UNIVERSITÀ DEGLI STUDI DI PADOVA

University of Padova

Department of Pharmaceutical and Pharmacological Sciences

PhD course in Pharmacological Sciences

Curriculum in molecular and cellular pharmacology

Cycle XXIX

INNOVATIVE STRATEGIES FOR TAILORING THERAPY IN CANCER PATIENT

PHARMACOGENETICS AND THERAPY PERSONALIZATION IN METASTATIC COLORECTAL CANCER PATIENTS TREATED WITH IRINOTECAN

Ph.D. COURSE COORDINATOR: Prof. Piero Maestrelli

INTERNAL SUPERVISOR: Prof. Pietro Giusti

EXTERNAL SUPERVISOR: Giuseppe Toffoli, MD

TUTOR: Erika Cecchin, Ph.D.

PhD STUDENT: Rossana Roncato

Alla mia famiglia

Questo lavoro di tesi è stato svolto presso la scuola di dottorato in Scienze Farmacologiche, indirizzo Farmacologia Cellulare e Molecolare dell'Università degli Studi di Padova, diretta dal Prof. Pietro Giusti in collaborazione con la Struttura Operativa Complessa di Farmacologia Sperimentale e Clinica del Centro di Riferimento Oncologico di Aviano (Istituto di ricerca e cura a carattere scientifico), diretta dal Dott. Giuseppe Toffoli.

Contents

Abstract

Pharmacogenetics focuses on inter-subject variation in drug therapeutic effects and toxicity depending on genetic polymorphisms. Irinotecan and fluoropyrimidine, currently used in cancer chemotherapy, are characterized by a sometimes unpredictably severe toxicity. Pharmacogenomics was largely applied in the last years to the irinotecan-based colorectal cancer (CRC) treatment personalization with limited data regarding validated marker of severe toxicity.

In the first part of my thesis, I have been focusing on the investigation of innovative pharmacogenetic markers of neutropenia or gastrointestinal toxicity irinotecan-related using the "tagging polymorphisms (SNPs)" (TagSNPs) approach. Since therapeutic implications of cancer-related inflammation have gained great attention in recent years, innovative prospects for the optimization of tailored therapy arose. Two hundred and fifty metastatic CRC patients, homogeneously treated with an irinotecan-including regimen (FOLFIRI), have been collected retrospectively for this study. Clinical parameters of toxicity (by NCI-CTC scale) and response to the therapy (by WHO criteria) were monitored all along the study. They were genotyped for 246 htSNPs characterizing 22 transcriptional regulators and cytokines inflammation-related genes; positive findings were replicated in a cohort of 167 metastatic CRC patients receiving FOLFIRI-based therapy. One polymorphism (rs1053004) in STAT-3 gene resulted predictive of severe GI toxicity in both discovery and replication cohort with a protective effect toward the risk of developing grade 3-4 events (OR=0.51 CI=0.27-0.99 p=0.045; OR=0.38 CI=0.15-0.95 p=0.038, respectively). Additional variants in NRs genes, especially HNF4α and VDR, although not validated, were suggested to contribute to determining the risk of developing neutropenia and GI toxicity. Preliminary pharmacokinetic data supported the observed genotype/phenotype clinical associations. A validated contribution of STAT-3 rs1053004 in determining GI toxicity risk after FOLFIRI therapy was pointed out. Further potential predictive markers of irinotecan-related toxicity were suggested. These findings could represent a further step towards personalized FOLFIRI therapy.

UGT1A1*28 polymorphism has been demonstrated in the last years to have an impact on irinotecan pharmacokinetics and toxicity to the treatment. Although, the adoption of a pre-emptive UGT1A1*28 genotyping to increase irinotecan safety and to better characterize patient "Diagnosis Related Groups", for therapy reimbursement purposes in clinical practice, is still limited. The second part of my thesis aimed to estimate the effect of UGT1A1*28 on the costs associated with irinotecan-related toxicity. A retrospective analysis of the costs of toxicity management was conducted on a subset of the

aforementioned population of 250 mCRC patients. 243 mCRC patients treated with FOLFIRI have been genotyped for UGT1A1*28. The mean predicted cost per patient was higher for *1/*28 (1,119€) and *28/*28 (4,886€), as compared to *1/*1 (812€) (P<0.001). This is consistent with a different grade 4 toxicity profile among the three groups of patients, and a higher frequency of costly interventions like hospitalization among patients with the *28 allele.

The aim of the third part of my thesis consisted of evaluating the implementation of the routine application of prospective *DPYD* risk variants and UGT1A1*28 screening at the National Cancer Center CRO of Aviano. A Pharmacogenetic implementation infrastructure has been set-up starting from January 2014 for the prevention of irinotecan (*UGT1A1*28* rs8175347) and/or fluoropyrimidine (*DPYD* rs3918290, rs55886062, rs67376798)-associated toxicity in the clinical routine of the National Cancer Center CRO of Aviano. Genotyping was performed by PCR-based methods, such as pyrosequencing, Sanger sequencing, and fragment analysis. A digital Pharmacogenomic report including the dose-adjustment recommended according to the published pharmacogenetics guidelines will finally be embedded in patients' clinical record and ultimately made available to the medical personnel. From September 2011 to September 2016, a total of 393 patients were genotyped for such variants at CRO-Aviano. Three hundred and eighty-six out of 393 patients were screened for at least one *DPYD* variants and 40 for *UGT1A1*28*. Of these patients, 9 patients (2.58%) were found to carry at least one *DPYD* variants, and two patients (5.00%) were found to carry two *28 risk alleles for *UGT1A1*. Moreover, twenty-three patients out of 393 (5.85%) were referred for toxicity from the CRO-Aviano oncologists.

In conclusion, in this work of thesis, interesting molecular markers with a predictive value on pharmacokinetics and pharmacodynamics of irinotecan were described. A possible application of these parameters in the clinical practice will be useful to design a tailored irinotecan dosing based on peculiar characteristics of each patient. In addition to the prevention of severe toxicity, pre-treatment UGT1A1*28 genotyping should be considered to save economic resources related to the management of irinotecan-related toxicities and for innovative reimbursement strategies. Plus, the implementation of pre-emptive pharmacogenetics tests is now part of a European Project (U-PGx) with the aim of providing the final proof of pharmacogenetics efficacy in increasing drug safety when fully integrated into the clinical practice.

Riassunto

I pazienti oncologici non sempre rispondono analogamente al medesimo trattamento farmacologico ed è il motivo per cui la somministrazione della stessa dose di un farmaco chemioterapico in una popolazione omogenea di pazienti può implicare la manifestazione di differenti risposte e tossicità. Tale variabilità intersoggettiva può essere determinata da interazioni complesse fra componenti fisiologiche, ambientali e fattori genetici individuali. Per valutare la correlazione tra genotipo e fenotipo si rendono necessari studi retrospettivi e prospettici per stabilire se alcuni polimorfismi (SNP) possano essere marcatori di tossicità od efficacia. Una volta stabilita la validità clinica di tali marcatori, la loro implementazione nella clinica può risultare un processo altrettanto ostico e articolato, così come la definizione della rimborsabilità di tale test genetico all'interno della Sanità Pubblica Italiana.

La presente tesi di dottorato descrive la mia attività di ricerca che si è diversificata in vari aspetti della personalizzazione della terapia nel paziente oncologico durante i tre anni di dottorato, in particolar modo focalizzandosi sul tumore al colon retto (CRC) metastatico. Nella prima parte della tesi mi sono occupata di investigare l'esistenza di innovativi marker farmacogenetici predittivi di neutropenia e tossicità gastrointestinale (GI) utilizzando l'approccio del "tagging polymorphisms (SNP)" (TagSNPs). Essendo ben riconosciuto il ruolo dell'infiammazione nello sviluppo del tumore, negli ultimi anni si sono aperte innovative prospettive di ottimizzazione della terapia in questo campo. Dallo studio di 250 pazienti caucasici omogeneamente trattati con FOLFIRI in prima linea per CRC metastatico sono stati genotipizzati 246 htSNPs in 22 geni regolatori della trascrizione e citochine legate all'infiammazione. Un polimorfismo nel gene di STAT-3 è risultato predittivo di tossicità GI severa sia nella coorte esplorativa sia in quella di validazione con un effetto protettivo nei confronti degli effetti tossici di grado 3 e 4.

A seguire, mi sono occupata della valutazione dell'effettiva utilità clinica del test genetico per UGT1A1*28 misurando l'effetto di tale polimorfismo sui costi associati alla tossicità da irinotecano all'interno dell'Istituto CRO di Aviano essendone già ampliamente riconosciuta la validità clinica. Su un sottogruppo di 243 pazienti della casistica descritta nel paragrafo precedente, è stata condotta un'analisi retrospettiva dei costi di gestione della tossicità in relazione al genotipo UGT1A1*28. Il costo medio previsto per paziente si è rivelato essere superiore per pazienti con genotipo *1/*28 (1.119 €), e *28/*28 (4.886 €), rispetto a quelli con genotipo *1/*1 (812 €) ($p < 0.001$).

Nell'ultima parte di questa tesi ho descritto il mio contributo nella messa a punto dei processi per istituire all'interno della struttura ospedaliera del CRO di Aviano il test preventivo di alcuni riconosciuti marcatori genetici di tossicità in principalmente due geni: *DPYD* (rs3918290, rs55886062, rs67376798) e *UGT1A1*28* (r8175347) per tutti i pazienti oncologici eleggibili a trattamento con fluoropirimidine o irinotecano. Le peculiarità dell'infrastruttura e le sue normative sono state prese in considerazione per una efficace integrazione del test genetico pre-trattamento nel flusso di lavoro clinico dell'ospedale. Il processo di diagnostica farmacogenetica inizia con una prescrizione del test farmacogenetico in formato digitale da parte dell'oncologo, successivamente personale medico formato si occupa di effettuare il prelievo e di farlo pervenire presso l'unità di Farmacologia Sperimentale e Clinica, dove personale tecnico altrettanto specializzato effettuerà l'analisi genetica e produrrà un referto seguendo un protocollo approvato ISO-9001. Il referto si costituisce di due parti: una tecnica e una clinica. Successivamente alla firma in digitale del personale tecnico per la validazione dell'analisi genetica effettuata, la firma digitale da parte di personale clinico assieme ad una interpretazione dei dati viene apportata. L'interpretazione dei dati sarà effettuata sulla base delle raccomandazioni di dosaggio farmacogenetiche nazionali ed internazionali disponibili.

La realizzazione di test di farmacogenetica preventivi è ora parte di un progetto europeo (U-PGx), che ha come obiettivo ultimo quello di fornire la prova definitiva dell'efficacia dell'approccio farmacogenetico quando completamente integrato nella pratica clinica.

1. Background

1.1. Pharmacogenetics and cancer chemotherapy

Every day millions of people take drugs that will not help them: numerous studies demonstrate that the obtained response rates can deeply vary among different therapeutic classes. For instance, the 75% of patients treated with anticancer drugs seems to have no benefits from the treatment, and unfortunately for most of them, it is ascribable to an inappropriate drug dosage (Figure 1). In addition, 2.2 million adverse drug reactions occur each year in the US, including more than 100'000 deaths [1].

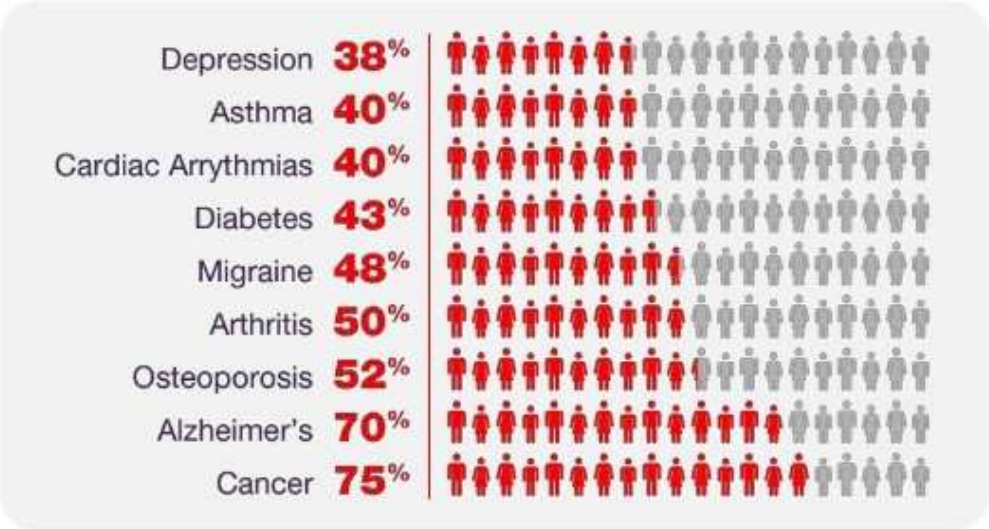


Figure 1. Average percentage of the patient population for which a particular drug in a class is ineffective. From Spear et al., 2001 Clinical application of pharmacogenetics *Trends in Molecular Medicine.*

These data fueled the debate in the scientific community and dramatically changed the approach to the drug administration, moving from "one dose fits all" model to the introduction of the "personalized medicine" model. The National Cancer Institute (NCI) has defined personalized medicine "[…]as a form of medicine that uses information about a person's genes, proteins and environment to prevent, diagnose and treat disease" [2]. Another definition of personalized medicine has been given by the US President's Council of Advisors on Science and Technology as "[…] the tailoring of medical treatment to the individual characteristics of each patient; to classify individuals into subpopulations that differ in their susceptibility to a particular disease or their response to a specific treatment so that preventive or therapeutic interventions can then be concentrated on those who will benefit, sparing expense and side effects for those who will not" [3]. As these descriptions suggested, this issue is extremely broad and

different points of view can highlight several facets of the same phenomenon.

Oncology is one of the medical specialty that strongly needs the adoption of the personalized medicine due to the complexity and to the lethality of the disease. Oncology practice continually struggles indeed with the challenge of matching the right therapeutic regimen to the right patient, balancing relative benefit with risk to achieve the most successful outcome. In 2016 about 1,685,210 new cancer cases are expected to be diagnosed in the US, and about 595,690 Americans are expected to die of cancer [4]. In January 2015, the American President Barack Obama founded with 215 million dollars a National Precision Medicine Initiative, with the aim of promoting the introduction of personalized medicine into the clinical practice, especially focusing on oncology and genetics. However, As said before, a large part of patients' treatment is indeed not only toxic but also ineffective, and it has been estimated that in only 25% of patients the expected response is achieved [5]. Several examples in literature report wide range of responses and toxicities (which in some cases can result even lethal) to the administration of the same dose of the same antineoplastic drug in patients affected by the same cancer disease. The new hurdle, in this century, is to find the key genetic pathogenic variants that drive specifically cancer growth in order to optimize the use of the therapies already used in the clinical practice and to develop new personalized drugs (figure 2).

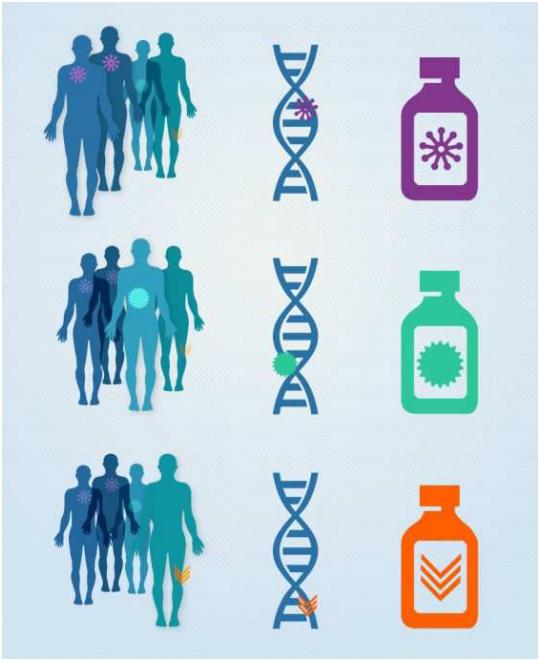


Figure 2. Using the genetic changes in a patient's tumor to determine their treatment is known as precision medicine. From www.cancergov.com

More broadly, “personalized medicine” may be referred to as the tailoring of medical treatment to the individual characteristics, needs and preferences of a patient during all stages of care, including prevention, diagnosis, treatment, and follow-up. At present, the dose of chemotherapeutic agents is generally adjusted by body surface area, but this is not sufficient to overcome inter-individual differences in drug disposition [6,7]. Moreover, most anticancer drugs are characterized by narrow therapeutic indexes [8], most impacted by inter-individual variability which could result in severe, even life-threatening toxicities, and response variability. Poor tolerability and therapeutic failure put in danger the adherence of the cancer patient to therapy as well as its possibility to get cured. An inevitable consequence of treatment interruption is relapsed. The explanation for this variability can be found in age, gender, environmental factors, hepatic and renal functions, comorbidities and co-medications and genetic determinants.

Personalized cancer therapy is considered a key challenge and several genomic technologies have been employed to achieve this goal. They have been directed both to the identification of targets for new drug development that uniquely attack a given tumor (i.e. imatinib mesylate, Gleevec, and trastuzumab, Herceptin) and to determine more focused applications for existing cancer therapeutics, many of which are very effective for subsets of cancer patients. Among these applications is included a genotype-based new dosing approach of existing cancer chemotherapeutics to account for their complex drug metabolism processes.

Pharmacogenomics (PGx) is a peculiar discipline of pharmacology with the aim of studying the role of inherited genetic differences in affecting individual responses to drugs, for overcoming patients’ variability both in terms of therapeutic and adverse effects. PGx is ultimately a promising field to choose the best drug or the best drug combination to achieve the highest efficacy and the minimal toxicity. Genetic profiling prior to pharmaceutical treatment has now started to be performed in order to cluster patients depending on the risk of developing severe toxicity and expected responsiveness to a standard dosage of pharmacological treatment, allowing better-targeted treatments delivery [9].

Genetic variability consists of several biological mechanisms: there are differences in transcription factor activity, gene expression, gene silencing (epigenetics), and benign genetic variants (polymorphisms). Pharmacogenetics intends to identify relationships between drug activity and the genetic variability derived from gene polymorphisms.

When referring to genetic influence on the treatment outcome, there is still some confusion between

the terms “pharmacogenetics” and “pharmacogenomics”. Indeed, they tend to be used interchangeably, and a precise consensus definition of either term remains equivocal. On the one hand, “pharmacogenetics” focuses on the association of candidate genes benign variants with drug activity, while “pharmacogenomics” considers the entire genome, through the broader application of new genomic technologies. On the other hand, in oncology “pharmacogenetics” is often regarded as concerning the germ-line polymorphisms and individual patient’s features, while “pharmacogenomics” usually refers to those of the tumor.

1.1.1 Polymorphisms

The vast majority of the human DNA sequence is identical among individuals, except for minor changes called polymorphisms constituted by nucleotide substitutions, deletions and insertions, repeats, gene copy number variations and sometimes more significant rearrangements. Polymorphisms are are often referred to as benign variants. Structure, expression, stability and activity of the proteins encoded by genes can be affected by DNA polymorphisms. Leading to minor phenotypic variations, benign variants explain the inter-individual differences, from eyes or hair color to disease susceptibility or drug sensitivity: this is why they present a major interest by a clinical point of view.

Substitution, deletion or insertion of nucleotides can arise from errors occurring during DNA replication or lesions induced by mutagenic agents which may lead to the replacement of a nucleotide by another one (substitution), to the loss (deletion) or to the addition of a nucleotide (insertion). When the coding sequence of a gene is affected by benign variants, the encoded protein may bear structural alterations, which possibly lead to its instability, enhancement or reduction in its activity or the loss of its functionality.

“Single Nucleotide Polymorphisms” or SNPs are one of the most common forms of genetic variations (>90%) and are characterized by the involvement of only one nucleotide (by substitution, insertion or deletion).

Other common types of polymorphisms concern the number of short series of nucleotides repeats (minisatellite or microsatellite), called Variable Number of Tandem Repeats (VNTR), or even the number of copies of a gene (CNV, Copy Number Variations). These polymorphisms are more likely to play a major role in the level of the mRNA and the protein produced.

Conventionally, an SNP is defined as a nucleotide variation having an allele frequency greater than 1%, whereas, when the frequency is lower, the genetic variation is indicated as a mutation. Mutations by substitution even if they are biochemically identical to SNPs (both involving the replacement of a nucleotide by another one), have a different meaning: they are rare and deleterious while SNPs are common and non-deleterious events. The frequency of benign variants is generally higher in introns than in exons.

Polymorphic genes usually have not essential functions for cell life: polymorphisms in genes involved in vital processes, if deleterious, would be rapidly eliminated by natural selection. In contrast, genes coding for enzymes involved in the metabolism of xenobiotics and drugs are often polymorphic since they have no major consequences on cell viability and so they are not eliminated from the genome. Nevertheless, an effect of these benign variants may arise in some special situations such as contact with xenobiotics or DNA damaging agents in which a polymorphic variant could allow some flexibility in front of environmental variations. Pharmacogenetics tries to find out the correlation between drug pharmacology and patient's genotype that results in an intermediate, low or even lack of activity of certain proteins, such as metabolizing enzymes, cellular receptors or other target proteins [10].

Several examples of genetic polymorphisms possibly affecting the pharmacodynamics of anticancer drugs have been described. They concern enzymes of phase I or II metabolism such as thiopurine S-methyltransferase, involved in metabolism of 6-mercaptopurine [11]; dihydropyrimidine dehydrogenase, which modulates 5-fluorouracil activity; methylenetetrahydrofolate reductase which might change patient susceptibility to methotrexate [12]; thymidylate synthase associated with pharmacoresistance to fluoropyrimidine derivatives [13]; glutathione-S-transferase responsible for detoxification of many anticancer drugs [14]; and N-acetyltransferase involved in the metabolism of amonafide [15]. Polymorphisms have been described also for transporter proteins belonging to the ATP-binding cassette (ABC) export pumps which are involved in detoxification from xenobiotics as well as in pharmacoresistance [16–18].

Heterogeneous systems are used for SNPs nomenclature, but the universal accepted one is that using the rs code (reference sequence). This system is the only one allowing the precise identification of a polymorphic variation within the most common used genetic databases (NCBI, Hap Map, SNP500 Cancer, 1000 genomes, etc.). If the SNP leads to the replacement of an amino acid by another one (non-synonymous SNP), usually the SNP nomenclature is completed by the name and the position of the amino-acid that is replaced in the protein, followed by the name of the novel amino acid (*i.e.*, L432V or Leu432Val).

1.1.2 Types of polymorphisms

SNPs lying in the coding region can be classified as synonymous or silent when there is no change of the amino acid encoded, still having potentially functional consequences. Indeed, the replacement of a nucleotide by another one may produce a different three-dimensional mRNA structure or a different stability of the variant transcript, and this could lead to a different translation rate into protein, therefore in a different amount of the synthesized protein. Missense are called the SNPs where one amino acid is replaced by another one. The potential consequences will differ basing upon the shared chemical properties of the substituted amino acid with the wild-type one. For example, if a valine is replaced by leucine or a glutamic acid replaced by an aspartic acid, usually the produced effect is modest. On the contrary, if a hydroxylated amino acid residue (potential substrate of a protein kinase) is replaced, for instance, by an aliphatic amino acid residue, an acidic by a basic one, functional effects may be more important.

An SNP may be nonsense if the polymorphism implies the occurrence of a stop codon leading to a truncated protein. The same effect could be obtained if the alteration occurs within a splicing site. On the contrary, a frameshift polymorphism takes place when insertions and deletions alter the reading frame, therefore generating completely different codons responsible for changing the entire downstream sequence of the protein. These variations have a high probability to induce the production of a truncated, totally inactive proteins when leading to the generation of early stop codons.

When SNPs occur at the splicing site the enzymes responsible for mRNA maturation no longer recognize it, and an abnormal amount of protein is produced. The resulting protein could either include an intronic aberrant sequence rapidly concluded by a stop codon or missing the portion encoded by the absent exon. The SNP rs776746 of the CYP3A5 gene is an example of a splicing site polymorphism, leading to the complete absence of the protein in about 90% of Caucasian subjects [19].

Many polymorphisms lie in introns and 5’ or 3’ UTR, which are non-coding regions, with possible phenotypic effect when these sequences have regulatory functions (promoters, silencers, enhancers, micro-RNA binding sites or micro-RNA genes).

The polymorphism is called Short Tandem Repeats polymorphism (STR) when variable number repeats of the CA or TA dinucleotide within a microsatellite that may be involved in gene regulation are ranging from 2 to 5.

The best known STR example is represented by 7 TA repeats in the promoter of the UGT1A1 gene

instead of 6 resulting in a 50% decrease in gene transcription rate, with hyperbilirubinemia and increased risk of irinotecan toxicity as major phenotypic consequences [20].

Finally, a polymorphism could also concern an entire gene: unequal chromosomal recombination may produce a variable copy number of that gene. Usually, if the gene copy number rises, the encoded protein is overexpressed: this is the case, for instance, of the CYP2D6 gene, which is duplicated (or more) in about 5% of Caucasian subjects [21].

1.2. FOLFIRI regimen

The standard of care for the first-line treatment of metastatic colorectal cancer (mCRC) is a combination of irinotecan, 5-fluorouracil (5-FU) and leucovorin (LV) called FOLFIRI regimen [22].

1.2.1. Irinotecan

Irinotecan Pharmacology and metabolism

Irinotecan (7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin, CPT-11, Camptosar) is a prodrug of the semisynthetic Camptothecin [23,24]. It acts as a topoisomerase I (Topo I) inhibitor, and it is approved worldwide for the treatment of mCRC [25]. Irinotecan is also active in ovarian, lung, breast, pancreatic and cervix cancer [26]. Two combination regimens of irinotecan, 5-fluorouracil and leucovorin have been in widespread use and approved for advanced CRC: IFL (irinotecan, 5-FU, and LV) is the USA schedule in which 5-FU bolus is used16, and the Europe FOLFIRI with 5-FU infusion [27]. These two studies demonstrated, in terms of overall response and survival, the superiority of irinotecan combination with 5-FU/LV compared to 5-FU/LV or irinotecan alone [27,28].

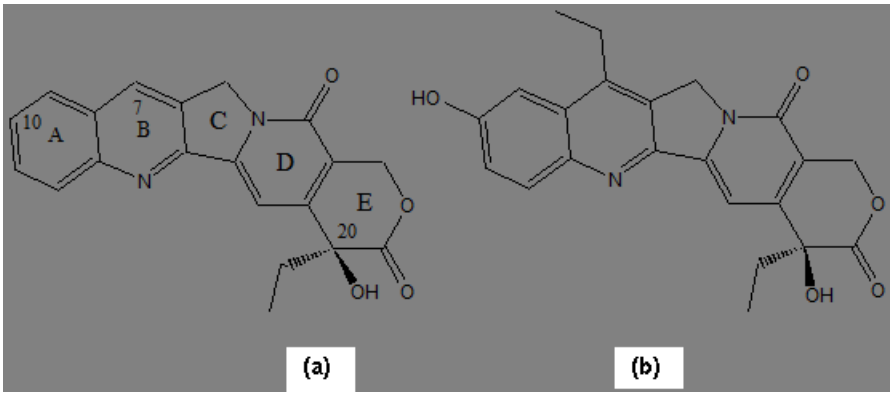


Figure 3. Chemical structures: (a) camptothecin, (b) semisynthetic camptothecin Topo I inhibitor SN-38 and (c) irinotecan prodrug derivative

Irinotecan is originated from the synthetic efforts of the Yakult-Honsha researchers originally oriented to make more feasible the clinical use of camptothecin, an alkaloid first isolated and characterized by Wall and co-workers from the wood of *Camptotheca acuminata* (*Nyssaceae*) with remarkable anti-tumor activity [29].The anti-cancer effect of Camptothecin, whose chemical structure is displayed in figure 3a, from where irinotecan derived it is ascribable to its exclusive intracellular target: Topo I [30,31], an enzyme required for swiveling and relaxation of DNA during molecular events such as replication and transcription [32]. In the absence of drug, Topo I unwinds supercoiled DNA ahead of the active

transcription/translation sites named replicating forks. Topo I first binds DNA to form a non-covalent complex of double-stranded DNA and Topo I, described as the "non-cleavable complex". This is in equilibrium with the so-called "cleavable complex", which forms when Topo I creates a transient break in one DNA strand and at the same time becomes covalently bound to the 3'-phosphoryl end of the nicked nucleic acid by phenolic hydroxyl group of a tyrosine (Tyr723 in human Topo I) [33] (figure 4). The whole DNA strand is allowed to unwind once and to pass through the break site, before Topo I re-ligates the cleaved DNA and re-establishes the double stranded configuration. Camptothecin is thought to interact with the “cleavable complex” and stabilize it, thus preventing religation of the nicked DNA strand and retarding/arresting the motion of Topo I along DNA [34]. Consequently, the advancing DNA polymerases operating in the replicating fork soon “collide” with the stabilized cleavable complex and create an irreparable double-stranded break in DNA that leads to a cascade of events culminating in cell death [35].

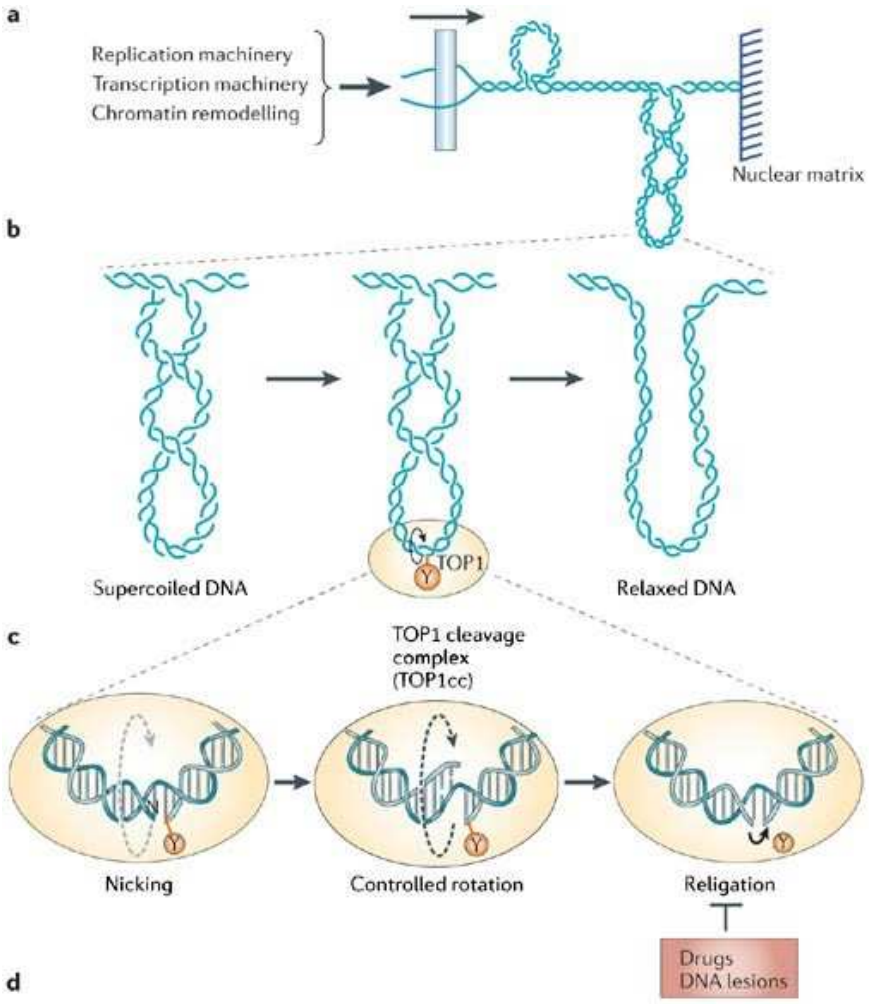


Figure 4. Relaxation of DNA supercoiling by TOP1-mediated DNA cleavage complexes, and the trapping of TOP1 cleavage complexes by drugs. From Pommier *Nature Reviews Cancer* 6, 789–802 (October 2006)

Irinotecan is activated in vivo to the more potent active metabolite SN-38, an inhibitor of Topo I, leading to his high therapeutic potential (figure 3b). After its initial approval as a second-line therapy for advanced CRC refractory to fluoropyrimidine, its high pharmacological activity as chemotherapeutic accelerate its use in clinical trials. Now Irinotecan has been evaluated for the treatment of other malignancies such as non-small cell lung cancer, breast carcinoma and cervical cancer [36] as well as metastatic esophageal carcinoma.

The terminal half-life of irinotecan is as long (7.9 to 14.2 hours) as that of SN-38 (13.0 to 13.8 hours). Both of them have a linear pharmacokinetics with increasing doses. Thus dose and regimen do not affect the total body clearance (14.3 to 15.3 l/h/m^2) and the volume of distribution (157 l/m^2) [37]. The main inactive metabolite of irinotecan is the glucuronide of SN-38, but at least 16 additional metabolites have been characterized in the plasma and blood [38].

A high inter-patient variability has been reported in term of toxicity to irinotecan, either by using irinotecan/5-FU infusion or irinotecan/5-FU bolus schedules. Several studies demonstrated that neutropenic fever is unusual, approximately 3% [39]. Nausea or acute cholinergic symptoms are more common with a tri-weekly schedule [40]. Diarrhea is common and can generally be well controlled with the aggressive use of loperamide or atropine. The variability occurring in toxicity development could be related to differential plasma levels of the active metabolite SN38 [41] among patients, as a result of the complex interplay of several metabolic pathways influenced by many factors, including genetic differences of the enzymes involved in drug metabolism [42] (figure 5).

In vivo activation to SN-38 is required for an antitumor effect by hydrolysis of the carbamate bound by carboxylesterase (CES) which are present prevalently in the liver and in many other normal tissues as well as tumors [43,44]. Two similar human liver microsomal CES, CES1 and CES2 were suggested to play a role in the activation of irinotecan [45]. Secondly, oxidative metabolism of irinotecan involves exclusively the CYP3A subfamily and is responsible for irinotecan inactivation [46]. There is competition for substrate between CYP3A enzymes, which inactivate irinotecan, and CES, which activates irinotecan. CYP3A4 and CYP3A5 are involved in these oxidative processes, but the role of CYP3A5 remains marginal and controversial compared to CYP3A4 isoform [38]. In patients receiving irinotecan, diverse oxidative metabolites have been observed in plasma, urine, and bile [71]. The most abundant are the 7-ethyl-10[4-N(5-aminopentanoic-acid)-1-piperidino] carbonyloxycamptothecin (APC) and the 7-ethyl-10-(4-amino-1-

piperidino) carbonyloxycamptothecin (NPC), which are supposed to be the result of further oxidation of the not fully characterized hydroxylated products of the intermediate oxidative reactions. Both NPC and APC are characterized by a lack of cytotoxicity, but NPC appears to be functionally important because it may be a substrate of CES and can be transformed into SN-38 [47]. Other de-ethylated and hydroxylated metabolites have been identified in the urine of some patients after irinotecan administration.

