA Data Files [Internet][cited 2018 Oct 11] Available from: <https://www.fda.gov/Drugs/InformationOnDrugs/ucm079750.htm>
- 19.** Simon N, Friedman J, Hastie T, et al: Regularization Paths for Cox’s Proportional Hazards Model via Coordinate Descent [Internet]. *J Stat Softw* 39, 2011[cited 2018 Oct 16] Available from: <http://www.jstatsoft.org/v39/i05/>
- 20.** Laimighofer M, Krumsiek J, Buettner F, et al: Unbiased Prediction and Feature Selection in High-Dimensional Survival Regression. *J Comput Biol J Comput Mol Cell Biol* 23:279–290, 2016

- 21.** Jardim DL, de Melo Gagliato D, Giles FJ, et al: Analysis of Drug Development Paradigms for Immune Checkpoint Inhibitors. *Clin Cancer Res Off J Am Assoc Cancer Res* 24:1785–1794, 2018
- 22.** Kaiser J, Couzin-Frankel J: Cancer immunotherapy sweeps Nobel for medicine. *Science* 362:13, 2018
- 23.** Hollebecque A, Massard C, Soria J-C: Implementing precision medicine initiatives in the clinic: a new paradigm in drug development. *Curr Opin Oncol* 26:340–346, 2014
- 24.** Massard C, Michiels S, Féré C, et al: High-Throughput Genomics and Clinical Outcome in Hard-to-Treat Advanced Cancers: Results of the MOSCATO 01 Trial. *Cancer Discov* 7:586–595, 2017
- 25.** Tsimberidou AM, Ringborg U, Schilsky RL: Strategies to overcome clinical, regulatory, and financial challenges in the implementation of personalized medicine. *Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Annu Meet* 118–125, 2013
- 26.** Yonchev D, Dimova D, Stumpfe D, et al: Redundancy in two major compound databases. *Drug Discov Today* 23:1183–1186, 2018
- 27.** Chang Y, Park H, Yang H-J, et al: Cancer Drug Response Profile scan (CDRscan): A Deep Learning Model That Predicts Drug Effectiveness from Cancer Genomic Signature. *Sci Rep* 8:8857, 2018
- 28.** Kuderer NM, Lyman GH: Evolving Landscape of US Food and Drug Administration Drug Approval in the Era of Precision Oncology: Finding the Right Balance Between Access and Safety. *J Clin Oncol Off J Am Soc Clin Oncol* 36:1773–1776, 2018
- 29.** Hwang TJ, Franklin JM, Chen CT, et al: Efficacy, Safety, and Regulatory Approval of Food and Drug Administration-Designated Breakthrough and Nonbreakthrough Cancer Medicines. *J Clin Oncol Off J Am Soc Clin Oncol* 36:1805–1812, 2018

- 30.** Shepshelovich D, Tibau A, Goldvaser H, et al: Postmarketing Modifications of Drug Labels for Cancer Drugs Approved by the US Food and Drug Administration Between 2006 and 2016 With and Without Supporting Randomized Controlled Trials. *J Clin Oncol Off J Am Soc Clin Oncol* 36:1798–1804, 2018
- 31.** Recondo G, Facchinetti F, Olausson KA, et al: Making the first move in EGFR-driven or ALK-driven NSCLC: first-generation or next-generation TKI? *Nat Rev Clin Oncol* 15:694–708, 2018
- 32.** Vermorken JB, Stöhlmacher-Williams J, Davidenko I, et al: Cisplatin and fluorouracil with or without panitumumab in patients with recurrent or metastatic squamous-cell carcinoma of the head and neck (SPECTRUM): an open-label phase 3 randomised trial. *Lancet Oncol* 14:697–710, 2013
- 33.** Hobbs BP, Barata PC, Kanjanapan Y, et al: Seamless Designs: Current Practice and Considerations for Early-Phase Drug Development in Oncology. *J Natl Cancer Inst* 111:118–128, 2019
- 34.** Cecchini M, Rubin EH, Blumenthal GM, et al: Challenges with Novel Clinical Trial Designs: Master Protocols. *Clin Cancer Res Off J Am Assoc Cancer Res* , 2019
- 35.** Wang P, Tech V, Li Y, et al: Machine Learning for Survival Analysis: A Survey. *ACM Comput Surv* 1:38
- 36.** DiMasi JA, Hermann JC, Twyman K, et al: A Tool for Predicting Regulatory Approval After Phase II Testing of New Oncology Compounds. *Clin Pharmacol Ther* 98:506–513, 2015
- 37.** Lo AW, Siah KW, Wong CH: Machine-Learning Models for Predicting Drug Approvals and Clinical-Phase Transitions [Internet]. Rochester, NY, Social Science Research Network, 2017[cited 2018 Oct 11] Available from: <https://papers.ssrn.com/abstract=2973611>



## Figure legends

### Figure 1. Flow-chart for drug selection.

ANA: antineoplastic agents; ECT: early clinical trial.

### Figure 2. RESOLVED2 model: Beta coefficients from Lasso-penalized Cox model.

For features definitions, see **Table S2** and **Supplementary Method 3**. The best Lasso penalization parameter ( $\lambda_{\min}$  value) was determined using a 100-fold cross-validated Cox regression model on the training set (**Figure S3**). Evolution of penalized  $\beta$  coefficients with  $\lambda$  values are depicted in **Figure S4**. A Cox regression model L1 (Lasso)-penalized with the  $\lambda_{\min}$  value identified by cross validation (6.59.10-2) was fit allowing feature selection and associated  $\beta$  coefficients computation.

### Figure 3. Food and Drug Administration approval-free-survival.

Kaplan-Meier curves of Food and Drug Administration (FDA) approval-free-survival. **A (left)**: FDA approval-free-survival according to the RESOLVED classification in the training set. Predicted approved and non-approved ANA refer to binary predictions performed using the RESOLVED2 classifier based on RESOLVED2 scores. In the present plot, the training set was split using the optimal cut-off calibrated on minimal log-rank derived  $p$ -value (see *Methods*). **B (right)**: FDA approval-free-survival according to the RESOLVED classification in the test set. The test set was split using the optimal cut-off in RESOLVED2 scores, identified in the training set (see *Methods*).

**For Supplementary Tables, Figures, and Methods, please refers to Supplementary Material list.**