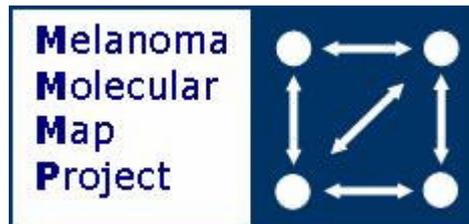


# Targeted Therapy Database (TTD)



**Melanoma Molecular Map Project (MMMP)**

<http://www.mmmp.org>

## **The Targeted Therapy Database**

The Melanoma Molecular Map Project (MMMP, [www.mmmp.org](http://www.mmmp.org)) hosts the Targeted Therapy Database (TTD), a systematic collection of the scientific knowledge regarding the development of targeted therapy for melanoma.

The TTD gathers in a standardized and research-oriented fashion the published evidence on the molecular features that can be exploited to foster the development of patient and melanoma specific therapies.

The TTD can be searched for the following three main objectives:

- 1) To provide both basic researchers and clinical investigators with an unprecedented synopsis of the available scientific literature on the development of targeted therapy for melanoma;
- 2) To obtain summaries of the current evidence on the relationship between each single molecule (or set of molecules) and the efficacy of a given therapeutic agent (or set of therapeutic agents);
- 3) To match the patient/cancer molecular profile with the available scientific evidence on the targeted therapy of melanoma, thus developing a drug ranking model for the personalized treatment of melanoma.

The sources of the information input in the TTD are the PubMed, Medline, Embase, Cancerlit and Cochrane databases. Literature search is aimed to identify scientific evidence about the relationship between:

A) Any molecule (in its particular functional state) and the anti-melanoma efficacy of a therapeutic agent being used or being investigated for the treatment of melanoma (relationship of sensitivity / resistance);

B) Any molecule (in its particular functional state) and the toxicity of any therapeutic agent being used or being investigated for the treatment of melanoma (toxicity relationship);

C) Any molecule that - after modulation of its particular functional state by a "modifier" - can synergistically increase the efficacy a therapeutic agent being used or being investigated for the treatment of melanoma (synergism relationship).

The data are extracted from each retrieved article according to the following driving principle: the Authors of each article describe their findings and virtually always come to a main conclusion, whether "positive" (e.g., a molecule in a specific state can favor tumor response to a given treatment), "negative" (e.g., a molecule in a specific state can oppose tumor response) or "null" (e.g., tumor response is unaffected by a given molecule in a specific state). In other words, each study sustains one *targeted therapy hypothesis*, whether positive (favorable for the patient), negative (unfavorable) or null (unimportant, not influential).

## The TTD structure

The data are organized in rows and columns using an MS Excel file. Each row contains the main data representing the targeted therapy hypothesis made by the Authors of a given article. Each column contains one type of data according to a standardized format.

The following 15 columns compose the database:

- 1) *ID*: this is a unique number identifying each record (that is, each row of the database).
  
- 2) *Source*: it indicates where the molecule under investigation (see next column) is expressed/ present. For instance, somatic mutations of BRAF are investigated in melanoma, polymorphisms of genes involved in drug metabolism can be studied in any patient's nucleated cell, and expression of cytokine receptors can be assessed in immune cells.
  
- 3) *Molecule*: this is the name of the molecule under investigation as a tumor specific target, or as a biomarker of resistance/ sensitivity of melanoma to therapeutic agents. The molecule's name generally is that reported by the Authors of the corresponding article. Since molecules are often studied at both protein and mRNA level, in this column no such distinction is made.
  
- 4) *Alias (molecule)*: since molecules often have multiple names, one alias of the molecule of interest is often reported in this column in order to clarify its identity. Aliases are chosen on the basis of international databases such as HUGO (<http://www.genenames.org>) and Uniprot (<http://www.uniprot.org>).
  
- 5) *State (molecule)*: it refers to the type of condition in which the molecule exerts the biological activity related to the targeted therapy hypothesis reported by the article. For instance, the expression "mut V600E" for the protein BRAF refers to its mutation V600E as opposed to the wild type protein or any other mutational status. The word "expressed" is utilized to indicate that the molecule is expressed or overexpressed as opposed to absent or underexpressed: the database does not reports other details (e.g., overexpressed with respect

to what reference cell/ tissue type), for which the users are referred to the original articles (see "Reference" column).

6) *Modifier*: it refers to any drug or drug-like compound or laboratory method that can modify the activity/expression of a molecule of interest. For instance, a small molecule inhibitor can decrease the biological activity of a target molecule; likewise, technology based on RNA interference (e.g. small interfering RNA) can downregulate the expression of a gene of interest.

7) *Alias (modifier)*: since modifiers often have multiple names, an alias of the modifier under investigation can be reported in this column in order to facilitate its identification.

8) *Relationship*: this column reports the hypothesized relationship between the molecule of interest and the corresponding treatment/drug (see "Drug" column). Three main types of relationships are considered: A) **Efficacy**: the molecule under investigation can be associated with either sensitivity or resistance to a therapeutic agent; B) **Synergism**: the modulation of a molecule activity by a modifier (see "Modifier" column) can be associated with an increased (synergism) or decreased (antagonism) therapeutic activity of a given drug/treatment; C) **Toxicity**: the molecule under investigation can be associated with either increased or decreased toxicity of a given drug/treatment. Of course, all these associations can be reported to be absent. For the purpose of prompt identification, positive (i.e. with positive effects), negative (i.e. with adverse effects) and null associations are highlighted with different colors (green, orange and blue, respectively).