The irinotecan detoxification occurs via glucuronidation, mainly catalyzed by the uridine diphosphate glucuronosyl transferase isoform 1 (UGT1A1), but also by the isoform 7 (UGT1A7) and 9 (UGT1A9). This specific enzyme is responsible for glucuronic acid transfer from the cofactor UDP-glucuronic acid (UDP-G) to SN-38 (forming SN-38-G), thereby inactivating SN-38. UGT1A products are exported outside the cell through the ATP-binding cassette isoform C2 (ABCC2), a membrane protein with substrate specificity. Polymorphic structures of the gene encoding for ABCC2 have been described and associated with a defective expression and functionality [16]. ABCC2 belongs to the ABC transporter genes superfamily, which also includes the ATP-binding cassette isoform B1 (ABCB1) gene, encoding for P-glycoprotein (P-gp). Irinotecan and SN-38 are also eliminated by the cell through the P-gp [48]. Polymorphic structures of the ABCB1 gene have been described [43]. SN-38-G is then disposed of in the bile.

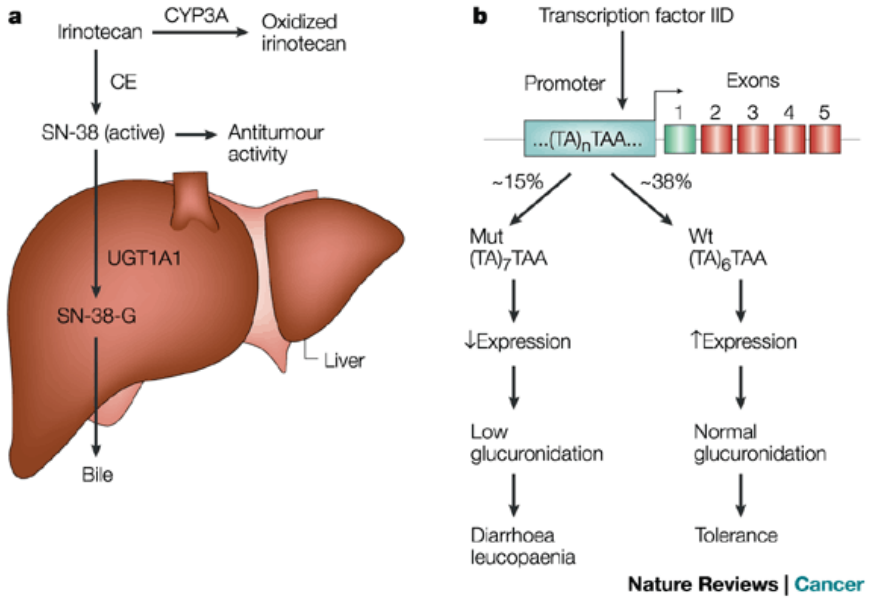


Figure 5. Polymorphisms that affect irinotecan therapy. From Mary V. Relling & Thierry Dervieux, 2001

Pharmacogenomics was widely applied in the last years to the irinotecan-based CRC treatment personalization providing a validated marker of severe toxicity. Significant results demonstrating the predictive role on FOLFIRI toxicity outcome of UDP-glucuronosyl transferase 1A (UGT1As) and ATP-binding cassette (ABC) and solute carrier (SLC) transporters genetic markers, in combination with clinic-demographic features have been published in our group of Experimental and Clinical Pharmacology at CRO-Aviano [49–52]. In particular, irinotecan PGx research has been mainly focusing on the UGT1A1 family, responsible for the conjugation of the active SN-38 to inactive SN-38G. Among the most studied SNPs within these genes, *UGT1A1*28* (rs81753479) SNP is surely one of the most well-known. In the case of low UGT1A1 activity, the accumulation of high levels of the irinotecan active metabolite SN-38 can cause diarrhea and leukopenia [53,54]. These are dose-limiting toxicities, responsible for a significant inter-individual variability [55,56]. The occurrence of severe and occasionally life-threatening complications is often caused by chemotherapy failure impacting negatively to the patients' care.

*1.2.1.1. UGT1A1*28*

The *UGT1A1*28* allele is characterized by seven thymine-adenine (TA) repeats within the promoter region (the binding site for transcription factor IID), as opposed to six that characterizes the wild-type allele (*UGT1A1*1*). These extra repeats impair proper gene transcription, resulting in decreased gene expression by approximately 70% The resulting reduced UGT1A1 prevents glucuronidation of SN-38, so SN-38 accumulates. Thus, patients homozygous or heterozygous for the UGT1A1*28 commonly develop dose limiting severe neutropenia and late diarrhea, and the current US package insert includes a warning for patients presenting homozygosity of *UGT1A1*28* as a risk factor for severe neutropenia [57].

*UGT1A1*28* is the most frequent benign variant in the Caucasian population (about 15% of homozygosity frequency and about 35% of allelic frequency), it is even more widespread in Africans (38.7-42.6% of allelic frequency), it is less prevalent in the Asian population (10.0-16.8 of allelic frequency) [25]. The wild-type allele, *UGT1A1*1*, has six thymine adenine (TA) repeats in the atypical TATA box region of the UGT1A1 promoter. The variant allele, *UGT1A1*28*, consists of seven TA repeats in the $A(TA)_7TAA$ motif [41]. An increase in the number of repeats is associated with a reduced expression of the gene *UGT1A1.*

The data from Innocenti et al[58], in 66 cancer patients, indicated *UGT1A1*28* as strongly associated with severe neutropenia, also in association with other promoter variants (-3156G>A, -3279G>T) in a linkage disequilibrium with *UGT1A1*28*. Finally, Marcuello et al., in 95 CRC patients treated with 4 different

regimens of irinotecan, found a statistically significant relationship between the appearance of severe diarrhea and homozygous *UGT1A1*28* polymorphism but found no relationship between the hematological toxicity and genotype [59]. Thus, cancer patients treated with irinotecan can be classified in carriers of a "high toxicity" genotype that is more likely associated with dose-limiting toxicities, such as severe neutropenia and diarrhea [50,50,58] and "low toxicity" genotype able to tolerate significantly higher doses [51]. Irinotecan indeed has also been a candidate for genotype-driven phase Ib studies. In particular, our group performed a dose-finding study in mCRC patients treated with FOLFIRI regimen and with the *UGT1A1*1/*1* and *UGT1A1*1/*28* genotypes. By dose escalating CPT-11 only in patients without the high-risk *UGT1A1*28/*28* genotype (10% on average in patients of European descent), they demonstrated that the recommended dose of 180 mg/m^2 for CPT-11 in FOLFIRI is considerably lower than the dose that can be tolerated by the non–*UGT1A1*28/*28* patients. In particular, patients with *UGT1A1*1/*1* genotype can safely be treated with a dose of 370 mg/m^2, while the maximum tolerated dose for *UGT1A1*1/*28* is assessed at 310 mg/m^2 [51].

Polymorphic structures with five (*UGT1A1*33*) or eight (*UGT1A1*34*) repeats have been described, in particular among the African population. They are correlated with a sharper variation in enzyme expression levels and with a more evident influence on irinotecan metabolism [60]. Recently Innocenti et al. described new *UGT1A1* polymorphisms, present in the phenobarbital-responsive enhancer module of the gene, in a linkage disequilibrium with *UGT1A1*28*, that demonstrated a role on irinotecan pharmacokinetics and pharmacodynamics [61]. Among them, 3156G>A and 3279G>T resulted the most frequent in the Caucasian population. It is important to point out that the polymorphic structures involved in Gilbert Syndrome, are the same that produce a reduction in the glucuronidation of irinotecan and SN-38. Some polymorphisms, involved in this metabolic defect, have been described, their distribution among the populations depends on the geographic area [55].

1.2.1.2. Innovative markers of irinotecan-related toxicity

The optimization of irinotecan-based therapy administration remains sub-optimal, and other under-explored elements might significantly contribute to determining the likelihood of severe complications after chemotherapy. Still limited are the data regarding markers of severe toxicity.

Therapeutic implications of cancer-related inflammation have gained considerable attention in recent years opening innovative prospects for the optimization of tailored therapy. Accumulating evidence have established a significant link between pro-inflammation state and CRC development, progression, invasion, and metastasis [62,63]. During inflammation, variable drug effects have been described and

attributed, at least in part, to changes in the expression level of ABC/SLC transporters and phase I and II drug-metabolizing enzymes [64–66]. A key role in this phenomenon was ascribable to transcription factors (TFs) such as signal transducers and activators of transcription (STAT-3) and nuclear factor κ B (NFκB1), that are well-known to be activated by pro-inflammatory cytokine-induced signaling pathways [66,67]. More recently, Nuclear Receptors (NRs), a further class of TFs [49,68,69], have emerged as additional crucial regulators of pharmacologically relevant proteins in the presence of cytokines released during inflammation process [49,64]. Specifically, NRs have been recently defined as xenosensors, due to their ability to mediate between the environmental stimuli and gene expression, overall regarding drug transforming genes[68,70]. Indeed, although polymorphic variants represent an important aspect of variability, they cannot explain all inter-individual variability observed in clinical practice. Environmental factors (such as oxidative stress, inflammation and dietary or pharmacological intake of exogenous compounds [64,71–73]) can interact with the genetic background and contribute to a drug pharmacological profile. Our group produced a review aimed at summarizing the state of knowledge about NR involvement in predictions of cancer therapy outcome [69].

NRs were implicated in the regulation of a broad spectrum of key ABC/SLC transporters and DM enzymes. NR activities are significantly influenced by endobiotics and xenobiotics, including chemotherapeutics.

Hence, NR proteins could represent the crucial link between environmental stimuli and the observed alterations in drug metabolism and disposition. A change in NR expression and the presence of polymorphic variants that affect protein functionality could potentially be responsible for differential individual responses to exogenous stimulation. In turn, these responses could confer variability in the systemic bioavailability and local accumulation of anticancer drugs.

PXR, CAR & HNFs

NR is a superfamily of transcription factors that can be categorized into four groups, based on their specific mechanisms of action [74]. In general, an NR binds to its ligand in the cytoplasm, which leads to its translocation to the nucleus. Then, it recognizes and binds its receptor-specific xenobiotic response element or hormone response element. The first group includes steroid hormone receptors, which rearrange to form homodimers before their translocation to the nucleus to bind a target DNA response element. The second group includes key NRs, as PXR and CAR and before their interaction with a specific target DNA sequence in the nucleus form heterodimers with an obligate partner RXR. The third group

includes other NRs, like RXRs and HNFs. They operate without ligands and form homodimers. The fourth and last group comprises NRs that function as monomers. CAR, PXR and HNFs interact with a variety of endogenous compounds (e.g., steroids, bile acid salts) and exogenous compounds (e.g., chemotherapeutic drugs, environmental chemicals), which can act as both agonists and antagonists [71,73,75–77]. The main activity of NRs is to regulate the transcription of target genes by recruiting co-activators or co-repressors, often in response to internal or environmental stimuli (Figure 6). PXR, CAR, and HNFs coordinately regulate the expression of a large proportion of genes in the liver, intestine and other organs Plus, since NRs have different and overlapping substrate specificity, they cooperate in modulating target gene transcription in a complex regulatory network producing produce reciprocal, interconnected crosstalk at the target promoter level (Figure 6).

PXR and CAR are expressed in the main drug excretory and detoxifying organs (liver, intestine, kidney only PXR). They are well-recognized as master regulators of a broad spectrum of Phase I/II enzymes, including CYPs (i.e., CYP3A, CYP2B and CYP2C), GSTs, SULTs, UGT1As, and UGT2Bs; they also regulate ABC/SLC carriers, such as MDR1, MRP1–4, BCRP and OATPs [75,77]. HNFs are localized predominately in the liver. They are among the most important controllers of liver-specific gene expression and hepatic drug metabolism [73,78].

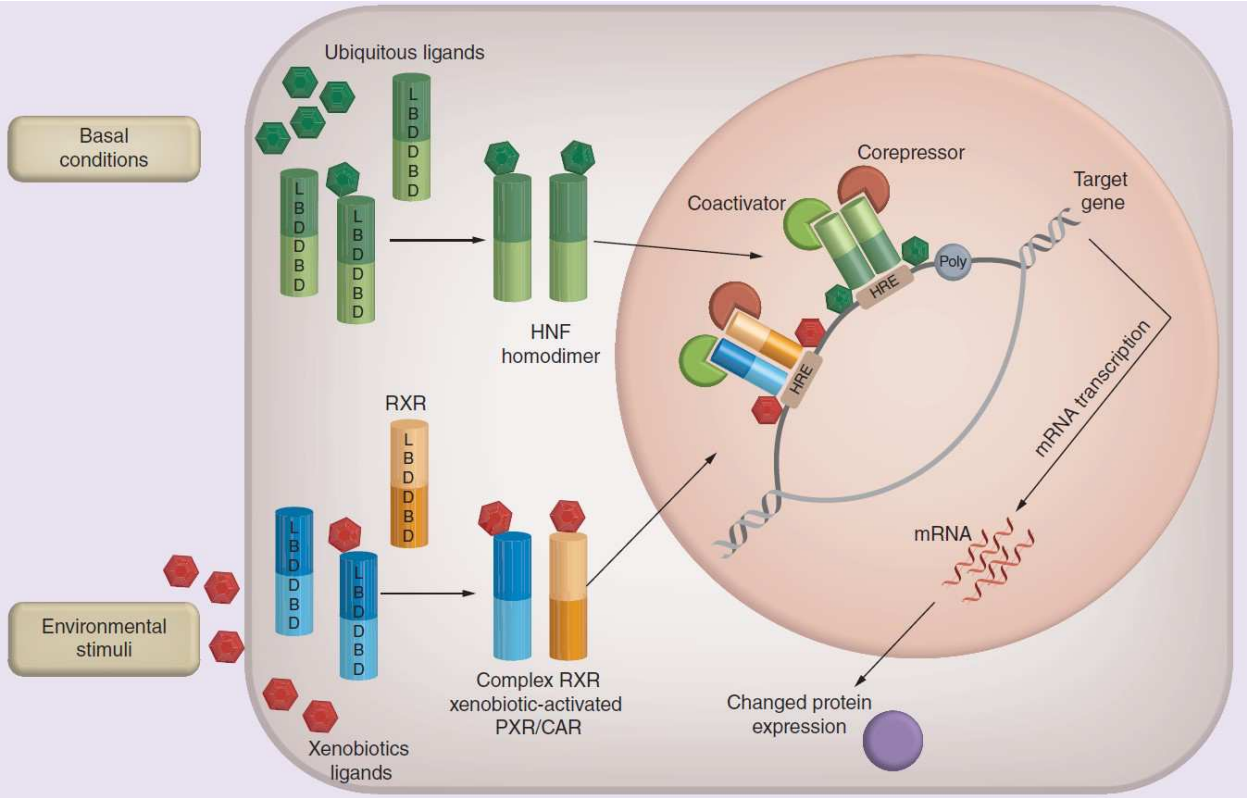


Figure 6. Cooperation between CAR/PXR and HNFs in the regulation of target gene expression. From De Mattia et al, 2016

In basal conditions, HNFs are activated by ubiquitous ligands and are implicated in basal control gene-expression while PXR and CAR enhance gene transcription in response to environmental stimuli (e.g., oxidative stress, proinflammatory cytokines, pharmacological modulators). NR binds to its ligand in the cytoplasm, which leads to its translocation to the nucleus. PXR and CAR form a heterodimer with the obligated partner RXR while HNFs can operate without ligand and act as homodimers. Once translocated to the nucleus, NR recognizes and binds the specific DNA response elements (i.e., HRE) in the promoter region of the target gene and recruits co-activators and co-repressor. The illustration highlights the interaction between CAR/PXR and HNF (i.e., HNF4a) to regulate the expression of some drug metabolism and transporters gene and the reciprocal interconnected crosstalk occurs at the target promoter level [73,75,76,78–80].

Interestingly, the induction of PXR, one the most studied NR, in response to environmental stimuli, was demonstrated to significantly affect the irinotecan metabolism, through the UGT1A-mediated detoxification enhancement as well as the regulation of other key irinotecan metabolic proteins (i.e. CYP3A4, CYP3A5) and transporters (i.e. MDR1, MRP1, MRP2 and BCRP) [81].

NRs markers in cancer therapy

These transcription factors play potentially key roles in the determination of treatment effectiveness and the toxicity profile. PXR, CAR, and HNFs potentially represent optimal markers for realizing a cancer precision medicine. Polymorphic variants in the genes encoding for TFs and NRs were reported (http://www.ncbi.nlm.nih.gov/snp/) and were possibly associated with changes in the transcriptional activity. Hence, an altered TFs and NRs functionality genetically determined could be supposed to impact the regulation of metabolic enzymes and transporters during inflammation stimuli, finally influencing the drug pharmacokinetic and pharmacodynamic profile. Recent data evidenced that PXR genetic markers (i.e. rs10934498, rs3814055, rs1523127, and rs2472677) were associated with SN-38 pharmacokinetic parameters and the risk of hematological toxicity [82].

However, despite the potential clinical relevance, the involvement of transcriptional proteins genetics in the irinotecan treatment modulation, especially in inflammatory diseases as CRC, is still an under-investigated topic.

1.2.2. Fluoropyrimidine

Fluoropyrimidines (FLs) are antimetabolite drugs widely used in the treatment of a range of cancers, including colorectal, breast, head and neck, and stomach cancer. However, despite the acknowledged efficacy of these drugs in the treatment of different solid tumors, the treatment with FLs remains challenging as a result of a considerable inter-patient variability in terms of efficacy and toxicity.

Fluoropyrimidine pharmacology and metabolism

FLs are analogs of the uracil base responsible for the inhibition of the nucleotide synthetic enzyme thymidylate synthase (TS). TS is responsible itself for the de novo synthesis of thymidylate, which is necessary for DNA replication and repair [83]. FLs include 5-FU, CAPE, and tegafur. CAPE is an oral FL that is absorbed unchanged through the gastrointestinal wall and is converted to 5'-deoxy-5-fluorouridine (5'-dFUR) in the liver by the sequential action of CE and cytidine deaminase. Tegafur is another prodrug administered per os which is enzymatically activated in the liver to 5'-hydroxytegafur and subsequently to 5-FU (Figure 7).

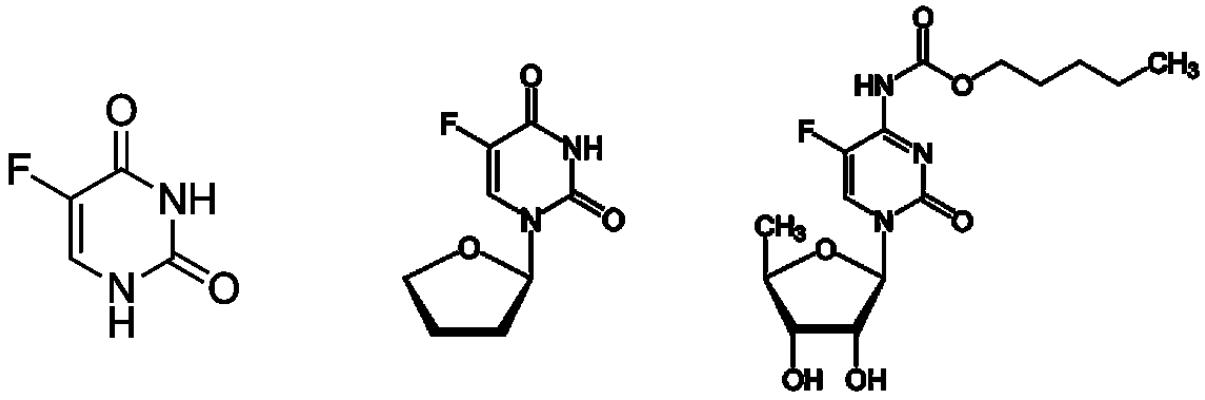


Figure 7. Chemical structures of 5-fluorouracil (a), tegafur (b) and capecitabine (c)

5-FU is an analog of uracil with a fluorine atom at the C-5 position in place of the hydrogen. It can thus rapidly enter inside the cells through the facilitated transport mechanism of uracil [84]. Once inside, it is intracellularly converted into several active metabolites such as fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine triphosphate (FdUTP), and fluorouridine triphosphate (FUTP), leading to RNA synthesis disruption and to the TS inhibition.

Actually, the 5-FU metabolite FdUMP binds to the TS nucleotide-binding site, forms a stable ternary complex with the reduced folate methylene tetrahydrofolate (5,10-CH_2THF), thereby blocking the binding of the normal substrate dUMP and inhibiting dTMP synthesis [85].

Dihydropyrimidine dehydrogenase (DPD) is the rate-limiting enzyme of FL catabolic pathway: inactivation of 5-FU and its prodrugs [86].

Fluoropyrimidine Pharmacogenetics

Approximately 10-30% of patients receiving 5FU or capecitabine experience severe (grade≥3) toxicity, such as diarrhea, mucositis and hand-foot syndrome [86]. The Fluoropyrimidine PGx research mainly focused on the DPD, encoded by the gene *DPYD*, for predictive markers of FL response discovery. Low DPD activity will result in decreased 5-FU inactivation and in the accumulation of high levels of 5-FU active metabolites. Indeed there is a strong correlation between reduced DPD activity and increased risk for severe and potentially fatal toxicity following treatment with a normal dose of 5FU [12,85,87,88]. Toxicity occurred in 73% of *DPYD**2A carriers, compared with 23% of wild-type ones [89].

In the case of DPD-deficient patients, toxicity could be limited by reducing the exposition to FL, performing a genotype-based dose adjustment. 4–5% of the population results DPD deficient and up to date, 167 SNPs altering the DPD aminoacidic sequence have been identified, and many clinical studies have investigated their association with FL-related severe toxicities. SNPs can appear in heterozygous form (one SNP on one allele), homozygous form (two identical SNPs on two alleles) or double heterozygous form (two different SNPs on either one or two alleles, the latter is also called compound heterozygous). Two SNPs on two alleles lead to a larger decrease in DPD enzyme activity, compared with the heterozygous form.

1.2.2.1. DPYD SNPs

*DPYD**2A/IVS14 + 1 G>A

First described by Vreken, the *DPYD* SNP rs3918290 (*DPYD**2A, *DPYD* IVS1411G>A or c.190511G>A) is surely one of the most well-known. Its variant allele frequency has been described to differ among ~0.1 and 1.0% in African-American and Caucasian populations, respectively [90]. *DPYD* rs3918290 is located at the intron edge of exon 14 resulting in a splicing defect responsible for the skipping of the entire exon. The enzymatic activity of the protein codified by the gene with *DPYD* rs3918290 was completely absent [91]. Heterozygous carriers, with one functional allele only, have ~50% of the normal DPD enzyme activity preserved. Several papers [92–96] have established the clinical impact of such SNP both for its deleterious effect on the mature protein.

*DPYD**13/c. 1679 T>G

DPYD rs55886062 (c.1679T>A), whose allele frequency varies from 0.07 to 0.1% in Caucasians [90], was first described by Collie-Duguid et al as 'T1679G' [97] as responsible for the Ile560Ser amino acid change in a flavine mononucleotide binding domain. Besides the functional consequence of this variant have not been unraveled yet, it is thought to be related to destabilization of a sensitive region of the protein [98]. Hypothesis reinforced from the finding that heterozygous patients for *DPYD* rs55886062 presented decreased enzyme activity [97]. Homozygous expression of the variant allele of this SNP has been demonstrated to result in a 75% reduction of DPD enzyme activity compared with the wild-type [91] from the aforementioned study of Offer et al. Patients with allele variants of *DPYD* rs55886062 showed severe toxic side effects in several studies [93,96,97].

DPYD c. 2846 A>T

The variant allele of *DPYD* rs67376798 (c. 2846A>T) was first described by van Kuilenburg in 2000 [12]. Reported variant allele frequencies vary from 0.1 to 1.1% in African-Americans and Caucasians, respectively [90]. This *DPYD* has been proposed to directly or indirectly interfere with cofactor binding or electron transport because of a Asp949Val amino acid change localized near an iron-sulfur motif and thus responsible for a structural change in DPD enzyme [98]. Besides the DPD enzyme activity with *DPYD* rs67376798 is significantly impaired, it is not comparable to the one observed with *DPYD* rs3918290, where homozygous expression resulted in a completely nonfunctional enzyme [99]. This finding suggests that a heterozygous carrier would have around 25% reduction in DPD activity.

DPYD rs3918290 and *DPYD* rs67376798 were found to be strongly associated with severe (grade more or equal to 3) toxicity 5-FU related in a large cohort of patients on a FL-based therapy [100] and also with capecitabine-related toxicity [92]. These associations were strengthened by a meta-analysis from Terrazzino and his colleagues [88].

1.3. Implementation of pharmacogenetic diagnostic in the clinic

This huge amount of information about the clinical impact of these SNPs fueled the discussion in the scientific community and gave rise to the publication of PGx irinotecan and FL dosing guidelines.

CPT-11 is one of the chemotherapeutic drug presenting a PGx warning inside the package insert of the drug comprising the homozygosity for the *UGT1A1*28* as a risk allele for severe neutropenia development [101]. The findings regarding *UGT1A1*28* genotype are now translated into specific pharmacogenetic guidelines suggesting a drug-specific dose adjustment according to the individual

genetic background [102,103]. Since 2011 the Dutch Pharmacogenetics Working Group of the Royal Dutch Association for the Advancement of Pharmacy (DPWG) has evaluated therapeutic dose recommendations for CPT-11 based on *UGT1A1* genotype: a reduction of the 30% of the standard dose was suggested for patients homozygous for the *UGT1A1*28* allele and treated with doses higher than 250 mg/m^2 [102]. More recently, a French joint working group including the National Pharmacogenetics Network (RNPGx) and the Group of Clinical Onco-pharmacology (GPCO-Unicancer) have published more complex guidelines according to *UGT1A1*28* genotype when prescribing CPT-11 [103]. At first cycle, the standard irinotecan dose between 180 and 230 mg/m^2 is recommended to be reduced by 30% for *UGT1A1*28/*28* patients because at increased risk of developing hematological and/or digestive toxicity, particularly in cases of associated risk factors (performance status >3). For initially scheduled doses ≥240 mg/m^2 every 2-3 weeks, *UGT1A1*28/*28* patients are at a much higher risk of hematological toxicity (neutropenia) as compared to other genotypes. These guidelines thus recommend, for the "FOLFIRI-HIGH" regimen, the administration of an intensified dose (240 mg/m^2) only in *UGT1A1*1/*1* patients. The administration of an intensified dose (240 mg/m^2) is only possible in *UGT1A1*1/*28* patients without additional risk factors and under strict medical surveillance (Figure 8).

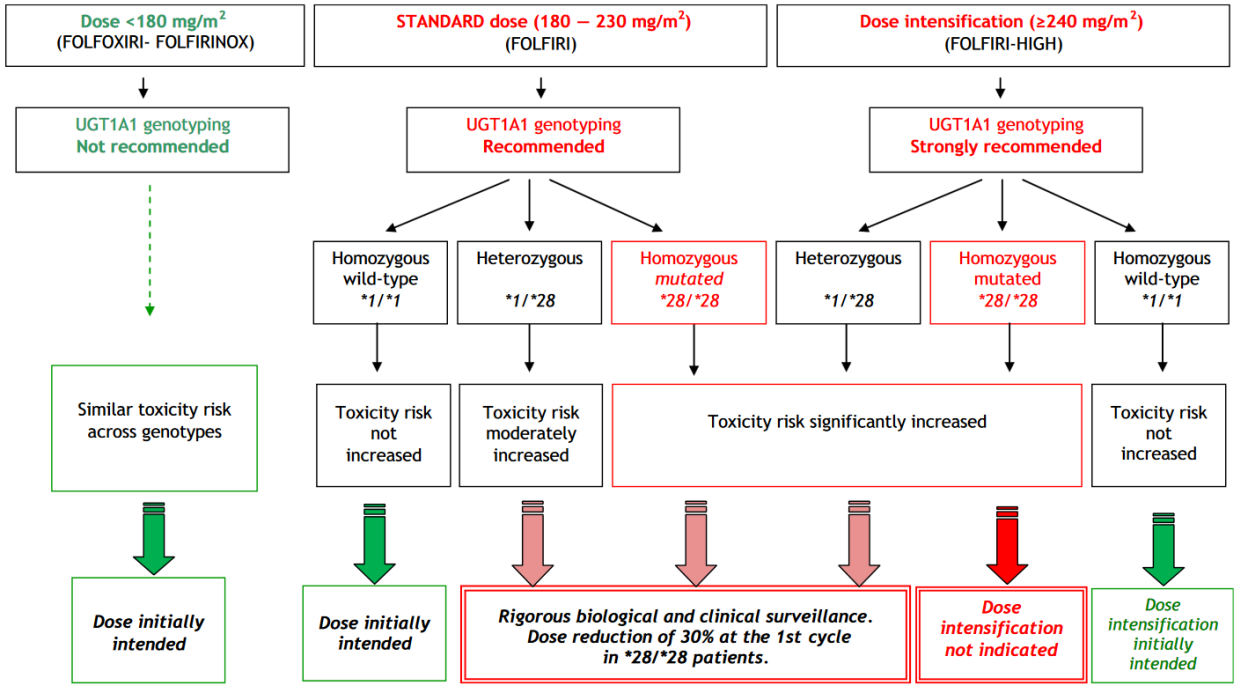


Figure 8. Decision tree for *UGT1A1* genotyping depending on initially intended irinotecan dose. From Etienne-Grimaldi, M.-C. *et al. UGT1A1* genotype and irinotecan therapy: general review and implementation in routine practice. *Fundam. Clin. Pharmacol.* 29, 219–237 (2015).

DPWG together with the Clinical Pharmacogenetics Implementation Consortium (CPIC) have also published FL-specific guidelines with recommendations regarding drug-related genetic tests and their integration in the clinical routine [102,104]. To date, personalization of the FL therapy can be achieved with the up-front test of three aforementioned *DPYD* genetic variants: rs3918290, rs55886062 and rs67376798. Heterozygous patients for at least one of these SNPs present intermediate or partial DPD enzyme activity. Thus a reduction of at least 50% of the initial dose for these patients is recommended in the CPIC guidelines. Also, the DPWG made available FL guidelines with the dose adjustments recommended to be applied in carriers of a *DPYD* variant allele. The Dutch group recently updated their online guidelines for FLs dose adjustments accordingly with a "gene activity score" [90]. Variant alleles consistent with this score can have a different weight according to the degree of enzyme activity. An FLs dose recommendation according to the gene activity scores for *DPYD* is shown in Table 1. After initial reduction, dosages can be further titrated based on clinical tolerance. Dose reductions are 75, 50 or 25% for gene activity scores of 0.5, 1, and 1.5, respectively. The gene activity score varies from 0 (no DPD activity) to 2 (normal DPD activity).

Table 1. Initial dose recommendation for *DPYD* gene activity score.

Gene activity score	**% of standard dose**
0	Alternative drug
0.5	25
1	50
1.5	75
2	100

Clues that other *DPYD* SNPs (e.g. rs2297595, rs1801160, rs1801158, rs1801159, and rs17376848) could have a role in the development of FL-related toxicities came from international guidelines [102] and from the most recent literature [93,96,105]. These SNPs have been previously observed in patients with low DPD enzymatic activity [91]. However, there is no final evidence that promotes them as a possible predictive biomarker of FL-related toxicity.

The pharmacogenetic diagnostic is not yet commonly implemented in Italian hospital practice as well as in the European one. Among the hospitals that started a routine *DPYD* screening program is the Leiden University Medical Center (LUMC; Leiden, The Netherlands).

1.4. Is the implementation of Pharmacogenetic diagnostic cost-effective?

Evidence of a robust association between the genotype and treatment response must be provided (analytical and clinical validity) before the testing and claiming clinical utility. Indeed, the association between specific pharmacogenetic patient features and clinical outcome (toxicity/efficacy) has been extensively demonstrated [106,107] but the potential economic consequences of this association have not been assessed for every drug with a pharmacogenetic recommendation [108]. Clinical utility regards the feasibility of the test and its value in the health-care although there has not always been a widely shared definition of its term [108,109]. Different definitions have been published from NIH, US Centers for Disease Control and Prevention and WHO but none of them specifically regard the pharmacogenetic test, and a general consensus is far from being reached. Although everyone could agree that clinical utility means evidence of a greater benefit than risk for the patient, the adequate level of proof needed to support the translation of pharmacogenetic biomarker in the clinic has yet to be established. "Considering the minimal direct risk and costs associated with pharmacogenetics testing, with the assumption of analytical and clinical validity, even a small benefit would probably outweigh the risks" [108]. A review from 2010 of pharmacoeconomic evaluation of pharmacogenetics test revealed extensive heterogeneity in study methodologies, even in the assessment of the same test [110]. Moreover, differences in costs for the PGx test can be substantial between countries, or even laboratories, and therefore the cost-effectiveness assessment. In addition test sensitivity or specificity can vary due to different ethnicities studied [111].

As already said the growing body of evidence concerning the clinical validity of genetic markers for the optimization of pharmacological treatment has led in recent years the pharmacogenetic community to publish guidelines regarding dose-adjustment regarding 27 drugs dose-adjustment based on the patient genotype [112]. Similar guidelines have also been recently published in Italy by the Italian Society of Pharmacology (SIF) and Italian Association for Medical Oncology (AIOM) (http://www.sifweb.org/docs/sif_aiom_position_paper_raccomand_farmacogen_gen15.pdf).

"Diagnosis Related Groups" (DRG) is a patient classification system to divide clinical cases into groups receiving similar health care services. In the DRG system, patient records are categorized into homogenous groups, according to the diagnosis and healthcare expenses involved, with the objective of providing appropriate reimbursement for the healthcare services. DRG represent nowadays the basis for hospital payment system in many western countries [113]. However, patient pharmacogenetic profiling has not been considered yet in this context. The definition of all the factors involved in the definition of the

costs of toxicity management, including pharmacogenetic markers should be considered in the DRG definition in order to reach a precision reimbursement strategy of standard therapy.

In cancer therapy, although the sensitivity of *DPYD* genotyping is low (<14.5% for *DPYD*2A* and c.2846A>T combined) [114], prospective screening for genetic variants in *DPYD* is a well-known strategy to detect patients with reduced DPD enzyme activity [89,92]. Up-front genotyping for *DPYD*2A* followed by 50% dose reduction for heterozygotes patients has been recently demonstrated to be cost-effective [9,89] Indeed genotype-guided dosing significantly reduced the incidence of severe (grade≥3) toxicity, from 73% in historical controls to 28% in the genotype-guided treatment cohort with an absolute risk reduction in the incidence of drug-induced death from 10% to 0% [9,89].

Regardless, *UGT1A1*28* polymorphism up-front genotyping before irinotecan [112] treatment is recommended since 2004 by FDA to increase irinotecan safety. Cancer patients with the **28/*28* genotype treated with irinotecan are at higher risk of developing severe neutropenia compared to the **1/*1* and **1/*28* genotype patients. To prevent irinotecan-related severe toxicities, irinotecan dosage can be reduced by 30% in **28/*28* patients [102]. Despite the robust data on the scientific validity of pre-emptive *UGT1A1*28* testing, the test is still not routinely adopted in the clinic [107]. Providing additional evidence of the potential utility of the test feasibility besides the definition of its clinical impact would represent a step forward to its integration in clinics. Similar attempts have been made in order to define the cost-effectiveness of the *UGT1A1*28* screening for irinotecan treatment but only computational, decision-analytic models have been applied to simulated data [115–118].

