

**A SYSTEMS TOXICOLOGY
FRAMEWORK FOR IMPROVING THE
IDENTIFICATION OF PARACETAMOL
OVERDOSE**

CHANTELLE LOUISE MASON

A thesis submitted in partial fulfilment of the
requirements of Liverpool John Moores

University for the degree of

Doctor of Philosophy

May 2019

Acknowledgements

I would like to thank my supervisor, Steve, firstly for introducing me to coffee. Without the coffee, there's no way this would have been possible. But seriously, I'm truly grateful for all the time and advice you've given to me over the past 3 years. I doubt I'll ever be lucky enough to have a boss like you again. To Joe, thank you so much for all the time you have given to me. That keen mathematical eye you have didn't miss a trick, and it's helped me to develop so much. Sotiris and Ian, thank you for your statistical insights!

To WebbLab, my family! Terry and Andrew, you're the annoying brothers I never got. Thanks for hurling abuse at me in the office every day. If anything, you've taught me how to discipline people, because you both absolutely need keeping in line. Ross, thank you for always being there to look up to, the original number 1! Adam, the sweetest of souls with the most brutal one-liners, thanks for keeping me sane. Rita, what a lady, thank you for keeping me laughing. WebbLab, thank you for putting up with me (especially during my dramatic, high-maintenance moments of panic). I'm going to miss each and every one of you.

Mum, thanks for every single time you've helped me when I couldn't help myself, I haven't forgotten and I never will. Pops, I've applied so many of your words of wisdom over the past few years, and I'm starting to feel the benefits. Thank you. N'Taya, thanks for giving me something else to stress about other than this thesis. But seriously, I haven't forgotten all those times you kept an eye on me when it was needed. To my best friends: Hollie, Gemma, April, Helena, Leah, Elle and Ashleigh, thank you all for being so patient. I appreciate every single one of you for sticking around. To Jo, yet again you've been with me every step of the way. Thanks for being stubborn with me when I needed it most. I couldn't have got here without you.

To Joshua, thanks for putting up with the meltdowns; for accepting the 5am wake-ups and the midnight switch offs; for the one million biscuits that have fueled this thesis; thank you for supporting me in everything, in every single way, always. I'm excited to see what our next installment brings the both of us.

To my Grandad, thank you for always wanting the best for me, even if some of my decisions have scared you; nothing makes me happier than updating you of every single step (even though you feel I don't update enough). Trust me, you have been, and remain at the top of my priority list. Grandpops, this thesis is dedicated to you.

Abstract

Paracetamol (APAP) overdose is the leading cause of acute liver failure and a concerning global health issue. However, the current clinical treatment framework is heavily criticized for its sub-optimality. Within this thesis, a systems toxicology approach is taken in an attempt to provide further insight into the APAP overdose problem, and propose potential improvements to the current treatment framework. In Chapter 2, a proof-of-concept pre-clinical pharmacokinetic-pharmacodynamic (PKPD) model describing APAP metabolism and corresponding toxicity biomarkers (ALT, HMGB1, full K18, fragmented K18) is defined. A statistical model is combined with the PKPD framework to simulate *in silico* population groups with the aim of predicting initial APAP dose, time since overdose, and probability of liver injury. In chapter 3, an identifiability analysis is performed on the PKPD model to identify parameter uncertainties. The model is also extended, enabling predictions for individuals deemed both “healthy” and “high-risk”.

In 2017 I was awarded the *in vitro* toxicology society mini-fellowship award, which funded 4 weeks of training in experimental wet-lab techniques. The training was used to investigate the effects of various combinations of APAP and its antidote, N’Acetylcysteine (NAC), on *in vitro* hepatocyte functionality. Subsequently, in chapter 4, the effect of the antidote (NAC) is incorporated into the PKPD model structure, and an additional toxicity measure is defined, describing severe loss of cell functionality. Different NAC regimens are tested, investigating their effect on both of our proposed toxicity measures. Through collaboration with the Royal Infirmary, Edinburgh, we obtained access to a clinical dataset of approximately 3,600 APAP overdose patients. In Chapter 5, a population-pharmacokinetic (Pop-PK) APAP model is defined, with PK parameters optimised based on this dataset. The framework has the ability to account for random inter-individual differences in PK parameter values. Current clinical toxicity thresholds are investigated and compared to those proposed by our model for various demographic groups.

Contents

Chapter 1: Background and introduction	1
1.1. Background	2
1.1.1. Clinical problem	2
1.1.2. Biological overview	4
1.1.3. Modelling overview	7
1.1.4. Review of relevant previously published models	9
1.2. Research novelty.....	13
1.3. Thesis overview.....	16
Chapter 2: A pre-clinical systems toxicology framework – predicting liver injury potential from novel biomarker concentrations	19
2.1. Background	20
2.2. Methods	23
2.2.1. Pharmacokinetic-Pharmacodynamic modelling	23
2.2.2. Statistical modelling	32
2.3. Results	45
2.3.1. PKPD model optimisation	45
2.3.2. PKPD model validation	50
2.3.3. Visualisation of <i>in silico</i> derived data	51
2.3.4. Classification of <i>in silico</i> derived data	51
2.3.5. Predicting initial dose and time since administration as continuous variables	51
2.3.6. Predicting the probability of liver injury	54
2.3.7. Sensitivity analysis.....	57
2.4. Chapter discussion.....	59
Chapter 3: A pre-clinical systems toxicology framework – Improving the confidence, biological relevance and clinical application potential.....	62
3.1. Background	63
3.2. Methods	65
3.2.1. Experimental data description	65
3.2.2. Identifiability analysis	66
3.2.3. Model refinement.....	67
3.2.4. Parameter optimisation.....	73
3.2.5. Virtual population simulation.....	73
3.2.6. Predicting time since administration and initial dose as continuous variables	75
3.2.7. Visualisation.....	76
3.2.8. Classification.....	77

3.2.9. Predicting the probability of liver injury	77
3.3. Results	77
3.3.1. Identifiability analysis – original model (defined in Chapter 2)	77
3.3.2. Model refinement.....	78
3.3.3. Identifiability analysis – refined model	83
3.3.4. Predicting initial dose and time since administration as continuous variables	87
3.3.5. Visualisation and classification of <i>in silico</i> derived data	89
3.3.6. Predicting the probability of liver injury	91
3.3.7. Visualising the probability of liver injury following an APAP dosed based on biomarker concentrations	93
3.4. Chapter discussion.....	94
Chapter 4: Pre-clinical systems toxicology approach to treating paracetamol overdose: predicting cell death dynamics and investigating antidote treatment regimens.....	97
4.1. Background.....	98
4.2. Methods	100
4.2.1. Experimental methods.....	100
4.2.2. Statistical methods.....	108
4.2.3. Mathematical modelling.....	111
4.3. Results	116
4.3.1. Wet-lab experiments.....	116
4.3.2. Multiple linear regression - predicting cell viability	125
4.3.3. Modelling the effect of NAC dosing regimen on the functionality of hepatocytes, and the probability of liver injury	126
4.4. Discussion.....	140
Chapter 5: Clinical paracetamol overdose model.....	145
5.1. Background.....	146
5.2. Methods	148
5.2.1. Clinical data description.....	148
5.2.2. Model selection criteria	148
5.2.3. Mixed effects modelling background.....	149
5.2.4. Model optimisation.....	150
5.2.5. Covariate model.....	156
5.2.6. The non-linear mixed effects model.....	157
5.3. Results	159
5.3.1. APAP Pop-PK model formulation	159
5.3.2. Identifying relationships between patient demographics and model parameters	162

5.3.3. Model parameterisation	163
5.3.4. Model evaluation.....	164
5.3.5. Model simulation.....	166
5.4. Chapter discussion.....	180
Chapter 6: Discussion	183
6.1. Conclusions.....	184
6.2. Future work	190
References	192
Appendices	213

Chapter 1: Background and introduction

1.1. Background

1.1.1. Clinical problem

Acetaminophen (paracetamol, APAP) is the most commonly used painkiller in the world [1] and, when taken in therapeutic doses, millions of people worldwide benefit from its effective and safe analgesic effects [2]. Whilst relatively safe and beneficial at therapeutic doses, APAP is also the most commonly used drug in overdose situations, accounting for 40% of all drug poisoning cases in the UK [3]. APAP overdose is the leading cause for acute liver failure (ALF) in the Western world [4] and therefore represents a concerning global health issue [5]. For example, in the USA, ALF is responsible for approximately 56,000 emergency room visits, 2,600 hospitalizations and 500 deaths per year [6]. Between 2015-2016, there was an 11% increase in deaths involving APAP in the UK [7]. In England and Wales, APAP poisoning results in approximately 40,000 hospital admissions, 20 liver transplants and 200 deaths per year [8].

There is, however, an antidote for APAP overdose, N'acetylcysteine (NAC). Current decisions on whether this antidote should be given are based on the simple nomogram treatment framework [9]. This nomogram treatment framework can be seen in Figure 1-1. The red treatment line shown in Figure 1-1 defines a relationship between APAP concentration and time-since-APAP-dose. If a patient falls above this threshold, NAC treatment is given. It should be noted however that this treatment line is heavily dependent on accurate knowledge of time since APAP ingestion [10] information, which is often provided by the patient, and can therefore be highly inaccurate. Decisions of whether or not to continue NAC treatment are then guided by measuring changes in blood-based markers of potential liver injury: APAP concentration is measured to determine how much drug is currently circulating within the system; alanine aminotransferase (ALT) concentration is measured to determine liver injury status; and the international normalised ratio (INR) is calculated to determine hepatic transplant appropriateness [11].

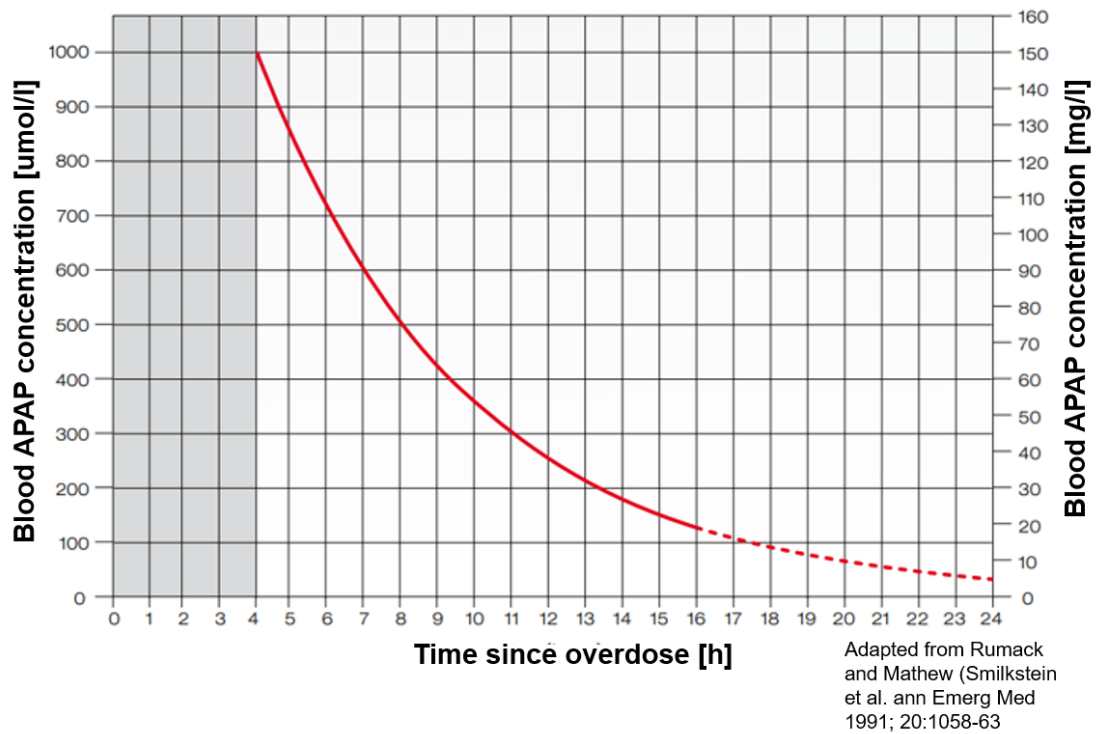


Figure 1-1: The nomogram treatment framework used in the clinic to determine antidote necessity in the APAP overdose case. *The red contour is the treatment line/threshold. Any patient with an APAP concentration – time since overdose related observation falling above the line should begin treatment, any patient with an observation below the line is deemed unnecessary for treatment. The nomogram treatment line has been reduced by 25% from the original threshold first defined in 1976 as a safety measure to account for patients who may potentially be high-risk prior to APAP overdose/measurement errors [9].*

ALT is an enzyme within the liver that catalyzes the transfer of an amino group (L-alanine) to a keto-group (α -ketoglutarate) and, the products of this reaction are crucial for respiration [12]. Measurements of this biomarker within the blood, show that the enzyme has been released from hepatocytes, indicating cell death/damage [13]. INR is a measure of how long it takes for the blood to clot (prothrombin time) that is normalised to account for any potential differences in laboratories [14]. Measuring the rate of blood clotting is important; the liver is a primary source of circulating coagulation factors and liver injury is associated with alterations in blood coagulation [15]. Whilst ALT is currently the most widely used blood-based biomarker for measuring drug-induced-liver-injury (DILI) [16], it's elevation is known to be non-DILI specific and only represents probable liver-injury post-occurrence [17]. Additionally, INR has been criticized for its lack of sensitivity; patients presenting with an acceptable INR measurement (too low to be deemed at risk), can be left untreated when they may in fact have a poor prognosis and could potentially benefit from a transplant [18].

Further increasing the complexity of the APAP toxicity identification problem, is the fact that some individuals may be more susceptible to toxicity than others, predominantly due to differences in the ability to synthesise or maintain sufficient glutathione (GSH) levels [10]. Potential factors that may increase a patient's susceptibility include age, pre-existing liver disease, concurrent use of alcohol or chronic APAP use [5]. A crucial factor is the patient's nutritional state [19]. The nomogram treatment line shown in Figure 1-1 has been lowered by 25% from the original line that was first defined in 1976, following recommendations from the Food and Drug Administration (FDA) in 1981 to account for such potential risk factors [9,20]. With an increase in both alcohol-related and eating disorder-related hospital admissions [21,22], there is an urgent need for further investigation of the utility of this "one-size-fits-all" threshold.

Since uncertainty currently exists around accurate APAP overdose identification (due to the dependence on insensitive biomarkers and the requirement of an accurate knowledge of time since overdose), antidote (NAC) administration decisions can be imprecise. These imprecise decisions can result in either treatment refusal when it is in fact necessary, or unnecessary administration of the antidote to patients who are not likely to experience liver injury. Unnecessary NAC administration can cause a range of side-effects spanning from nausea and vomiting to anaphylactoid reactions, thereby exacerbating the problem of ill-informed treatment [23]. Such inaccurate treatment decisions have led to an estimated cost of £58.1 million in the UK since 2012 [24]. It is clear that the sub-optimal treatment framework for APAP overdose requires improvement.

1.1.2. Biological overview

In an attempt to improve the APAP overdose treatment framework, we build a mathematical model that is representative of the biological processes. This biological representation can be investigated computationally to provide both insight into the APAP overdose problem, and suggestions for improving the treatment framework. The biological processes that occur following an APAP dose are well defined in the literature. In this section, we define the key processes that form the basis of our mathematical model.

1.1.2.1. APAP metabolism

Following APAP ingestion, metabolism predominantly takes place within the liver, in the phase II pathway, with a small proportion being metabolised in the Phase I pathway [25]. Within these two metabolic pathways, three principle mechanisms are involved: conjugation with a glucuronosyl group (UDP-glucuronic acid) in a process called glucuronidation within the phase II pathway; conjugation with a sulfo-group (PAPS) in a process called sulfation in the phase II pathway; and oxidation via the cytochrome P-450 dependent, mixed function phase 1 oxidative pathways [26]. These processes can be visualised in Figure 1-2.

In therapeutic cases, metabolism occurs predominantly via glucuronidation and sulfation, with a small amount (<10%) being oxidized into the toxic metabolite, N-acetyl-p-benzoquinone imine (NAPQI). Approximately 3-5% of APAP is excreted unchanged in the urine and less than 1% is recovered in the bile (an aqueous solution that is produced and secreted by the liver and consists mainly of bile salts, phospholipids, cholesterol, conjugated bilirubin, electrolytes and water [27]). Detoxification of NAPQI occurs via conjugation with hepatic stores of glutathione (GSH) [28]. In overdose cases, the UDP-glucuronic acid and PAPS pools deplete, leaving more APAP available for oxidation such that NAPQI can accumulate. If NAPQI levels become too high, the GSH pools will eventually be overwhelmed, resulting in excess levels of NAPQI which can bind to intracellular macromolecules causing lipid peroxidation of membranes and eventual hepatic toxicity [29]. The antidote currently used to treat APAP overdose, NAC, provides cysteine to hepatocytes, acting as a pre-cursor to GSH synthesis [30]. An increase in GSH synthesis increases detoxification capacity and reduces the likelihood of macromolecule binding, cell death, and subsequent progression into drug-induced liver injury (DILI).

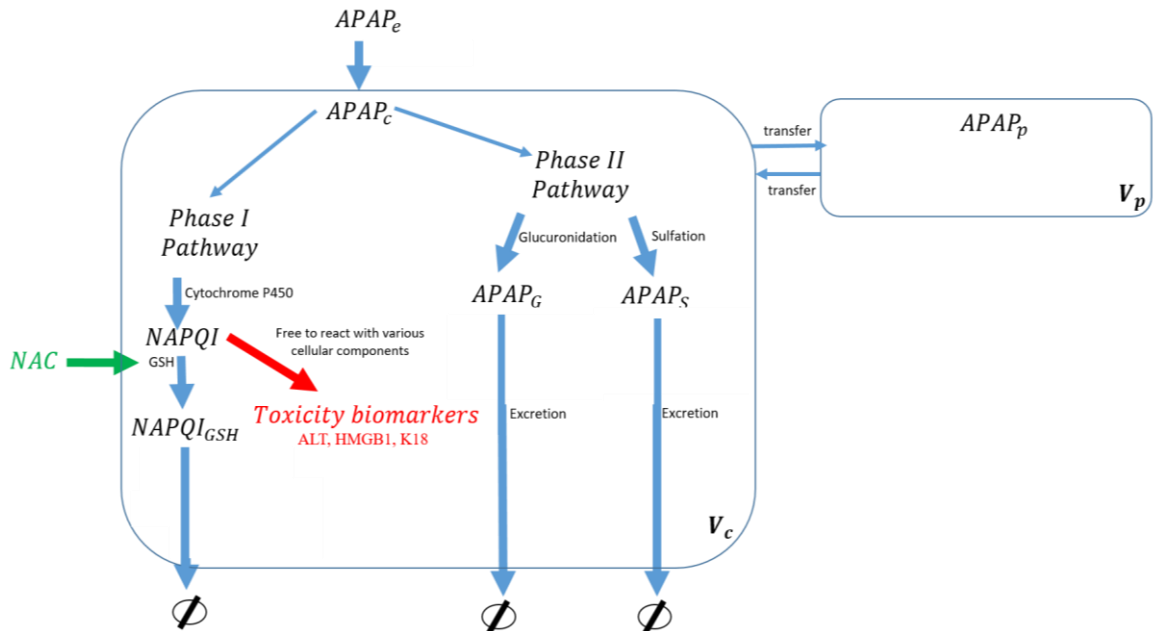


Figure 1-2: APAP metabolism process. Extracellular APAP ($APAP_e$) enters hepatocytes. Intracellular APAP ($APAP$) is predominantly metabolised through the phase II pathway. APAP conjugates with either a glucuronosyl group (UDP-glucuronic acid) to form APAP-glucuronide ($APAP_G$) or a sulfo-group (PAPS) to form APAP-sulfate, $APAP_S$. When phase II pathways are saturated, APAP will also undergo phase I reactions. Here, APAP combines with the cytochrome P450 enzymes resulting in the production of a highly toxic metabolite, NAPQI. NAPQI is detoxified via conjugation with GSH and then excreted. In overdose cases, the GSH stores deplete and NAPQI is able to accumulate and react with various cellular components, forming toxic adducts and causing hepatocyte damage, which can result in toxicity biomarkers. The antidote for APAP overdose, NAC, works as a pre-cursor to GSH, therefore increasing the detoxification capacity in hepatocytes.

1.1.2.2. Hepatocellular toxicity

When toxicity occurs, cell death can usually be described by one of two forms: necrosis or apoptosis [31]. Necrosis is a rapid and irreversible form of cell death which occurs when an external trauma causes cells to quickly die. Apoptosis is a much slower physiological process, occurring when the cell has the energy (ATP) to programme its own death [32]. In the APAP toxicity scenario, necrosis is thought to be the dominant form of hepatocellular cell death [33–35]. However, for mild APAP toxicity cases, the apoptotic form of cell death is also observed [33]. Having the capability to not only predict potential cell death in toxicity cases, but also gain further insight into the mechanism of cell death taking place can help to better understand the severity level of a toxicity case. In less severe cases, ATP stores may not be critically affected and therefore the cell may have the capability to programme its own death via apoptosis. If, however, the case of overdose is severe, this external threat may be too large, causing rapid inflammation of the cells and eventual necrosis.

1.1.2.3. Biological toxicity markers

In terms of identifying potential liver toxicity, there have been many promising novel mechanistic biomarker candidates recently identified. The mitochondria within cells is the core of ATP and therefore energy production [36]. The presence of biomarker Glutamate dehydrogenase (GLDH) in circulation indicates mitochondrial dysfunction and subsequent loss of mitochondrial membrane integrity, which occurs during hepatocellular necrosis [37]. GLDH correlates with the currently used ALT, but is more tissue specific. Although GLDH has improved specificity when compared with ALT, this biomarker is not sensitive enough to predict APAP induced toxicity prior to the presence of ALT [38]. Another biomarker that correlates well with ALT and is released into the serum following tissue damage is an enzyme known as malate dehydrogenase (MDH), but this biomarker is known to be less tissue-specific [37]. Molecular forms of High-mobility group box 1 (HMGB1) and Keratin-18 (K18) have recently been experimentally established as blood-based biomarkers for APAP-hepatotoxicity [39]. Circulating hypo-acetylated HMGB1 and full-length K18 illustrates cell necrosis. However, during apoptotic structural rearrangements, full-length K18 undergoes caspase-mediated-cleavage and, subsequently, fragmented-K18 can be released into the circulation, representing cell apoptosis [37]. Additional to these mechanism-based biomarkers, are microRNAs (miRNA) which are non-coding RNAs involved in regulation of gene expression [40]. The most abundant hepatic miRNA, miR-122, is known to be released into circulation following liver damage and has been shown to have improved organ specificity over ALT [37].

1.1.3. Modelling overview

1.1.3.1. Quantitative Systems Toxicology modelling

As the number of drugs developed to improve the health of society quickly increases, so too does the requirement for efficient and accurate toxicity predictions [41]. The well-known Sulfanilamide tragedy (1937) [42] caused over 100 deaths unexpectedly from nephrotoxicity (kidney toxicity), resulting in the Food Drug and Cosmetics Act (FDCA) being passed by congress in 1938. This set the precedent for the current toxicity testing strategy (assessing effects of drugs on animals prior to initial administration to humans) [43]. However, identifying toxicity following drug administration through such traditional preclinical testing strategies, particularly in the cases of rare adverse drug reactions (ADRs), has proven to be a difficult task

[44]. This issue has been recently improved by the advancement of technologies that now enable simultaneous measurements of transcriptomic, proteomic, and imaging readouts as well as organelle and cellular phenotypes [45]. In the informatics-accelerated healthcare sciences era, the goal is to combine systems modelling approaches with this increasing amount of data in order to efficiently and robustly make toxicity predictions [44]. Quantitative systems toxicology (QST) modelling comprises a useful tool to reduce and refine animal testing and is now considered as both an essential component of modern toxicity testing and a foundation for individualised therapeutic treatment [43]. However, since this approach is a computational representation of biology, and not the true biology itself, it is recognized that any uncertainty in predictions must be quantified and limitations specified, so that the models may be utilised to their full potential [46].

The three main current modelling approaches for toxicity identification are: Quantitative Structure Activity Relationship (QSAR) modelling, network based modelling, and pharmacokinetic-pharmacodynamic (PKPD) modelling [43]. These modelling techniques can be further separated into three main categories [47]: (i) “top down” modelling and simulation, which aims to understand the system’s characteristics based on observed data; (ii) “bottom up” highly mechanistic modelling, which is based on knowledge of the human biology and utilises *in vitro* experimentation for input data; and (iii) “middle out” which combines the “top down” and “bottom up” approaches, utilising *in vivo* data to determine unknown/uncertain model parameters.

QSAR modelling is often used at the pre-clinical stage, due to its ability to screen through vast amounts of compounds and identify those that are likely to be active at a target site [48]. QSAR models are useful in situations when the mechanism of the toxicity is well understood, such as skin sensitization/mutagenicity [49,50]. In terms of DILI prediction, Cheng et al. [51] reported the first QSAR model in 2003. They built a training set of 382 drugs and drug-like compounds with dose-response data. They applied the model to a set of 54 external compounds, with 81% of the compounds being correctly classified. Typically, however, since this early study, most published QSAR models suffer from a low statistical performance or are limited to small, inaccurate datasets [52]. For example, a QSAR study of the permeability of drugs in the central nervous system (CNS) is based on incorrect descriptor values (including molecular weights differing by 14) for Acetylsalicylate and Aspirin [53]. Although guidance is provided, there are many issues that still exist within QSAR

modelling. Dearden et al. [54] reviewed 21 potential issues that exist within the approach which include: poor transferability; unacknowledged omission of data-points; and replication of compounds within a dataset.

Network based modelling aims to understand how the integration of molecular events allows biological processes to occur [43]. The modelling is typically derived from the mathematical formalism of graph theory, aiming to describe complex qualitative relationships between multiple biological components. The resulting dynamical networks can be built using ordinary/partial differential equations, Boolean algebra or Bayesian inferences, ranging from continuous to discrete and deterministic to stochastic descriptions [43].

The goal of pharmacokinetic (PK) data analysis is to estimate the parameters that determine the rates of drug absorption, distribution, metabolism and elimination [55]. Such models often contain both a pharmacokinetic component for the dose-plasma relationship and a pharmacodynamic component for the concentration-effect relationship [56]. Validated PKPD models have the capability, in theory, to predict responses at any time, following any dose administered, via any route. PKPD modelling now has a significant role in all phases of drug development: from pre-clinical stages, to carry out toxicokinetic studies of chemical entities; to phase 1 studies, to enable crucial decisions regarding tolerability/efficacy; and phase 2 studies, to help analyse the dose-response relationship [57]. Successful PK studies resulting in clinical application span a wide variety of patients and conditions, including diabetes, clotting disorders, arthritis, organ transplantation and self-poisoning [58]. In this thesis, we use a PKPD modelling approach to quantify the biochemical dynamics in APAP overdose.

1.1.4. Review of relevant previously published models

The drug induced liver injury (DILI)sim initiative is a partnership that has created a software (DILIsym), which applies various QST methods, aiming to better understand and predict liver toxicity [59]. Hepatotoxicity is modelled predominantly via the mechanisms of bile acid transporter inhibition, mitochondrial dysfunction, oxidative stress, and lipotoxicity [60]. DILIsym comprises of multiple sub models of these toxicity mechanisms, whilst additionally having the capacity to predict cell death mechanisms (e.g. via hepatic apoptosis, necrosis or regeneration) and also novel liver injury biomarkers concentrations [61]. The tool can be used to investigate the effect of various treatment regimens. For example, Woodhead et al. proposed

alterations to a novel antibiotic (BAL30072) treatment regimen in order to reduce potential liver injury side-effects using DILIsym [62]. Additionally, Longo and colleagues. used DILIsym to simulate hepatotoxic potential following treatment of adjunct therapies for Parkinson's disease (Tolcapone and Entacapone) [63].

Relevant to the work within this thesis, investigations through DILIsym suggest that newly proposed NAC dosing regimens are superior to the regimen currently used in the APAP overdose case, therefore recommending that modification of the current treatment nomogram should be considered [64]. Although promising, the modelling framework within DILIsym makes large simplifications of the underlying biology, and the predictivity of its results are heavily reliant on the (often *in vitro*) data used during parameter optimisation. We create a detailed mechanistic framework, for which the confidence of the data used during optimisation is tested and improved where necessary. Our final result is a modelling framework that accounts for individual variability between humans, and is optimised based on a clinical dataset of 3,616 patients. Whilst DILIsym is currently the most comprehensive computational approach for integrating data into a DILI safety model, the approach is still in its infancy; it is based on a number of simplified modelling assumptions and has a reliance on parameterisation via only *in vitro* experimental data [65]. The addition of more detailed mechanistic descriptions and *in vivo* validation will further improve the potential of this computational framework.

Promisingly, there has been much focus on further developing mechanistic models in the area of DILI. For example, researchers in the field have created many mathematical models to assess APAP metabolism and its effects on biological functionality. Reith et al. [66] provide a system of ordinary differential equations (ODEs). The parameters within the system are obtained by fitting to human data consisting of patients dosed with pain relief. The system has the aim of providing clarification of the role of the glucuronidation and sulfation pathways, outlining a basis for examining APAP metabolism in various disease states. Ochoa et al. [67] employ a multi-scale approach. They firstly create a spatiotemporal prediction of drug and metabolite concentrations within the liver, and then, at the whole-body level, include blood flow between organs. Krauss et al. [68] further extend an already published physiologically based-pharmacokinetic (PBPK) model to include a genome scale network reconstruction of a human hepatocyte within the liver tissue compartment of the model. The model has the aim of quantifying the impact of paracetamol-induced liver failure by monitoring changes in 67 hepatic functions (for

example, production of cysteine, creatine, GSH and ATP) during a paracetamol overdose.

Geenen et al. [69] also further develop a PBPK model by introduction of an additional sub-model. The group combine a PBPK model of APAP metabolism and disposition with a refined version of their previously developed mathematical systems model of hepatic glutathione homeostasis [70]. The combined model predicts that APAP-induced hepatic GSH depletion results in elevated concentrations of the biomarkers within their model (5-oxoproline and ophthalmic acid in blood and 5-oxoproline in urine). They also provide additional insight; the biomarker concentrations are also likely to be influenced by prolonged administration of APAP and by the concentrations of intracellular metabolites such as methionine. Using a system of ODEs with Michaelis-Menten kinetics to investigate the link between APAP, hepatic GSH status and concentrations of related intermediary metabolites, Stahl et al. [71] confirm the importance of methionine availability on hepatocellular GSH status. Leclerc et al. [72], investigate the effect of GSH depletion on reactive oxygen species (ROS) production by NAPQI by coupling a drug PK model with a systemic biology (SB) model (which uses a system of ODEs to represent the Nrf2 pathway). The group identify a transition between 0.025-0.25 μM APAP, from a NAPQI detoxification phase, to a NAPQI and ROS accumulation phase, due to depletion of GSH.

Remien et al. [73] provide a predictive modelling framework for acetaminophen-induced liver damage (MALD) by deriving ODEs describing changes in AST, ALT and INR following an APAP dose. The authors optimise initial APAP dose amount and time since overdose by fitting the resulting ODEs to clinical data (from 53 overdose patients). The model represents the first dynamical rather than statistical approach to determining poor prognosis in patients with APAP-induced-liver-injury, by using commonly measured patient biomarkers to estimate initial APAP dose amount, time since dose, and subsequent outcome. Remien et al. [74] extend this framework to a cell-based model. While the original MALD model has the ability to predict for acute APAP overdose cases, the extended model is aimed at also accounting for chronic overdose cases. The author's results suggest that there is a threshold relating NAPQI and GSH production rates, where liver injury occurs rapidly. That is, when the rate of NAPQI production is less than that of GSH, the liver is unlikely to suffer damage. If, however, NAPQI production exceeds GSH, NAPQI accumulation will occur, leading to subsequent liver damage.

Ben-Shachar et al. [75] propose a retrospective study complementary to Remien's model. Both models describe the detailed mechanisms by which APAP is detoxified in the liver in both therapeutic and overdose situations. Whilst Remien's model aims to predict overdose occurrence, Ben-Shachar's model aims to determine whether an overdose is likely to lead to fatal liver damage. The framework they propose is a multi-compartmental model (gut, plasma, liver, tissue, urine) consisting of 21 ODEs that describe APAP transport and metabolism. The liver compartment describes metabolism of APAP, with ODEs representing sulfation, glucuronidation, conjugation with GSH, production of toxic metabolites and subsequent liver damage. The study strengthens the evidence that APAP metabolism has a dependence on the expression level of certain key liver enzymes, but further extends the knowledge to include the effects of antidote treatment (NAC) on these aforementioned liver enzyme expression levels, and effects on subsequent toxic potential.

Reddyhoff et al. [76] provide a cell-based model that describes the three major metabolic pathways that impact upon APAP clearance (sulfation, glucuronidation and oxidation). The aim of their model is to reduce the complexity of the APAP metabolism system so that analytical methods can be used for investigation. Sensitivity analysis is applied to investigate which parameters have the largest effect on progression to toxicity, identifying rates for glucuronidation and oxidation as key parameters. Results from this study suggest that the sulfation pathway is less influential on the dynamics of the systems when compared to the glucuronidation/oxidation pathways. Whilst predictive models for improving APAP overdose seem promising to date, most models rely on insensitive biomarkers, such as ALT, for the prediction of DILI [17]. Going forward, mathematical models that include the novel biomarkers discussed in Section 1.1.2 would be a significant addition to the current state of the art. In terms of these novel biomarkers for APAP toxicity there is currently only one modelling study within the literature that focuses on HMGB1, namely, Shoda et al. [60]. However, although insightful to the understanding of HMGB1, the focus of this work is on the role of HMGB1 with regards the innate immune response rather than APAP toxicity.

Often, mathematical modelling approaches are applied to animal data, particularly at the proof-of-concept stage due to the quantity of data available and also the reduced amount of variability from the lab animals [77]. Translatability of these models to the human case is often then of particular interest [43,78]. In one such

translation study, Pery et al [79], find that toxicokinetic/toxicodynamic approaches based on alternative methods and modelling only, have the potential to predict *in vivo* liver toxicity with good accuracy comparable to *in vivo* methods. They conclude this result through the combination and calibration of a PBPK model with a toxicodynamic model for cell viability. There are also some human PBPK models already published which aim to improve treatment in the APAP overdose case. For example, Zurlinden et al [80] propose a mathematical model optimised against humans under overdose conditions. The model aim is to predict an APAP dose whilst elucidating the effects of blood sampling time and additional blood sampling requirements. Whilst the model can accurately predict initial dose for training and test datasets, it has poor predictability when applied to real clinical data.

1.2. Research novelty

Within this thesis, a new framework is provided for APAP toxicity identification using novel biomarkers to predict initial dose, time since ingestion and subsequent liver injury probability. Whilst Remien et al. [73] also suggest a predictive framework, the biomarker they use for toxicity prediction (ALT) is currently criticised for its lack of sensitivity. Within our framework, we combine this biomarker with additional novel biomarkers that are believed to have the potential to predict APAP-induced liver injury pre occurrence (namely, HMGB1, full K18 and fragmented K18). Whilst Shoda et al. [60] model HMGB1 in relation to the immune response, our focus with HMGB1 differs, namely by using HMGB1 within a panel of DILI biomarkers, aiming specifically to improve APAP toxicity identification rather than immune function. In comparison to other biomarker based APAP toxicity studies, this study covers both overdose and non-overdose cases, identifying the key biomarkers that discriminate between the two situations.

Whilst there is currently one APAP toxicity model which takes into consideration individuals that may have depleted GSH stores [81], the focus of their study is to understand how depleted nutritional levels and certain lifestyle choices affect GSH regeneration. We incorporate the fact that nutritional levels have subsequent effects on GSH levels, and use this to extend our model framework, enabling the quantification of how probable liver injury differs for nutritionally depleted individuals. In order to reduce the skepticism that often arises around QST models, we also propose an identifiability analysis technique, based on the profile likelihood approach [82]. This allows the assessment and improvement of any uncertainties in

the model predictions, whilst also enabling the modelling framework to be as biologically relevant as possible given the accessible data.

Additionally, although Remien's model [74] accounts for cell functionality dynamics following APAP dose, there is no accountability for effects of the NAC antidote. Our modelling framework extends previous work by predicting cell functionality dynamics in both APAP, and APAP plus NAC antidote cases. As a result, we have the ability to understand NAC effects on APAP dose likely to induce probable liver injury (based on novel biomarker concentrations) and also loss of cell functionality. Finally, we propose a new APAP overdose population pharmacokinetic model, optimised against approximately 3,600 overdose patients. We investigate various demographic groups (e.g. age, weight) with regards to their maximum tolerable APAP dose and find the demographic group that is predicted to have the lowest tolerance to APAP. We also compare current clinical toxicity thresholds against those predicted from our model and propose some potential adjustments that could provide substantial clinical impact.

There have been limited mathematical approaches aimed at improving the APAP overdose problem to date. In the literature, there are deterministic approaches specifically aimed at modelling APAP metabolism. For example, the mechanistic models underpinning both DILIsym and the framework defined by Geenen et al. [69,83] are built using ordinary differential equations (ODEs) representing metabolism processes over time within various compartments of the body, including transfer of the drug between these compartments within a physiologically based PK (PBPK) framework. Parameters within these models are either fixed values from the literature, or optimised by fitting to experimental data. The predictivity of these models is therefore heavily dependent on the data used during optimisation, and with DILIsym, for example, having a large dependence on *in vitro* data, and Geenen et al basing their model on only limited data for clearance, uncertainty in parameter values remains a huge issue for these types of models.

Although we also take an ODE approach to modelling APAP metabolism and corresponding biomarkers, we consider these potential limitations by defining an uncertainty quantification technique. We define an identifiability analysis technique, which is able to determine areas of uncertainty within our model structure that require refinement/additional data during optimisation. We are subsequently able to define the uncertainty within our predictions, and improve confidence in our framework. Additional to this, we optimise our final clinical model against human

data from approximately 3,600 patients, rather than depending on *in vitro/in vivo* information. Mathematical models are often used as a basis for numerical analysis of the system to investigate the effects of variations in the mechanistic processes. For example, Remien et al. [74] conduct a local sensitivity analysis by perturbing parameters by 50% and analysing changes in subsequent biomarker outcomes. Additionally, in a secondary publication, they conduct a bifurcation analysis to better understand the thresholds for the transition from no liver injury, to severe liver injury. In a similar fashion, we conduct a local sensitivity analysis on our pre-clinical model parameters to determine those of particular importance prior to translation to the human clinical case, but we also apply an identifiability analysis to quantify the confidence in the parameters being perturbed.

As well as modelling drug absorption, distribution, metabolism and excretion (ADME) processes over time, it is possible to also model spatial effects (e.g. spatial heterogeneity within tissues/organs). For example, Ochoa et al. [67] provide a model framework that uses coupled partial differential equations to model both hepatocyte and sinusoid functionality (with distribution described by an advection-diffusion process), subsequently linking this spatial liver model to an ODE based PBPK whole body model. However, since our focus is to specifically model plasma (blood) toxicity biomarkers with the aim to better guide patient treatment, this organ level spatial detail is neglected to focus instead on identification of key relationships between plasma glutathione (GSH) as well as novel biomarkers. We particularly focus on the combination of biomarker simulations with statistical analyses to provide an improved APAP treatment framework.

The key mathematical novelty of this study, is the combination of mathematical and statistical modelling approaches. Deterministic and statistical models are formulated, providing key mechanistic insight, particularly with regard to novel biomarker mechanisms, and also potential causes for inter-individual variability in the APAP overdose setting. Simulation of the mathematical models is utilised throughout to create virtual populations, which are combined with statistical visualisation and classification techniques, with the aim of predicting patient toxicity outcome. A more detailed description of the combination of mathematical and statistical approaches is represented by the schematic in Figure 1-3. Firstly, a dose is randomly selected from a uniform distribution, and the corresponding biomarker time-course profile is simulated using the mechanistic model (A). Similar to the dose selection, a time-point is randomly selected from a uniform distribution. From the

previously simulated biomarker time-course profile, the randomly selected time point is used as an index to extract the corresponding biomarker measurements (B). This information is combined to provide a single combination of biomarker concentrations for a simulated individual given a random dose of APAP, and having their biomarker concentrations measured at a random time-point (C).

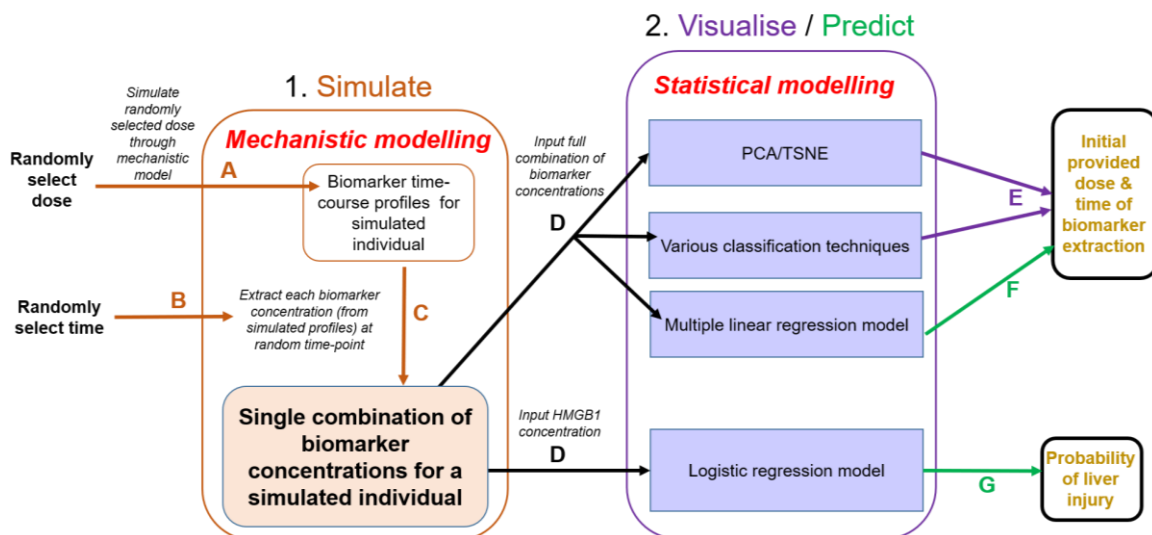


Figure 1-3: Schematic representation of process combining mechanistic and statistical modelling. (A,B,C) represent how the mechanistic model is used to retrieve simulated biomarker concentrations for an individual receiving a random dose, and having concentrations extracted at a random time-point. (D) represents how these biomarker concentrations are fed into the statistical models. (E,F,G) represents how the statistical models are used to visualise/predict the initial dose, time of biomarker extraction, and probability of liver injury.

The mechanistic modelling is now combined with the statistical modelling (D) in one of two ways: the full combination of biomarker concentrations are used in conjunction with various statistical techniques to visualise /predict the initial dose/time of biomarker extraction (E,F), or the HMGB1 concentration alone is combined with a logistic regression model to predict the probability of liver injury.

1.3. Thesis overview

This thesis details the development of a systems toxicology approach with the aim of improving the current paracetamol overdose clinical framework. A PKPD modelling approach is taken, using systems of ODEs to represent paracetamol metabolism and the corresponding toxicity, represented by novel biomarker concentrations. The thesis consists of 4 research chapters, 3 of which have a pre-clinical focus, with the final chapter focusing on the clinical application.