9) *Drug*: this is the drug (or more generally the treatment) whose effectiveness can be influenced (positively, negatively or not significantly) by the molecule listed in column "Molecule". The drug's name generally is that reported by the Authors of the corresponding article.

10) *Alias (drug)*: since drugs often have multiple names, one alias of the drug of interest is often reported in this column in order to clarify its identity.

11) *Model*: this column reports the model used by the Authors to generate the hypothesis. Seven different models are considered:

- 1) animal, in vitro (e.g., murine melanoma cell line)
- 2) animal, in vivo (e.g., syngeneic murine melanoma model)
- 3) human in vitro (e.g., human melanoma cell line)
- 4) human xenograft (e.g., human melanoma xenogeneic model)
- 5) clinical study / non randomized clinical trial
- 6) randomized controlled trial
- 7) meta-analysis

This order is dictated by the "distance" of the model from the human-in vivo condition, or - in other words - by the level of evidence of the published data. This order will play a key role in the "weight" assigned to each study, as described in detail later on.

12) *H (hypothesis)*: As above mentioned, each article can be classified according to the main conclusions of its Authors supporting a "positive" hypothesis (e.g., a molecule in a specific state can favor tumor response to a given treatment), "negative" hypothesis (e.g., a molecule in a specific state can oppose tumor response) or "null" hypothesis (e.g., tumor response is unaffected by a given molecule in a specific state). Following this principle, each record (row) of the TTD is assigned a value that identifies the corresponding hypothesis (+1, -1 or 0, respectively).

13) *Cases*: this is the number of cases (e.g., patients, animals, cell lines) examined. At present, this information is only available for clinical studies/ trials (i.e., number of patients).

14) *Reference*: the citation of the source of information is reported.

15) *Notes*: additional information on the study results/ features can be found in this column in order to facilitate the interpretation of the data reported in the previous columns. This information can help users understand whether or not the molecular condition described in the record applies to their research/clinical question.

The information found in the TTD regards cutaneous melanoma, except for drug toxicity data (which are independent of the tumor type). If data regard uveal melanoma, this is specified at the beginning of the column "Notes" by the bolded expression "Uveal melanoma". Therefore, should one be interested exclusively in targeted therapy for uveal melanoma, data must be ordered by column "Notes": this way the information contained in this column is rearranged in the alphabetical order and data on uveal melanoma will appear towards the end of the database as a sequence of rows tagged by the expression "Uveal melanoma" written in the column "Notes".

Likewise, information on specific subtypes of melanoma (e.g., acral lentiginous melanoma, mucosal melanoma) can be easily retrieved using the same method.

Information on gene polymorphisms and drug toxicity can derive from non melanoma specific models, as specified in the "Notes" column in bold character.

## Synopsis of the evidence

As above mentioned, the TTD enables investigators to find targeted therapy related information organized in a standardized and research oriented fashion. Since data are collected in an Excel file, they can be ordered by each of the 15 columns and also by any combination of three columns in sequential order.

For instance, by ordering the database by "Molecule", "State" and "Drug" (in this sequence), one can easily obtain for each molecule (and its state) the list of therapeutic agents whose efficacy is influenced by that molecule (in that particular state), as shown in **Figure 1**.

On the other hand, by ordering the database by "Drug", "Molecule" and "State" (in this sequence), one can easily obtain for each therapeutic agent the list of molecules (and their state) that can modulate its efficacy, as shown in **Figure 2**.

Likewise, by ordering the database by "Drug", "Relationship" and "Modifier" (in this sequence), one can easily obtain for each therapeutic agent the list of compounds that can modulate its efficacy.

Obviously, many other searches can be performed by ordering the columns on the basis of a specific interest (e.g., evidence only from human models) or research question (e.g., "what gene polymorphisms affect the toxicity of cisplatin?").

