Planning to be Healthy: Towards Personalized Medication Planning

Lee-or Alon*, **Hana Weitman, Alexander Shleyfman and Gal A. Kaminka**

Computer Science Department, Bar-Ilan University, Israel

Abstract. Personalized medication plans determine the selection, dosage, and administration schedule of medications, to achieve medical goals that are *specific to the patient and to its individual health constraints*. This paper introduces medication planning as a novel domain for artificial intelligence planning, using PDDL+. We evaluate the suggested representation via experiments based on data collected from medical studies conducted on mice and rats.

1 Introduction

Personalized medication planning is the process of generating a plan of drug administrations that meets a given set of medical goals that are specific to the individual patient. The planning process must take into account general health safety constraints, helpful or harmful interactions between drugs, and individual physiological differences in responses to medications. The resulting personalized medication plan defines *what drugs* are administered, *when*, and at *what dosage*: too little is ineffective; too much is toxic.

The behaviour of medicine administrations in the patient's body can be estimated by combining pharmacokinetic and pharmacodynamic models. *Pharmacokinetic* models [17, 21] describe the time-changing biodistribution (concentration) of the medicine in the patient's body. *Pharmacodynamic* models [15, 45] describe the effect of the drug on various biochemical properties in the body. In other words, pharmacokinetic models assess *how much* of the drug is present at a certain time point in across the body, while pharmacodynamic models describe *what* the medicine does to the body.

Medication planning is a complex process, manually carried out by healthcare professionals. Its complexity is often encountered in mitigating harmful drug interactions in patients with multiple diseases [12], or in *combination therapy*, where multiple medications are used to synergistically improve therapeutic effects while minimizing side effects [39, 41]. Indeed, a combination of drugs can result in effects no drug can achieve alone [42].

Alaboud and Coles [3] introduced a restricted case of medication planning, where the goal is to maintain a level of a single medication in the body of a patient. Their work uses PDDL+ [16] to model the non-linear effects of the drug by assuming it follows an exponential decay curve, parameterized by the drug half-life (a common pharmacokinetic model in medicine). Recently, Alon et al. [4] described a more general case, whereby the planning process considers multiple drugs, arbitrary non-linear effects, and the interacting bio-chemical properties of drugs and the body; these are considered with respect to patient safety and the achievement of medical goals. However, they did not report on a representation approach, nor on any planners capable of carrying out such planning.

We present a general comprehensive approach for general medication planning (GMP), using PDDL+ [16]. In contrast to previous work, the approach accommodates the administration of multiple medications, even repeatedly. In addition, it facilitates a representation of pharmacokinetic-pharmacodynamic models and the consideration of multi-drug effects. It supports the differentiation of medications with respect to various biological sites, which we collectively refer to as *bio-sites*. Consequently, goals can be articulated in terms of whole-body effects or tailored to target a particular target area, using *targeted medications*.

Through experiments conducted on problem sets using real-world data, we examine the challenges presented by the suggested representation. Additionally, we evaluate the interaction between the representation and various search algorithms and heuristics utilized in numerical planning, analyzing their effectiveness and potential limitations.

2 Medication Planning: Background

Our work falls within a general trend of using artificial intelligence (AI) tools to assist and personalize medical care. We focus on *medication planning*, which is concerned with selecting drugs to be administered, as well as determining the dosage and schedule of the chosen medications. Below, we discuss medication planning in detail; the broader context in AI and medicine is discussed separately at the end of the paper.

2.1 Pharmacokinetics and Pharmacodynamics

Once a drug is introduced into the body, it is generally carried by the bloodstream circulating throughout the body. The drug reaches tissues in various *bio-sites*—including organs and tumors—and may accumulate for some time, before it is eventually cleared out of the body. The concentration of a drug in various bio-sites, known as its *biodistribution*, undergoes changes over time, which can be described by *pharmacokinetic* (PK) models of varying complexity. These models range from simple 1-3 compartment exponential decay models [21, 40] to more advanced models that account separately for multiple kinetic processes (see [17]). Alternatively, biodistribution trajectories can also be represented by explicit curves [1, 24], obtained from clinical trials.

For example, in Figure 1, we see the biodistribution trajectories of a specific drug administered to a mouse (nanoparticle #11, in [24]). Drug concentrations (percentage of initial dosage per gram of tissue)

* Corresponding Author. Email: alonlee1@biu.ac.il.

were measured in four bio-sites (*kidney, lung, spleen, liver*), at several time points (measured in hours since the administration at time $t_0 = 0$). Such trajectories change between medicines, but may also change between patients. The horizontal axis shows the time since administration. The vertical axis shows the concentration per gram of tissue as a percentage of the injection dosage. Each line shows the PK trajectory at a different bio-site.

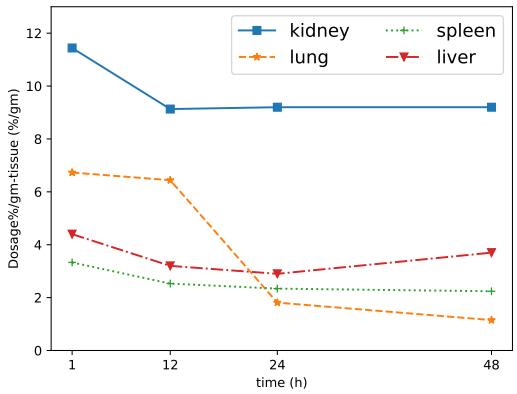

Figure 1: Biodistribution trajectories of drug #11 in mice (from [24]).

When the drug reaches a target bio-site, it may affect the properties of that bio-site. These effects can be characterized using *pharmacodynamic* (PD) models [45]. PD models describe the relationship between the drug concentration at any given bio-site, or the body as a whole (in simple models), and the resulting therapeutic effect.

PK and PD models are combined to form *PKPD* models [45], which predict the expected magnitude of drug therapeutic effect over time. Figure 2 illustrates the connection between PK and PD models. The PK model yields the concentration of the administrated drug in the body at a specific time post-administration. Subsequently, the PD model utilizes this data to calculate the biochemical effect of that drug on the patient's body. PKPD models (and their component models) have been and continue to be an active area of research in medicine and pharmacology, with entire journals devoted to their investigation.

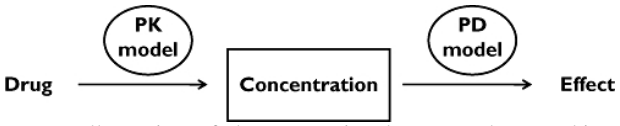

Figure 2: Illustration of the connection between pharmacokinetic models and pharmacodynamic models. Scheme was taken from [45].