Chapter 2 describes the development of a pre-clinical 2-compartment APAP PKPD model. The PK element represents APAP metabolism within the liver and the PD

element represents conventional (ALT) as well as novel (HMGB1, full K18 and fragmented K18) APAP toxicity biomarkers. Within our model, we assume that APAP toxicity is dependent upon the depletion of GSH. The panel of biomarkers are tested for significance in predicting the probability of liver injury (measured by experimental mouse histology scores), with HMGB1 alone being identified as the most significant predictor. The PKPD model is also simulated to create a virtual population of mice receiving an APAP dose between 0-600mg/kg. Biomarker concentrations are extracted at a random time-point to mimic the clinical scenario and are used to make predictions of initial dose, time since APAP administration, and subsequent liver injury.

In Chapter 3 we refine the model developed in Chapter 2. The model refinement is led by the results of an identifiability analysis technique which is developed based on a profile likelihood approach [82]. I was introduced to this technique by Professor Jens Timmer's research group during a two-week training visit, funded by the LJMU International Mobility Award that I received in 2017. In Chapter 2, all biomarkers are assumed to follow similar dynamics, which are represented by an indirect PD model (i.e. APAP is assumed to inhibit a pre-cursor, GSH, which then regulates the dynamics of the biomarkers). In Chapter 3, in order to improve the structural unidentifiability found in the original model, we develop piecewise linear equations to represent the biomarker dynamics of full and fragmented K18 (which are necrotic and apoptotic versions of the same biomarker). We also optimise the model against additional data, including fasted-mouse data. This data enables us to model the biomarker dynamics of a GSH-depleted mouse, relevant for the high-risk GSH-depleted population that often present within the clinic. Robust, biologically relevant predictions for initial dose amount, time since administration and probability of liver injury are calculated for both "healthy" and "high-risk" populations using numerical and statistical methods.

The framework described in Chapters 2 and 3 may be sufficient to make toxicity predictions upon initial investigation of a suspected APAP overdose case. However, the dynamics of those models do not take into account the effect of the APAP overdose antidote, NAC. Using the IVTS Mini-Fellowship Award that I obtained in 2017, I undertook a four week wet-lab training placement at the University of Liverpool. The *in vitro* experimental data obtained is used within this thesis with two primary aims: (1) better understanding of the influence of NAC on the biomarker dynamics; and (2) better understanding of toxicity at the cellular level. The results

from the experiments are detailed within Chapter 4. The potential of biomarker data predicting cell functionality is discussed and modelled, as well as the effect of NAC administration schedules on our proposed toxicity thresholds (for both probable liver injury and loss of cell functionality).

Chapters 2, 3 and 4 provide a predictive pre-clinical framework. However, having formed a new collaboration with Dr. James Dear (Queens Medical Research Institute QMRI, Edinburgh) we obtained access to a clinical dataset of approximately 3,600 APAP overdose patients from 3 UK hospitals. Chapter 5 describes our clinical two compartment Pop-PK APAP model based on this data. We carry out a covariate correlation analysis and propose 3 human covariate factors that we predict are correlated with various PK parameters. We analyse differences in the maximum tolerable APAP dose in various patient groups, and find that younger, lower weight individuals have a lower tolerance to APAP. We also compare current clinical toxicity thresholds against those predicted from our model and propose some potential screening adjustments that with further investigation, could significantly impact the clinical environment.

Chapter 2: A pre-clinical systems toxicology framework – predicting liver injury potential from novel biomarker concentrations

The research defined in this chapter has been published in Clinical Pharmacology and Therapeutics: Pharmacometrics and Systems Pharmacology (CPT: PSP):

Mason CL, Leedale J, Tasoulis S, Jarman I, Antoine DJ, Webb SD. Systems Toxicology Approach to Identifying Paracetamol Overdose. CPT Pharmacometrics Syst Pharmacol. 2018;7(6):394–403. Available from: <http://doi.wiley.com/10.1002/psp4.12298>

2.1. Background

Whilst paracetamol/APAP is the most commonly prescribed painkiller worldwide [1], in the western world, this analgesic is also the leading cause of acute liver failure (ALF) [4]. There is an antidote to treat APAP overdose cases, N-acetylcysteine (NAC), which if provided in a timely manner following APAP doses, has been shown to reduce the likelihood of progression into drug-induced liver injury (DILI) [30]. The effectiveness of the antidote is thought to diminish in cases where it is administered beyond 8 to 10 hours of initial APAP dose [84]. Although NAC is currently the most effective APAP overdose treatment, the antidote also has many associated adverse side effects. Typical side effects are those such as rash, vomiting and more uncommonly, anaphylactoid reactions [30]. The current decision to administer NAC is predominantly based upon the nomogram treatment line [9]. A visualisation of the treatment framework can be seen in Chapter 1, Figure 1-1. A patient's point on this framework is found by determining the relationship between their measured APAP blood concentration and the time elapsed since APAP overdose. If the patient's point on the framework falls above the nomogram treatment line, NAC administration should begin, however if they fall below the line, NAC treatment is deemed unnecessary. This initial decision to begin treatment is heavily dependent on the patient's knowledge of time elapsed since ingestion [20]. However, this information is often highly unpredictable within the clinical setting. Once NAC treatment begins, the decision to continue treatment is based on measuring changes in the following biomarker concentrations – APAP and alanine aminotransferase (ALT). Additionally, the patient's international normalised ratio (INR) is calculated [11].

ALT elevation represents probable liver injury post-occurrence [17] and is the most widely used blood based biomarker for measuring DILI [16]. INR is an adjusted version of the prothrombin time (PT) test. The PT test evaluates an individual's ability to appropriately form blood clots, and INR slightly adjusts this measurement to account for changes in PT reagents, allowing results from different laboratories to be compared [85]. AST is another DILI biomarker [16]. AST could be used for toxicity predictions as it accumulates in the blood due to liver damage, but it is not liver specific as it is also linked to other pathologies, e.g. heart injury [86]. Increased serum total bilirubin (TBL) is another potential marker; it is indicative of the substantial loss of functional hepatocytes [17]. However, since the marker identifies loss of hepatocytes, it does not predict hepatotoxicity potential but instead is a post-

occurrence indicator, similar to ALT. In order to improve the treatment of APAP-induced DILI via NAC therapy, biomarkers which can predict liver damage *a priori* are required. Recently, biomarkers K18 and HMGB1 have been shown to add value to the measurement of ALT [87] and have the potential to predict DILI pre-occurrence. However, such new biomarkers are often examined singly and clarification of their functional relationships is required to aid clinical implementation [88]. In this chapter, we aim to combine these novel biomarkers with conventional biomarkers, with the objective of improving the prediction accuracy of APAP overdose cases.

In silico modelling allows for the development of mechanistic understanding of biological systems which may not always be possible from *in vitro/in vivo* experiments alone. An inter-disciplinary, systems toxicology approach is a cost-effective way of understanding and predicting drug efficacy and toxicology whilst additionally complying with the 3Rs (a scientific framework for the effective use of animals in research) [89]. There are many examples of successful *in silico* modelling frameworks that have been previously developed to study APAP metabolism and its associated toxic potential. Many are discussed in Chapter 1, but frameworks relevant to the work in this chapter are redefined here.

For example, Reith et al. [66] propose a system of ODEs. The parameters within the system are obtained by fitting to human clinical data. The system has the aim of providing clarification of the role of the glucuronidation and sulfation pathways, outlining a basis for examining APAP metabolism in various disease states. Ochoa et al. [67] approach their modelling in a multi-scale manner, firstly creating a spatiotemporal prediction of drug and metabolite concentrations within the liver, and then, at the whole-body level, including blood-flow between organs. Remien et al. [73] derive ordinary differential equations (ODEs) describing changes in AST, ALT and INR, creating a model (based on a single-time overdose) with the ability to predict acetaminophen-induced liver damage. The authors optimise initial APAP dose amount and time since overdose by fitting the resulting ODEs to clinical data (from 53 overdose patients). Remien et al. [74] subsequently provide a cell-based model extension to this framework. Reddyhoff et al. [76] also construct a cell-based model which describes major metabolic pathways impacting on APAP clearance. Sensitivity analysis determines the parameters that have the largest effect on the progression to toxicity. Ben-Shachar et al. [75] create a retrospective study complementary to Remien's model. Whilst Remien's model has the aim of predicting

overdose occurrence, Ben-Shachar's has the aim of determining whether an overdose would lead to fatal liver damage. Our study also further extends Remien's approach by combining ALT (used in their study) with additional novel biomarkers (K18 and HMGB1) that have the potential to predict APAP-induced liver injury pre-occurrence. Additionally, the study is extended to non-overdose and overdose cases, attempting to identify the key biomarkers that discriminate between the two situations. Shoda et al. [60] also mechanistically model the biomarker HMGB1. They focus on the role of HMGB1 with regards to the innate immune response, concluding that HMGB1 is a key input for immune cell activation. Our focus is to instead investigate HMGB1 within a panel of DILI biomarkers, with the aim of predicting APAP toxicity in mice.

In this chapter, we propose a novel framework with the aim of predicting initial APAP dose, time since administration and the probability of APAP-induced liver injury. The platform is distinctive primarily due to the use of novel biomarkers, optimised within the PKPD framework by combining the use of deterministic modelling with statistical PKPD analysis.

The mouse is widely considered to be a good model for APAP toxicity prediction in humans [90] and due to the rich data sets available, we utilise mouse-derived data in this chapter to develop our new *in silico* framework. At this initial stage of model development, we wish to avoid the uncertainties associated with APAP human overdose data. The human clinical APAP model is described in Chapter 5. The basis of this current chapter is to demonstrate the development and validation of our new predictive framework using the more amenable mice data, prior to translation to the human situation.

The results from our investigation define previously undocumented PK parameters for APAP in mice, which can now be used for further application throughout the field. Additionally, as mentioned above, novel biomarkers are included in the framework as a panel, rather than the current approach which is to consider them individually. The key aim of the modelling development at this initial stage is to determine which biomarkers are most predictive of APAP induced-DILI.

Chapter aims

- Create a mechanistic PKPD model that represents APAP metabolism and corresponding biomarkers (ALT, HMGB1, K18) in mice.
- Test biomarkers as a panel to determine the most significant in predicting DILI.
- Combine analyses to provide a predictive DILI framework based on biomarker measurements only, eradicating the current requirement for verbal information regarding overdose amount and time elapsed since overdose.

2.2. Methods

2.2.1. Pharmacokinetic-Pharmacodynamic modelling

2.2.1.1. Pharmacokinetic model development

To understand how to effectively model the ADME processes, the observed APAP concentrations (from four datasets from two separate published studies [39,91] recording APAP concentration over time in mice following intraperitoneal administration of 50, 150, 500 and 530 mg/kg doses) were transformed into log-space and plotted over time. From Figure 2-1, for the lower doses particularly (green and black profiles), and also the highest dose (red profile), it is clear to see that the distribution and elimination of APAP within these mice occurred over two phases. For the 500mg/kg dose, although this result did not seem so clear, the data had no observations beyond four hours and therefore the full course of distribution and elimination was difficult to conclude. Based on the results in Figure 2-1, we assumed that the distribution and elimination of APAP occurred over two-phases, and therefore a two compartmental modelling approach was necessary. A two compartmental model assumes that tissues within the body can be separated into two different compartments [92]. The first compartment mainly consists of the blood and well-perfused tissues such as heart, liver and brain etc. and the secondary compartment consists of poorly perfused tissues such as fat and bone. [93]. As mentioned, the liver is a well-perfused tissue, and therefore any liver-related dynamics within the model (sections 2.2.1.2 and 2.2.1.3) will occur within the central compartment. A two-compartmental model schematic is described in Figure 2-2.

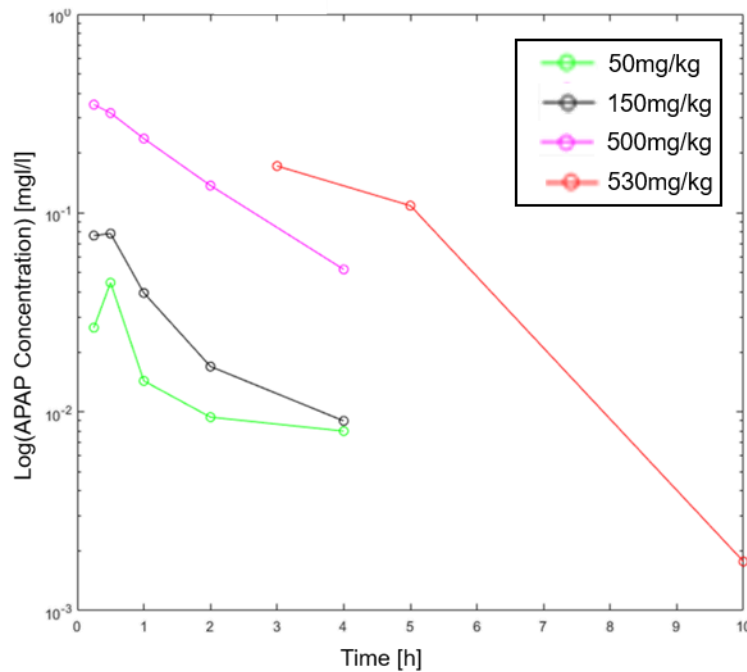


Figure 2-1: Log-transformed paracetamol concentration versus time. The green profile represents the log-transformed APAP time course following an initial dose of 50mg/kg, the black profile represents the log-transformed APAP time course following an initial dose of 150 mg/kg, the pink profile represents the log-transformed APAP time course following an initial dose of 500 mg/kg, and the red profile represents the log-transformed APAP time course following an initial dose of 530 mg/kg.

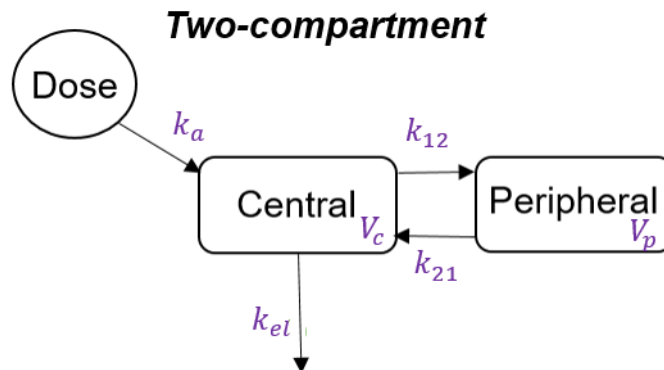


Figure 2-2: Two-compartment PK model schematic. A dose of APAP is absorbed into the central compartment at a rate k_a . The central compartment has a theoretical volume, V_c . APAP is secondarily transferred to the peripheral compartment (theoretical volume V_p) at a rate k_{12} . APAP is transferred from the peripheral compartment to central compartment at a rate k_{21} . APAP is then eliminated from the central compartment at a rate k_{el} .

Based on the model schematic in Figure 2-2, which is assumed to represent APAP metabolism in mice, two ordinary differential equations (ODEs) were used to represent changes in APAP concentration within the two PK compartments (central and peripheral) of the mice in the following system,

$$\frac{d[C_c]}{dt} = \frac{k_a D_0 e^{-k_a t}}{V_c} + k_{21}[C_p] \frac{V_p}{V_c} - k_{12}[C_c] - k_{el}[C_c], \quad (2-1)$$

$$\frac{d[C_p]}{dt} = k_{12}[C_c] \frac{V_c}{V_p} - k_{21}[C_p], \quad (2-2)$$

where $[C_c]$ represents the central compartment concentration of APAP ($\mu\text{mol/l}$), $[C_p]$ represents the peripheral compartment concentration of APAP ($\mu\text{mol/l}$), k_a represents the absorption rate from the peritoneal cavity (h^{-1}), D_0 represents the initial dose (mg), k_{21} represents the transfer rate from the peripheral to the central compartment (h^{-1}), k_{12} represents the transfer rate from the central to the peripheral compartment (h^{-1}), V_p represents the theoretical volume of the peripheral compartment (l/kg), V_c represents the theoretical volume of the central compartment (l/kg), k_{el} represents the overall elimination rate (summation of both excretion and metabolism processes) (h^{-1}), and t represents the time variable (h). Note that our model is representative of intraperitoneal administration. For applications to oral administration, the absorption rate parameter, k_a , would be multiplied by a bioavailability fraction to implicitly take into account effects of gastric emptying and absorbed fraction.

We used the method of Laplace transforms [94] to solve the equations for APAP in the central and peripheral compartment. Briefly, Laplace transformation works by transforming differential equations on the time domain, t , to a frequency domain, s . It is then possible to analytically solve these simpler equations, and transform them back to the original time domain. The methodology of solving equations (2-1) and (2-2) via Laplace transformation is described below.

If we wish to transform a function of time, for example, $A(t)$, via Laplace transformation \mathcal{L} , the Laplace notation can be applied:

$$\mathcal{L}(A(t)) = \int_0^{\infty} e^{-st} A(t) dt, \quad (2-3)$$

where s is the frequency domain we wish to transform the variable A onto.

The Laplace transform of a first order differential is, for example, defined by:

$$\frac{dA}{dt} \simeq s\mathcal{L}(A) - A(0), \quad (2-4)$$

where $\frac{dA}{dt}$ is the rate of change of variable A over the time (t) domain, s is the frequency domain we wish to transform the variable A onto, $\mathcal{L}(A)$ is the Laplace transform of variable A and $A(0)$ is the initial concentration (at $t = 0$) of variable A .

We can now apply this theory and transform our differential equations ((2-1) and (2-2)) into Laplace form:

$$s\mathcal{L}([C_c]) - [C_c](0) = \frac{k_a D_0 \mathcal{L}(e^{-k_a t})}{V_c} + \frac{k_{21} \mathcal{L}([C_p]) V_p}{V_c} - k_{12} \mathcal{L}([C_c]) - k_{el} \mathcal{L}([C_c]), \quad (2-5)$$

$$s\mathcal{L}([C_p]) - [C_p](0) = \frac{k_{12} V_c \mathcal{L}([C_c])}{V_p} - k_{21} \mathcal{L}([C_p]) \quad (2-6)$$

The Laplace transform of the exponential is found using the following:

$$\mathcal{L}(e^{-k_a t}) = \int_0^{\infty} e^{-(s+k_a)t} dt = \frac{1}{s+k_a} \quad (2-7)$$

For simplicity, we let

$$Y_1 = \mathcal{L}([C_c]), \quad Y_2 = \mathcal{L}([C_p])$$

Therefore equations (2-5) and (2-6) can now be re-written:

$$sY_1 - [C_c](0) = \frac{k_a D_0}{V_c(s+k_a)} + \frac{k_{21} V_p Y_2}{V_c} - k_{12} Y_1 - k_{el} Y_1 \quad (2-8)$$

$$sY_2 - [C_p](0) = \frac{k_{12} V_c Y_1}{V_p} - k_{21} Y_2 \quad (2-9)$$

Initially (at $t = 0$), there is no APAP in either the central or peripheral compartment, therefore,

$$[C_c](0) = 0, \quad [C_p](0) = 0$$

We can now use this information to solve equation (2-9) for Y_2 , namely:

$$sY_2 - 0 = \frac{k_{12} V_c Y_1}{V_p} - k_{21} Y_2. \quad (2-10)$$

Rearranging gives,

$$Y_2 = \frac{k_{12}V_c Y_1}{V_p(S+k_{21})}. \quad (2-11)$$

Substituting this result for Y_2 ((2-11)) into equation (2-8)) we can solve for Y_1 as follows:

$$SY_1 = \frac{k_a D_o}{V_c(S+k_a)} + \frac{k_{21}V_p k_{12}V_c Y_1}{V_c V_p(S+k_{21})} - k_{12}Y_1 - k_{el}Y_1. \quad (2-12)$$

Rearranging to give,

$$SY_1 + (k_{12} + k_{el})Y_1 - \frac{k_{21}k_{12}Y_1}{(S+k_{21})} = \frac{k_a D_o}{V_c(S+k_a)}, \quad (2-13)$$

and factorising for Y_1 gives,

$$Y_1 \left(S + k_{12} + k_{el} - \frac{k_{21}k_{12}}{(S+k_{21})} \right) = \frac{k_a D_o}{V_c(S+k_a)}. \quad (2-14)$$

If we allow

$$\alpha + \beta = k_{21} + k_{12} + k_{el} \quad (2-15)$$

and

$$\alpha\beta = k_{el}k_{21}, \quad (2-16)$$

then

$$Y_1 = \frac{k_a D_o}{V_c} \left(\frac{1}{S+k_a} \right) \left(\frac{S+k_{21}}{(S+\alpha) + (S+\beta)} \right). \quad (2-17)$$

Taking the inverse Laplace transform of Y_1 , we get:

$$[C_c](t) = \mathcal{L}^{-1}(Y_1), \quad (2-18)$$

from which we get:

$$[C_c](t) = \frac{k_a D_o}{V_c} \left[\frac{(k_{21} - \alpha)}{(k_a - \alpha)(\beta - \alpha)} e^{-\alpha t} + \frac{(k_{21} - \beta)}{(k_a - \beta)(\alpha - \beta)} e^{-\beta t} + \frac{(k_{21} - k_a)}{(\alpha - k_a)(\beta - k_a)} e^{-k_a t} \right], \quad (2-19)$$

where α and β are related to the model parameters as follows,

$$\alpha = \frac{1}{2} \left(k_{12} + k_{21} + k_{el} + \sqrt{(k_{12} + k_{21} + k_{el})^2 - 4k_{21}k_{el}} \right),$$

and

$$\beta = \frac{1}{2} \left(k_{12} + k_{21} + k_{el} - \sqrt{(k_{12} + k_{21} + k_{el})^2 - 4k_{21}k_{el}} \right).$$

We fitted equation (2-19) to the four aforementioned datasets simultaneously using a Nelder-Mead search algorithm [95] (Fminsearch tool in Matlab [96]). Parameters k_a, k_{21}, V_c, α and β were optimised in order to minimise the difference between the model output and the observed APAP mice data. Note that all subsequent data fitting in this thesis also employs this algorithm. Optimised parameter values (from the model defined in this chapter) are defined in the results section of this chapter.

2.2.1.2. Glutathione depletion model

APAP is metabolised predominantly by the phase II pathway via glucuronidation and sulfation, but at high doses, the sulfation process saturates and an increased amount of APAP is then metabolised by the phase I (CYP) pathway. In this case, APAP combines with cytochrome P450 to create N-acetyl-p-benzoquinoneimine (NAPQI), a highly toxic metabolite [97]. NAPQI can be detoxified by glutathione (GSH). However, with large doses of APAP, GSH stores deplete [98] and NAPQI accumulates leading to the possibility of DILI. For this reason, the biomarker response dynamics were modelled to be dependent on GSH depletion. For a visualisation of the key APAP metabolism processes, please refer to Figure 1-2.

To formulate the rate of change of GSH, the following mass-action-based ODE was considered

$$\frac{d[\text{gsh}]}{dt} = k_i - k_o[\text{gsh}] - k_G[\text{NAPQI}][\text{gsh}], \quad (2-20)$$

where k_i (h^{-1}) is the background production rate of GSH, $[\text{gsh}]$ ($\mu\text{mol/l}$) is the concentration of GSH, k_o (h^{-1}) is the natural decay/background-usage rate of GSH, $[\text{NAPQI}]$ is the concentration of NAPQI ($\mu\text{mol/l}$) and k_G ($\mu\text{M}^{-1}\text{h}^{-1}$) is the decay rate of GSH due to binding with NAPQI. The model is extended by assuming the rate of change of NAPQI as follows,

$$\frac{d[\text{NAPQI}]}{dt} = \xi k_{el}[C_c] - k_G[\text{NAPQI}][\text{gsh}] - k_p[\text{NAPQI}], \quad (2-21)$$

where ξ is the proportion of CYP-activated APAP that is transformed into NAPQI, k_{el} is the total rate of APAP elimination as described above, and k_p (h^{-1}) is the rate at which NAPQI binds to other (non-GSH) cellular proteins.

Because NAPQI is short-lived and the associated reactions are known to be rapid on the time-scale of APAP depletion, we assume $[\text{NAPQI}]$ is at a quasi-steady state on the time-scale of APAP PK, namely,

$$0 = \xi k_{el}[C_c] - k_G[\text{NAPQI}][\text{gsh}] - k_p[\text{NAPQI}], \quad (2-22)$$

and thus,

$$[\text{NAPQI}] = \frac{\xi k_{el}[C_c]}{k_G[\text{gsh}] + k_p}. \quad (2-23)$$

Substituting this term into equation (2-20) we obtain,

$$\frac{d[\text{gsh}]}{dt} = k_i - k_o[\text{gsh}] - \frac{[\text{gsh}]\xi k_{el}[C_c]}{[\text{gsh}] + k_{pr}}, \quad (2-24)$$

where $k_{pr} = \frac{k_p}{k_G}$, represents the ratio of NAPQI forming other protein adducts relative to detoxification by GSH. We then assume a constant GSH background level to be gsh_0 , so that in the absence of APAP and at steady-state we have $k_i = k_o gsh_0$. Using this, we obtain

$$\frac{d[\text{gsh}]}{dt} = k_o gsh_0 - k_o[\text{gsh}] - \frac{\xi k_{el}[\text{gsh}][C_c]}{[\text{gsh}] + k_{pr}}. \quad (2-25)$$

In summary, the first term in equation (2-25) relates to the natural production of GSH. The second term in the equation represents the background usage of GSH. The final term in the equation represents the loss of free GSH due to interaction with NAPQI.

2.2.1.3. Pharmacodynamic biomarker model

In order to determine the most relevant model to represent the biomarker dynamics following an APAP dose, biomarker effect vs drug concentration plots were produced for each biomarker. This helped to visualise the relationship between the drug (APAP) and subsequent biomarkers (ALT, HMGB1, full K18, and fragmented K18). Effect versus concentration plots for each biomarker can be seen in Figure 2-3. The results from Figure 2-3 identified a temporal delay between drug and

biomarker accumulation. This hysteresis relationship implies that APAP likely regulates a pre-cursor, (in this case GSH) which then influences each biomarker response, and therefore an indirect PD model was chosen to account for this hysteresis delay [99].

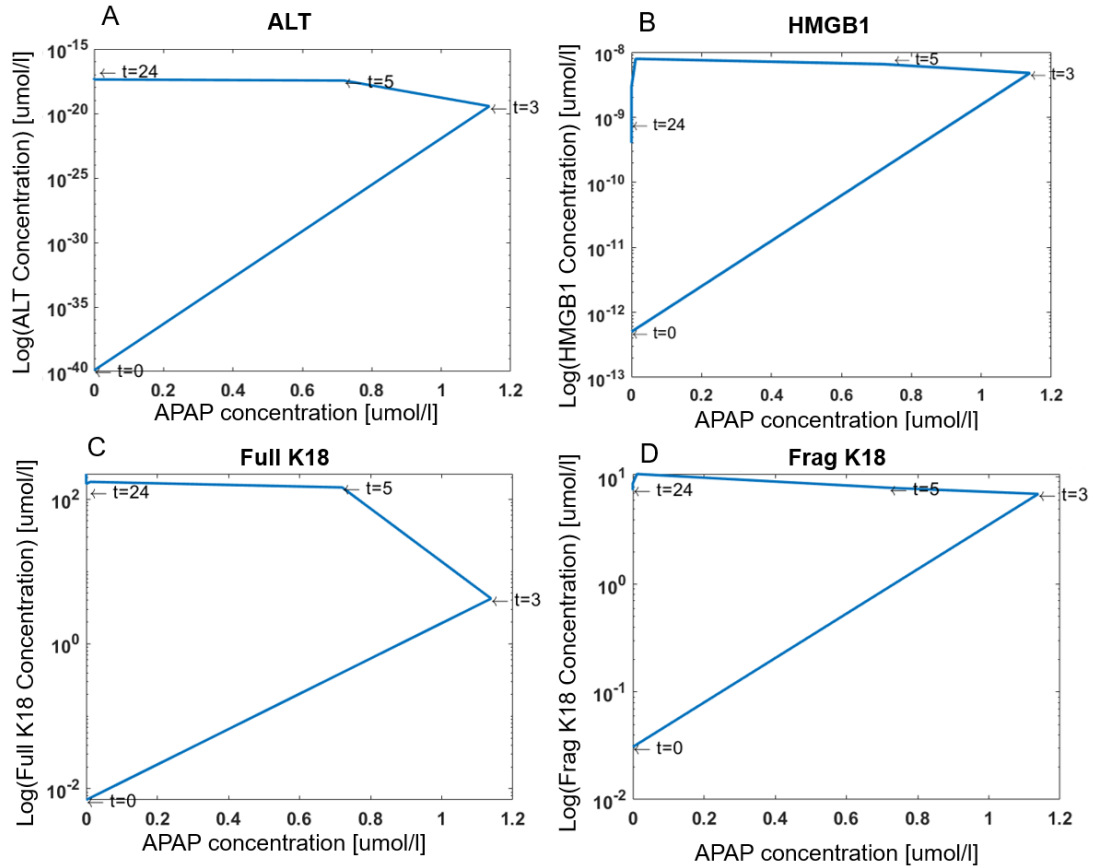


Figure 2-3: Biomarker effect versus concentration plots. A visualisation of the relationship between APAP and each biomarker. (A) shows the relationship between APAP and ALT, (B) shows the relationship between APAP and HMGB1, (C) shows the relationship between APAP and full K18 and (D) shows the relationship between APAP and fragmented K18.

The toxic response to APAP overdose was mathematically described with individual, indirect PD models representing biomarker concentrations ($[r]$ = ALT, HMGB1, K18 and Fragmented K18) over time, as described in equation (2-26),

$$\frac{d[r]}{dt} = r_0 k_{out} \left(\frac{R_{50}^n + gsh_0^n}{R_{50}^n} \right) \left(1 - \frac{[gsh]^n}{R_{50}^n + [gsh]^n} \right) - k_{out}[r], \quad (2-26)$$

where r_0 is the biomarker baseline concentration, k_{out} is the natural decay rate of the biomarker (h^{-1}), R_{50} represents the concentration of (GSH) which causes the biomarker production (response) to be half its maximal value ($\mu\text{mol/l}$), and n is a parameter that reflects the steepness of the biomarker production term [100]. Whilst parameter values r_0 and gsh_0 can be identified directly from the data (for each

biomarker, the baseline value was fixed at the average of all control values provided for that biomarker within the study by Antoine et al. that was used during optimisation [39]). Parameters k_{out} , R_{50} and n were optimised by individually fitting the model output to the data measuring biomarker concentration over time following a 530 mg/kg dose of APAP [39].

All initial conditions for the PKPD model are detailed in Table 2-1.

Model initial conditions	
Variable	Initial condition ($\mu\text{mol/l}$)
[C _c]	0
[C _p]	0
[gsh]	696.9136
[alt]	0.7621
[hmgb1]	0.0005
[fullk18]	0.0146
[fragk18]	0.0642

Table 2-1: Model initial conditions. Initial conditions for each variable within the dynamical system. Biomarker baseline values were fixed at the average values of control measurements from the data used for optimisation [39].

2.2.1.4. Sensitivity analysis

The output of the model was analysed to determine its sensitivity with regard to any local perturbations in parameters. This was carried out using the sensitivity analysis tool in Mathworks' Simbiology [101]. Parameters were varied in ranges such that the lowest bound was one order of magnitude lower than the parameter value, and the highest bound was one order of magnitude higher than the parameter value. Sensitivity outputs were normalised to allow comparison.

The general methodology of this analysis tool can be described with the case of an example.

If we define the model output as $x(t)$ and a model parameter as y , then the time-dependent sensitivity of output x with respect to parameter y is given by the scaled time-dependent derivative,

$$\frac{\partial x/x}{\partial y/y'} \quad (2-27)$$

where the numerator is the sensitivity output and the denominator denotes the sensitivity input to the analysis given by parameter y . We then calculate the corresponding time-integral sensitivity coefficients (S_q), which give an indication of the total sensitivity of the model parameter y on output x over the entire time course of the simulation

$$S_q = \int_{t=0}^{t=t(end)} \left| \frac{\partial x/x}{\partial y/y} \right| dt. \quad (2-28)$$

Further details can be found in [102]; however, the ultimate result is a number for each parameter, representative of how sensitive the output is to perturbations in this parameter. A larger value indicates a higher sensitivity, and therefore only small changes in this parameter would result in large changes in the output, and vice-versa.

2.2.2. Statistical modelling

The PKPD model was used to create an *in silico* population for statistical testing. The aim was to determine whether or not we could predict the initial APAP dose, time elapsed since dose, and probability of liver injury for an *in silico* individual with a suspected overdose based on a single measurement of the panel of biomarkers.

2.2.2.1. *In silico* population simulation

Each computationally-derived dataset used within the analysis consisted of 1,000 *in silico* individuals simulated under different scenarios: given a random APAP dose, and selected from a uniform distribution of range 0-1,000 mg/kg. Once a random dose was chosen, the mechanistic model (with all parameters other than dose fixed at their optimised value) was simulated for this dose. As visually displayed in the mechanistic modelling section of Figure 1-3, this resulted in corresponding concentration-time profiles (APAP, GSH, ALT, HMGB1, full K18, fragmented K18). A random time-point from a uniform range between 0-24 hours was selected. This time point was used as an index to extract a single observation from each biomarker time course. The result was a single combination of biomarker concentrations for an *in silico* individual (where the initial dose and time of biomarker extraction are random). This process was repeated 1000 times to create various scenarios for an *in silico* individual for testing. All biomarker concentrations in the computationally-derived dataset were normalised in the range [0,1] using the min-max normalisation

method [103] to account for varying orders of magnitude. Experimental noise was replicated *in silico* by applying observed *in vivo* standard deviations in biomarker concentrations from an APAP study performed by Antoine et al. [39] (ALT s.d = 11.22, HMGB1 s.d = 0.00097, K18 s.d = 2.39, fragmented K18 s.d = 0.12 $\mu\text{mol/l}$).

2.2.2.2. Visualisation

To examine class structure and separability whilst retaining model variation, we applied Principal Component Analysis (PCA) and constructed two-dimensional scatter plots of the *in silico* derived data projected onto the first two principal components. Additionally, to visually expose class structure, a fast tree-based implementation of the T-distributed Stochastic Neighbour Embedding (T-SNE) method was employed [104].

Principal component analysis (PCA)

PCA is a common statistical technique used to reduce the dimensionality of high dimensional data [105]. This is a useful technique when data has more than 3 dimensions (or variables) and is therefore difficult to visualize. The method works by identifying patterns in high dimensional data, expressing such data in a way that highlights any similarities and differences in the data points, and compressing in relation to the identified patterns. This subsequently reduces the dimensionality of the data with little loss of information [106].

In order to visualise the data in a new subspace with reduced dimensions, the following steps must be followed [107]:

- Remove the label from each variable and assume that each is representative of a dimension. Include each dimension in a matrix, for example, A .
- Compute the mean for every dimension within the dataset (column within matrix A)
- Compute the covariance matrix of the whole dataset (A) using the following formula for calculating the covariance between two variables:

$$\text{cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})(Y_i - \bar{y}), \quad (2-29)$$

where X and Y are the two variables we wish to calculate the covariance between. The total number of observations is represented by n , the current observation for variable X is represented by X_i , the mean of all of the observations in variable X is represented by \bar{x} , the current observation for

variable Y is represented by Y_i and the mean of all of the observations in variable Y is represented by \bar{y} .

The covariance between each variable is calculated to form a covariance matrix for (A) , $(\text{cov}(A))$.

- Compute eigenvectors and corresponding eigenvalues of $\text{cov}(A)$. The eigenvalues of $\text{cov}(A)$ are roots of the characteristic equation,

$$\det(A - \lambda I) = 0 \tag{2-30}$$

For each eigenvalue, λ , we have:

$$(A - \lambda I)\underline{v} = 0, \tag{2-31}$$

where \underline{v} is the eigenvector associated with eigenvalue λ .

- Sort the eigenvectors by decreasing eigenvalues. For a 2-dimensional visualisation, choose the 2 eigenvectors with the largest eigenvalues, these two eigenvectors should then be used to form a new matrix, B .
- Use the transposition of this new matrix (B) to transform each original sample onto a new subspace by simply calculating:

$$C = B'A, \tag{2-32}$$

where B' is the transformation matrix, A is the original dataset matrix and C is the transformed dataset matrix (mapped onto a new subspace).

While the original dataset can have any number of dimensions (or variables), matrix C will have 2 dimensions. These dimensions represent linear combinations of the original variables, with the first vector in matrix C accounting for most of the variance in the original data, and the second vector in matrix accounting for the second most amount of variance in the original data. This explains why the eigenvectors with the highest eigenvalues are chosen to be included in matrix C . The transformed data can now be visualised in two dimensions, plotting one component of matrix C against the other, retaining most of the information (or variation) of the data.

While the original dataset can have any number of dimensions (or variables), matrix C will have 2 dimensions. These dimensions represent linear combinations of the original variables, with the first vector in matrix C accounting for most of the variance in the original data, and the second vector in matrix accounting for the second-most

amount of variance in the original data. This explains why the eigenvectors with the highest eigenvalues are chosen to be included in matrix C . The transformed data can now be visualised in two dimensions, plotting one component of matrix C against the other, retaining most of the information (or variation) of the data.

t-distributed Stochastic Neighbour Embedding (TSNE)

TSNE is a visualisation technique used to reduce high dimensional data, so that it can be visualised in lower dimensions. The method works by minimizing the divergence between two distributions: one that measures pairwise similarities of the input objects (in high-dimensional space) and one that measures pairwise similarities of the corresponding low-dimensional points found from the embedding [104]. Calculating the conditional probability that a high-dimensional point $\Omega 1_i$ would choose the high-dimensional point $\Omega 1_j$ as its neighbour (if neighbours are picked in proportion to their probability density under a Gaussian/Normal distribution), results in determining the similarity of points $\Omega 1_i$ and $\Omega 1_j$ [108]. For nearby points, the conditional probability is relatively high, whereas for points that are widely separated, the conditional probability is almost infinitesimal [109].

Mathematically, the conditional probability, $p_{j|i}$, that the high-dimensional $\Omega 1_i$ would pick the high dimensional $\Omega 1_j$ is represented by:

$$p_{j|i} = \frac{e^{-\left(\frac{\|\Omega 1_i - \Omega 1_j\|^2}{2\sigma_i^2}\right)}}{\sum_{k \neq i} e^{-\left(\frac{\|\Omega 1_i - \Omega 1_k\|^2}{2\sigma_i^2}\right)}}, \quad (2-33)$$

where $\|\dots\|$ is the norm of the element, ensuring a strictly positive resultant vector, and σ_i is the variance of the Guassian that is centred on data point $\Omega 1_i$.

If we wish to define the pairwise similarities in the high-dimensional space, p_{ij} , the obvious representation would be

$$p_{ij} = \frac{e^{-\left(\frac{\|\Omega 1_i - \Omega 1_j\|^2}{2\sigma^2}\right)}}{\sum_{k \neq l} e^{-\left(\frac{\|\Omega 1_k - \Omega 1_l\|^2}{2\sigma^2}\right)}}, \quad (2-34)$$

where $\Omega 1_i, \Omega 1_j, \Omega 1_k, \Omega 1_l$ are all different points within the high-dimensional data set $\Omega 1$, and σ is the total variance of the high-dimensional dataset. This makes the pairwise distance relative to every other pairwise distance present within the dataset.

However, if the high-dimensional data point $\Omega 1_i$ is an outlier (all pairwise distances $\| \Omega 1_i - \Omega 1_j \|^2$ are large for $\Omega 1_i$) then p_{ij} will be extremely small for all values of j , so the location of its low dimensional map-point, $\Omega 2_i$, has very little effect on the cost function, meaning that the position of the map point is not well determined by the positions of the other map points. To resolve this issue, we assume that the joint probabilities p_{ij} in the high-dimensional space, are symmetrized conditional probabilities. This allows each data point to make a significant contribution to the cost function. The joint probabilities in the high dimensional space can therefore be defined as

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N}, \quad (2-35)$$

where p_{ij} are the pairwise similarities in the high-dimensional space, $p_{j|i}$ is the conditional probability that high-dimensional $\Omega 1_i$ would choose high-dimensional $\Omega 1_j$, $p_{i|j}$ is the conditional probability that high-dimensional $\Omega 1_j$ would choose high-dimensional $\Omega 1_i$ and N is the total number of observations. Since we are only interested in pairwise comparisons, we set $p_{i|i} = 0$.

The similarities between the low dimensional counterpart points, $\Omega 2_i$ and $\Omega 2_j$ (which are the original data points $\Omega 1_i$ and $\Omega 1_j$ having been mapped onto a lower dimensional space) can be represented by the conditional probability (or pairwise similarity) q_{ij} , which is normalized by the student t-kernel with a single degree of freedom. This normalization is preferred since the heavy tails of a student t-kernel allow dissimilar high-dimensional input objects $\Omega 1_i$ and $\Omega 1_j$ to be modelled by low dimensional counterparts $\Omega 2_i$ and $\Omega 2_j$ that are too far apart. More space is created during this normalization so that small pairwise distances can be accurately modelled in the low-dimensional embedding [104].

Mathematically, the conditional probability or pairwise similarity between the two points in low-dimensional space, q_{ij} , is represented by

$$q_{ij} = \frac{(1 + \| \Omega 2_i - \Omega 2_j \|^2)^{-1}}{\sum_{k \neq l} (1 + \| \Omega 2_k - \Omega 2_l \|^2)^{-1}}. \quad (2-36)$$

Again, since we are only interested in pairwise comparisons, we set $q_{ii} = 0$.

The locations of the embedding points are then determined by minimizing the Kullback-Liebler divergence between the joint distributions P and Q ($KL(P||Q)$):

$$C = KL(P||Q) = \sum_i^N \sum_j^N p_{ij} \log \left(\frac{p_{ij}}{q_{ij}} \right), \quad (2-37)$$

where C is the cost function to be minimized. This cost function is minimized by the gradient descent method. The gradient of C can be given by

$$\frac{\delta C}{\delta \Omega_{1i}} = 4 \sum_j (p_{ij} - q_{ij})(\Omega_{2i} - \Omega_{2j})(1 + \|\Omega_{2i} - \Omega_{2j}\|^2)^{-1}. \quad (2-38)$$

While the original dataset can have any number of dimensions (or variables), the embedded points will have 2 dimensions, making visualisation much easier, so that clusters may be better identified. These dimensions should represent the data in a way such that there is minimal difference to the structure (in terms of points that are nearby each other) in higher dimensions.

2.2.2.3. Classification

In order to test the predictive potential of biomarker concentrations, critical ranges for predicting time since administration were defined as (0-2], (2-5], (5-10], (10-15] and (15-24] hours. For dose, the ranges were [0-200], [201-400] and [401-1,000] mg/kg, capturing therapeutic, small, and large (overdoses), respectively. Various classification techniques were applied to determine the utility of the biomarkers. Multinomial logistic regression [110] was used as a method that fits well when multiple response categories are available. Since response categories were in an ordinal manner, ordinal multinomial logistic regression was also used. Both linear and quadratic discriminant analysis were employed [111]. A naïve Bayes classifier [112] was also used to predict class probability. Additional model-free classification techniques (k-nearest neighbour (k-nn) and optimal weighted nearest neighbour) were also employed [113,114] to test for robustness. Since classes are not previously defined for model-free classification techniques, observations group together based solely on their similarity, and the aim is to determine whether similar observations automatically group into our desirable classes. The methodology of each classification technique is described below.