The cost effectiveness of *UGT1A1*28* screening prior irinotecan-based chemotherapy has never been tested in a clinical context. To our knowledge, only a few studies approached this topic applying simulated decision analytic models to define the cost-effectiveness of a pre-emptive *UGT1A1*28* genotyping [115–118]. All of them demonstrated that *UGT1A1*28* upfront testing might be a cost-effective strategy to prevent severe drug-related toxicity. Three of them suggested reducing irinotecan dosage according to the patients' genotype in order to prevent toxicity occurrence. Gold et al. brought evidence that the dose reduction on *28/*28 individuals may be cost-effective as long as the treatment efficacy remains above 98.4% of full dose efficacy, compatible with a dose reduction of 25% [116]. The study by Pichereau was the only one to suggest that a prophylactic GCSF treatment based on genotype in *28/*28 patients can be cost-effective as compared to a genotype-driven irinotecan dose reduction, without affecting treatment efficacy [115]. The *UGT1A1*28* genotype-driven dose reduction evaluated in

different ethnicities has shown to potentially be cost-effective just for Africans and Caucasians populations and not for Asians, where the *UGT1A1*28* allele frequency is considerably lower [117,118].

All the mentioned studies showed encouraging evidence of the cost-effectiveness of *UGT1A1*2*8 genetic screening before treatment with irinotecan in different economic contexts and in different ethnicities. Regardless of their inhomogeneities, they univocally agree in considering the test cost-effective, but they provide only theoretical data.

The high sensitivity of UGT1A1*28 test due to its prevalence, its clinical validity and the related ethical issues of not applying a dose reduction in patients carrying a known risk variant, make this pre-emptive test a good candidate for a broadly integrated widespread screening into a daily routine.

2. Rationale

Tailoring medicines is considered a key challenge for the pharmaceutical market, but still, it hobbles to become common practice in medical oncology. Administering the therapeutic dose for each patient is crucial and, so far, the current drug-dosing method has been relying basically on the patients' body mass index (BMI) along with other factors as gender and age. In the last years, it has become increasingly evident the need for a new dosing approach accounting for the complex processes of drug metabolism. One of the most promising fields to achieve better-targeted treatments consists on Pharmacogenomics. Indeed, genetic profiling prior to pharmaceutical treatment can classify patients into clusters depending on the risk of developing severe toxicity and the expected responsiveness to a standard dosage of pharmacological treatment, allowing personalized interventions.

Inter-individual variability in patient's response is often the reason of severe toxicities after treatment and of a reduced effectiveness, causing damage both to the patient and to the healthcare system. Especially, chemotherapeutics are characterized by a narrow therapeutic index which is even more impacted by inter-individual variability representing a challenge in the management of therapy outcome. Adverse drug reactions to chemotherapies are a major cause of morbidity and mortality worldwide, and their cost in health care is substantial. For these drugs, several published data highlighted the clinical validity of several genetic biomarkers useful to optimize their dosage. Prescribing dosing guidelines based on patients' genotype were generated to help the clinicians in the definition of the right dosage for each patient basing on studies whose relative strength of evidence ranged from high (large studies and/or replicated findings) to low (a single study with 20/200 subjects). Several studies report that pharmacogenomic drug advice to a patient in clinical practice brings measurable clinical benefit in terms of reduced toxicity and improved outcome.

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in males and the second in females, and cause of almost 700,000 cancer-related deaths worldwide. Despite the introduction of new therapeutic agents and the great improvement achieved in the response rate and patient's survival, great research efforts have been focused on elucidating the contribution of the host genetic variability on the outcome of backbone chemotherapy such as irinotecan and fluoropyrimidines. A clear defined biomarker for irinotecan-related severe toxicity has been identified in UDP glucuronosyltransferase 1, specifically the *UGT1A1*28* SNP. Its clinical validity, along with the availability of a feasible PGx test, and of peer-reviewed guidelines represents an important opportunity when systematically introduced in the

daily clinical practice. Fluoropyrimidines-related severe toxicities could be predicted from another recently defined biomarker: Dihydropyrimidine dehydrogenase gene, *DPYD*. Limited data are available in CRC treatment personalization regarding new validated marker of severe toxicity.

Improving the efficacy-toxicity balance in oncology remains an unmet need. Indeed, pharmacogenomics represents a powerful new approach not only in leading to optimal care but also to improve chemotherapy cost-effectiveness.

3. Aims

The aim of my thesis consisted in establishing new pharmacogenomic markers of the irinotecan and/or fluoropyrimidines-associated toxicity to increase treatment safety in colorectal cancer patients. To this purpose, I have explored different stages of the PGx markers development: from the exploratory setting to the clinical implementation setting.

Specifically, aims of my activity have been:

1. To explore the existence of innovative PGx biomarker of toxicity to a FOLFIRI regimen in mCRC patients. For the discovery of novel genetic predictors of neutropenia and gastrointestinal (GI) toxicity risk following a FOLFIRI-based treatment, a haplotype-tagging polymorphism strategy (htSNPs) and an independent replication analysis have been used. It is well acknowledged that inflammation plays an important role in the tumor development with repercussions also on chemotherapeutics' pharmacokinetics and pharmacodynamics. We, thus, investigated 250 Caucasian metastatic CRC patients homogeneously treated with first-line FOLFIRI for 246 htSNPs in 22 transcriptional regulators and cytokines inflammation-related genes. One polymorphism in STAT3 gene resulted predictive of severe GI toxicity with a protective effect toward the risk of developing grade 3-4 events and was validated in a replication set.
2. To analyze the clinical utility of profiling *UGT1A1*28,* an acknowledged PGx biomarker in irinotecan-treated mCRC patients in the clinical practice. Specifically, we aimed at estimating *UGT1A1*28* clinical efficacy by measuring its relationship with the financial costs associated with irinotecan-related toxicity at the National Cancer Center CRO of Aviano. Of the previously described population we identified a subset of 243 patients, we conducted a retrospective analysis of the cost of toxicity management in relation with *UGT1A1*28* genotyping. The mean predicted cost per patient was higher for *1/*28 (1,119€), and *28/*28 (4,886€), as compared to *1/*1 (812€) ($P<0.001$). This resulted consistent with a differential grade 4 toxicity profile among the three groups of patients, and with a higher frequency of costly interventions like hospitalization among patients with the *UGT1A1*28* allele.
3. To collaborate to the set-up of a PGx implementation infrastructure for the prevention of irinotecan and/or fluoropyrimidines-associated toxicity in the clinical routine of the National Cancer Center CRO of Aviano. The establishment of a successful implementation of pre-emptive PGx testing consisted of testing the presence of *DPYD* (rs3918290, rs55886062, rs67376798)

and/or *UGT1A1*28* (rs8175347) risk variants, for patients eligible for fluoropyrimidine and/or irinotecan treatment respectively, in the National Cancer Center CRO of Aviano routine practice. For its efficient integration in the hospital workflow we considered the peculiar infrastructure and regulatory conditions: starting from the digital pharmacogenetic inquiry from the oncologist, to the medic personnel training and sensitization, to the elaboration of a ISO-9001 approved protocol for the sample processing and eventually to the generation of a digital PGx report (both technical and clinical) according to the published pharmacogenetics dosing guidelines that once is embedded in patients' clinical record. The implementation of pre-emptive pharmacogenetics test is now part of a European UPGx Project with the aim of providing the final proof of pharmacogenetics efficacy when fully integrated into the clinical practice in increasing drug safety.

4. Methods

4.1. Patients enrollment and drug administration

Three study set populations have been taken into consideration for the present study. In study sets one and two, the severities of neutropenia and GI toxicities were evaluated prospectively, and according to National Cancer Institute Common Terminology Criteria for Adverse Events version 3.0 criteria and the Institutional Review Board (IRB) of each participating institution approved the study protocol. All the patients signed a written informed consent for the genetic analysis before entering the study.

4.1.1. Study set one

The study set one regarded two hundred and fifty CRC patients, afferent to the CRO-National Cancer Institute of Aviano, Italy, who were enrolled in the study. Details on eligibility, modalities of treatment and toxicity data collection are herein reported from the study published by our group in 2006 [50]. Patients were treated with either the modified FOLFIRI regimen (>90% of patients) as described by Tournigand (irinotecan 180 mg/m^2 intravenously for 2 hours on day 1 + FU 400 mg/m^2 bolus followed by FU 2,400 mg/m2 continuous infusion during 46 hours + LV 200 mg/m^2 on day 1 every 2 weeks) or the FOLFIRI regimen (irinotecan 180 mg/m^2 intravenously for 2 hours on day 1 + FU400mg/m^2 bolus followed by FU 600mg/m^2 continuous infusion during 22 hours on days 1 and 2 + LV 200 mg/m^2 on days 1 and 2 every 2 weeks). Before starting irinotecan administration, patients were treated with atropine 0.5 mg, dexamethasone 8 mg, and granisetron 3 mg or ondansetron 8 mg. Diarrhea was treated promptly with loperamide 4mg at the onset, and then with 2 mg every 2 hours until the patient was diarrhea-free for at least 12 hours. One cycle of treatment consisted of two-week therapy. Eligibility criteria included Caucasian ethnicity, histologically metastatic CRC (presented unresectable metastases); no prior chemotherapy for metastatic disease (adjuvant chemotherapy was allowed, except for irinotecan); age between 18 and 75 years; absolute neutrophil count ≥ 2,000µL; platelets ≥ 100,000µL; performance status (WHO) of 0 to 2; life expectancy more than 3 months; at least one measurable cancer lesion; normal renal function (creatinine clearance > 65 mL/min by Cockcroft formula); and ALT, AST < 1.25x normal value or < 2x for Gilbert's Syndrome.

All of the initially enrolled subjects were found eligible and included both in the exploration and in the cost analysis for toxicity. Toxicity was classified as hematological (neutropenia, anemia, leucopenia, and thrombocytopenia) or non-hematological (diarrhea, nausea, vomiting, asthenia, alopecia, mucositis, anorexia, and non-neutropenic infection). In the aforementioned studies, particular attention was given

to neutropenia and GI (diarrhea, nausea, vomiting) toxicities that represent the major irinotecan-related side effects; the worst event recorded during the entire course of chemotherapy was considered.

Objective clinical evaluation, blood counts, and hepatic and renal function tests were performed within 48 hours before each cycle. Patients were questioned specifically about nausea and vomiting, mucositis, diarrhea, malaise, and appetite at every cycle. A single cycle of chemotherapy administration was considered sufficient for evaluation of acute toxicity, whereas the response to treatment was evaluated only in patients who had received at least four cycles of chemotherapy. Clinical evaluations were performed blindly with respect to the genetic results, and clinical data were monitored by the study sponsor. Chemotherapy was delayed until recovery if neutrophils were ≤ 1,500µL or in the presence of significant, persisting, non-hematologic toxicity. In the event of grade 3 or 4 neutropenia, thrombocytopenia, and diarrhea, the irinotecan dose was reduced (from 180 mg/m^2) to 90 to 150 mg/m^2 based on the physician's assessment. Treatment was discontinued in the event of repeated grade 3 to 4 toxicity, despite dose reduction, or because of patient refusal.

4.1.2. Study set two

The study set two was composed of 167 Eastern Canadian mCRC patients receiving FOLFIRI-based regimens and characterized for hematologic and GI toxicity profile. Details on eligibility, modalities of treatment and toxicity data collection are herein reported from Levesque et al in 2013 and in the two study from Chen in 2015 [119–121].

Patients were treated with one of the following FOLFIRI-based chemotherapies. Patients treated with the modified FOLFIRI regimen received irinotecan (180 mg/m^2 i.v.) for 2 hours on day 1 plus a bolus of 5-fluorouracil (400 mg/m^2) followed by continuous infusion of 5-fluorouracil (2400 mg/m^2) plus leucovorin (200 mg/m^2) over 46 hours. Patients received this treatment cycle every two weeks. Sixty-nine patients also received the monoclonal antibody bevacizumab (Avastin; Genentech, San Francisco, CA) in co-administration with their regimen, and 6 patients received either an experimental drug or placebo. All patients received an 180 mg/m^2 intravenous dose of irinotecan every 2 weeks, and 75 patients also received co-treatments-bevacizumab, an experimental drug, or a placebo.

The toxicity endpoints consisted of both GI and hematologic toxicities and were analyzed separately. For GI toxicities, all patients completed a daily report of GI toxicities during the first 14 days of each cycle to record the incidence and severity of nausea, vomiting, and diarrhea. For hematologic toxicities, laboratory parameters were collected before each cycle of chemotherapy and/or when the treatment

was delayed. The most severe toxicity reported was used for data analysis. GI toxicity was assessable for all patients except for one who died before toxicity assessment, and another who did not fill out the GI toxicity diary, while hematologic toxicity was evaluable for 166 of 167 patients.

4.1.3. Study set three

The study set three was composed of 393 Caucasian patients with a cancer diagnosis at our institute, CRO-Aviano eligible for treatment with either irinotecan or fluoropyrimidine-based chemotherapy. Details on demographics, treatment, toxicity data collection have been collected retrospectively from a survey submitted to the prescriber oncologist (see appendix 1).

The pharmacogenetic diagnostic service started from October 2011 and systematically from January 2014. The diagnostic service could answer to both a pre-emptive inquiry and to a toxicity inquiry. In the first case, patients were profiled for genetic variants known to impact the risk of severe toxicity development after either fluoropyrimidine (*DPYD* variants rs3918290, rs55886062, rs67376798) or irinotecan-treatment (*UGT1A1* SNP rs8175347) in order to recommend the oncologist a dose-adjustment in order to avoid the development of severe and also life-threatening toxicities. In the second case, patients were screened for the same risk variants however in response to an unexpected development of severe toxicity after either fluoropyrimidine or irinotecan administration.

4.2. Molecular analysis

4.2.1. Sample storage

Whole blood samples were collected from the patients, cataloged in the proper database and stored in a freezer at –80°C.

4.2.2. Genomic DNA extraction

The automated extractor BioRobot EZ1 (Qiagen SPA, Milano, Italy) with the Card "EZ1 DNA Blood" was used in association with the Kit "EZ1 DNA Blood Kit 200 µl", for the extraction of genomic DNA from 350µl of whole blood obtaining 200 µl as final volume, corresponding approximately to 5-12 µg of DNA. With this procedure, blood cells are lysed during a short incubation (10 minutes at 70°C) with proteinase K in the presence of a chaotropic salt (guanidine-HCl), which immediately inactivates nucleases, and DNA binds to the surface of the silica-coated magnetic particles. The particles are then separated from the lysates using a magnet, and the DNA is efficiently washed and eluted in the elution buffer. In this way, the DNA is held and purified from the blood sample.

DNA extracted is then kept at 2-8°C.

4.2.3. Polymerase chain reaction (PCR)

PCR is an in vitro reaction exploiting DNA polymerase catalytic activity to dramatically amplify a fragment of DNA starting from a small quantity of template DNA in the presence of deoxynucleotides triphosphate (dNTPs), PCR oligonucleotides used as primers and a DNA polymerase in a proper reaction buffer. Three distinct events must occur during a PCR cycle: denaturation of the template, primer annealing and DNA synthesis by a thermostable polymerase as Taq polymerase. DNA is initially heated to temperatures close to boiling, in order to denature it and thus obtain a single-strand. DNA denaturation occurs when the reaction is heated to 92-96°C. The time required to denature the DNA depends on its complexity, the geometry of the tube, the thermal cycle and the volume of the reaction, usually a 30 seconds denaturation time is used. For DNA sequences that have a high G+C content, longer denaturation times have been used to improve the yield of PCR

After denaturation, the oligonucleotide primers (sense and antisense) hybridize to their complementary single-stranded target sequences in a process called annealing. The temperature of this step varies from 37°C to 65°C, depending on the homology of the primers for the target sequences as well as the base composition of the oligonucleotides. Primers are used at a higher concentration than the target DNA,

and are shorter in length; as a result, they hybridize to their complementary sequences at an annealing rate several orders of magnitude faster than the target duplex DNA can re-anneal.
The last step is the elongation of the oligonucleotide primers by a thermostable polymerase, Taq polymerase responsible for the catalysis of the parental strand duplication. This portion of the cycle is usually carried out at 72°C. The time required to copy the template fully depends on the length of the PCR products. To obtain the amplification of the desired DNA sequence, the cycle of denaturation/annealing/elongation must be repeated several times, typically from 25 to 40 times.

The reagents used in a PCR are: reaction buffer, magnesium ions supplied by the magnesium chloride ($MgCl_2$), deoxynucleotides triphosphate (dNTPs), the specific primers, DNA polymerase and the template. In particular, for each sample, a reaction mixture, containing the reaction buffer, a solution of $MgCl_2$, the dNTPs, primers and DNA polymerase, is made before adding genomic DNA.

This series of thermal cycles is carried out thanks to a programmable instrument, the thermal cycler, capable of changing the temperature very quickly and keep it constant for a given period of time. The result of a PCR is that, at the end of *n* cycles of amplification, the reaction mixture contains a theoretical maximum number of double-stranded DNA equal to 2^n (where "*n*" represents the number of amplification cycles).

The most common issue with PCR is the contamination of reactions with target nucleic acids. This can occur during steps prior to the actual amplification reaction and can be avoided mainly by trying to operate in a DNA-free, clean environment.

4.2.4. Methodologies for polymorphisms analysis

NCBI (National Center for Biotechnology Information) databases and tools were used to develop the polymorphisms analyses. The NCBI presents a website showing links to databases containing information about genes (Gene), polymorphisms (dbSNP), scientific literature (PubMed), besides search and analysis tools. These and other additional databases (SNP500, PharmGKB (The Pharmacogenomics Knowledge Base), and 1000 Genomes Browser) were consulted for assay design (genetic sequences, polymorphisms description, primer design), during this work of thesis.

Subsequently, according to the type of polymorphism and to the specific characteristics of the nucleotide sequence to be analyzed, the most suitable method of genotyping has been chosen.

In particular, in this PhD thesis, semi-automated, recently developed genotyping methods have also been used. These are based on PCR reactions and allow the identification of genetic polymorphisms in a very simple and easy way: Pyrosequencing (PSQ), allelic discrimination based on TaqMan chemistry, the Fragment Analysis (Gene Scan).

In a second phase, analytical platforms were implemented with the introduction of Illumina BeadXpress® Reader, based on GoldenGate chemistry and VeraCode Beads technology.

4.2.4. Pyrosequencing

Pyrosequencing is an analytical technology for SNP identification consisting of a real-time pyrophosphate detection method [122].

This technique is based on indirect bioluminometric assay of the pyrophosphate (PPi) that is released from each dNTP upon DNA chain elongation. Following Klenow polymerase mediated base incorporation, PPi is released and used as the substrate, together with adenosine 5'-phosphosulfate, for the ATP sulfurylase, which results in the formation of ATP. Subsequently, the ATP accomplishes the conversion of luciferin to its oxi-derivative by luciferase. The ensuing light output is proportional to the number of added bases, up to about four bases. To allow processivity of the method, dNTPs in excess are degraded by apyrase, which is also present in the starting reaction mixture and continuously degrades ATP and unincorporated dNTPs. This switches off the light and regenerates the reaction solution. The dNTPs are added one by one to the template during sequencing procedure. It should be noted that deoxyadenosine alfa-thio triphosphate is used as a substitute for the natural dATP since it is efficiently used by the DNA polymerase, but not recognized by the luciferase. The process is fully automated and adapted to a 96-well format, which allows rapid screening of a large panel of samples.

Following the first phase of sample preparation, the plate is loaded on an instrument, the PSQ 96MA Pyrosequencing, which determines and provides the genotype directly at the level of the analyzed SNP.

Pyrosequencing analysis is performed on PCR-amplified DNA. One of the PCR primers must be biotin-labeled for immobilization to streptavidin-coated Sepharose beads. This allows the separation of the two DNA strands produced by PCR, since the assay must be carried out on single-stranded DNA. If the reverse primer is biotinylated we have the forward assay, otherwise, if the forward primer is biotin-labeled, the assay is called reverse.

PCR reaction product is mixed with streptavidin coated High-Performance Sepharose beads (Amersham Biosciences, Uppsala, Sweden) in the presence of a binding buffer (Tris 10 mM, Sodium Chloride 2 M, EDTA 1 mM and Tween 20 0.1%, pH 7.6). The mixture is allowed to shake for 10 minutes at room temperature. The samples are subsequently transferred to a 96-well filter plate, and vacuum (vacuum manifold for 96 well filter plate, Millipore) is applied to remove all liquid. Denaturation solution (Sodium Hydroxide 0.2 M) is added to denature double-stranded PCR product DNA. After 1-minute incubation, a vacuum is applied to remove the solution and the non-immobilized DNA. The beads are washed twice with a washing buffer (Tris 10 mM, pH 7.6) in the presence of the vacuum. The beads with the

immobilized template are resuspended by adding 45 µl annealing buffer (Tris 20 mM, Magnesium Acetate Tetra-Hydrate 2 mM, pH 7.6), and sequencing primer (2 µM) is added to each sample. The design of sequencing primers for Pyrosequencing follows the same criteria as for the PCR primers, except that the Tm of this primer may, if necessary, be lowered. The sequencing primer could thus be shorter than the PCR primers, typically 15 bp. The position of the primer is flexible within 5 bases from the SNP and can be designed on both the positive (reverse assay) or on the negative (forward assay) strand. Thirty-five µl of this mixture is transferred to a Pyrosequencing 96 wells plate (PSQ 96 Plate Low).

The plate is incubated for 5 minutes at 60° C to allow complete sequencing primer annealing on the template DNA. After samples cooling, the plate is transferred to the Pyrosequencing instrument. The biotin labeled DNA template, annealed to the sequencing primer, is incubated with enzymes (DNA polymerase, ATP sulfurylase, luciferase, and apyrase) and the substrates (adenosine 5'phosphosulfate and luciferin).

The first of four dNTPs is added to the reaction. DNA polymerase catalyzes the incorporation of the dNTP into the DNA strand, if complementary to the base in the template strand. Each incorporation event is accompanied by the release of pyrophosphate in quantity equimolar to the amount of incorporated nucleotide. ATP sulfurylase converts PPi to ATP in the presence of adenosine 5'phosphosulfate. This ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin that generates visible light in amounts that are proportional to the amount of ATP. The light produced in the luciferase-catalyzed reaction is detected by a charge coupled device (CCD) camera and seen as a peak in a pyrogram. The height of each peak (light signal) is proportional to the number of nucleotides incorporated.

Apyrase, a nucleotide degrading enzyme, continuously degrades ATP and unincorporated dNTPs. This switches off the light and regenerates the reaction solution. The next dNTP is then added. The addition of dNTPs is performed one at a time. It should be noted that deoxyadenosine alfa-thio triphosphate is used as a substitute for the natural dATP since it is efficiently used by the DNA polymerase, but not recognized by the luciferase. As the process continues, the complementary DNA strand is built up, and the nucleotide sequence is determined from the signal peaks in the pyrogram.

"PSQ Assay Design" software was used for the planning of the described assays: it allows to easily choose the set of primers (sense and antisense primers for PCR and sequencing primer for subsequent

analysis at PSQ) most suitable for the study of each SNP The analysis of the results is accomplished with the"PSQ™ 96 MA software".

The reagents and solutions used in Pyrosequencing analysis are:

- Aqua B. Braun Ecotainer, sterile water for injection (B. Braun, Melsugen AG, Germany);
- Streptavidin Sepharose™ High Performance (Amersham Biosciences AB, Uppsala, Sweden);
- Sequencing primer provided in lyophilized form (Sigma Genosys, Cambridge, UK) and then resuspended in sterile water to obtain a final concentration of 100 μM;
- Pyro Gold Reagents Kit (Biotage AB, Uppsala, Sweden) constituted by:
  - Enzyme mixture (luciferase, DNA polymerase, apyrase, sulfurylase and proteins binding to single-stranded DNA- provided in lyophilized form and then resuspended in sterile water;
  - Substrate mixture (adenosine 5'fosfosulfato [APS] and luciferin) provided in lyophilized form and then resuspended in sterile water;
  - dATPαS in solution;
  - dCTP in solution;
  - dGTP in solution;
  - dTTP in solution.
- PSQTM 96 Sample Preparation Kit (Biotage AB, Uppsala, Sweden) constituted by:
  - Binding Buffer (10 mM Tris-HCl, 2M NaCl, 1 mM EDTA, 0,1% Tween 20; pH=7,6);
  - Denaturation Solution (0,2 M NaOH);
  - Washing Buffer (10 mM Tris-acetate; pH=7,6);
  - Annealing Buffer (20 mM Tris-acetate, 2 mM Mg^{2+}-acetate; pH=7,6).

Other materials and instruments used in Pyrosequencing methodology are:

- PSQ 96 Plate Low (Biotage AB, Uppsala, Sweden);
- 96-well filter plates (Millipore, MA, USA);
- PSQTM 96 Reagent Cartridge (Biotage AB, Uppsala, Sweden);
- PSQTM 96 Sample Prep Tool Thermoplate (Biotage AB, Uppsala, Sweden);
- Multichannel Pipette (Matrix Technologies Corporations, NH, USA);
- Vacuum pump (Millipore, MA, USA);
- Shaker (Analitica De Mori, MI, Italia);
- PyroMark™ Vacuum Prep Workstation (Biotage AB, Uppsala, Sweden);

- PSQ96 MA Pyrosequencing (Biotage AB, Uppsala, Sweden), software PSQTM 96 MA;
- Pyrosequencing Assay Design Software, version 1.0.6 (Biotage, Westbrough, MA, USA).

4.2.5. TaqMan ® assay

TaqMan® allelic discrimination is based on the use of a Real Time PCR (RT PCR), that, in addition to the sense and antisense primers needed for the amplification of the SNP containing fragment, involves the use of an oligonucleotide *(probe)* able to pairing with the template. The probe pairs in an intermediate position between the sense and the antisense primer. The probe is functionalized at the two ends: in one part there is a "quencher" fluorophore (TAMRA) which acts as a silencer of fluorescence, the other one is tied to a "reporter" constituted by a fluorescent fluorophore (FAM or VIC). The action of silencing by the quencher occurs by transfer of energy from one fluorochrome to the other one when they are near to each other. In the reaction two different allele-specific probes, labeled with different fluorophores (fluorochrome FAM or VIC), are placed: one contains a perfect match to the wild type (allele 1) and the other one presents a perfect match to the mutation (allele 2). The allelic discrimination assay classifies unknown samples as homozygous and heterozygous.

TaqMan probe-based chemistry uses a fluorogenic probe to detect specific PCR product as it accumulates during PCR cycles. During the denaturation step, the reporter (R) and the quencher (Q) are attached to the 5' and 3' ends of a TaqMan probe. When both dyes are attached to the probe, reporter dye emission is quenched. During each extension cycle, the hot-start DNA polymerase system cleaves the reporter dye from the probe. After being separated from the quencher, the reporter dye emits its characteristic fluorescence which is recorded by a detector.

The probes are chosen according to certain characteristics:

- The T_m must be at least 5° C higher than the T_m of the two PCR primers because they must bind to the nucleotide sequence when executing the synthesis of the complementary strand;
- The oligonucleotide must have a length of about 20-30 bp and 50% of G and C;
- The extension phase must be performed at a temperature lower than the 72°C used in the PCR, in order not to cause the detachment of the probe from the template (for this reason we use high concentrations of $MgCl_2$);
- The probe must not form dimers or even pair with itself.

Samples are analyzed using the Applied Biosystems 7500 Real-Time PCR System instrument. The allelic discrimination was performed with the SDS software 2.3 (Applied Biosystems).

For SNP assay a preformed “TaqMan® SNP Genotyping Assay” is employed: it is available on-line in the catalog of Applied Biosystems (http://www3.appliedbiosystems.com/AB_Home/index.htm). As an alternative, you can use the service offered by the same company that, on sending the gene sequence containing the nucleotide variation, develops and tests specifically an essay called "Custom SNP Genotyping assay TaqMan®.

The practical procedure of the TaqMan® technology is really very simple and allows to analyze the genotype quickly using a universal mix (master mix) and a solution containing PCR primers and the two allele-specific probes. The step of sample preparation involves the use of 96-well plates with specific optical properties. The reaction mixture is prepared by combining the specific mix for the SNP under investigation (SNP Assay 20X or 40X), containing primers (sense and antisense) and the two probes labeled with FAM or VIC, to the Master Mix (TaqMan Genotyping Master Mix 2X) universal for all genotyping analyses, containing dNTPs, Taq Polymerase, $MgCl_2$ and salts in a suitable concentration creating an adequately buffered environment. The solution is dispensed into wells and, finally, genomic DNA is added (approximately 20 ng of DNA for each sample).

Once set up, the plate is covered with an adhesive film and centrifuged for a few minutes to eliminate the presence of any air bubbles at the bottom of the wells. Then the plate is loaded into the ABI PRISM 7900HT machine, at this stage, the RT-PCR conditions (temperature, duration, and cycles) and the test volumes (20 µl) are determined, and the markers (FAM and VIC) are assigned to polymorphism’s alleles. The amplification is carried out with a thermal cycler integrated into the instrument using the following thermal profile:

- 50° C for 2 minutes;
- 95° C for 10 minutes;
- 40 cycles for (92° C for 15 seconds; 60° C for 1 minute)

At the end of the PCR reaction, an endpoint scanning of the 96-well plate containing the samples is carried out, in order to detect the fluorescence signal produced in each well by the two fluorophores (FAM and VIC) associated with the allele-specific probes. Finally, thanks to the processing of obtained data by software SDS 2.3, the assignment of the genotype corresponding to each sample occurs.

For the analysis with TaqMan® technology were used the following reagents:

- 2X TaqMan Genotyping Master Mix (Applied Biosystems, CA, USA);
- 20X or 40X "TaqMan® SNP Genotyping Assay" or "CustomTaqMan® SNP Genotyping Assay" (Applied Biosystems, CA, USA);
- MicroAmp® Optical 96-Well Reaction Plate (Applied Biosystems, CA, USA);
- Optical Adhesive Covers (Applied Biosystems, CA, USA);
- Real-Time ABI PRISM 7900HT instrument (Applied Biosystems, CA, USA);
- SDS 2.3 software (Applied Biosystems, CA, USA).

4.2.6. Automated fragment analysis and Sanger Sequencing

Automated fragment analysis is performed to detect small variation in the length of a DNA fragment. It is based on the capillary electrophoresis coupled with fluorescence detection. Capillary electrophoresis occurs when an electric field is applied to an electrolyte solution within a capillary, causing ions migration. DNA fragments, having a negative charge, move toward the anode (+) and are separated by size (figure 9).

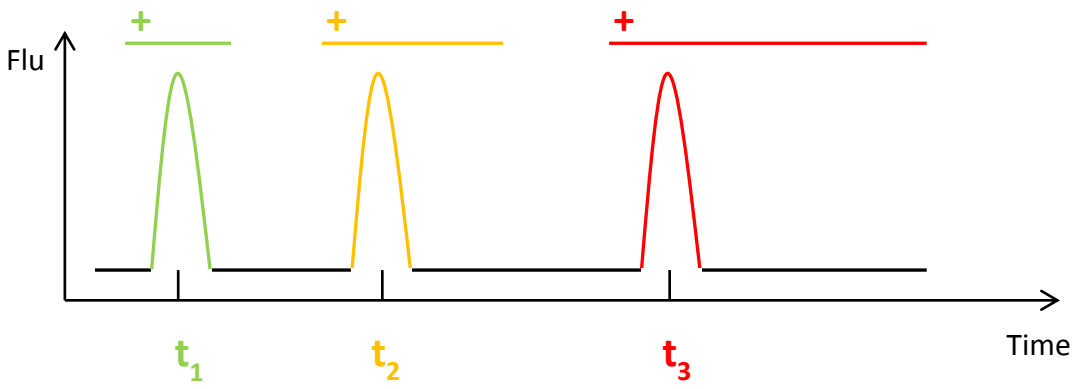


Figure 9. Scheme of the separation of different sized DNA fragments labeled with different fluorophores (ROX=red, JOE=green, LIZ=yellow) by capillary gel electrophoresis.

DNA fragments are produced by PCR employing one 5' fluorescence-labeled primer with the HEX (isomer-free succinimidyl ester of 6-carboxy-2',4,4',5',7,7'-hexachlorofluorescein, excitation and emission maxima of 535 and 556 nm respectively) fluorophore. In the analysis, a marker of DNA molecular weight labeled with a different fluorophore, the ROX, is also employed. It serves as an internal standard.

These dye-labeled fragments are detected by fluorescence and in turn rendered into a sequence or sized fragment. The pherogram analyzed by the software presents on the abscissa the separated molecular weight fragments, while on the ordinate the intensity of the fluorescence peak. The samples are

analyzed in the Genetic Analyzer ABI Prism 3100 instrument (Applied Biosystems). Gene Scan analysis software (Applied Biosystems) allows data extraction and elaboration.

The process is very simple. The first phase consists in the amplification of the gene fragment containing the polymorphism of interest by means of a PCR that presents one of the two primers labeled with the fluorophore HEX covalently linked in 5' (not reactive extremity).

Since this method is very sensitive, a small concentration of amplified fragment to conduct the analysis is sufficient. Consequently, the samples, after being analyzed by electrophoresis on agarose gel, are suitably diluted. The mix needed to perform the analysis consists of 14.5 µl of deionized and purified formamide and 0.5 µl of Internal Lane Size Standard [ROX] for each sample. Once prepared the mix, this is aliquoted into a 96-well plate and, subsequently, 1 µl of the diluted sample is added to obtain a total of 16 µl per well. The formamide is a strong denaturant and is sufficient the contact with the DNA to exert its effect. The plate is covered, to prevent evaporation of the solution, and denaturated (2 min at 95° C). Immediately after denaturation, the plate is placed in ice to avoid the rewinding of DNA strands, and it is loaded into the Genetic Analyzer ABI PRISM 3100 instrument.

Once started the instrument, the 16 capillaries penetrate in the plate and take samples. The loading of the samples takes place through electrokinetic injection, i.e. through the application of a potential of 15 KV for about 5 seconds which moves all the charged molecules within the capillary. There are activities of competition by charged molecules or ions, present in the sample, which can interfere with this delicate phase of the process. The sample dilution in sterile water and purified formamide is also useful to reduce these interference phenomena. The phenomenon of stacking, which allows to the fragments and the mix to be loaded into a restricted and compact zone of the capillary, ensures the correct injection of the samples in the capillaries avoiding DNA diffusion phenomena. Stacking permits to produce an area of low conductivity, and this is made possible from the immersion of the capillary in water before loading the samples. After the first phase of injection, the samples are separated by an electrophoretic run and, at the exit of the capillary, they are blasted by a laser that excites all fluorophores emitting fluorescence in different regions of the spectrum. An analyzer of multiple wavelengths CCD camera (charged-coupled device) identifies the emissions of each fragment passing through the detector. The analysis of fluorescence occurs both for unknown fragments and for the standard internal fragments.

The ABI 3100 data collection software allows to control the conditions of electrophoresis and manages the creation of samples files and lists of injections. The extraction and processing of data are managed by the Gene Scan analysis software that allows the conversion of the data into appropriate colored peaks which have assigned values of fragment length, based on the time of output and the type of emission. The instrument Genetic Analyzer ABI PRISM 3100 (Applied Biosystems, Foster City, CA) was also employed, managed by the Gene Scan analysis software (Applied Biosystems, Foster City, CA).

The reagents and solutions used in this methodology are:

- Aqua B. Braun Ecotainer, sterile water for injection (B. Braun, Melsugen AG, Germany);
- Hi-Di™ Formamide (Applied Biosystems, Foster City, CA);
- Fluorophore ROX™ DYE (5-carboxy-X-rhodamine, succinimidyl ester) Gene Scan™ 400HD [ROX] Size Standard, (Applied Biosystems, Foster City, CA).