**Figure 1**

Molecule	State (molecule)	Relationship	Drug	Alias (drug)	Model	Cases	Reference	Notes
BRAF	mut V600E	sensitivity to	U0126			3	Smalley KS, Oncogene 2009, 28:85-94	U0126: MEK inhibitor
BRAF	mut V600E	sensitivity to	U0126			4	Sharma A, Cancer Res 2006, 66:8200-9	U0126 is a MEK inhib
BRAF	mut V600E	no relationship with	Sorafenib + Rapamycin			3	Molhoek KR, J Transl Med 2005, 3:39	BRAF mutational stat
BRAF	mut V600E	no relationship with	Sorafenib	Nexavar (R)		5	34 Eisen T, Br J Cancer 2006, 95:581-6	Tumor response was i
BRAF	mut V600E	no relationship with	Sorafenib	Nexavar (R)		5	37 Min CJ, J Clin Oncol 2008, 26: abstract 9072	Sorafenib: BRAF, CR/
BRAF	mut V600E	no relationship with	Sorafenib	Nexavar (R)		4	Sharma A, Cancer Res 2006, 66:8200-9	Sorafenib: BRAF, CR/
BRAF	mut V600E	resistance to	Sorafenib	Nexavar (R)		3	McDermott U, Proc Natl Acad Sci USA 2007, 104:19936-41	Sorafenib: BRAF, CR/
BRAF	mut V600E	resistance to	Sorafenib	Nexavar (R)		3	Smalley KS, Oncogene 2009, 28:85-94	Sorafenib: BRAF, CR/
BRAF	mut V600E	sensitivity to	Sorafenib	Nexavar (R)		4	Karasarides M, Oncogene 2004, 23:6292-8	Sorafenib: BRAF, CR/
BRAF	mut V600E	sensitivity to	RAF-265			4	Fecher LA, Pigment Cell Melanoma Res 2008, 21:410-1	RAF 265 is a BRAF ir
BRAF	mut V600E	sensitivity to	PLX4032			3	Sala E, Mol Cancer Res 2008, 6:751-9	PLX4032 is a BRAF ir
BRAF	mut V600E	sensitivity to	PLX 4720			4	Tsai J, Proc Natl Acad Sci USA 2008, 105:3041-6	PLX 4720 is a BRAF i
BRAF	mut V600E	sensitivity to	PD0325901 + ABT-737			4	Cragg MS, J Clin Invest 2008, 118:3651-9	U0126: MEK inhibitor.
BRAF	mut V600E	sensitivity to	IGFBP7			4	Wajapeyee N, Mol Cancer Ther 2009, 8:3009-14	In human melanomas
BRAF	mut V600E	resistance to	Daunorubicin			3	Sheridan C, J Biol Chem 2008, 283:22128-35	Oncogenic BRAF cau
BRAF	mut V600E	resistance to	Cisplatin	CDDP		3	Sheridan C, J Biol Chem 2008, 283:22128-35	Oncogenic BRAF cau
BRAF	mut V600E	sensitivity to	CI-1040			4	Solit DB, Nature 2006, 439:358-62	CI-1040 is a MEK inhi
BRAF	mut V600E	sensitivity to	BRAF siRNA			3	Hingorani SR, Cancer Res 2003, 63:5198-202	When present, BRAFi
BRAF	mut V600E	sensitivity to	AZ628			3	McDermott U, Proc Natl Acad Sci USA 2007, 104:19936-41	AZ628 is a BRAF inhi
BRAF	mut V600E	sensitivity to	AZ628			3	Montagut C, Cancer Res 2008, 68:4853-61	AZ628: RAF selective
BRAF	mut V600E	resistance to	Actinomycin-D			3	Sheridan C, J Biol Chem 2008, 283:22128-35	Oncogenic BRAF cau
BRAF	mut V600E	no relationship with	17-AAG	Tanespimycin		3	Grbovic OM, Proc Natl Acad Sci USA 2006, 103:57-62	17-AAG is a HSP90 ir
BRAF	mut V600E	no relationship with	17-AAG	Tanespimycin		5	15 Solit DB, Clin Cancer Res 2008, 14:8302-7	No objective response
BRAF	mut V600E	sensitivity to	17-AAG	Tanespimycin		3	da Rocha Dias S, Cancer Res 2005, 65:10686-91	17-AAG is a HSP90 ir