Multinomial logistic regression

Multinomial logistic regression was applied in an attempt to predict time since administration and initial dose categories. This method uses more than one predictor variable (in this instance, multiple biomarkers) to predict the probability of falling in the class of a nominal outcome variable (in this instance, multiple categories for time since administration or initial dose). The observed data ($\text{logit}(Y)$), representing either time since administration or initial dose) is assumed to be a linear function of the predictor variables (X_1, X_2, X_3, X_4 and X_5 representing biomarker concentrations APAP, ALT, HMGB1, Full K18 and Fragmented K18, respectively). The equation for the model is written in terms of the logit equation, since we are interested in predicting the probability of falling into a category rather than predicting a continuous value. If all of the predictor variables are significant, and therefore included in the model, the relationship between predictor and outcome variables can be represented using a generalised linear model:

$$\text{logit}(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4. \quad (2-39)$$

The objective of multinomial logistic regression is to optimise the parameters $B_i, i = 1 \dots 4$, so that the model provides the best fit to the observed (either time since administration or initial dose) data. Since we are using a probabilistic model, Ordinary least squares will not suffice as a fitting method and maximum likelihood estimation should be used to optimise for the parameters. For various values of X_1, X_2, X_3 and X_4 , the probability of falling into a particular class can be estimated by applying a logistic transformation to equation (2-39) according to the logistic cumulative distribution function, resulting in the following:

$$\pi = \frac{1}{1 + e^{\text{logit}(Y)}}. \quad (2-40)$$

The key assumptions are:

- *The dependent variable is measured on a nominal level.* This assumption holds since the time since administration classes are sub-categories ranging from 0-24 hours and the initial dose classes are sub-categories ranging from 0-600 mg/kg.
- *One or more of the independent variables are continuous, ordinal or nominal.* The predictor variables in our model are represented by biomarker

concentrations (APAP, ALT, HMGB1, full K18 and fragmented K18), all of which are measured on a continuous scale.

- *Observations must be independent. Additionally, the dependent variable should have mutually exclusive/exhaustive categories (i.e. each observation must belong to an outcome group, and they must only belong to one outcome group, not many).* In the data used for this analysis, each *in silico* mouse is independent of one another, and each *in silico* mouse must fall into category [0-2, 2-5, 5-10, 10-15, 15-24] hours for time since administration or [0-200, 201-400, 401-600] mg/kg for initial dose. For both models, each *in silico* mouse can only be assigned one score.
- *Multi-collinearity should not be present between any predictor variables.* This assumption was tested using the Variance Inflation Factor (VIF) test. The VIF is indicative of the amount of variance an estimated regression coefficient is increased by due to multi-collinearity. Since the VIF is representative of increased variance, the square root of this result would be indicative of the increased standard error. As an example, a variable with VIF 9 has a standard error 3 times larger than it would be if the variable was uncorrelated with the other variables. A recommended reasonable range for the VIF is 1 – 10 [115]. Predictor variables for all models were tested and remained within this range, therefore multi-collinearity does not exist between the predictor variables and the assumption is met.
- *Outliers/highly influential points should not be present in the observed data.* No outliers were detected.
- *Adequate sample size;* a recommended sample size is 10 times the number of predictor variables. There were 5 predictor variables, and therefore 50 observations would be sufficient. The *in silico* derived data consists of 1,000 observations.

Once multinomial logistic regression was deemed a feasible test for the data, variations of two models were investigated; one for predicting initial dose category and one for time since administration category. Biomarkers APAP, ALT, HMGB1, full K18 and fragmented K18 were the independent variables tested for predictivity. A forward stepwise logistic regression analysis was applied to test the panel of predictors. This means that the most significant predictor (biomarker) is entered into the model first, and one additional predictor is included at each stage, but only if the overall significance of the model is improved by its inclusion.

Multinomial logistic regression assumes that the dependent variable is nominal, and makes no explicit use of the fact that categories may be ordered. Ordinal logistic

regression accounts for the fact the dependent variable is categorised on an ordered scale [116].

Discriminant analysis

Linear Discriminant Analysis (LDA) works by transforming high dimensional data into a lower dimensional space by maximizing the ratio of between-class variance and within-class variance [117]. Tharwat et al. [117] detail the three necessary steps for carrying out linear discriminant analysis:

1. Calculate the between-class variance, S_B :

$$S_B = \sum_c (\mu_c - \bar{x}) (\mu_c - \bar{x})^T, \quad (2-41)$$

where μ_c represents the mean of the class (c), and \bar{x} represents the total mean of the data. Therefore, S_B represents the separation distance between the mean of the i th class and total mean (or the between class variance of the i th class). The superscript T represents the transpose operation.

2. Calculate the within-class variance (covariance), S_W :

$$S_W = \sum_c (x_{ic} - \mu_c)(x_{ic} - \mu_c)^T, \quad (2-42)$$

where x_{ij} represents the i th sample in the j th class, and μ_j represents the mean of the class. Therefore, S_W represents the separation distance between the sample of interest, and the mean of the class to which the sample belongs (or the within-class variance of the j th class).

3. Construct the lower-dimensional space

Once the between-class variance and within-class variance have been calculated, the data can then be transformed into a lower dimensional space which shows the maximal variance between classes relative to the variance within classes. This is done by maximizing Fisher's criterion:

$$W_{lda} = \arg \max_W \frac{W^T S_B W}{W^T S_W W}, \quad (2-43)$$

where W is the transformative (or projection) matrix of the LDA technique.

While LDA assumes a common covariance matrix between all of the classes, quadratic discriminant analysis (QDA) assumes that each class has its own covariance matrix [118].

Naïve Bayes

The Naïve Bayes classifier is used to predict the probability that a sample of interest belongs to a particular class [112]. The classifier is based on Bayes' theorem, which computes the probability of an event, based on the probabilities of other events that influence it [119]. This classifier adds strong (naïve) independence assumptions to Bayes' theorem, assuming that the presence (or absence) of a feature of interest within a class is completely unrelated to the presence (or absence) of any other feature [120].

Mathematically, the probability of a class is found using Bayes' theorem

$$p(C|F_1, \dots, F_n) = \frac{p(C)p(F_1, \dots, F_n|C)}{p(F_1, \dots, F_n)}, \quad (2-44)$$

where C is the dependent class variable, which is conditional on several feature variables F_1, \dots, F_n . The values of the features F_i are given, so the denominator of equation (2-44) is effectively constant, and the numerator is equivalent to the joint probability model:

$$p(C, F_1, \dots, F_n). \quad (2-45)$$

Using applications of the definition of conditional probability, (2-45) can be rewritten as

$$p(C)p(F_1|C)p(F_2|C, F_1)p(F_3|C, F_1, F_2) \dots p(F_n|C, F_1, F_2, F_3, \dots, F_{n-1}). \quad (2-46)$$

With naïve Bayes, it is assumed that each feature F_i , is conditionally independent of every other feature F_j , for $j \neq i$. This means that

$$p(F_i|C, F_j) = p(F_i|C) \quad (2-47)$$

for $i \neq j$, and so the joint model (equation (2-47)) can be expressed as

$$p(C, F_1, \dots, F_n) = p(C)p(F_1|C)p(F_2|C)p(F_3|C) \dots p(F_n|C), \quad (2-48)$$

or, equivalently,

$$p(C, F_1, \dots, F_n) = p(C) \prod_{i=1}^n p(F_i|C). \quad (2-49)$$

This means that under the independence assumptions used in naïve Bayes, the conditional distribution over the class variable (C) can be expressed as:

$$p(C|F_1, \dots, F_n) = \frac{1}{Z} p(C) \prod_{i=1}^n p(F_i|C) \quad (2-50)$$

where Z (the evidence, or $p(F_i)$) is a scaling factor dependent only on F_1, \dots, F_n and therefore a constant when the features are known [120].

K-nearest neighbour (KNN)

The KNN method uses the local neighbourhood to make a prediction, and samples are classified based on their level of similarity [121]. The level of similarity between two samples is measured by their distance between one another. The number of samples chosen to be added to the class is defined by the user as K . When a new sample is added, the centre of the class is updated, and it is this point that is used to find the neighbour with the shortest distance. We used Euclidean distance as a distance measure, D :

$$D(x_1, x_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2}, \quad (2-51)$$

where x_1 and x_2 are the points of interest we wish to measure the distance between, x_{1i} is the i th observation we wish to test as the current point, x_{2i} is i th observation we wish to test as the new point and n is the total number of pairwise combinations of points.

Optimal weighted nearest neighbour

The optimal weighted nearest neighbour extends the KNN classification method by assigning neighbours that are particularly close, a higher weight in the decision than neighbours that are further away.

Rather than finding k nearest neighbours, in the weighted classification, $k + 1$ nearest neighbours are found. This $(k + 1)$ th nearest neighbour is used to standardise all other distances and results in the normalised distance function, D_N [122]. This via

$$D_N(x, x_i) = \frac{d(x, x_i)}{d(x, x_{k+1})} \text{ for } i = 1, \dots, k, \quad (2-52)$$

where D_N is the normalised distance, d is the current distance of interest, x_i is the current observation of interest, and x is a new observation. The normalised distance D_N is then transformed into a weight W with the use of a kernel function, $K(D_N)$, with the maximum weighted observation being assigned to the class at each iteration [122].

2.2.2.4. Predicting initial dose and time since administration as continuous variables

Multiple linear regression

To test the use of biomarkers in predicting time since administration and initial dose separately, multiple linear regression analysis was applied. This method takes a very similar approach to the multinomial logistic regression method, details of which can be found in section 2.2.2.3. However, in the linear regression setting, the outcome variable must be continuous rather than nominal/categorical. Since we are interested in predicting a continuous value, the relationship between predictor and outcome variables can be represented using a generalised linear model. The assumptions required for this analysis to be valid are similar to those necessary for multinomial logistic regression. There is one additional assumption required for linear regression analysis:

- *Residuals should be approximately normally distributed.* A histogram and normal P-P plot of the standardized residuals in the models predicting both time since administration and initial dose were inspected, as shown in Figure 2-4.

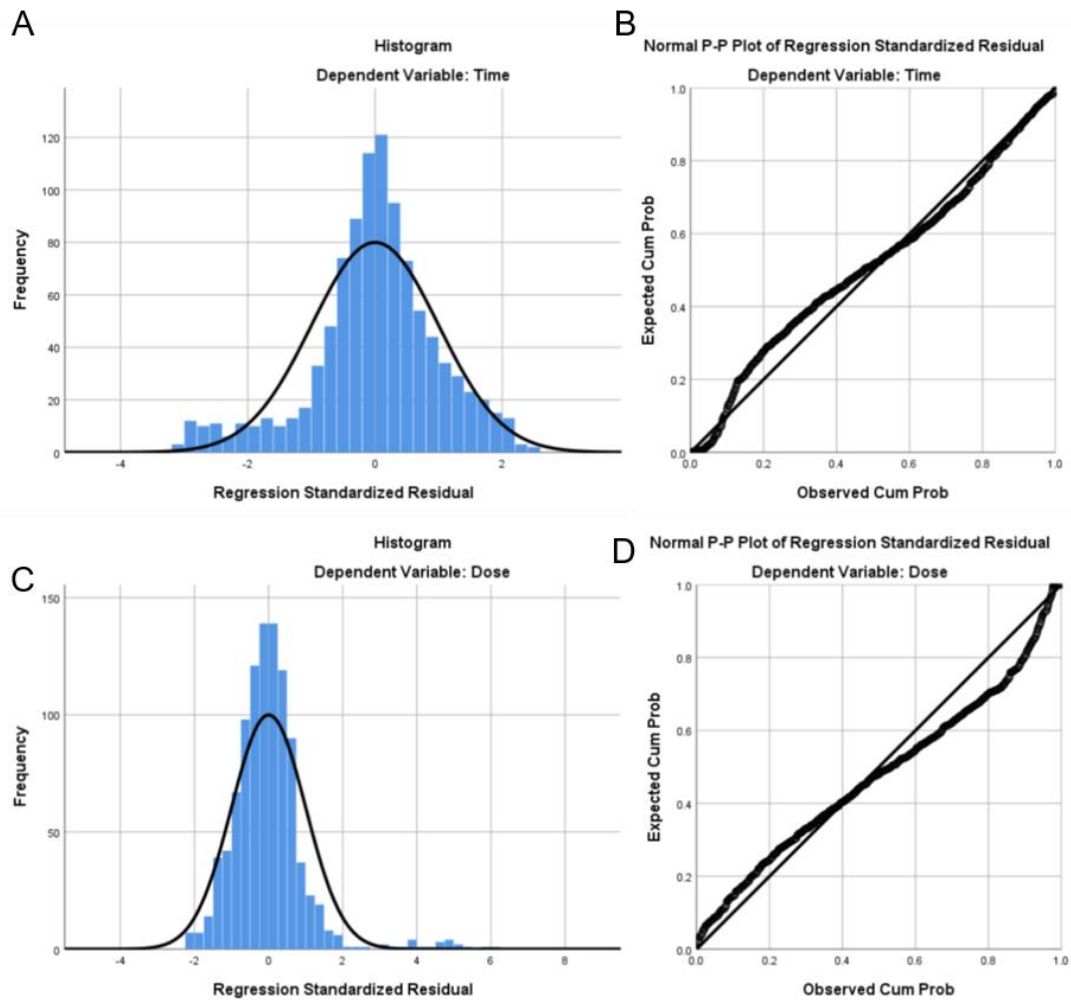


Figure 2-4: Data diagnostic plots. (A,C) Histogram – Visualisation of how the standardized residuals are distributed for the multiple linear regression models predicting time since administration and initial dose respectively. For data to be diagnosed as normally distributed, the mean value should be approximately 0 and the standard deviation should be approximately 1. (B,D) Normal P-P plot of regression standardised residual for the models predicting time since administration and initial dose respectively. The cumulative probability expected from the model is plotted against the observed cumulative probability. For the data to be diagnosed as approximately normally distributed, most results (circles) should lie along the diagonal line.

2.2.2.5. Predicting the probability of liver injury

The biomarker time-course experimental data [39] used to create the PD model in this chapter also provided a corresponding histology score for each mouse from the range [0, 1, 2, 3]. These histology scores were binarised based upon previously published criteria [39]. Within this criteria, histology scores range from [0-4]: 0-normal histology; 1–minimal to mild necrosis; 2–mild to moderate necrosis; 3–moderate to severe necrosis; and 4–severe necrosis. With advice from biological collaborators, we defined any mouse with an observed histology score of 0, 1 or 2 as having a corresponding binary score of 0 (no liver injury) and any mouse with an observed histology score of 3 or above as having a binary score of 1 (liver injury), since our interest was in discriminating between liver injury or no liver injury.

Forward-stepwise binary logistic regression [123] was applied in order to understand the most significant biomarker, or panel of biomarkers for DILI. The most significant biomarkers were then used in combination with PK-PD model simulations to predict the DILI probability [124].

2.3. Results

2.3.1. PKPD model optimisation

Simulations of the PKPD models using the optimised parameters can be seen in Figure 2-5. Visually, the model simulations provide a good fit to the experimental data. Additionally, with an R^2 value of 0.8304 for the PK model, and values of 0.7513, 0.9634, 0.7413, and 0.6526 for the PD models for ALT, HMGB1, K18 and fragmented K18, respectively, the *in silico* model appears to recapitulate the *in vivo* experimental dynamics very well. Optimised parameters for all of the PK-PD models can be found in **Error! Reference source not found.** A key parameter of interest is the R_{50} parameter in the biomarker PD models. This parameter defines a concentration of GSH at which the biomarker reaches half of its maximal production rate (MPR), and can therefore be used to determine the fastest responding biomarker following an APAP dose. For biomarkers ALT, HMGB1, K18 and fragmented K18, the R_{50} values are 227.67, 399.08, 212.87 and 72.09 $\mu\text{mol/L}$ respectively. Therefore, in the model, as GSH depletes from a baseline of 696.91 $\mu\text{mol/L}$ [125] and reaches a concentration 399.08 $\mu\text{mol/L}$ (42.73% depletion), HMGB1 reaches half of its MPR and is therefore considered to be the fastest responding biomarker. GSH must be further depleted to 227.67 $\mu\text{mol/L}$ and 212.87 $\mu\text{mol/L}$ (67-69% depletion), respectively, before biomarkers ALT and K18 reach half of their MPR. Approximately 90% GSH depletion is required for fragmented K18 to reach half of its MPR in the model.

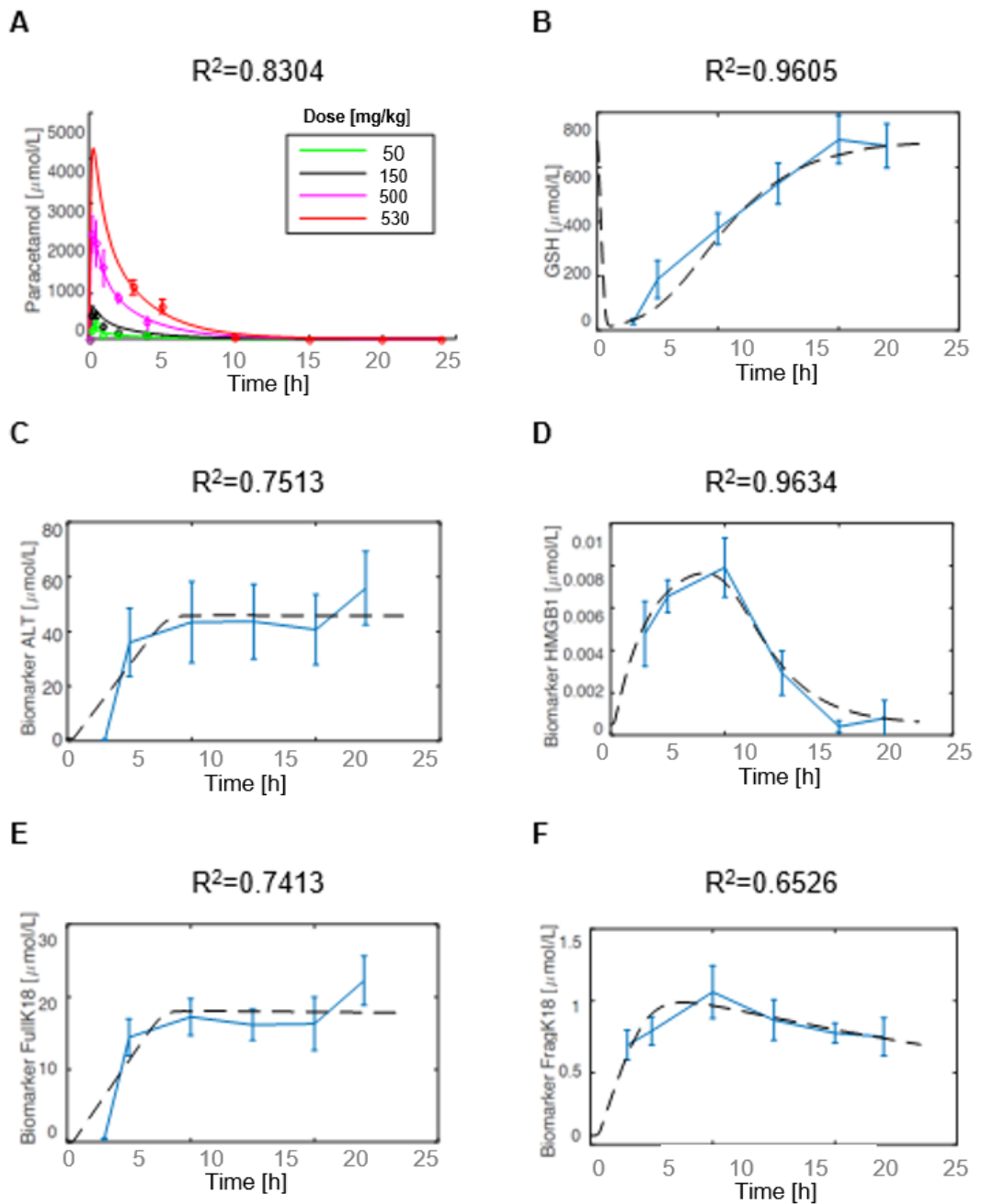


Figure 2-5: *In silico* simulation outputs from the optimised model compared with the experimental data. (A) APAP PK simulations (solid lines) comparable to original data values with green, black, magenta and red representing APAP time-course following a 50, 150, 500, and 530 mg/kg dose respectively. (B) GSH simulations (black dashed lines) comparable to original data (blue). Individual PD simulation (black dashed lines) comparable to data (blue) for biomarkers ALT (b), HMGB1 (C), Full K18 (D), and Fragmented K18 (E).

Table 2-2: Parameters and corresponding standard error estimates for PKPD model.

Model	Parameter	Optimised or Fixed?	Value	Standard Error	Parameter Description	Optimising Dataset
PK	k_a (h ⁻¹)	Optimised	9.05	1.85	Absorption rate from peritoneal cavity	[39,91]
	k_{el} (h ⁻¹)	Optimised	5.2×10^{-1}	4×10^{-2}	Total elimination rate	
	k_{12} (h ⁻¹)	Optimised	4.2×10^{-1}	6×10^{-4}	Distribution rate from central to peripheral compartment	
	V_c (l/kg)	Optimised	2×10^{-2}	1×10^{-3}	Theoretical volume of central compartment	
	V_p (l/kg)	Optimised	1×10^{-2}	Fixed	Theoretical volume of peripheral compartment	
	k_{21} (h ⁻¹)	Optimised	1.01	4.44×10^{-1}	Distribution rate from central to peripheral compartment	
PD- GSH	gsh_0 (μmol/l)	Fixed	696.9136		Baseline value of GSH [125]	[39]
	k_{pr}	Optimised	71.06	2848.6441	Ratio of NAPQI forming other protein adducts relative to detoxification	
	ξ	Optimised	6.8×10^{-1}	1.73×10^{-1}	Proportion of CYP activated APAP that is transformed into NAPQI	
	k_o (h ⁻¹)	Optimised	2.5×10^{-1}	2.3196	Natural decay/background usage rate of GSH	
	k_{el} (h ⁻¹)	Fixed	5.2×10^{-1}		Total elimination rate	
PD-ALT	R_0 (μmol/l)	Fixed	7.62×10^{-1}		Baseline value of ALT	[39]

	n	Optimised	9.26	1.7422	Reflects the steepness of the biomarker production term	
	gsh_0 ($\mu\text{mol/l}$)	Fixed	696.9136		Baseline value of GSH	
	R_{50} ($\mu\text{mol/l}$)	Optimised	227.67	14.0245	Concentration of GSH which causes ALT concentration to be half its maximum value	
	k_{out} (h^{-1})	Optimised	2×10^{-4}	$< 1 \times 10^{-4}$	Natural decay/background usage of ALT	
PD-HMGB1	R_0 ($\mu\text{mol/l}$)	Fixed	5×10^{-4}		Baseline value of HMGB1	[39]
	n	Optimised	4.90	1.57×10^{-1}	Reflects the steepness of the biomarker production term	
	gsh_0 ($\mu\text{mol/l}$)	Fixed	696.9136		Baseline value of GSH	
	R_{50} ($\mu\text{mol/l}$)	Optimised	399.08	24.8957	Concentration of GSH which causes HMGB1 concentration to be half its maximum value	
	k_{out} (h^{-1})	Optimised	3.5×10^{-1}	1.59×10^{-1}	Natural decay/background usage of HMGB1	
PD-Full K18	R_0 ($\mu\text{mol/l}$)	Fixed	1.46×10^{-2}		Baseline value of full K18	[39]
	n	Optimised	10.42	4.3×10^{-3}	Reflects the steepness of the biomarker production term	
	gsh_0 ($\mu\text{mol/l}$)	Fixed	696.9136		Baseline value of GSH	
	R_{50} ($\mu\text{mol/l}$)	Optimised	212.87	1.23×10^{-1}	Concentration of GSH which causes Full K18 concentration	

					to be half its maximum value	
	k_{out} (h^{-1})	Optimised	7×10^{-4}	1.23×10^{-1}	Natural decay/ background usage of Full K18	
PD-Frag-K18	R_0 ($\mu\text{mol/l}$)	Fixed	6.42×10^{-2}		Baseline value of Fragmented K18	[39]
	n	Optimised	2.30	1.46×10^{-2}	Reflects the steepness of the biomarker production term	
	gsh_0 ($\mu\text{mol/l}$)	Fixed	696.9136		Baseline value of GSH	
	R_{50} ($\mu\text{mol/l}$)	Optimised	72.09	4.03×10^{-1}	Concentration of GSH which causes Frag-K18 concentration to be half its maximum value	
	k_{out} (h^{-1})	Optimised	2×10^{-2}	1.17×10^{-2}	Natural decay/backgr ound usage of Fragmented K18	

2.3.2. PKPD model validation

In order to test the accuracy of the *in silico* model, experimental data (from collaborators at University of Liverpool) consisting of 4 different APAP doses [0, 150, 300, 530] mg/kg in mice and their corresponding biomarker concentrations at 5 h was used for validation. The *in silico* model was simulated using identical doses and the resultant biomarker concentrations were extracted at 5 h post-dose and compared to the corresponding *in vivo* mouse data. Both datasets comprised of CD-1 type mice. For GSH, there appeared to be an adaptive response in the validation data at low doses which was not included in the *in silico* model and therefore was not portrayed in the simulation. However, this is a minor discrepancy and given the large dose range in this validation, the *in silico* output matched the validation data very well. Results can be seen in Figure 2-6.

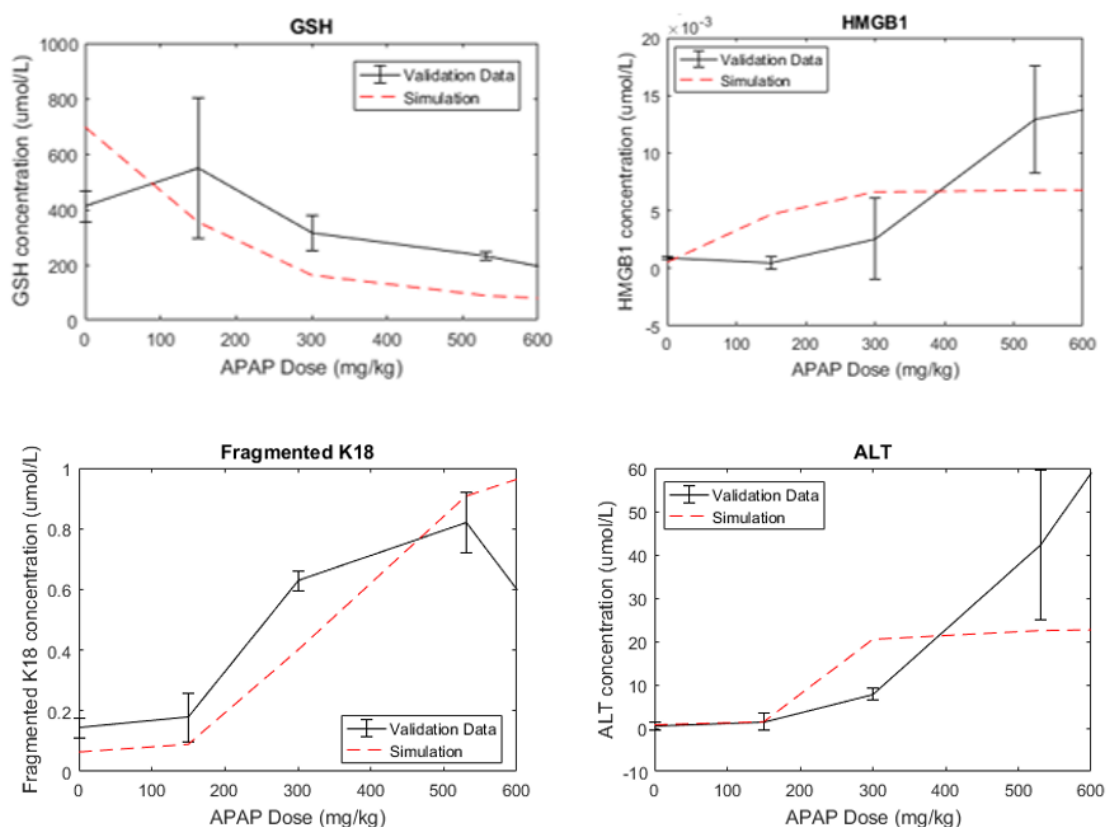


Figure 2-6: *In silico* simulated data versus dose/response validation data used to test the accuracy of the *in silico* model in new scenarios.

2.3.3. Visualisation of *in silico* derived data

The utilisation of biomarkers in class prediction (time since administration and initial dose) can be seen by the projection of the *in silico* derived biomarker data on to the first two principal components (which account for 97% of the variability within the dataset, Figure 2-7 (A)-(B)). Each PCA plot was separated with respect to time since administration and dose amount. Classes were clearly distinguished in both instances, however, the level of class overlap with respect to dose was much lower. Visualising the data with the T-SNE method (Figure 2-7 (C)-(D)) further enhances the PCA visualisation, allowing initial dose to be separated more clearly.

Additionally, the time-since-administration classes were more separable with the T-SNE method, particularly with earlier time ranges.

2.3.4. Classification of *in silico* derived data

The classification results were consistent across the different methodologies (Table 2-3). Should a new observation arise, this framework could predict in which ‘time-since-administration’ and ‘dose’ category it belongs in with 73.7% and 86.5% accuracies respectively.

Classification Method	Time Accuracy	Dose Accuracy
Multinomial Logistic Regression	72.8%	86.5%
Ordinal Multinomial Logistic Regression	57%	85.9%
Naïve Bayes	68.9%	84.4%
Linear Discriminant Analysis	65.7%	86%
Quadratic Discriminant Analysis	73.7%	85.3%
K-nearest neighbour	66.4%	85.9%
Optimal Weighted Nearest Neighbour	67.6%	85.8%

Table 2-3: Classification results for predicting time since administration and initial dose based upon biomarker concentrations.

2.3.5. Predicting initial dose and time since administration as continuous variables

Virtual animals each have two values assigned to them: time since administration, and initial dose. These values can be within the ranges shown in Table 2-4. Using multiple linear regression analysis, the exact time-since-administration value could

be predicted with a residual standard error and accuracy of 3.6 h, whilst the exact dose could be predicted with only an error of 56.81 mg/kg (Table 2-5).

In both models (predicting time, and dose) the APAP concentration was the highest model coefficient, meaning it was this biomarker that had the most influence on the resulting prediction. Whilst Full K18 was a predictive biomarker for both outcomes, fragmented K18 was not a significant biomarker for predicting time since administration. It was however, the second most important biomarker in predicting initial dose amount. HMGB1 was the second most important biomarker in determining the time since APAP dose, however was insignificant for predicting initial dose amount. For both models (predicting time since overdose, and initial dose amount) although the conventional biomarker (ALT) was significant, it had the least impact on the output when compared to the other biomarkers included in each model.

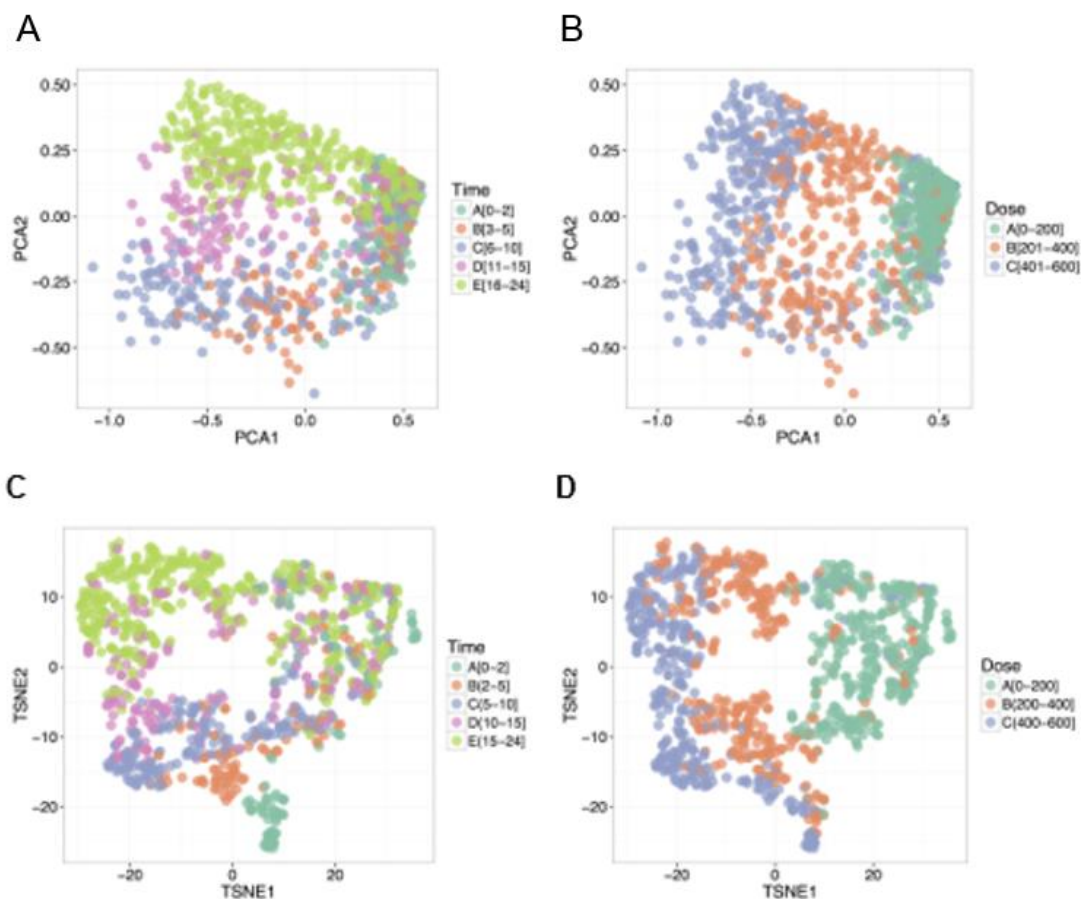


Figure 2-7: Visualisation of time since administration and dose results. For time since administration, dark green represents class [0-2), orange represents [2-5), blue represents [5-10), pink represents [10-15) and pale green represents [15-24) hours. For dose, green represents [0-200], orange represents [201-400] and blue represents [401-600] mg/kg. (A)-(B) 2-dimensional PCA visualisation of *in silico* mouse observations with respect to time since administration and dose respectively. (C)-(D) 2-dimensional TSNE visualisation of *in silico* mouse observations with respect to time since administration and dose respectively.

Predicted value	Potential range
Time since administration	[0-24h]
Initial dose	[0-600mg/kg]

Table 2-4: Ranges of potential time since administration and initial dose values for virtual animals.

	Dependent Variable (coefficient and related error)	
	Time (1)	Dose (2)
APAP Concentration	-18.141*** (1.095)	445.602*** (13.865)
ALT concentration	2.402** (0.988)	94.724*** (12.830)
HMGB1 concentration	-15.928*** (0.636)	
Full K18 concentration	8.964*** (0.837)	241.527*** (12.958)
Fragmented K18 concentration		310.574*** (13.260)
Constant	14.812*** (0.268)	67.068*** (3.193)
Observations	1,000	1,000
Residual Std. Error (df == 994)	3.593	56.805

*Note: *p<0.1; **p<0.05; ***p<0.01*

Table 2-5: Multiple linear regression analysis results. *Summary statistics for models used to predict both time since administration and dose. The first number in each element of the table represents the biomarker coefficient in the regression model, whilst the second number represents the coefficient's corresponding error. For example, -18.141 is the APAP concentration coefficient in the model predicting time since administration, and this coefficient has an error of 1.095. The significance of each biomarker in the model is indicated by the number of asterisks (see note).*

2.3.6. Predicting the probability of liver injury

From the forward-stepwise logistic regression analysis, it appeared that many panels of biomarkers could provide significant predictions of liver injury potential. As expected, the currently used panel (APAP and ALT) was significant (p-value 0.008). APAP alone was also significant at the same level (p-value 0.008). APAP and full K18 combined was also a significant panel (p-value 0.03). However, the model which used HMGB1 concentration alone as a predictor had the highest significance (p-value 0.003).

Upon applying a binary logistic regression analysis in SPSS statistical software, the output was a logit equation based solely on HMGB1 concentration,

$$L_{eq} = 0.635[\text{hmgb1}] - 3.870. \quad (2-53)$$

This equation was substituted into a standard probability equation, resulting in the final liver damage probability equation used in our analysis,

$$\text{Probability of liver damage} = \frac{1}{1 + e^{-L_{eq}}}. \quad (2-54)$$

Figure 2-8 (A)-(F) represents the fold-changes in biomarker concentrations with respect to time following various doses. As expected, for higher doses, APAP and related toxicity biomarker concentrations were significantly increased during the time course, whilst conversely, GSH was significantly decreased, representing depletion of stores. Figure 2-8 (G) shows how the probability of serious liver injury (dependent only on HMGB1 concentration as predicted by the logistic regression model) changes over time for doses between 0-600 mg/kg. A threshold probability of 0.5 (i.e. 50% liver injury likelihood) was used to determine likeliness of DILI. Any observation within the white contour boundary was therefore predicted likely to be a concentration representative of liver injury (i.e. greater than 50% chance). For lower toxic doses, according to the model, HMGB1 concentrations that likely indicate liver injury were most apparent between 5 to 10 h post-dose. As the dose increased, HMGB1 concentrations appeared to remain higher for longer, and the time-frame for probable liver injury increased to approximately 5 to 15 h.

Currently, toxicity is thought to be apparent in mice after a 300 mg/kg dose, shown by the red line in Figure 2-8 (G). Application of our binary logistic regression model

(based solely on HMGB1 concentration) suggests that there is more than 50% chance of liver injury at a 200 mg/kg dose, shown by the white contour in Figure 2-8 (G). The currently used toxic dose (300 mg/kg) coincided with around 90% GSH depletion which can be seen in Figure 2-8 (B). This toxicity threshold and GSH depletion level is a relationship well known in the literature [97]. This toxic level was also the dose at which fragmented K18 elevations began, as shown in Figure 2-8 (F). The toxic dose proposed by the *in silico* model (200 mg/kg) was the dose at which ALT and full K18 began to elevate (Figure 2-8 (C) and Figure 2-8 (E) respectively) and HMGB1 first reached peak concentration (Figure 2-8 (D)).

With the aim to use biomarker concentrations to not only predict initial dose and time since ingestion, but also to get a quantitative prediction for probability of liver injury progression from one combined biomarker sample, the PCA/T-SNE analysis was combined with our proposed framework for predicting the probability of liver injury. Figure 2-9 shows the clear clustering of observations with biomarker concentrations that represent a high probability of liver injury (bottom left-hand corner of the parameter space). High probability cases were distinctively separable from low probability cases.

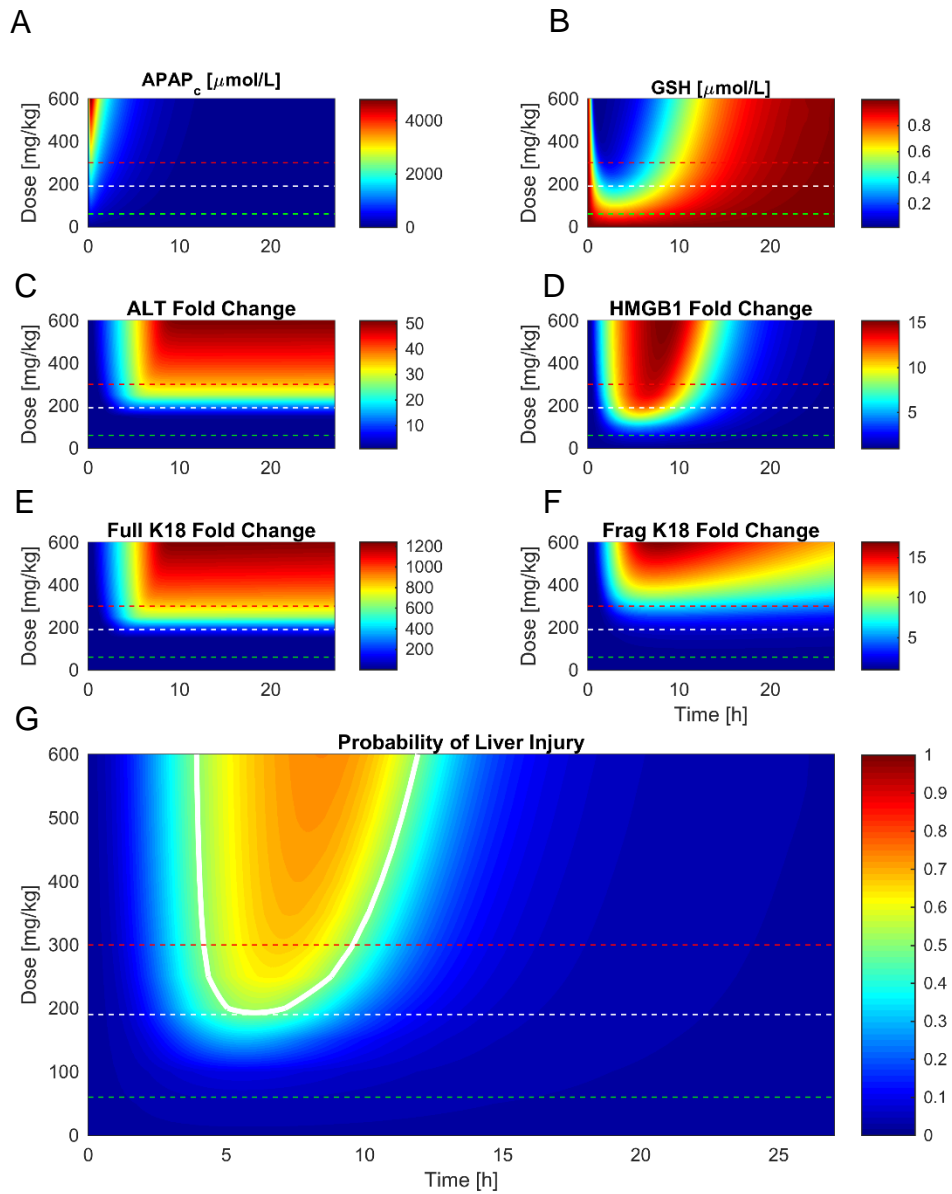


Figure 2-8: (A)-(F) Fold-changes in biomarker concentration relative to their baseline values over time [0-24] hrs for APAP, GSH, ALT, HMGB1, Full K18 and Fragmented K18 respectively, following APAP doses ranging from 0-600 mg/kg. (G) Proposed framework for predicting probability of liver injury dependent upon dose, time and HMGB1 concentration. The white contour indicates the threshold of probability 0.5 of liver injury, the red dashed-line represents currently used APAP dose for toxicity studies in mice, the white dashed-line represents toxic dose proposed by our model, the green dashed-line indicates current known therapeutic dose for mice.