2.2 *Medication Planning*

Medication planning involves medical goals that are specified in terms of properties of different bio-sites (or the body taken as a whole), taking into account temporal pharmaceutical dynamics and kinematics. It combines information about the rate of accumulation and clearance of drugs in different bio-sites with information about toxicity and personal health constraints and patient activities to meet target levels of the drug or its biological effects.

The process involves *selecting*, *determining dosage*, and *scheduling* medication administrations to a patient. It is therefore a generalization of dosing regimen planning [38], which assumes drugs have already been selected, and deals with administration dosages and schedules.

A good example of medication planning in mice is presented in a study dealing with a new approach to cancer treatment conducted by von Maltzahn et al. [42]. It consists of two steps, each using a different drug type. In the first step, a targeted drug that attaches to tumors and induces coagulation is administered. Seventy-two hours later, the drug naturally clears from the body, leaving the tumor bio-sites with elevated coagulation. Subsequently, a second drug is administered, which exhibits an affinity to accumulate in bio-sites with elevated levels of coagulation. This approach utilizes two different drug types, which can provide targeted treatment only in combination. The first medicine type can target tumors, but cannot be used for treatment. The second drug cannot target tumors on its own, but can selectively accumulate in the coagulated tumor sites.

Alaboud and Coles [3] and Alaboud [2] introduced an early attempt at automating medication planning, using AI planning to solve problems involving a *single* medication, whose target levels in the patient's body (*single bio-site*) must be maintained in face of the patient's daily activities, influencing the desired drug levels.

Alon et al. [6] suggested several PDDL+ representations for the personalized medication planning problem where multiple medicines may be administered. Additionally, they consider multiple bio-sites in their representations. However, their representations support only PK models.

We present general personalized medication planning as a domain for planning. The approach we take utilizes PKPD models to predict drug concentration and effects that vary across an arbitrary number of bio-sites, over time (Fig. 1). It allows for an arbitrary number of medicines, each may be administered repeatedly if needed. The interactions of the drugs are modeled, so that the planner can avoid harmful interactions, and replace one drug with another (or with a combination of drugs).

3 The Medication Planning Problem

In this section, we present the representation of the general medication planning problem (GMP) in PDDL+. We begin with a brief reminder of PDDL+, before continuing by describing how a patient's medical state (and drug distribution within it) is represented (Sec. 3.1). We then show how pharamacokinetic models and measurements may be represented (Sec. 3.2), and used together with a pharmacodynamic model to represent drug administrations and (personalized) medical goals, under safety constraints (Sec. 3.3).

We follow the common definition of a PDDL+ planning problem [16]. A tuple $\langle \mathcal{V}, \mathcal{S}, s_0, \mathcal{C}, \mathcal{G}, \mathcal{A}, \hat{\mathcal{E}}, \hat{\mathcal{P}} \rangle$ where $\mathcal{V}$ is a set of state variables either propositional or numeric, $\mathcal{S}$ is a set of states, where each state is a complete assignment of values to all variables $v \in \mathcal{V}$, $s_0 \in \mathcal{S}$ is an initial state, $\mathcal{C}$ is a set of constraints on possible assignments of values, and $\mathcal{G}$ is a goal description (a set of conditions over variables). $\mathcal{A}$ is a set of instantaneous actions that change the values of variables when selected by the agent, and $\hat{\mathcal{E}}, \hat{\mathcal{P}}$ sets of events and processes (resp.) that change the values of variables instantaneously or overtime, outside of the control of the agent.

Every action $a \in \mathcal{A}$ has a set of preconditions pre_a and a set of effects eff_a. The precondition pre_a is a set of propositional and arithmetic (in)equalities that are defined over the variables in $\mathcal{V}$.[1] If a state $s \in \mathcal{S}$ satisfies the preconditions of an action $a \in \mathcal{A}$, i.e., $s \models pre_a$, then the action a is *applicable* in state s, and the planner may consider adding it to the plan. When an applicable action a is executed in state s, the state variables are updated according to the effect formulas eff_a, where propositional variables are set to specific values, and numeric variables are assigned values determined by arithmetic

[1] An arithmetic formula is defined as an application of operators $\{+, -, \times, \div\}$ to state variables and constants.

formulas instantiated in the given state. In contrast, if a state s satisfies the precondition of an event in $\hat{\mathcal{E}}$ or a process in $\hat{\mathcal{P}}$, it is always executed; it is not subject to the planners choice. Syntactically, the preconditions and effects of events and processes are similar to the ones of actions, the only difference is the specific variable $\#t$ that represents the individual time of the process, it is used to calculate continuous effects and intermediate values of some numeric fluents.

A valid solution (a plan) is an ordered set of consecutively applicable *actions* that starts at the initial state s_0, transforms fluent values with each ordered action, and reaches a goal state, i.e., a state compatible with $\mathcal{G}$, such that no state in the state trajectory violates any of the constraints in $\mathcal{C}$. Typically, an optimal plan minimizes an objective such as number of actions, or their accumulating costs.

3.1 Biochemical Properties: Variables and States

From a medical perspective, a patient's body can be viewed as a set B of bio-sites, such as organs and blood. Basic pharmacological models often depict the entire body as a single bio-site ($|B| = 1$), but in more complex models multiple bio-sites are represented. The approach we take here follows the more general view.

Each bio-site is represented as a set of P biochemical properties. Their values indicate concentration levels or other measures of interest and generally vary between bio-sites. Here, we represent each property of a given bio-site as a numeric fluent, whose value at any given time is measured in relevant standard units such as nanograms per gram of tissue, for example. These measurements are familiar to anyone who had a blood test.

The variables in $\mathcal{V}$ encompass various elements, including but not limited to the fluents describing all properties across all bio-sites. Therefore, we divide $\mathcal{V}$ based on bio-sites and their respective properties, where $b[j]$ denotes property $j \in P$ in bio-site $b \in B$. Additionally, $\mathcal{V}$ contains auxiliary variables that handle the tracking of different events and processes, e.g., monitoring medicine administration, violated constraints, etc.

A state is a complete assignment of values to all variables in $\mathcal{V}$. Table 1 shows an example of 12 property variables, used in the experiments. The property m_{11} (first row) measures the concentration of drug #11 (Fig. 1) in six different bio-sites, at a specific time (e.g., *liver*$[m_{11}] = 2.97$, and *kidney*$[m_{11}] = 9.2$). A second property, measuring the level of mu-opioid receptor (MOR) activity, is shown in the second row. Its values in this case are derived from PKPD model parameters reported elsewhere [33]. The *initial state* s_0 of a patient's body may be represented by setting the values of properties, in each bio-site, to current values. For properties measuring drug concentration, initial values in all bio-sites are zero.