**Figure 2**

Source	Molecule	State (molecule)	Modifier	Relationship	Drug	Model	Cases	Reference	Notes
Melanoma	APEX1	expressed		no relationship with	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	PARP1, POLB, APEX1, LIG1
Melanoma	bFGF	expressed		resistance to	Temozolomide		3	Fontijn D, Mol Cancer Ther 2007, 6:2807-2814	bFGF overexpression can
Melanoma, endothelial	BRAF, CRAF, VEGFR, PDGFR	inhibited by	Sorafenib	synergism with	Temozolomide		5 167	Amaravadi RK, Clin Cancer Res 2009, 15:151-157	4-arm phase II trial (metas
Melanoma	CDK	inhibited by	PHA-848125	synergism with	Temozolomide		3	Caporali S, Pharmacol Res 2010, 61:437-442	The cyclin-dependent kina
Melanoma	DHFR	inhibited by	Pyrimethamine	synergism with	Temozolomide		3	Chen M, Mol Cancer Res 2009, 7:703-712	The authors screened 2,000
Melanoma	Galectin-1	downregulated (by)	Galectin-1 siRNA	sensitivity to	Temozolomide		2	Mathieu V, J Invest Dermatol 2007, 127:2303-2308	Galectin-1 protects from TI
Melanoma	Gene set: ABCC3, APLF	expressed (high)		sensitivity to	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	This set of genes resulted
Melanoma	Gene set: ARHGAP29, C	expressed (high)		resistance to	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	This set of genes resulted
Melanoma	hTERT	downregulated (by)	hTERT dominant negative	resistance to	Temozolomide		3	Tentori L, Mol Pharmacol 2003, 63:192-200	Inhibition of telomerase in
Melanoma	hTERT	active (activity)		no relationship with	Temozolomide		3	Tentori L, Mol Pharmacol 2003, 63:192-200	Susceptibility to TMZ of me
Immune cells	IFNAR	activated by	IFN alpha 2b pegylated	no synergism with	Temozolomide		5 124	Spieth K, Ann Oncol 2008, 19:801-6	Overall response rate: 18%
Melanoma	IGF1R	downregulated (by)	IGF1R siRNA	sensitivity to	Temozolomide		3	Yeh AH, Oncogene 2006, 25:6574-81	These results support dev
Melanoma	IKKB	inhibited by	BMS-345541	synergism with	Temozolomide		4	Yang J, Mol Cancer Ther 2009, 8:636-47	BMS-345541 is an IKKB in
Melanoma	IL29R	activated by	IL-29	synergism with	Temozolomide		3	Guenterberg KD, Mol Cancer Ther 2010, 9:1039-43	Bortezomib-induced and te
Immune cells	IL2R	activated by	IL2	no synergism with	Temozolomide		5 38	Tarhini AA, Cancer 2008, 113:1632-40	IL-2: high dose. The overal
Melanoma	Integrin alpha V	inhibited by	Cilengitide	synergism with	Temozolomide		2	Tentori L, Oncol Rep 2008, 19:1039-43	The integrin antagonist cil
Melanoma	KIT, PDGFR	inhibited by	Imatinib	antagonism with	Temozolomide		3	Triozzi PL, Melanoma Res 2008, 18:420-3	Imatinib causes TMZ resis
Melanoma	KIT, PDGFR	inhibited by	Imatinib	synergism with	Temozolomide		4	Triozzi PL, Melanoma Res 2008, 18:420-3	Imatinib causes TMZ resis
Melanoma	LIG1	expressed		no relationship with	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	PARP1, POLB, APEX1, LIG1
Melanoma	MGMT	upregulated (by)	MGMT genetic engineering	resistance to	Temozolomide		3	Passagne I, Toxicol Appl Pharmacol 2006, 210:106-11	O(6)-methylguanine DNA-r
Melanoma	MGMT	methylated (gene promoter)		no relationship with	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	No correlation was observ
Melanoma	MGMT	methylated (gene promoter)		no relationship with	Temozolomide		5 49	Rietschel P, J Clin Oncol 2008, 26:2299-305	In this phase II study of ext
Melanoma	MGMT	inhibited by	Lomeguatrib	no synergism with	Temozolomide		5 32	Kefford RF, Br J Cancer 2009, 100:1245-9	In this phase I study of exte
Melanoma	MGMT	inhibited by	O6 benzylguanidine	synergism with	Temozolomide		3	Mhaidat NM, Br J Cancer 2007, 97:1225-30	Inhibition of MGMT with O6
Melanoma	MGMT	inhibited by	O6 benzylguanidine	synergism with	Temozolomide		3	Naumann SC, Br J Cancer 2009, 100:322-30	MGMT inactivation by O(6)-
Melanoma	MGMT	inhibited by	Lomeguatrib	no synergism with	Temozolomide		6 104	Ranson M, J Clin Oncol 2007, 25:2540-5	RCT of the combination of
Melanoma	MGMT	inhibited by	IL24	synergism with	Temozolomide		3	Zheng M, Mol Cancer Ther 2008, 7:3842-5	Interleukin-24 overcomes i
Melanoma	MGMT	inhibited by	4BTG	synergism with	Temozolomide		4	Middleton MR, Int J Cancer 2000, 85:248-51	4BTG was less toxic (but e
Melanoma	MGMT	inhibited by	O6 benzylguanidine	synergism with	Temozolomide		4	Ueno T, Mol Cancer Ther 2006, 5:732-8	Isolated limb infusion moc
Melanoma	MGMT	inhibited by	O6-benzylguanidine	synergism with	Temozolomide		3	Fontijn D, Mol Cancer Ther 2007, 6:2807-2814	MGMT causes resistance i
Melanoma	MGMT	expressed (high)		resistance to	Temozolomide		3	Mhaidat NM, Br J Cancer 2007, 97:1225-30	Inhibition of MGMT with O6
Melanoma	MGMT	expressed		no relationship with	Temozolomide		5 50	Middleton MR, Br J Cancer 1998, 78:1199-204	Measurements of pretreat
Melanoma	MGMT	expressed		no relationship with	Temozolomide		5 49	Rietschel P, J Clin Oncol 2008, 26:2299-305	In this phase II study of ext
Melanoma	MGMT	expressed		resistance to	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	MGMT showed a significan
PBMC	MGMT	expressed		toxicity decreased for	Temozolomide		5 30	Middleton MR, Int J Cancer 2000, 88:469-74	Original: underexpression
Melanoma	MGMT	demethylated by	5-azacytidine	antagonism with	Temozolomide		3	Fontijn D, Mol Cancer Ther 2007, 6:2807-2814	MGMT causes resistance i
Melanoma	MGMT	active (high activity)		resistance to	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	MGMT showed a significan
Melanoma	MGMT	active (high activity)		resistance to	Temozolomide		3	Naumann SC, Br J Cancer 2009, 100:322-30	The O(6)-methylguanidine-C
Melanoma	MPG	expressed		no relationship with	Temozolomide		3	Augustine CK, Clin Cancer Res 2009, 15:151-157	PARP1, POLB, APEX1, LIG1

## **Figure 1**

Example of evidence synopsis regarding the targeted therapy of melanoma, as obtained by searching the Targeted Therapy Database (TTD). The available evidence on the relationship between a molecule state (BRAF mutation V600E) and its effects on different therapeutic agents is shown. Due to space considerations, not all columns or rows are displayed.

## **Figure 2**

Example of evidence synopsis regarding the targeted therapy of melanoma, as obtained by searching the Targeted Therapy Database (TTD). The available evidence on the relationship between a drug (temozolomide) and the molecular determinants of its therapeutic effect is shown. Due to space considerations, not all columns or rows are displayed.

## Summary of the evidence

The TTD offers the opportunity of making summaries of the available evidence on a given subject.

The standard way of making a quantitative review of the available scientific knowledge is performing a meta-analysis, which is considered the highest level of evidence in medicine, particularly when based on randomized controlled trials.