The Short Tandem Repeats (STRs) polymorphisms analyzed with the fragment analysis method are reported in Table 2.

Table 2. STRs analyzed by automated fragment analysis.

GENE	STR NAME	Rs ID	FUNCTION	AA CHANGE	SEQUENCE
UGT1A1	*28	rs8175347	5'UTR	NA	CTTGGTGTATCGATTGGTTTTTGCCA[(TA)6/7]AGTAGG AGAGGGCGAACCTCTGG

Sanger sequencing (Sanger *et al.*, 1977) is a method of DNA sequencing based on the selective incorporation of chain-terminating modified di-deoxynucleotides (ddNTPs) by DNA polymerase for detection in automated sequencing machines. These chain-terminating nucleotides lack a 3'-OH group required for the formation of a phosphodiester bond between two nucleotides, causing a stop in DNA extension when a modified ddNTP is incorporated. This process generates a pool of DNA fragment with different length, each one terminating with a ddNTP. These are fluorescently labeled with four dye-terminators, each one emitting at different wavelengths. The resulting DNA fragments are denatured both by heat and formamide and subsequently separated performing a capillary electrophoretic run, similarly as for the automated fragment analysis. To perform the PCR required for the Sanger Sequencing, a mixture, containing reaction buffer, $MgCl_2$, dNTPs, forward and reverse primers, Taq polymerase, and water has to be prepared (see details on 4.4. materials & methods section).

4.2.7. Beadexpress reader coupled with Veracode® Technology and Goldengate® Assay

Illumina BeadXpress Reader (Illumina, La Jolla, CA) is a dual-color laser scanning system allowing users to analyze several genetic markers in a multiplexing manner exploiting the VeraCode™ microbeads digital technology. This technology allows several types of multiplex testing ranging from genotyping, gene expression, RNA and protein-based assays, methylation and expression studies of 1 to 384 biomarkers per well at the same time. The VeraCode™ system is based on the VeraCode Beads, glass microcylinders (240 µm in length by 28 µm in diameter), each inscribed with a unique digital holographic code to unambiguously designate and track the specific analyte or genotype of interest throughout the multiplex reaction.

Unlike traditional microarrays, the VeraCode microbeads are used in solution, which takes advantage of solution-phase kinetics for more rapid hybridization times.

The microbeads highly pure glass, stable at high temperatures and chemical agents, represents an optimal surface for biomolecules attachment. In the GoldenGate Genotyping® Assay, each microbead is functionalized with a specific oligonucleotide which univocally identifies a single SNP. VeraCode™ beads are used for analyzing up to 384 genetic markers per sample in plates containing 96 samples each.

To perform a BeadXpress analysis, a sample preparation phase is required. In this phase, a unique multiple PCR reaction is performed. Subsequently, each SNP-containing fragment produced is conjugated with a specific VeraCode microbead for the genotyping attribution analyses.

The first step in the GoldenGate Assay is DNA activation by biotinylation, which enables genomic DNA samples to bind (by biotin-streptavidin interaction) to paramagnetic particles. This activation process is highly robust and requires only 250 ng of genomic DNA.

Three oligonucleotides are designed for each SNP locus. For each SNP site, there are two allele-specific oligos (ASOs). They have exactly the same sequence but differ only by the last nucleotide, which matches the polymorphic one found at the SNP site in the sequence of interest. So, for each DNA strand, only one ASO hybridizes, depending on the SNP genotype. A third oligo, the locus-specific oligo (LSO), instead, hybridizes several bases downstream from the SNP site. All three oligonucleotide sequences contain universal PCR primer sites (that is, complementary sequences recognized by the universal PCR primers P1, P2, P3); the LSO contains a unique address sequence (or “Lumicode”) that targets a particular oligonucleotide-coated VeraCode microbead type. Assay oligonucleotides (ASOs and LSOs),

hybridization buffer, and paramagnetic particles are then combined with the activated DNA in the hybridization step. During the primer hybridization process, ASOs and LSOs hybridize to the genomic DNA sample bound to paramagnetic particles. Because hybridization occurs prior to any amplification steps, no amplification bias is introduced into the assay. Following hybridization, several wash steps are performed, removing excess and mis-hybridized oligonucleotides. Extension of the appropriate ASO and ligation of the extended product to the LSO join information about the genotype present at the SNP site to the address sequence on the LSO. The ligation products (containing the SNP and the address sequence) serve as the PCR templates for universal PCR primers P1, P2, and P3. Primers P1 and P2 are fluorophore-labeled with Cy3- and Cy5-dyes, respectively, so, depending on the allele, the instrument will detect one color (in the case of homozygosis) or a two colors fluorescence (in the case of heterozygosis). P3 primer is the only reverse primer at the locus specific site, allowing the amplification of the address sequence for the binding with a specific bead. After downstream processing, the single-stranded, dye-labeled PCR products are hybridized to their complementary bead type through their unique address sequences. Hybridization of the GoldenGate Assay products onto the VeraCode beads separates the assay products for individual SNP genotype readout.

After hybridization, the BeadXpress® Reader is used for microbead code identification and fluorescent signal detection.

The plate is loaded in the BeadXpress Reader and beads from 8 wells at a time are drawn up and aspirated onto the 8-chambered transparent groove plate in which, thanks to a combination of fluid flow, gravity, and capillary force, they populate and align closely within the grooves. Once the beads are aligned, the entire fluidic cell is actuated across the optical system and scanned for fluorescent intensity and code classification. Here, a dual-color laser detection system identifies, on one hand, the unique holographic code embedded in each VeraCode bead and on the other hand the signal intensity associated with each bead discriminating the genotype. Assays developed with VeraCode microbeads typically include up to 30 replicates of each bead type. Each microbead is optically scanned up to a dozen times providing about 300 independent data point for each analyte ensuring reliable and accurate results.

The plate preparation process lasts about two days, and the workflow is summarized in Table 3.

Table 3. BeadXpress workflow.

Process	Time	Day
DNA activation	1h 20 min	1
Oligonucleotides addition and DNA binding to paramagnetic particles	3h	1
Oligonucleotides-DNA binding	50 min	1
PCR - amplification	2h 30 min	1
Amplicons isolation	1h e 40 min	2
Amplicons hybridization with Veracode Bead	3h	2
Veracode Bead Plate washing	10 min	2
Veracode Bead Plate reading	1h 10 min-96 polymorphisms 3h 30 min-384 polymorphisms	2

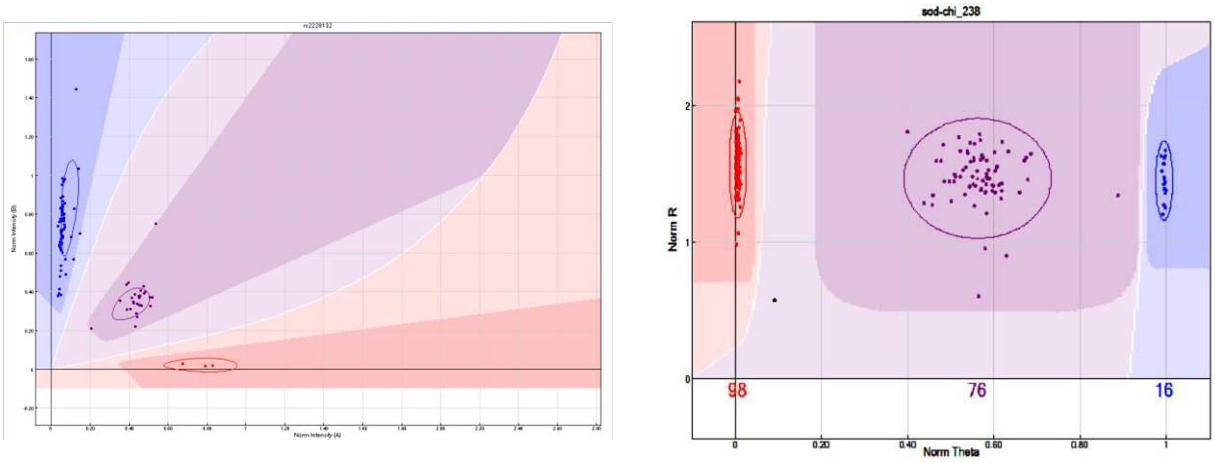
Data generated using the BeadXpress Reader can be analyzed with Illumina's GenomeStudio™ data analysis software, which performs automated genotype clustering and calling.

The software permits the association between the fluorescence data and the correspondent genotype. A clusterization algorithm assembles in three groups the fluorescence values related to each sample based on the presence of only one (in the case of homozygous genotypes) or two (heterozygous genotype) fluorescence signals. This process lets the software call the genotypes for each SNP investigated. The holographic code links the genotype call to a specific sample.

The graphical display of genotypes in GenomeStudio is a Genoplot, with data points color-coded for the call (red = AA genotype, purple = AB genotype, blue = BB genotype).

Genotypes are called for each sample with a dot by two coordinates representing their signal intensity (norm R) on the y-axis and Allele Frequency (Norm Theta) on the x-axis relative to canonical cluster positions (dark shading) for a given SNP marker.

Genome Studio normalizes the intensity of each fluorescence, so the Theta angle between the sample dot and the x-axis is converted into a value on the x-axis (Norm Theta), while in the y-axis the fluorescence intensity is reported as a Theta angle normalized value (Norm R), approximately ranging from 0 to 1 (with some exceptions for outliers which can reach values of 2 or more) (Figure 10).



A **B**

Figure 10. A) Samples clusterization according to their genotypes: intensity values are not normalized. **B)** Normalized graphical representation: the theta angle between the sample and the x-axis is converted into a value on the x-axis and on the y-axis is reported the fluorescence intensity as a theta-normalized angle function.

According to the fluorescence distribution, three clusters are shown: usually, homozygous genotypes creates vertical clusters while heterozygous ones generate a more spread cluster. Ideally, Theta angle has to assume a value of 0 for the AA homozygous genotype (red dots), 0.5 for the AB heterozygous genotypes (purple dots) and 1 for BB homozygous genotypes (blue dots). Generally, an analysis is considered good if homozygous dots lie between 0 and 0.2 and between 0.8 and 1 (this means that BeadXpress Reader detects a 20% fluorescence from one dye and the 80% from the other one), while heterozygous dots are positioned between 0.2 and 0.8 on the x-axis. A good Norm R value is around 1: conventionally, if it is lower than 0.4 the analysis is considered failed (Figure 11).

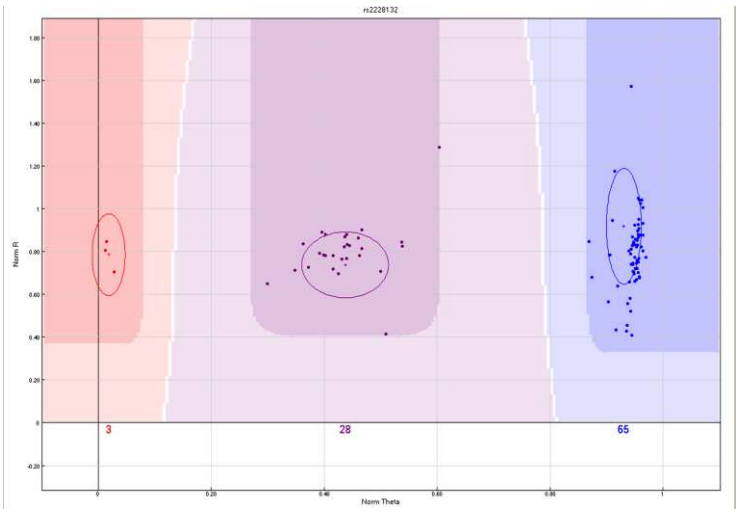


Figure 11. Example of a good result: Norm R is higher than 0.4, and heterozygous dots lie between 0.2 and 0.8 while the homozygous dots have values lower than 0.2 and higher than 0.8.

Based on the obtained clusterization quality, the software assigns to each SNP a score called “gene train score” ranging from 0 to 1 (index of an optimal clusterization).

Data analysis

Genome Studio software performs a basic data analysis; then operators can improve call rates and evaluate assay performance, sample quality, and locus performance by following simple guidelines.

Analysis begins with an overall evaluation of the assay performance and determination of which samples, if any, require reprocessing or removal. Clustering should be done after inclusion of reprocessed samples and removal of failed or suboptimal samples, allowing for a more detailed evaluation of sample quality. Each locus can then be evaluated for editing or zeroing (excluding) to optimize call rates. In particular, these parameters need to be analyzed: controls and Gencall score.

Controls

48 sample-dependent, sample-independent, and contamination controls are all built into the GoldenGate assay. These controls provide a way to assess the overall performance of samples, reagents, equipment, and BeadChips. During preliminary sample quality evaluation, samples falling outside the expected performance parameters should be highlighted for additional analysis. Failure in these controls could indicate a processing failure in a specific step or poor DNA quality.

Gencall score

Before evaluating the quality of SNP clusters, it is important to highlight samples that have poor performance in the genotyping assay. The GenCall score is a quality metric, ranging from 0–1, calculated for each genotype (data point). GenCall scores generally decrease in value the farther a sample data point is from the center of its cluster.

Each SNP is evaluated based on the angle, dispersion, and overlap of clusters and intensity.

Problematic samples are identified by a scatter plot of the call rate as a function of the 10% GenCall score (10% GC or p10 GC) (Fig. 12).

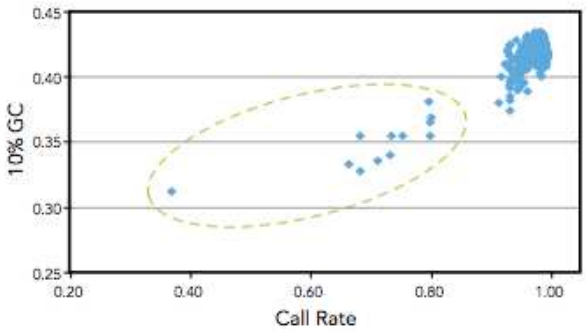


Figure 12. Poorly performing samples are obvious outliers from the majority of samples when 10% GC Score is plotted against sample call rate (green oval).

Poorly performing samples—those with low sample call rates, low 10% GC scores, or outliers from the main population cluster—should be considered for reprocessing or exclusion from the project.

4.2.8. Candidate Gene and Polymorphisms Selection

For this analysis, both cohorts from study sets one and two were used. Target genes were firstly selected on the basis of literature search (Pubmed-MEDLINE) prioritizing TFs, NRs, and cytokines clearly implicated in the regulation of transporters and phase I and II enzymes during inflammation stimuli. Of particular attention was a point to the modulation of membrane carrier (i.e. ABCB1, ABCC1, ABCC2, ABCG2, SLCO1B1) and metabolic proteins (i.e. UGT1A, CES, CYPs) strictly involved in the FOLFIRI (irinotecan, 5-FU) drug pathway. Successively, genetic variants for each candidate genes were chosen using the tagging polymorphisms (TagSNPs) approach. The selection of the TagSNPs, covering the genetic diversity of the targeted genes, was performed using the genotype frequencies data downloaded from HapMap website (http://hapmap.ncbi.nlm.nih.gov/); filter parameters were HapMap CEU database and minor allele frequency (MAF) ≥ 0.05. This search permitted to obtain records about variants located in the exonic and intronic region of the gene. The downloaded genotype data were then used to predict the tagSNPs using the Tagger program implemented in Haploview (http://www.broad.mit.edu/mpg/haploview, Broad Institute, Cambridge, MA, USA).The panel of selected TagSNPs was further integrated with additional variants located in the 5'- and 3' untranslated region (5'-UTR, 3'-UTR) of the gene and chosen by a screening of NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/) through the following criteria: Pubmed citation and MAF ≥ 0.05 in HapMap CEU population. At the end a set of 246 molecular markers in 22 candidate genes encoding for NRs (PXR, LXR-A/B, FXR, RXR-A/B/G, CAR, VDR, PPAR-A/G/D, HNF4A, HNF1A), TFs and related pathway (STAT-3, NFkB1, IKBKB, CHUK) and key pro-inflammatory cytokines (TNF, IL-1B, IL-6, INFG), were selected (see appendix 2) and introduced in the pharmacogenetic analysis.

Only the samples and polymorphisms with call rates > 90% were retained in the final report. The markers excluded because not compatible with a successful GoldenGate genotyping and the residual six polymorphisms of the selected pool were tested through an allelic discrimination reactions using predesigned TaqMan single-nucleotide polymorphism (SNP) genotyping assays. All the commercial TaqMan assays were purchased from Applied Biosystems (www.appliedbiosystems.com), and the analyses were performed with the Applera TaqMan Universal Master mix on ABI 7500 (AB Applied Biosystems Foster City, CA) according to the manufacturer's instructions. Positive and negative control samples were included in each analysis.

4.3. Toxicities and their management costs

For the purpose of this analysis, the cohort of study set one was used. Although all the clinical and pathological data of the patients were prospectively collected; patients' clinical records were reviewed to evaluate all interventions performed during FOLFIRI therapy to manage the chemotherapy-related toxicity. Toxicity data related to the chemotherapy according to the physician assessment, were recorded at each chemotherapy cycle, until treatment discontinuation for any reason, and classified according to the NCI-CTC v.3.0. Since this analysis was conducted on the basis of the Italian Public Health Care System payer, only direct costs reimbursed by the Health Care System were considered, such as those directly associated with the management of toxicity during the entire trial. They include hospitalization, outpatient or ambulatory services, i.e. laboratory analyses, supportive therapies, physician visits, instrumental examinations, facility fees and other health practitioner fees. Indirect costs (transportation to and from care centers, accommodation for family members, etc.) and productivity costs (lost or impaired work or leisure time due to morbidity), were not analyzed.

Since Italian Public Health Care System reimbursements are managed at the Regional level, and the study participating centers are settled in the North East part of Italy, mostly in the Friuli Venezia Giulia Region, the cost estimates used in the analysis were derived from 1) the Health Agency of the Friuli Venezia Giulia Region, (www.egas.sanita.fvg.it/); 2) the current version of the Friuli Venezia Giulia Regional Health System website (Table 4).

Together with an oncologist from the hospital, a database has been built to consider all direct expenses commonly associated with a specific toxic event of a precise NCI-CTC grade. An example with neutropenia febrile or not is visible in figure 13. The obtained esteem of the cost for each toxic event was then adjusted to the medical record information of every patient available.

The current estimated costs for hospitalization for Diagnosis Related Groups (DRG) in Italy (defining groups of patients having the same clinical-pathological condition) were used**.**

The costs used in the analysis were measured in Euros and are based on the year 2015.

Table 4. Costs used in the model.

Medical Intervention	Cost	Source
Hospitalization		
1 day in hospital (ordinary recovery)	1800	Local Economic Data*
1 day in day-hospital	470	Local Economic Data*
Blood/Platelet transfusion (per die)	39	FVG Regional Health System website
Support therapy costs		
GCSF standard treatment for 2 days	22	Health Agency of FVG Region
Mucositis support therapy for 1 week	17	Health Agency of FVG Region
Instrumental exams costs		
Ultrasound abdomen	80	FVG Regional Health System website
Colonoscopy	120	FVG Regional Health System website
Echocardiogram	66	FVG Regional Health System website
Blood analysis (hemochrome)	51	FVG Regional Health System website
Health practitioner hourly rate		
Physician (minimum hourly rate)	27	FVG Regional Health System website
Nurse (minimum hourly rate)	12	FVG Regional Health System website
*UGT1A1*28* genetic test	163	FVG Regional Health System website

*Based on the definition of Diagnosis Related Groups in Italy.
GCSF, Growth Colony Stimulating Factor; FVG, Friuli-Venezia Giulia

	Cost neutropenia G1	Cost neutropenia G2	Cost neutropenia G3	Cost neutropenia G4	Cost neutropenia G5
Tot		€ 199.22	€ 319.87	€ 3,120.75	€ 3,114.37
	febrile	€ 207.77	€ 328.43	€ 3,129.31	
Prescribed therapy (posologic unit per die)		1	1	1	1
Prescribed therapy (days)		2	3	4	4
Oncologist (work hours per day of hospitalization)		0.5	0.5	1	1
Nurse (work hours per day of hospitalization)		1	1	1	1
Hospitalization (days)		0	0	1.5	1.5
Total number of additional blood analyses		3	5	6	6
Ultrasound abdomen		0	0	0	0
Colonscopy		0	0	0	0
Gastroscopy		0	0	0	0
Echocardiogram		0	0	0	0
Holter		0	0	0	0
Stress test		0	0	0	0

Figure 13. Neutropenia example from database

4.4. Implementation of the Pharmacogenetic diagnostic service

4.4.1. Specific Pyrosequencing and Sanger sequencing conditions and materials

Operative instructions for the Pharmacogenetic diagnostic service have been written in order to perform a robust analysis. Following are reported all the materials and conditions to perform sequencing via two techniques: Pyrosequencing and Sanger Sequencing.

*4.4.1.1. DPYD*2A/ IVS14 + 1 G > A*

rs3918290

For both Pyrosequencing and Sanger Sequencing protocols, the PCR reagents are the same (see methods section 4.2.3. for PCR details). Exception is made for PCR primers and reagents concentrations in the final PCR mix. Details could be found below in table 5.

Table 5. Reagents concentration adopted for the PCR mix of *DPYD**2A analysis.

	Stock concentrations	**Mix concentrations**	
Reagents		Pyrosequencing Protocol	Sanger Sequencing
GeneAmp® PCR Gold Buffer	10X	1X	1X
AmpliTaq Gold DNA polymerase	5Units/µL	0,02Units/µL	0,02Units/µL
$MgCl_2$ solution	25mM	2mM	2,5mM
Forward primer	50µM	0,2µM	0,5µM
Reverse primer	50µM	0,2µM	0,5µM
Template DNA	-	1µL	1,2µL
dNTPs	25mM	0,125mM	0,25mM
H_2O	-	Up to 50µL	Up to 30µL (Optional 5% DMSO)

To the mix concentration is added at the end 1µL of patient's DNA in each well and in duplicate. Sequences of the primers used in both cases are reported below in table 6.

Table 6. Sequence for the forward and reverse primers adopted for the PCR mix of *DPYD**2A analysis.

Primers name Pyrosequencing	**Sequence 5'-3'**	**T°melt**	**Length bases**	**%GC**
Forward Primer	CGGCTGCATATTGGTGTCAA	66,9°C	20	50
[Biotinylated] Reverse Primer	[Btn]CACCAACTTATGCCAATTCTCTTGT	65,9°C	25	40
Primers name Sanger Sequencing				
Forward Primer	ATGTATGGCCCTGGACAAAG	63,6°C	20	50
Reverse Primer	ATGCATCAGCAAAGCAACTG	64,0°C	20	45

PCR thermic profiles to be applied to the thermocycler are herein reported for the two different techniques (table 7).

Table 7. Thermic profile for PCR amplifications for the *DPYD**2Aanalysis (both Pyro and Sanger sequencing protocol).

		Pyrosequencing		**Sanger sequencing**	
Steps	**Process**	Temperature	Time	Temperature	Time
0	Lid Pre-Heat	105°C		105°C	
		95°C	On hold	95°C	On hold
1	Enzyme activation	95°C	5 min	95°C	10 min
2	Denaturation	95°C	30 secs	95°C	30 secs
3	Annealing	60°C	30 secs	61°C	30 secs
4	Elongation	72°C	1 min	72°C	30 secs
5		Go to step 3 and repeat 35 cycles		Go to step 3 and repeat 40 cycles	
6	Final Step	72°C	5 min	72°C	7 min
7		16°C	On hold	16°C	On hold

*4.4.1.2. DPYD*13/ 1679 T > G*

rs55886062

Likely for *DPYD*2A* both Pyrosequencing and Sanger Sequencing materials and conditions are reported below (see Methods section 4.2.3. for PCR details) (Table 8).

Table 8. Reagents concentration adopted for the PCR mix of *DPYD**13 analysis.

	Stock concentration	**Mix concentration**	
Reagents		Pyrosequencing	Sanger sequencing
GeneAmp® PCR Gold Buffer	10X	1X	1X
AmpliTaq Gold DNA polymerase	5Units/µL	0,02Units/µL	0,02Units/µL
$MgCl_2$ solution	25mM	2mM	2,5mM
Forward primer	50µM	0,25µM	0,5µM
Reverse primer	50µM	0,25µM	0,5µM
Template DNA	-	1µL	1µL
dNTPs	25mM	0,125mM	0,25mM
H_2O	-	Up to 50µL	Up to 30µL

To the mix concentration is added at the end 1µL of patient's DNA in each well and in duplicate. Sequences of the primers used in both cases are reported below in table 9.

Table 9. Sequence for the forward and reverse primers adopted for the PCR mix of *DPYD**13 analysis.

Primers name Pyrosequencing	**Sequence 5'-3'**	**T°melt**	**Length bases**	**%GC**
[Biotinylated] Forward Primer	[Btn]CCTTTTGGTCTTGCTAGCGC	65,9°C	20	55%
Reverse Primer	AGTTTTGGTGAGGGCAAAACC	66°C	21	47,6%
Primers name Sanger Sequencing				
Forward Primer	CGGATGCTGTGTTGAAGTGATTT	67,1°C	23	43,4%
Reverse Primer	GTGTAATGATAGGTCTTGTCAAA TAGT	59,4°C	27	33,3%

PCR thermic profiles to be applied to the thermocycler are herein reported for the two different techniques (table 10).

Table 10. Thermic profile for PCR amplifications for *DPYD**13 analysis (both Pyro and Sanger sequencing protocol).

		Pyrosequencing		**Sanger Sequencing**	
Steps	**Process**	Temperature	Time	Temperature	Time
0	Lid Pre-Heat	105°C		105°C	
		95°C	On hold	95°C	On hold
1	Enzyme activation	95°C	10 min	95°C	10 min
2	Denaturation	95°C	30 secs	95°C	30 secs
3	Annealing	52,5°C	30 secs	63°C	30 secs
4	Elongation	72°C	1 min	72°C	30 secs
5		Go to step 3 and repeat 35 cycles		Go to step 3 and repeat 37 cycles	
6	Final Step	72°C	5 min	72°C	7 min
7		16°C	On hold	16°C	On hold

4.4.1.3. DPYD 2846 A > T

rs67376798

Likely to for *DPYD*2A and *13* both Pyrosequencing and Sanger Sequencing materials and conditions are reported below (see methods section 4.2.3. for PCR details) (table 11).

Table 11. Reagents concentration adopted for the PCR mix of *DPYD* 2846 A>T analysis.

	Stock concentration	**Mix concentration**	
Reagents		Pyrosequencing	Sanger Sequencing
GeneAmp® PCR Gold Buffer	10X	1X	1X
AmpliTaq Gold DNA polymerase	5Units/µL	0,02Units/µL	0,02Units/µL
$MgCl_2$ solution	25mM	2,5mM	2,5mM
Forward primer	50µM	0,25µM	0,5µM
Reverse primer	50µM	0,25µM	0,5µM
Template DNA	-	1µL	1µL
dNTPs	25mM	0,125mM	0,25mM
H_2O	-	Up to 50µL	Up to 50µL

To the mix concentration is added at the end 1µL of patient's DNA in each well and in duplicate. Sequences of the primers used in both cases are reported below in table 12.

Table 12. Sequence for the forward and reverse primers adopted for the PCR mix of *DPYD* 2846 A>T analysis.

Primers name Pyrosequencing	**Sequence 5'-3'**	**T°melt**	**Length bases**	**%GC**
[Biotinylated] Forward Primer	[Btn]GCAGTACCTTGGAACATTTGGT	63,9°C	22	45,4%
Reverse Primer	AGGTCATGTAGCATTTACCACAGT	62,5°C	24	41,6%
Primers name Sanger Sequencing				
Forward Primer	GGTCCAAAAATGAGAAAAAGTTAGCC	66,3°C	26bp	38,4%
Reverse Primer	TCTCTCTAATGTTGTGGCTGATGA	64,8°C	24bp	41,6%

PCR thermic profiles to be applied to the thermocycler are herein reported for the two different techniques (table 13).

Table 13. Thermic profile for PCR amplifications for *DPYD* 2846 A>T analysis (both Pyro and Sanger sequencing protocol).

		Pyrosequencing		Sanger Sequencing	
Steps	**Process**	Temperature	Time	Temperature	Time
0	Lid Pre-Heat	105°C		105°C	
		95°C	On hold	95°C	On hold
1	Enzyme activation	95°C	10 min	95°C	10 min
2	Denaturation	95°C	30 secs	95°C	30 secs
3	Annealing	56°C	30 secs	61°C	30 secs
4	Elongation	72°C	30 secs	72°C	30 secs
5		Go to step 3 and repeat 35 cycles		Go to step 3 and repeat 40 cycles	
6	Final Step	72°C	5 min	72°C	7 min
7		4°C	On hold	16°C	On hold

4.4.2. PCR check

The PCR amplification was verified by the electrophoretic run in a 3% agarose gel before both the Pyrosequencing and Sanger sequencing analysis.

4.4.3. Sequencing

As seen in section 4.2.4., Pyrosequencing need a further primer as reagent: the "sequence primer".

Herein are reported the sequence of *DPYD* sequence primers (table 14).

Table 14. Sequence for the *DPYD* sequence primers for *DPYD**2A, *DPYD**13 and *DPYD* 2846 A>T analyses.

Sequence primers name	**Sequence 5'-3'**	**T°melt**	**Length bases**	**%GC**
*DPYD*2A*	GGCTGACTTTCCAGACA	57,1°C	17	52,9%
*DPYD*13*	CTTCAAAAGCTCTTCG	51,3°C	16	43,7%
DPYD 2846	CACAGTTGATACACATTTTCT	52,2°C	20	35%

*4.4.4.1. UGT1A1*28*

*UGT1A1*28*_TA (6/7) (rs8175347) is a STR polymorphism characterized by a variable number of dinucleotide (TA) repeats on the promoter region ranging from 5 to 8.

Herein are reported the reaction mix (table 15), the forward and reverse primers sequences (table 16) and the optimal PCR conditions employed for the analysis (table 17), according to the best results obtained by the setting up procedure.

Table 15. Reagents concentration adopted for the PCR mix for *UGT1A1*28* analysis.

Reagent	Stock concentration	Mix concentration	
		Sanger Sequencing	Automated fragment analysis
GeneAmp® PCR Gold Buffer	10X	1X	1X
AmpliTaq Gold DNA polymerase	5Units/µL	0.02Units/µL	PolyTaq 1U
$MgCl_2$ solution	25mM	2mM	3mM
Forward primer	50µM	0.2µM	0.5µM
Reverse primer	50µM	0.2µM	0.5µM
Template DNA	-	1µL	1µL
dNTPs	25mM	0.125mM	0.25mM
H_2O	-	Up to 50µL	Up to 20µL

Table 16. Sequence for the forward and reverse primers adopted for the PCR mix for *UGT1A1*28* analysis.

Primers name Sanger Sequencing	Sequence 5'-3'	T°melt	Length bases	%GC
*UGT1A1*28* FW primer	GTCACGTGACACAGTCAAACATTAACTTGG	71,3°C	30	
*UGT1A1*28* FW primer	[HEX]TTTGCTCCTGCCAGAGGTT	64,9°C		
Primers name Automated fragment analysis				
Forward Primer	GTCACGTGACACAGTCAAACATTAACTTGG	67,1°C	23	43,4%
Reverse Primer	[HEX]TTTGCTCCTGCCAGAGGTT	59,4°C	27	33,3%

Table 17. Thermic profile for PCR amplifications for *UGT1A1*28* analysis (both Pyro and Sanger sequencing protocol).

Steps	Process	Sanger Sequencing		Automated fragment analysis	
		Temperature	Time	Temperature	Time
0	Lid Pre-Heat	105°C		105°C	
		95°C	On hold	95°C	On hold
1	Enzyme activation	95°C	5 min	95°C	10 min
2	Denaturation	95°C	30 secs	95°C	30 secs
3	Annealing	60.0°C	30 secs	62.5°C	30 secs
4	Elongation	72°C	1 min	72°C	30 secs
5		Go to step 3 and repeat 35 cycles		Go to step 3 and repeat 38 cycles	
6	Final Step	72°C	5 min	72°C	10 min
7		16°C	On hold	16°C	On hold

In Table 18 are listed the molecular weights, expressed in base pairs (bp), of each fragment analyzed.

Table 18. Number of TA repeats and the corresponding molecular weight expressed in bp.

Number of TA repeats	Molecular weight (bp)
5	94
6	96
7	98
8	100

Figure 14 represents the typical UGT1A1 fragment analysis electropherograms and illustrates the difference in the length depending on the number of TA repeats.

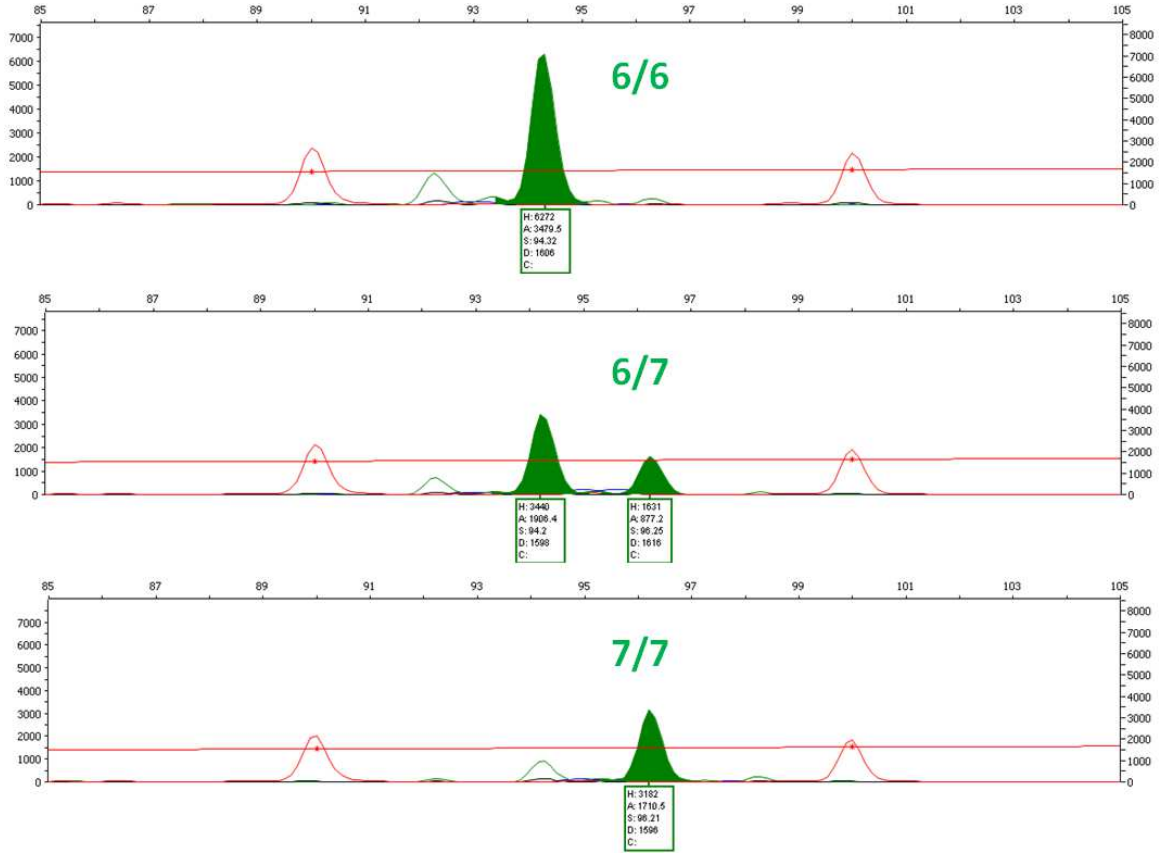


Figure 14. Electropherograms corresponding to the UT1A1 6/6; 6/7; 7/7 genotypes.