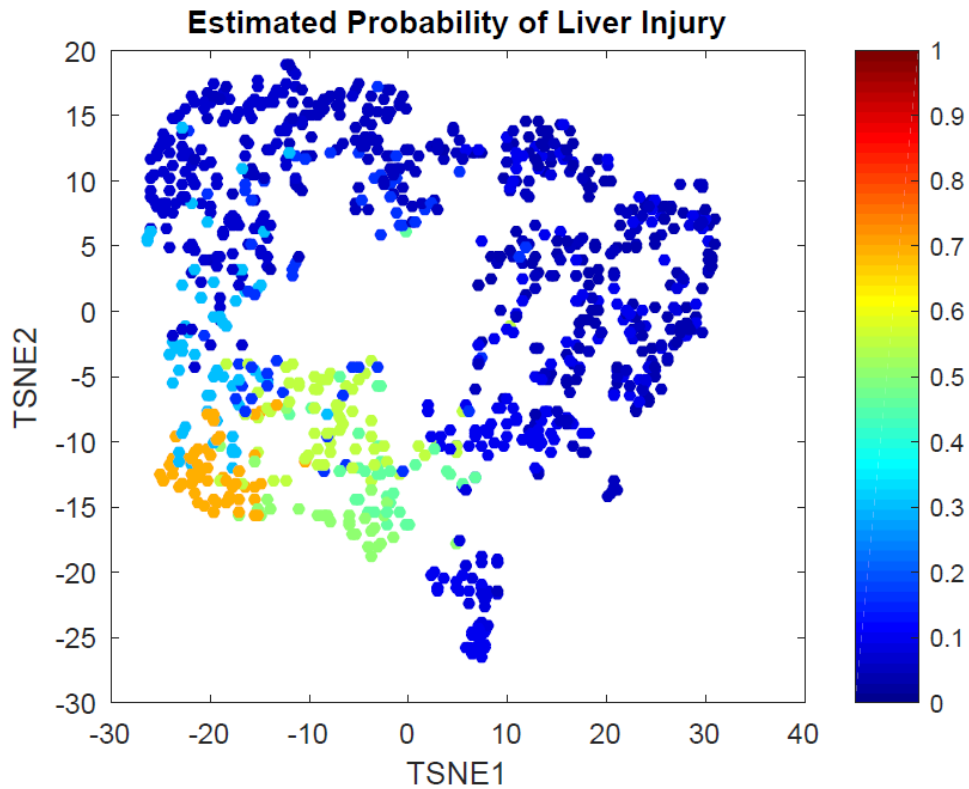


Figure 2-9: Two-dimensional TSNE visualisation of *in silico* mouse observations with respect to estimated probability of liver injury.

2.3.7. Sensitivity analysis

Whilst the results of the pre-clinical framework appear to be positive at this stage, it is important to think forward to translation. By determining sensitive parameters within the pre-clinical model structure, we can focus on ensuring we have confidence in these sensitive parameters at the translation stage. We therefore calculated time-dependent sensitivities of all model variables (APAP, GSH, ALT, HMGB1, Full K18 and Fragmented K18 concentrations) with respect to all model parameters using the sensitivity analysis tool in Mathworks' Simbiology [101]. Details of the sensitivity analysis methodology can be found in Section 2.2.1.4.

The results of the analysis in Figure 2-10 portray the level of sensitivity each of the *in silico* outputs has with regard to perturbations in each of the model parameters. With reference to Figure 2-10, we observe that any parameters deemed to be sensitive by this analysis also have small standard errors for the parameter estimates, providing confidence in the robustness of these predicted values. Although some parameters had greater standard error estimates e.g. k_{pr} ; and k_o , from the GSH model; the sensitivity analysis showed that these were not highly sensitive and therefore impact minimally on the outputs of the model.

Importantly, this sensitivity analysis allowed us to identify the most sensitive parameters which would potentially require most attention if translating this model to a human clinical Pop-PKPD framework. The most sensitive parameter was the baseline level of GSH, gsh_0 (Figure 2-10). In addition to this parameter, the total APAP elimination rate, k_{el} and the theoretical volume of the central compartment, V_c were also highlighted as being sensitive and so they should therefore be given special consideration when translating into the clinical context. For the PD element of the model, the proportion of CYP-activated APAP which is transformed into NAPQI, ξ , was also deemed sensitive. Since this quantity will be dependent on baseline GSH this result is not particularly surprising, and the importance of further investigation into this parameter has already been highlighted. Biomarkers ALT and Full K18 were identified as being sensitive to parameter changes. Biomarkers HMGB1 and Fragmented K18, however, were identified as being relatively less sensitive and more robust to parameter changes.

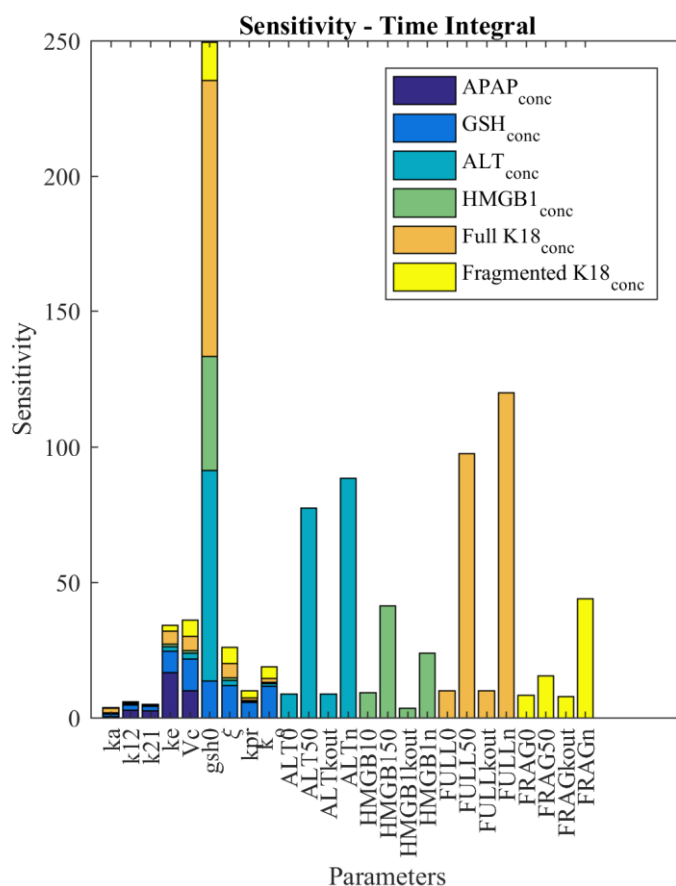


Figure 2-10: Sensitivity analysis of the *in silico* model parameters – visualising the change in model output with regard to perturbations in model parameters.

2.4. Chapter discussion

The current clinical framework for predicting whether or not APAP antidote treatment is necessary is highly dependent upon information provided by the patient such as when the dose was taken and in what quantity. This information is often vague and/or unreliable. Consequently, critically vulnerable patients are often left untreated or, conversely, NAC is unnecessarily administered. Changes in legislation have already led to an estimated increased cost of £8.3 million per year due to overused NAC treatment [24]. Within this chapter, we have used mathematical and statistical analysis to provide a proof-of-concept tool which has the ability of predicting individuals requiring treatment, based on a single measurement of biomarkers with improved sensitivity over those currently used within the clinic.

Using a systems toxicology approach, we have developed an optimised PK-PD model for APAP and corresponding liver injury biomarkers. The model can be used to conduct various investigations within an APAP dosing range of 0-600 mg/kg without the requirement for further *in vivo* testing. Importantly, this provides greater scope for reducing the dependency on animal testing in toxicity and complying with 3Rs principles [89]. A key result from our analysis could be used to refine experiments (i.e. our model proposed that toxicity in mice could be seen following any dose above 200mg/kg rather than the currently used 300mg/kg). That is, not only may experimentalists be dosing mice at amounts higher than necessary, they may also be missing vital information apparent at lower doses.

We aimed to determine the most significant, and fastest responding biomarkers to be used in toxicity predictions. Currently, APAP-induced liver toxicity is thought to occur when GSH depletes by around 80-90% [97]. Our framework suggests that this coincides with elevated fragmented K18 levels. The *in silico* PD model, and its reported R_{50} values, suggest that levels of HMGB1, ALT and Full K18 elevate prior to this depletion level, elevating at 43%, 67% and 69% respectively. As a result, HMGB1 in particular could be considered as an earlier indicator of DILI.

The identification of more accurate predictions of dose timing and amount, informed by biomarker concentration samples, will improve nomogram treatment line accuracy [20]. Predictions for the time since administration were successfully categorised into (0-2], (2-5], (5-10], (10-15], and (15-24] hour ranges based on APAP, ALT, HMGB1 and full K18 concentration values with 73.7% accuracy. Should

this framework be translated to a similar level of efficiency in the human clinical case, this information will have impact regarding the determination of the potential liver injury, with less dependency on patient information. Additionally, an exact time since overdose value was predicted with an accuracy of 3.6 h. Similarly, initial dose was able to be classified into [0-200], [201-400], [401-600] mg/kg categories with 86.5% accuracy and an exact dose predicted with an expected error of ± 56.81 mg/kg. A panel of biomarker measurements could be used in this manner to provide the dose and time information, which will identify a (time-dose) point on the liver injury framework, provided in Figure 2-8, from which one can read off an instantaneous probability of liver injury and how this probability is predicted to change as time progresses. Obtaining dose and time information based on biomarker concentrations and combining this with our proposed liver injury framework shows the utility of these biomarkers in predicting dose amount, time since ingestion and most importantly, the subsequent probability of liver injury.

Although ALT concentration is currently used as a clinical measure to inform potential toxicity, it was found to have the least importance in the regression model for predicting time since administration and initial dose amount as continuous variables. Out of all the biomarkers used in the multiple linear regression analysis, HMGB1 was found to be the most predictive. This analysis suggests therefore, that not only is HMGB1 an earlier indicator of DILI, but it is also an important biomarker in accurately predicting the time elapsed since administration. Furthermore, logistic regression analysis identified HMGB1 as the most significant predictor for liver injury, in line with recent studies defining HMGB1 as a more sensitive DILI predictor [126]. As noted above, the focus of this chapter has been the biomarkers that work well for DILI prediction due to APAP with the aim to provide an improved proof-of-concept framework. HMGB1 is highlighted by our analysis as the most important in DILI prediction.

The pre-clinical framework proposed within this chapter has the potential for substantial clinical impact once translated to human. The analysis was applied to mice due to the relative abundance and quality of data (especially for toxicity cases) and the quantity of relevant biomarker data required to properly characterise such a mathematical and statistical predictive framework. The aim within this chapter, was to investigate the utility of the panel of biomarkers in DILI prediction, and to determine whether or not they warrant investigation in a human clinical study. Prior to translation to human studies, identifying sensitive parameters within the model is

vital. From our sensitivity analysis, the most sensitive parameter was the baseline level of GSH, gsh_0 (Figure 2-10). Since this value was based on experimental estimates in this study [125], we are confident that the results are robust for the mouse situation. However, this parameter is clearly of importance for the predictivity of this model structure so will need to be of particular focus when translating to the clinical case. Additionally, biomarkers ALT and Full K18 were identified as being sensitive to parameter changes, highlighting the need for sufficiently dense data for these biomarkers if they are to be used in a translated human model. Our clinical research is discussed in Chapter 5.

An advantage of the framework proposed in this chapter is that the same biomarkers can be measured in both humans and animals via the same methodologies. Moreover, the model hepatotoxin we have employed, APAP, is directly comparable between human and mice with respect to both the toxicity mechanism and the action of the antidote, NAC. Taking these points into consideration, in its current form, our framework is highly predictive and provides promise for clinical use in discriminating time since administration, initial dose amount and subsequent probability of liver injury. This would be a significant application and could instruct the determination of NAC intervention in patients suspected of APAP overdose. In the next chapter, we carry out an uncertainty analysis on our model framework to study the robustness of its predictions and quantify any parameter uncertainties.

Chapter 3: A pre-clinical systems toxicology framework – Improving the confidence, biological relevance and clinical application potential

The research defined in this chapter has been published in the Journal of Computational Toxicology.

Mason CL, Leedale J, Tasoulis S, Jarman I, Webb SD. A systems toxicology paracetamol overdose framework: accounting for high-risk individuals. 2019;12. Available from: <https://doi.org/10.1016/j.comtox.2019.100103>

3.1. Background

In Chapter 2, we describe a modelling framework that uses a single sample approach to biomarker quantification (rather than the multiple sample approach currently required within the clinic) with the aim of predicting the probability of liver injury. The research in Chapter 2 identifies which biomarkers are the most significant in the prediction of liver injury, and provides a framework which, if extended to the human case, would be easily interpretable for clinicians, increasing the effectiveness of APAP overdose treatment.

However, something which must be crucially noted, is that the utility of a mathematical model is limited by assumptions which are often necessary for simplification. This is due to mathematical models being an abstract representation of the true biology. Consequently, there are often multiple aspects of a model that potentially contain uncertainty. Not all states of a dynamical model can be directly measured experimentally, and conversely, not all experimental data may be useful for model calibration, since the data itself may contain errors not accounted for by the model. These limitations can raise skepticism around the employment of model predictions. While it would be unrealistic to attempt to completely eradicate every level of error, it is crucial that any parameter uncertainties should be assessed, reported and minimised in order for model predictions to be truly useful [46]. There are many existing and developing techniques to quantify uncertainty, and the chosen method often depends on the aims of the model. For example, if there is a small level of uncertainty in parameters and the model of interest is relatively simple, a local sensitivity analysis may be sufficient (perturbing parameters individually and monitoring changes in output) [127]. For larger models, where parameters are reasonably known within some sort of range, Monte-Carlo (MC) probabilistic methods are usually the first method of choice, particularly in the health and environmental sectors [128]. The most highly used probabilistic method for stochastic inverse problems (attempting to estimate original parameters from noisy data) is Bayesian inference [129]. If both the structure of the model, and the data used in optimisation require testing for uncertainty, identifiability analysis can be employed to determine whether model parameters can be uniquely identified based upon the structure of both the model and data used [130].

In this chapter, we aim to quantify any uncertainties in the model defined in Chapter 2 with the following ideas in mind: APAP is predominantly metabolised in the liver

via glucuronidation and sulfation pathways, with a small fraction being oxidised into the toxic metabolite, NAPQI; and detoxification of NAPQI occurs via conjugation with hepatic stores of glutathione (GSH) [28]. Therefore, although initial dose and time since ingestion are known to be the most important indicators of overdose severity level, additional factors affecting an individual's ability to synthesise or maintain sufficiently high levels of GSH should also be considered [10]. Such factors may include age, pre-existing liver disease, concurrent use of alcohol and/or other liver-metabolised medications, genetic predispositions and acuity/chronicity of APAP use [5]. In 1981 the FDA recommended the original nomogram line should be reduced by 25% to account for potentially high-risk patients/measurement errors [9].

The model defined in Chapter 2 optimises parameters by fitting to fed-mouse data and therefore is analogous to applications relating to individuals with unimpaired clearance capacity. Through the application of an identifiability analysis technique, in this chapter we identify areas within the model structure that require improvement and use this knowledge to make the structure more relevant to the current APAP toxicity clinical environment. There are many other *in silico* models that focus on describing and understanding APAP-induced toxicity (as described in Chapter 1). However, there is currently only one *in silico* APAP model which takes into consideration individuals that may have depleted GSH stores. Navid et al. (2013) define a multi-compartmental PBPK model of APAP metabolism, with the aim of understanding how nutritional deficiencies and certain lifestyle choices (such as alcohol consumption) may affect GSH regeneration [81]. Whilst they focus on how various factors may affect GSH synthesis, their model does not extend to how depleted GSH stores can subsequently alter toxic effects. Within the mathematical model defined in this chapter, we account for the fact that certain factors may affect GSH regeneration, but we also model how toxic effects may differ in individuals with both normal/suppressed GSH restoration ability.

This chapter provides an extension to our modelling approaches defined in Chapter 2, this time quantifying the effects of various factors impacting upon GSH availability. The mathematical model is refined using techniques from uncertainty analysis to account for mechanistic changes indicative of suppressed GSH capacity, and optimised against additional data to improve its scope and predictive potential. Crucially, the availability of GSH is known to be heavily dependent on the nutritional level of a patient. When optimising against additional data, we therefore choose to

include fasted mouse data to reflect this. The model defined in this chapter provides predictions of initial dose, time since ingestion and probability of liver injury for both healthy and high-risk populations, with a much greater level of confidence than the predictions provided in Chapter 2.

Chapter aims

- Carry out identifiability analysis on the model framework defined in Chapter 2 to determine areas of uncertainty.
- Refine the model structure dependent on results from identifiability analysis, and optimise against additional data where necessary.
- Provide an improved model framework which is more biologically relevant, and which cannot only make predictions for the general individual, but also for individuals who are already deemed high risk.

3.2. Methods

3.2.1. Experimental data description

The mathematical model in this chapter was parameterised against multiple experimental datasets. For the APAP PK element of the model, four datasets from two separate published studies [25, 26] recording APAP concentration over time in mice following intraperitoneal administration of 50, 150, 500 and 530 mg/kg doses were used for optimisation of both the framework defined in Chapter 2 and the framework defined within this chapter. For the biomarker PD element of the model, in Chapter 2, one experimental dataset was used during parameter optimisation [39] which recorded biomarkers (GSH, ALT, HMGB1, full and fragmented K18) over time following a 530 mg/kg APAP dose. In this chapter, this dataset was also used for optimisation of the PD model parameters, but with the addition of two other datasets from two separate studies by Antoine et al. and Mason et al. [35,131]. The first, [35], provided biomarker concentrations GSH, ALT, HMGB1, full and fragmented K18 at 5 and 24 hours for both fed and fasted mice following a 530 mg/kg APAP dose. The second, [131], provided dose-response data for mouse biomarker concentrations GSH, ALT, HMGB1 and fragmented K18 at 5 hours following APAP doses [0,150,300,530,750,1,000] mg/kg.

3.2.2. Identifiability analysis

Identifiability analysis was performed to visualise changes in the sum of squared errors (SSE) following parameter perturbations and, subsequently, to determine the identifiability of each parameter in the model. Since expertise around this methodology was not available at my home institution, I used funding from the LJMU International Mobility Award, obtained in 2017, to fund an uncertainty analysis training placement at the Freiburg Institute for Advanced Studies (FRIAS). I spent two weeks with Professor Jens Timmer's research group, who introduced me to the methodology and applications of identifiability analysis; particularly, the profile likelihood approach they had developed [130]. Within this chapter, we apply a method of identifiability analysis similar to the profile likelihood approach defined by the (FRIAS) research group. Parameter estimates are either deemed identifiable, practically unidentifiable, or structurally unidentifiable [132]. During this analysis, each individual parameter was tested separately for identifiability. This "test parameter" was varied by 20% intervals (from -50% to +200% of its original value). In each iteration, the modified test parameter was fixed, while all the other parameters in the model were varied in two ways: either fixed at the optimum values (as found from previous multi-start optimisation); or randomly sampled from a Latin hypercube (bounds for sampling can be found in section 3.2.4 of this chapter). For each test parameter iteration, the parameter set corresponding to the lowest SSE value was then identified. During both sensitivity and identifiability analyses, model parameters are perturbed and subsequent changes in model output are studied. However, parameters are not re-optimised during sensitivity analysis. Identifiability analysis seeks to determine whether distinct model parameterisations could provide the same model solution.

Examples of each resultant case for a parameter (identifiable, practically unidentifiable, structurally unidentifiable) can be visualised in Figure 3-1. Similar to the profile likelihood approach, an identifiable parameter is defined as a parameter which, when perturbed from its initial (optimal) value (both positively and negatively), results in an increased SSE and therefore predicts a greater error between the model output and the data. If the SSE increases on only one side (i.e. in the positive or negative direction) of the test parameter, and remains relatively unchanged on the other side, this parameter is defined as practically unidentifiable; that is, either increasing or decreasing the test parameter value causes an increased error between the model output and the data. However, since the error between model

output and data reduces as we head towards the test parameter, but then remains relatively unchanged for further perturbations in that direction, we cannot be confident that the parameter is uniquely optimal, since there are multiple values that provide the same approximate SSE. Since the optimum path does change in one direction, however, we have confidence that we are capturing some of the structure, and often these kind of ‘practical unidentifiabilities’ can be resolved by including more calibration data [130]. If the SSE value does not change either side (i.e. in the positive or negative direction) of the original test parameter value, this parameter is known as structurally unidentifiable. This means that parameter optimisation via data-fitting is relatively insensitive to the choice of this parameter; the parameter cannot be uniquely determined and therefore even if removed entirely, values of other parameters could be suitably adjusted to compensate for the change in the model structure.

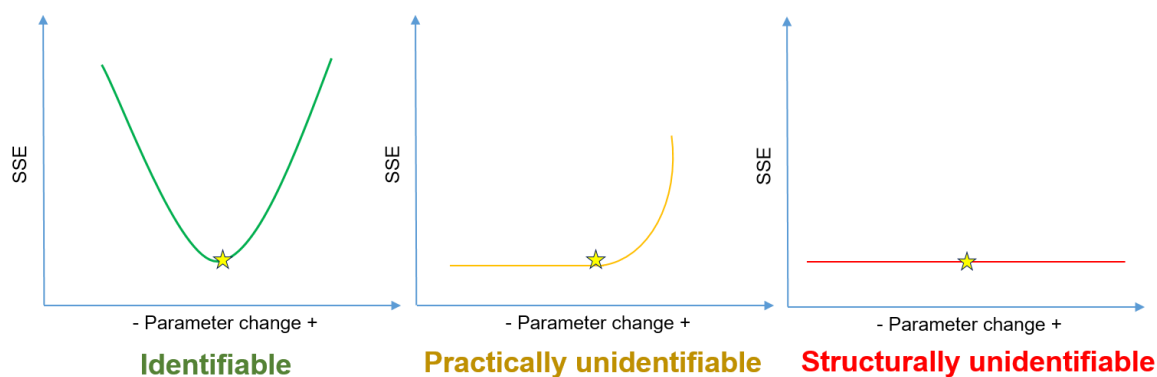


Figure 3-1: Identifiability definitions in relation to parameter perturbations and corresponding minimum SSE profile. As a parameter is perturbed (in both the positive and negative direction), if the minimum SSE increases, this parameter will be deemed identifiable. If the minimum SSE only increases in response to parameter changes in one direction (either negative or positive) this parameter is deemed practically unidentifiable. If a parameter is changed in both the positive and negative direction, and there is no corresponding change in the minimum SSE, this parameter is deemed structurally unidentifiable.

3.2.3. Model refinement

3.2.3.1. APAP pharmacokinetic model

The PK model structure remained unchanged from the model defined in Chapter 2 (see equations (2-1) and (2-2)), but is summarised below for completeness. Two ordinary differential equations (ODEs) were used to represent changes in APAP concentration within two PK compartments (central and peripheral) in the following system,

$$\frac{d[C_c]}{dt} = \frac{k_a D_0 e^{-k_a t}}{V_c} + k_{21}[C_p] \frac{V_p}{V_c} - k_{12}[C_c] - k_{el}[C_c], \quad (3-1)$$

$$\frac{d[C_p]}{dt} = k_{12}[C_c] \frac{V_c}{V_p} - k_{21}[C_p], \quad (3-2)$$

where, as in Chapter 2, $[C_c]$ is the central compartment concentration ($\mu\text{mol/l}$) of APAP, $[C_p]$ is the peripheral compartment concentration ($\mu\text{mol/l}$) of APAP, and k_a represents the absorption rate (h^{-1}) from where APAP is administered (the peritoneal cavity in this case). The initial dose ($\mu\text{mol/kg}$) is given by D_0 , k_{21} represents the transfer rate (h^{-1}) from peripheral to central compartment, k_{12} represents the transfer rate (h^{-1}) from central to peripheral compartment, V_p is the theoretical volume (l/kg) of the peripheral compartment, V_c is the theoretical volume (l/kg) of the central compartment, k_{el} represents the overall elimination rate (summation of excretion and metabolism processes) (h^{-1}), and t represents the time variable (h).

3.2.3.2. Pharmacodynamic models

In Chapter 2, the pharmacodynamic (PD) element of the model was parameterised using a dataset consisting of GSH and biomarker (ALT, HMGB1, full and fragmented K18) time-course concentrations following a 530 mg/kg intraperitoneal APAP dose [39]. In this chapter, we extended this optimisation to also include dose-response data consisting of plasma biomarker concentrations at 5 hours following APAP doses ranging from 0-1,000 mg/kg [131] and a dataset consisting of biomarker concentrations at 5 and 24 hours for both fed and fasted mice following a 530 mg/kg APAP dose [35]. This extension was necessary in order to account for differing mechanisms of cell death, i.e. apoptosis versus necrosis, and also to account for an increased dosing range more representative of the clinical environment.

Glutathione depletion dynamics

As in Chapter 2, paracetamol toxicity biomarker dynamics were assumed to be directly dependent on GSH depletion; i.e., during APAP overdose cases when GSH pools deplete, NAPQI accumulates, potentially leading to liver toxicity and associated biomarker release. The GSH model component remained identical to the model previously defined in Chapter 2 (see equation (2-25)), namely:

$$\frac{d[\text{gsh}]}{dt} = k_o(\text{gsh}_0 - [\text{gsh}]) - \frac{\xi k_{el}[C_c][\text{gsh}]}{[\text{gsh}] + k_{pr}}, \quad (3-3)$$

where, as in Chapter 2, k_o is the basal removal rate (h^{-1}) of GSH (including background usage), gsh_0 is the baseline concentration ($\mu\text{mol/l}$) of GSH in the APAP-free steady state, ξ is the proportion of eliminated APAP that is transformed into NAPQI, k_{el} is the APAP elimination rate, k_{pr} can be thought of as the proportion of NAPQI forming other protein adducts, relative to being detoxified by GSH, although the parameter represents the GSH concentration at which GSH degradation due to binding with NAPQI has reached 50% of its maximal value ($\mu\text{mol/l}$). The term $[\text{gsh}]$ represents the concentration ($\mu\text{mol/l}$) of GSH. Further details on the derivation of this GSH model can be found in Chapter 2. For the fasted case, basal GSH dynamics are modified to simulate a delay in GSH repletion due to depleted co-factors stemming from the reduced food intake. These slower dynamics are incorporated by rescaling k_o by an additional parameter, $\delta < 1$ (so that k_o becomes $\delta k_o < k_o$). That is, the GSH model for the fasted case is now defined below:

$$\frac{d[\text{gsh}_f]}{dt} = \delta k_o (gsh_0 - [\text{gsh}_f]) - \frac{\xi k_{el} [C_c] [\text{gsh}_f]}{[\text{gsh}_f] + k_{pr}}. \quad (3-4)$$

ALT and HMGB1 dynamics

Following an APAP overdose, the toxic response of biomarkers ALT and HMGB1 were mathematically modelled in the same way as the framework detailed in Chapter 2, namely:

$$\frac{d[r]}{dt} = r_0 k_{out} \left(\frac{R_{50}^n + gsh_0^n}{R_{50}^n} \right) \left(1 - \frac{g_{max} [\text{gsh}]^n}{R_{50}^n + [\text{gsh}]^n} \right) - k_{out} [r], \quad (3-5)$$

where, as in Chapter 2, $[r]$ is the biomarker concentration ($\mu\text{mol/l}$), r_0 is the respective biomarker baseline ($\mu\text{mol/l}$), k_{out} is the natural decay rate (h^{-1}) of the biomarker, R_{50} represents the concentration ($\mu\text{mol/l}$) of (GSH) at which the biomarker response to GSH is half its maximal rate, and n is a Hill coefficient indicating the steepness of the biomarker response [100]. Parameter g_{max} is a new addition to this model and represents the maximal regulatory effect of GSH on biomarker production. In the fed case, we set $g_{max} = 1$ such that GSH depletion is solely responsible for augmented biomarker production and therefore for the APAP-free steady state, the biomarker is only produced at low, basal levels maintaining the background steady state value, r_0 . For the fed case, this element of the model

was therefore identical to that employed in Chapter 2. However, experimental observations showed that background biomarker levels were higher for fasted animals. We therefore allowed $g_{max} < 1$ in the fasted case to account for these higher background values.

K18 dynamics

Keratin-18 (K18) is an intermediate filament protein expressed in abundant levels in hepatocytes [39]. This protein undergoes caspase-mediated cleavage during apoptosis, resulting in the release of fragmented K18 upon cell death [133]. This feature makes K18 a useful biomarker to distinguish between apoptosis and necrosis as the presence of full (as opposed to fragmented) K18 instead suggests the occurrence of necrosis [134]. Full and fragmented K18 were therefore modelled in this chapter as necrotic and apoptotic forms of the same single biomarker, K18. However, we could not find sufficient data for K18 to parameterise a model of the form shown in equation (3-5) for both full and fragmented K18. We were limited to data for full K18 following a 530mg/kg APAP dose; although we had access to dose-response data spanning a wide range of doses for biomarkers ALT, HMGB1 and fragmented K18, this was unavailable for full K18 at the time of investigation. We therefore adopted a simple form of dynamics for K18 using piecewise linear representations of the non-linear terms in (3-5), as illustrated in Figure 3-2. That is, the dynamics of the necrotic marker, full K18, were modelled in the following way,

$$\frac{d[\text{k18}]}{dt} = r_0^{18}k_{out}^{18} + k_{max}k_{in}^{18}H(gsh_{\theta_2} - [\text{gsh}]) - k_{out}^{18}[\text{k18}], \quad (3-6)$$

where r_0^{18} is the baseline concentration ($\mu\text{mol/l}$) of full K18; the natural decay rate of the biomarker is represented by k_{out}^{18} ; the production rate of the biomarker is represented by k_{in}^{18} ; the GSH threshold below which additional K18 production is induced due to cell death is represented by gsh_{θ_2} ; the concentration of full K18 ($\mu\text{mol/l}$) is represented by $[\text{k18}]$; and k_{max} is a measure of the production capability of full K18 ($0 < k_{max} < 1$); since there is a finite quantity of cells, there is a maximum amount of biomarker that can be produced. In the fed case, we take $k_{max} = 1$. Since fasting in mice causes extensive cell loss at early time points [35], the amount of biomarker able to be produced from a smaller amount of cells will therefore be smaller. To account for this, in the fasted case, we allowed $k_{max} < 1$. The Heaviside function $H(x) = 1$ when $x \geq 0$ whereas $H(x) = 0$ when $x < 0$.

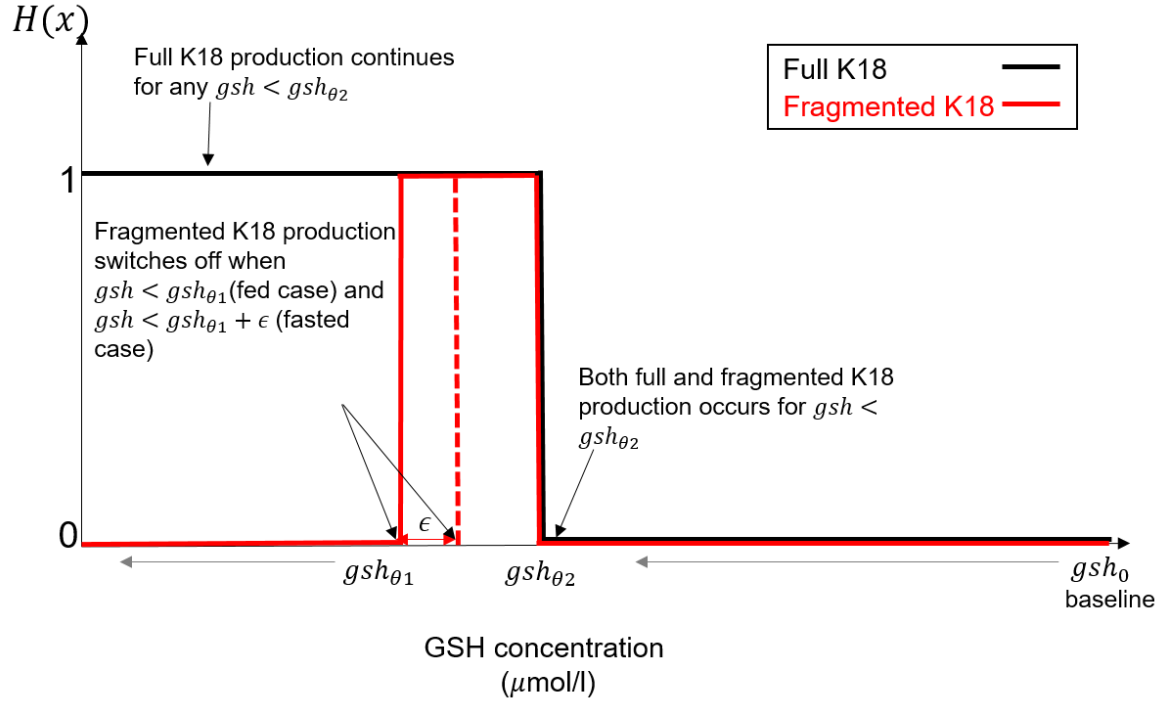


Figure 3-2: Relationship between full and fragmented K18 biomarker production and GSH concentration. Black lines represent the full K18 relationship, and red lines represent the fragmented K18 relationship (solid lines represent the fed case and dashed lines represent the fasted case). As GSH depletes from baseline, gsh_0 , and reaches a certain threshold, gsh_{θ_2} , production of full and fragmented K18 begins. Full K18 production continues for any GSH concentration below this threshold, however, fragmented K18 production ceases beyond a GSH concentration of gsh_{θ_1} in the fed case and $gsh_{\theta_1} + \epsilon$ in the fasted case.

Similarly, the dynamics of fragmented K18 were modelled in the following way,

$$\frac{d[\text{fk18}]}{dt} = r_0^{f18} k_{out}^{f18} + k_{in}^{f18} \left(H([\text{gsh}] - gsh_{\theta_1} - \epsilon) - H([\text{gsh}] - gsh_{\theta_2}) \right) - k_{out}^{f18} [\text{fk18}], \quad (3-7)$$

Where $[\text{fk18}]$ is the concentration ($\mu\text{mol/l}$) of fragmented K18, r_0^{f18} is the baseline concentration ($\mu\text{mol/l}$) of fragmented K18; the natural decay rate (h^{-1}) of the biomarker is represented by k_{out}^{f18} , the production rate of the biomarker (h^{-1}) is represented by k_{in}^{f18} is the production rate, the GSH threshold below which augmented production of fragmented K18 production is initiated is represented by gsh_{θ_2} , and the GSH threshold below which augmented production ceases due to a switch from apoptosis to necrosis is represented by gsh_{θ_1} . Since there is this switch from apoptosis to necrosis, we do not need a parameter within this model to represent the maximum production capability of fragmented K18 (fk_{max}); as extensive cell loss occurs, the mode of cell death switches from apoptosis to necrosis, and therefore production of apoptosis will completely cease. For the fed

case, production begins at the GSH concentration represented by gsh_{θ_2} ($\mu\text{mol/l}$) and ceases at the GSH concentration represented by gsh_{θ_1} ($\mu\text{mol/l}$). For the fasted case however, necrosis will be more apparent than apoptosis, and less GSH depletion will be required before apoptosis is no longer sustainable. To account for this, ε represents the change in GSH threshold denoting the switch to necrosis, i.e., $\varepsilon = 0$ in the fed case and $\varepsilon > 0$ in the fasted case.

All initial conditions for this PKPD model are detailed in Table 3-1. The fed and fasted models will begin at different initial conditions. Lack of nutrition is known to have effects on GSH synthesis levels [81], therefore, even with no APAP present, GSH concentration is likely to be depleted at time 0. Since the biomarkers within our model are directly dependent on GSH depletion, this will in turn result in higher baseline levels for the biomarkers in the fasted case when compared to the fed case. Although our fed and fasted biomarker models begin at different concentrations, since food was re-introduced to the mice used in the experiments for optimisation, the biomarker profiles will eventually return to the fed steady-state.

Model initial conditions		
Variable	Initial condition (Fed case - $\mu\text{mol/l}$)	Initial condition (Fasted case - $\mu\text{mol/l}$)
[C _c]	0	0
[C _p]	0	0
[gsh]	559.47497	374.0909
[alt]	0.7626	0.9528
[hmgb1]	0.0005	0.0007
[fullk18]	0.0088	0.0113
[fragk18]	0.0977	0.1634

Table 3-1: Model initial conditions. *Initial conditions for each variable within the dynamical system. For the fed case, the initial condition for GSH is an optimised value, found by fitting the fed GSH model to 3 different datasets for fed mice dosed with APAP [35,39,131]. The initial conditions for ALT, HMGB1, full K18 and fragmented K18 are fixed as the average of the control values for each respective biomarker in the fed case from each study [35,39,131]. For the fasted case, the initial conditions for GSH ALT, HMGB1, full K18 and fragmented K18 are fixed as the average of the control values for each respective biomarker in the fasted case from Antoine et al.'s study [35].*

3.2.4. Parameter optimisation

A multi-start technique was applied during parameter optimisation in an attempt to find the globally optimal parameter set following refinement of the original model. All dynamical models were optimised by minimising the sum of squared errors (SSE) between model simulated output and experimental data (Fminsearch, Matlab [96]). As described in Chapter 2, the Matlab minimisation function uses a Nelder-Mead search algorithm to iteratively search the parameter space until a local minimum is found [95]. Although the algorithm is a local minimiser, Latin hypercube sampling (LHS) was used to generate 1,000 different initial estimates for each parameter. Ranges for the bounds of each parameter are defined in Table 3-2.

3.2.5. Virtual population simulation

Model predictions were made for three virtually simulated populations: healthy, high-risk, and a mixture of healthy and high-risk individuals. Healthy populations were based on biomarker concentrations simulated with parameter values derived from the fed mouse data. High-risk populations were based on biomarker concentrations simulated with parameter values derived from fasted mouse data. The mixed population was based on a weighting of the biomarker concentrations simulated with parameter values derived from fed/fasted data, with respect to the proportion of healthy/high risk patients that are seen in the clinic. Namely, Craig et al. [135] analysed overdose patterns in 663 patients over 16 years and found that 42.3% patients had psychiatric history, 45.3% had alcohol abuse, and 44.7% combined the overdose with alcohol. We take the mean of these values (44.1%), and in line with this, in our mixed population case, 44.1% of the population were assumed to be high-risk and 55.9% were assumed to be healthy.

As in Chapter 2, each virtual population dataset consisted of 1,000 independent and individually distributed *in silico* individuals. Within this chapter, the *in silico* mice were administered a random dose selected from a uniform distribution of range 0-1,000 mg/kg. As in Chapter 2, biomarker concentrations were subsequently extracted at a random time-point from a uniform range between 0-24 hours. Simulated concentrations were normalised in the range [0,1] using the min-max normalisation method [103] to account for varying orders of magnitude. As in Chapter 2, experimental noise was replicated in the *in silico* data by applying observed *in vivo* standard deviations in biomarker concentrations from an APAP study performed by

Antoine et al. (2009) (ALT s.d = 11.22, HMGB1 s.d = 0.00097, K18 s.d = 2.39, fragmented K18 s.d = 0.12 $\mu\text{mol/l}$).

Latin Hypercube Sampling (LHS) parameter bounds			
Model	Parameter	Lower bound	Upper bound
PK	k_a (h^{-1})	0	20
	V_c (l/kg)	0	1
	V_p (l/kg)	0	1
	k_{12} (h^{-1})	0	1
	k_{21} (h^{-1})	0	10
	k_{el} (h^{-1})	0	1
GSH	gsh_0 ($\mu\text{mol/l}$)	0	700
	k_{pr}	0	100
	ξ	0	1
	k_o (h^{-1})	0	1
	δ (h^{-1})	0	1
ALT	n	1	6
	R_{50} ($\mu\text{mol/l}$)	0	700
	k_{out} (h^{-1})	0	1
	g_{max}	0	0.99998
HMGB1	n	1	3
	R_{50} ($\mu\text{mol/l}$)	0	700
	k_{out} (h^{-1})	0	1
	$r g_{max}$	0	1
K18	k_{in}^{18} (h^{-1})	0	6
	k_{out}^{18} (h^{-1})	0	1
	k_{max}	0	1
	$gsh_{\theta 2}$ ($\mu\text{mol/l}$)	0	700
f-K18	k_{in}^{f18} (h^{-1})	0	6
	k_{out}^{f18} (h^{-1})	0	1
	$gsh_{\theta 2}$ ($\mu\text{mol/l}$)	0	700
	$gsh_{\theta 1}$ ($\mu\text{mol/l}$)	0	700
	ε ($\mu\text{mol/l}$)	0	100

Table 3-2: Parameter bounds used for Latin Hypercube Sampling of initial estimates for the parameter optimisation. For each parameter within the mathematical model, initial estimates for parameter optimisation were chosen through Latin hypercube sampling. The lower and upper bounds for each parameter in this Latin hypercube is detailed.

3.2.6. Predicting time since administration and initial dose as continuous variables

To test the use of biomarkers in predicting time since administration and initial dose separately, multiple linear regression analysis was applied to the *in silico* derived data, as in Chapter 2 (section 2.2.2). As described in Chapter 2, there are various assumptions that should be tested in order for multiple linear regression analysis to be applied. As a reminder, the key assumptions are:

- *The outcome variable is measured on a continuous level;*
- *One or more of the predictor variables are continuous, ordinal or nominal;*
- *Observations must be independent; and*
- *Multi-collinearity/Outliers/highly influential points should not be present in the observed data.*

All of these assumptions held for the *in silico* derived data from this chapter.

An additional assumption of multiple linear regression is:

- *Residuals should be approximately normally distributed;* a histogram and normal P-P plot of the standardized residuals in the models predicting both time since administration and initial dose were inspected. Results for the 'healthy' dataset are shown in Figure 3-3, but results were similar in the 'average' and 'high-risk' datasets.

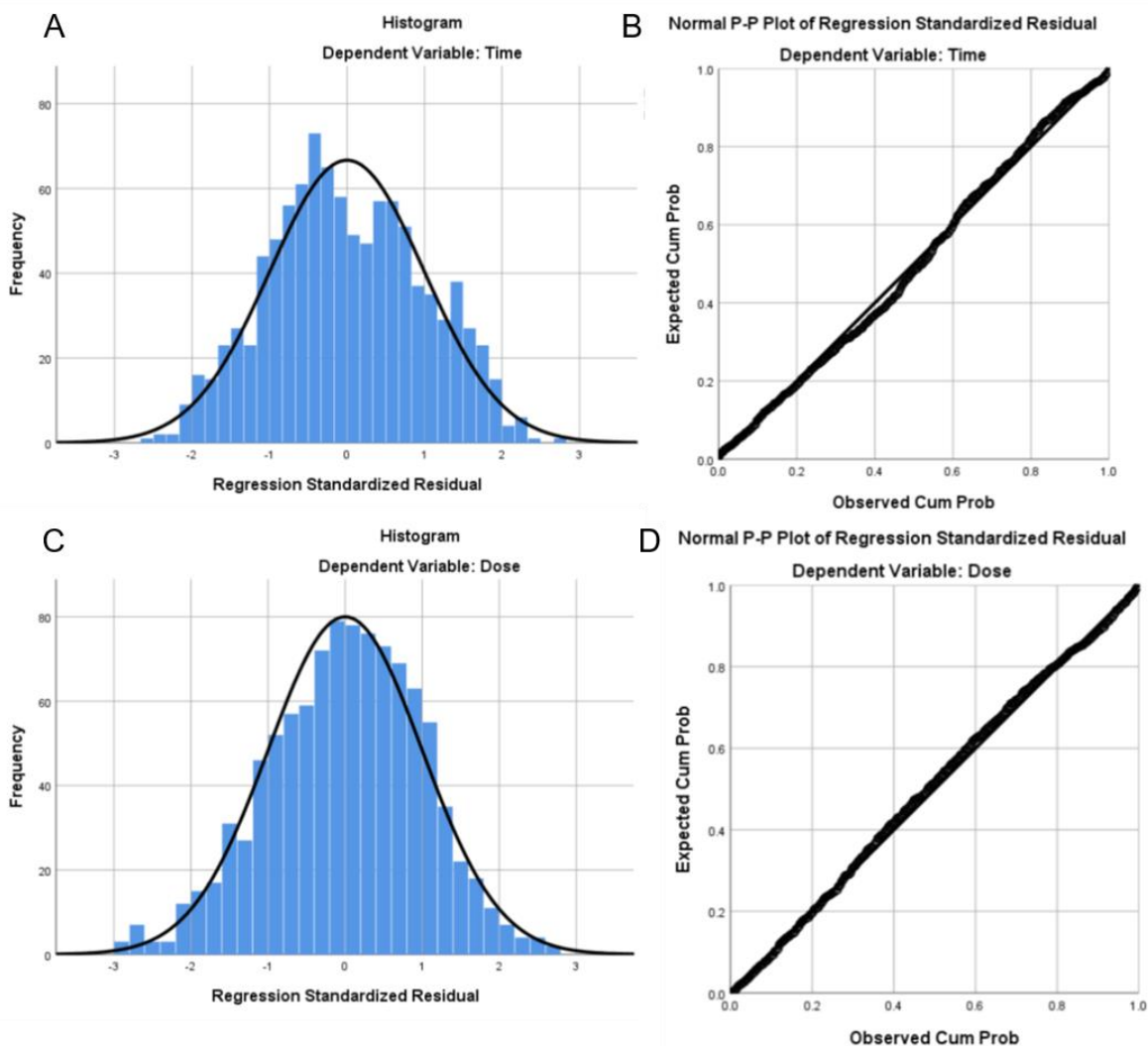


Figure 3-3: Data diagnostic plots. (A, C) Histogram – Visualisation of how the standardized residuals are distributed for the models predicting time since administration and initial dose respectively. For data to be diagnosed as normally distributed, the mean value should be approximately 0 and the standard deviation should be approximately 1. (B, D) Normal P-P plot of regression standardised residual for the models predicting time since administration and initial dose respectively. The cumulative probability expected from the model is plotted against the observed cumulative probability. For the data to be diagnosed as approximately normally distributed, most results (circles) should lie along the diagonal line.