The basic idea behind a meta-analysis is to calculate the weighted mean of the results reported by different studies regarding a particular subject; to this aim, the following key steps must be taken: an effect measure (e.g., odds ratio, hazard ratio, relative risk, risk difference, mean, rate) common to all the studies must be identified, the effect size (and its variance) must be extracted from each study, and then the weighted mean of the effect sizes can be calculated. In a therapeutic perspective, this overall effect quantifies the benefit (or the harm) of a given treatment, its confidence interval (CI) representing the measure of uncertainty about the estimate of the overall effect (which in turn determines the statistical significance in terms of type I error, based on the predefined alpha level of significance).

In the light of these considerations, one can understand why meta-analysis is not appropriate for making summaries of the information contained in the TTD: in fact, the different effect measures adopted by the Authors to describe the results obtained in different models (ranging from animal in vitro models to randomized clinical trials) cannot be pooled together. Moreover, even though the effect measures were the same, different experimental models cannot be considered equally informative and reliable: obviously, human and in vivo models provide higher level of evidence as compared to animal and in vitro models.

Therefore, the TTD cannot be exploited to calculate an overall effect size for a given therapeutic approach: this is why the TTD does not report any data regarding the effect sizes of the single studies.

So, what do we mean by "summary of the evidence" within the TTD ? As above mentioned, each study (which is represented by a row of the database) can be envisaged as a working hypothesis about a targeted therapy against melanoma. When more than one record (i.e., one row of the database) exists for a given hypothesis (e.g., BRAF activated by mutation V600E

modulates the efficacy of small molecule inhibitor sorafenib), a score-based approach can be proposed to make the summary of the available evidence.

This method aims to identify the "prevalent" hypothesis, a process taking the following steps (see also **Figure 3**):

**1)** Each record (i.e., each row of the database) is assigned one of the integer numbers "+1", "-1" or "0", based on the fact that it represents a piece of evidence in support of one of the three possible hypotheses (as expressed by the Authors of the corresponding manuscript):

**A)** *positive relationship* (green color in the database): the study supports the hypothesis that the molecule (e.g. BRAF) in a particular state (e.g. mutation V600E) is associated with increased efficacy of a drug, synergism between drugs or decreased toxicity of a drug. On the practical ground, a patient carrying this molecule (in this specific state) would benefit of the given treatment;

**B)** *negative relationship* (orange color): the study supports the hypothesis that the molecule can oppose the efficacy of the drug; a patient (tumor) carrying this molecule (in this specific state) would be refractory to the given treatment

**C)** *null relationship* (blue color) if the study supports the hypothesis that the molecule does not affect the efficacy of the drug; knowing that a patient (tumor) carries this molecule (in this specific state) would be uninformative in terms of responsiveness to the treatment.

**2)** Each record is also assigned a score (**model score**), based on the experimental/clinical model used to generate the targeted therapy hypothesis. Clearly, the evidence coming from an in vitro study carried out with murine melanoma cell lines cannot have the same "weight" as the evidence derived - for instance - from a study performed in a human trial model. The closer the model to the in vivo human condition, the higher the level of evidence and thus the greater is the weight assigned to that study.

Within the frame of this arbitrary score, the proportion between the weights of "adjacent" models is fixed: in particular, the score of each model is twice that of the immediately precedent model. The starting score (model: animal, in vitro) was set to 6 because this is the smallest natural number that meets the decision rule below described (in case a single study based on such a model supported a given hypothesis).

The evidence score is then adjusted according to an additional weight (**size score**), which is based on the number of cases analyzed ("Cases" column): this way, studies describing results obtained from larger series are assigned a higher score.

The total **evidence score** ( $ES_i$ ) for each hypothesis  $i$  is computed according to the following formula:

$$ES_i = (\text{Model score}) \times (\text{Size score}),$$

where Size score =  $n/10$  ( $n$  is the sample size [e.g., number of patients enrolled] of the study under evaluation).

**3)** The percentage of the evidence score (**score percentage, SP**) in favor of each of the three above mentioned hypotheses is simply defined as the proportion between the evidence score in favor of each hypothesis  $i$  and the sum of the evidence score of all hypotheses:

$$SP = ES_i / \sum (ES_i)$$

**4)** At this point, a decision rule must be applied to determine whether or not a prevalent hypothesis exists: we chose 50% (0.5) of the evidence score as the minimum value to define the prevalent hypothesis. In other words, if one of the three possible hypotheses (i.e., positive, negative, null) is associated with more than 50% of the available evidence score and the lower level of the 95% CI of this proportion does not cross this decision rule value, one can reasonably suppose this is the prevalent hypothesis in the scientific literature.

The 95% CI of the score percentage (SP) can be calculated according to the Agresti-Coull formula (which provides a substantial improvement over the widely used Wald method especially for proportion values near 0 and 1 and for small sample sizes, as it can occur in the TTD):

$$\text{Score percentage 95\% CI} = SP_c \pm 1.96 * SE$$

where:

$SP_c = [ES_i + (1.96^2/2)]/TSc$  , that is the score percentage (SP) corrected according to the Agresti-Coull method

$SE = \sqrt{[SP_c * (1 - SP_c) / TSc]}$  , that is the standard error of  $SP_c$

$TSc = \sum (ES_i) + (1.96^2)$  , that is the total evidence score (supporting any given hypothesis  $i$ ) corrected according to the Agresti-Coull method.