4.5. Statistical methods

The distribution of individual characteristics was evaluated by simple descriptive statistics and variable frequency, between the groups, was compared using the Fisher exact tests.

To assess the relative excess risk of G3/4 WHO FOLFIRI-related toxicity between patients with different polymorphic status and to control for confounding factors, it was carried out by multivariate logistic regression procedure (including all available prognostic factors: gender, age, primary sites). Odds ratios (OR) and the corresponding 95% confidence intervals (CI) were calculated and adjusted for the patient's clinical-demographic characteristics (i.e. gender, age, first tumor site, radical surgery and adjuvant chemotherapy). Dominant, recessive, and additive genetic models were considered for each polymorphism by combining heterozygous with homozygous genotypes; the best-fitting genetic model was selected according to Wald chi-square test. A $p<0.05$ was considered as statistically significant (two-sided). All the analyses were carried out with Stata 12.1.

The overall cost for each patient in the study was calculated as the sum of the costs of the medical interventions performed to manage the chemotherapy-related toxicity, from the beginning of treatment with FOLFIRI until the discontinuation of FOLFIRI.

Since the distribution of costs data is usually skewed or do not follow normal distribution (due to a small number of cases whose cost of toxicity management was very high), generalized linear models with a gamma distribution and log-link function were generated to assess the association between the cost per patient and *UGT1A1*28* genotype, as well as the occurrence of at least a grade 4 toxicity during the first 6 cycles of treatment. To adjust the costs predicted by the model by the patient characteristics, age, gender, adjuvant chemotherapy (yes/no), and the total number of chemotherapy cycles, were included. The statistical difference between groups of patients is expressed as Regression Coefficient and respective 95% CI. Recycled predictions method was used to convert the Regression Coefficient values into costs (adjusted for the above specified patient characteristics) per patient in each group. The same method was used to calculate the difference between costs in the different groups (incremental costs). Costs and incremental costs were reported as the mean and 95% CI.

The association between any type of grade 4 toxicity and genotype was tested by univariate Fisher's exact test. Odds Ratio and 95% CI were computed. Differences between groups of patients for continuous variables (i.e. the number of cycles) were assessed using Kruskal-Wallis test.

5. Results

Results are described according to the three aims defined in chapter three.

5.1. Study sets

5.1.1. Study set one

The discovery set

Study set two is composed of 250 Italian mCRC patients receiving FOLFIRI regimen in first-line setting and characterized for hematologic and GI toxicity profile. From July 2002 to November 2004, 250 patients who met all the inclusion criteria were enrolled in a prospective multicenter study, involving 15 institutions from north-eastern Italy [50]. The study was coordinated and sponsored by the Centro di Riferimento Oncologico, National Cancer Center of Aviano (Aviano, Italy). All subjects were Caucasian. A total of 1715 cycles were administered (median 7 range 1-20). Reduction of treatment occurred in 25 patients (12.19%) whereas treatment was stopped in 23 patients (11.22%) based on protocol criteria. The most common cause for discontinuation of treatment was disease progression (34.78%), other causes were toxicity (30.43%), death during treatment (17.39%), refusal of the patient to continue therapy (8.70%) and loss at follow-up (8.70%).

In order to make toxicity evaluations homogeneous and more consistent and relevant from a clinical perspective, the tolerance to the treatment was evaluated both during the 1st cycle of chemotherapy (acute toxicity), before the 6 cycles of chemotherapy (cumulative toxicity) and by the end of treatment. Toxicity related to the 1st cycle was evaluated in all the 250 patients that entered the study, whereas toxicity developed over the 6 cycles of chemotherapy was evaluated in 216 patients (86.4%). For the cost-analysis also the toxicities developed during the entire duration of the patient's treatment were evaluated. Among them, 206 completed the 6 chemotherapy cycles, and 1 patients discontinued the therapy due to toxicity developed before the 6th cycle. Generally, patients well tolerated the treatment with severe toxicity (G3-G4) and occurred in only 17 out of the 250 patients (6.8%) during the 1st cycle, and in 41 out of the 216 patients (19.0%) before the 6 cycles of chemotherapy. The most frequent severe adverse effect was neutropenia, experienced by 9 (3.6%) patients during the 1st cycle (6 patients exhibited G3, 2 patients had G4 neutropenia and 1 G4 febrile neutropenia).

Before the 6 cycles, G3-G4 neutropenia was observed in 27 out of the 216 (12.5%) patients, 22 of which exhibited G3 and 7 others exhibited G4. Six patients, experiencing G3-G4 hematological toxicity, exhibited also G3 non-hematological toxicity (2 with G3 diarrhea and neutropenia and 1 of them also with G3 mucositis, 1 with G3 mucositis and anemia, 1 with nausea and vomiting with neutropenia G3 and another one with neutropenia G4 and leukopenia G3 and the last one with asthenia G3 concomitant with neutropenia and leukopenia G4). The occurrence of febrile neutropenia was observed in 3 patients (all patients over the 6 cycles). Grade 3 anemia was observed in one patient during the 1st cycle, and in other four cases considering the entire duration of treatment. whereas G3-G4 thrombocytopenia was never observed. The predominant non-hematological toxicities before the 6 cycles of therapy were diarrhea and nausea/vomiting, which affected a total of 97 (44.9%) and 93 (43.1%) patients, respectively. However, G3-G4 of these side effects were only observed in 12 (5.6%) and 3 patients (1.2%), respectively. During the 1st cycle, 4 G3 diarrheas (1.6%) and 1 G3 nausea and vomiting (0.4%) were observed. Details on eligibility, modalities of treatment and toxicity data collection are reported in the methods section [50]. Patient's demographics (age and gender) and clinical information (treatment, toxicity, tumor site) are reported in table 19.

	N	(%)
All the eligible patients	250	(100.0)
Gender		
Male	162	(64.8)
Female	88	(35.2)
Age (Median, range)	60.6	26-75
First Tumor Site		
Left colon	100	(55.9)
Right colon	79	(44.1)
Rectum	71	(28.4)
Stage at diagnosis		
I	5	(2.0)
II	20	(8.0)
III	65	(26.0)
IV	160	(64.0)
Radical Surgery of primary tumor		
Yes	200	(80.0)
No	50	(20.0)
Adjuvant Chemotherapy		
Yes	82	(32.8)
No	168	(67.2)
Adjuvant Radiotherapy		
Yes*	33	(46.5)
No	38	(53.5)
Toxicity		
Diarrhea (grade 3-4)	35	(14.2)
Neutropenia (grade 3-4)	26	(10.5)

Table 19. Demographic and Clinical Characteristics at Study Entry of Study set one.

*only rectum

5.1.2. Study set two

The replication set

Study set two is composed of 167 Eastern Canadian mCRC patients receiving FOLFIRI-based regimens and characterized for hematologic and GI toxicity profile.

All patients received an 180 mg/m^2 intravenous dose of irinotecan every 2 weeks, and 75 patients also received co-treatments with bevacizumab, an experimental drug, or a placebo. This multi-institution prospective study involved patient recruitment from 2003 to 2012 at three medical centers in eastern Canada: Hotel-Dieu de Québec in Québec City, QC; Hotel-Dieu de Lévis in Lévis, QC; and The Ottawa Hospital in Ottawa, ON. The ethics committee of each participating institution approved the study protocol, and all patients signed a written informed consent before entering the study. Eligibility criteria included patients (18–90 years old) initiating their first irinotecan-based chemotherapy with a histologically confirmed metastatic colorectal cancer, a life expectancy of at least 3 months, and a good performance status (Eastern Cooperative Oncology Group ≤ 2). Details on eligibility, modalities of treatment and toxicity data collection are reported in the methods section [75,120,121]. Table 20 summarizes patient demographics (age and gender) and clinical information (treatment, toxicity, tumor site).

Table 20. Demographic and Clinical Characteristics at Study Entry of Study set two.

	N	**(%)**
All the eligible patients	167	(100.0)
Gender		
Male	110	(65.9)
Female	57	(34.1)
Age (Median, range)	61.5	
First Tumor Site		
Colon	122	(73.1)
Rectum	42	(25.1)
Unknown	3	(1.8)
Regimen		
FOLFIRI	167	(100)
Co-treatment	69	(41.3)
Avastin/Bevacizumab	6	(3.6)
Other drugs		
Toxicity		
Diarrhea (grade 3-4)	24	(14.4)
Neutropenia (grade 3-4)	28	(16.8)

The toxicity endpoints consisted of both GI and hematologic toxicities and were analyzed separately. For GI toxicities, all patients completed a daily report of GI toxicities during the first 14 days of each cycle to record the incidence and severity of nausea, vomiting, and diarrhea. For hematologic toxicities, laboratory parameters were collected before each cycle of chemotherapy and/or when the treatment was delayed. GI toxicity was assessable for all patients except for one who died before toxicity assessment, and another who did not fill out the GI toxicity diary, while hematologic toxicity was evaluable for 166 of 167 patients. Severe neutropenia was reported in the 16.8% of cases (28/167) while grade 3-4 GI side effects in the 14.4 % (24/167). Genetic markers predictive of severe neutropenia and GI toxicity were determined separately. Toxicities' severity was evaluated prospectively and according to National Cancer Institute Common Terminology Criteria for Adverse Events version 3.0 criteria.

5.2. The exploration setting

5.2.1. Study design

The study design for the exploration analysis is described in figure 15. We genotyped all the 250 mCRC patients from the study set one for 246 tagSNP, although we retained only the samples and polymorphisms with call rates >90%. Thus, three samples were excluded from the study and genotype data of 247 patients were eventually available. We then attempt to replicate the statistically significant associations ($p<0.05$) between genetic markers and risk of severe neutropenia and GI toxicity found in the study set one also in the study set two. If it was confirmed a concordant effect in the two populations, we found validated markers ($p<0.05$) of severe neutropenia or GI toxicity. If not, we tried to look at some pharmacokinetic parameters as a proof of concept. If such an analysis found positive results, validated markers with concordant effect on pharmacokinetic parameter were found.

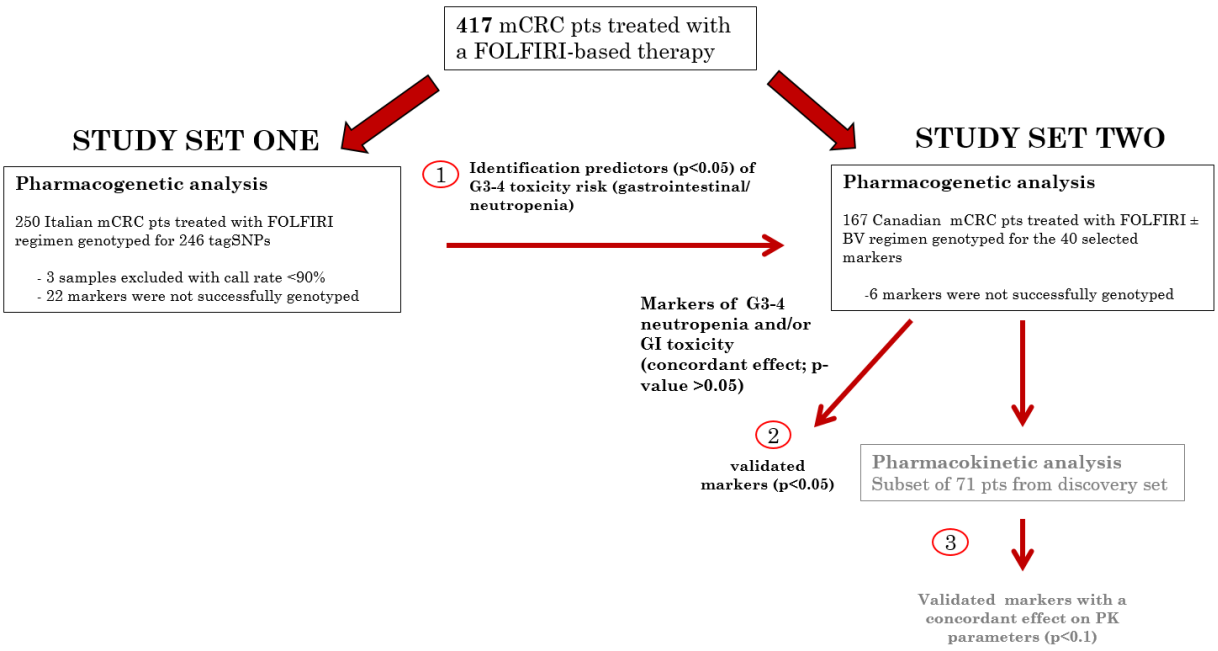


Figure 15. Exploration study design

5.2.1. Genetic Analyses

After genomic DNA had been extracted from peripheral blood using the High Pure PCR Template Preparation Kit (Roche Diagnostics GmbH, Mannheim, Germany), each DNA sample was genotyped using the Illumina BeadXpress platform, based on Golden Gate technology, and the allelic discrimination method based on the TaqMan system. In particular, a 192-plex and 48-plex Illumina VeraCode

GoldenGate Genotyping Assay (Illumina Inc., San Diego, CA) were developed using the Assay Design Tool (ADT) available through the Technical Support in Illumina website (https://illumina.com). The list of selected polymorphisms was uploaded and the bioinformatics tool that assigned for each variant a final score, correlated with the quality of the assay (ranging from 0 to 1.1), and a designability score, associated with the capability to design the selected assay (ranging from 0 to 1). Only the assays with high final score (≥ 0.6) and an optimal designability (=1) were considered compatible with successful GoldenGate genotyping and introduced in the ordered custom panel. VeraCode GoldenGate system uses the high-density BeadArray technology in combination with an allele-specific extension, adapter ligation, and amplification assay procedure; the analytical practice was performed according to manufacturer's protocol. The VeraScan software (version 2.0) was employed for fluorescence detection; GenomeStudio V2011.1 tool (Illumina Inc.,) was applied for genotype clustering with a polymorphism call-threshold of 0.25 (on a scale 0-1). The cluster generated by the program were further visually inspected and manually reviewed to ensure high-quality data. The control dashboard was checked to evaluate the overall quality of the performed analyses and to exclude samples with low performance. Sample replicates were introduced in each analysis to assess the robustness of output records and to provide duplicate data helping in the clustering redefinition. The genotype data were further randomly validated by direct Sanger sequencing.

Genotyping analyses with BeadXpress platform lead to success in 224 assays. The markers excluded and the residual six polymorphisms of the selected pool were tested through an allelic discrimination reactions using predesigned TaqMan single-nucleotide polymorphism (SNP) genotyping assays. 16 markers were successfully retested with the TaqMan method as well as the six remaining polymorphisms of the selected set. Positive and negative control samples were included in each analysis.

Only the samples and polymorphisms with call rates > 90% were retained in the final report. Three samples were excluded from the study since they didn’t reach the call rate threshold of 90% probably due to a low DNA quality. Eventually, genotype data of 247 patients were available. Replicate samples included in the analyses presented an average concordance rate of 100%; the random verification of the obtained genotypes through direct sequencing resulted concordant in all cases.

The genotyping on the study set two was performed with Sequenom platform. It was successfully designed for all candidate polymorphisms except three. All 167 samples were effectively genotyped, and the average genotype call rate was 0.98% (range: 0.94%-100%).

5.2.2. Patients and Toxicity

The main demographic and clinical characteristics of the two study population (study set 1 and 2) are reported in Table 19 and 20. In the subset of 247 successfully genotyped cases from study set one, severe neutropenia (G3-G4) was still the most frequent severe toxicity (35/247 cases, 14.2%); diarrhea, nausea and/or vomiting of high grade (G3-G4) occurred in 26 out of 247 cases (10.5%) and represented the predominant non-hematological toxicities.

5.2.3. Markers of Neutropenia

Eighteen genetic variations in genes encoding for four NRs (HNF4α, PXR, PPARs, VDR), one TFs (NFkB) and one cytokine (TNF) emerged as significant predictors ($p<0.05$) of severe neutropenia over the entire course of chemotherapy in the cohort of study set one. The genotype distribution by neutropenia severity was reported in Table 21; the minor allele frequencies (MAFs) of these polymorphisms were checked and resulted in line with the data reported for Caucasian population (http://www.ncbi.nlm.nih.gov/snp).Of the 18 markers identified, ten were found to be associated with an increased risk of developing grade 3-4 neutropenia with ORs ranging from 1.68 to 16.08 while the remaining eight were indicated to have a protective effect on toxicity development with ORs ranging from 0.16 to 0.60. None of these associations were replicated in the study set two. Although not statistically significant ($p>0.05$) probably due to a population-specific effect, three markers resulted to have in both study populations a concordant predictive effect on severe neutropenia risk according to the same genetic model. Particularly, two HNF4A variants (rs6093976, rs745975) and one VDR markers (rs12717991) emerged as protective factors toward the likelihood of developing grade 3-4 neutropenia (dominant model). The results of logistic regression analysis were summarized in Table 22.

Table 21. Genotype frequency distribution for G3-G4 vs. G0-G2 cumulative neutropenia in the subset of study set one (n=247 mCRC patients) according to gene polymorphisms.

Genes	SNP	Base change	Genotype frequency					
			G0-G2 (n=212)			G3-G4 (n=35)		
			AA	Aa	aa	AA	Aa	aa
HNF4A	rs2425637	G>T	0.269	0.509	0.222	0.171	0.429	0.400
HNF4A	rs3212183	T>C	0.332	0.455	0.213	0.229	0.371	0.400
HNF4A	rs3212197	C>T	0.852	0.143	0.005	0.771	0.143	0.086
HNF4A	rs6093976	C>T	0.632	0.335	0.033	0.886	0.114	0.000
HNF4A	rs6093978	C>T	0.445	0.403	0.152	0.657	0.257	0.086
HNF4A	rs6130615	C>T	0.741	0.250	0.009	0.657	0.257	0.086
HNF4A	rs745975	G>A	0.561	0.401	0.038	0.771	0.200	0.029
HNF4A	rs2425640	G>A	0.434	0.415	0.151	0.543	0.429	0.028
NR1I2	rs16830505	A>G	0.816	0.170	0.014	0.686	0.286	0.028
NR1I2	rs7643645	A>G	0.345	0.476	0.179	0.588	0.177	0.235
PPARD	rs2076169	T>C	0.793	0.193	0.014	0.686	0.257	0.057
PPARG	rs2972164	T>C	0.246	0.531	0.223	0.412	0.412	0.176
PPARG	rs880663	T>C	0.585	0.396	0.019	0.800	0.200	0.000
NFKB1	rs230539	A>G	0.521	0.379	0.100	0.343	0.457	0.200
TNF	rs3093662	A>G	0.848	0.138	0.014	0.686	0.314	0.000
VDR	rs11168287	A>G	0.260	0.514	0.226	0.114	0.714	0.172
VDR	rs11574026	C>T	0.802	0.179	0.019	0.657	0.343	0.000
VDR	rs12717991	G>A	0.393	0.422	0.185	0.600	0.171	0.229

Table 22. Cumulative neutropenia in the study set one and two according to gene polymorphisms (SNP). Only the associations with p-value <0.05 are reported in the study set one; markers with the same predictive effect and genetic model in both cohorts are evidenced in grey.

OR and 95% CI are meant for G3-G4 vs G0-G2								
			Study set one			Study set two		
Genes	**SNP**	**Base change**	**Model**	**OR (95% CI) [a]**	**p-value**	**Model**	**OR (95% CI) [a]**	**p-value**
HNF4A	rs2425637	G>T	Additive	1.88 (1.05-3.38)	0.035	Dominant	1.51 (0.53-4.27)	0.442
HNF4A	rs3212183	T>C	Recessive	2.50 (1.11-5.56)	0.026	Additive	1.20 (0.65-2.21)	0.558
HNF4A	rs3212197	C>T	Additive	2.29 (1.07-4.91)	0.033	Dominant	0.47 (0.09-2.35)	0.357
HNF4A	rs6093976	C>T	Dominant	0.16 (0.05-0.56)	0.004	Dominant	0.67 (0.27-1.64)	0.379
HNF4A	rs6093978	C>T	Dominant	0.41 (0.18-0.93)	0.033	Recessive	1.83 (0.59-5.68)	0.292
HNF4A	rs6130615	C>T	Recessive	16.08 (2.89-89.62)	0.002	Additive	0.70 (0.27-1.79)	0.457
HNF4A	rs745975	G>A	Dominant	0.33 (0.13-0.85)	0.021	Dominant	0.68 (0.29-1.61)	0.380
HNF4A	rs2425640	G>A	Additive	0.60 (0.36-0.98)	0.042	Recessive	3.35 (0.82-13.66)	0.092
NR1I2	rs16830505	A>G	Dominant	2.33 (1.01-5.34)	0.046	Additive	1.87 (0.83-4.23)	0.132
NR1I2	rs7643645	A>G	Dominant	0.28 (0.13-0.61)	0.001	Recessive	3.01 (1.01-9.01)	0.049
PPARD	rs2076169	T>C	Recessive	8.99 (1.31-61.85)	0.026	Dominant	0.18 (0.02-1.42)	0.103
PPARG	rs2972164	T>C	Dominant	0.36 (0.16-0.82)	0.015	Dominant	1.34 (0.49-3.69)	0.565
PPARG	rs880663	T>C	Additive	0.36 (0.16-0.82)	0.015	---	---	---
NFKB1	rs230539	A>G	Additive	1.68 (1.02-2.78)	0.043	Recessive	0.42 (0.09-1.87)	0.254
TNF	rs3093662	A>G	Dominant	2.56 (1.10-5.96)	0.029	---	---	---
VDR	rs11168287	A>G	Dominant	3.12 (1.02-9.56)	0.046	Recessive	1.98 (0.73-5.43)	0.182
VDR	rs11574026	C>T	Dominant	2.36 (1.06-5.23)	0.035	Recessive	1.55 (0.63-3.77)	0.339
VDR	rs12717991	G>A	Dominant	0.36 (0.16-0.82)	0.015	Dominant	0.84 (0.35-2.00)	0.695

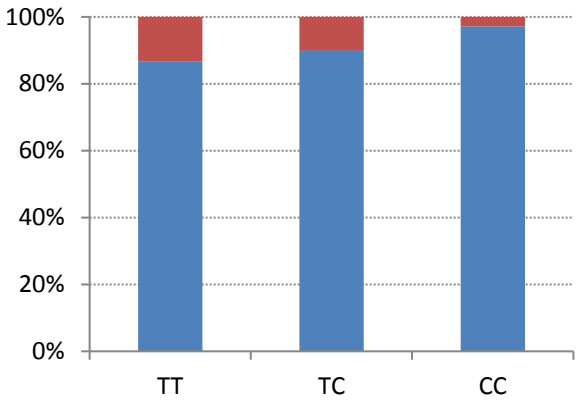
[a] Estimated from the unconditional logistic regression, adjusted for gender, age, cancer site, stage at diagnosis, radical surgery, and adjuvant chemotherapy.

5.2.4. Markers of Gastrointestinal toxicity

Twenty-two polymorphic variants in genes encoding for five NRs (HNF4α, CAR, PPARs, RXRs, VDR), one TF (STAT-3) and three cytokines (TNF, INFγ, IL-6) were found to be significantly ($p<0.05$) associated with the risk of developing severe GI toxicity during the entire course of chemotherapy. The genotype distribution by GI grade was reported in Table 23; the MAFs of these polymorphisms resulted in line with the data reported for Caucasian population (http://www.ncbi.nlm.nih.gov/snp). Of the 22 markers emerged, the majority (n=14) were predictors of an increased chance of having grade 3-4 GI toxicity with ORs ranging from 1.72 to 20.74 while the remaining (n=8) were correlated to an inferior risk of severe GI toxicity with ORs ranging from 0.12 to 0.51. One STAT-3 marker (rs1053004) was successfully replicated in the Canadian cohort and was inversely associated with grade 3–4 severe GI toxicity. The polymorphic rs1053004-C allele was significantly associated with a decreased toxicity risk in the study set one (OR=0.51, p=0.045) according to the additive model (figure 16 A). Similarly, rs1053004-C resulted to exert a protective dominant effect towards toxicity development in the validation cohort (OR=0.39, p=0.043) (figure 16 B). All other associations between genetic markers and risk of severe GI toxicity found in the cohort of study set one were not replicated in the study set two ($p >0.05$).

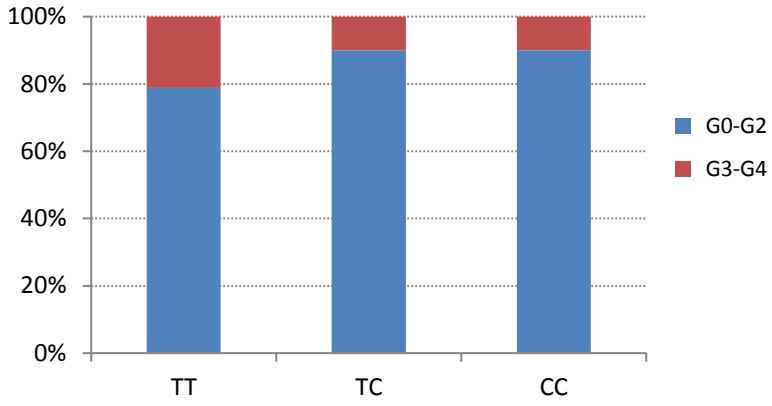
Among others association between genetic markers and risk of severe GI toxicity, although not significant, probably due to a population-specific effect, eight markers resulted to have in both study populations a comparable predictive effect on GI toxicity according to the same genetic model. Particularly, three HNF4A variants (rs1800961, rs2071197, rs6031587), PPARD-rs2076169, RXRG-rs380518, and VDR-rs11574077 were found to predict an increased likelihood of developing grade 3-4 GI toxicity while INFG-rs2069716 and RXRG-rs380518, resulted in a detrimental factor for severe GI side effects (VDR rs11574077 allele frequency visual displayed in figure 17 A and B). The data obtained by logistic regression analysis were shown in Table 24.

A) Study set one



A.M. OR 0.51, 95%CI (0.27-0.99), P=0.045

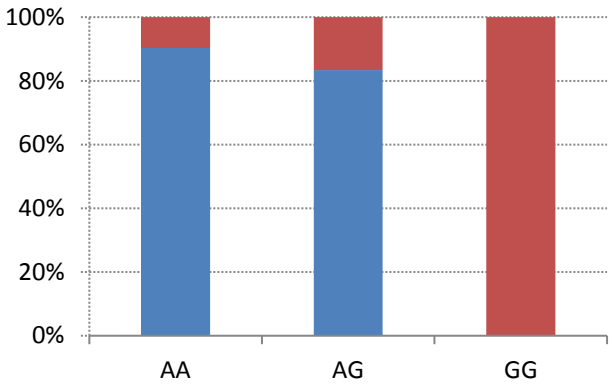
B) Study set two



D.M. OR 0.39, 95%CI (0.15-0.97), P=0.043

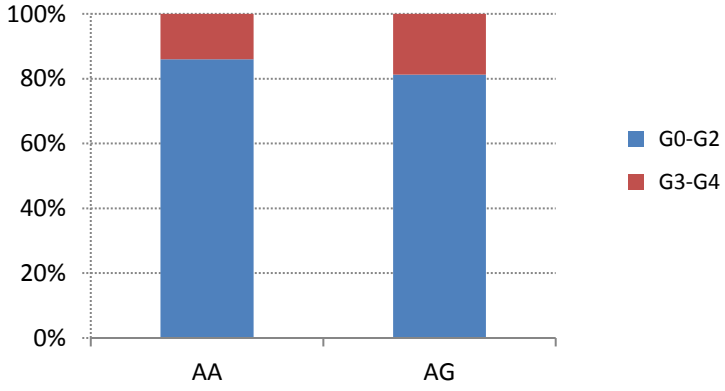
Figure 16. Genotype frequency (in percentage) for *STAT3 rs1053004.*
Figure 16A and B regard the STAT3 rs1053004 allele frequency in study sets one and two respectively.

A) Study set one



A.M. OR 4.46, 95%CI (1.43-13.96), P= 0.010

B) Study set two



A.M. OR 1.44, 95%CI (0.37-5.63), P= 0.601

Figure 17. Genotype frequency (in percentage) for *VDR rs11574077*
Figure 17A and B regard VDR rs11574077 allele frequency in study set one and two respectively

Table 23. Genotype frequency distribution for G3-G4 vs. G0-G2 cumulative gastrointestinal toxicity in the study set one (n=247 mCRC patients) according to gene polymorphisms.

Genes	SNP	Base change	Genotype frequency					
			G0-G2 (n=221)			G3-G4 (n=26)		
			AA	Aa	aa	AA	Aa	aa
HNF4A	rs1800961	C>T	0.986	0.014	--	0.885	0.115	--
HNF4A	rs2071197	G>A	0.869	0.131	0.000	0.654	0.308	0.038
HNF4A	rs6031587	C>T	0.896	0.104	0.000	0.731	0.192	0.077
HNF4A	rs6093976	C>T	0.651	0.317	0.032	0.808	0.192	0.000
NR1I3	rs2307424	C>T	0.471	0.407	0.122	0.192	0.731	0.077
NR1I3	rs4073054	T>G	0.388	0.416	0.196	0.423	0.539	0.038
PPARA	rs9626736	A>G	0.416	0.425	0.159	0.154	0.654	0.192
PPARD	rs2076169	T>C	0.783	0.203	0.014	0.731	0.192	0.077
RXRG	rs380518	T>C	0.832	0.145	0.023	0.615	0.385	0.000
RXRG	rs4657437	C>A	0.394	0.468	0.138	0.500	0.462	0.038
RXRG	rs283695	G>A	0.389	0.475	0.136	0.231	0.577	0.192
RXRG	rs3767344	G>C	0.249	0.498	0.253	0.385	0.385	0.230
RXRG	rs157880	C>T	0.751	0.235	0.014	0.769	0.154	0.077
RXRB	rs2744537	G>T	0.670	0.298	0.032	0.654	0.231	0.115
VDR	rs11574077	A>G	0.932	0.068	0.000	0.846	0.115	0.039
VDR	rs4760648	C>T	0.356	0.480	0.164	0.154	0.577	0.269
VDR	rs2853564	T>C	0.332	0.514	0.154	0.539	0.423	0.038
TNF	rs1800629	A>G	0.227	0.773	--	0.080	0.920	--
TNF	rs1800630	C>A	0.654	0.223	0.123	0.461	0.308	0.231
STAT3	rs1053004	T>C	0.441	0.405	0.154	0.577	0.385	0.038
INFG	rs2069716	A>G	0.808	0.192	--	0.960	0.040	--
IL6	rs2069840	C>G	0.430	0.448	0.122	0.615	0.308	0.077

Table 24. Cumulative gastrointestinal toxicities in the cohort of study set one and two according to gene polymorphisms (SNP). Only the associations with p-value <0.05 are reported in the study set one; markers with the same predictive effect and genetic model in both cohorts are evidenced in gray. Validate marker are underlined.

OR and 95% CI are meant for G3-G4 vs. G0-G2								
Study set one						Study set two		
Genes	**SNP**	**Base change**	**Model**	**OR (95% CI) [a]**	**p-value**	**Model**	**OR (95% CI) [a]**	**p-value**
HNF4A	rs1800961	C>T	Dominant	11.51 (2.02-65.61)	0.006	Dominant	3.27 (0.57-18.69)	0.182
HNF4A	rs2071197	G>A	Dominant	4.42 (1.75-11.17)	0.002	Dominant	1.43 (0.45-4.58)	0.546
HNF4A	rs6031587	C>T	Additive	3.24 (1.40-7.51)	0.006	Additive	2.21 (0.86-5.70)	0.100
HNF4A	rs6093976	C>T	Additive	0.36 (0.13-0.98)	0.046	Recessive	1.92 (0.35-10.65)	0.457
NR1I3	rs2307424	C>T	Dominant	3.38 (1.23-9.27)	0.018	Additive	0.66 (0.35-1.24)	0.196
NR1I3	rs4073054	T>G	Recessive	0.14 (0.02-0.92)	0.041	Dominant	0.63 (0.26-1.50)	0.293
PPARA	rs9626736	A>G	Additive	2.14 (1.14-4.00)	0.018	Dominant	1.90 (0.71-5.08)	0.203
PPARD	rs2076169	T>C	Recessive	20.74 (2.88-149-18)	0.003	Recessive	6.26 (0.36-109.01)	0.208
RXRG	rs380518	T>C	Dominant	4.69 (1.73-12.70)	0.002	Dominant	2.15 (0.82-5.66)	0.120
RXRG	rs4657437	C>A	Additive	0.48 (0.24-0.99)	0.047	Dominant	2.54 (0.90-7.22)	0.080
RXRG	rs283695	G>A	Additive	1.80 (1.04-3.12)	0.036	Recessive	0.43 (0.10-1.98)	0.282
RXRG	rs3767344	G>C	Dominant	0.35 (0.13-0.93)	0.034	Dominant	0.75 (0.28-2.03)	0.574
RXRG	rs157880	C>T	Recessive	11.70 (20.03-67.58)	0.006	Dominant	1.44 (0.58-3.59)	0.435
RXRB	rs2744537	G>T	Recessive	5.34 (1.42-20.07)	0.013	---	---	---
VDR	rs11574077	A>G	Additive	4.46 (1.43-13.96)	0.010	Additive	1.44 (0.37-5.63)	0.601
VDR	rs4760648	C>T	Additive	2.09 (1.13-3.84)	0.018	Additive	0.71 (0.36-1.42)	0.338
VDR	rs2853564	T>C	Additive	0.38 (0.18-0.78)	0.008	Additive	1.42 (0.72-2.78)	0.310
TNF	rs1800629	A>G	Dominant	4.07 (1.13-14.72)	0.032	---	---	---
TNF	rs1800630	C>A	Additive	1.72 (1.01-2.94)	0.047	---	---	---
STAT3	rs1053004	T>C	Additive	0.51 (0.27-0.99)	0.045	Dominant	0.39 (0.15-0.97)	0.043
INFG	rs2069716	A>G	Dominant	0.12 (0.01-0.97)	0.047	Dominant	0.88 (0.20-3.88)	0.862
IL6	rs2069840	C>G	Dominant	0.40 (0.17-0.95)	0.039	---	---	---

[a] Estimated from the unconditional logistic regression, adjusted for gender, age, cancer site, stage at diagnosis, radical surgery, and adjuvant chemotherapy.

5.3. The cost-analysis

5.3.1. Study set one

Patients, rates of toxicity, and *UGT1A1*28* genotypes

From the study set one population, we identified a subgroup of patients free of well acknowledged genetic risk factor for FOLFIRI-related toxicity in which study irinotecan-related toxicity management costs. Therefore, since the study population one received homogeneously a first-line FOLFIRI treatment, we excluded patients carrying the variants conferring DPD deficiency (rs3918290, rs67376798, rs55886062 in *DPYD*), thus a genetic predisposition to develop fluoropyrimidine-related toxicities. Seven heterozygous patients for *DPYD* alleles increasing the risk of 5-fluorouracil toxicity were excluded from the present analysis to avoid the confounding effect of other genetically-related toxicities. Among these seven patients, one grade 4 toxicity, two grade 3 toxicities, two grade 2 toxicities requiring treatment interruption/delay, were reported, while two patients did not develop any toxicity.