Properties P \ Organs B	Blood	Heart	Liver	Spleen	Lung	Kidney
m_{11}	1.6	0.73	2.97	2.34	1.81	9.2
MOR activity	26.4	20.28	30.003	28.79	27.27	33.55

Table 1: Illustration of the state of a patient's body 24 hours post-administration. Columns represent bio-sites. Rows represent property values.

3.2 Pharmacokinetics: Actions and Events

When a drug is administered, it directly affects the corresponding properties tracking the levels of the drug itself in different bio-sites. These drug accumulation effects are highly non-linear (as discussed in the previous section). Every administration affects multiple bio-sites, continuously, over time, until the drug clears the body.

We use PDDL+ *action templates* $adm(m, d)$ to represent administering dosage d of a drug m. We record its dosage as $d(m)$. The time within the plan, t, is constantly tracked by a dedicated process. Each administration instance is associated with a fluent that represents the time of administration, $t_0(m)$. This fluent is initialized on the administration action to the current time in the system at the administration moment. Thus, $t - t_0(m)$, the difference between the current time in the system (being tracked by the process) and the time of that instance administration, gives the time that has passed since m was administered. For simplicity, we omit the representation of the time-tracking process in the discussion.

The selection of an administration action triggers a PDDL+ event computing its pharmacokinetic effects over time. For every medicine m, there is *at least* a single property m in every bio-site b, i.e., $b[m] \in \mathcal{V}$, which represents the concentration level of medicine m in bio-site b. These levels can be computed by parametric PK models (see, e.g., [3, 21]), or estimated directly from explicitly-represented trajectories (e.g., Fig. 1). We follow this latter approach as it is more general, and explicitly represents biodistribution trajectories in the problemm description in PDDL+.

We begin with the simpler case, where a drug is administered at most once. In this case, the event computes the grounded direct effects of an action $adm(m, d)$ as a function of the initial dosage $d(m)$ and the time since injection $t - t_0(m)$, which yields a percentage in the associate biodistribution trajectory. We follow common practices and assume that the biodistribution trajectories are given in percentages of the initial dosage, for a standard mass unit.

For every bio-site $b \in B$, the value of property $b[m]$ for any time $t \geq t_0(m)$ is given as:

$$b[m] := g(b, m, t - t_0(m)) \cdot d(m), \tag{1}$$

where $d(m)$ is the dosage selected for the administration of m, and $g(b, m, t - t_0(m))$ is the sampled value of the biodistribution trajectory of medicine type m for bio-site b, relative to the time of administering m. Both t and $t_0(m)$ are given in absolute time (since the beginning of the plan). For drugs that directly affect more than a single property, the event may be used to compute the direct effects in P properties. Different medicines may be administered simultaneously, and their effects will be computed by different events.

Repeating Administrations We can allow for repeating administrations of the same medication type, at different times. This may be useful for achieving medical goals while not violating safety constraints (discussed later).

When more than a single administration of the same drug are present in the patient's body, the change in the medicine level due to the administrations cannot be simply overwritten (as is done in Eq. 1). Let us denote by m_i the i^{th} repeating administration of drug m. Then, the total concentration level for a drug m should be computed by adding the relative contribution of each administered instance m_i. Naïvely, the property value should be computed as:

$$b[m] := \sum_i [g(b, m_i, t - t_0(m_i)) \cdot d(m_i)]. \tag{2}$$

There are a number of difficulties with Eq. 2. First, PDDL+ does not allow the generation of new m_i as needed. Instead, we permit up to N administrations of the same type of medicine $(m_1, \ldots, m_N)$, where N is finite and known in advance (in the problem description). Second, every administration action triggers its own event for computing the effect. This event cannot assign a value to $b[m]$ (using the PDDL+ assignment keyword), as it will overwrite values from other events. To correctly sum the concentration levels, we must differentiate between events triggered by distinct administrations of the same

medicine, and sum their contributions. We expand the representation to address these issues.

First, the administration action of instance m_i has two additional pre-conditions: (i) instance m_i has not been administered yet, and (ii) the state has not administered another instance of the same medicine type in the current time step. To prevent symmetries between identical administration instances, we keep a counter for each medicine type. An administration action may consider administering instance m_i only if all instances m_j where $j < i$ have already been administered. After administrating a medicine instance, the action increases the counter value of this medicine to allow the administration of the next instance, if any.

To keep track of which medicine type was administered in a time step, the state holds a predicate for each medicine type. At the initial state, all these predicates are initialized to *False*. When an instance i of medicine m is administered, the corresponding predicate is marked as *True*. An event resets these predicates to enable the administration of instances of these medicine types in the next time step. For further discussion see supplementary materials.

To correctly sum the concentration levels from repeating instances, we change the computation of $b[m]$ by each event triggered by an action. Below, we show the representation of *one* such events. Each PK event, even for the same medication type, has its own set of separate variables. The only shared variables $b[m]$ are medicine dependent. They represent the total concentration level of m, regardless of the number of repeating administrations.

For each administered instance m_i, the event calculates the *difference* between the current property level (that is, $d(m_i) \cdot g(b, m, t - t(m_i))$) and the previous property level created by this medicine instance, i.e., $\text{prev}(b, m_i)$, and adds it to the current medicine level value. This way, the event considers only the change from the previous effect by this specific instance, without overwriting the PK effect by other instances. That allows multiple medicine instances to affect the level of the medicine in a bio-site simultaneously.

The PDDL+-like pseudo-code of the event is as follows (for readability, we avoid the Lisp-like syntax),

$$b[m] \mathrel{+}= g(b, m, t - t(m_i)) \cdot d(m_i) - \text{prev}(b, m_i) \quad (3)$$

$$\text{prev}(b, m_i) := g(b, m, t - t(m_i)) \cdot d(m_i) \quad (4)$$

Eqs. 3–4, triggered by events for each administered instance m_i, collectively carry out the computation outlined in Eq. 2.

3.3 Pharmacodynamics: Events, Safety, and Goals

As a drug is accumulated in a bio-site (measured by its concentration level), it causes changes in other biochemical properties within the same bio-site. These changes can be predicted using PD models. The combination of the PK and PD model types, known as a PKPD model, allows for the estimation of how the accumulation and clearance of a drug change in biochemical properties influence various bio-sites over time [21, 45, 15].