A formal comparison between a given score percentage (SP) and the 50% (0.5) decision rule value can be made using a Z-test, according to the following formula:

$$Z = (SP - 0.5) / SE$$

where:

$$SP = ES_i / \sum (ES_i)$$

$$SE = \sqrt{[(SP_c * 1 - SP_c) / \sum (TSc)]}$$

For a two-tailed test, the P-value is given by:

$$P\text{-value} = 2 [1 - \Phi (|Z|)]$$

where  $\Phi (|Z|)$  = standard normal cumulative distribution.

Of course, the decision rule value (0.5) can be shifted up or down so to make it more or less stringent respectively, thus rendering more or less conservative the conclusion on the relationship between the patient's profile and the response to treatment.

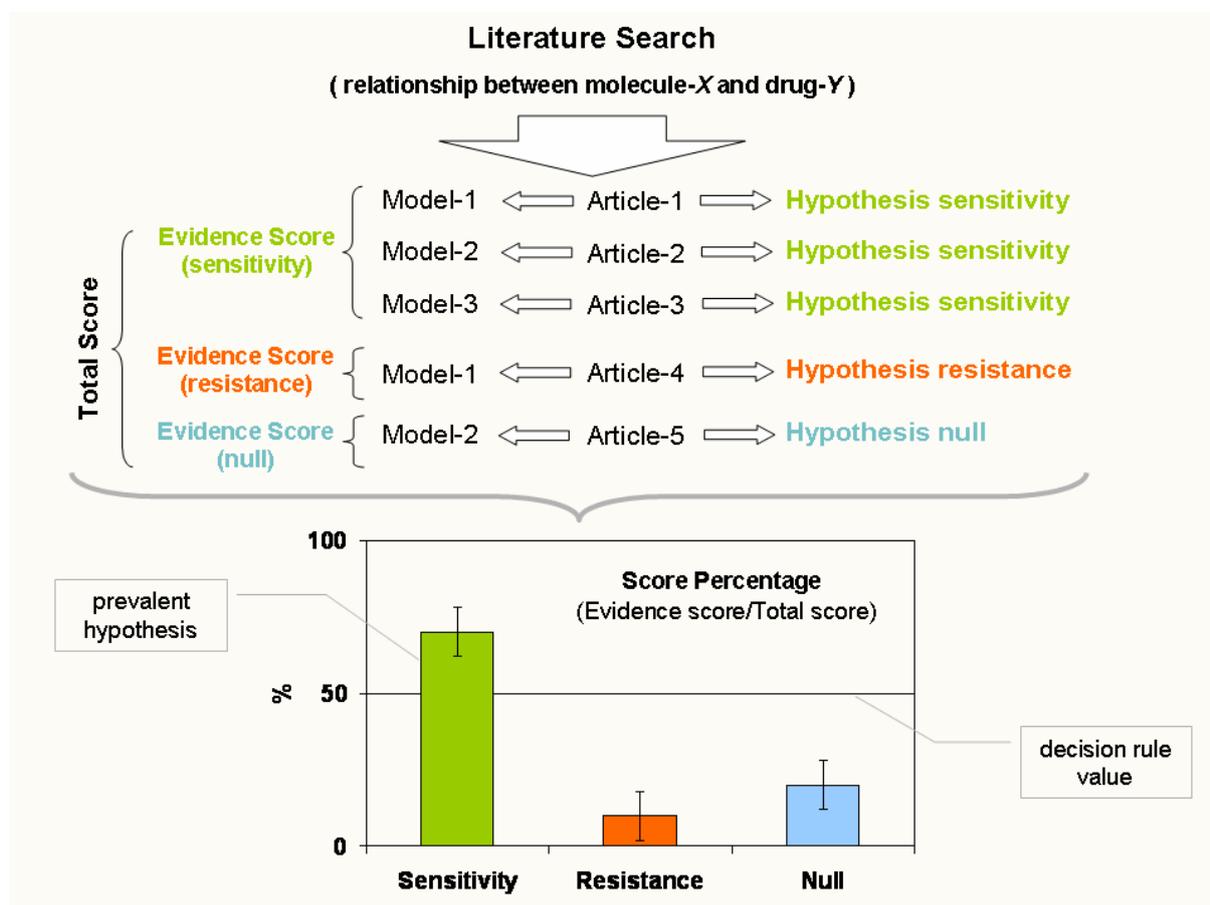
If none of the three hypotheses meets the decision rule, we can reasonably suppose that there is no prevalent hypothesis, that is, there is not enough evidence to link a given molecule (in a particular state) to the efficacy/synergism/toxicity of a given drug.

**5)** Once we know that there is enough evidence to support the hypothesis that no relationship exists between a molecule and a drug, or that not enough evidence exists to support any hypothesis on this relationship, this molecule is eliminated from the list of molecules useful to predict drug responsiveness. Importantly, this is not a definitive elimination, because new data will likely be published on this relationship and thus the result of the summary can change any time. Since the TTD is routinely updated, the selection of relevant molecules is a dynamic process that can provide different results over time as the scientific knowledge grows.

**6)** If the summary of evidence is instead in favor of the hypothesis that a molecule (in a particular state) can modulate (either positively or negatively) the activity of a treatment, then that molecule is added to the list of molecules potentially useful (i.e., informative) to predict the responsiveness to the treatment.

Of note, using the same methodology just above exposed, one can assess whether or not there is a prevalent hypothesis regarding the relationship between any modifier (or a set of modifiers) and the therapeutic activity of a drug (or set of drugs).