The final group considered for the current cost-analysis included 243 patients. The main demographic and clinical characteristics of the subset population of the study set one are reported in table 25.

The distribution of the **28* genotypes among 243 patients is: 109 (44.9%) **1/*1*, 112 (46.1%) **1/*28*, and 22 (9.1%) **28/*28* (Hardy-Weinberg equilibrium test, P>0.05).

During the entire course of therapy, 28 patients (11.5%) did not develop toxicity of any grade, 144 (59.3%) developed grade 1 or 2 as maximum grade toxicity, and 17 (7.0%) developed grade 4 toxicity (Table 26). No toxic death was reported.

The number of cycles of FOLFIRI administered to the patients did not differ by genotype: the median is 8 (range, 2-20) for **1/*1*, 8 (1-20) for **1/*28*, and 7 (1-16) among **28/*28*.

Table 25. Patient demographic and clinical characteristics.

	N	**(%)**
All the patients	243	(100.0)
Gender		
Male	160	(65.8)
Female	83	(34.2)
BMI (Median, IQR)	24.5	22.4 – 27.5
Age (Median, IQR)	61	55-68
First Tumor Site		
Left colon	96	(39.5)
Right colon	78	(32.1)
Rectum	69	(28.4)
Stage at diagnosis		
I	5	(2.1)
II	20	(8.2)
III	63	(25.9)
IV	155	(63.8)
Radical Surgery of primary tumor		
Yes	195	(80.2)
No	48	(19.7)
Adjuvant Chemotherapy		
Yes	80	(32.9)
No	163	(67.1)
Adjuvant Radiotherapy		
Yes*	32	(13.2)
No	37	(15.2)
Number of cycles of first-line FOLFIRI (Median, IQR)	8	6-12

*only rectum

BMI, Body Mass Index; IQR, Inter-Quartile Range; FOLFIRI, 5-Fluorouracil, Leucovorin, Irinotecan

Table 26. Common toxicities among the 243 patients in the study.

Toxicity	**Patients**			
	Any grade		Grade 4*	
	N of patients	%	*N of patients*	*%*
Any kind of toxicity	215	88.5	17	7.0
Non-haematological				
Diarrhoea	110	45.3	-	-
Alopecia	35	14.4	-	-
Nausea/ Vomiting	117	48.1	-	-
Mucositis	60	24.7	1	0.4
Infection without severe neutropenia	11	4.5	2	0.8
Asthenia	62	25.5	-	-
Hepatic (hyperbilirubinemia)	3	1.2	-	-
Cardiac	15	6.2	2	0.8
Neurologic	5	2.1	1	0.4
Hematological				
Neutropenia	89	36.6	8	3.3
Leukopenia	62	25.5	3	1.2
Severe neutropenia with fever	7	2.9	3	1.2

* according to the NCI-CTC version 3.0
N, number

5.3.2. Genetic Analyses

*UGT1A1*28* polymorphism was analyzed by automated fragment analysis based on capillary gel electrophoresis comparing the amplified product weight with an internal standard added to each sample. The *UGT1A1*28* allele was present at the homozygous status in 22 out of 243 patients (9.0%), at the heterozygous status in 112 patients (46.1%). One hundred nine patients out of 243 (44.9%) were wild type for the considered polymorphism. The frequency of the *UGT1A1*28* allele was 32%.

All 250 patients enrolled in this study were genotyped for three *DPYD* variants (*DPYD*-rs3918290, *DPYD*-rs67376798 and *DPYD*-rs55886062). All the *DPYD* variant alleles were detected only in heterozygosity. Seven patients (2.8%) carried at least one variant allele for any polymorphism (four for *DPYD*-rs3918290, three for *DPYD*-rs67376798, none for *DPYD*-rs55886062).

5.3.3. Association between *UGT1A1*28* and costs of toxicity management

The median cost of chemotherapy-related toxicity management per patient, including patients that do not develop toxicity of any grade, was56€ (inter-quartile range 1-324; range 0-30,500). Among all patients, the cumulative cost of all toxicities management was 283,768€. The cost associated with hospitalization has the biggest impact on the overall cost of toxicity management (Figure 18). This accounts for a total of 125 days of hospitalization, with 22/243 (9.0%) patients hospitalized for toxicity during the course of the entire treatment.

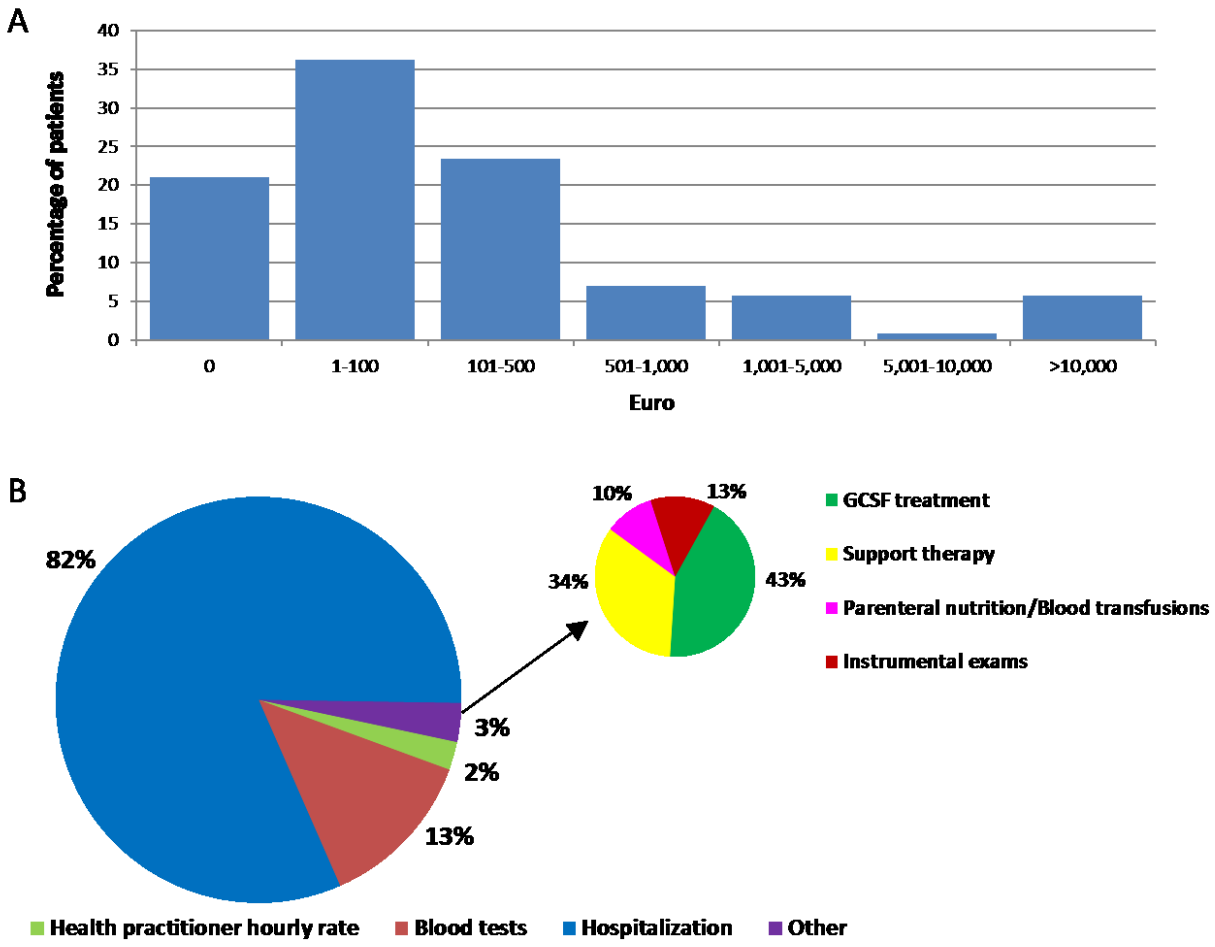


Figure 18: Distribution and composition of the cost of toxicity management. *Figure 18A shows the distribution of the cost of toxicity management calculated for each patient. Figure 18B shows the fraction of the cost related to each medical intervention.*

The association between the *UGT1A1*28* genotype and the cost of toxicity management is herein assessed. The mean cost per patient is 812€ for **1/*1* patients; 1,119€ for **1/*28* patients, and 4,886€ for **28/*28* patients (Table 27). The difference is statistically significant for **28/*28* versus **1/*1 (P*=0.024), for **1/*28* versus **1/*1* (*P*<0.001), and for **1/*28* versus **28/*28* patients (*P*<0.001) (Table 27). The incremental cost per patient for the **28/*28* patients when compared to the **1/*1* patients is 4,074€ (95% CI 1,808 – 6,340), and 307€ (95% CI 232 – 383) when compared to the **1/*28* patients.

Table 27. Association between *UGT1A1*28* polymorphism and the costs for toxicity.

***UGT1A1* Genotype**	**Number of patients**	**Mean Predicted Cost per patient* (95% CI) (Euro)**		**Regression Coefficient**	**95%CI**	***P*-value**	**Regression Coefficient**	**95%CI**	***P*-value**
**1/*1*	109	812	(653 – 970)	Ref[#]					
**1/*28*	112	1,119	(885 - 1,353)	0.32	0.04 – 0.60	0.024	Ref[#]		
**28/*28*	22	4,886	(2,611 - 7,160)	1.79	1.31 - 2.28	<0.001	1.47	0.99-1.95	<0.001

*by generalized linear model, adjusted for age, gender, adjuvant chemotherapy, and total number of chemotherapy cycles

[#] Reference Category for Regression Coefficients calculation

CI, Confidence Interval; Ref, Reference Category

5.3.4. Time course of toxicity events and cost of toxicity management

The evaluation of the occurrence of toxicity over time is the basis for a correct analysis of the impact of genotype on toxicity and hence cost of toxicity management and treatment. In Figure 19, a visual display of the course of toxicity grade by patient over time, considering the maximum toxicity of any type, is provided. The graph suggests a different distribution of grade 4 toxicities among the three genotypes. Six cycles are the standard duration for a treatment course in these patients, after which treatment decisions (like continuing treatment, treatment holidays, and others) are taken.

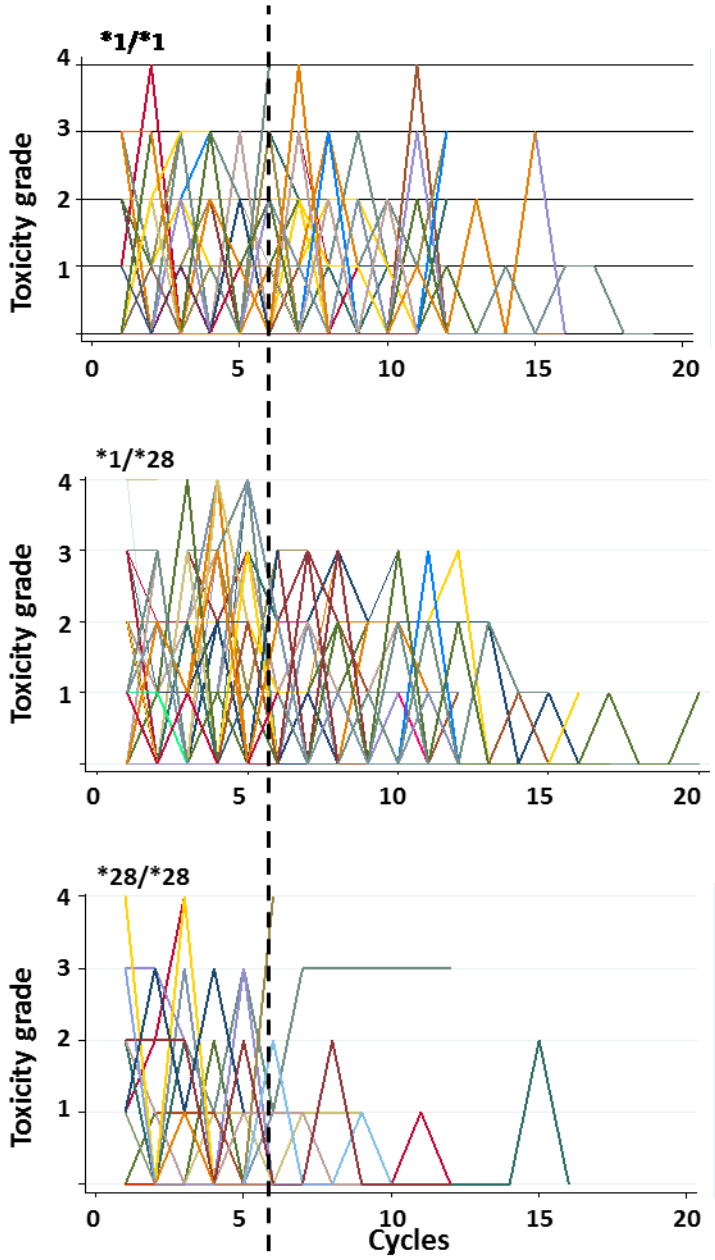


Figure 19. Visual display of toxicity occurrence by cycle of chemotherapy and *UGT1A1*28* genotype.

Figure 19 has three graphs, one for each group of patients stratified by *UGT1A1*28* genotype. Each line represents the maximum toxicity grade by cycle for each patient included in the study. The vertical dashed line corresponds to the 6th cycle. The x-axis represents the number of cycles of chemotherapy; the y-axis represents the maximum grade of toxicity of any type occurred during each cycle.

Hence, the evaluations of the occurrence of grade 4 toxicity within the first 6 cycles shows that **1/*28* and **28/*28* patients appeared more likely to develop grade 4 toxicity (11/134, 8.2%), compared to **1/*1* patients' (1/109 patients, 0.9%). The difference is no longer significant when considering grade 4 toxicities developed during the entire course of treatment (Table 28).

Table 28. Occurrence of any kind of grade 4 toxicities based on *UGT1A1*28* polymorphism.

***UGT1A1* Genotype**	**Number of patients**	***Toxicity Before Sixth Cycle***		**OR (95% CI)**	***P*-value**	***Toxicity Entire course***		**OR (95% CI)**	***P*-value**
		G0-G3	**G4 (%)**			**G0-G3**	**G4 (%)**		
**1/*1*	109	108	1 (0.9)	Ref		104	5	Ref	
**1/*28*	112	103	9 (8.0)	9.4 (1.2-75.8)	0.019	103	9	1.8 (0.6-5.6)	0.409
**28/*28*	22	20	2 (9.1)	10.8 (0.9-124.9)	0.073	19	3	3.3 (0.7-14.9)	0.131
**1/*28/ and*28/*28*	134	123	11 (8.2)	9.7 (1.2-76.1)	0.014	122	12	2.0 (0.7-6.0)	0.214

G, Grade; *CI, Confidence Interval; Ref, Reference Category*

The earlier occurrence (within the first 6 cycles) of grade 4 toxicity in *1/*28* and **28/*28* patients as compared to **1/*1* was hypothesized to impact the overall cost of chemotherapy-related toxicity management, due to a differential clinical management of patients experiencing grade 4 toxicity within or after 6 cycles of chemotherapy (Table 29). Indeed, the mean cost per patient (calculated by a generalized linear model) was higher for patients experiencing at least one grade 4 toxicity event within the first 6 cycles of chemotherapy as compared to the others (*P*<0.001) (Table 30).

Table 30. Association between the development of at least one grade 4 toxicity before the sixth cycle and the cost of toxicity management.

*****Maximum grade of toxicity***	**Number of patients**	**Mean Predicted Cost per patient* (95% CI) (Euro)**		**Regression Coefficient**	**95%CI**	***P*-value**
Grade 0-3	231	696	(606 – 785)	Ref[#]		
Grade 4	12	10,254	(4,452 – 16,056)	2.7	2.1 – 3.3	<0.001

*by generalized linear model, adjusted for age, gender, adjuvant chemotherapy, and total number of chemotherapy cycles

**developed within the first 6 cycles of treatment

[#] Reference Category for Regression Coefficients calculation

CI, Confidence Interval; Ref, Reference Category

To better account for the difference in cost of toxicity management observed between the three genotype groups, and related to a differential risk of grade 4 toxicity, the impact of hospitalization is evaluated. Hospitalization was the most expensive medical intervention related to toxicity management (Table 4). Out of 109 **1/*1* genotype patients, 6 (5.5%) were hospitalized for toxicity (mean days of hospitalization 6.4, range 2-12), versus 10 (8.2%) out of 122 **1/*28* patients (5.7 days, 2-16), and 6 (27.3%) out of 22 **28/*28* patients (4.8 days, 2-8). During the entire course of treatment, **28/*28* patients were more likely to be hospitalized compared to **1/*1* patients' (OR 6.4, 95% CI 1.8-22.4; *P*=0.005). The median number of days of hospitalization per patient in the three genotype groups was also statistically different (*P*=0.0051, by ANOVA Kruskal Wallis test).

Table 29. Description of the follow up toxicities of patients with a grade 4 toxicity episode of any type.

Patient	***Genotype***	***Cycle with first grade 4 toxicity episode***	***Grade 4 toxicity type***	***Additional cycles after toxicity***	***Additional toxicities, maximum grade***
Before Sixth Cycle					
1	**1/*1*	2	Neutropenia	14	Leukopenia G3, Constipation G1, Neutropenia G1, Nausea G2
2	**1/*28*	1	Neutropenia	1	Febrile Neutropenia G4
3	**1/*28*	1	Neutropenia	11	Alopecia G2, Constipation G1, Asthenia G2, Diarrhea G3, Anemia G2
4	**1/*28*	3	Neutropenia	14	Leukopenia G3*, Leukopenia G2, Neutropenia G3, Nausea and Vomiting G1, Diarrhea G3
5	**1/*28*	4	Mucositis	4	Diarrhea G3*, Neutropenia G2, Leukopenia G2, Mucositis G2
6	**1/*28*	4	Acute pancreatitis	8	Neutropenia G2, Diarrhea G2
7	**1/*28*	4	Neutropenia	4	Proctitis G2, Neutropenia G2, Leukopenia G2, Mucositis G2, Alopecia G2, Diarrhea G2
8	**1/*28*	5	Cardiotoxicity	0	None
9	**1/*28*	5	Neutropenia	10	Asthenia G2, Infection G1
10	**1/*28*	5	Febrile Neutropenia	7	Leukopenia G3*, Infection G1, Thrombocytopenia G1, Asthenia G1, Anemia G1, Neutropenia G2, Leukopenia G2, Anemia G1, Diarrhea G1
11	**28/*28*	1	Febrile Neutropenia	7	Anemia G2*, Cholinergic Syndrome G1*, Infection G4, Anemia G2, Diarrhea G2, Cardiotoxicity G2
12	**28/*28*	3	Hematochezia	9	Mucositis G1*, Diarrhea G1*, Nausea G1, Diarrhea G1, Anemia G1
From Sixth Cycle On					
13	**1/*1*	6	Neutropenia	0	Leukopenia G3*, Nephrotoxicity G2, Asthenia G3
14	**1/*1*	7	Cardiotoxicity	1	None
15	**1/*1*	7	Neutropenia	5	HF syndrome G1, Infection G1, Mucositis G1, Anemia G2, Diarrhea G3, Neurotoxicity G1
16	**1/*1*	11	Infection	1	Diarrhea G3*, Mucositis G3*
17	**28/*28*	6	Neurotoxicity	0	None

**Concomitant toxicities*

G, grade; HF syndrome, Hand-Foot Syndrome.

5.4. The Implementation setting

Sample route

The implementation modalities of the service of pharmacogenetic diagnostic at the unit of Clinical and Experimental pharmacology at the Institute of CRO-Aviano are herein illustrated. The sample route for the genotyping of relevant risk variant could be divided into three sections: the pre-analytical, analytical and post-analytical one. The genotyping service is performed by authorized service personnel on duty at the Clinical Pharmacology laboratory or fellow graduate after adequate educational training and supervised by the Clinical and Experimental Pharmacology unit executive in charge of the process. Seven working days from the sample acceptance are needed to complete the analysis.

Operating procedure

1. Arrival of patient's whole blood in a test tube Sarstedt S-Monovette® from 7.5ml containing anticoagulant (EDTA Potassium) at the laboratory of Experimental and Clinical Pharmacology (figure 20) located in the basement of the Institute, coming from the clinical chemistry laboratory located on the first floor. The clinical chemistry laboratory collects samples to be genotyped from: Oncologic unit A, B and C, other organizational structures of the CRO and eventually external centers to CRO.

Each biological sample has to be transported on ice to our unit accompanied by the completed survey with all the patients' characteristics and the type of analysis required for which treatment (see appendix 1).

2. The technicians first verify the integrity of the sample then the sample is noted in the "Diagnostic register", completed of: arrival date, first and last name, birth date, medical record number, genetic analysis requested and eventually other notes. In the digital "Diagnostic Laboratory notebook" is filled out a report of the technical analysis completed with the initials of the technician that is in charge to carry out the analysis (figure 20). The sample arrival is eventually recorded in the digital repository through the software DNLab where the request of pharmacogenetics analysis for that patient is listed as pending. The blood sample, awaiting processing, is temporarily stored at +4 °C.
3. Extraction of the DNA take place within 72h from the arrival with the automatic extractor BioRobot® EZ1 workstation system. 350µL of whole blood are used for the DNA collection. The remains are kept in an "ad-hoc" criobox named "Criobox diagnostic blood" stored at -80°C.
4. Long term are to be stored the following biological materials:

   - 1 tube of whole blood (1.5 ml) in the freezer at -80°C;

- 1 tube of buffy coat in the freezer at -80°C;

- 3 plasma aliquots (at least 500µL each) in the freezer at -80°C;

- 1 tube of DNA in the refrigerator at + 4°C.

5. Amplification by PCR following the operational instructions (see method's section 4.4.1.)
6. Verification of the PCR success through agarose gel at 3%. In the eventuality of unexpected results from the positive and/or negative controls, the previous step is repeated.
7. Genotyping is carried out using direct pyrosequencing or Sanger sequencing (fragment analysis) following the operational instructions (see method's section 4.2.4. and 4.2.6.).
8. Technical report filing. The technical report is filed using DNLab software and signed by the operator who performed the analysis.
9. Clinical report filing. The clinical interpretation of the genotyping results is signed by the clinical personnel of the Clinical and Experimental Pharmacology unit (Director or its delegate) with DNFirma4. The genotype-driven dose adjustment recommendations are given in accord with national and international guidelines reported in table 31.

Table 31. Recommended reductions of initial 5-fluorouracil (or capecitabine) and irinotecan dose.

***DPYD* variant**	**Initial dose reduction (%)**	**Ref**
*2A (c.1905+1G>A)	50	[104]
*13 (c.1679T>G)	50	[104]
c.2846A>T	50 -> 25	[104]
Compound *DPYD* heterozygosity	Switch drug	[104]
*UGT1A1*28*	70	[102]

10. Pharmacogenetic report access. The report is accessible from the G3 system to all medical personnel in the hospital unit. If requested, paper reports are issued as well, and upon the request enveloped and addressed to the doctor requiring it.
11. Documents storage. The "Diagnostic Laboratory notebook" as well the "diagnostic register" are stored in a dedicated collector in a dedicated cabinet supplied with keys, accessible only to authorized personnel and stored for at least 1 year.

The gel electrophoresis images, acquired through GelDoc Biorad system are stored on a dedicated PC in the room "robotics workshop 2" for at least one year in a special folder reserved exclusively for the pharmacogenetic diagnostic service. A paper copy is attached to the report in the "Diagnostic Laboratory Notebook".

All the graphs resulting from the sequencing and pyrosequencing analysis are stored in computers in the "Diagnostic" folder on the PC dedicated to the instrument in the room "Pharmacogenetics laboratory section instrumental" SOC-FSC.

Workflow

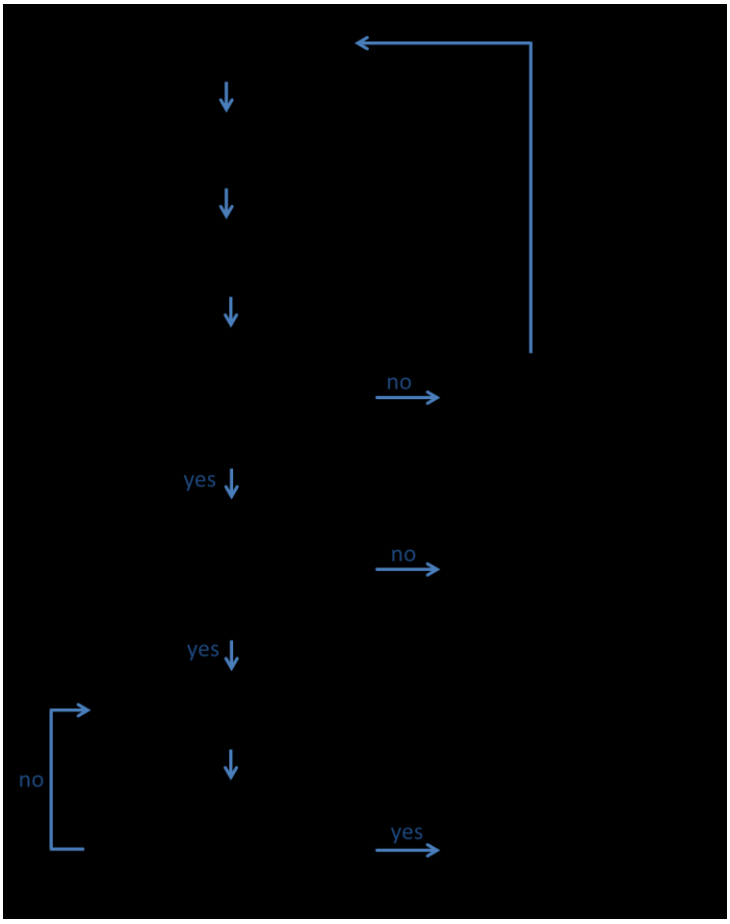


Figure 20. Workflow of the pharmacogenetic diagnostic service

5.4.1. Study set three (the diagnostic cohort)

Patients, rates of toxicity, and genotypes

Although patients for the presence of *DPYD*2A* and *UGT1A1*28* variants have been genotyped starting from September 2011, an institutionally acknowledged diagnostic PGx service was implemented at the national cancer Institute, CRO-Aviano on the first of January 2014. The prospective screening program includes *DPYD* SNPs (rs67376798, rs3918290, rs55886062) and the *28 allele risk variant in *UGT1A1* (rs8175347) conferring DPD and UGT1A1 deficiency, respectively and is performed according to the procedure previously described. Study set three is composed of 393 Caucasian patients who received a cancer diagnosis at Centro di Riferimento Oncologico, National Cancer Center of Aviano (Aviano, Italy) and resulted eligible for treatment with either fluoropyrimidine or irinotecan-based treatment. These patients were profiled for genetic variants known to impact the risk of severe toxicity development after either fluoropyrimidine (*DPYD* variants) or irinotecan-treatment (*UGT1A1* SNP). Twenty-three patients out of 393 patients instead were screened for the same genetic variants to investigate further the reason of severe toxic events development reported after chemotherapy administration. The most common toxicity cause for genetic profiling inquiry was severe diarrhea (65.2%) followed by severe neutropenia (35%) and mucositis (13%). The reported information was collected from a survey delivered to the prescriber oncologist (see appendix 1) after the request for genetic profiling. The median age of the patients was 61 years with a range of 19-92 years. Most frequently this service was given for neoadjuvant therapeutic setting (51.98%) followed by first line treatment for the metastatic disease (33.73%). The most common therapeutic regimen assigned, for whom it was requested the genetic diagnostic service, was Fluoropyrimidine + Platinum (Pt)-based chemotherapy with or without other drugs (53.3%), followed by Fluoropyrimidine in monotherapy (15.93%) or in association with RT (Capecitabine) (15.38%).

All the *DPYD* variant alleles were detected only in heterozygosity. Nine patients (2.58%) carried at least one variant allele for any polymorphism (four for *DPYD*-rs3918290, three for *DPYD*-rs67376798, none for *DPYD*-rs55886062). On 40 patients tested for *UGT1A1*28*, 27 carried 1 *28 allele (67.5%) and only 2 carried 2*28 risk allele (5.00%) for severe toxicity irinotecan-related.

Patient's characteristics and demographics are reported in table 32.

Table 32. Demographic and Clinical Characteristics at Study Entry of Study set three.

	N	(%)
All the Patients	393	100.00
Male	250	63.61
Female	143	36.39
Age (median, range)	61	19 - 92
Administration Setting		
Adjuvant	28	11.11
Neoadjuvant	131	51.98
First line	85	33.73
Other	8	3.17
Regimen		
FP +/- LV	29	15.93
FP + irinotecan +/- other	11	6.04
FP + Pt-CT	22	12.09
FP+ Pt-CT + other	75	41.21
FP + other	4	2.20
RT + Capecitabine	28	15.38
RT + Capecitabine +Other	9	4.95
Other	4	2.20

5.4.2. Outcome of the prospective genetic screening program

Before the program started, such genetic analyses were requested in the case of signaling from the medical personnel of an unexpected development of severe toxicity after either fluoropyrimidine or irinotecan administration. A visual display of the implementation of a prospective screening program for *DPYD* and *UGT1A1*28* at CRO-Aviano cancer Institute is reported in Figure 20.

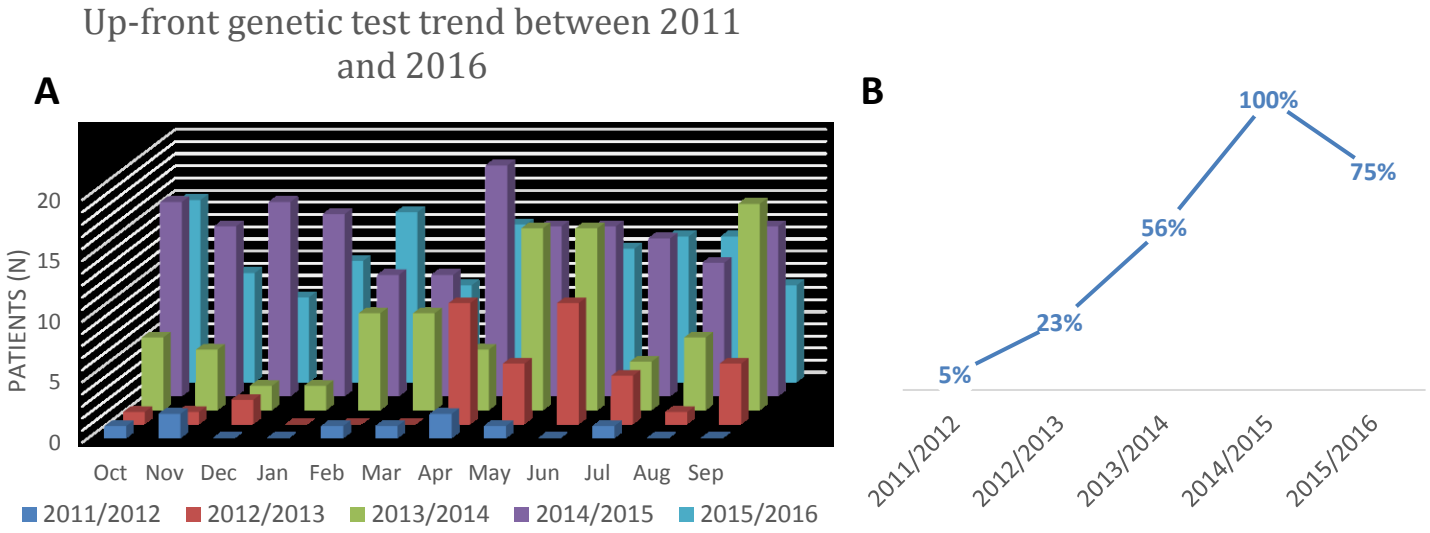


Figure 20. Entity and trend of up-front genetic tests implementation of in the oncologic clinical practice of CRO-Aviano

Figure 20A shows the distribution of genotyped patients per month since the up-front genetic test has been performed in the Institute. Figure 20B shows the percentage of patients genotyped per year.

As the implementation of the genetic screening program became fully integrated into the clinical routine also *DPYD* *13 and c.2846A>T were added to the *DPYD* screening. An overview is shown in table 33.

Table 33. Recommended reductions of initial 5-fluorouracil (or capecitabine) and irinotecan dose.

***DPYD* variant**	**Initial dose reduction (%)**	**Inclusion in screening program**	**Patients screened**
*2A (c.1905+1G>A)	50	Sep-11	386
*13 (c.1679T>G)	50	Jan-14	324
c.2846A>T	50 -> 25	Jan-14	324
*UGT1A1*28*	70	Apr-12	40

Of these 393 patients genotyped for risk variants at CRO-Aviano 374 (95.2%) were intended to be treated at the CRO-Aviano, whose oncologists personally prescribed the genetic profile both in a pre-therapeutic setting and to investigate the reason a precise patient developed a severe, life-threatening, toxicity. The dose reductions

recommended by the medical personnel for each individual *DPYD* and *UGT1A1* variants are displayed in table 33. Although part of the DPWG guidelines[102] *DPYD* 1236G>A is not yet part of the prospective screening program for *DPYD*.

5.4.2. Test results with a dose recommendation

Since the program started, 386 patients were screened for at least one *DPYD* variants and 40 for *UGT1A1*28*. Of these patients, 9 patients (2.58%) were found to carry at least one *DPYD* variants, and two patients (5.00%) were found to carry two *28 risk alleles for *UGT1A1*. Shown in table 34 are the *DPYD* variants that were screened for and of each variant the frequency in comparison to the literature. In table 35 are reported the recommendations for *UGT1A1*28*.

Table 34. *DPYD* variants in the pharmacogenetic diagnostic service implemented at CRO-Aviano.

***DPYD* variant**	**Tested pts, n**	**SNPs, n**	**CRO-Aviano (%)**	**Literature (%)**	**Ref**
*2A (c.1905+1G>A)	386	4	1.04	1.0 - 1.8	[92,123]
*13 (c.1679T>G)	324	0	0	0.1	[100]
c.2846A>T	324	5	1.54	1.0 - 1.4	[92,100]
Total	386	9	2.58	2.1 – 3.3	

Table 35. *UGT1A1*28* variant in the pharmacogenetic diagnostic service implemented at CRO-Aviano.

UGT variant	**Tested pts, n**	**SNPs, n**	**CRO-Aviano (%)**	**Literature (%)**	**Ref**
*UGT1A1*1/*28*	40	27	67.50	40 - 60	[124]
*UGT1A1*28/*28*	40	2	5.00	5 - 15	[124]

Regarding *DPYD* variants, four patients with *DPYD*2A* variant and five patients with *DPYD c.2846A>T* received a recommendation to reduce the dose by 50%. No patients with compound heterozygosity were found. Regarding *UGT1A1*28* SNP instead, within the five patients carrying two *28 variants, one received a 25% and four a 30% recommendation of dose reduction.