3.2.7. Visualisation

For each *in silico* individual, as in Chapter 2, the t-Distributed Stochastic Neighbour Embedding (T-SNE) method [104] was applied to visualise the dataset constituted by the aforementioned variables (APAP, GSH, ALT, HMGB1, K18, fragmented K18). Two-dimensional scatter plots of the embedded data were employed in order to examine class structure and separability whilst retaining model variation. For a more detailed description of the TSNE method, please refer to section 2.2.2 (Chapter 2).

3.2.8. Classification

In order to test the predictive potential of biomarker concentrations, critical ranges for predicting time since administration were defined as (0-2], (2-5], (5-10], (10-15] and (15-24] hours. For dose, the ranges were [0-200], [201-400] and [401-1,000] mg/kg, capturing therapeutic, small, and large (overdoses), respectively. The same classification techniques applied in Chapter 2 were applied and compared here. Please refer to section 2.2.2 (Chapter 2) for a more detailed description of each technique.

3.2.9. Predicting the probability of liver injury

The binary logistic regression framework described in Chapter 2 has been recently published [131]. The framework uses experimental biomarker time-course data, predicting a corresponding histology score for each mouse. When developing the framework, the whole panel of biomarkers (ALT, HMGB1, full and fragmented K18) were tested, and the analysis found HMGB1 concentration to be the most significant in predicting the probability of liver injury. The resultant logistic regression model was used within this chapter, in combination with PK-PD model simulations (from the newly refined model) to predict the drug-induced liver injury (DILI) probability.

3.3. Results

3.3.1. Identifiability analysis – original model (defined in Chapter 2)

In order to determine the robustness of the parameter set defined for our previous model (Chapter 2), we carried out an identifiability analysis on the model framework. This analysis allowed us to determine whether the parameters we proposed were unique within our system, helping to analyse whether or not these parameters are in fact representative of the mechanisms we desire. Additionally, the analysis provided information that allowed us to determine whether sufficient data had been included in the parameter optimisation, ensuring the parameters can be defined with confidence. Parameters in the model structure (defined in Chapter 2) were individually perturbed to visualise resultant differences between model output and experimental data. These changes were assessed to determine the identifiability of

each parameter in the original model structure. As seen in Figure 3-4, 10 out of 21 parameters were identifiable. The ALT component of the model was identifiable, and all but one of the parameters in the PK component were identifiable. However, the HMGB1 component of the model was structurally unidentifiable, and practical unidentifiabilities exist in all other elements of the model. Five parameters were concluded to be practically unidentifiable. There were 3 practical unidentifiabilities within the GSH component: the ratio of NAPQI forming other protein adducts relative to being detoxified by GSH (k_{pr}), the proportion of eliminated APAP that is transformed into NAPQI (ξ), and the basal removal rate of GSH (k_o). The remaining practical unidentifiabilities were found within the K18 and fragmented K18 components of the model: the decay rate of full K18 ($K18-k_{out}$), and the fragmented K18 Hill coefficient (fragmented k18 - n). The remaining 6 parameters were structurally unidentifiable. These included the theoretical volume of the peripheral compartment (V_p) from the PK component and all parameters from the HMGB1 component: the Hill coefficient (HMGB1 - n), the GSH concentration at which the augmented production rate of HMGB1 reaches 50% of its maximum (HMGB1 - R_{50}), and the decay rate (HMGB1 - k_{out}). Two out of three parameters within the fragmented K18 component were structurally unidentifiable: the GSH concentration at which the augmented production rate of fragmented K18 reaches 50% of its maximum (fragmented K18 - R_{50}), and the decay rate (fragmented K18 - k_{out}).

3.3.2. Model refinement

Following model refinement and re-parameterisation against the multiple datasets (increased dose range and fasted data), we identified a number of parameter adjustments. In cases where parameters are common to both model structures, percentage changes in their optimised values following re-parameterisation are shown in Table 3-3. The baseline level of GSH, gsh_0 , was optimised against the data in the new model structure, rather than fixed as in the previous model (Chapter 2). The resultant value reduced by almost 20% to 559.47 $\mu\text{mol/l}$. The proportion of CYP-activated APAP that is transformed into NAPQI, ξ , increased to around 80%, but there was also an increased level of NAPQI detoxification resulting from the re-parameterisation (represented by a 22.1% decrease in parameter k_{pr} , the ratio of NAPQI forming other protein adducts relative to detoxification). The new parameter, δ , incorporated an effective delay in GSH repletion in the fasted case due to

depleted co-factors. The optimised value was 0.0483 and considerably reduced the timescale of GSH dynamics in the fasted case.

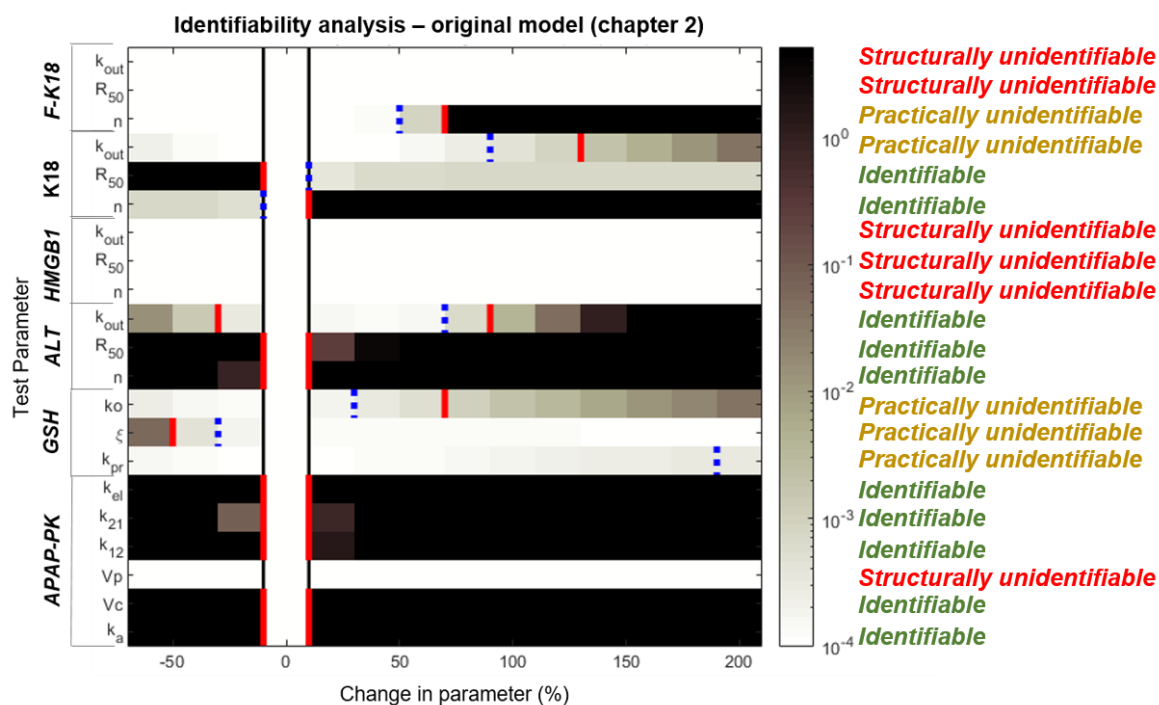


Figure 3-4: Identifiability analysis of the model structure defined in Chapter 2. Each test parameter is fixed at 20% intervals, and the other parameters in the model are allowed to vary. The percentage change of resultant optimised function values (SSE) are plotted at each interval (on log-scale). The lowest SSE change is represented by white; darker colours show an increase in SSE change. The highest SSE change is represented by black. Blue dashed bounds indicate where the parameter is identifiable at the 0.5% level. Red bounds indicate where the parameter is identifiable at the 1% level. A parameter is identifiable at the 1% level if it is bounded by red in both the positive and negative parameter-change directions. If the parameter is bounded by red in one direction and blue in the other direction, it is identifiable at the 0.5% level but practically unidentifiable at the 1% level. A parameter is practically unidentifiable if it is bounded by red/blue in either the positive or negative parameter change direction, and unbounded in the opposite direction. A parameter is structurally unidentifiable if it is unbounded in both positive and negative parameter-change directions (no red/black bounds exist for the whole range of parameter changes). For example (ALT- R_{50}) is identifiable at the 1% level, (K18- R_{50}) is identifiable at the 0.5% level but practically unidentifiable at the 1% level, (fragmented K18- n) is practically unidentifiable, V_p is structurally unidentifiable.

Following model refinement, we found that an increased amount of GSH depletion was required for the GSH-induced ALT response to be half of its maximal value in the refined model structure. Namely, 67% GSH depletion was required in the framework described in Chapter 2, whereas almost 94% was required in the framework described within this chapter. Although more GSH depletion was also required for HMGB1 induction within the new framework, this biomarker response was still faster than that of ALT, since it had reached 50% maximal response rate at

around 86.5% GSH depletion. For Full-K18, results from the optimisation suggest that augmented production of the biomarker will occur when GSH decreases below 174.5205 $\mu\text{mol/l}$ (~68.8% depletion). Fragmented K18 was also induced at this level of GSH; however, if GSH was further depleted to 167.3663 $\mu\text{mol/l}$, augmented production of this apoptosis marker would cease, and the necrotic, full version of the biomarker would then dominate. In the fasted case, this threshold increased by 5.0457 $\mu\text{mol/l}$ (ϵ) such that the necrotic switch (as indicated by the absence of K18 fragmentation) occurred when GSH is depleted beyond 172.412 $\mu\text{mol/l}$. Optimised model simulations were plotted and compared with the APAP PK time-course data (Figure 3-5), the biomarker PD time-course data (Figure 3-6) and the APAP dose-biomarker response data (Figure 3-7). The refined model provided a much better replication of the dose-response data, which is a key result from this chapter (Figure 3-7).

Model	Parameter	Optimised value		% change
		Original (Chapter 2)	Refined model	
APAP	k_a (h ⁻¹)	9.05	8.6152	-4.8%
	k_{el} (h ⁻¹)	0.52	0.5459	+5%
	k_{12} (h ⁻¹)	0.42	0.4502	+7.2%
	V_c (l/kg)	0.02	0.0220	+10%
	V_p (l/kg)	0.01	0.2102	+2000%
	k_{21} (h ⁻¹)	1.01	1.0315	+2.1%
GSH	gsh_0 (μmol/l)	696.9136	559.47497	-19.7%
	k_{pr}	71.06	55.33401	-22.1%
	ξ	0.68	0.80571	+18.5%
	k_o (h ⁻¹)	0.25	0.78807	+215%
ALT	R_0 (μmol/l)	0.7626	0.7626	0%
	n	9.26	4.3324	-53.2%
	R_{50} (μmol/l)	227.67	35.6531	-84.3%
	k_{out} (h ⁻¹)	0.0002	0.0004	+1%
HMGB1	R_0 (μmol/l)	0.0005	0.0005	0%
	n	4.90	2.3445	-51.2%
	R_{50} (μmol/l)	399.08	75.2828	-81.1%
	k_{out} (h ⁻¹)	0.35	0.0964	-72.5%
Full K18	r_0^{18} (μmol/l)	0.0146	0.0088	-39.73%
	k_{out}^{18} (h ⁻¹)	0.0007	0.0031	+342.9%
Fragmented-K18	r_0^{f18} (μmol/l)	0.0642	0.0977	+52.2%
	k_{out}^{f18} (h ⁻¹)	0.02	0.0031	-84.5%

Table 3-3: Parameter changes following model refinement. Any parameter that remained within the new model structure is defined, with its original value and the re-parameterised value. Percentage changes are also defined.

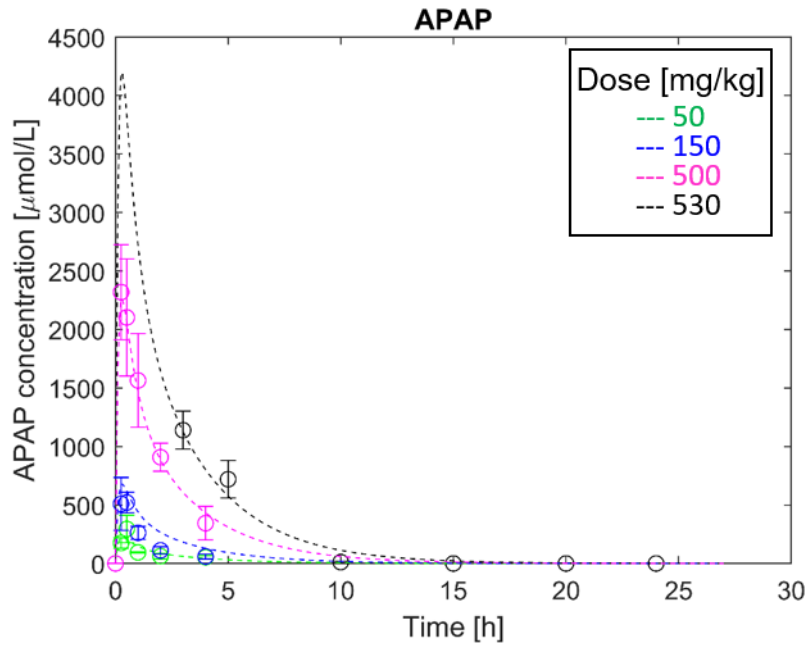


Figure 3-5: APAP PK Model simulations versus time-course data. APAP concentration simulations are plotted for 50 (green), 150 (blue), 500 (pink) and 530 (black) mg/kg doses of APAP.

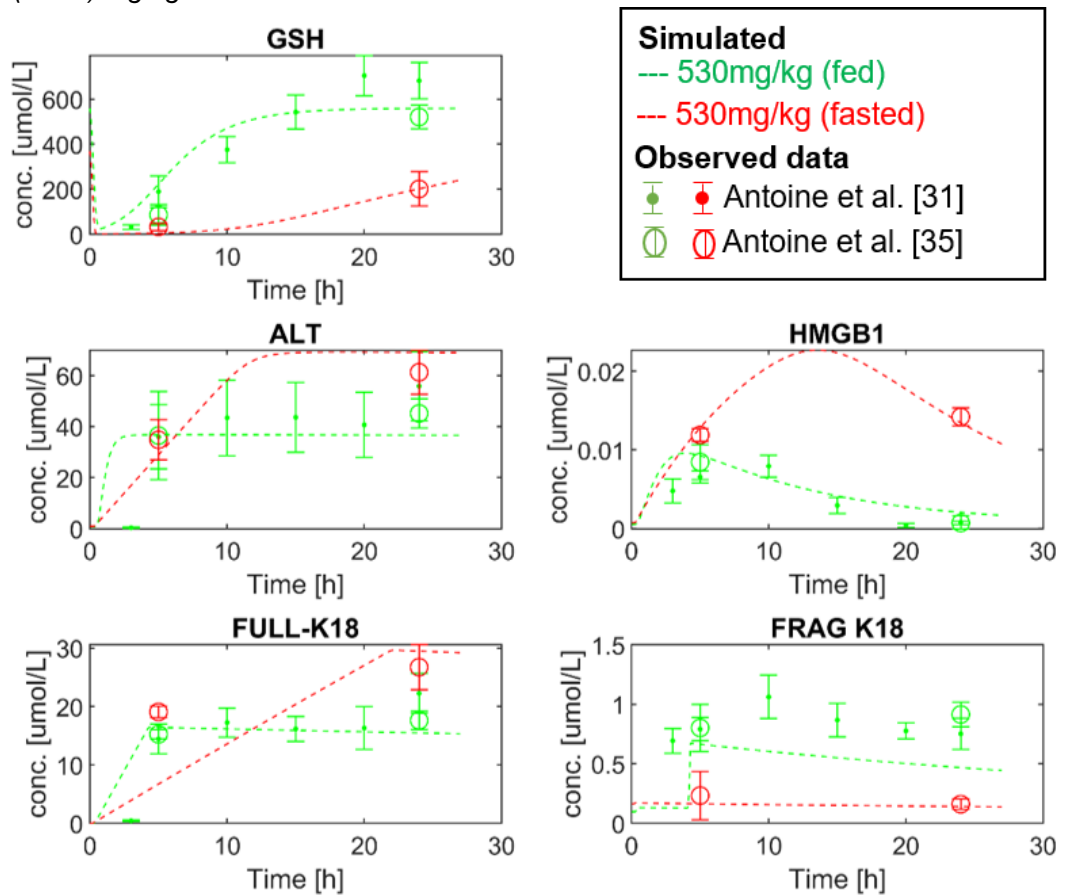


Figure 3-6: Biomarker PD Model simulations versus time-course data. Fed and fasted simulations (green and red respectively) are plotted for GSH and biomarkers (ALT, HMGB1, Full and fragmented K18) following a 530 mg/kg dose of APAP.

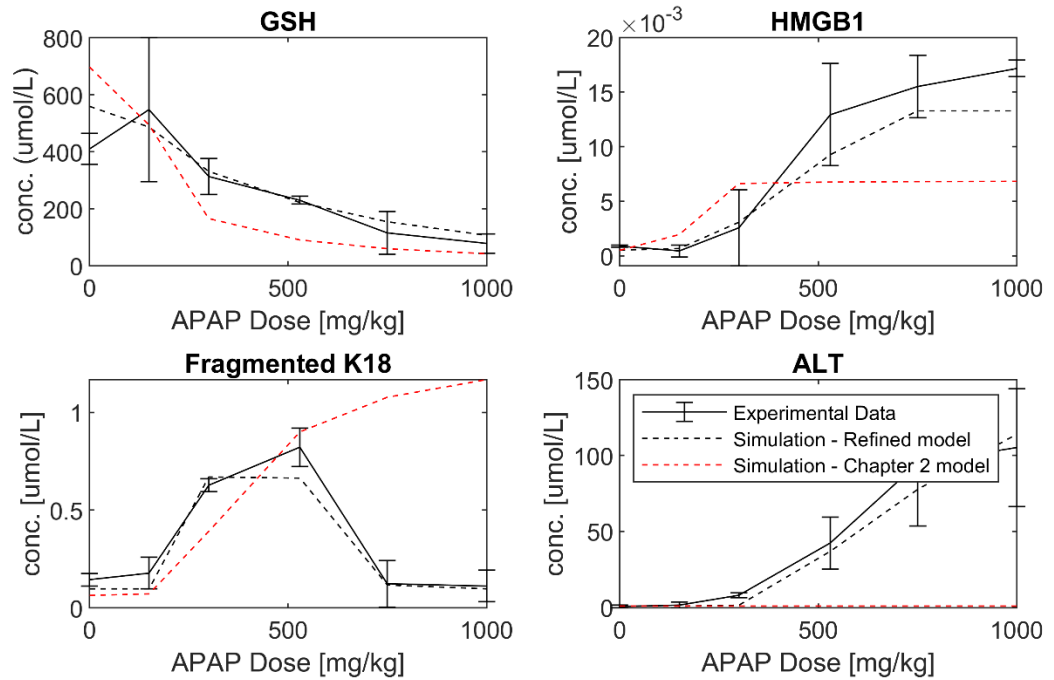


Figure 3-7: Model simulations versus dose-response data. *Solid lines represent the experimental data used for calibration. Black dashed lines represent dose-response simulations of GSH and biomarkers (ALT, HMGB1 and fragmented K18) in fed mice for a range of APAP doses (0, 150, 300, 530, 750, 1,000 mg/kg) using the refined model. Red dashed lines represent dose-response simulations of GSH and biomarkers (ALT, HMGB1 and fragmented K18) in fed mice for a range of APAP doses (0, 150, 300, 530, 750, 1,000 mg/kg) using the original model (defined in Chapter 2).*

3.3.3. Identifiability analysis – refined model

Following an identifiability analysis of the original model structure (results of which are defined in section 3.3.1 of this chapter), we refined the original model in an attempt to resolve some of the structural unidentifiabilities found from the analysis. This included, for example, rather than assuming similar dynamics for K18, modelling biomarkers full K18 and fragmented K18 to be necrotic and apoptotic versions of the same biomarker. All modelling refinements made in an attempt to improve structural unidentifiability issues are defined in section 3.3.2 of this chapter. As well as attempting to improve the structural unidentifiabilities concluded from the analysis, we also worked to improve the practical unidentifiabilities found within the model. To do so, we optimised the refined model against an extended experimental dataset. Whilst the initial model (Chapter 2) was optimised against data from one study [39], the refined model was optimised against data from three studies [35,39,131]. Since additional data was required, we chose to use fasted mouse data within optimisation, so that our model may be extended to make predictions for

individuals with suppressed GSH levels. While the identifiability analysis of the original model framework provided interesting insight, it was important to determine whether the changes made to improve the system were justifiable and had improved the confidence within the system. We therefore conducted an identifiability analysis on the refined model structure for comparison.

To conduct this analysis, parameters in the refined model were individually perturbed to visualise resultant differences between model output and experimental data. These changes were assessed to determine the identifiability of each parameter. Results can be seen in Figure 3-8. In the refined model, 16 out of 27 parameters were now identifiable. Parameter identifiability remained for the ALT component, and the GSH component of the model was now also completely identifiable. The PK component remained identifiable, other than the volume of the peripheral compartment (V_p), which remained structurally unidentifiable. There were some unidentifiabilities still present within this updated model. Other structurally unidentifiable parameters included: R_{50} , k_{out} and g_{max} from the HMGB1 component; and $K18_{kout}$, f_{kin} , f_{kout} , gsh_{θ_1} and ε from the full and fragmented K18 component. Practical unidentifiabilities remained for two parameters: the HMGB1 Hill coefficient (n) and the production capability of full K18 (k_{max}).

We now compare the results of applying an identifiability analysis to the original model structure defined in Chapter 2, and also the refined model defined here within this chapter. Comparative results can be seen in Figure 3-9. Less than half of the parameters (47%) were identifiable in the original model structure (only the ALT component of the model was completely identifiable), 24% of the model parameters were practically unidentifiable and 29% were structurally unidentifiable. However, following model refinement, the identifiability results improved; the percentage of identifiable parameters increased to above half of the parameters (60%). Parameter identifiability remained for the ALT component of the model and the percentage of practical unidentifiabilities reduced from 24% to 7%. The percentage of structural unidentifiabilities was approximately the same in both the original and refined model (29% compared to 33%). Whilst the GSH component was unidentifiable in the original model, it was completely identifiable in the refined model. Although unidentifiabilities improved for parameters in the HMGB1, K18 and f-K18 components, some unidentifiabilities remained.

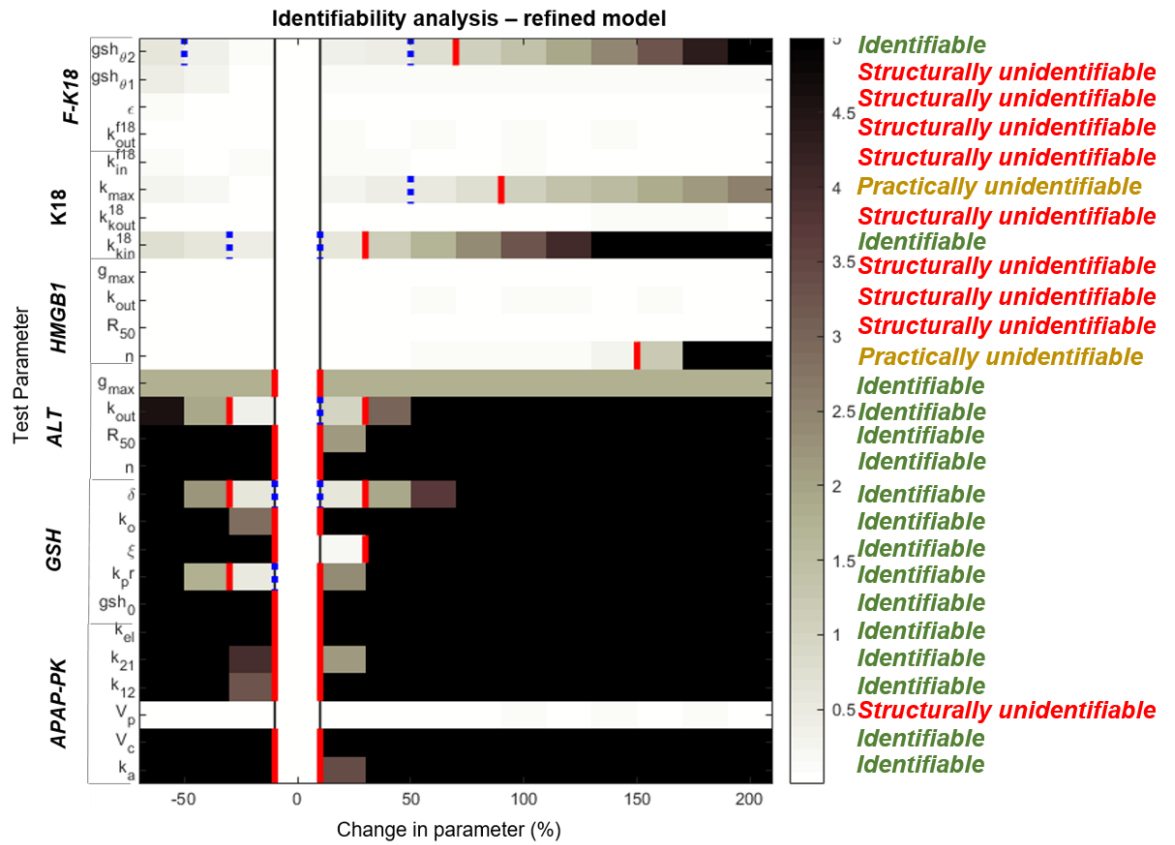


Figure 3-8: Identifiability analysis of the refined model structure. Figure annotation is the same as for Figure 3-4. Example results from figure: The ALT R_{50} parameter is identifiable at the 1% level, the HMGB1 n parameter is practically unidentifiable, the PK V_p is structurally unidentifiable.

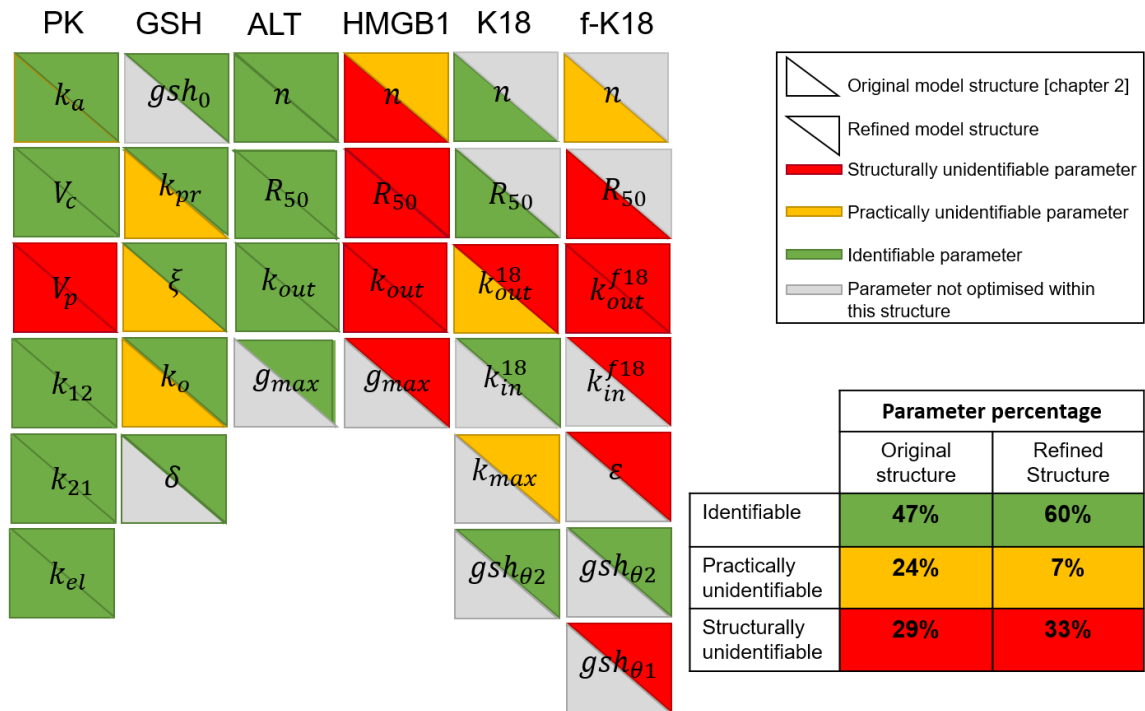


Figure 3-9: Parameter identifiability comparison between the original and refined model. Each column represents a different sub-component of the APAP PKPD model. Each element (or square) represents the parameter's identifiability within the model. Triangles in the bottom left-hand corner of each element represent the parameter's identifiability in the original model structure. Triangles in the top right-hand corner of each element represent the parameter's identifiability in the refined model structure. If the triangle is red, the parameter is structurally unidentifiable; if the triangle is yellow, the parameter is practically unidentifiable; if the triangle is green, the parameter is identifiable. If the triangle is grey, the parameter was not present within the corresponding model structure, and therefore could not be tested for identifiability.

3.3.4. Predicting initial dose and time since administration as continuous variables

We used the refined PKPD framework to simulate biomarker concentrations representative of 3 sub-populations; a “healthy” population, “high-risk” population and “mixed” population (virtual population details are described in section 3.2.5 of this chapter). Within each sub-population, each individual was provided a random dose of APAP, and biomarker concentrations were extracted at a random time-point. As in Chapter 2 (section 2.3.5) a multiple linear regression analysis was applied to the *in silico* derived data in an attempt to predict the exact time since ingestion and initial APAP dose. For the healthy population, an exact time since administration could be estimated with an error of approximately 3.9 hours, and an exact initial dose amount could be estimated with an error of approximately 66.14 mg/kg (Table 3-4).

Time since administration could be predicted more accurately in the high-risk population, with a standard error of 2.118 hours. Initial dose was much harder to predict in the high-risk population, however, with the standard error being almost triple that of the model for the healthy population (184.1 mg/kg). Predictions for an assumed mixed population were similar to that of the healthy population, with a slight improvement in predicting time since administration (standard error 3.485 hours) and a slight reduction in accuracy for predicting initial dose (standard error 73.73 mg/kg). When predicting time since administration, in the healthy population, all biomarkers were significant, in the high-risk population, all biomarkers were significant, however, in the mixed population, ALT and fragmented K18 were no longer significant. When predicting dose, in the healthy population, all biomarkers except HMGB1 were significant, in the high-risk population, all biomarkers except fragmented K18 were significant, and in the mixed population, all biomarkers were significant.

Predictors	Healthy Population		High-risk Population		Mixed Population	
	Time	Dose	Time	Dose	Time	Dose
APAP Conc.	- 16.549*** (1.1785)	673.985*** (17.0745)	-9.5093*** (0.5963)	1295.366*** (47.8525)	-18.122*** (0.9542)	909.341*** (17.8005)
ALT conc.	8.972*** (0.9533)	515.625*** (13.8110)	5.5034*** (0.6537)	695.2118*** (52.4551)		655.0837*** (17.2145)
HMGB1 conc.	- 22.035*** (0.9113)		-8.8119*** (0.5440)	367.0774*** (43.6578)	-17.869*** (0.7705)	289.2333*** (14.3736)
FullK18 conc.	7.521*** (1.0632)	479.137*** (15.4035)	21.0427*** (0.4743)	- 154.2403*** (38.0652)	22.3833*** (0.9145)	229.0486*** (17.0605)
Fragmented K18 conc.	-4.599* (1.3895)	151.975*** (20.1309)	3.8472*** (0.5962)			33.8383* (15.5231)
Constant	14.918*** (0.3074)	77.247*** (4.4529)	5.8083*** (0.2527)	-80.145*** (20.2759)	12.6308*** (0.3694)	-65.6354*** (6.8914)
Residual Std. Error (df == 994)	3.907	66.14	2.118	184.1	3.485	73.73
Note:	*p<0.05		**p<0.01		***p<0	

Table 3-4: Multiple linear regression analysis results. Independent variable coefficients for predicting time since administration (Time) and initial dose (Dose) respectively for healthy, high-risk and mixed populations. The first number in each element of the table represents the biomarker coefficient in the regression model and the second number represents the corresponding error. For example, -16.549 is the APAP concentration coefficient in the healthy population model predicting time since administration, and this coefficient has an error of 1.1785. The standard error of overall model predictions is provided. The significance of each biomarker in the model is indicated by the number of asterisks (see note).

3.3.5. Visualisation and classification of *in silico* derived data

As in Chapter 2, T-SNE visualisation was applied to the *in silico* derived data to investigate time/dose class structure/separability. Subsequently, various classification techniques were also employed, using the biomarker concentrations of the *in silico* observations in an attempt to classify a time/dose category. Differences in levels of discrimination regarding initial dose and time since administration for healthy, high-risk and mixed populations can be seen for each case by embedding the *in silico* derived data in two-dimensions using T-SNE (Figure 3-10). Both variables, time since administration and initial dose, were reasonably separated in all healthy, high-risk and mixed populations. This result was supported by the reasonable accuracy rates of the classification techniques. Time since administration was difficult to discriminate from the biomarker concentrations in the healthy population, as shown by the cluster of observations of the right-hand side of Figure 3-10(A)). However, these same observations corresponded to low-dose situations, as shown in Figure 3-10(B) which visualises this same cluster of observations with regards to the initial dose,

From measuring biomarker concentrations, a correct time since administration category could be predicted with 69.9% accuracy for the healthy population, and a correct initial dose category can be predicted with 91.5% accuracy. These results are indicated in Table 3-5. A time since administration category was easier to predict in the high-risk population (80.4% accuracy). However, predictions for initial dose in the high-risk population were less accurate than in the healthy population (79%). If an observation was assumed to be taken from a mixed population, the prediction accuracy was similar to that of a healthy population, with results of 69.94% and 89.5% for time since administration and initial dose respectively. In our previous study, a correct time category could be predicted with 72.8% accuracy and a correct dose category could be predicted with 86.5% accuracy (Chapter 2, section 2.3.4). For a healthy population, the time classification results slightly worsened with our new framework. However, the dose classification results slightly improved.

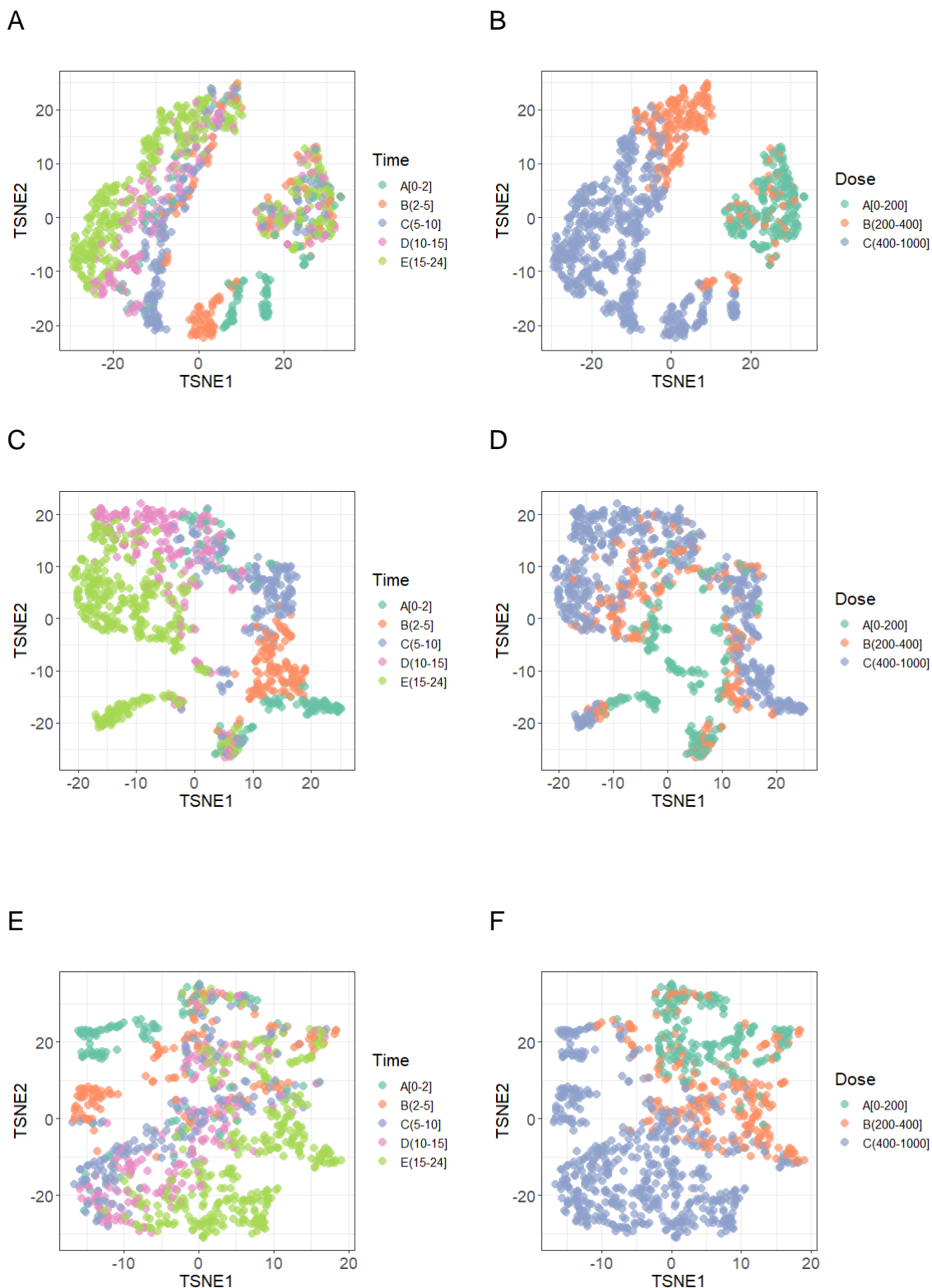


Figure 3-10: Visualisation and classification of time-since-administration and dose results for healthy, high-risk and mixed populations. For time-since-administration, dark green represents class [0-2), orange represents [2-5), blue represents [5-10), pink represents [10-15) and pale green represents [15-24) hours. For dose, green represents [0-200), orange represents [201-400) and blue represents [401-1,000) mg/kg. TSNE visualisations of *in silico* mouse observations with respect to time since administration and initial dose can be seen in (A)-(B) for the healthy population, (C)-(D) for the high-risk population and (E)-(F) for the mixed population.

Classification Method	Healthy Population		High-risk Population		Mixed Population	
	Time	Dose	Time	Dose	Time	Dose
KNN Regression	66.2%	91.5%	80.4%	79%	69.5%	87.4%
Naïve Bayes	64.4%	91.3%	76.2%	72.7%	68.2%	88.5%
Multinomial logistic regression	68%	90.8%	73.1%	77.1%	75.7%	89.5%
Ordinal logistic regression	53.7%	90%	67.8%	78.1%	57.2%	87.5%
Linear discriminant Analysis	59.3%	90.7%	72.5%	77.9%	65.8%	87.8%
Quadratic discriminant analysis	69.9%	90.4%	76.8%	71.7%	69.94%	86.8%

Table 3-5: Classification accuracy of predictions based on biomarker concentrations for healthy, high-risk and mixed populations. For example, the Quadratic discriminant analysis model can predict the correct time since administration category in the healthy population with 69.9% accuracy.

3.3.6. Predicting the probability of liver injury

Figure 3-11 shows how the probability of liver injury changes over time for both healthy and high-risk populations, for doses between 0-1,000 mg/kg. The model-derived probabilities were dependent only on HMGB1 concentration (as predicted by our previous logistic regression model in Chapter 2). A threshold probability of 0.5 was used to determine the threshold likelihood of DILI (white boundaries in Figure 3-11). Any observation within the white contour boundary was therefore predicted to be representative of probable liver injury.

For the healthy population, the time-frame for biomarker concentrations representing those of probable liver injury was around 2 to 18 hours. The APAP dose threshold predicted to induce toxicity was approximately 350 mg/kg. Note that this prediction is only slightly above the currently used toxic dose (300 mg/kg) [97] whereas the framework provided in the previous chapter, predicted toxicity would be apparent at an APAP dose above 200mg/kg. The framework for a healthy population suggests that there is almost 100% chance of liver injury when the dose is only slightly higher than 350mg/kg.