A common PKPD model in medical literature is the *direct action* (direct effect) model [45, 15]. This model describes the relationship between the time-dependent drug concentration, and its effects, measured in relevant units that vary between drugs. The model relies on two parameters, E_{max} (the maximal effect achieved by the drug), and EC_{50} (the drug concentration level needed to achieve 50% of E_{max}). These parameters, defined by the drug's PD model, remain constant over time.

Following the computation of $b[m]$—the total concentration level of drug m at bio-site b at time t—separate events compute the effect of the drug on the value of the effect property p (at time t), using direct effect formula [45, 15]:

$$b[p] \mathrel{+}= E_{max}(m, b, p) \cdot \frac{b[m]}{b[m] + EC_{50}(m, b, p)}, \quad (5)$$

where $E_{max}(m, b, p)$ and $EC_{50}(m, b, p)$ are given by the PD model for the drug m, when active in bio-site b, affecting property p. As a reminder, $b[m]$ measures the total concentration level of m in bio-site b, even when resulting from repeating administrations of m. Unlike in the PK event, we do not need to model each administration instance of the same medicine type with separate events. That is due to the fact that a PKPD event considers the total level of the medication currently in the bio-site, instead of considering the relative value of each instance by itself.

Note that Eq. 5 recomputes the drug effects for each new t. At the beginning of each time step, events nullify the values of all properties that do not represent medicine levels. This allows the simple summation of the PKPD effect. For simplicity, we omit the representation of this event in the discussion.

Different drugs may affect the same property simultaneously. As these will be handled by different events, their effects will increase the value of the property according to the PKPD effect of the associated medicine type. This naturally follows the Loewe *additive* drug interaction model [10, 45], whereby drugs can affect the same property, but at different "strength". Contra-indicated drugs (may not be taken together) are handled by constraints (see below).

Goals $\mathcal{G}$ and Safety Constraints $\mathcal{C}$ Given the definitions of states and actions above, it seems a simple matter to define goal states in terms of target levels for properties of interest, at a specific set of bio-sites (therapeutic sites). However, medically, we must also ensure that the levels of all properties are maintained at safe levels, *before* the target levels are reached, as well as *after*.

We use events to impose limits on the maximal and/or minimal values of a property at any moment. These limits can come from medical defaults, or they may be personalized for specific health conditions of a patient. For example, if a patient has diabetes, the glucose level must stay below a given threshold h at all times. Such a constraint on the property j of bio-site b can be expressed as $b[j] \bowtie h$, where $\bowtie \in \{>, \geq, =, \leq, <\}$.

Constraints can be placed to prevent interactions between drugs. For example, we may represent a constraint that if a property value i in a bio-site b is greater than a given threshold h_i, the value of property j in the same bio-site must be less than a threshold h_j, i.e., $b[i] > h_i \Rightarrow b[j] < h_j$.

The goal description $\mathcal{G}$ has two components in the PDDL+ representation of GMP. The first involves specifying target levels for properties in the set of therapeutic sites. These target levels can be personalized and differ between patients. The second component ensures that constraints are maintained after these target levels are achieved.

Once the goal conditions are first satisfied at time t_g, safety constraints should be upheld not only in the interval $[0, t_g]$ but also in the extended interval $[t_g, \infty)$, bearing in mind that action effects have finite durations. Thus, a second subgoal introduced using PDDL+ checks that all administered medication had been eliminated from the patient's body *after* the first component has been achieved.

Personalization Patients, who seek treatment for the same goal, vary in their medical history (e.g., background medical conditions leading to differences in safety constraints) and treatment preferences (e.g., due to age, sex, levels of activity). Two patients with the same medical goal may still require different treatment plans due to dif-

ferences in background medical conditions, such as diabetes, pregnancy, etc. Patient diversity may cause differences in their PKPD responses, both in biodistribution trajectories, as well as E_{max} and EC_{50} parameters.

4 Experiments

We empirically evaluate the use of PDDL+ planning for medication planning using medical PK and PD data from mice and rats. Specifically, PK data was taken from an extensive database reporting on the biodistribution trajectories of more than 200 nanoparticle-based drugs in up to 15 different bio-sites in mice [24]. In principle, such nanoparticles can be used to carry many different kinds of therapeutic drugs, and as such the database does not report on PD (therapeutic effects, other than concentration). We therefore used PD parameters (E_{max}, EC_{50}) for various medications in mice and rats [33, 25, 31]. All experiments were run using the ENHSP-20 planner [37], which supports PDDL+, including non-linear conditions and effects.

We evaluated the representation under various planning algorithms and heuristics, on 261 medication planning problems. Each was attempted by up to six different planning algorithms/heuristic combinations (described below), for a total of more than 1500 trials.[2]

Personalized Medication Planning Example To demonstrate the difficulty of solving GMP problems, let us examine a relatively simple instance. Suppose the planner is to consider only two types of medicine reported in [33], *Nalbuphine* and *Buprenorphine*, each allowed to be given at most once. Suppose the clearance time for Nalbuphine is 25 hours, while for Buprenorphine it is 49 hours. The biodistribution trajectories were taken from [24]. Five different dosages are available for each medicine. The medical objective is to have the mu-opioid receptor selectivity reach a level of at least 51 in the spleen, but not exceed 53.8.

While this problem seems small, solving it could be challenging due to the many possible administration times ($25 \cdot 49$) and dosage options ($5 \cdot 5$) resulting in over $30,500$ combinations. However, The ENHSP planner solved this problem in under 3 seconds using A^* and the h^{max} heuristic (see below on the various algorithms and heuristics used in the experiments). It recommended administering Buprenorphine with a dosage of 30, waiting 24 hours, then administering Nalbuphine with the same dosage, and waiting until both medicines clear the body, totaling 49 hours.

We examined the ability to personalize the treatment to a patient's individual responses. Román et al. [31] reports that the same medicine (*Dantrolene*) affects rats with two different medical conditions, in a different manner. The PD parameters for the drug's effects on *5-HT* (a particular property) were $E_{max} = 0.67,\ EC_{50} = 8.43$ for one type of rat. For the other type, the PD parameters were: $E_{max} = 0.82,\ EC_{50} = 20.39$ in the same bio-site. The different behaviour also occurs with a different drug, *Nimodipine*: its PD parameters for 5-HT were $E_{max} = 0.42,\ EC_{50} = 7.49$ for the first type of rat, and $E_{max} = 0.35,\ EC_{50} = 12.4$ for the second.

We gave the planner the task of reaching a 5-HT value above 0.3. All other parameters, including safety constraints, were identical, except for PKPD responses. Using the same settings, ENHSP suggests different plans for the two rat types. It advised Dantrolene for the first type and Nimodipine for the second.