Twenty-three patients out of 393 (5.85%) were referred for toxicity from the CRO-Aviano oncologists. In table 36 are reported the main characteristics for this subgroup and in table 37 is reported the complete list of these patients. Two out of the 22 tested for *DPYD* presented heterozygosity for at least one *DPYD* variant (9.1%) and 4 out of the 6 tested for *UGT1A1*28* presented 1 *28 allele (66.7%). None was found to carry compound *DPYD* heterozygosity nor 2 copies of the *28 risk allele in *UGT1A1*. The most common toxicity the patients were reported for was diarrhea (60.9%) followed by neutropenia (34.8%). This percentage could be underestimated since not all the doctors specified the toxic event developed.

Table 36. Characteristics of patients referred for pharmacogenetic diagnostic service implemented at CRO-Aviano.

	N (n=393)	(%)
Genetic Diagnostic Inquiry		
Up-front	248	89.86
Toxicity	23	8.33
Hematologic	9	
Neutropenia	8	
Not hematologic	16	
Diarrhea	14	
Mucositis	3	
Other	5	1.81
First FP Administration		
Yes	223	88.49
No	29	11.51
First Irinotecan Administration		
Yes	19	70.37
No	8	29.63

Table 37. Genotype and Clinical characteristics for patients referred for toxicity in the pharmacogenetic diagnostic service implemented at CRO-Aviano. Patients presenting either *DPYD* or/and *UGT1A1* aforementioned variants are evidenced in gray.

Patient's characteristics			***DPYD* genotype**			***UGT1A1***	**Therapy**		**First administration**		**Toxicity**		
n	**Gender**	**Age**	**IVS14 + 1G>A**	**c.2846A>T**	**1679T>G**	***28**	**Regimen**	**Setting**	**FP**	**Irinotecan**	**Hematologic**	**Non-Hematologic**	**Other**
1	F	42	GG				FP+ Pt-CT +/- other					Diarrhea G3	
2	F	51	GA	AA	TT		FP +/- LV	Other	Yes		Leukopenia G4, Neutropenia G4, Thrombocytopenia G2	Diarrhea G2	
3	M	68	GG	AA	TT	*1/*28	FP+ Pt-CT +/- other	First line	Yes			Diarrhea	
4	M	74				*1/*28	Irinotecan	Other		No	Leukopenia G4, Neutropenia G4	Diarrhea G3	
5	M	61	GG	AA	TT	*1/*1	FP + irinotecan +/- other	Neoadjuvant	Yes	Yes		Diarrhea G3, Mucositis G3	Asthenia G3
6	F	66	GG	AA	TT			Other	No		Neutropenia G3	Gastroenteric (diarrhea) G3	
7	M	64	GG	AA	TT	*1/*1	FP + irinotecan +/- other	First line			Neutropenia Gx	Diarrhea G3	
8	M	62	GG	AA	TT	*1/*28		First line	No	Yes		Diarrhea G3	
9	F	59	GG	AA	TT			First line	Yes				
10	F	57	GG	AA	TT		FP+ Pt-CT +/- other	Adjuvant	No		Thrombocytopenia	Mucositis	
11	M	57	GG	AA	TT			First line	Yes				
12	M	47	GG	AA	TT			First line	Yes				

13	M	66	GG	AA	TT			First line	Yes			
14	F	68	GG	AA	TT		FP +/- LV	Adjuvant	Yes			Cardiotoxicity
15	M	59	GG	AA	TT			First line	Yes			
16	M	71	GG	AA	TT		FP +/- LV	Other	No	Neutropenic infection		Sepsi G4, Alopecia
17	F	58	GG	AT	TT		FP +/- LV	First line	No	Leukopenia, Neutropenia	Diarrhea	
18	F	70	GG	AA	TT		FP+ Pt-CT +/- other	Adjuvant	Yes		Diarrhea G4	Nephrotoxicity
19	M	85	GG	AA	TT		FP +/- LV	First line	No	Leukopenia G4, Neutropenia G4, Thrombocytopenia G4	Mucositis G4	
20	F	61	GG	AA	TT		RT/FP+ Pt-CT +/- other	Neoadjuvant	No		Gastroenteric (diarrhea) G1-G3	
21	F	53	GG	AA	TT		FP + irinotecan +/- other	First line	Yes	Yes	Leukopenia G4, Neutropenia	Diarrhea G4
22	M	65	GG	AA	TT	*1/*28	FP + irinotecan +/- other	First line	No	No		Diarrhea
23	M	78	GG	AA	TT		FP +/- LV				Diarrhea, Fever, Vomiting	Dermatologic toxicity, Ulcers G4

6. Discussion

Pharmacogenetics represents one of the main strategies for precision medicine in cancer. Although, despite all the advancements made in the mCRC pharmacological treatment, combinatorial chemotherapy regimens still represent the backbone of each therapeutic scheme, and still burden the patients with unpredictable side effects. The significant inter-subjects variability in the neutropenia and GI toxicity profile represents a critical aspect of the management of mCRC patients treated with irinotecan-containing regimens (FOLFIRI) [37,55,56,125]. In the last years, several pharmacogenetic studies have been performed in order to identify genetic markers that could help clinicians in personalizing irinotecan-based therapy. Despite these efforts, the only validated predictor marker of severe toxicity remains up to date the *UGT1A1*28* polymorphism adopted by international guidelines for irinotecan dose adjustment [102,126]. Additional work is still required to define further molecular markers to better identify patients at risk for severe complications.

My thesis objective was to develop pharmacogenomic markers of the irinotecan and/or fluoropyrimidines-associated toxicity by exploring different stages of the PGx markers: from the exploratory setting to the clinical implementation setting. Indeed, my thesis explored three different aspects of Pharmacogenetic. The first one consisted in the discovery of novel genetic predictors of neutropenia and gastrointestinal (GI) toxicity risk following a FOLFIRI-based treatment in mCRC patients. For this purpose, a haplotype-tagging polymorphism strategy (htSNPs) and an independent replication analysis have been used. The second aim of this thesis was to analyze the clinical utility of profiling *UGT1A1*28,* which is an acknowledged PGx biomarker in irinotecan-treated mCRC patients in the clinical practice. Specifically, evidence of *UGT1A1*28* up-front genetic screening clinical efficacy was provided by measuring its relationship with the financial costs associated with irinotecan-related toxicity at the National Cancer Center CRO of Aviano. Eventually, this thesis describes the process setting-up of a PGx implementation infrastructure for the prevention of irinotecan and/or fluoropyrimidines-associated toxicity in the clinical routine of the National Cancer Center CRO of Aviano.

Exploratory analysis of new markers

The main result obtained in the exploration setting of my thesis is the identification of STAT-3-rs1053004 polymorphism as a crucial predictor of severe irinotecan-induced GI side effects over the entire course of chemotherapy; this association resulted validated in two independent FOLFIRI-treated mCRC patient cohorts comprising over 400 patients. A growing body of evidence has demonstrated that the cancer-related inflammation could have crucial therapeutic implications [62,63]. Inflammation response has been shown to impact the pharmacokinetic and pharmacodynamic parameters of various chemotherapeutics possibly through the

modulation of drug metabolic enzymes and ABC/SLC transporters [64,65,127]. This effect has been recognized to be mediated by transcriptional regulators, such as STAT3 whose activity is controlled by pro-inflammatory cytokine-induced signaling pathways [64,75,127].

In particular, the polymorphic STAT-3 rs1053004-C allele resulted in a protective factor toward the likelihood of experiencing severe GI side effects in both discovery and replication cohorts. This polymorphism, located in the 3'-UTR region of the gene in a putative miRNA-binding site for miR-423-5p, was reported to regulate the expression of STAT3- protein, by probably altering mRNA degradation, and to significantly affect the STAT-3 transcriptional activity [128]. Once activated, STAT-3 transduces signal across the cytoplasm and into the nucleus mediating gene transcription often in cooperation with NF-κB [129–132]. STAT-3 is triggered in response to the binding of numerous pro-inflammatory cytokines and represents the critical factor in interleukin-6 (IL-6) induced gene regulation [130,132]. Interestingly, IL-6/STAT3 cascade has been reported to be indirectly implicated in the regulation of crucial irinotecan-pathway related proteins, such as cytochromes (i.e. CYP3A4) [133] and ABC/SLC transporters (i.e. MDR-1) [66,134], that have a crucial role in drug disposition. An altered STAT-3 activity associated to the polymorphic variant could thus potentially impact the multifactorial mechanism involved in the determination of drug bioavailability and consequently the individual predisposition to experience severe side effects. In addition to the effect on the irinotecan metabolic proteins regulation, STAT-3, in view of its role as a mediator of cytokine signaling, could also be directly involved in the complex mechanism underlining the toxic damage on the GI epithelium (i.e. mucositis) induced by some chemotherapeutics including irinotecan. Accumulating published data have indicated that the pathobiology of mucositis drug-related involves the mucosal immune system with an important role played by pro-inflammatory cytokine release [135]. This mechanism was reported to be specifically implicated also in the onset of intestinal injury subsequent to the administration of irinotecan.

IL-6/STAT3 signaling could be thus supposed to contribute to the entity of mucosal damage FOLFIRI-related by controlling the proliferation and survival of intestinal epithelial cells as that observed in another inflammatory context like colitis [136,137]. Furthermore, it is interesting that an interaction between IL-6/STAT3 cascade and the gut microbiome, that was known to contribute to the local accumulation of the cytotoxic active metabolite SN-38 [41], was also described [138]. Hence, a change in the STAT-3 activity genetically determined (i.e. rs1053004), could significantly alter the mediation of the pro-inflammatory cytokines signaling modifying the susceptibility of the GI mucosa to the toxic effect of irinotecan-based treatment in agreement with the findings of the present work. Unfortunately, the missing genotype data for the validation set, did not permit to elucidate the potential predictive role on GI toxicity of other pro-inflammatory cytokines markers analyzed in the present study.

This study pointed out also VDR rs11574077 polymorphism as a noteworthy predictor of GI toxicity over the entire course of chemotherapy. Even if its effect was not replicated at a significant level in the replication set of this study, a concordant detrimental effect on the risk to develop severe toxicity was demonstrated, and this effect was supported by a consistent effect on irinotecan pharmacokinetics. VDR rs11574077 polymorphism, associated in this study with an increased likelihood to experience GI side effects, was significantly correlated with an increased BI and inferior GR, parameters indicating an inferior efficacy of the glucuronidation and detoxification pathway. VDR, beyond its physiological role in calcium and phosphate homeostasis, was shown to cooperate in the transcriptional regulation of drug transporters and metabolic enzymes [139], thus possibly affecting the drug disposition profile. rs11574077 represents an intronic variant that, although was reported to impact the risk of developing some tumors and cardiovascular diseases, was of unknown functional significance [140,141]. However, a potentially altered VDR activity determined by the polymorphism could be supposed to influence the drug bioavailability with significant consequences in the irinotecan-based GI toxicity modulation. It is also important to note that since the pharmacokinetic analysis was performed in a subgroup of discovery cohort patients, a specific population/ethnic effect could not be excluded.

The limitations of this study need to be considered. First, FOLFIRI regimen is now commonly given to mCRC patients in combination with monoclonal antibodies. A subset of the patients in the validation set was however treated with a combination scheme of FOLFIRI and bevacizumab supporting a predictive effect of the markers also in this poly-chemotherapy treated patients. A second limit is that the functional meaning of the markers highlighted in the present study is unknown. Formal functional analyses should be performed in order to understand better the molecular mechanism underlying the observed associations. However, the concordant pharmacokinetic effect found for some of the highlighted markers support a functional effect of the variant in affecting the efficiency of the encoded protein enzymatic activity. The last limitation must be pointed out related to the inclusion in the study of common genetic variants with a MAF>0.5%. As recently pointed out rarer genetic variants could account for a high percentage of inter-individual variability in drug metabolism, including NRs genes [142] as well as for the observed inter-individual heterogeneity in the drug toxicity and pharmacokinetics. Future pharmacogenetic approaches should, therefore, include also these emerging markers in order to better describe patient's phenotype regarding response to pharmacological treatment.

The lack of homogeneous studies with different clinical and genetic end-points renders it difficult to understand the real value of the single analyzed variants. The advancement of next generation sequencing technologies points out another emerging problem in pharmacogenetics, still not taken into account, that is the role of rare genetic variants. A recent revision of the 1K genome project results highlighted that about 30 to 40% of the

functional variability in ADME genes and NRs is imputable to rare variants not commonly included in the genetic screening panels [142] and can account for the observed variability among individuals in the drug toxicity and pharmacokinetics. The scale, affordability, efficiency and sensitivity of NGS that can now be achieved is providing unprecedented progress in oncology. It is likely that for the future new approaches based on next generation sequencing, also for the study of NRs pharmacogenomics, should be undertaken in order to better describe patient's phenotype regarding response to pharmacological treatment. Including NGS as a comprehensive approach to completely screen all pharmacogenes will be extremely useful to identify novel putatively pathogenic pharmacovariants,

The cost analysis

Although pharmacogenetics has been extensively applied for defining the best therapy for each patient, the economic impact of its implementation in the clinical practice needs to be better investigated. In this second part of my thesis, I have defined the cost of toxicity management in patients based on the *UGT1A1*28* genotype. A final consensus on the clinical utility of *UGT1A1*28* test implementation in the oncological clinical practice has not been reached yet, despite several publications and meta-analyses [143–146] clearly supporting its clinical validity. A differential cost in the toxicity management of patients based on the *UGT1A1*28* genotype could also be applied to the reimbursement strategies that the national regulatory agencies consider in the DRG-based payment system.

The aim consisted in addressing, for the first time, the economic impact of the management of chemotherapy-related toxicities in a large group of mCRC patients treated with the standard of care therapy irinotecan-based (FOLFIRI regimen), based upon the Italian Public Health System. The main finding was that the toxicity-related cost per patient significantly increased according to the **28* genotype, in an allele-dependent fashion. The cost of all toxicities management per patient increased (1.4 folds) for **1/*28* carriers in comparison to **1/*1*, with the biggest effect observed for the **28/*28* patients, for whom there was a 6-fold increase.

This finding is consistent with the association herein demonstrated between **28* genotype and occurrence of grade 4 toxicity. It is noteworthy that grade 4 toxicity occurred earlier (during the first part of the treatment course), in **28* carriers vs. non-carriers. The consequence of this finding is that, probably because of a longer period of supportive therapies, the cost of toxicity management is higher in a **28* carrier than a non-carrier. Hospitalization is the most expensive medical intervention for toxicity management in this patient population and is also related to the **28* genotype, with **28* carriers experiencing more frequent and longer periods of hospitalization than **28* non-carriers. This effect is unlikely to be biased by the knowledge of a patient genotype,

as the information on the test results and its association with toxicity has been analyzed after the completion of the study.

Every year, about 16,000 patients in Italy undergo treatment with an irinotecan-containing regimen based on the epidemiological data on the incidence of mCRC in Italy in 2012 [147]. The results of the current analysis can be used to provide an estimate of the impact of the incremental cost of toxicity management associated to *UGT1A1*28* at a national level, based on the prevalence of the *UGT1A1*28* allele in populations residing in Italy [124,148]. Based upon the results of our study, an approximate calculation of the total incremental cost spent by the Italian Public Health System to manage the chemotherapy-related toxicity for CRC patients carrying at least one **28* allele and being treated with irinotecan would be around 8 million €, every year. The cost for genotyping all the 16,000 irinotecan-candidate patients, based upon a genotyping cost of 163€ per patient (according to the Regional Health System, Table 4), would be around 2.6 million €. The cost of toxicity management exceeds, by far, the costs that requested to genotype all these patients before an irinotecan treatment. Considering that a pre-emptive genotyping approach in CRC patients candidate to an irinotecan treatment could lead to alternative therapeutic modifications in patients with the **28* allele, limiting the risk and/or the severity of toxicity, we conclude that pre-emptive genotyping could be a cost-effective strategy. However, because of the assumptions made, the clinical and economic impact of pre-emptive genotyping should be prospectively verified.

This estimate is based on the assumption that all the patients would be treated within the Public Health System, as usually, this is the case in Italy for the oncological setting. This estimate might go up significantly in the case of treatment in private hospital settings, where the cost of hospitalization might be significantly higher than in the public hospital setting.

It could be assumed that similar findings could be obtained in other countries within the European Community, where a national health system supported by the government is predominant. A retrospective survey of the cost of chemotherapy-related toxicity management in a large population of cancer patients in Germany and a direct comparison with the same costs in Italy reported similar results between the two countries, in terms of cost of toxicity management [149,150]. In a group of patients treated in Sweden, costs for medical intervention in the case of drug-related toxicity were similar to those reported in the present study, supporting the potential transferability of the results.[151] It must also be considered that the genotyping cost in Italy is, at the time of writing, higher than that reported by other Health Care sources as US Medicare laboratory fee schedule [116] that reports a cost of 103$ (equal to 92€ at the time of writing). It is also likely that the genotyping costs will drop in the near future, as single variants are likely to be included in pre-emptive genotyping panels or obtained by DNA sequencing, both approaches becoming less expensive and more widespread [152].

Strategies for the use of the results of the *UGT1A1*28* testing have been described elsewhere. In **28/*28* patients, they vary from dose reduction of irinotecan at first cycle with adjustment of dosing based upon toxicity to the exclusion of patients from treatment with irinotecan and administration of other equally effective regimens [115–118]. From an economic point of view, these data suggest focusing not only on the **28/*28* patients but also on the **1/*28* patients, due to the observed increase in toxicity-management costs also in this class of patients (Table 27).

Among the study limitations, there is the investigation of a retrospective group of patients treated with cytotoxic chemotherapy, before the introduction of targeted agents and biologic such as anti-EGFR monoclonal antibodies (i.e. cetuximab, panitumumab), and anti-angiogenetic drugs (i.e. bevacizumab, aflibercept, and regorafenib). However, FOLFIRI is still a highly used, standard backbone chemotherapy in patients with mCRC, and *UGT1A1*28* has been demonstrated to impact the toxicity risk also in patients treated with the combination regimens [153–155].

Another limitation is that not all the toxic events can be ascribed to irinotecan, as some of its toxicities overlap with those typical of this drug, such as neutropenia and diarrhea. To overcome, at least in part, this limitation, patients heterozygous for one of the three *DPYD* variants included in the Clinical Pharmacogenetics Implementation Consortium and the Dutch Pharmacogenetics Working Group of the Royal Dutch Association for the Advancement of Pharmacy guidelines [102,126], therefore at increased risk of developing fluoropyrimidine-related toxicity, were excluded from the analysis. A further limitation is that the competing effect of disease progression, potentially interfering with the occurrence of toxicity, was not considered in the present analysis and it should be addressed by using competing risk computational modeling. It should be considered that previous meta-analyses excluded an interaction between **28* genotype and patient time to progression [156,157]. Plus, the quality of life was not assessed in the current cost-analysis.

It has been claimed that, following the introduction of DRG-based payment, the quality of care reduced in many European countries because this system was not completely able to reflect differences in the complexity of treating different patients [158–160]. Particularly, high-cost outlier cases, that have a strong influence on the average costs of cases within a DRG are not represented [161]. Almost all European countries attempted to refine the DRG-based payment system to ensure adequate hospital reimbursement with an expansion over time in the number of groups. Some other countries developed mechanisms to identify outlier cases and to pay hospitals separately for the extra costs of treating such patients otherwise hospitals would experience incentives to avoid these high-cost cases (dumping phenomenon). In this, Pharmacogenetics could play a key role in intercepting high-cost outlier cases embodied by different genotypes, well acknowledged to influence the severe and costly toxicity occurrence. Plus, considering that genotype data will become widely accessible the introduction of pharmacogenetics profiling

in the DRG-based reimbursement system could deliver an efficient care that could be ultimately functional to save costs.

Implementation of pharmacogenetic diagnostic in the clinic

5-Fluorouracil (5-FU) and capecitabine (CAP) are among the most frequently prescribed anticancer drugs. They are inactivated in the liver by the enzyme DPD. Up to 5% of the population is DPD deficient, and these patients have a significantly increased risk of severe and potentially lethal toxicity when treated with regular doses of 5-FU or CAP.

Certainly, of the thousands of genetic markers of toxicity reported in the literature, only a few thrive to be introduced in the everyday clinical practice [108,162]. Although prospective *DPYD* and *UGT1A1*28* genotyping is a valuable tool to identify patients with DPD or UGT1A1 deficiency, and thus those at risk for severe and potential life-threatening toxicity, prospective genotyping has not yet been implemented in daily clinical care. Among the factors preventing the clinical implementation of the pharmacogenetic markers in the clinical practice, including *UGT1A1*28* and *DPYD*, is the lack of their clinical utility assessment and consequently formal health technology assessment studies, including cost-effectiveness and cost-consequences evaluation and also the low awareness of the pharmacogenetic issues among clinical practitioners [108]. Prospective randomized clinical trials comparing PGx-guided therapy to the normal clinical practice would solve a lot of issues regarding the clinical utility of the test. Even though performing prospective randomized clinical trials comparing the PGx-guided therapy to the normal clinical practice would face ethical issues since in the control arm it would be administered a standard dosage to a patient known to carry a risk variant for severe toxicity.

Although the clinical validity of *DPYD* testing has been largely demonstrated, a specific limiting factor is the low test sensitivity, relative to low *DPYD* variants frequency [95]. Contrariwise, *UGT1A1*28* test presents relatively high sensitivity prediction values due to its prevalence in oncologic population (compared to other pharmacogenetics tests, i.e. *DPYD*). Immediate benefit in patient care can be expected through decreasing toxicity, while maintaining efficacy as well 5-FU, the other drug included in the FOLFIRI regimen [114]. Besides, a lot of economic studies claimed the pharmacogenetic diagnostic cost-effectiveness [89,111] because of the rapidity and the reduced invasiveness in regard to phenotyping tests and the substantial screening's affordability and the avoiding of costly toxicities. Moreover, once genotypes have been obtained, the results are valid for lifetime.

The main result obtained in the implementation setting of my thesis was the successful routine application of a prospective *UGT1A1*28* and *DPYD* screening program followed by pharmacogenetically guided dose recommendations. Although not available, it would be useful to know the percentage of patients of all eligible

(newly prescribed 5FU or CAP or irinotecan) patients in which genetic screening was performed. In the study period between 2011 and 2016, 89.86% of the patients were not screened prior to receiving 5FU or CAP therapy. Since the pharmacogenetic diagnostic service systematically started by January 2014, 331 patients were genotyped, which on average comes down to ten patients per month.

Due to the retrospective design of data collection, available data may not always have been fully complete. Considering that the clinical data were collected in a real world clinical setting, from a survey submitted to the oncologist. From these data, we were not able to determine the level of routine application of *DPYD* and *UGT1A1*28* screening in daily practice. An increase in the level of routine application of these tests would indicate that prescribers eventually underwent a learning or acceptance curve following the initial start, and were getting used to applying *DPYD* genotyping increasingly in their daily routine. Therefore, we can conclude that prospective *DPYD* and *UGT1A1*28* screening was successfully implemented in the reality of CRO-Aviano hospital. Although, a revision of the clinical records of these patients would be useful to better understand the oncologists' sensitization to the utility of such pharmacogenetics tests. In order to support the clinical implementation, clinical acceptance should increase, and the use of a clinical decision support system might be suitable. In LUMC Netherland Hospital Pharmacy a clinical decision support system entitled adverse drug event alerting system (ADEAS) is used in daily practice in the hospital pharmacy of LUMC [163]. This system could be useful also in Italian Hospital Pharmacies to systematically select patients at risk of possible adverse drug events. It retrieves data from several information systems and uses clinical rules to select the patient at risk of adverse drug events.

Even if all patients with a *DPYD* and *UGT1A1*28* risk allele are identified and treated with an appropriately reduced dose, not all fluoropyrimidine and irinotecan-related toxicities can be prevented. Adding a DPD/UGT phenotyping test may increase sensitivity, but is expensive and logistically challenging to implement in clinical practice [164]. Even though these pharmacogenetic diagnostic cannot prevent all fluoropyrimidine and irinotecan-related toxicity, we feel that the available evidence strongly supports implementation in clinical practice and can prevent fluoropyrimidine- induced deaths [89,114].

The presence of one of the four *DPYD* variants that were pre-emptively tested resulted in a recommendation to the oncologist to reduce the initial dose of 5FU or CAP by 25–50% depending on the known variant. In February 2015 the recommended dose decrease for c.2846A>T was changed from 50 to 25%, following the updated guidelines of the DPWG [102].

The increase in oncologists' and patients' awareness is required to successfully succeed in the integration of pharmacogenetic diagnostic test in the clinical routine.

7. Conclusion

To summarize, by a discovery-validation approach it was demonstrated for the first time that STAT-3 rs1053004 could give a crucial contribution in the definition of the GI toxicity risk after irinotecan therapy (FOLFIRI regimen) in mCRC patients. This finding further highlights the importance of inflammatory response mediators in the pathobiology of the drug-induced mucosal injury, a topic that needs more attention by pharmacogenetic investigations. VDR rs11574077 marker further emerged as a novel promising determinant of GI toxicity risk FOFLFIRI-related soliciting future research efforts in this direction. The discovery of novel markers predicting the individual predisposition to develop severe GI toxicity after chemotherapy is of great interest considering the increasing recognized clinical and economic implications of the mucosal damage induced by treatment [165]. Altogether, the findings of this work could represent a further step towards personalized FOLFIRI treatment.

Secondly, evidence was provided for differential costs of FOLFIRI-related toxicity management according to the *UGT1A1*28* allele of patients and its potential nationwide impact. Such data provide an innovative perspective on the implementation of pre-emptive *UGT1A1*28* testing in the clinic, with new insights to include genotype information in the DRG-based reimbursement system to support a more comprehensive stratification approach. Since the clinical validity of the *UGT1A1*28* test is widely acknowledged in the scientific community and our data support its cost-effectiveness, considerations based on the results of the test should be made to improve precision in the reimbursement models related to toxicity management.

Lastly, this thesis project shows that a systematic *DPYD* and *UGT1A1*28* prospective screening can be successfully implemented in daily real-world clinical practice. Although a more active follow-up of adherence and toxicity development would provide more data to improve patient safety in irinotecan and fluoropyrimidine administration even further.

In conclusion, this thesis presents a substantial novel research in three different settings: the exploration, the cost-analysis, and the implementation setting. Interesting and innovative molecular markers of toxicity irinotecan-related with a predictive value on pharmacokinetics and pharmacodynamics of irinotecan were discovered. In addition to the prevention of severe toxicity, up-front *UGT1A1*28* genotyping should be considered as an important tool to save economic resources related to the management of irinotecan-related toxicities and for innovative reimbursement strategies. A possible application of these parameters in the clinical practice will be useful to design a tailored irinotecan dosing based on peculiar characteristics of each patient. Moreover, the pre-emptive pharmacogenetics tests can be successfully translated in the clinical Italian routine of the CRO-Aviano,

although further evidence is required to better understand the size of oncologists' and patients' sensitization to the clinical utility of pharmacogenetics tests introduction in the clinic. UPGx project will provide the final proof of pharmacogenetics efficacy in increasing drug safety when fully integrated in the clinical practice of many European countries.

Bibliography

1. Lazarou, J., Pomeranz, B. H. & Corey, P. N. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA* **279**, 1200–1205 (1998).
2. Kalia, M. Biomarkers for personalized oncology: recent advances and future challenges. *Metabolism.* **64**, S16-21 (2015).
3. Kalia, M. Personalized oncology: recent advances and future challenges. *Metabolism.* **62 Suppl 1**, S11-14 (2013).
4. Cancer Facts & Figures 2016 | American Cancer Society. at <http://www.cancer.org/research/cancerfactsstatistics/cancerfactsfigures2016/>
5. Spear, B. B., Heath-Chiozzi, M. & Huff, J. Clinical application of pharmacogenetics. *Trends Mol. Med.* **7**, 201–204 (2001).
6. Efferth, T. & Volm, M. Pharmacogenetics for individualized cancer chemotherapy. *Pharmacol. Ther.* **107**, 155–176 (2005).
7. Marsh, S., Van Booven, D. J. & McLeod, H. L. Global pharmacogenetics: giving the genome to the masses. *Pharmacogenomics* **7**, 625–631 (2006).
8. Muller, P. Y. & Milton, M. N. The determination and interpretation of the therapeutic index in drug development. *Nat. Rev. Drug Discov.* **11**, 751–761 (2012).
9. Lunenburg, C. A., Staveren, M. C. van, Gelderblom, H., Guchelaar, H.-J. & Swen, J. J. Evaluation of clinical implementation of prospective DPYD genotyping in 5-fluorouracil- or capecitabine-treated patients. *Pharmacogenomics* **17**, 721–729 (2016).
10. Ratain, M. J. Irinotecan dosing: does the CPT in CPT-11 stand for 'Can't Predict Toxicity'? *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **20**, 7–8 (2002).

11. Ando, M. *et al.* Genetic polymorphisms of thiopurine S-methyltransferase and 6-mercaptopurine toxicity in Japanese children with acute lymphoblastic leukaemia: *Pharmacogenetics* **11**, 269–273 (2001).
12. Kuilenburg, A. B. van *et al.* Clinical implications of dihydropyrimidine dehydrogenase (DPD) deficiency in patients with severe 5-fluorouracil-associated toxicity: identification of new mutations in the DPD gene. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **6**, 4705–4712 (2000).
13. Iacopetta, B., Grieu, F., Joseph, D. & Elsaleh, H. A polymorphism in the enhancer region of the thymidylate synthase promoter influences the survival of colorectal cancer patients treated with 5-fluorouracil. *Br. J. Cancer* **85**, 827–831 (2001).
14. Stoehlmacher, J. *et al.* Association between glutathione S-transferase P1, T1, and M1 genetic polymorphism and survival of patients with metastatic colorectal cancer. *J. Natl. Cancer Inst.* **94**, 936–942 (2002).
15. Butcher, N.J., Boukouvala, S., Sim, E. & Minchin, R.F. Pharmacogenetics of the arylamine N-acetyltransferases. *Pharmacogenomics J.* (1905).
16. Iyer, L. *et al.* Biliary transport of irinotecan and metabolites in normal and P-glycoprotein-deficient mice. *Cancer Chemother. Pharmacol.* **49**, 336–341 (2002).
17. Iyer, L. *et al.* UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J.* **2**, 43–47 (2002).
18. Luo, F. R., Paranjpe, P. V., Guo, A., Rubin, E. & Sinko, P. Intestinal transport of irinotecan in Caco-2 cells and MDCK II cells overexpressing efflux transporters Pgp, cMOAT, and MRP1. *Drug Metab. Dispos. Biol. Fate Chem.* **30**, 763–770 (2002).
19. Xie, H.-G., Wood, A. J. J., Kim, R. B., Stein, C. M. & Wilkinson, G. R. Genetic variability in CYP3A5 and its possible consequences. *Pharmacogenomics* **5**, 243–272 (2004).
20. Innocenti, F. & Ratain, M. Irinotecan treatment in cancer patients with UGT1A1 polymorphisms. *Oncol. Williston Park N* **17**, 52–55 (2003).

21. Sachse, C., Brockmöller, J., Bauer, S. & Roots, I. Cytochrome P450 2D6 variants in a Caucasian population: allele frequencies and phenotypic consequences. *Am. J. Hum. Genet.* **60**, 284–295 (1997).
22. Fujita, K., Kubota, Y., Ishida, H. & Sasaki, Y. Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J. Gastroenterol.* **21**, 12234–12248 (2015).
23. Kunimoto, T. *et al.* Antitumor activity of 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxy-camptothec in, a novel water-soluble derivative of camptothecin, against murine tumors. *Cancer Res.* **47**, 5944–5947 (1987).
24. Sawada, S. *et al.* Synthesis and antitumor activity of 20(S)-camptothecin derivatives: carbamate-linked, water-soluble derivatives of 7-ethyl-10-hydroxycamptothecin. *Chem. Pharm. Bull. (Tokyo)* **39**, 1446–1450 (1991).
25. Slatter, J. G. *et al.* Pharmacokinetics, metabolism, and excretion of irinotecan (CPT-11) following I.V. infusion of [(14)C]CPT-11 in cancer patients. *Drug Metab. Dispos. Biol. Fate Chem.* **28**, 423–433 (2000).
26. Gerrits, C. J., Jonge, M. J. de, Schellens, J. H., Stoter, G. & Verweij, J. Topoisomerase I inhibitors: the relevance of prolonged exposure for present clinical development. *Br. J. Cancer* **76**, 952–962 (1997).
27. Douillard, J. Y. *et al.* Irinotecan combined with fluorouracil compared with fluorouracil alone as first-line treatment for metastatic colorectal cancer: a multicentre randomised trial. *Lancet Lond. Engl.* **355**, 1041–1047 (2000).
28. Saltz, L. B. *et al.* Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. Irinotecan Study Group. *N. Engl. J. Med.* **343**, 905–914 (2000).
29. Wall, M. E., Wani, M. C., Natschke, S. M. & Nicholas, A. W. Plant antitumor agents. 22. Isolation of 11-hydroxycamptothecin from Camptotheca acuminata Decne: total synthesis and biological activity. *J. Med. Chem.* **29**, 1553–1555 (1986).
30. Hsiang, Y. H. & Liu, L. F. Identification of mammalian DNA topoisomerase I as an intracellular target of the anticancer drug camptothecin. *Cancer Res.* **48**, 1722–1726 (1988).
31. Hsiang, Y. H., Hertzberg, R., Hecht, S. & Liu, L. F. Camptothecin induces protein-linked DNA breaks via mammalian DNA topoisomerase I. *J. Biol. Chem.* **260**, 14873–14878 (1985).

32. Wang, J. C. DNA topoisomerases. *Annu. Rev. Biochem.* **65**, 635–692 (1996).

33. Stewart, L., Redinbo, M. R., Qiu, X., Hol, W. G. & Champoux, J. J. A model for the mechanism of human topoisomerase I. *Science* **279**, 1534–1541 (1998).

34. Hertzberg, R. P., Caranfa, M. J. & Hecht, S. M. On the mechanism of topoisomerase I inhibition by camptothecin: evidence for binding to an enzyme-DNA complex. *Biochemistry (Mosc.)* **28**, 4629–4638 (1989).

35. Hsiang, Y. H., Lihou, M. G. & Liu, L. F. Arrest of replication forks by drug-stabilized topoisomerase I-DNA cleavable complexes as a mechanism of cell killing by camptothecin. *Cancer Res.* **49**, 5077–5082 (1989).

36. Rothenberg, M. L. Topoisomerase I inhibitors: review and update. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **8**, 837–855 (1997).

37. Kudoh, S. *et al.* Phase II study of irinotecan combined with cisplatin in patients with previously untreated small-cell lung cancer. West Japan Lung Cancer Group. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **16**, 1068–1074 (1998).

38. Innocenti, F. & Ratain, M. J. Update on pharmacogenetics in cancer chemotherapy. *Eur. J. Cancer Oxf. Engl. 1990* **38**, 639–644 (2002).

39. Pitot, H. C. *et al.* Phase II trial of irinotecan in patients with metastatic colorectal carcinoma. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **15**, 2910–2919 (1997).

40. Rothenberg, M. L. *et al.* Superiority of oxaliplatin and fluorouracil-leucovorin compared with either therapy alone in patients with progressive colorectal cancer after irinotecan and fluorouracil-leucovorin: interim results of a phase III trial. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **21**, 2059–2069 (2003).

41. Mathijssen, R. H. *et al.* Clinical pharmacokinetics and metabolism of irinotecan (CPT-11). *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **7**, 2182–2194 (2001).

42. Toffoli, G., Cecchin, E., Corona, G. & Boiocchi, M. Pharmacogenetics of Irinotecan. *Curr. Med. Chem. - Anti-Cancer Agents* **3**, 225–237 (2003).