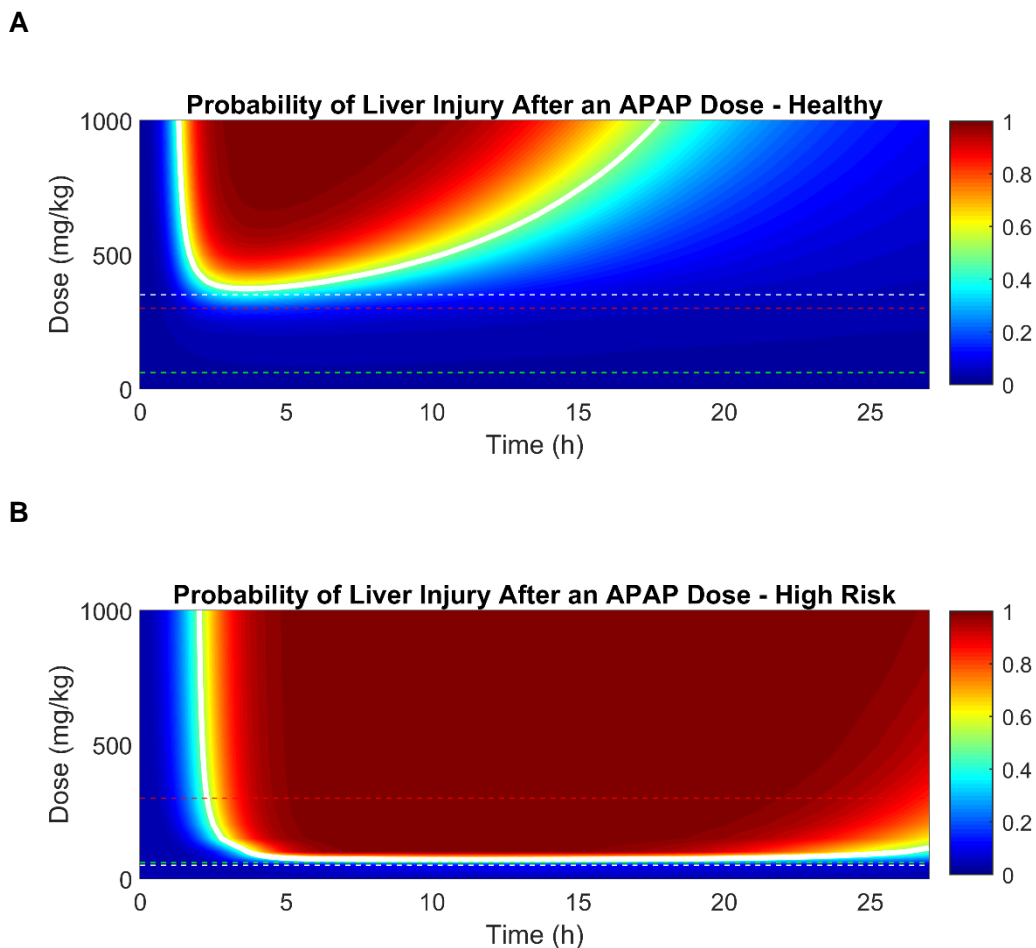


Figure 3-11: Proposed framework for predicting the probability of liver injury dependent upon dose, time and HMGB1 concentration. *The plotted frameworks are representative of both healthy (A) and high-risk (B) populations. In each, the white contour indicates the threshold of 0.5 probability of liver injury; the red dashed-line represents the currently used APAP dose for toxicity studies in mice; the white dashed-line represents toxic dose as proposed by our model; and the green dashed-line indicates currently considered therapeutic dose for mice. In the high-risk population, the toxic dose proposed by our model and the therapeutic dose are identical.*

For the fasted population, however, at the currently recommended therapeutic APAP dose (60 mg/kg), there was approximately 50% chance of liver injury progression. For any dose above the known therapeutic threshold, liver injury progression was predicted to be highly probable (above 60%); beyond approximately 2.5 h post-dose, HMGB1 concentrations remained indicative of highly probable liver injury for the whole time-course.

3.3.7. Visualising the probability of liver injury following an APAP dosed based on biomarker concentrations

In Chapter 2, we carried out a forward step-wise binary logistic regression analysis to determine the most significant panel of biomarkers in predicting the probability of liver injury. The analysis concluded that HMGB1 was the most significant biomarker for predicting DILI probability, with the output from SPSS software being a logit equation based solely on this biomarker concentration,

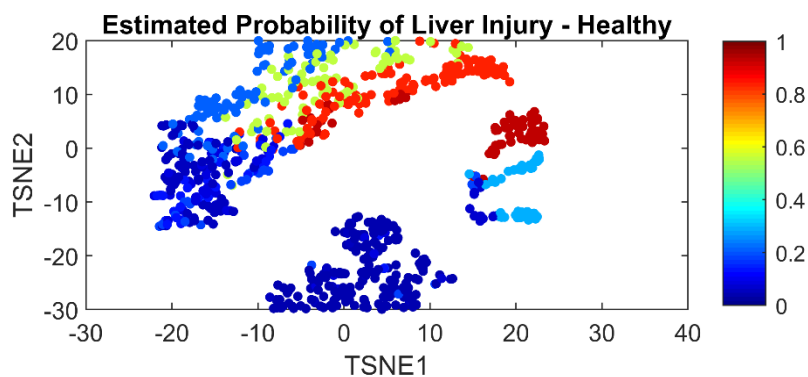
$$L_{eq} = 0.635[\text{hmgb1}] - 3.870. \quad (3-8)$$

Within this chapter, using the refined model, we simulated multiple *in silico* populations. For each observation within both the “healthy” and “high-risk” populations, simulated HMGB1 concentrations were substituted into equation (3-8). Then, using the standard probability equation, we calculated for each *in silico* individual the probability of liver damage as follows

$$\text{Probability of liver damage} = \frac{1}{1 + e^{-L_{eq}}}. \quad (3-9)$$

Each *in silico* observation was then visualised and discriminated by their resultant liver injury probability using the TSNE method (Figure 3-12). Relatively safe/unsafe observations were identifiable in both healthy and high-risk populations. For the healthy population (Figure 3-12(A)) most of the observations had less than around 35% chance of DILI progression. The small group of observations representing at-risk individuals can be identified at the top, right-hand side of the image, portrayed by red markers. In stark contrast, most of the *in silico* observations had almost 100% chance of liver injury progression in the high-risk population, as can be seen by the dominance of red markers in Figure 3-12(B). The observations corresponding to lower injury potential are indicated in both the central and the left-hand side of the plot.

A



B

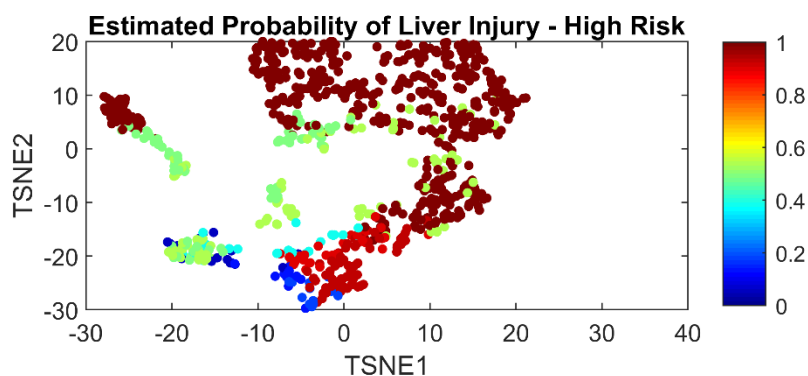


Figure 3-12: two-dimensional TSNE visualisation of *in silico* observations with respect to their estimated probability of liver injury. Separated liver injury probabilities are visualised for a range of *in silico* generated inputs for both healthy (A) and high-risk (B) populations. Predicted DILI probabilities (0-1) for each individual observation are indicated by the colour bar. Dark blue represents 0% chance of liver injury progression, while dark red represents 100% chance.

3.4. Chapter discussion

The current clinical APAP treatment framework is known to be inaccurate, mainly due to a dependency on known time of APAP ingestion, but also due to inaccurate estimates of initial dose and being heavily based on biomarkers that are currently criticised for being insensitive [136]. In Chapter 2, we detailed a computational framework for predicting the probability of liver injury based on novel biomarker concentrations. Whilst mathematical models are increasingly being accepted as an efficient tool in toxicity testing [44,78], they are only a representation of the biology, and due to the necessity of assumptions for simplification, many uncertainties can arise (in both parameter estimates and subsequent output). For this reason, the most useful mathematical models are provided with a level of quantified uncertainty in their predictions, so that the frameworks can be utilised in the safest, most

efficient manner possible. We therefore used this chapter to determine any levels of uncertainty that were present in the framework defined in Chapter 2, and used the results to try and improve any uncertainties where possible. The identifiability analysis performed within this study did in fact highlight parameter unidentifiabilities within the original model that required addressing. The identifiability analysis established where the model defined in Chapter 2 required structural changes and also where more data was required in order to increase confidence in predictions. Using the results, we have re-structured the model where necessary and re-parameterised against additional data (from both fed and fasted mice). A comparative identifiability analysis of the newly defined model structure found that only 7% of parameters are now practically unidentifiable (compared with 24% previously).

As well as improving confidence in the mathematical representation of the system, we have now provided a model that is more representative of the clinical environment by firstly including a wider range of APAP doses (previously the model could account for an APAP dose range [0-600] mg/kg, whereas now we can account for a range of [0-1,000] mg/kg). Furthermore, the new treatment framework can now be adjusted for individuals considered to be at high risk of APAP-induced liver injury [5]. To account for the fact that many overdose patients tend to have an increased underlying susceptibility to APAP-induced liver injury, we chose to optimise our model against both fed and fasted mouse data. When training the biomarker models against multiple datasets and comparing to the model defined in Chapter 2, which is calibrated against only 530 mg/kg APAP dose time-course data, there were noticeable adjustments in the parameter values, particularly for the GSH model. Many of these changes are intuitive and may be representative of changes in mechanisms due to fasting. Additionally, changes may be attributable to the fact that the model can now better account for a larger variety of dosing scenarios. The significant changes in parameter values highlight the importance of optimising against a range of both therapeutic and toxic datasets and both healthy and unhealthy populations in the modelling field.

The model we propose in this chapter is more biologically relevant than the previous model defined in Chapter 2. Originally, we assumed that full and fragmented K18 had similar mechanisms and acted as independent biomarkers. However, full and fragmented K18 are known to be effectively necrotic/apoptotic versions of the same biomarker [88]. Incorporating this into our model framework we confirmed that

necrosis was the pre-dominant form of cell death in mice APAP overdose cases [137]. Also, an identifiability analysis on the new model structure established an increased confidence in parameter estimates for the GSH, HMGB1, full and fragmented K18 components of the model structure. There are, however, some unidentifiabilities remaining, particularly within the HMGB1 and fragmented K18 components of the model, indicating that additional data and model development is still required in order to have full confidence in the uniqueness of the chosen mechanistic parameter values.

Chapter 4: Pre-clinical systems toxicology approach to treating paracetamol overdose: predicting cell death dynamics and investigating antidote treatment regimens

4.1. Background

Up until this point in the thesis, we propose an *in silico* framework based on mouse data which can provide accurate predictions regarding whether or not an APAP overdose has taken place (by quantifying dose amount and time since administration), and whether an overdose is likely to progress into liver injury. There are, however, aspects of the model which require improvement prior to contemplating eventual translation to humans.

Throughout the analyses until this point, HMGB1 outperforms all other biomarkers in predicting overdose and probability of liver injury, including the current clinical favourite ALT. As well as lacking in sensitivity, ALT is also considered sub-optimal in the APAP overdose case due to its non-liver specificity [138]. Although the novel biomarkers used in our *in silico* framework provide enhanced mechanistic information relating to the underlying basis of APAP toxicity [87], and our results thus far suggest that HMGB1 is particularly sensitive, the non-liver specificity issue remains for all biomarkers.

An additional issue is that currently, the mathematical model is built assuming that cell death is one general mechanism. In fact, cell death can usually be described by one of two mechanisms: necrosis and apoptosis [31]. Necrosis is rapid and irreversible and occurs when an external trauma causes cells to quickly inflame and become damaged. Apoptosis, however, is a much slower physiological process and occurs when the external stimuli has not caused so much injury to the cell, that is has the energy (Adenosine Triphosphate, ATP) to programme its own demise, with cellular metabolism and membrane integrity being maintained until very late stages [32]. There is extensive research-based evidence to demonstrate that necrosis is the dominant form of APAP-induced hepatocellular death in overdose cases [33–35]. However, mild levels of APAP toxicity can cause apoptosis [33] and, in both low and high APAP dose cases, there is evidence that these modes of cell death can simultaneously exist until ATP levels deplete, with necrosis dominating for higher doses [139]. Identifying the mode of cell death can therefore be beneficial in determining whether or not a patient is a mild toxicity/overdose case. Coupling this with our dose and time since administration predictions will strengthen our framework. Our framework can predict, for the mouse case, individuals likely to develop liver injury, and therefore those who may benefit from intervention with

antidote therapy. This project has the aim of providing a proof-of-concept framework which cannot only predict cases where intervention may be necessary, but also provide the optimal treatment strategy. By incorporating information regarding the mode of cell death at different doses/time points, the accuracy of the model with regard to predicting when to intervene may be improved.

The PKPD model is currently developed based upon mice dosed with various amounts of APAP and no treatment with the antidote, N'acetylcysteine (NAC). This means that up until this stage, we propose a proof-of-concept framework for initial patient presentation, however the model may not be useful as a platform for monitoring changes in liver injury potential following an antidote. Since NAC is not initially included in the modelling framework (defined in Chapters 2 and 3), we require additional data and model development.

The *In vitro* Toxicology Society (IVTS) provide Mini-Fellowship awards, which allow early-career scientists to learn techniques which are useful for their research aims, but are currently unavailable at their home institution. The aim is to promote further career development within the *in vitro* toxicology field, and also to build strong collaborations between *in silico* and *in vitro* researchers. In 2017 I was successful in winning the IVTS Mini-Fellowship award, and used the £1.5k to fund experimental work in collaboration with the University of Liverpool to fill the aforementioned data gap and allow further model refinement. As well as providing additional data, the fellowship allowed me to obtain hands-on experience of the wet-lab environment, and form collaborations with *in vitro* experimentalists. The experiments were carried out in collaboration with a University of Liverpool PhD student, Nathalie De Bois-Brillant. To obtain the necessary data, we exposed primary mouse hepatocytes to a range of antagonist (APAP) and antidote (NAC) concentrations to investigate their interaction, whilst conducting experimental tests to understand their effect on cell dynamics and mode of cell death.

In order to provide experimental insight into the cell dynamics, we provide results of an ATP assay used to quantify cell viability. In order to determine the mode of cell death we provide results from experiments based on Western Blotting for Caspase 3 (an apoptosis marker) and quantitative polymerase chain reaction (qPCR) analysis for miR-122 (a necrosis marker). The necrosis marker is chosen to be miR-122 due to its high level of liver-specificity, something the model is hitherto lacking. Additional to these experimental findings, we provide results from a fluorescence-activated cell-sorting (FACS) analysis which compares the amount of apoptosis

versus necrosis present at 24h. In order to test the predictivity of biomarkers in an APAP overdose case, and also in an APAP plus antidote treatment (APAP+NAC) case, we apply multiple linear regression analysis. The mathematical model defined in Chapters 2 and 3 is extended within this chapter in order to simulate cell viability dynamics in both the APAP and APAP+NAC cases. Administration schedules of NAC are investigated and compared, determining the effects on biomarker output, probability of liver injury and severity of cell damage.

Chapter aims

- Generate and analyse *in vitro* experimental data to better understand the influence of NAC administration on biomarker concentrations and their subsequent predictivity.
- Obtain greater understanding of liver injury at the cellular level.
- Investigate the effects of NAC administration on defined APAP toxicity measures.

4.2. Methods

All experiments were carried out at the University of Liverpool under the supervision of Nathalie De Bois-Brillant. The protocols described were undertaken in accordance with criteria outlined in a license granted under the Animals (Scientific Procedures) Act 1986 and approved by the University of Liverpool Animal Ethics Committee.

4.2.1. Experimental methods

4.2.1.1. Mouse hepatocyte isolation

Hepatocyte isolation was carried out by Nathalie De Bois-Brillant at the University of Liverpool. Again, methods are detailed here for completeness.

Mice were ordered 7 days prior to the planned experiment to allow for acclimatisation. Wash, perfusion and digestion buffers were prepared one day prior to the planned experiment and stored in the fridge until required. When preparing perfusion and digestion buffers, HEPES was added first to aid salt dissolution.

The perfusion buffer was made by the following method:

Mouse 250 ml perfusion buffer

5 ml HEPES (1M)
217.6 ml ddH₂O
25 ml 10X HBSS
1.2 ml 7.5% (w/v) NaHCO₄
1.25 ml EDTA (50 mM)
pH to 7.35-7.46

The digestion buffer was made by the following method before being split into two bottles: 150 ml (referred to as DGI) and 100 ml (referred to as DGII)

Mouse 250 ml digestion Buffer

5 ml HEPES
216.3 ml ddH₂O
25 ml 10X HBSS
1.2 ml 7.5% (w/v) NaHCO₄
25 mg MgSO₄
1.25 ml CaCl₂
pH to 7.35-7.46

DGI was kept on ice and used only after hepatocyte isolation. DGII was warmed in the water bath, as was 2.5 mg of 20X collagenase which was then added to DGII immediately before tissue digestion. DGI did not contain collagenase and could therefore be used to spin cells whereas DGII could be used for digestion.

4.2.1.2. Liver Perfusion

Liver perfusion was carried out by Nathalie De Bois-Brillant at the University of Liverpool.

The centrifuge was set to a 4°C cycle. The perfusion pump was set up to run at a speed of 8 ml/min (14G tube). Pentobarbitone was taken out of the fridge and diluted down 1:3 in 0.9% saline from a 200 mg/ml stock solution to 50 mg/ml and left aside to reach room temperature. Mice were dosed into the peritoneal cavity (intraperitoneally, i.p), which is situated just below the abdomen, with 1 µl/g of diluted solution. The level of anesthesia was determined by loss of corneal reflex response and foot withdrawal response. Once the mouse was completely sedated, the

perfusion was quickly performed via cannulation of the inferior vena cava and the portal vein was cut to relieve pressure. The liver was perfused with wash buffer until cleared and then perfused with digestion buffer until digested. The liver was then transferred to a petri-dish for release of cells.

4.2.1.3. Cell release

Cell release was carried out by Nathalie De Bois-Brillant at the University of Liverpool. Methods are detailed here for completeness.

The liver was placed into a 50 ml falcon tube containing ice cold 1X DGI. On ice, the liver hepatocytes were cleaned by stripping the capsule at the edges. Using curved forceps, holes were made in the four lobes of the liver. The liver was flipped to the underside and this process was repeated. The liver was punctured many times, before being gently shaken in buffer until the biliary tree became visible. The crude cell mixture was then put through a 100-micron filter to remove connective tissue and aggregates. The hepatocytes were collected and 50 g were spun for 2 minutes. The supernatant was discarded and the hepatocytes were then re-suspended using ice cold digestion buffer (without collagenase). The spin was then repeated. This process (discarding, washing and spinning) was repeated to obtain a pellet of pure hepatocytes. The hepatocytes were then cultured in the following medium

Hepatocyte culture medium (/100 ml)

1 ml P/S

1 ml L-Glutamine

1 ml ITS

100 µl of 100 µM DEX

96.9 ml Williams Medium E

10% FBS omitted since performing ATP assay

Once the liver was perfused and cells released, I was then involved in carrying out each of the experiments that follow, under the supervision of Nathalie. We first investigated the effects of APAP and NAC on cell viability. Secondly, we investigated the effects of APAP and NAC on the mode of cell death (apoptosis/necrosis).

4.2.1.4. Dosing

All cells were incubated for 24 hours. Following this, the primary mouse hepatocytes were exposed to a range of concentrations of APAP (0-20 mM) and one hour later were further exposed to varying concentrations of NAC (0-5 mM).

4.2.1.5. Cell viability – ATP

Adenosine Triphosphate (ATP) is the main energy source for all cells [140]. Measuring cellular ATP levels is therefore often used to measure cell functionality; reduction in cellular energy levels can in turn be assumed to represent reduction in cell viability. We therefore carried out an ATP assay to determine how different dose combinations of APAP+NAC effected cell functionality over time.

ATP content of the cells was measured using a kit ATP bioluminescent somatic cell assay kit (Sigma-Aldrich) following the manufacturer's instructions. 20 µl of assay reagent was added to each well and shaken for one minute. 100 µl from each well was then plated onto a 96-well flat-bottomed plate and left to incubate for five minutes. Luminescence was measured using VarioSkan flash plate reader (Thermo-Fisher).

4.2.1.6. Bradford Assay

A Bradford assay was used to measure the total protein concentrations in the cells. Within our experiments, the Bradford protein quantification assay was performed in accordance with the manufacturer's instructions (BioRad, UK). Six 500 µg/ml dilutions of a protein standard (bovine serum albumin, BSA) were prepared. Duplicates of each protein standard and sample (20 µl) were pipetted into separate microtitre wells. Following this, 200 µl of diluted dye reagent (prepared by diluting 1 part dye reagent concentrate with 4 part DDI water) was added to each well. They were then incubated at room temperature for 10 minutes. A standard curve was constructed using BSA (Sigma-Aldrich) between the concentrations 0-1 mg/ml. Absorbance was measured at 595 nm.

4.2.1.7. Gene quantification– Western blotting

Western blotting is an experimental technique that can detect and quantify particular proteins among a mixture of proteins [141]. Firstly, all proteins are separated into their linear chains and then separated according to their size using a method called gel electrophoresis. Once separated, each protein is transferred onto a blotting

membrane. The membrane subsequently undergoes a blocking procedure which disables any non-specific reactions. A primary antibody is then used to incubate the membrane, which specifically binds to the protein of interest. Any unbound primary antibody is washed away and the membrane is then incubated with a secondary antibody. The secondary antibody not only specifically recognises and binds to the primary antibody, but is also linked to a reporter enzyme that produces colour or light, allowing for the detection of the protein of interest. We carried out Western Blotting analysis to detect levels of protein markers, actin and caspase.

Sample preparation

A master mix of lithium dodecyl sulfate (LDS) sample buffer and reducing agent (70:30) was created. Each sample was added to 5 μ l of the master mix, centrifuged for 30 seconds and placed in a heating block at 80°C for 10 minutes.

Loading Precast Gel

Precast 4-12% Bis-Tris gel was locked into the electrophoresis unit. 1X MOPS running buffer was created by diluting 50 ml of 20X MOPS in 950 ml dH₂O. The MOPS running buffer was mixed and poured into the central chamber of the electrophoresis unit. The outside of the chamber was filled with MOPS running buffer to approximately half of the total volume. The heated samples were pulse-spun and loaded into the appropriate wells of the gel. 5 μ l of protein marker were loaded in the first well whilst 10 μ l of protein marker were loaded into each sample well. The gel was then run for 10 minutes at 90V followed by 1 hour at 170V.

Electrophoretic transfer

1X transfer buffer was prepared by mixing 100 ml 10X transfer buffer + 200 ml MeOH + 700 ml dH₂O. Each gel had 3x sponges, 2x filter papers, 1x transfer buffer and 1x nitrocellulose membrane (GE Healthcare, Little Chalfont, UK). The plastic plates were opened using a spatula and the gel remained attached to one plate. The wells were cut away from the top of the gel, and the ridge cut away from the bottom of the gel. Pre-soaked filter paper and a nitrocellulose membrane were placed onto the gel, and the gel was carefully peeled off the plate with a spatula. The gel was placed onto the bench (filter-paper down) and the pre-soaked membrane was placed on top of the gel. An additional pre-soaked filter paper was placed on top of the membrane. The transfer cassette was then assembled and locked into the transfer chamber. Both the cassette and cool pack were inserted into

the transfer tank. The transfer tank was filled with 1x transfer buffer until almost full and stirred for 1 hour at 230mA.

Protein detection

The membrane was rinsed in 0.1% TBST and stained with ponceau red for 10 seconds. The membrane was then cut at 37 kDa marker for β -Actin (42 kDa) and for caspase 3 (35 kDa). The actin membrane was blocked in 10% milk blocking solution overnight whilst the caspase membrane was blocked for 1 hour. The primary caspase 3 antibody was then added at the required dilution of 1:2000 in 2% milk and incubated overnight at 4°C. The actin membrane was incubated in 10% milk blocking-solution on a rocker overnight at 4°C. The primary actin antibody was added at a dilution of 1:10,000 in 10% milk blocking-solution and the actin membrane was incubated on the rocker for 30 minutes at room temperature. Both actin and caspase 3 membranes were washed for 4 x 10 minutes in 0.1% T-TBS. The secondary antibody solutions were prepared (anti-mouse HRP at a dilution of 1:10,000 in 2% milk for actin and anti-rabbit HRP at a dilution of 1:10,000 in 2% milk for caspase 3) and added to the necessary membrane, each being incubated on the rocker for 1 hour at room temperature. Both membranes were then washed for 4 x 10 minutes in 0.1% T-TBS.

Visualisation of immune reactivity was completed using Enhanced Chemiluminescence Plus (ECP) (Perkin Elmer, MA, U.S.A) according to the manufacturer's instructions. Membranes were blotted dry and covered with a chemiluminescence reagent for 2 minutes. Membranes were then covered in Saran wrap and exposed to ECL film in a dark room for 1-15 minutes. All western images were scanned and analysed using Quantity One 1-D analysis software (Bio-Rad, Hercules, CA, USA).

4.2.1.8. microRNA quantification– qPCR

We aim to measure genetic changes over time in order to quantify protein markers. During Deoxyribonucleic acid (DNA) analysis, amplification is required in order for there to be enough DNA to provide a detectable signal for quantification [142]. Polymerase Chain Reaction (PCR) is a technique for amplifying DNA. Firstly, all double stranded DNA is “melted” into single strands, with each half representing a template for a new molecule. A polymerase enzyme then synthesises a complementary sequence of bases to each single strand of DNA. Adding small pieces of DNA, complimentary to the gene of interest (known as primers), prepares

the DNA sample, ensuring the polymerase can bind. This eventually results in a copy of the gene of interest. In order to quantify this reaction, probes (or fluorescently labelled DNA oligonucleotides) can be added to the mixture, which bind downstream to one of the primers and give a fluorescent signal during the reaction. As the number of gene copies increases, so too does the fluorescence. This real-time analysis is often referred to as “quantitative” PCR or qPCR.

Within our study, measuring the number of gene copies over time allowed for quantification of an apoptotic marker (Caspase) and a necrotic marker (miR-122).

miRNA extraction

For amplification to occur, miRNA must first be extracted and “melted” into single strands, with each half representing a molecule template for copying. All samples were stored in frozen conditions to ensure all DNA information remained intact and to ensure no enzymatic reactions could take place. Therefore, before miRNA extraction could occur, the samples were firstly defrosted and spun down. The amount of supernatant used for the extraction was up to 200 µl. 700 µl of Qiazol reagent was added and then mixed by vortexing. The mixture was incubated at room temperature for 3 minutes to ensure thorough nucleo-disruption. 140 µl of chloroform was added and the tubes were shaken vigorously by hand for 15 seconds, before being left at room temperature for a further 5 minutes. Once centrifuged at 12,000 rpm for 15 minutes at 4°C, the samples were allowed to equilibrate at room temperature. 350 µl of supernatant was transferred to RNeasy mini-spin columns which were then placed inside collection tubes and centrifuged at 8,000 rpm for 15 seconds at room temperature. The flow-through was then finally transferred to a 2 ml micro-tube.

Purification

Purification of RNA is central to any gene expression/regulation investigation; it is essential to reduce the chance of degradation which could dramatically affect the quality of the qPCR data. Once extracted from the samples, the miRNA was then able to be purified.

Within our experiments, in order to purify the miRNA, 450 µl of 100% ethanol was added to the previously collected flow-through and mixed by vortexing. Samples were applied (700 µl at a time) to RNeasy MinElute columns until there was no sample remaining. They were then centrifuged at 8,000 rpm for 15 seconds at room

temperature before the flow-through was discarded. 500 µl of RPE buffer was added to the MinElute column and then centrifuged at 8,000 rpm for 15 seconds at room temperature. The flow-through and collection tubes were then discarded. The column was then placed in a fresh collection tube ensuring MinElute column had no residual flow-through to avoid contamination. The columns were then centrifuged with their lids open at 8,000 rpm for 5 minutes to ensure they were completely dry. The MinElute column was placed into a 1.5 µl micro-tube and 14 µl of RNase-free water was pipetted onto the spin column membrane and centrifuged at 8,000 rpm for 1 minute at room temperature to elute the purified miRNA fraction.

Reverse Transcription

The reverse transcription (RT) stage is critical for accurate and sensitive RNA quantification [142]. It is at this point that a polymerase enzyme then synthesises a complementary sequence of bases to each single strand of RNA. Within our experiments, reverse transcription was performed using the TaqMan miRNA reverse transcription kit (Applied Biosystems) following manufacturer instructions. The primers (or small pieces of DNA complimentary to the gene of interest) used were miR-22 and U6 (endogenous control). A GeneAmp PCR9700 machine was used to carry out the analysis. 2 µl of purified miRNA was used to synthesise complementary DNA (cDNA) with a total reaction volume of 15 µl in a 96-well PCR plate which was held in an ice block. RT was performed via thermal cycling (30 min at 16°C, 30 min at 42°C, 5 min at 85°C and then held at 4°C).

miRNA PCR

Primers (or small pieces of DNA complimentary to the gene of interest) allow the polymerase enzymes to bind and form a copy of the gene. In PCR, probes are added to the mixture which provide a fluorescent signal during the reaction, so that the number of copies can be quantified. For the quantification to take place within our experiments, 1.33 µl of cDNA from the RT were used in duplicate, with the use of corresponding specific qPCR primers (Life Technologies) and Taqman PCR Master mix (Life Technology) according to manufacturer's instructions. 1.33 µl of cDNA was used and plated by a QIAgility (Qiagen) machine for high-precision automated PCR setup and analysed on a ViiA7 machine (Life Technologies). After undergoing thermal cycling (40 cycles of 15 sec at 95°C and 60 sec at 60°C), levels of miRNA were measured by the fluorescent signal produced from the Taqman assay probes.

4.2.1.9. Cell sorting– FACS

Flow cytometry is a powerful tool used for analysing complex cell populations within a short space of time [143]. The process works by suspending cells into a stream that creates a laminar flow, allowing each cell to pass an interrogation point, where a beam of monochromatic light intersects the cells. The directions in which the emitted light from the cells are given off allows the computer to conclude various things about the cell, with particular focus on its shape and size. A derivative of flow cytometry, fluorescence-activated cell sorting (FACS) allows a researcher to physically sort a heterogeneous mixture of cells into different populations due to the use of highly specific antibodies tagged with fluorescent dyes [144]. With this in mind, we used FACS to sort our cell population with regard to the mode of cell death (apoptosis or necrosis).

The following method was adapted from eBiosciences staining protocol. Supernatant, including cell debris, was harvested and transferred into 1.5 ml EP tubes on ice. The cells were then rinsed with PBS to remove serum and the adherent cells were harvested by using trypsin in the same 1.5 ml EP tube. The cells were then spun twice, firstly for 5 minutes at 3000 rpm and secondly for 8 seconds at 12,000 rpm. The supernatant was removed and the cell pellets were put on ice, re-suspended in 200 µl of cold PBS and transferred to round bottom 96-well plates. The 96-well plates were spun at 2300 rpm for 3 minutes at 4°C. The PBS was then discarded and the cell pellets re-suspended in cold PBS. The spinning process was repeated but once the PBS was discarded, the cells were this time re-suspended in 2.5 µl annexin-FITC per well. The cells were mixed with a multichannel pipette and incubated in the dark, whilst shaking, for 10 minutes at room temperature. Following incubation, the cells were washed in 200 µl of binding buffer and centrifuged at 2300 rpm for 3 minutes at 4°C. The supernatant was discarded and the cell pellets were re-suspended in 190 µl of binding buffer. 200 µl of binding buffer was added to the FACS tube and then 190 µl of cell suspension was also transferred. The samples were kept on ice. Just prior to acquiring the samples in the FACS, 10 µl of propidium iodide-A (PI) per tube was added. The FACS machine then acquired information regarding cell survival; flow cytometry was used to sort any dead cells with regard to their mode of cell death (apoptosis vs necrosis).

4.2.2. Statistical methods

4.2.2.1. Multiple linear regression: predicting cell viability

To test the use of biomarkers in predicting cell viability, multiple linear regression analysis was applied. This method uses more than one predictor variable (in this instance, multiple biomarkers) to predict an outcome variable (in this instance, cell viability). In the linear regression setting, the outcome variable must be continuous rather than nominal/categorical. The observed data (Y , representing cell viability) is a linear function of the predictor variables (X_1, X_2, X_3, X_4 and X_5 , representing biomarker concentrations APAP, ALT, HMGB1, Full K18 and Fragmented K18, respectively). If all of the predictor variables were significant, and therefore included in the model, the relationship between predictor and outcome variables can be represented using a generalised linear model:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5. \quad (4-1)$$

The objective of multiple linear regression is to optimise the parameters, β_i ($i = 0 \dots 5$), so that the model provides the best fit to the observed (cell viability) data. Since a generalised linear model is assumed, ordinary least squares is a sufficient fitting method (minimising the distance from the model output and the data points [145]). In order to apply multiple linear regression analysis, there are several data assumptions that must be tested in order to ensure the method is appropriate for use and therefore the results are interpretable. The key assumptions are:

- *The outcome variable is measured on a continuous level.* This assumption holds since cell viability is measured on a continuous level.
- *One or more of the predictor variables are continuous, ordinal or nominal.* The predictor variables in our model are represented by biomarker concentrations (APAP, ALT, HMGB1, Full K18 and Fragmented K18), all of which are measured on a continuous level.
- *Observations must be independent.* The Durbin-Watson statistic was used to test independence of the observations. This statistic tests for autocorrelation in the residuals from a statistical regression analysis. The test statistic created by Durbin and Watson summarises how each of the residuals vary in comparison to one another and the formula can be found in their study [146]. For independence to be present, the Durbin-Watson value can range between 0-4 but should be approximately 2 to indicate that there are no correlations between residuals (values from $[0,2)$ indicate positive autocorrelation and values more than $(2,4]$ indicate negative autocorrelation).

The Durbin-Watson statistic resulting from this dataset is 2.048 and therefore the independence of observations can be assumed.

- *Multi-collinearity should not be present between any predictor variables.* This assumption was tested using the Variance Inflation Factor (VIF) test. The VIF is indicative of the amount of variance an estimated regression coefficient is increased by due to multicollinearity. Since the VIF is representative of increased variance, the square root of this result would be indicative of the increased standard error. As an example, a variable with $VIF = 16$ has a standard error 4 times larger than it would be if the variable was uncorrelated with the other variables. A recommended reasonable range for the VIF is [1,10]. Predictor variables were tested and remained within this range (all values within [1,3]). Therefore, multicollinearity does not exist between the predictor variables and the assumption is valid.
- *Outliers/highly influential points should not be present in the observed data.* No outliers were identified.
- *Residuals should be approximately normally distributed.* A histogram and normal P-P plot of the standardized residuals were inspected, and the data was found to be approximately normally-distributed.

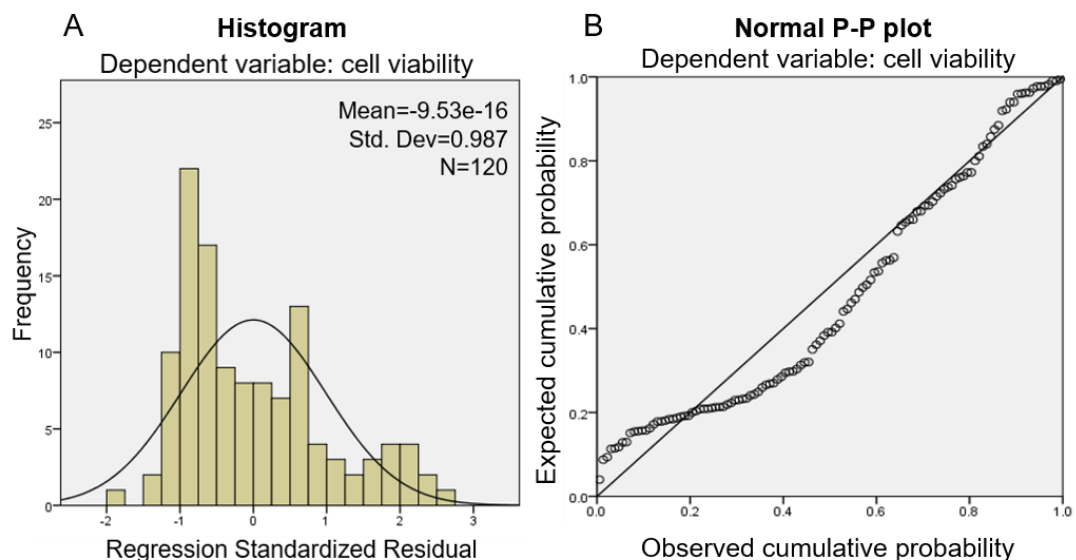


Figure 4-1: Data diagnostic plots. (A) Histogram – Visualisation of how the standardized residuals are distributed. For data to be diagnosed as normally distributed, the mean value should be approximately 0 and the standard deviation should be approximately 1. (B) Normal P-P plot of regression standardised residual. The cumulative probability expected from the model is plotted against the observed cumulative probability. For the data to be diagnosed as approximately normally distributed, most results (circles) should lie along the diagonal line.

Various regression models were created for subsets of the dataset (no antidote, 1mM NAC, 2mM NAC, 3mM NAC, 4mM NAC, 5mM NAC) and then a regression model was created for the whole dataset

A forward stepwise approach was used for each regression model. This means that the method (employed in SPSS) creates iterative multiple linear regression models. At each iteration, only the most significant predictor is included. An additional predictor is only included in the model if it provides improvement to the overall significance of the model. The resultant predictive biomarkers for each model created can be found in Table 4-1.

4.2.3. Mathematical modelling

4.2.3.1. Formulation

We extended the PKPD framework described in Chapters 2 and 3 to account for cell functionality following an APAP dose. We employed a mathematical model defined by Remien et al. [73] which describes how functional and damaged cells change over time. In our model, we assumed that functional cells could be represented by viable cells, and therefore optimised the parameters against the cell viability data obtained from the experiments I carried out with my IVTS fellowship prize.

The model for functional and damaged hepatocytes defined by Remien et al. is given by

$$\frac{dH}{dt} = rH \left(1 - \frac{H + Z}{H_{max}} \right) - \eta[N]H, \quad (4-2)$$

$$\frac{dZ}{dt} = \eta[N]H - \delta_z Z, \quad (4-3)$$

where H is the number of functional hepatocytes, Z is the number of damaged hepatocytes, r is the functional hepatocyte regeneration rate (h^{-1}), H_{max} is the maximum number of hepatocytes, η is the functional hepatocyte damage rate ($l/\mu\text{mol}/h$), $[N]$ is the NAPQI concentration ($\mu\text{mol}/l$) and δ_z is the damaged hepatocyte lysis rate (h^{-1}). Since our experimental data is expressed in terms of proportional cell viability, we set $H_{max} = 1$ and effectively rescale such that H becomes the proportional number of functional hepatocytes and Z becomes the proportional number of damaged hepatocytes relative to the maximum. In Chapter

2, we described how the dynamics of NAPQI are assumed to be short-lived and the associated reactions are rapid on the time-scale of APAP depletion. In the model framework described in Chapters 2 and 3, we assume that NAPQI is at a quasi-steady state on the time-scale of APAP PK. We re-introduce our NAPQI model here in equation (4-4):

$$\frac{d[N]}{dt} = \xi k_{el}[C_c] - k_G[N][gsh] - k_p[N], \quad (4-4)$$

where ξ is the proportion of CYP-activated APAP that is transformed into NAPQI, k_{el} (h^{-1}) is the total rate of APAP elimination, k_G ($l/\mu mol/h$) is the decay rate of GSH due to binding with NAPQI, and k_p (h^{-1}) is the rate at which NAPQI binds to other non-GSH proteins. We can derive the expression for NAPQI concentration by making use of the quasi-steady state approximation, $d[N]/dt = 0$. By substitution into equation (4-4), it follows that

$$N = \frac{\xi k_{el}[C_c]}{k_G[gsh] + k_p}. \quad (4-5)$$

In vivo, hepatocytes have the ability to regenerate, whereas in the *in vitro* environment, they do not. Our aim was to create a mathematical model describing cell functionality following APAP dose, for which the parameters were representative of the *in vitro* scenario. Bearing this in mind, we adjusted Remien's *in vivo* functional hepatocyte model (equation (4-5)) to result in the model shown in equation (4-6):

$$\frac{dH}{dt} = -\eta[N]H. \quad (4-6)$$

The dynamics of functional hepatocytes *in vitro* could now be described by the model defined in equation (4-6). The dynamics for damaged hepatocytes *in vitro* could be modelled in the same way as the *in vivo* scenario (equation (4-3)).

As well as modelling the effects of APAP on cell functionality, it is also essential to model the effects of antidote (NAC) on cell functionality. The distribution of NAC throughout the body can be represented mathematically by a 3-compartment model [147], with one (central) compartment consisting of well-perfused tissues and two peripheral compartments that contain poorly perfused tissues such as fats and blood binding proteins for example. Whilst APAP quickly distributes to well perfused tissues (central compartment) and secondarily to poorly perfused tissues (represented as one peripheral compartment), the PK of NAC has been shown to require an additional peripheral compartment as physiological which accounts for

different affinity sequestrations (to peripheral tissues, binding proteins etc.). In line with these findings, we employed a 3-compartment model previously described by Shen et al. [148] (Figure 4-2).

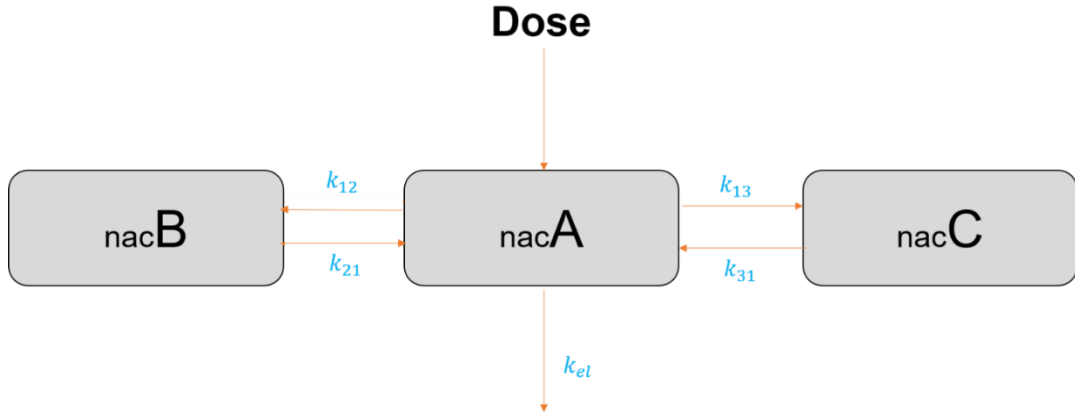


Figure 4-2: Schematic of N'acetylcysteine (NAC) dynamical model described by Shen et al. [148]. A dose of NAC is provided and instantaneously reaches the first compartment (*nacA*). There are then continual rates of transfer to and from other compartments, *nacB* (k_{12} and k_{21}) and *nacC* (k_{13} and k_{31}). NAC is eliminated at rate k_{el} from the first compartment (*nacA*).

The model schematic portrayed in Figure 4-2 can be mathematically described as follows,

$$\frac{d[N_A]}{dt} = \frac{NDose}{V_{NA}} + k_{N21}[N_B] \left(\frac{V_{NB}}{V_{NA}} \right) + k_{N31}[N_C] \left(\frac{V_{NC}}{V_{NA}} \right) - (k_{N12} + k_{N13} + k_{Nel})[N_A], \quad (4-7)$$

$$\frac{d[N_B]}{dt} = k_{N12}[N_A] \left(\frac{V_{NA}}{V_{NB}} \right) - k_{N21}[N_B], \quad (4-8)$$

$$\frac{d[N_C]}{dt} = k_{N13}[N_A] \left(\frac{V_{NA}}{V_{NC}} \right) - k_{N31}[N_C], \quad (4-9)$$

where $[N_A]$ is the total NAC concentration in the central compartment ($\mu\text{mol/l}$), $[N_B]$ and $[N_C]$ are the total NAC concentration in either peripheral compartment B or peripheral compartment C ($\mu\text{mol/l}$), $NDose$ is the initial dose of NAC provided (μmol), V_{NA} is the theoretical volume of the central compartment (l/kg), V_{NB} and V_{NC} are the theoretical volumes of peripheral compartments B or C respectively (l/kg), k_{N12} is the rate of transfer of NAC from the central compartment to peripheral compartment B (h^{-1}), k_{N21} is the rate of transfer of NAC from peripheral compartment B to the central compartment (h^{-1}), k_{N13} is the rate of transfer of NAC from the central

compartment to peripheral compartment C (h^{-1}), k_{N31} is the rate of transfer of NAC from peripheral compartment C to the central compartment (h^{-1}), and k_{Nel} represents the overall elimination rate of NAC (summation of excretion and metabolism processes) (h^{-1}).