Parameters affecting computation: Baseline experiments We now turn to evaluating the representation proposed in computational terms. Obviously, the search algorithm and heuristics used affect the results. We used the following ENHSP-implemented numeric heuristics: *blind* (1 for non-goal states, 0 for goals), the admissible heuristics h^{max} and h^{rmax}, and inadmissible h^{add} and h^{radd} are all due to [36]. Their treatment of processes and events is based on the inadmissible *AIBR* heuristic that is due to [34]. The search algorithms used are the A^* search [19] and the Greedy Best First Search (*GBFS*) [13]. *Planners*, therefore, are a combination of a search algorithm and a heuristic, where opt-blind is $A^* + blind$, opt-hmax is $A^* + h^{max}$, opt-hrmax is $A^* + h^{rmax}$, sat-aibr is $A^* + AIBR$, sat-hadd is $GBFS + h^{add}$, and sat-hradd is $GBFS + h^{radd}$.

We created 37 problem instances, using the data from [24, 25, 33]. The experiments differ in the number of medicine types M (1,2,4,6,7), potential repetitive administrations of each type N (1–4), the number of bio-sites B (1–3) and properties of interest P (1 or 2). Medications vary in their biodistributions and clearance time, i.e., the duration of the biodiversity trajectory curves (25 or 49 hours); longer clearance times require the planner to maintain its associated processes over more states. Medications also vary in their E_{max} and EC_{50} parameters between bio-site properties. The desired values in the goal may differ between problem instances.

We tested every problem with each of the different planners (algorithm+heuristic combination). Runtime was measured in CPU-seconds, using the Linux time command. If no solution was found after 10 CPU minutes, the planning process was stopped. We limited the memory consumption to 16GB RAM. In both scenarios, the inability of the planner to find a solution was considered a failure. The experiments were performed on a machine with Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20GHz CPU and 258GB RAM. Tests were conducted in parallel using 20 hyperthreaded cores (a total of 40 logical cores). No more than 20 tests were run in parallel.

The initial experiment set served as our baseline. These problems were devoid of any medical safety constraints, rendering them relatively simpler. By utilizing this experiment set, we were not only able to establish a baseline for subsequent experiments but also isolate the individual effects of each parameter on planner runtime.

A comparison of the results from planners in this experiment are summarized in Table 2 in the first row (marked *baseline*). We see that opt-blind was practically unable to solve the medication planning problems used in the evaluation (it solved only three out of 37), though opt-hmax and opt-hrmax were able to solve all and most of the problems, respectively. Perhaps surprisingly, the satisficing planners, *sat-*$*$, performed very badly, in comparison: sat-aibr solved three problems; sat-hadd and sat-hradd solved none.

Fig. 3 shows the effect of various parameters on opt-hmax and opt-hrmax runtimes, in the baseline experiment. The vertical axis shows the runtime in CPU seconds. Each marker in the graph represents a single run. Unsuccessful runs are marked as a runtime of 600 seconds for spacing. Goal difficulty was determined by sampling reachable goal values in a few problem instances and comparing their runtimes.

Computational effects of medical safety constraints A second set of experiments sought to establish the role of the safety constraints on the difficulty of the problems. We created two more sets of 37 problems each, where each problem from the baseline was duplicated and modified, such that it now contained a set of *loose* (set 1), or *tight* constraints (set 2). Loose constraints were picked such that their threshold values were far above the target property values in a valid solution, so a wider range of dosages and schedules could achieve the goal. Tight constraints were close to the target property values, and thus the number of valid solutions is much reduced, i.e., harder to find a plan.

Rows 2 and 3 in Table 2 show the results from this second set of

[2] For all instances, see https://bitbucket.org/lee-ora/ecai24-supplementary/.

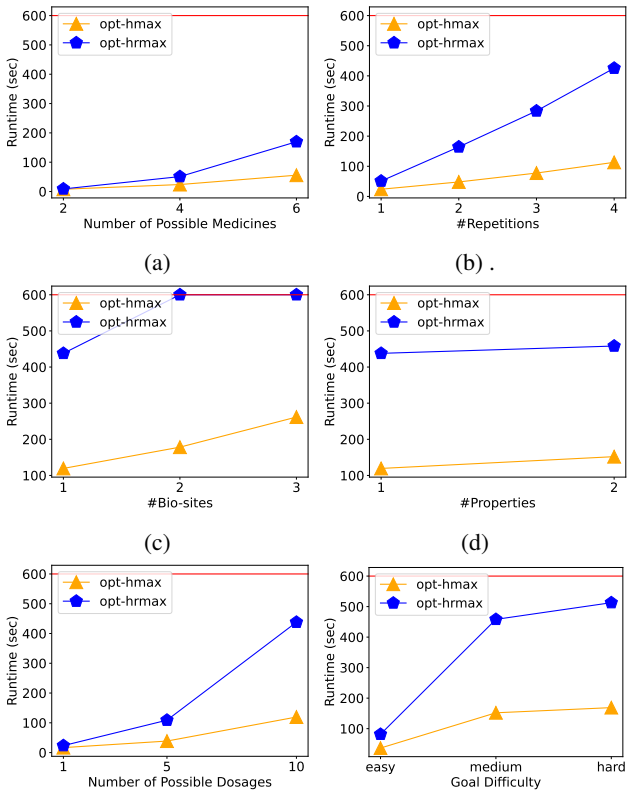

Figure 3: Runtimes of opt-hmax and opt-hrmax in various baseline settings: (a) 2–6 medicine types, (b) 1–4 possible repetitive administrations of each medicine type, (c) 1–3 bio-site to consider in the plan, (d) 1–2 bio-site properties to consider, (e) 1–10 available possible dosages for each medicine type, and (f) the goal difficulty.

experiments. opt-hmax continues to maintain its advantage over opt-hrmax. However, we see that both planners suffer in performance as constraints become tighter.

Computational effects of the PK clearance times A final set of experiments tests the impact of PK duration, i.e., the time from administering a drug to until it is cleared. For these experiments, we had to synthesize PK curves, by shrinking and stretching the real PK data, compared to the original baseline. To shrink a problem (shorten the PK duration), we divide the time of each sample point in the biodistribution by shrinkage factor, and then remove redundant sampling points that have the same time after this division. Similarly, for stretching duration by a stretching factor, we added time points between the original samples and evaluated the values at these new points using linear interpolation. We used factors of 2 and 4 for both shrinking and stretching. By exploring these different experimental settings, we aimed to understand the impact of each such parameter on the coverage of the problem instances and later on their runtimes.