43. Cascorbi, I. *et al.* Frequency of single nucleotide polymorphisms in the P-glycoprotein drug transporter MDR1 gene in white subjects. *Clin. Pharmacol. Ther.* **69**, 169–174 (2001).
44. Kaneda, N., Nagata, H., Furuta, T. & Yokokura, T. Metabolism and pharmacokinetics of the camptothecin analogue CPT-11 in the mouse. *Cancer Res.* **50**, 1715–1720 (1990).
45. Stoops, J. K. *et al.* Carboxylesterases (EC 3.1.1). Kinetic studies on carboxylesterases. (2002).at <http://pubs.acs.org/doi/abs/10.1021/bi00833a037>
46. Khanna, R., Morton, C. L., Danks, M. K. & Potter, P. M. Proficient metabolism of irinotecan by a human intestinal carboxylesterase. *Cancer Res.* **60**, 4725–4728 (2000).
47. Kuehl, P. *et al.* Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat. Genet.* **27**, 383–391 (2001).
48. Huang, C. *et al.* A novel compound heterozygous variation of the uridine-diphosphoglucuronosyl transferase 1A1 gene that causes Crigler-Najjar syndrome type II. *Pharmacogenetics* **11**, 639–642 (2001).
49. De Mattia, E. *et al.* Pharmacogenetics of ABC and SLC transporters in metastatic colorectal cancer patients receiving first-line FOLFIRI treatment. *Pharmacogenet. Genomics* **23**, 549–557 (2013).
50. Toffoli, G. *et al.* The role of UGT1A1*28 polymorphism in the pharmacodynamics and pharmacokinetics of irinotecan in patients with metastatic colorectal cancer. *J. Clin. Oncol.* **24**, 3061–3068 (2006).
51. Toffoli, G. *et al.* Genotype-driven phase I study of irinotecan administered in combination with fluorouracil/leucovorin in patients with metastatic colorectal cancer. *J. Clin. Oncol.* **28**, 866–871 (2010).
52. Cecchin, E. *et al.* Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan. *J. Clin. Oncol.* **27**, 2457–2465 (2009).
53. Ando, Y. *et al.* Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis. *Cancer Res.* **60**, 6921–6926 (2000).

54. Kawato, Y., Aonuma, M., Hirota, Y., Kuga, H. & Sato, K. Intracellular roles of SN-38, a metabolite of the camptothecin derivative CPT-11, in the antitumor effect of CPT-11. *Cancer Res.* **51**, 4187–4191 (1991).

55. Gupta, E. *et al.* Metabolic fate of irinotecan in humans: correlation of glucuronidation with diarrhea. *Cancer Res.* **54**, 3723–3725 (1994).

56. Gupta, E., Wang, X., Ramirez, J. & Ratain, M. J. Modulation of glucuronidation of SN-38, the active metabolite of irinotecan, by valproic acid and phenobarbital. *Cancer Chemother. Pharmacol.* **39**, 440–444 (1997).

57. Table of Pharmacogenomic Biomarkers in Drug Labeling (FDA). at <http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm>

58. Innocenti, F. *et al.* Genetic variants in the UDP-glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. *J. Clin. Oncol.* **22**, 1382–1388 (2004).

59. Marcuello, E. *et al.* A genotype-directed phase I-IV dose-finding study of irinotecan in combination with fluorouracil/leucovorin as first-line treatment in advanced colorectal cancer. *Br. J. Cancer* **105**, 53–57 (2011).

60. Ciotti, M., Basu, N., Brangi, M. & Owens, I. S. Glucuronidation of 7-ethyl-10-hydroxycamptothecin (SN-38) by the human UDP-glucuronosyltransferases encoded at the UGT1 locus. *Biochem. Biophys. Res. Commun.* **260**, 199–202 (1999).

61. Innocenti, F. *et al.* Haplotype structure of the UDP-glucuronosyltransferase 1A1 promoter in different ethnic groups. *Pharmacogenetics* **12**, 725–733 (2002).

62. Janakiram, N. B. & Rao, C. V. The role of inflammation in colon cancer. *Adv. Exp. Med. Biol.* **816**, 25–52 (2014).

63. Markman, J. L. & Shiao, S. L. Impact of the immune system and immunotherapy in colorectal cancer. *J. Gastrointest. Oncol.* **6**, 208–223 (2015).

64. Cressman, A. M., Petrovic, V. & Piquette-Miller, M. Inflammation-mediated changes in drug transporter expression/activity: implications for therapeutic drug response. *Expert Rev. Clin. Pharmacol.* **5**, 69–89 (2012).

65. Petrovic, V., Teng, S. & Piquette-Miller, M. Regulation of drug transporters during infection and inflammation. *Mol. Interv.* **7**, 99–111 (2007).

66. Ho, E. A. & Piquette-Miller, M. Regulation of multidrug resistance by pro-inflammatory cytokines. *Curr. Cancer Drug Targets* **6**, 295–311 (2006).

67. Reuter, S. *et al.* Thiocolchicoside exhibits anticancer effects through downregulation of NF-κB pathway and its regulated gene products linked to inflammation and cancer. *Cancer Prev. Res. Phila. Pa* **3**, 1462–1472 (2010).

68. Cecchin, E., De Mattia, E. & Toffoli, G. Nuclear receptors and drug metabolism for the personalization of cancer therapy. *Expert Opin. Drug Metab. Toxicol.* **12**, 291–306 (2016).

69. De Mattia, E., Cecchin, E., Roncato, R. & Toffoli, G. Pregnane X receptor, constitutive androstane receptor and hepatocyte nuclear factors as emerging players in cancer precision medicine. *Pharmacogenomics* **17**, 1547–1571 (2016).

70. De Mattia, E., Dreussi, E., Cecchin, E. & Toffoli, G. Pharmacogenetics of the nuclear hormone receptors: the missing link between environment and drug effects? *Pharmacogenomics* **14**, 2035–2054 (2013).

71. Chai, X., Zeng, S. & Xie, W. Nuclear receptors PXR and CAR: implications for drug metabolism regulation, pharmacogenomics and beyond. *Expert Opin. Drug Metab. Toxicol.* **9**, 253–266 (2013).

72. Mukherjee, S. & Mani, S. Orphan Nuclear Receptors as Targets for Drug Development. *Pharm. Res.* **27**, 1439–1468 (2010).

73. Hwang-Verslues, W. W. & Sladek, F. M. HNF4α -- role in drug metabolism and potential drug target? *Curr. Opin. Pharmacol.* **10**, 698–705 (2010).

74. Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).

75. Chen, Y. *et al.* Nuclear receptors in the multidrug resistance through the regulation of drug-metabolizing enzymes and drug transporters. *Biochem. Pharmacol.* **83**, 1112–1126 (2012).

76. Kobayashi, K., Hashimoto, M., Honkakoski, P. & Negishi, M. Regulation of gene expression by CAR: an update. *Arch. Toxicol.* **89**, 1045–1055 (2015).

77. Wallace, B. D. & Redinbo, M. R. Xenobiotic-sensing nuclear receptors involved in drug metabolism: a structural perspective. *Drug Metab. Rev.* **45**, 79–100 (2013).

78. S. Azmi, A., W. Bao, G., Gao, J., M. Mohammad, R. & H. Sarkar, F. Network Insights into the Genes Regulated by Hepatocyte Nuclear Factor 4 in Response to Drug Induced Perturbations: A Review. *Curr. Drug Discov. Technol.* **10**, 147–154 (2013).

79. Pondugula, S. R. & Mani, S. Pregnane xenobiotic receptor in cancer pathogenesis and therapeutic response. *Cancer Lett.* **328**, 1–9 (2013).

80. Gonzalez, F. J. Regulation of hepatocyte nuclear factor 4 alpha-mediated transcription. *Drug Metab. Pharmacokinet.* **23**, 2–7 (2008).

81. Basseville, A. *et al.* Irinotecan induces steroid and xenobiotic receptor (SXR) signaling to detoxification pathway in colon cancer cells. *Mol. Cancer* **10**, 80 (2011).

82. Mbatchi, L. C. *et al.* Effect of Single Nucleotide Polymorphisms in the Xenobiotic-sensing Receptors NR1I2 and NR1I3 on the Pharmacokinetics and Toxicity of Irinotecan in Colorectal Cancer Patients. *Clin. Pharmacokinet.* **55**, 1145–1157 (2016).

83. Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330–338 (2003).

84. Wohlhueter, R. M., McIvor, R. S. & Plagemann, P. G. Facilitated transport of uracil and 5-fluorouracil, and permeation of orotic acid into cultured mammalian cells. *J. Cell. Physiol.* **104**, 309–319 (1980).

85. Diasio, R. B. & Harris, B. E. Clinical pharmacology of 5-fluorouracil. *Clin. Pharmacokinet.* **16**, 215–237 (1989).

86. Rosmarin, D. *et al.* A candidate gene study of capecitabine-related toxicity in colorectal cancer identifies new toxicity variants at DPYD and a putative role for ENOSF1 rather than TYMS. *Gut* **64**, 111–120 (2015).

87. Heggie, G. D., Sommadossi, J. P., Cross, D. S., Huster, W. J. & Diasio, R. B. Clinical pharmacokinetics of 5-fluorouracil and its metabolites in plasma, urine, and bile. *Cancer Res.* **47**, 2203–2206 (1987).

88. Terrazzino, S. *et al.* DPYD IVS14+1G>A and 2846A>T genotyping for the prediction of severe fluoropyrimidine-related toxicity: a meta-analysis. *Pharmacogenomics* **14**, 1255–1272 (2013).

89. Deenen, M. J. *et al.* Upfront Genotyping of DPYD*2A to Individualize Fluoropyrimidine Therapy: A Safety and Cost Analysis. *J. Clin. Oncol.* **34**, 227–234 (2016).

90. Henricks, L. M. *et al.* Translating DPYD genotype into DPD phenotype: using the DPYD gene activity score. *Pharmacogenomics* **16**, 1277–1286 (2015).

91. Offer, S. M. *et al.* Comparative functional analysis of DPYD variants of potential clinical relevance to dihydropyrimidine dehydrogenase activity. *Cancer Res.* **74**, 2545–2554 (2014).

92. Deenen, M. J. *et al.* Relationship between single nucleotide polymorphisms and haplotypes in DPYD and toxicity and efficacy of capecitabine in advanced colorectal cancer. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **17**, 3455–3468 (2011).

93. Froehlich, T. K., Amstutz, U., Aebi, S., Joerger, M. & Largiadèr, C. R. Clinical importance of risk variants in the dihydropyrimidine dehydrogenase gene for the prediction of early-onset fluoropyrimidine toxicity. *Int. J. Cancer* **136**, 730–739 (2015).

94. Johnson, M. R., Wang, K. & Diasio, R. B. Profound dihydropyrimidine dehydrogenase deficiency resulting from a novel compound heterozygote genotype. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **8**, 768–774 (2002).

95. Lee, A. M. *et al.* DPYD variants as predictors of 5-fluorouracil toxicity in adjuvant colon cancer treatment. *J. Natl. Cancer Inst.* **106**, (2014).

96. Loganayagam, A. *et al.* Pharmacogenetic variants in the DPYD, TYMS, CDA and MTHFR genes are clinically significant predictors of fluoropyrimidine toxicity. *Br. J. Cancer* **108**, 2505–2515 (2013).

97. Collie-Duguid, E. S., Etienne, M. C., Milano, G. & McLeod, H. L. Known variant DPYD alleles do not explain DPD deficiency in cancer patients. *Pharmacogenetics* **10**, 217–223 (2000).

98. Kuilenburg, A. B. P. van *et al.* Novel disease-causing mutations in the dihydropyrimidine dehydrogenase gene interpreted by analysis of the three-dimensional protein structure. *Biochem. J.* **364**, 157–163 (2002).

99. Offer, S. M., Wegner, N. J., Fossum, C., Wang, K. & Diasio, R. B. Phenotypic profiling of DPYD variations relevant to 5-fluorouracil sensitivity using real-time cellular analysis and in vitro measurement of enzyme activity. *Cancer Res.* **73**, 1958–1968 (2013).

100. Morel, A. *et al.* Clinical relevance of different dihydropyrimidine dehydrogenase gene single nucleotide polymorphisms on 5-fluorouracil tolerance. *Mol. Cancer Ther.* **5**, 2895–2904 (2006).

101. FDA Camptostare package insert PGx warning. at <http://www.fda.gov/ohrms/dockets/AC/04/briefing/2004-4079B1_03_Topic1-TabA.pdf>

102. Swen, J. J. *et al.* Pharmacogenetics: from bench to byte--an update of guidelines. *Clin. Pharmacol. Ther.* **89**, 662–673 (2011).

103. Etienne-Grimaldi, M.-C. *et al.* UGT1A1 genotype and irinotecan therapy: general review and implementation in routine practice. *Fundam. Clin. Pharmacol.* **29**, 219–237 (2015).

104. Caudle, K. *et al.* Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process. *Curr. Drug Metab.* **15**, 209–217 (2014).

105. Meulendijks, D. *et al.* Clinical relevance of DPYD variants c.1679T>G, c.1236G>A/HapB3, and c.1601G>A as predictors of severe fluoropyrimidine-associated toxicity: a systematic review and meta-analysis of individual patient data. *Lancet Oncol.* **16**, 1639–1650 (2015).

106. Evans, W. E. & Relling, M. V. Pharmacogenomics: Translating Functional Genomics into Rational Therapeutics. *Science* **286**, 487–491 (1999).

107. Goldstein, D. A., Shaib, W. L. & Flowers, C. R. Costs and effectiveness of genomic testing in the management of colorectal cancer. *Oncol. Williston Park N* **29**, 175–183 (2015).

108. Gillis, N. K. & Innocenti, F. Evidence required to demonstrate clinical utility of pharmacogenetic testing: the debate continues. *Clin. Pharmacol. Ther.* **96**, 655–657 (2014).

109. Janssens, A. C. J. W. & Deverka, P. A. Useless until proven effective: the clinical utility of preemptive pharmacogenetic testing. *Clin. Pharmacol. Ther.* **96**, 652–654 (2014).

110. Beaulieu, M., Denus, S. de & Lachaine, J. Systematic review of pharmacoeconomic studies of pharmacogenomic tests. *Pharmacogenomics* **11**, 1573–1590 (2010).

111. Berm, E. J. J. *et al.* Economic Evaluations of Pharmacogenetic and Pharmacogenomic Screening Tests: A Systematic Review. Second Update of the Literature. *PLoS ONE* **11**, (2016).

112. List Dosing Guidelines Pharmgkb. at <https://www.pharmgkb.org/view/dosing-guidelines.do?source=CPIC>

113. Busse, R. *et al.* Diagnosis related groups in Europe: moving towards transparency, efficiency, and quality in hospitals? *BMJ* **346**, f3197 (2013).

114. Lunenburg, C. A. T. C. *et al.* Prospective DPYD genotyping to reduce the risk of fluoropyrimidine-induced severe toxicity: Ready for prime time. *Eur. J. Cancer Oxf. Engl. 1990* **54**, 40–48 (2016).

115. Pichereau, S. *et al.* Cost-effectiveness of UGT1A1*28 genotyping in preventing severe neutropenia following FOLFIRI therapy in colorectal cancer. *J. Pharm. Pharm. Sci.* **13**, 615–625 (2010).

116. Gold, H. T., Hall, M. J., Blinder, V. & Schackman, B. R. Cost effectiveness of pharmacogenetic testing for uridine diphosphate glucuronosyltransferase 1A1 before irinotecan administration for metastatic colorectal cancer. *Cancer* **115**, 3858–3867 (2009).

117. Obradovic, M., Mrhar, A. & Kos, M. Cost-effectiveness of UGT1A1 genotyping in second-line, high-dose, once every 3 weeks irinotecan monotherapy treatment of colorectal cancer. *Pharmacogenomics* **9**, 539–549 (2008).

118. Butzke, B. *et al.* The cost-effectiveness of UGT1A1 genotyping before colorectal cancer treatment with irinotecan from the perspective of the German statutory health insurance. *Acta Oncol.* 1–11 (2015).doi:10.3109/0284186X.2015.1053983

119. Chen, S. *et al.* ABCC5 and ABCG1 polymorphisms predict irinotecan-induced severe toxicity in metastatic colorectal cancer patients. *Pharmacogenet. Genomics* **25**, 573–583 (2015).

120. Chen, S. *et al.* A novel UGT1 marker associated with better tolerance against irinotecan-induced severe neutropenia in metastatic colorectal cancer patients. *Pharmacogenomics J.* **15**, 513–520 (2015).

121. Lévesque, E. *et al.* Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. *J. Pharmacol. Exp. Ther.* **345**, 95–101 (2013).

122. Fakhrai-Rad, H., Pourmand, N. & Ronaghi, M. Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. *Hum. Mutat.* **19**, 479–485 (2002).

123. Kuilenburg, A. B. van *et al.* Lethal outcome of a patient with a complete dihydropyrimidine dehydrogenase (DPD) deficiency after administration of 5-fluorouracil: frequency of the common IVS14+1G>A mutation causing DPD deficiency. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **7**, 1149–1153 (2001).

124. Ruzzo, A. *et al.* Pharmacogenetic profiling in patients with advanced colorectal cancer treated with first-line FOLFIRI chemotherapy. *Pharmacogenomics J.* **8**, 278–288 (2008).

125. Rothenberg, M. L. *et al.* Phase I dose-finding and pharmacokinetic trial of irinotecan (CPT-11) administered every two weeks. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO* **12**, 1631–1641 (2001).

126. Caudle, K. E. *et al.* Clinical Pharmacogenetics Implementation Consortium guidelines for dihydropyrimidine dehydrogenase genotype and fluoropyrimidine dosing. *Clin. Pharmacol. Ther.* **94**, 640–645 (2013).

127. Harvey, R. D. & Morgan, E. T. Cancer, inflammation, and therapy: effects on cytochrome p450-mediated drug metabolism and implications for novel immunotherapeutic agents. *Clin. Pharmacol. Ther.* **96**, 449–457 (2014).

128. Permuth-Wey, J. *et al.* STAT3 polymorphisms may predict an unfavorable response to first-line platinum-based therapy for women with advanced serous epithelial ovarian cancer. *Int. J. Cancer* **138**, 612–619 (2016).

129. Qi, J. *et al.* β1,6 GlcNAc branches-modified PTPRT attenuates its activity and promotes cell migration by STAT3 pathway. *PloS One* **9**, e98052 (2014).

130. Aggarwal, B. B. *et al.* Signal transducer and activator of transcription-3, inflammation, and cancer: how intimate is the relationship? *Ann. N. Y. Acad. Sci.* **1171**, 59–76 (2009).

131. Grivennikov, S. I. & Karin, M. Dangerous liaisons: STAT3 and NF-kappaB collaboration and crosstalk in cancer. *Cytokine Growth Factor Rev.* **21**, 11–19 (2010).

132. Yu, Y., Zhao, Q., Wang, Z. & Liu, X.-Y. Activated STAT3 correlates with prognosis of non-small cell lung cancer and indicates new anticancer strategies. *Cancer Chemother. Pharmacol.* **75**, 917–922 (2015).

133. Jover, R., Bort, R., Gómez-Lechón, M. J. & Castell, J. V. Down-regulation of human CYP3A4 by the inflammatory signal interleukin-6: molecular mechanism and transcription factors involved. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **16**, 1799–1801 (2002).

134. Zhu, H. *et al.* The synergistic effects of low-dose irinotecan and TRAIL on TRAIL-resistant HT-29 colon carcinoma in vitro and in vivo. *Int. J. Mol. Med.* **30**, 1087–1094 (2012).

135. Sonis, S. T. A biological approach to mucositis. *J. Support. Oncol.* **2**, 21-32-36 (2004).

136. Moriasi, C., Subramaniam, D., Awasthi, S., Ramalingam, S. & Anant, S. Prevention of colitis-associated cancer: natural compounds that target the IL-6 soluble receptor. *Anticancer Agents Med. Chem.* **12**, 1221–1238 (2012).

137. Nguyen, N., Furuta, G. T. & Menard-Katcher, C. Recognition and Assessment of Eosinophilic Esophagitis: The Development of New Clinical Outcome Metrics. *Gastroenterol. Hepatol.* **11**, 670–674 (2015).

138. Hainzl, E. *et al.* Intestinal Epithelial Cell Tyrosine Kinase 2 Transduces IL-22 Signals To Protect from Acute Colitis. *J. Immunol. Author Choice* **195**, 5011–5024 (2015).

139. Prakash, C. *et al.* Nuclear Receptors in Drug Metabolism, Drug Response and Drug Interactions. *Nucl. Recept. Res.* **2**, (2015).

140. Muindi, J. R. *et al.* Serum Vitamin D Metabolites in Colorectal Cancer Patients Receiving Cholecalciferol Supplementation: Correlation with Polymorphisms in the Vitamin D Genes. *Horm. Cancer* **4**, 242–250 (2013).

141. Verschuren, J. J. W. *et al.* Systematic testing of literature reported genetic variation associated with coronary restenosis: results of the GENDER Study. *PloS One* **7**, e42401 (2012).

142. Kozyra, M., Ingelman-Sundberg, M. & Lauschke, V. M. Rare genetic variants in cellular transporters, metabolic enzymes, and nuclear receptors can be important determinants of interindividual differences in drug response. *Genet. Med. Off. J. Am. Coll. Med. Genet.* (2016).doi:10.1038/gim.2016.33

143. Hu, Z.-Y., Yu, Q., Pei, Q. & Guo, C. Dose-dependent association between UGT1A1*28 genotype and irinotecan-induced neutropenia: low doses also increase risk. *Clin. Cancer Res.* **16**, 3832–3842 (2010).

144. Hu, Z.-Y., Yu, Q. & Zhao, Y.-S. Dose-dependent association between UGT1A1*28 polymorphism and irinotecan-induced diarrhoea: a meta-analysis. *Eur. J. Cancer* **46**, 1856–1865 (2010).

145. Liu, X., Cheng, D., Kuang, Q., Liu, G. & Xu, W. Association of UGT1A1*28 polymorphisms with irinotecan-induced toxicities in colorectal cancer: a meta-analysis in Caucasians. *Pharmacogenomics J.* **14**, 120–129 (2014).

146. Campbell, J. M. *et al.* Irinotecan-induced toxicity pharmacogenetics: an umbrella review of systematic reviews and meta-analyses. *Pharmacogenomics J.* (2016).doi:10.1038/tpj.2016.58

147. Globocan, 2012. at <http://globocan.iarc.fr/Default.aspx>

148. Falvella, F. S. *et al.* DPD and UGT1A1 deficiency in colorectal cancer patients receiving triplet chemotherapy with fluoropyrimidines, oxaliplatin and irinotecan. *Br. J. Clin. Pharmacol.* **80**, 581–588 (2015).

149. Herold, M. & Hieke, K. Costs of toxicity during chemotherapy with CHOP, COP/CVP, and fludarabine. *Eur. J. Health Econ.* **3**, 166–172 (2002).

150. Magdelijns, F. J. H., Stassen, P. M., Stehouwer, C. D. A. & Pijpers, E. Direct health care costs of hospital admissions due to adverse events in The Netherlands. *Eur. J. Public Health* **24**, 1028–1033 (2014).

151. Gyllensten, H. *et al.* Economic impact of adverse drug events--a retrospective population-based cohort study of 4970 adults. *PloS One* **9**, e92061 (2014).

152. Ratain, M. J. & Johnson, J. A. Meaningful use of pharmacogenetics. *Clin. Pharmacol. Ther.* **96**, 650–652 (2014).

153. Lu, C.-Y. *et al.* Clinical Implication of UGT1A1 Promoter Polymorphism for Irinotecan Dose Escalation in Metastatic Colorectal Cancer Patients Treated with Bevacizumab Combined with FOLFIRI in the First-line Setting. *Transl. Oncol.* **8**, 474–479 (2015).

154. Yeh, Y.-S. *et al.* Prospective analysis of UGT1A1 promoter polymorphism for irinotecan dose escalation in metastatic colorectal cancer patients treated with bevacizumab plus FOLFIRI as the first-line setting: study protocol for a randomized controlled trial. *Trials* **17**, 46 (2016).

155. Toffoli, G. *et al.* Genotype-guided dosing study of FOLFIRI plus bevacizumab in metastatic colorectal cancer patients. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* (2016).doi:10.1158/1078-0432.CCR-16-1012

156. Liu, X., Cheng, D., Kuang, Q., Liu, G. & Xu, W. Association between UGT1A1*28 polymorphisms and clinical outcomes of irinotecan-based chemotherapies in colorectal cancer: a meta-analysis in Caucasians. *PloS One* **8**, e58489 (2013).

157. Dias, M. M., McKinnon, R. A. & Sorich, M. J. Impact of the UGT1A1*28 allele on response to irinotecan: a systematic review and meta-analysis. *Pharmacogenomics* **13**, 889–899 (2012).

158. Farrar, S. *et al.* Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. *BMJ* **339**, b3047 (2009).

159. Ljunggren, B. & Sjödén, P.-O. Patient-reported quality of life before, compared with after a DRG intervention. *Int. J. Qual. Health Care J. Int. Soc. Qual. Health Care* **15**, 433–440 (2003).

160. Diagnosis Related Groups in Europe - Moving towards transparency, efficiency and quality in hospitals. European Observatory on Health Systems and Policies Series. at <http://www.euro.who.int/__data/assets/pdf_file/0004/162265/e96538.pdf>

161. Cots, F., Elvira, D., Castells, X. & Sáez, M. Relevance of outlier cases in case mix systems and evaluation of trimming methods. *Health Care Manag. Sci.* **6**, 27–35 (2003).

162. Church, D. *et al.* 'Toxgnostics': an unmet need in cancer medicine. *Nat. Rev. Cancer* **14**, 440–445 (2014).

163. Rommers, M. K., Zwaveling, J., Guchelaar, H.-J. & Teepe-Twiss, I. M. Evaluation of rule effectiveness and positive predictive value of clinical rules in a Dutch clinical decision support system in daily hospital pharmacy practice. *Artif. Intell. Med.* **59**, 15–21 (2013).

164. Staveren, M. C. van, Guchelaar, H. J., Kuilenburg, A. B. P. van, Gelderblom, H. & Maring, J. G. Evaluation of predictive tests for screening for dihydropyrimidine dehydrogenase deficiency. *Pharmacogenomics J.* **13**, 389–395 (2013).

165. Gibson, R. J. *et al.* Chemotherapy-induced gut toxicity and pain: involvement of TLRs. *Support. Care Cancer Off. J. Multinatl. Assoc. Support. Care Cancer* **24**, 2251–2258 (2016).

Appendix 1

FORM FOR THE PHARMACOGENETIC ANALYSIS

Origin:

- Oncology unit A ☐
- Oncology unit B ☐
- Oncology unit C ☐
- Radiotherapy ☐
- Other (Specify) ☐ .....

Patient:

Given name .....
Surname .....
Birth date ...../...../.....
Medical record num. .....

Patient's label

Blood draw date: ...../...../.....

Intended therapeutic choice (specify) .....

.....

Intended data for the first cycle of therapy: ...../...../.....

Administration setting:

- Neoadjuvant ☐
- Adjuvant ☐
- First line ☐
- Other .....

Pharmacogenetic test requested:

- DPYD (Fluoropyrimidine) ☐
- UGT1A1*28 (Irinotecan) ☐
- Other (specify) .....

Is it the first time the patient is receiving the following drugs?

	SI	NO	N.A.
Fluoropyrimidine	☐	☐	☐
Irinotecan	☐	☐	☐

Diagnostic inquiry:

- Pre-therapy inquiry ☐
- Toxicity inquiry ☐
  (Specify type and grade)

- Hematological ☐
  - Neutropenia ☐
  - Thrombocytopenia ☐
  - Anemia ☐
- Gastrointestinal ☐
  - Diarrhea ☐
  - Mucositis ☐
- Hand-Foot syndrome ☐
- Cardiac toxicity ☐
- Other (specify) .....

.....

Concomitant therapies (specify)

.....

.....

.....

The prescriber oncologist

.....

Appendix 2

Candidate genes and related TagSNPs/polymorphisms selected for pharmacogenetic analysis.

GENE	TagSNPs (exonic/introic region)	SNPs (5'/3' UTR)
PXR; NR1I2	rs1403527 rs13059232 rs3814055 rs1403526 rs3732357 rs11929668 rs16830505 rs3814057 rs6784598 rs3732359 rs3732360 rs1054190 rs7643645	
FXR; NR1H4	rs35724 rs1030454 rs11110415 rs11110390 rs11610264 rs17030285 rs4764980	
RXR-A; NR2B1	rs4917353 rs3118536 rs7864987 rs10881582 rs1805352 rs877954 rs11103482 rs3132294 rs11185659 rs7038018 rs4240705 rs1045570 rs11103473 rs6537944 rs10776909 rs7039190	
RXR-B; NR2B2	rs2744537 rs2072915 rs2076310	
RXR-G; NR2B3	rs157864 rs157869 rs3767344 rs3767333 rs157880 rs10800098 rs3767339 rs4657437 rs10489747 rs285480 rs285481 rs1123944 rs380518 rs746332 rs2651860 rs10489745 rs100537 rs283690 rs752739 rs283695 rs157862 rs285482 rs283694	
LXR-A; NR1H3	rs7120118	rs11039149 rs10838681
LXR-B; NR1H2	rs1405655	rs4802703
CAR; NR1I3	rs2501873 rs4073054 rs2307424 rs2307418 rs6686001	
VDR; NR1I1	rs2248098 rs11168292 rs10875695 rs757343 rs2239180 rs11574012 rs1544410 rs10783219 rs886441 rs12717991 rs3782905 rs2853564 rs2239182 rs11574046 rs2107301 rs2238136 rs3819545 rs2239186 rs4328262 rs11574026 rs7299460 rs11168287 rs11168275 rs4760648 rs2189480 rs2239179 rs2254210 rs11574077	rs11574143 rs7139166 rs4516035 rs11568820
PPARP-A; NR1C1	rs14842 rs5766743 rs6008259 rs135551 rs4253711 rs135547 rs135538 rs11703495 rs9627046 rs4253701 rs4253623 rs9615264 rs12330015 rs6007662 rs9626736 rs1555208 rs6008197 rs4253662 rs11090819 rs4253655 rs4253755	
PPAR-D; NR1C2	rs9658100 rs2076167 rs2016520	

	rs7744392 rs9470001 rs9658119 rs4713854 rs1053046 rs2076169	
PPAR-G; NR1C3	rs880663 rs1175540 rs1801282 rs2938392 rs4135284 rs17793951 rs17036281 rs13099828 rs7626560 rs4135275 rs2120825 rs2972164 rs3856806 rs4135268 rs4135247 rs1797912	
HNF4A; NR2A1	rs3212198 rs6031551 rs11574738 rs8116574 rs2425637 rs6065725 rs4812829 rs2868094 rs3212208 rs6093978 rs3212183 rs745975 rs717248 rs6031595 rs3092370 rs6031587 rs2425640 rs6093976 rs4364072 rs11574736 rs6103716 rs8114057 rs2071200 rs6031580 rs11574730 rs4812831 rs3818247 rs3212197 rs1800961 rs6017335 rs2071197 rs6073418 rs11574733	rs6130615
HNF1A	rs2244608 rs2393791 rs1169293 rs1169300 rs1169302 rs1169286 rs1169307 rs1882149 rs3999413 rs1169303 rs12427353	
	rs735396 rs2071190	
STAT-3	rs1026916 rs1053005 rs12949918 rs17593222 rs6503695 rs1053004 rs8069645	
NFκB1	rs230539 rs230496 rs4647992 rs4648110 rs11722146 rs4648135 rs3774934 rs4648022 rs1598861 rs4648127 rs4648090	
IKBKB	rs6474388 rs2272733 rs9694958 rs17875671 rs5029748 rs3747811 rs10958713	
CHUK	rs3818411 rs11591741 rs12570957 rs11595324 rs3818411 rs11190430	
IL-1B	rs1143643 rs3136558 rs1143634	rs1143627 rs1143623 rs16944
IL-6	rs2069845 rs2069840	rs2069861 rs2069827 rs1800797 rs1800795
INFG	rs1861494 rs2069716	rs2069727

Abbrevations: 3/5'-UTR; 3'/5' untranslated region; CAR, constitutive androstane receptors (NR1I3); CHUK, conserved helix-loop-helix ubiquitous kinase; FXR, farnesoid X receptor; HNF1A, HNF1 homeobox A; HNF4A; hepatocyte nuclear factors 4 (NR2A1); IL-1B, interleukin-1 beta; IL-6, interleukin-6; INFG, interferon gamma; IKBKB, inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta, LXR-A,B, liver X receptors (NR1H3 and NR1H2); NFκB1, nuclear factor κ B; PPAR-A,D,G, peroxisome proliferator-activated receptors (NR1C1, NR1C2 and NR1C3), PXR, pregnane X receptor; RXR-A,B,C, retinoid X receptors (RXR; NR2B1, NR2B2 and NR2B3); STAT-3, signal transducers and activators of transcription; TagSNPs: tagging polymorphisms; TNF, tumor necrosis factor; VDR, vitamin D receptor (NR1I1).

Appendix 3

List of publications arising from the thesis

Journal

1. "Cost evaluation of irinotecan-related toxicities associated with the UGT1A1*28 patient genotype"

   Roncato R, Cecchin E, Montico M, De Mattia E, Giodini L, Buonadonna A, Solfrini V, Innocenti F, Toffoli G.

   Clin Pharmacol Ther. 2017 Jan; 11

2. "Ubiquitous Pharmacogenomics (U-Pgx): the time for implementation is now. An H2020 program to drive pharmacogenomics into clinical practice"

   Cecchin E, Roncato R, Toffoli G, Guchelaar HJ

   Curr Pharm Biotechnol. 2017 Jan; 2

3. "Pregnane X receptor, constitutive androstane receptor and hepatocyte nuclear factors as emerging players in cancer precision medicine"

   De Mattia E, Cecchin E, Roncato R, Toffoli G.

   Pharmacogenomics. 2016 Sep; 17

4. "Candidate germline polymorphisms of genes belonging to the pathways of four drugs used in osteosarcoma standard chemotherapy associated with risk, survival and toxicity in non-metastatic high-grade osteosarcoma"

   Hattinger CM, Biason P, Iacoboni E, Gagno S, Fanelli M, Tavanti E, Vella S, Ferrari S, Roli A, Roncato R, Giodini L, Scotlandi K, Picci P, Toffoli G, Serra M.

   Oncotarget. 2016 Aug; 22.

5. "Clinical validity of a DPYD-based pharmacogenetic test to predict severe toxicity to fluoropyrimidines"

   Toffoli G, Giodini L, Buonadonna A, Berretta M, De Paoli A, Scalone S, Miolo G, Mini E, Nobili S, Lonardi S, Pella N, Lo Re G, Montico M, Roncato R, Dreussi E, Gagno S, Cecchin E.

   Int J Cancer. 2015 Dec; 15

Oral communications at conferences

1. 28 national congress of the Italian Society of Chemotherapy (SIC). Firenze, 26-28 November 2015

   "Economic burden associated with irinotecan-related toxicities according to UGT1A1*28"

2. 37th national congress of the Italian Society of Pharmacology (SIF), awarded with the 2016 best poster award

   "Economic burden associated with irinotecan-related toxicities according to UGT1A1*28"