These equations are applicable when a bolus dose of NAC is assumed, with the dose amount being converted to a concentration as described in equation (4-7) but redefined here for completeness,

$$NAC\ concentration_{bolus} = \frac{NDose}{V_{NA}}. \quad (4-10)$$

However, NAC regimens are typically administered via intravenous infusion, and therefore an infusion rate k_{inf} ($\mu\text{mol/kg/h}$) needed to be defined in the model based upon both dose amount and infusion time. For example, if a dose of 50 mg/kg is administered over 4 hours, $k_{inf} = 12.5$ mg/kg/h. In this scenario the infusion parameter, k_{inf} , replaces the bolus parameter, $NDose$, such that the initial concentration term of equation (4-7) becomes,

$$NAC\ concentration_{IVinf} = \frac{k_{inf}}{V_{NA}}. \quad (4-11)$$

For the conventional NAC regimen (initial administration of 150 mg/kg over 1 hour, followed by 50mg/kg over 4 hours, followed by a final infusion of 100 mg/kg over 16 hours) the infusion rate (k_{inf} , mg/kg/h) was defined as follows:

$$k_{inf} = \begin{cases} 150 & 0 \leq t < 1 \\ \frac{50}{4} & 1 \leq t < 5 \\ \frac{100}{16} & 5 \leq t < 21 \end{cases} \quad (4-12)$$

For the shortened NAC (SNAP) regimen (initial administration of 100 mg/kg over 2 hours, followed by 200 mg/kg over 10 hours, the infusion rate (k_{inf} , mg/kg/h) was defined as follows:

$$k_{inf} = \begin{cases} \frac{100}{2} & 0 \leq t < 2 \\ \frac{200}{10} & 2 \leq t < 12 \end{cases} \quad (4-13)$$

Regardless of whether administered as a bolus dose or intravenous dose, NAC is an effective pre-cursor of GSH [149], provided in the APAP overdose case to increase GSH levels and subsequently increase the rate of detoxification of NAPQI.

With this in mind, we amended the GSH model described in Chapters 2 and 3, assuming that GSH production is a function of NAC concentration in the central compartment:

$$\frac{d[\text{gsh}]}{dt} = k_o(\text{gsh}_o - [\text{gsh}]) + \zeta[\text{N}_A] - \frac{\xi k_{el}[\text{C}_c][\text{gsh}]}{[\text{gsh}] + k_{pr}} \quad (4-14)$$

Where gsh is the GSH concentration ($\mu\text{mol/l}$), k_o is the basal removal rate of GSH (including background usage) (h^{-1}), gsh_o is the baseline concentration of GSH in the APAP-free steady state ($\mu\text{mol/l}$), ζ is the NAC-induced production rate of GSH (h^{-1}), $[\text{N}_A]$ is the NAC concentration in compartment nacA ($\mu\text{mol/l}$), ξ is the proportion of eliminated APAP that is transformed into NAPQI, k_{el} is the total elimination rate of APAP (h^{-1}), $[\text{C}_c]$ is the APAP concentration and k_{pr} is the ratio of NAPQI forming other protein adducts relative to being detoxified by GSH. It should be noted that NAC and GSH will have different apparent volumes of distribution because of their different chemistries.

Parameters for the all of the APAP PK and some of the PD (ALT, HMGB1, full and fragmented K18) elements of the model remained fixed at the optimised values used in our framework in Chapter 3. Parameters for the NAC PK model were equal to the values provided by Shen et al. [148]. Since we aimed to provide parameter values that could describe cell functionality changes following APAP/NAC doses, and provide profiles replicative of the cell viability data found from our experiments, we optimised the following parameters: from the cell functionality component of the model, the functional hepatocyte damage rate (η) and the damaged hepatocyte lysis rate (δ_z); from the NAPQI component of the model, the GSH decay due to binding with NAPQI rate (k_G); and from the GSH component of the model, the NAC-induced production rate of GSH (ζ). All parameters were optimised by fitting the aforementioned (cell functionality, NAPQI and GSH) models to the cell viability time-course data from my experiments (ATP assay).

Parameters within the cell functionality component of our model were optimised by fitting to our *in vitro* cell viability time-course data. When using the model to simulate for the *in vivo* scenario, we assumed these parameters (found from fitting to *in vitro* data) would be similar in the *in vivo* scenario. However, *in vivo*, hepatocytes have the ability to regenerate, and therefore we did not assume that the cell functionality model for the *in vivo* scenario should be completely identical to the *in vitro* scenario. The *in vivo* model structure for functional hepatocytes over time was identical to

equation (4-2). The functional hepatocyte damage rate, η , was optimised from our *in vitro* data, but since no additional *in vivo* data was available, the functional hepatocyte regeneration rate, r , remained fixed as the value provided in Remien's *in vivo* study [73]. The *in vivo* model employed for damaged hepatocytes is described by equation (4-3), with the functional hepatocyte damage rate, η , and the damaged hepatocyte lysis rate, δ_z , optimised based on our *in vitro* study.

4.2.3.2. Simulation

The model was simulated for various NAC doses provided at 1h post-APAP dose to visualise any effects on the thresholds (APAP dose and time since administration) required for our probability of liver injury toxicity measure. This toxicity measure was defined by our previous framework, details of which are described in Chapter 3, section 3.2.9. Using the extended framework defined within this chapter, we also assessed how amending the dose of NAC impacted the thresholds for an additional toxicity measure, APAP-induced loss of cell functionality. Loss of cell functionality was defined as any simulated proportion of functional cells lower than 20%; Acute liver failure is thought to take place beyond 80% hepatocyte loss [150]. The exact threshold for when a liver should be deemed "healthy" or "unhealthy" is currently unknown and of particular clinical interest [151].

4.3. Results

4.3.1. Wet-lab experiments

4.3.1.1. Microscopic images of hepatocyte culture

Microscopic images of the hepatocyte culture were taken to visualise the viability of the cells. Images were taken at 24h, 48h and 72h (Figure 4-3). A lower and higher dose of APAP was compared (1 mM and 20 mM respectively), as well as the effects of providing antidote treatment 1 hour post-APAP dose. For the control images (No APAP or NAC) the cells became more confluent as time progressed, with little indication of any cell death.

The smaller dose of APAP (1 mM) was comparable to the control images. For the higher dose of APAP (20 mM), however, negative effects upon the hepatocytes could be clearly be visualised. At the 24 h time-point, dead hepatocytes were visible and identified by dark brown masses. As time progressed, so too did loss of

hepatocyte functionality, with the 72 h time-point providing a clear visualisation of hepatocyte death. Providing antidote treatment (5 mM) to the 1 mM APAP case provided no real positive/negative effect, i.e. both the 1 mM APAP case and the 1mM APAP + 5 mM NAC case were visually comparable to the control case. Providing antidote treatment to the 20 mM APAP dosed cells did, however, provide a positive visual effect on the cells. At the 24 h and 48 h time-points in the 20 mM APAP dose without antidote case, the cells were already dying/dead. In comparison, at these time-points the cells treated with antidote were visually healthy and comparable to the 1 mM APAP dose and control cases. Although the cells were visibly dying at 72 h in the 20 mM APAP + 5 mM NAC case, the amount of time taken for the cells to die was longer in the antidote case (72 h compared with 24 h in the no antidote case).

4.3.1.2. Cell viability – ATP assay

In order to quantify cell viability, we measured the intracellular ATP and calculated the changes relative to baseline ATP. We see in (Figure 4-4) how varying doses of APAP and NAC affected the cell viability of each case as time progressed.

Figure 4-4(F) shows that the average proportion of viable cells decreased over time in all cases (\pm NAC). At 6 h, the NAC treated data had 80% of the hepatocytes in a viable state, reducing to approximately 10% at 72 h. For the hepatocytes which were not treated with the antidote, approximately 70% were viable at 6 h, reducing to around 10% at 72 h. At all time-points, the cells which had 0 mM NAC treatment provided (shown by the blue line) had reduced viability when compared to the cells who were given a dose of NAC (green line). At 72 h, regardless of whether the cells were provided with NAC treatment, cell viability was low, around 10%. This result was consistent with that found by visualising the microscopic hepatocytes (Figure 4-3); NAC appears to slow down the onset of cell death, but not completely prevent it.

At the 6-hour time-point (Figure 4-4(A)), cell viability was reasonably high for the 1mM and 5mM APAP dose cases, averaging around 80% across all NAC doses. Providing 1 mM and 3 mM NAC in both of these cases seemed to protect the cells and allowed the cells to remain at full viability (100%) 6 h post-dose. For the higher APAP doses, 10 mM and 20 mM, viability was reduced to an average of approximately 60% and 40% respectively. In both of these instances, NAC did not seem to provide any clear protection, regardless of the dose.

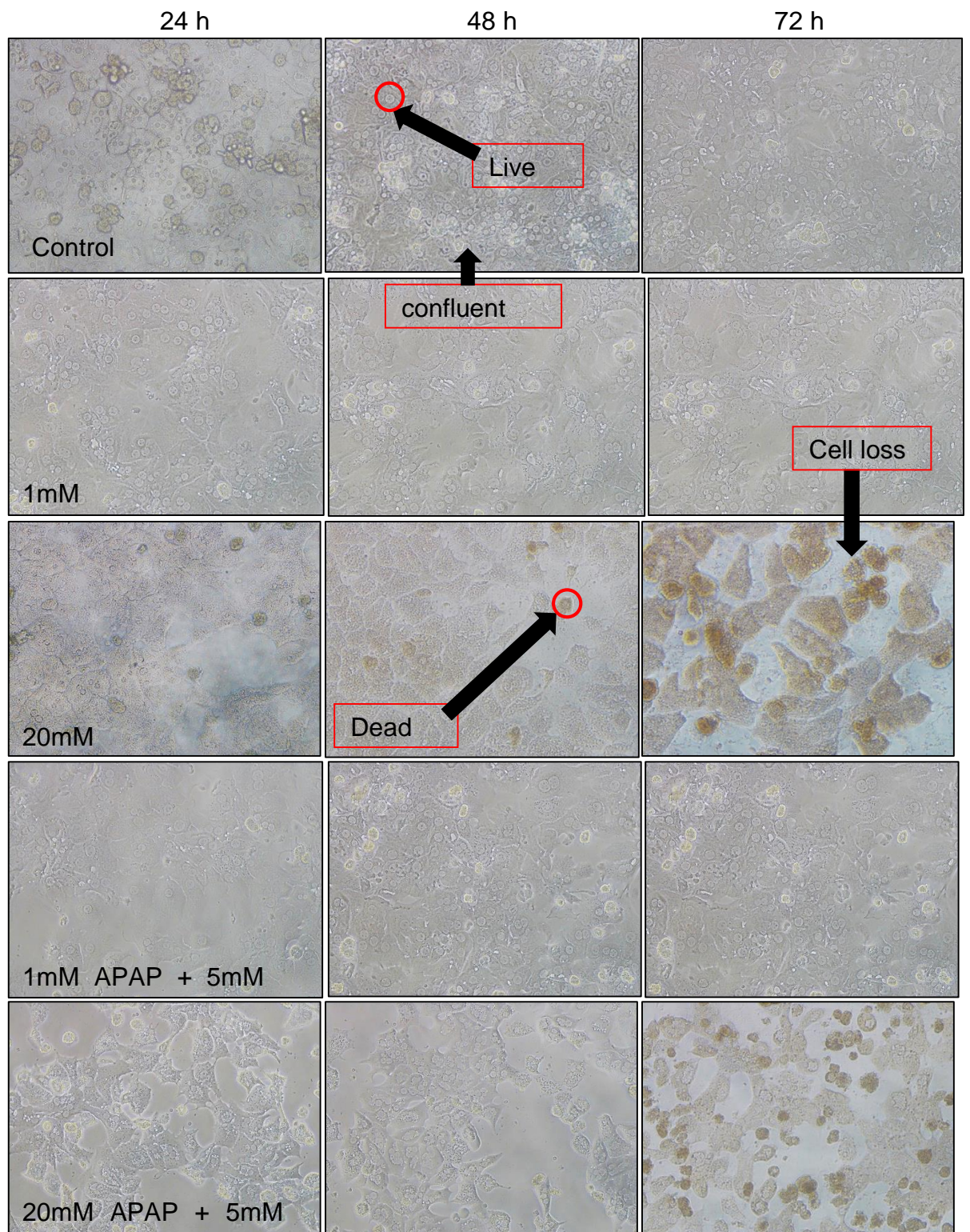


Figure 4-3. Microscopic images at various time-points for various doses of APAP/NAC, visualising the effects of both the toxin (APAP) and antidote (NAC) on primary mouse hepatocyte viability. *Live cells are indicated by pale coloured shapes with visible nuclei. Dead cells are represented by shapes of a much darker brown shade, with less visible nuclei. Images which show a high population of condensed live cells are representative of cell confluence, whilst images with distinct empty spaces represent wells where cell loss has occurred.*

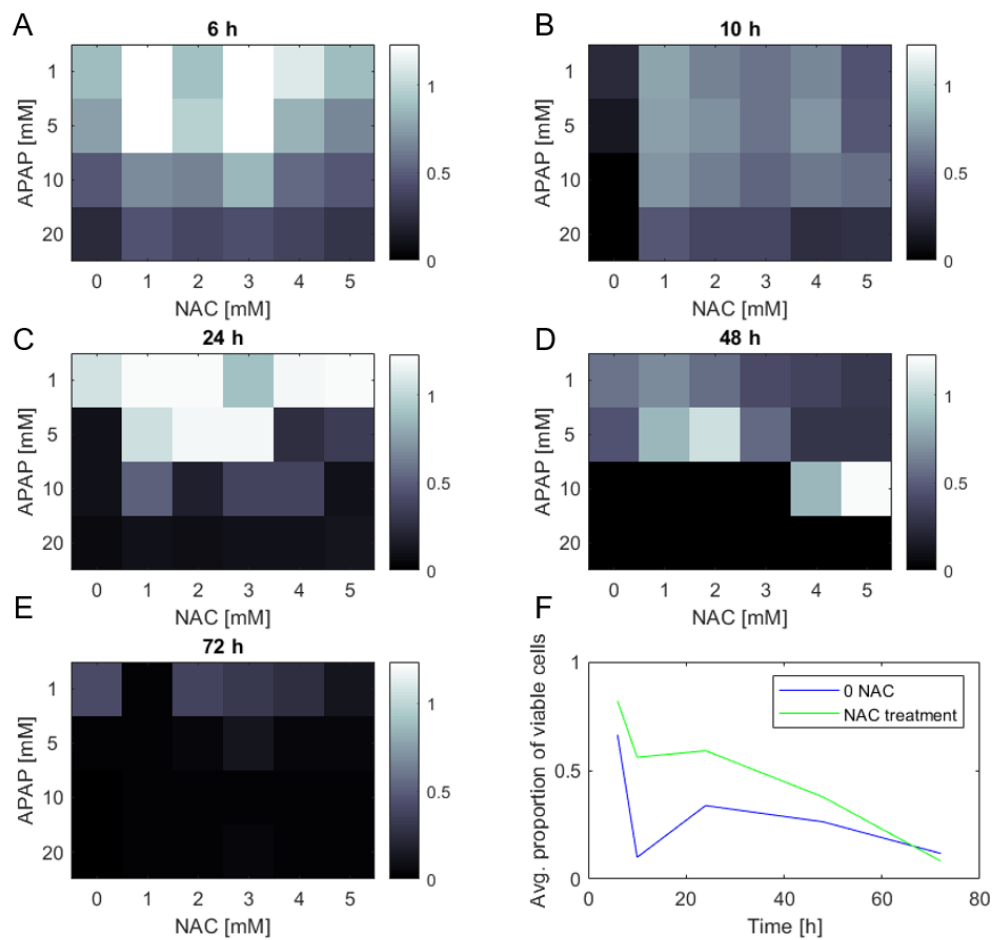


Figure 4-4: The effect of varying APAP/NAC doses on hepatocyte viability. The primary hepatocytes were dosed with four APAP doses, [1, 5, 10, 20] mM. Each scenario was given a varying dose of NAC [0,1,2,3,4,5] mM 1 hour post-APAP dose. Proportion of viable cells for each scenario was measured and represented as a score [0,1]: 1 for 100% viable (white); 0 for 0% viable (black). (A) shows the effect of varying APAP/NAC doses on hepatocyte viability 6 hours post-APAP dose, (B) shows the effects 10 hours post-APAP dose, (C) shows the effects 24 hours post-APAP dose, (D) shows the effects 48 hours post-APAP dose, (E) shows the effects 72 hours post-APAP dose and (F) shows the average proportion of viable cells across the whole time course, for scenarios given no NAC (green), and for scenarios provided with NAC treatment (blue).

At the 10 h time point (Figure 4-4(B)), a general reduction in cell viability could be clearly seen. At this time-point, no cases were completely viable (viability in all cases is less than 1). However, NAC did seem to have some sort of protective effect. For the [1, 5, 10] mM APAP dose cases treated with NAC, cell viability on average was approximately 70%. For the highest APAP dose (20 mM) cell viability was approximately 40% across all NAC dosed cases. Providing no NAC in the 1 mM and 5 mM APAP cases reduced viability from approximately 70% in the NAC dosed case to approximately 20%. For the 10 mM and 20 mM APAP cases, cell viability was completely reduced to 0% if no NAC was provided.

At 24 hours (Figure 4-4(C)), cell viability was reduced to 0% for the 5 mM, 10 mM and 20 mM APAP cases which were not given NAC. For the highest APAP dose (20 mM), regardless of the amount of NAC provided, cell viability remained at 0%. For the 10 mM APAP case, cell viability was measured to be up to 50% at 24 h if 3 mM NAC treatment was provided. Providing the same dose of NAC in the 1 mM APAP case allowed cells to be measured as completely viable (100%) at 24 h, this was an increase of 40% from the 10 h time-point. For all of the cells dosed with 1 mM APAP, whether dosed with NAC or not, cell viability had recovered to above 90%, suggesting that for the 1 mM dose of APAP, NAC may not be required for ATP recovery to occur. However, we did begin to see the protective effects of NAC for other cases. There was no improved viability at 24 h for the higher doses of APAP ([10, 20] mM) without NAC protection; for these doses, viability was 0% at 10 h and remained 0% at 24 h. For the 5 mM APAP case, providing NAC between 1-3 mM improved the viability almost completely at 24 h, increasing the viability from an average of 70% at 10 h to above 90% at 24 h. However, higher doses of NAC did not protect the viability of the cells at 24 h as viability was reduced from approximately 70% at 10 h to 20% at 24 h. For the 10 mM APAP case, NAC did not protect the cells at this time-point; for the NAC-dosed cases, viability reduced from an average of 70% at 10 h to approximately 30% at 24 h. Interestingly, providing the highest NAC dose (5 mM) resulted in 0% viability at 24 h, an identical result to cell viability at 24 h following no NAC dose at all.

At the 48-hour time point, viability started to generally decrease again (Figure 4-4(D)). The positive effects of NAC for the 1 mM APAP case at 24-hours were no longer seen here. If a NAC dose was provided, viability was measured at 100% at 24 h, whereas at 48 h the maximum measured cell viability was approximately 70%. Although there was generally a decrease in viability for the 5 mM APAP case (an

average decrease of around 20% from the 24 h measurements), at 24 h, the cells were measured to be completely viable. Subsequently, a decrease to 80% cell viability at 48 h, when compared to a measurement of 40% viability for the case without antidote, suggests that the protective effects of NAC are still seen at 48-hours for the 5 mM APAP case. For the 1 mM and 5 mM APAP cases, cell recovery was identified (40% and 20% increase in cell viability respectively) between 10 and 24 h, but this recovery was not seen for the 10 mM and 20 mM APAP cases. At 48 hours, we began to see recovery for the 10 mM APAP case (increase of approximately 80% cell viability from 24 h to 48 h). This increase was only seen for cells provided with the highest doses of NAC, [4, 5] mM, and this late recovery suggests it takes longer for the positive effects of NAC to be seen as the dose of APAP increases. This could correspond to the severe depletion of GSH stores in high APAP dose cases that would take longer to restore. For the 20 mM case, cell viability remained at 0%, regardless of the NAC dose given.

At the 72-hour time point, the protective effects of NAC could not be seen (Figure 4-4(E)). Cell viability completely reduced for the 10 mM and 20 mM APAP dose cases. Whilst providing a high NAC dose allowed cells to be measured as 100% viable at 48 h, at 72 h, regardless of whether NAC was provided or not, there were 0% viable cells for the 10 mM and 20 mM APAP cases. For the 5 mM APAP case, all NAC-dosed cases had cell viability measurements less than 20%, with most at 0%. Even for the lowest APAP case (1 mM), cell viability reduced by an average of 40% from 48 h to 72 h, regardless of the NAC dose initially provided.

4.3.1.3. Cell sorting - FACS

FACS analysis was carried out to investigate the mode of cell death (apoptosis or necrosis (Figure 4-5)). The clearest result from this analysis was that in almost all APAP dose cases, providing no NAC treatment resulted in the percentage of dead cells being greater than live cells. This result is reflected by the summation of the black and white (necrotic and apoptotic regions) in Figure 4-5, for the 0 mM NAC dose in all APAP cases being greater than 50%. Comparing the amount of necrosis against apoptosis when no NAC was provided, for the 1 mM APAP case (Figure 4-5(A)), 42% of cells were necrotic whilst 8% were apoptotic. For the 5 mM APAP case (Figure 4-5(B)), 54% were necrotic whilst 2% were apoptotic. For the 10 mM and 20 mM APAP case, (Figure 4-5(C-D)) 56% of cells were necrotic whilst 0% were

apoptotic. The amount of necrotic cells dominated the amount of apoptotic cells in each of these cases.

For the 1 mM APAP dose case with no NAC, approximately half of the cells died with around 42% of the cells being necrotic and 8% apoptotic. Providing any dose of NAC and comparing this to the case with no NAC, showed an increased amount of apoptosis and decreased amount of necrosis: 2 mM NAC decreased the necrotic percentage to 20% and increased the apoptotic percentage to 10%, 3 mM NAC decreased the necrotic percentage to 20% but increased the apoptotic percentage to 22%, 4 mM NAC decreased the necrotic percentage to 22% but increased the apoptotic percentage to 12% and 5 mM NAC decreased the necrotic percentage to 8% but increased the apoptotic percentage to 22%. When providing 2 mM, 3 mM and 4 mM NAC, the proportion of apoptosis and necrosis were similar. Providing the highest dose of NAC (5 mM) resulted in apoptosis dominating with reduced levels of necrosis.

For the 5 mM APAP case, when no NAC was provided, there was a greater amount of necrosis (52% versus 3% apoptosis) than seen in the 1 mM APAP case (40% necrosis versus 10% apoptosis). Providing the antidote in this 5 mM APAP case caused a distinctive switch to apoptosis for all doses of NAC. Specifically, providing 2 mM NAC resulted in an almost negligible amount of necrosis (2%). The cases dosed with 3 mM and 4 mM NAC provided similar results to one another – approximately 10% of the cells were necrotic whilst approximately 25% were apoptotic. For the highest dose of NAC however, whilst apoptosis still dominated (25% of cells), it was to a lesser extent since the amount of necrosis had increased from 10% to 15%.

Providing no antidote in the 10 mM APAP case did not allow any apoptosis to occur, with around 57% of the cells being necrotic. Providing 2, 3 or 4 mM NAC allowed a decreased percentage of necrotic cells (10%, 15% and 17% respectively), and increased percentage of apoptotic cells (22%, 22% and 25% respectively). Apoptosis remained the main form of cell death until the 5 mM NAC dose, where the percentage of apoptosis was very similar to necrosis, at approximately 30% each.

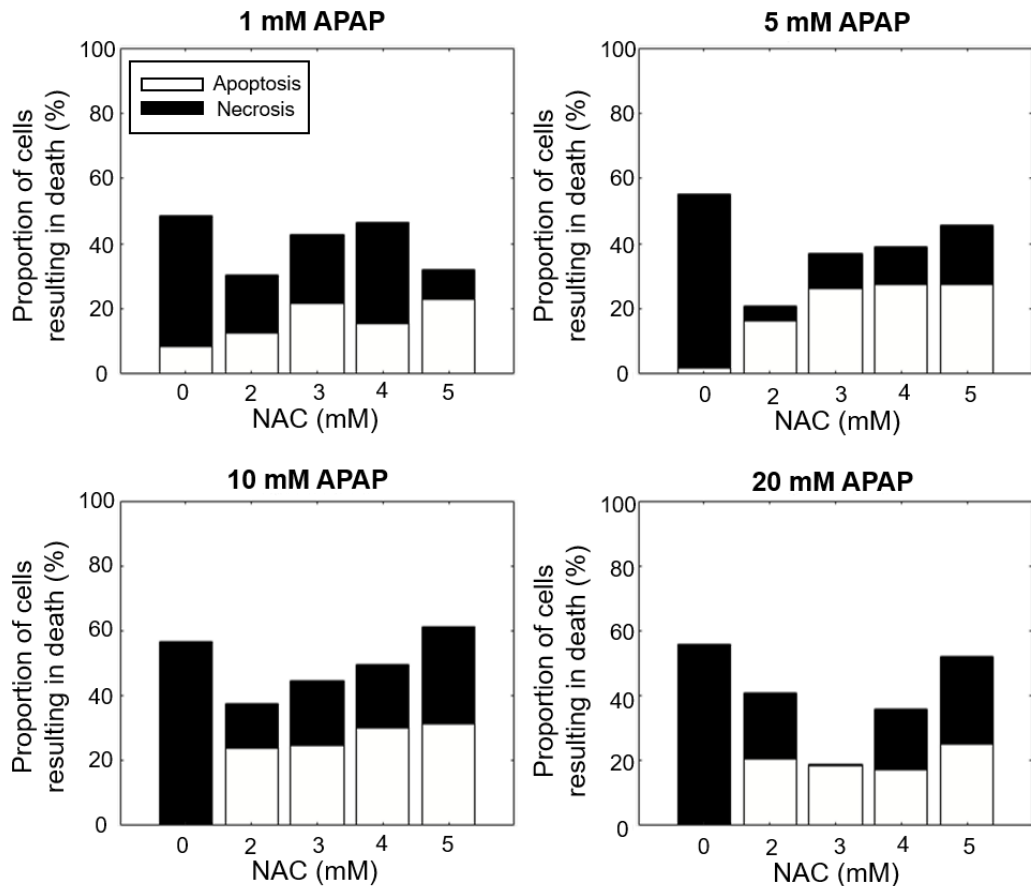


Figure 4-5: Quantifying the mode of cell death 24 hours post-APAP dose through FACS analysis. The mode of cell death (apoptosis vs necrosis) was compared for four APAP doses [1, 5, 10, 20] mM, at 24 hours post-dose. Additionally, the effect of NAC on this mode of cell death was analysed. Varying doses of NAC were applied and compared, [0, 2, 3, 4, 5] mM, 1 hour post-APAP dose. Apoptotic cells are indicated by white regions on each bar, whilst necrotic cells are represented by black regions on each bar.

Providing no antidote in the 20 mM case provided similar results to the 10 mM case with no antidote; around 57% cells were necrotic with no visible apoptosis. When providing 2 mM NAC in the 20 mM APAP case, the apoptotic and necrotic forms of cell death were identical (20% each). Interestingly, providing a 3 mM NAC dose seemed to completely reduce the number of necrotic cells, whilst 20% of the cells remained apoptotic. Increasing the NAC dose beyond 4 mM in the 20mM APAP case, however, did result in necrotic cells, however, the apoptotic dominance was not apparent; both apoptosis and necrosis levels were similar (20% each in the 4 mM NAC case and 25% each in the 5 mM NAC case).

4.3.1.4. Gene quantification - qPCR

qPCR was carried out in order to quantify and compare the amount of an apoptotic marker (caspase) and a necrotic marker (miR-122) following two different doses of APAP (5 mM and 20 mM), and how various doses of NAC ([0,1,2,3,4,5] mM) affected this subsequent output.

Figure 4-6(A) clearly shows that apoptosis levels were higher for the 5 mM APAP case, and lower for the 20 mM APAP case. This was expected, since a higher dose of APAP is likely to result in relatively higher levels of necrosis, and therefore lower levels of apoptosis. NAC treatment did not cause a distinctive difference to the apoptosis levels except for the 20 mM APAP dose case, where NAC > 4 mM seemed to increase the amount of apoptosis.

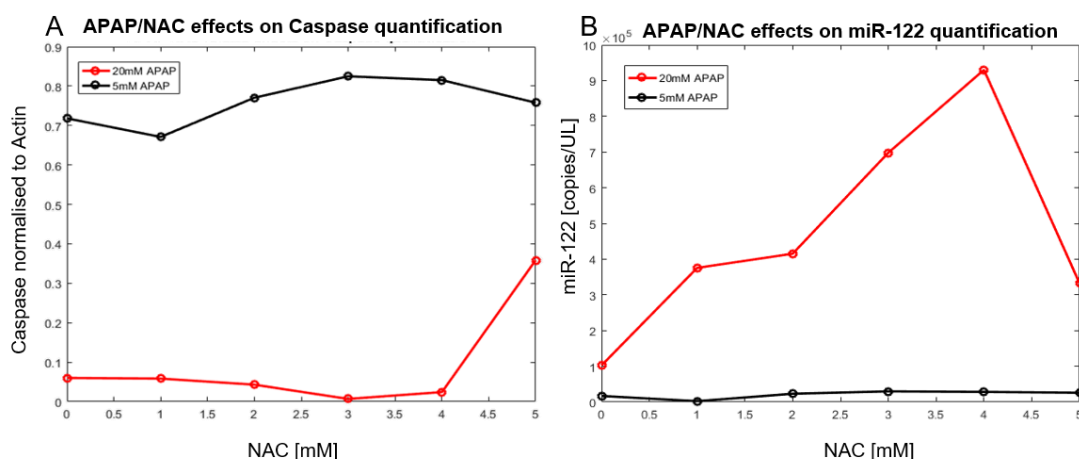


Figure 4-6: Comparing mode of cell death through quantification of apoptotic and necrotic markers. The effects of 5 mM and 20 mM doses of APAP (black and red lines respectively) in combination with various doses of NAC, [0,1,2,3,4,5] mM (1 hour post-APAP dose) were analysed at 24 hours post APAP dose. Firstly, with respect to the quantification of an apoptotic marker, (caspase, A) and secondly, a necrotic marker (miR-122, B).

The levels of necrosis were higher for the 20 mM APAP case than they were for the 5 mM case, as expected (Figure 4-6(B)). For the 5 mM APAP case, NAC did not seem to have any effect on necrosis levels. However, for the 20 mM case, it seemed that increasing the amount of NAC from 0-4 mM had a negative effect, increasing the amount of necrosis, until beyond 4 mM when necrosis levels started to decrease. Although this decrease in necrosis for the 20 mM case is mirrored by an increase in the caspase apoptosis marker in Figure 4-6(A), the overall results in Figure 4-6 are not consistent with the results from the FACS analysis carried out above in Figure 4-5. When quantifying apoptosis by analysing changes in caspase for the 5 mM APAP case (Figure 4-6(A)), NAC seemed to have no impact. The FACS analysis

(Figure 4-5(B)) showed minimal levels of apoptosis for the 0 mM NAC case, but an increased level for NAC [2, 5] mM. This protection is not mirrored in the caspase quantification results. For the 20 mM APAP case, caspase remained stable and increased when NAC >4 mM. Again, this result does not reflect the result from the FACS analysis (Figure 4-5(D)), where the increase in apoptosis was seen only at 2 mM, and then remained stable for all higher doses of NAC.

With regards to the 5 mM APAP dose, the FACS analysis showed high necrotic levels when 0mM NAC was provided, with reduced levels when a dose of [2, 5] mM NAC was provided. Again, this result is not mirrored with the results from the miR-122 quantification; miR-122 was low when 0 mM NAC was provided and remained stable throughout all doses of NAC. For the 20 mM APAP dose, Figure 4-6(B) shows that NAC did not impact on miR-122 as we may have expected – increasing the NAC dose also increased the amount of the miR-122 necrosis marker. This result is contrary to the result found in the FACS analysis, Figure 4-5(D), where high necrotic levels were found for the 0 mM NAC case, with a reduced amount for higher NAC doses. Carrying out more repeats of the experiments may help to explain this disparity.

4.3.2. Multiple linear regression - predicting cell viability

All biomarkers (APAP, ALT, HMGB1, full K18 and fragmented K18) were initially employed in the statistical analysis. Significant predictors from the forward stepwise multiple linear regression analysis for each subset of data are detailed in Table 4-1.

Dataset	Significant predictors	P-value
0 mM NAC subset	Full K18, Fragmented K18	0.001
1 mM NAC subset	APAP, HMGB1, Fragmented K18	<0.001
2 mM NAC subset	APAP, HMGB1, Fragmented K18	<0.001
3 mM NAC subset	APAP, HMGB1, Fragmented K18	<0.001
4 mM NAC subset	APAP, HMGB1, Fragmented K18	0.001
5 mM NAC subset	-	-
Full dataset	ALT, Full K18, Fragmented K18	<0.001

Table 4-1: Table of significant biomarkers for predicting cell viability for various subsets of data.

For the subset of data which had not been dosed with any antidote, cell viability could be significantly predicted by the full and fragmented versions of biomarker K18. For each subsequent NAC dosed subset, the combination of biomarkers APAP, HMGB1 and fragmented K18 could significantly predict cell viability. For the

subset of data that was dosed with 5 mM NAC, none of the biomarkers could significantly predict viability of the cells. When all of the subsets of data were combined into one full dataset (\pm NAC) viability of cells could be significantly predicted by ALT, full and fragmented K18 biomarker concentrations.

Interestingly, currently used clinical biomarkers (APAP/ALT) were predictive of cell viability when NAC dosed data was tested. However, if the analysis was tested on data with no NAC provided, these clinical biomarkers were not significant, and only the novel biomarkers could significantly predict cell viability.

4.3.3. Modelling the effect of NAC dosing regimen on the functionality of hepatocytes, and the probability of liver injury

4.3.3.1. Model optimisation

By extending our mathematical model within this chapter, our framework now has the ability to account for APAP, APAP toxicity biomarkers, and cell functionality profiles over time. For parameters affecting the cell functionality component of the model, so that model output could replicate our cell viability time-course data, the parameters were optimised by fitting the models to the cell viability data (resulting from the ATP assay). Optimised parameters affecting the cell functionality component of the model were as follows: from the functional hepatocyte model, functional hepatocyte damage rate, $\eta = 3.584 \text{ l}/\mu\text{mol}/\text{h}$; from the damaged hepatocyte model, damaged hepatocyte lysis rate, $\delta_z = 2.1891 \text{ h}^{-1}$; from the NAPQI model, decay of GSH due to NAPQI, $k_g = 7.1437 \text{ l}/\mu\text{mol}/\text{h}$; and from the GSH model, increasing factor of GSH due to NAC, $\zeta = 165.1162 \text{ h}^{-1}$. The *in vivo* functional hepatocyte regeneration rate was fixed as the value from Remien's study [73], $r = 0.0417 \text{ h}^{-1}$. The predictive output versus the data used for optimisation can be seen in Figure 4-7.

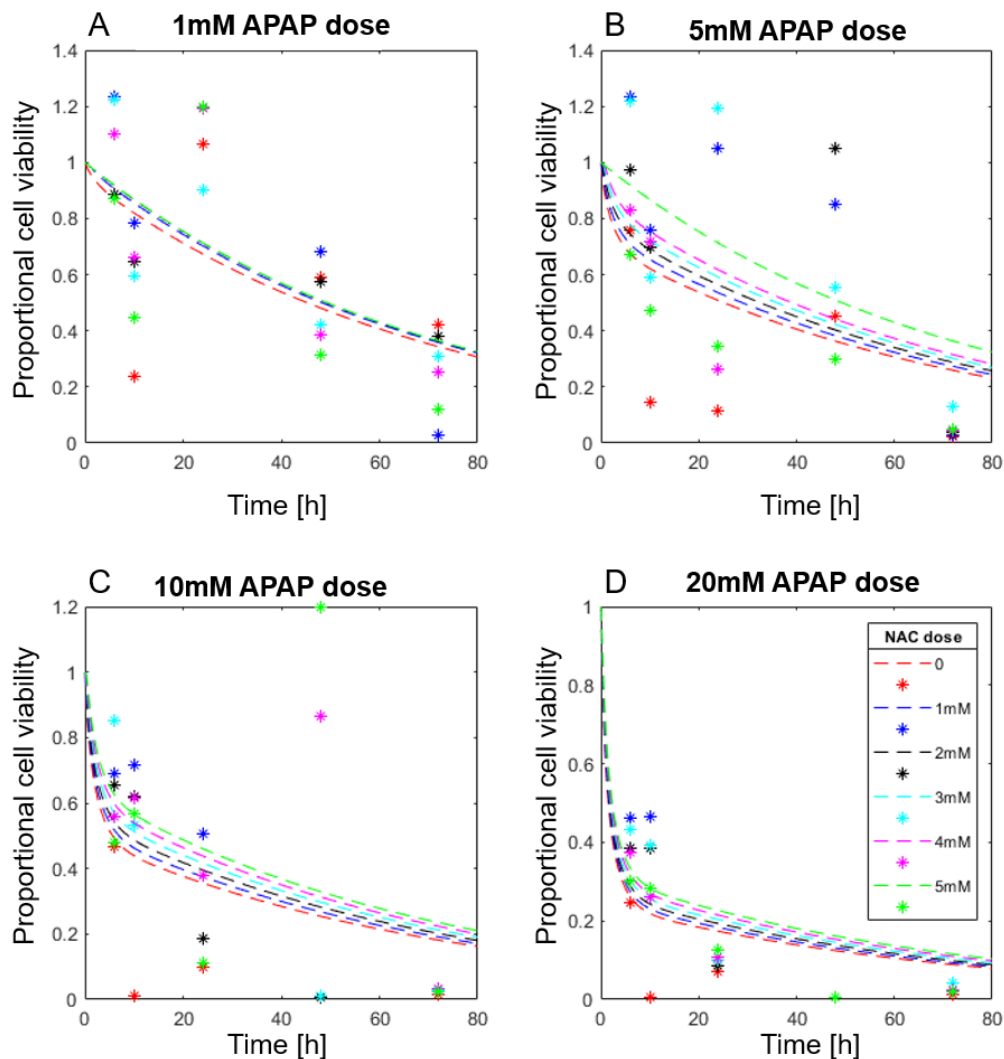


Figure 4-7: Predicted model output versus experimental data used for optimisation. (A) Proportional cell viability over time (*in vitro*) following a dose of 1 mM APAP. (B) Proportional cell viability over time (*in vitro*) following a dose of 5 mM APAP. (C) Proportional cell viability over time (*in vitro*) following a dose of 10 mM APAP. (D) Proportional cell viability over time (*in vitro*) following a dose of 20 mM APAP. In each instance, stars represent the experimental data used for optimisation and dashed lines represent the predicted model output.

4.3.3.2. The effect of NAC administration time on biomarker profiles

The optimal timing of NAC administration for the best protective effects is currently heavily debated within the clinic. Therefore, we used our model to simulate various scenarios, with differing times of NAC administration, and analysed subsequent effects on biomarker and cell functionality profiles (Figure 4-8). Note that the model was simulated to provide time-courses over a 24-hour period, since this was the latest time point observed in the biomarker data that was originally used for biomarker-related parameter optimisation.

For an APAP overdose of 350 mg/kg, providing no NAC at all resulted in very different predicted profiles compared to when 150 mg/kg NAC 1-hour-post-APAP is provided. These results can be seen in Figure 4-8. When no NAC was provided, the predicted GSH profile (Figure 4-8(B)) remained low for the entire time-course, whereas providing a dose of NAC at 1 hour caused great increases in the GSH concentration. The benefits of providing NAC can also be seen by the NAC dosed cases having lower predicted biomarker (ALT, HMGB1, NAPQI) concentrations than if no NAC was provided. While biomarkers ALT and HMGB1 began to increase following an APAP dose of 300 mg/kg, a slightly higher APAP dose (350 mg/kg) was required to see full and fragmented K18 profiles perturbed. However, at APAP doses higher than 350 mg/kg, all biomarkers followed a similar pattern— lower concentrations for the NAC-dosed case and higher concentrations when no NAC was provided. We report the findings for a 350 mg/kg APAP dose since this is the dose thought to induce toxicity in mice (predicted by our previous framework in Chapters 2 and 3). NAC dose is predicted to also have a positive effect on cell functionality: Figure 4-8(H) shows that when provided with a 350 mg/kg APAP dose and no NAC protection, the predicted proportional cell viability reduced to around 10%, and only began to slightly recover at late time-points (beyond 24 hours). In the case where 150 mg/kg NAC was provided 1 hour post-APAP-dose however, the proportion of viable cells did not decrease beyond approximately 50%, and immediately after NAC intervention the cells began to recover.

NAC is usually administered beyond 4 hours of the initial APAP overdose, and so therefore we simulated the administration of 150 mg/kg NAC 5 hours post APAP-dose to visualise the effects on biomarker concentrations and cell functionality (Figure 4-9).

Providing the NAC dose 5-hours-post-APAP-dose still provided favourable biomarker concentration/cell functionality profiles when compared to the cases where no NAC was provided. This is reflected firstly in Figure 4-9(B) where the GSH concentration profile in the NAC-dosed scenario (green) is much higher than the non-NAC-dosed scenario (red). These positive results are also mirrored by ALT and HMGB1 concentrations being lower for the NAC dosed case than the non-NAC dosed case Figure 4-9(C,D); and also the proportion of viable cells being slightly higher in the NAC case compared to the non-NAC case Figure 4-9(H). Generally though, NAC was predicted to provide less improvement if administered 5 hours post-dose (Figure 4-9) than if it is administered 1 hour post dose (Figure 4-8).

It is thought that NAC is no longer beneficial if provided longer than 8 hours post-overdose [23]. We therefore simulated the administration of 150 mg/kg NAC 15-hours-post-APAP-dose to investigate the effects on biomarker concentrations and cell functionality (Figure 4-10).

Although Figure 4-10 (B) shows the expected positive effects of NAC administration on GSH concentration (the NAC-dosed profile is much higher than the non-dosed profile) beyond 15 hours, NAC administration only very slightly lowered ALT and HMGB1 concentrations (Figure 4-10(C,D)). Providing NAC 15-hours-post-APAP dose was predicted to have no protective effects on cell viability Figure 4-10 (H,I).