The results are shown in Table 2, in the last four rows. We see that the longer the duration of the PK modeled process (the longer the drug stays in the body), the harder it is for the planner to solve medication planning problems containing the drug in question.

For the two planners with the highest coverage, opt-hmax and opt-hrmax, we also report on the geometric means of runtime in seconds (T), the expanded nodes (N), and the evaluated states (E). Means were calculated on the problems both opt-hmax and opt-hrmax successfully solved in each row.

While the means of the expanded nodes and evaluated states are the same for both planners in each row, the runtimes for opt-hmax are always lower than those of opt-hrmax. This might indicate that the redundant constraints of h^{rmax} did not help solve the GMP problems faster. Instead, the additional computation required for these redundant constraints made opt-hrmax run slower compared to opt-hmax.

	sat-aibr	opt-blind	opt-hmax		opt-hrmax			
Experiment Setting	C	C	C	T	C	T	N	E
Baseline (No Const.)	3	3	**37**	**30.96**	34	65.83	289.1	8212.17
Constraints (Loose)	3	3	**37**	**32.17**	33	67.62	338.49	9303.06
Constraints (Tight)	1	2	**19**	**33.06**	14	70.14	245.56	8810.45
Time Shrinkage (X4)	16	13	**37**	**12.84**	**37**	19.48	117.73	3447.51
Time Shrinkage (X2)	8	3	**37**	**20.64**	**37**	37.37	210.98	6266.07
Time Stretching (X2)	0	3	**34**	**33.2**	20	73.17	357.47	7013.89
Time Stretching (X4)	0	0	**21**	**44.67**	13	127.3	672.95	8971.08

Table 2: Comparison of planners in different experiment settings (rows). Coverage (C) is measured by number of problems solved (out of 37). Planners sat-hadd and sat-hradd did not solve any problem (columns omitted). Columns marked (T), (E) and (N) denote the geometric means of the run-time, expanded nodes and evaluated states, respectively, for opt-hmax and opt-hrmax by row. Means were calculated on the problems both opt-hmax and opt-hrmax successfully solved in each row. Best values in each row are in bold.

Focusing on the leading planners, Fig. 4 contrasts the runtime of opt-hmax and opt-hrmax on all problems from all experiment sets. The two axes measure the runtime for each problem on a logarithmic scale. Each marker represents a single problem. A marker below the main diagonal indicates that opt-hrmax ran faster than opt-hmax on the specific problem. A marker above the main diagonal indicates that opt-hmax had a lower runtime. When a run was unsuccessful for a planner, it is given a runtime of 600 sec., denoted by the red lines. Multiple markers can be hidden by another marker if they share the same runtime. For example, by the marker at the top right corner at (600,600), which indicates all the unsuccessful runs by both planners. Generally, opt-hmax performed better in comparison to opt-hrmax in both terms of coverage and runtime.

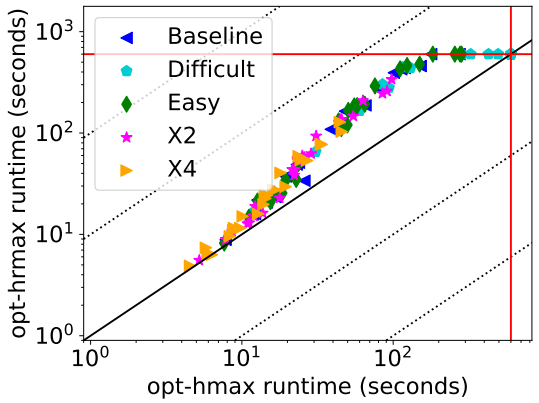

Figure 4: The runtimes of opt-hmax compared to opt-hrmax runtimes, measured on *logarithmic* scale axes. Each marker represents a single problem instance. Unsuccessful runs are marked as a runtime of 600 seconds for spacing. All problems where both opt-hmax and opt-hrmax failed are located at the upper right corner of the graph.

5 Discussion

Heuristic planners face significant increases in runtime when dealing with scenarios involving a greater number of medicine types, higher repetitions of each drug, consideration of additional bio-sites, or allowance for more dosage options. Additionally, tighter constraints around the goal and the utilization of PK biodistribution trajectories spanning longer periods also increase planner runtimes.

However, runtime is not solely determined by the number of parameters, such as the number of medicines, but also by their pharmacokinetic-pharmacodynamic (PKPD) responses. For instance, Morphine, a potent pain-relief drug, consistently shortened runtime for both opt-hmax and opt-hrmax planners. Despite increasing the number of optional plans, the inclusion of Morphine reduced runtime by almost half in various instances. To illustrate, in a problem with a six-drug set but without Morphine, opt-hmax and opt-hrmax solved it in 55.68 sec. and 169.71 sec., respectively. Yet, when Morphine was added to the drug set, the runtimes decreased to 19.52 sec. for opt-hmax and 37.23 sec. for opt-hrmax.

The planner chooses Morphine whenever possible, because the representation presented here currently lacks the means to represent its long-term health complications, which would prohibit doctors from prescribing it unless absolutely necessary: It is an opioid drug; frequent usage may cause the patient to become addicted. This type of information is very much relevant to the medication plan, and in future work, we will address means to represent it.

The experiments presented in Table 2 show that opt-hmax yielded the best results in terms of coverage, while sat-hadd and sat-hradd performed the worst among the tested planners. This disparity in performance may be due to the fact that the representation is not of a simple numeric planning (SNP) form. In SNP, effects can only increase/decrease a variable by a constant. However, the GMP representation does not adhere to this restriction. For non-SNP representations like the GMP, sat-hadd and sat-hradd use a blind heuristic [35].

In this representation, we used biodistribution trajectories and the direct-effect model to assess the time-changing effect of the drug across various bio-sites. However, we did not model the impact of other biochemical properties that can affect drug distribution and effect. The von Maltzahn [42] example previously discussed serves to illustrate: The baseline behavior of the second medicine changes in response to coagulation presence, which is why the first drug is used initially to increase coagulation in tumors. To describe the PK changes of the second drug, we could naively use conditional effect, stating that the drug's biodistribution changes in the presence of coagulation. Yet, to prevent conflicting effects from overlapping conditions, we would have to explicitly consider all possible combinations of the conditions. Given r conditions, this means specifying 2^r condition combinations. We are currently investigating parametric PKPD models that consider the impact of additional properties on drug behavior, without the need for explicit conditions.

6 Related Work

Personalized medication planning is a recent area of research within the more general area of personalized treatment planning, tailoring therapeutic interventions to individual patient and medical needs.