**Figure 3**



### Figure 3

A scheme of the evidence score method to synthesize the literature evidence and identify prevalent hypotheses regarding the relationship of sensitivity/resistance between a given molecule (in a specific state) and a given drug. Each study is assigned an evidence score based on the experimental model used to generate the findings reported in each article. In this example, 70% of the total score (that is, 70% of the published evidence rated according to the experimental model used to generate the findings reported in each article) supports the hypothesis that molecule-*X* (in a particular state, here not specified for the sake of simplicity) is associated with responsiveness to drug-*Y*. To be defined as "prevalent", the hypothesis must be characterized by the fact that the lower bound of the 95% confidence interval of its score percentage does not cross the decision rule value (50%).

The same method can be used to identify prevalent hypotheses regarding the relationship of toxicity and synergism (see text for more details).

## Drug ranking system

Once a list of molecules for which "consistent" evidence is available in favor of their role in predicting the responsiveness (or refractoriness) to a specified therapeutic agent, as assessed by means of the above described summary of the evidence, one might be willing to test the relevant biospecimens from a given patient for these molecules and match the patient's molecular profile with the currently available evidence on targeted therapy.

This opens the avenue to the use of the already available scientific knowledge for generating hypothesis of personalized treatment based on the fundamental principle of molecular medicine: to use the patient (disease) molecular profile for designing the treatment most effective and least toxic.

Before entering the technical details, one crucial issue must be clearly addressed. The TTD has exclusively research purposes, and thus neither the information nor the analytical models included in this databank should be used for the clinical decision making process by any means. In fact, this way of summarizing the evidence across (sometime very) different models has never been reported before and thus it requires adequate validation before it can be considered reliable on the clinical ground.

Having clarified this issue, the following steps can be taken to match the patient's molecular profile with the current evidence on targeted therapy (see also **Figure 4**):

**1)** Using the above described score-based system, the informative molecules (each along with a particular state of expression/ function) are extracted from the TTD along with their score percentage (SP) and 95% CI. *Each SP can be viewed as a measure of strength of the hypothesis sustaining the relationship between the molecule and the drug efficacy (toxicity, synergism) based on the available literature as rated by the evidence score above described.*

**2)** Score percentages (SP) of molecules associated with sensitivity to treatment are initially assigned a "+" sign (e.g. BRAF mutation V600E increases the efficacy of drug Sorafenib), whereas molecules associated with resistance to treatment are assigned a "-" sign (e.g. BRAF mutation V600E decreases the efficacy of drug Sorafenib). Then, the concordance (or discordance) between the molecular state of the prevalent hypothesis and that of the patient

(tumor) must be assessed. In particular, the sign of the SP will be left unchanged if the patient carries the same molecular state as that of the SP (e.g. BRAF mutation V600E); in contrast, if the patient carries the "opposite" molecular state (e.g. BRAF wild type), the SP will be assigned the opposite sign.

**3)** At this point, an **overall score (OS)** can be calculated as the weighted average of the score percentage calculated for each informative molecule. The OS and its confidence interval can be calculated using the inverse variance method as follows:

$$\text{OS} = \sum (W_i * SP_i) / \sum W_i$$

and

$$\text{Overall score 95\% CI} = \text{OS} \pm 1.96 * \text{SE} ,$$

where:

$SP_i$  : score percentage for each molecule  $i$  (in a specific state)

$W_i = 1/V_i$  , the weight assigned to each molecule based on the variance of the SP

$V_i = [\text{SPc} * (1 - \text{SPc}) / \text{TSc}]$  , i.e. the variance of the SPc calculated for each molecule  $i$

SE = standard error =  $\sqrt{\text{OV}}$

OV = overall variance =  $1 / \sum W_i$

The interpretation of the result obviously depends upon the decision rule one adopts. Using the 50% decision rule (as we suggested for the summary of the evidence), two outcomes can occur:

**A)** if the overall score for a given patient is greater than 50% (0.5) and its 95% CI does not cross the 50% decision rule value, one can reasonably conclude that the available evidence supports the hypothesis that this specific profile is associated with sensitivity (or resistance, depending on the "direction" of the overall score) to the treatment under evaluation;

**B)** if the overall score for a given patient either is lower than (or equal to) 50% (0.5) or its 95% CI crosses the 50% decision rule value, one can reasonably conclude that there is not enough evidence linking this specific profile to the responsiveness (or refractoriness) to the treatment under evaluation.

A formal comparison between the calculated overall score (OS) and the 50% (0.5) decision rule value can be made using a Z-test, according to the following formula:

$$Z = (OS - 0.5) / SE$$

where OS and SE are defined as above reported. For a two-tailed test, the P-value is given by:

$$P\text{-value} = 2 [1 - \Phi (|Z|)]$$

where  $\Phi (|Z|)$  = standard normal cumulative distribution.

Of course, the decision rule value (0.5) can be shifted up or down so to make it more or less stringent respectively, thus rendering more or less conservative the conclusion on the relationship between the patient's profile and the response to treatment.