A prime example of treatment planning is *radiation therapy planning*, whereby motion planning and machine learning methods aid in the design of radiation treatments by simulating precise configurations, such as visualizing affected tissues, and suggesting treatment parameters that align with medical objectives. This contributes significantly to the continuum of patient care, as highlighted in recent reviews by [43], [11], and [22]. Due to significant variations in tumor morphology, position, and other patient-specific factors, personalized radiotherapy plans must be formulated in order to minimize damage to normal tissue while persevering sufficient tumor control.

Another example of this general treatment plan is in planning treatments for patients with multiple diseases, by merging available multiple single-disease clinical guidelines. This intricate process includes substituting drugs when adverse or redundant interactions occur, adjusting and scheduling tests to monitor for such interactions, and other related tasks. Techniques for automating this process utilize constraint satisfaction [44, 29], model-based reasoning [30, 28], and planning [32, 14, 26]. While these investigations touch on medicine choices, they do not provide personalized dosage or hourly medication schedule. Instead, they produce plans that span weeks or months, rather than hours. Similarly, the use of HTN planning to schedule chemo-therapy treatments over weeks, considering doctor availability [18], differs from GMP.

A final example area of treatment planning is discussed by Amir et al. [7, 8]. They use teamwork theory to enhance and improve collaboration between patients and various caregivers: family, medical professionals, and support organizations; the improvements allow the caregivers to better coordinate their human decision-making.

7 Conclusion

We introduced general medication planning (GMP), a relatively new and underexplored area in personalized medical treatment planning. This domain requires planning for actions whose effects are durative, multi-dimensional and non-linear, which may overlap in time. We present a PDDL+ representation for GMP problems, using events to describe the behaviour of drugs in the patient's body.

Experiments with clinical data from rodents demonstrate that the suggested representation is capable of modeling GMP scenarios. However, they also reveal a rapid increase in planning runtime with various problem parameters. Consequently, even seemingly simple problems from a medical perspective remain impractical for application. In addition to the challenges outlined in Section 5, a key objective for future work is the development of heuristics and representation changes to facilitate solving medically-relevant problems. Furthermore, it is pertinent to evaluate the performance of other planners such as Metric-FF [20] and OPTIC [9]. This necessitates the translation of the suggested representation from PDDL+ using dedicated translators, such as the one proposed by Percassi et al. [27].

GMP presents an exciting and promising opportunity, albeit with significant challenges. Ideally, GMP planning involves the consideration of dozens or even hundreds of drugs, some of which exhibit effects that accumulate over days and weeks due to frequent administrations. This is particularly true for many psychiatric drugs, which pose unique complexities to the planning process.

Acknowledgements

The research was partially supported by ISF grants #2443/23 (Shleyfman) and #1373/24 (Kaminka). Alon is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship. As always, thanks to K. Ushi.

Ethics/Broad Impact Statement

This paper presents a first step towards a medical decision-support system. The motivation for personalized medication planning is to re-

move biases that may pre-exist in data or general medication plans. Creating treatment plans from scratch using personalized data ensures that we do not apply (possibly biased) general treatment plans. This paper focuses on using AI planning to carry out such personalization. While the data we use is taken from published sources, the system has not been evaluated clinically, even with rodents. Actual clinical use would require mitigating automation bias by the user, consideration of execution robustness (e.g., actual timing of medication), reasoning about long term effects, and other factors influencing the safe use of automated medication planning. In addition, we foresee the need for enhancing clinicians' understanding of the proposed plans and easing their workload, such as visualization and explainability [23, 5]. All of these challenges we leave for future work.

References

[1] S. Akhtar, Q. Khan, S. Anwar, G. Ali, M. Maqbool, M. Khan, S. Karim, and L. Gao. A comparative study of the toxicity of polyethylene glycol-coated cobalt ferrite nanospheres and nanoparticles. nanoscale res lett 14: 386, 2019.

[2] F. K. Alaboud. *Personalising Medication & Activity Regimes Using Novel State Progression Models for Forward Search with PDDL+*. PhD thesis, King's College London, 2022.

[3] F. K. Alaboud and A. Coles. Personalized medication and activity planning in pddl+. In *ICAPS*, pages 492–500, 2019.

[4] L. Alon, H. Weitman, and G. A. Kaminka. First steps towards planning for targeted medicine. In *ICAPS Workshop on KEPS*, 2023.

[5] L. Alon, H. Weitman, A. Shleyfman, and G. A. Kaminka. Towards explainable general medication planning. In *ECAI Workshop on Explainable AI in the Medical Domain*, 2024.

[6] L. Alon, H. Weitman, A. Shleyfman, and G. A. Kaminka. Towards personalized medication planning. In *ICAPS Workshop on KEPS*, 2024.

[7] O. Amir, B. Grosz, K. Gajos, S. Swenson, and L. Sanders. From care plans to care coordination: Opportunities for computer support of teamwork in complex healthcare. In *CHI*, 2015.

[8] O. Amir, B. Grosz, and K. Gajos. Mutual influence potential networks: Enabling information sharing in loosely-coupled extended-duration teamwork. In *IJCAI*, pages 796–803, 2016.

[9] J. Benton, A. Coles, and A. Coles. Temporal planning with preferences and time-dependent continuous costs. In *ICAPS*, volume 22, pages 2–10, 2012.

[10] M. C. Berenbaum. Synergy, additivism and antagonism in immunosuppression: a critical review. *CEI*, 28(1):1–18, 1977.

[11] J. C. L. Chow. Artificial intelligence in radiotherapy and patient care. In *Artificial Intelligence in Medicine*, pages 1–13. Springer, 2021.

[12] M. Dawes. Co-morbidity: we need a guideline for each patient not a guideline for each disease. *Family Practice*, 27(1), 2010.

[13] J. E. Doran, D. Michie, and D. G. Kendall. Experiments with the graph traverser program. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 294(1437):235–259, 1966.

[14] J. Fdez-Olivares, E. Onaindia, L. Castillo, J. Jordán, and J. Cózar. Personalized conciliation of clinical guidelines for comorbid patients through multi-agent planning. *AIME*, 96:167–186, 2019.

[15] M. A. Felmlee, M. E. Morris, and D. E. Mager. *Mechanism-Based Pharmacodynamic Modeling*, pages 583–600. Humana Press, 2012.

[16] M. Fox and D. Long. Modelling mixed discrete-continuous domains for planning. *JAIR*, 27:235–297, 2006.

[17] L. E. Gerlowski and R. K. Jain. Physiologically-based pharmacokinetic modeling: principles and applications. *Journal of pharmaceutical sciences*, 72(10):1103–1127, 1983.

[18] A. Gonzáez-Ferrer, A. ten Teije, J. Fdez-Olivares, and K. Milian. Automated generation of patient-tailoed electronic care pathways by translating computer-interpretable guideslines into hierarchical task networks. *AI in Medicine*, 57:91–109, 2013.

[19] P. E. Hart, N. J. Nilsson, and B. Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cybern.*, 4(2):100–107, 1968.

[20] J. Hoffmann. The metric-ff planning system: Translating "ignoring delete lists"to numeric state variables. *JAIR*, 20:291–341, 2003.

[21] C. Hull. Pharmacokinetics and pharmacodynamics. *British Journal of Anaesthesia*, 51(7):579–594, 1979.

[22] S. Jones, K. Thompson, B. Porter, M. Shepherd, D. Sapkaroski, A. Grimshaw, and C. Hargrave. Automation and artificial intelligence in radiation therapy treatment planning. *JMRS*, 2023.

[23] A. Korikov, A. Shleyfman, and C. Beck. Counterfactual explanations for optimization-based decisions in the context of the gdpr. In *ICAPS workshop on explainable AI planning*, 2021.

[24] M. Kumar, P. Kulkarni, S. Liu, N. Chemuturi, and D. K. Shah. Nanoparticle biodistribution coefficients: A quantitative approach for understanding the tissue distribution of nanoparticles. *Advanced Drug Delivery Reviews*, 194:114708, 2023.

[25] D. Lezama-Martinez, M. Elena Hernandez-Campos, J. Flores-Monroy, I. Valencia-Hernandez, and L. Martinez-Aguilar. Time-dependent effects of individual and combined treatments with nebivolol, lisinopril, and valsartan on blood pressure and vascular reactivity to angiotensin ii and norepinephrine. *JCPT*, 26(5):490–499, 2021.

[26] M. Michalowski, M. Rao, S. Wilk, W. Michalowski, and M. Carrier. Mitplan 2.0: Enhanced support for multi-morbid patient management using planning. In *AIME*, pages 276–286, 2021.

[27] F. Percassi, E. Scala, and M. Vallati. A practical approach to discretised PDDL+ problems by translation to numeric planning. *JAIR*, 76:115–162, 2023.

[28] L. Piovesan and P. Terenziani. A mixed-initiative approach to the conciliation of clinical guidelines for comorbid patients. In *AIME: Knowledge Representation for Health Care*, volume 9485 of *LNCS*, pages 95–108, 2015.

[29] L. Piovesan and P. Terenziani. A constraint-based approach for the conciliation of clinical guidelines. In *IBERAMIA*, volume 10022 of *LNCS*, pages 77–88, 2016.

[30] D. Riaño and A. Collado. Model-based combination of treatments for the management of chronic comorbid patients. In *AIME*, volume 7885 of *LNCS*, pages 11–16, 2013.

[31] M. Román, L. García, M. Morales, and M. J. Crespo. The combination of dantrolene and nimodipine effectively reduces 5-ht-induced vasospasms in diabetic rats. *Scientific Reports*, 11(1):9852, 2021.

[32] I. Sánchez-Garzón, J. Fdez-Olivares, E. Onaindía, G. Milla, J. Jordán, and P. Castejón. A multi-agent planning approach for the generation of personalized treatment plans of comorbid patients. In *AIME*, pages 23–27, 2013.

[33] E. J. Santos, N. Nassehi, E. W. Bow, D. R. Chambers, E. S. Gutman, A. E. Jacobson, J. A. Lutz, S. A. Marsh, K. C. Rice, A. Sulima, et al. Role of efficacy as a determinant of locomotor activation by mu-opioid receptor (mor) ligands in female and male mice. ii. effects of novel mor-selective phenylmorphans with high-to-low mor efficacy. *Pharmacology Research & Perspectives*, 11(4):e01111, 2023.

[34] E. Scala, P. Haslum, S. Thiébaux, and M. Ramirez. Interval-based relaxation for general numeric planning. In *ECAI*, pages 655–663, 2016.

[35] E. Scala, P. Haslum, S. Thiébaux, et al. Heuristics for numeric planning via subgoaling. In *IJCAI*, pages 3228–3234, 2016.

[36] E. Scala, P. Haslum, S. Thiébaux, and M. Ramírez. Subgoaling techniques for satisficing and optimal numeric planning. *JAIR*, 68:691–752, 2020.

[37] E. Scala, A. Saetti, I. Serina, and A. E. Gerevini. Search-guidance mechanisms for numeric planning through subgoaling relaxation. In *ICAPS*, pages 226–234, 2020.

[38] F. Sime, M. Roberts, and J. Roberts. Optimization of dosing regimens and dosing in special populations. *CMI*, 21(10):886–893, 2015.

[39] S. K. Singh, A. Mohammed, O. A. Alghamdi, and S. M. Husain. New approaches for targeting drug resistance through drug combination. In *Combination Therapy Against Multidrug Resistance*, pages 221–246. Academic Press, 2020.

[40] P.-L. Toutain and A. Bousquet-mélou. Plasma terminal half-life. *JVPT*, 27(6):427–439, 2004.

[41] O. Turan, P. Bielecki, K. Tong, G. Covarrubias, T. Moon, A. Rahmy, S. Cooley, Y. Park, P. M. Peiris, K. B. Ghaghada, et al. Effect of dose and selection of two different ligands on the deposition and antitumor efficacy of targeted nanoparticles in brain tumors. *Molecular pharmaceutics*, 16(10):4352–4360, 2019.

[42] G. von Maltzahn, J.-H. Park, K. Y. Lin, N. Singh, C. Schwöppe, R. Mesters, W. E. Berdel, E. Ruoslahti, M. J. Sailor, and S. N. Bhatia. Nanoparticles that communicate in vivo to amplify tumour targeting. *Nature materials*, 10(7):545–552, 2011.

[43] C. Wang, X. Zhu, J. C. Hong, and D. Zheng. Artificial intelligence in radiotherapy treatment planning: Present and future. *Technology in Cancer Research & Treatment*, 18:1–11, 2019.

[44] S. Wilk, W. Michalowski, M. Michalowski, K. Farion, M. M. Hing, and S. Mohapatra. Mitigation of adverse interactions in pairs of clinical practice guidelines using constraint logic programming. *JBI*, 46(2): 341–353, 2013.

[45] D. F. Wright, H. R. Winter, and S. B. Duffull. Understanding the time course of pharmacological effect: a PKPD approach. *BJCP*, 71(6):815–823, 2011.