4) If the above procedure is performed for more than one treatment (i.e., the patient's molecular profile is matched with more than one therapeutic agent), it is also possible to create a drug rank based on the overall score obtained for each drug as above outlined. A formal comparison between two overall scores (e.g.,  $OS_a$  and  $OS_b$ ) relative to the matching of the patient's profile with drug A and drug B can be computed using a Z-test, according to the following formula:

$$Z = (OS_a - OS_b) / SE_{a-b}$$

where:

$$SE_{a-b} = \sqrt{(OV_a + OV_b)}$$

$OV_a$  : variance of the overall score for the matching of patient's profile with drug A

$OV_b$  : variance of the overall score for the matching of patient's profile with drug B

For a two-tailed test, the P-value can be calculated using the following formula:

$$P\text{-value} = 2 [1 - \Phi (|Z|)]$$

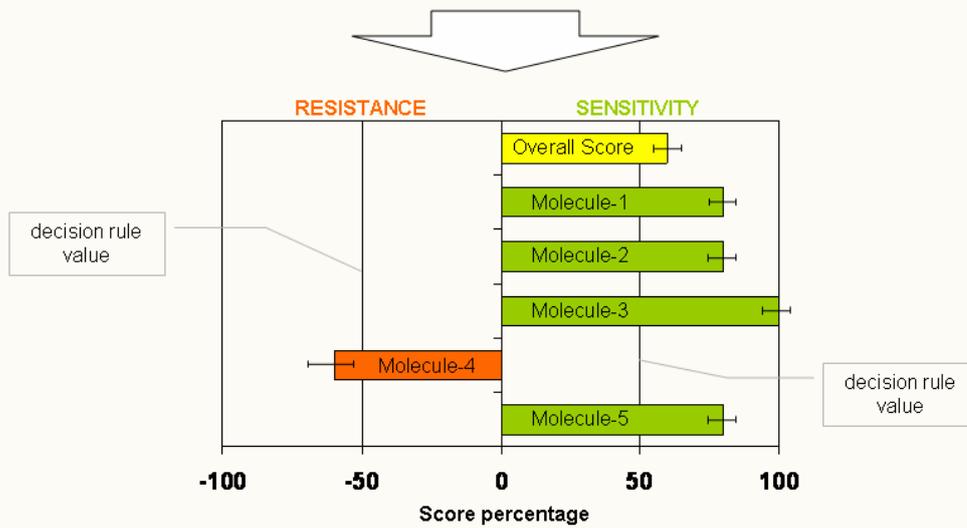
where  $\Phi (|Z|)$  = standard normal cumulative distribution.

Of course, the same procedure can be used to match the patient's molecular profile with the available evidence regarding drug/treatment toxicity.

Finally, the synthesis of the evidence regarding synergisms between drugs/treatments can be used to explore the hypothesis that a combination regimen can increase the likelihood of tumor response, which is of special value in case the patient's molecular profile results incompatible (based on the available evidence) with the response to the drugs/treatments tested.

**Figure 4**

	State	Prevalent hypothesis	Patient profile	Score percentage
<b>Molecule-1</b>	Mutated	Sensitivity to Drug-Y	Mutated	(+)80%
<b>Molecule-2</b>	Overexpressed	Resistance to Drug-Y	Downregulated	(+)80%
<b>Molecule-3</b>	Methylated (promoter)	Sensitivity to Drug-Y	Methylated (promoter)	(+)100%
<b>Molecule-4</b>	Mutated	Sensitivity to Drug-Y	Wild type	(-)60%
<b>Molecule-5</b>	Downregulated	Resistance to Drug-Y	Overexpressed	(+)90%



#### Figure 4

A scheme of the drug ranking system to match the patient's molecular profile with the available scientific evidence regarding the relationship of sensitivity/resistance between a set of molecules (each in a specific state) and a given drug. After identifying the prevalent hypothesis (along with its score percentage) for each molecule according to the evidence score method (see text and Figure 3 for more details), the same molecules (and their state) are tested in the tumor of a patient. Each molecule is said to be concordant (positive sign) or discordant (negative sign) according to whether the molecule state found in the patient's tumor is identical or opposite to the state reported in the literature, respectively. Then, a weighted mean of the score percentages is calculated to obtain the overall score for the patient. In this example, the overall score indicates that on average 60% of the available evidence (that is, 60% of the published evidence rated according to the experimental model used to generate the findings reported in each article) is in favor of the hypothesis that the patient's molecular profile is associated with responsiveness to drug-*Y*.

To be defined as "sensitive" (or "resistant"), a molecular profile must be characterized by an overall score with a lower bound of its 95% confidence interval that does not cross the decision rule value (+50% or -50%, respectively).

The same method can be used to assess whether the available evidence supports the hypothesis that a molecular profile is associated with higher/lower toxicity for a given drug-*Y* (see text for more details).

## Intended use of TTD

The TTD and the data interpretation model above described are only intended for research use.

In fact, since the drug ranking system above described is based on a theoretical model, it is important to remember that it should only be used to generate hypotheses and not to make clinical decisions.

In other words, the findings obtained from the above presented model should only be used a posteriori (after the patients has been treated with a regimen chosen independently of the model results) in order to determine the actual performance of the model itself.

Only this validation of the model will allow to verify whether our theoretical computations are accurate enough to be clinically valuable, and thus to propose the implementation of the model in the routine clinical setting for choosing the therapeutic regimen most likely to benefit the single patient under evaluation.

Moreover, it should be clear that scoring the hypotheses reported in the literature cannot replace in any way the standard rules of research, including meta-analysis methods to pool effect sizes of therapeutic interventions as well as clinical phases of treatment evaluation.

The model presented above can only speed up the identification of the most promising targeted therapy hypotheses that must be then tested on the clinical ground according to the well known rules of the clinical investigation.