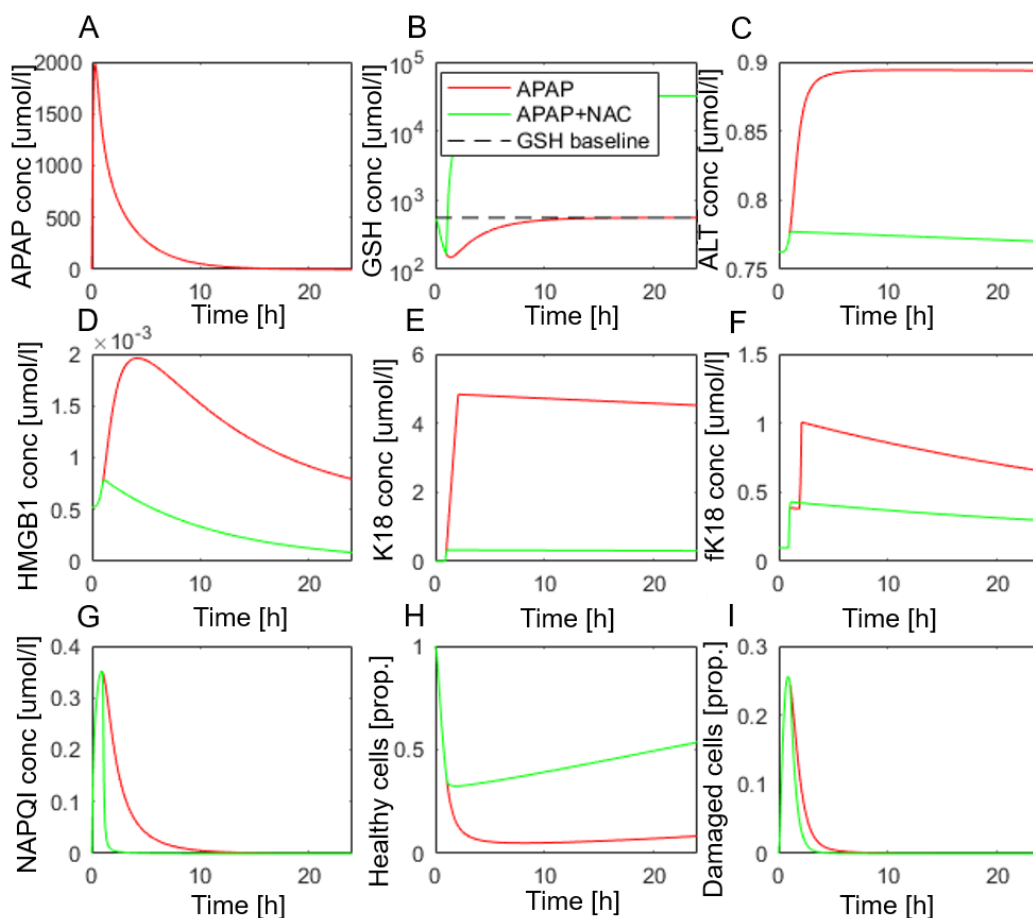


Figure 4-8: Effect of administering antidote (NAC) 1-hour-post-APAP overdose on predicted biomarker and cell functionality profiles. Red lines show the simulated output if no NAC dose is provided, while green shows the simulated output if 150 mg/kg NAC is provided. If only a green simulation is present, this indicates it overlaps the non-NAC simulation, and therefore they are identical. (A) Predicted APAP concentration over time, (B) predicted GSH concentration over time, (C) predicted ALT concentration over time, (D) predicted HMGB1 concentration over time, (E) predicted full K18 concentration over time, (F) predicted fragmented K18 concentration over time, (G) predicted NAPQI concentration over time, (H) proportion of healthy cells over time, (I) proportion of damaged cells over time.

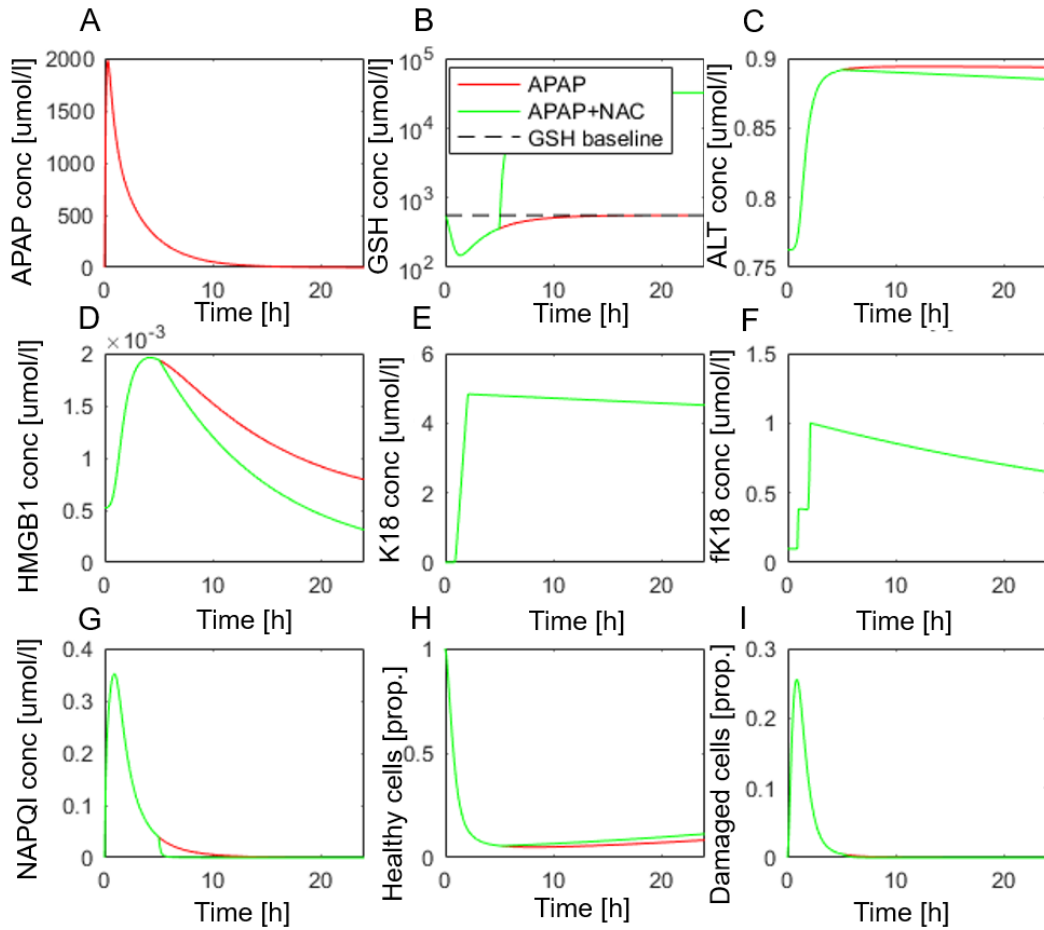


Figure 4-9: Effect of administering antidote (NAC) 5-hours-post-APAP overdose on predicted biomarker and cell functionality profiles. Red lines show the simulated output if no NAC dose is provided, while green shows the simulated output if 150 mg/kg NAC is provided. If only a green simulation is present, this indicates it overlaps the non-NAC simulation, and therefore they are identical. (A) Predicted APAP concentration over time, (B) predicted GSH concentration over time, (C) predicted ALT concentration over time, (D) predicted HMGB1 concentration over time, (E) predicted full K18 concentration over time, (F) predicted fragmented K18 concentration over time, (G) predicted NAPQI concentration over time, (H) proportion of healthy cells over time, (I) proportion of damaged cells over time.

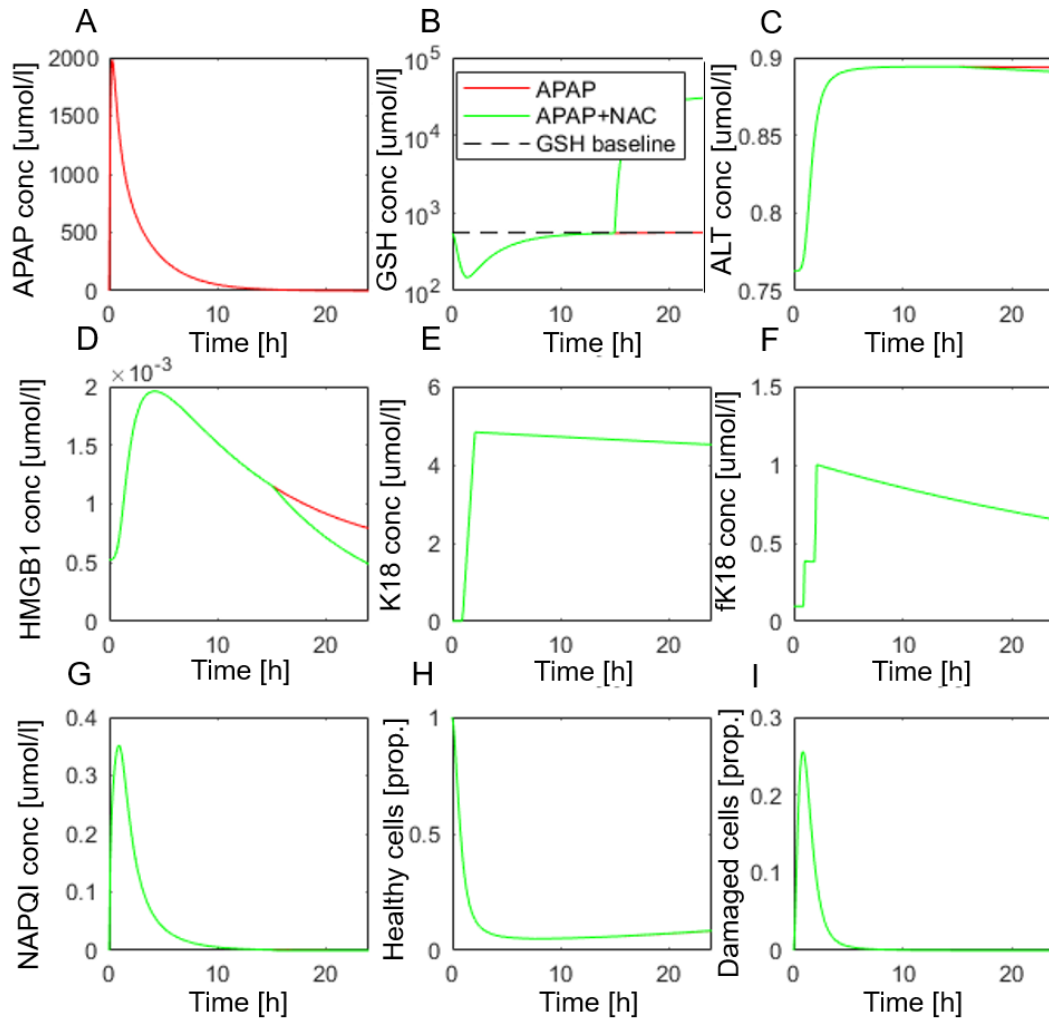


Figure 4-10: Effect of administering antidote (NAC) 15-hours-post-APAP overdose on predicted biomarker and cell functionality profiles. Red lines show the simulated output if no NAC dose is provided, while green shows the simulated output if 150 mg/kg NAC is provided. If only a green simulation is present, this indicates it overlaps the non-NAC simulation, and therefore they are identical. (A) Predicted APAP concentration over time, (B) predicted GSH concentration over time, (C) predicted ALT concentration over time, (D) predicted HMGB1 concentration over time, (E) predicted full K18 concentration over time, (F) predicted fragmented K18 concentration over time, (G) predicted NAPQI concentration over time, (H) proportion of healthy cells over time, (I) proportion of damaged cells over time.

4.3.3.3. The effect of NAC dose amount on toxicity thresholds

As well as investigating effects of the administration time of NAC, the effects of NAC dose alone (without changing the time of administration) was investigated with regard to the effect on both of our toxicity measures: probability of liver injury and proportion of healthy cells. Each NAC dose under investigation, [0, 150, 400] mg/kg, was administered 1-hour post-dose, and we calculated which APAP dose would indicate probable liver injury and severe reduction in the proportion of healthy cells.

If the NAC administration was beneficial, then a higher dose of APAP would be required to result in probable liver injury/severe reduction in cell viability.

NAC resulted in protecting against much higher APAP doses when using the probability of liver injury based on biomarker concentrations as a toxicity measure, in comparison to using loss of cell functionality as a toxicity measure. This can be clearly visualised in Figure 4-11 by the order of magnitude difference in the scale of APAP doses reaching up to 10,000 mg/kg in the probable liver injury cases (Figure 4-11(A,C,E)) and reaching only 1,000 mg/kg in the cell functionality cases (Figure 4-11(B,D,F)). If no NAC was provided, the APAP dose threshold for toxicity was of the same order of magnitude. Using probability of liver injury as a toxicity measure (Figure 4-11(A)), toxicity was predicted for any APAP dose above approximately 550 mg/kg. This is shown by the black contour – any observation above this contour is predicted to be highly probable to progress into liver injury. When using loss of cell functionality (or the percentage of dead/damaged cells being greater than 80%) as a toxicity measure (Figure 4-11(B)), toxicity was predicted for any APAP dose above approximately 400 mg/kg, with any APAP dose above approximately 500 mg/kg predicted to result in 100% of cells being dead/damaged.

Providing a dose of NAC (either 150 mg/kg or 200 mg/kg) appeared to make a slight change to the predictions when using cell functionality as a toxicity measure (Figure 4-11(D,F)). The contour increased to a slightly higher APAP dose, and the progression from 80% to 100% hepatocyte death/damage was more gradual as the APAP dose increased, rather than any APAP dose above 500 mg/kg resulting in 100% death/damage. When providing NAC and using biomarker-based-probable liver injury as a toxicity measure, the antidote benefits could be clearly seen. Treatment with the recommended 150 mg/kg dose of NAC increased the APAP dose predicted to cause liver injury from approximately 550 mg/kg (Figure 4-11(A)), to a much greater 7000 mg/kg (Figure 4-11(B)). This APAP dose was even further increased when 200 mg/kg NAC is provided (Figure 4-11(C)).

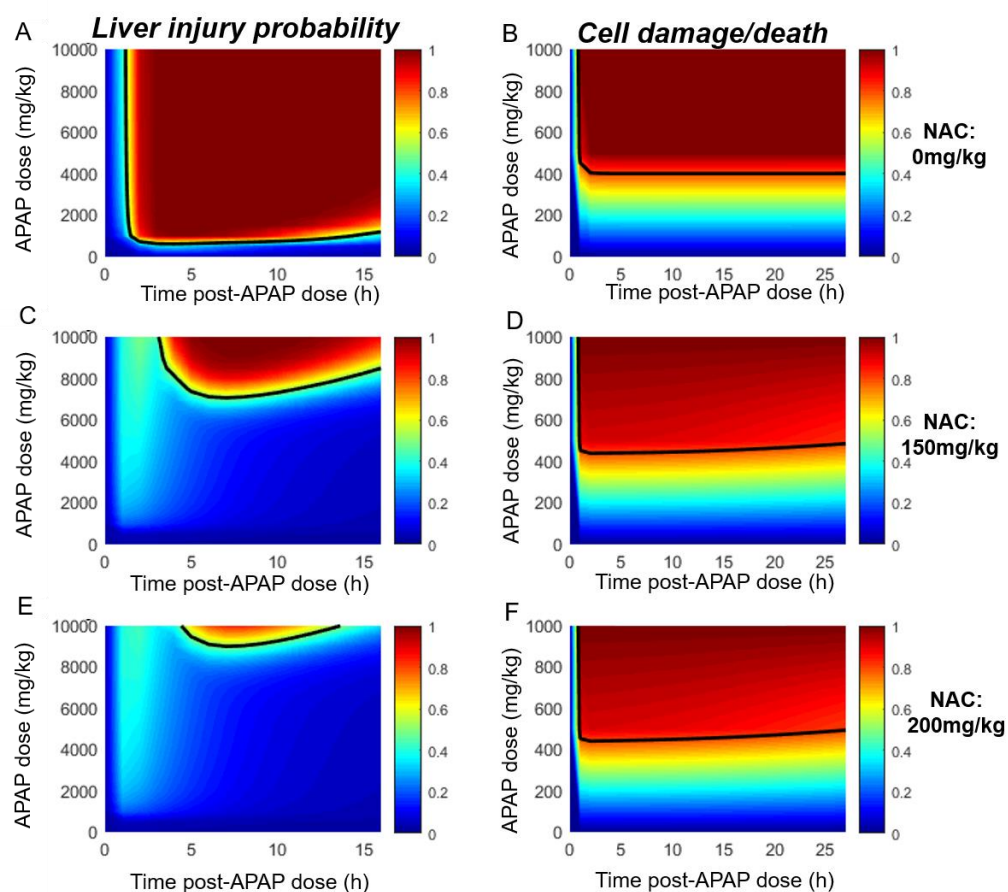


Figure 4-11: Effect of administering different doses of antidote (NAC) 1-hour-post-APAP overdose on the predicted probability of liver injury and proportion of healthy cells for various APAP doses. *Black contours are thresholds where any observation above such would be predicted to have a high probability of liver injury (A,C,E) or a severe loss of cell functionality (B,D,F). Red regions are representative of 1 and blue regions are representative of 0. Therefore, in the liver injury probability plots (A,C,E) the red regions are representative of a predicted probability of liver injury of 100% and therefore severe toxicity. The red region in the cell functionality plots (B,D,F), is representative of cells being 100% dead/damaged, and therefore indicative of severe toxicity. (A,B) show the effects of providing no NAC dose on the APAP dose thresholds for probability of liver injury and cell functionality, respectively. (C,D) show the effects of providing a 150 mg/kg NAC dose on the APAP dose thresholds probability of liver injury and cell functionality, respectively. (E,F) show the effects of providing 200 mg/kg NAC dose on the APAP dose thresholds for probability of liver injury and cell functionality, respectively.*

The benefits of providing NAC at 1-hour-post-APAP dose (when using probable liver injury as a toxicity measure) is clearly represented in Figure 4-12(A) by the large difference in thresholds for the “No NAC” case (black contour) when compared to the NAC cases (blue and orange contours).

The APAP dose predicted to induce severe loss of cell functionality when no NAC was provided was approximately 400 mg/kg (Figure 4-11(B)). Providing a NAC dose of 150 mg/kg slightly increased the APAP dose predicted to induce severe loss of cell functionality, however increasing the NAC dose from 150 mg/kg to 200 mg/kg

did not have any additional effect (Figure 4-12(B)). Our results suggest although NAC may slightly increase the APAP dose required to cause toxicity, this effect may be limited for cell functionality and increasing the NAC dose will not necessarily increase the APAP dose likely to cause toxicity.

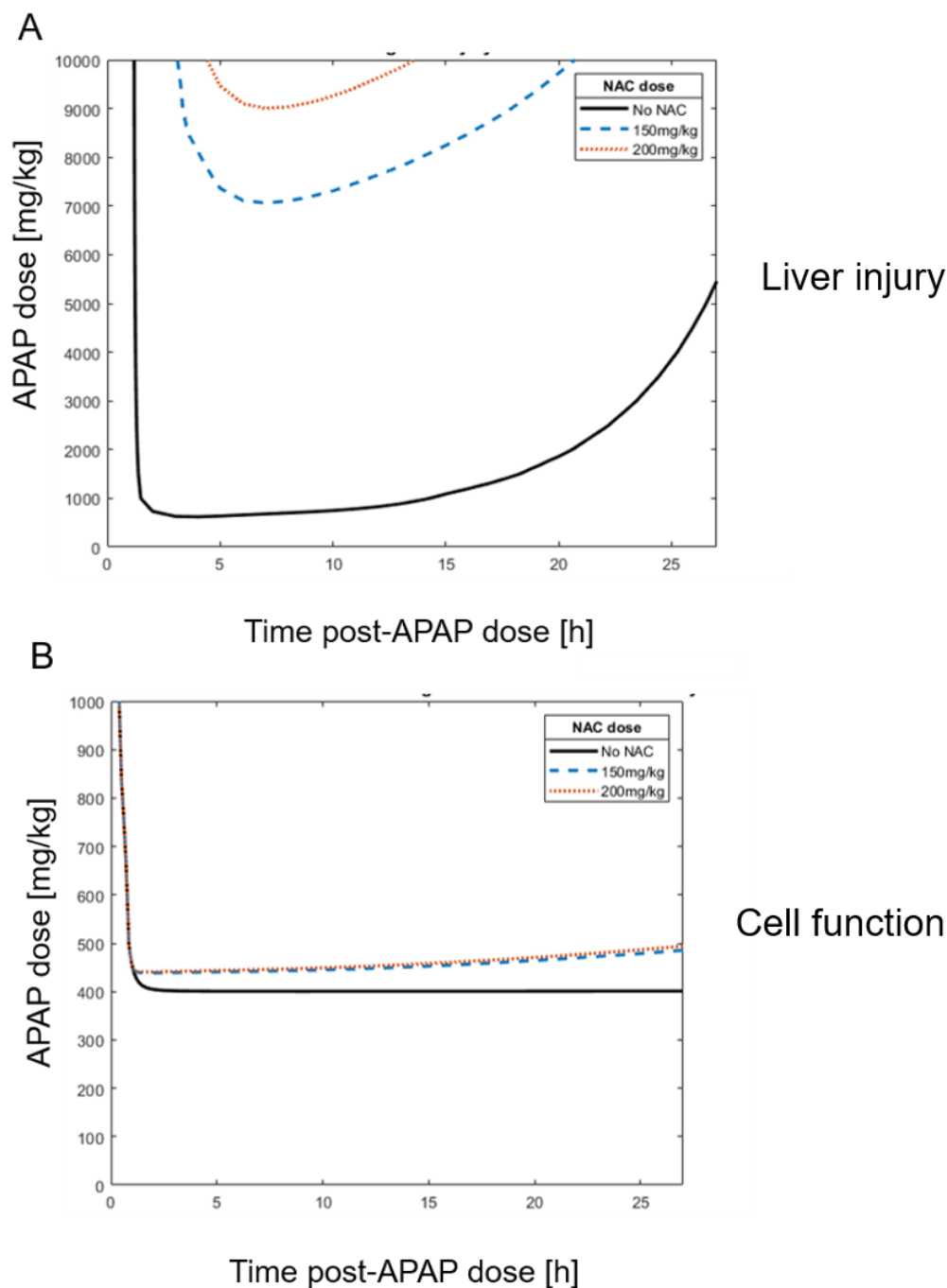


Figure 4-12: NAC dose effects on the APAP dose threshold for inducing liver toxicity based on two different measures: (A) probable liver injury based on HMGB1 concentration (probability greater than 0.5) and (B) severe loss of cell functionality (80% reduction of cell viability). Any observation above any of the thresholds is predicted to be a toxic case. The black contour represents the toxic threshold when no NAC is provided. The blue dashed contour represents the toxic threshold when 150 mg/kg NAC is provided. The orange dotted contour represents the toxic threshold when 200 mg/kg NAC is provided.

4.3.3.4. Comparing the effects of two NAC regimens on biomarker profiles

The current NAC regimen [24] (providing an initial dose of 150 mg over 1h, followed by a second dose of 50 mg over 4h, followed by a final dose of 100 mg over 16h) was compared to a NAC regimen in which the NAC is administered over a shorter time period (100 mg over 2h followed by 200 mg over 10h). This shorter NAC regimen is often referred to as SNAP and is proposed by Bateman et al. [152]. Both NAC protocols were applied and corresponding biomarker outputs were simulated (Figure 4-13).

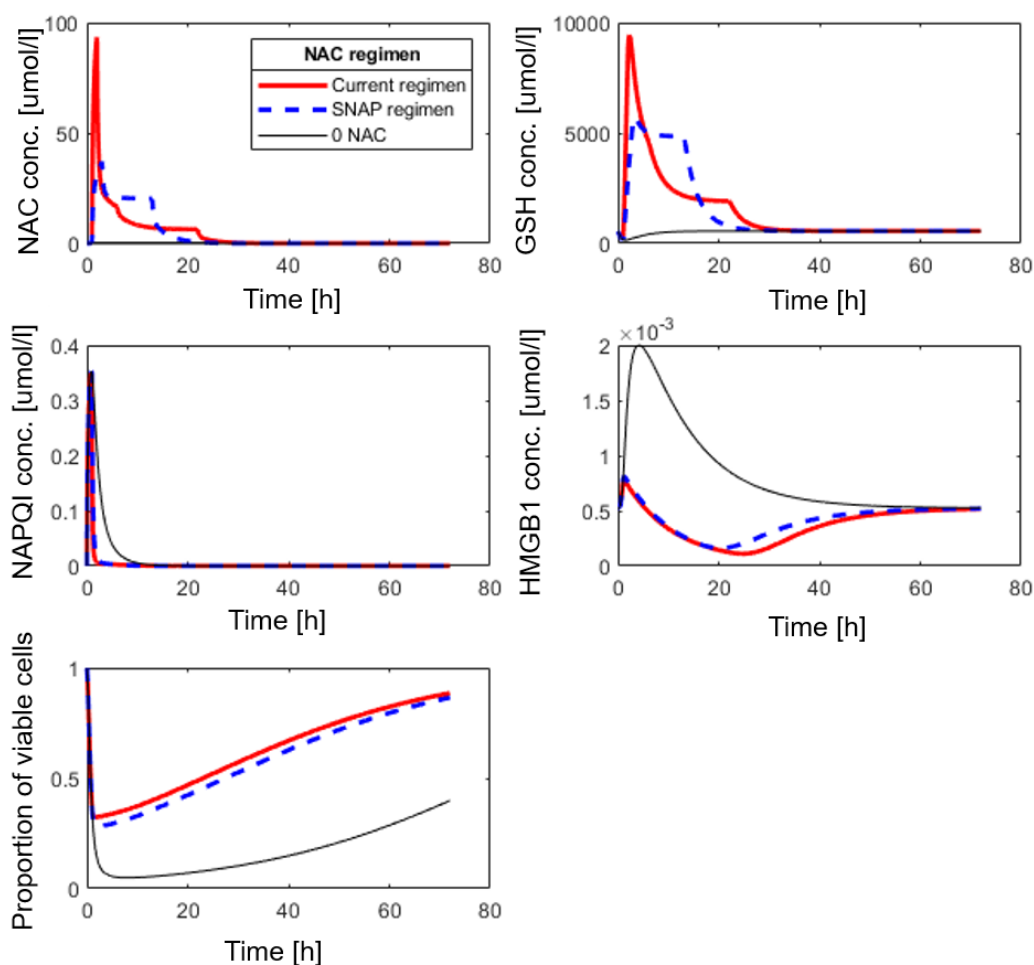


Figure 4-13: The effect of different NAC regimens on various time-course profiles (black profile – no NAC provided, red profile – current NAC regimen, blue dashed profile – SNAP regimen). (A) NAC concentration profile, (B) GSH concentration profile, (C) NAPQI concentration profile, (D) HMGB1 concentration profile, (E) Proportional cell viability profile.

Of course, the NAC concentration remained at 0 throughout the whole time-course when no NAC was provided (Figure 4-13(A)). The current regimen allowed a high peak concentration which quickly declined. The SNAP regimen allowed a peak

concentration around half of that provided by the current regimen, but the concentration remained around this level for around 17 hours before declining. Both NAC regimens allowed the GSH concentration to far exceed its baseline (Figure 4-13(B)) before decreasing to its baseline after approximately 24 hours. NAPQI concentration rapidly decreased following administration of both NAC regimens (Figure 4-13(C)). When no NAC was provided, as expected, HMGB1 increased from baseline, before returning to baseline around 50 hours post APAP-dose. When either NAC regimen was introduced, the biomarker decreased below the baseline value, before returning at around 50 hours. Both regimens maintained HMGB1 concentration at around the same value for the first 20 hours post-dose, however, since the shorter NAC regimen did not continue beyond this point, HMGB1 concentrations returned to baseline at a faster rate than the conventional regimen. If no NAC was provided, cell viability was predicted to almost completely diminish before starting to slowly recover beyond approximately 20 hours. Both NAC regimens were predicted to protect the function of cells, with viability reducing to approximately 40% in both scenarios, and recovering quicker than the case in which no NAC was provided.

4.3.3.5. Comparing the effect of two NAC regimens on toxicity thresholds

Both regimens (conventional and SNAP) were simulated and compared to investigate any differences in APAP dose threshold predicted to induce probable liver injury probability and severe loss of cell functionality (Figure 4-14).

NAC administration had a much greater influence on the toxicity measure based on probable liver injury in comparison to the toxicity measure based on loss of cell functionality (Figure 4-14). This is reflected by APAP doses likely to induce toxicity being on a much greater scale (up to 10,000 mg/kg) in the probable liver injury plots (Figure 4-14(A,C)) when compared to the lower scale (up to 800 mg/kg) on the cell functionality plots (Figure 4-14(B,D)). This result (NAC having a greater effect on probable liver injury as a toxicity measure, rather than loss of cell functionality) mirrors the result found in (Figure 4-11). Our results suggest that if cell functionality is used as the toxicity measure, (Figure 4-14(B,D)), both NAC regimens have similar protective effects, with severe loss of cell functionality apparent for any APAP dose above approximately 450 mg/kg. When using probable liver injury based on HMGB1 concentration as a measure however (Figure 4-14(A,C)), the NAC regimens

provided clearly different results. The current regimen only indicated probable toxicity when an APAP dose above 8,000 mg/kg had been taken, and this probability was only around 0.5. The SNAP regimen however, identified potential for cases where the APAP dose was above a much lower threshold of 4,000 mg/kg, with most cases having a probability greater than 0.8.

Whilst the differences in the NAC effects on both toxicity measures can be seen in Figure 4-14, we looked more closely at comparing each regimen against the measures individually, with Figure 4-15(A) showing the regimen effects if probable liver injury based on HMGB1 was used as a toxicity measure, and Figure 4-15(B) showing the regimen effects if loss of cell functionality was used as a toxicity measure.

When using probable liver injury based on HMGB1 concentration as a toxicity measure (Figure 4-15(A)), both NAC regimens were predicted to be highly protective (toxicity not predicted to occur for any APAP dose below 3,500 mg/kg). However, the protective effects were predicted to be much greater for the current regimen. Simulating the current regimen (black threshold) identified HMGB1 concentrations that resulted in probable liver injury levels for APAP doses above 8,000 mg/kg. It appears that the maximum APAP dose the SNAP regimen (blue dashed threshold) could protect against was 3,500 mg/kg APAP, although this is still a very high threshold.

When cell functionality was used as a toxicity measure (Figure 4-15(B)), both NAC regimens were predicted to have similar protective potential, especially at early time-points following the APAP dose. The current regimen (black threshold) enabled functionality to recover at a slightly faster rate than the proposed regimen (blue dashed threshold). For example, if two individuals were to take 430 mg/kg, with one being administered NAC according to the current regimen, and the other being administered NAC according to the SNAP regimen, at 2 hours-post APAP dose, both would be deemed toxic, since more than 80% cells would be predicted to be dead/damaged. At 25-hours-post-APAP-dose however, whilst the person following the SNAP schedule would remain in the toxic region (above the blue dashed threshold), the person following the current NAC regimen would no longer be deemed toxic.

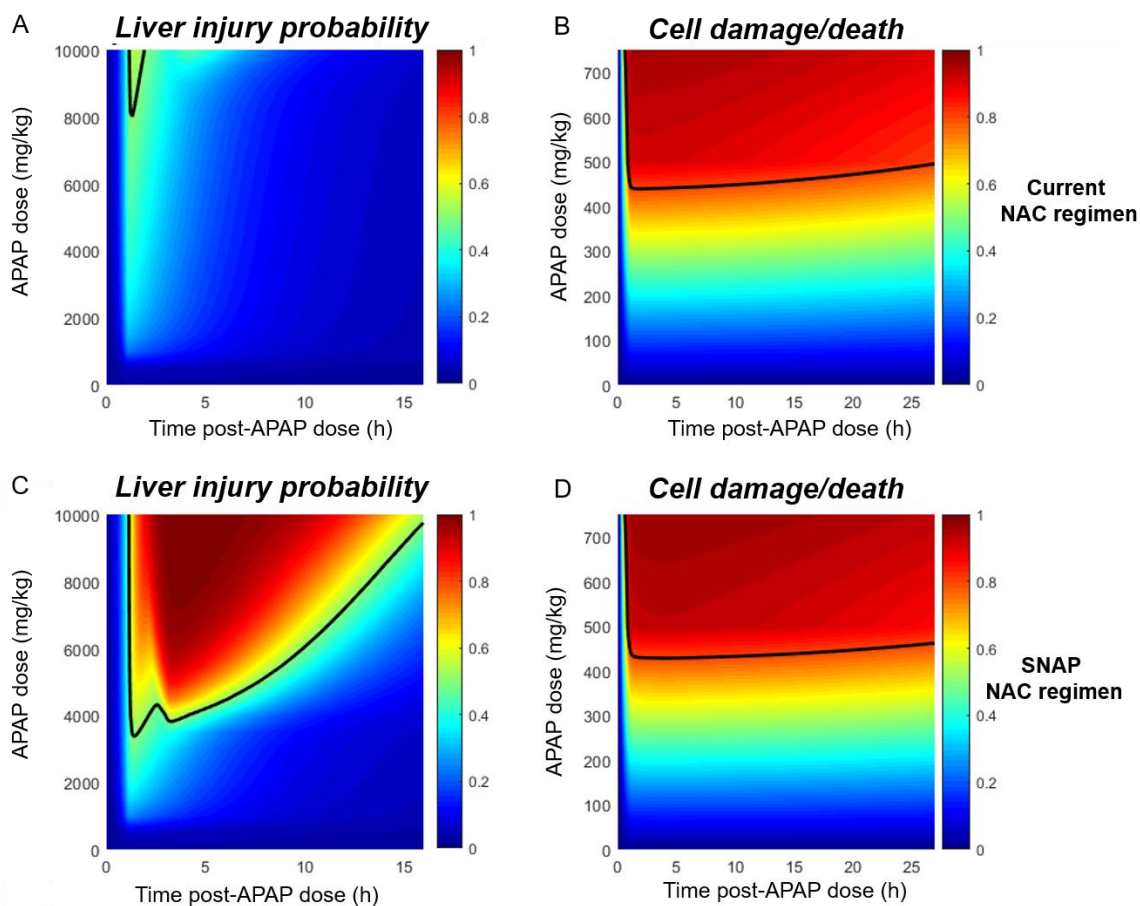


Figure 4-14: Effect of two comparative NAC dosing regimens on the predicted probability of liver injury and proportion of damaged/dead cells for various APAP doses. *Black contours are thresholds where any observation above such would be predicted to have a high probability of liver injury or a severe loss of cell functionality. Red regions are representative of 1 and blue regions are representative of 0. Therefore, in the liver injury probability plots (A, C) the red regions are representative of a predicted probability of liver injury of 100% and therefore severe toxicity. The blue region in the cell functionality plots (B, D), is representative of cells being 100% damaged/dead, and therefore indicative of severe toxicity. (A,B) show the effects of providing the conventional NAC dosing regimen on the APAP dose thresholds for probability of liver injury and cell functionality, respectively. (C,D) show the effects of providing the proposed SNAP dosing regimen on the APAP dose thresholds for probability of liver injury and cell functionality, respectively.*

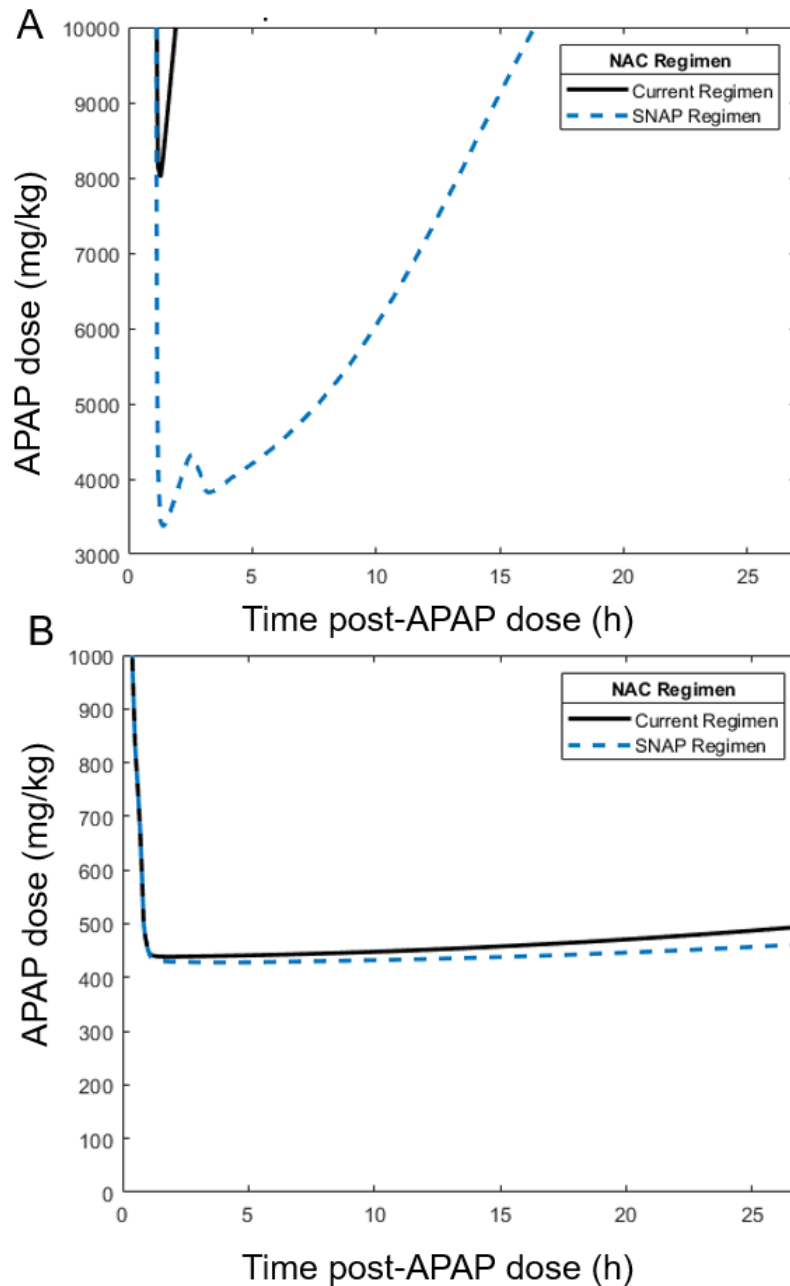


Figure 4-15: NAC dosing regimen effects on the APAP dose threshold for inducing liver toxicity based on two different measures: (A) probable liver injury based on HMGB1 concentration (probability greater than 0.5) and (B) severe loss of cell functionality (80% reduction of cell viability). Any observation above the thresholds is predicted to have a HMGB1 concentration indicating probable liver injury. The black contour represents the threshold when the current clinical NAC regimen is provided. The blue dashed contour represents the proposed shorter NAC (SNAP) regimen is provided.

4.4. Discussion

In previous Chapters (2 and 3), we have been interested in the relationship between APAP and toxicity biomarkers and their subsequent predictive potential. Should the interest be to make predictions upon initial investigation of a suspected APAP overdose case, this may be sufficient. However, if we aim to use the toxicity biomarkers throughout a patient's hospital admission, the relationship between APAP dose and these novel biomarkers requires extension to include the antidote (NAC). The aim with this chapter was to generate and analyse my own experimental data to better understand the influence NAC had on these biomarker concentrations and their subsequent predictivity, and also to obtain a further understanding of liver injury at the cellular level (cell death via necrosis or apoptosis). We firstly discuss whether NAC shows protective effects, and if so, how this changes the mechanism of cell death.

Looking at the cell-viability time-course (Figure 4-4), as expected, there was a general decrease in viability as time progressed following an APAP dose. Also, the protective effects of NAC could be seen, especially for lower doses of APAP. Providing a high dose of NAC (5 mM) had similar effects as providing no antidote treatment at all, this was an interesting result. This may give further justification for the requirement to better optimise NAC treatment; our results suggest that providing too much antidote could cause as much damage as not intervening at all. This result was further justified by the conclusions from our FACS analysis; NAC provided a protective switch, where we observed the dominant form of cell death change from necrosis to apoptosis. However, the amount of NAC required to cause this switch was different for different APAP doses, and the relationship between NAC and APAP dose with regards to this switch was not clear from these experiments. It must be noted that these experiments have not yet been repeated; further repeats could help to understand the relationship between NAC and APAP dose and the switch from necrosis to apoptosis.

Although an attempt was made to further investigate the relationship between APAP, NAC and mode of cell death by comparing quantification of apoptosis marker (caspase) and necrotic marker (miR-122), the results were inconsistent with our FACS analysis. These opposing results could be due to interference when the assays were carried out, and since we only have $n=1$, more repeats are required in order to clarify this discrepancy. At present, since NAC is known to have protective

effects in the APAP overdose case, the FACS analysis results are in line with what would be expected, and therefore can be included in the model with more confidence than the caspase/miR-122 quantification data, until more repeats are carried out.

With the aim of further model development, note that the frameworks identified in the two previous chapters may only be applicable for a patient at initial hospital presentation with a suspected APAP overdose. However, following hospital admission, patients will likely be given NAC treatment. Additional to investigating NAC effects on hepatocyte functionality, we used multiple linear regression analysis to determine if introduction of therapy could alter the predictivity of the biomarkers. When looking at the subset of data which was not given any NAC therapy, biomarkers full and fragmented K18 could significantly predict cell viability. Upon initial presentation, before a patient has been given the antidote, these biomarkers may prove useful to predict how well the cells are functioning. Measuring these two values would make sense since they are necrotic/apoptotic versions of the same biomarker; quantifying both levels could therefore give insight into cell damage – high levels of necrosis compared to low levels of apoptosis for example may indicate the functionality of the cell is diminishing. For each NAC dose (other than 5 mM) the same biomarkers – APAP, HMGB1 and fragmented K18 could significantly predict cell viability. The fact that no biomarkers could significantly predict cell viability for the 5 mM NAC dosed subset could be due to the fact that the dose of NAC is too high and could have interaction effects with APAP.

Combining all datasets (no antidote and various antidote dosed subsets) resulted in biomarkers ALT, full and fragmented K18 being significant predictors for cell viability. ALT is a biomarker already utilised clinically which gives credence to the proposal of our panel. Therefore, measuring these biomarkers (ALT, full and fragmented K18), whether previously given a dose of the antidote or not, could predict the level of cell viability and therefore give insight into cell functionality following an APAP dose. The fact that current clinical biomarkers only became significant in cases where the analysis was applied to data which included NAC doses, may suggest that these biomarkers are only useful for monitoring purposes following antidote treatment. However, if cell functionality is of interest at initial presentation when no NAC therapy has been provided, our analysis predicts that the novel biomarkers outperform those currently used in the clinic.

Whilst in previous Chapters (2 and 3) we describe a predictive framework for the probability of liver injury based on biomarker concentrations, we here provide an extension to this framework, which additionally uses severe loss of cell functionality as a marker for potential toxicity. Remien et al. [73] published a mathematical model which describes the dynamics of functional and damaged hepatocytes in the APAP overdose setting. Whilst their parameters are representative of cell functionality following APAP overdose, parameters within our model can describe cell functionality when both APAP plus NAC antidote are provided, rather than APAP alone.

Simulating the administration of NAC at 1-hour-post-APAP dose provided lower predicted biomarker concentrations than the case that was simulated without any NAC administration. The longer the time period between APAP dose and NAC dose, the lower the effect NAC had on subsequent biomarker concentrations, which could further support the idea that NAC is no longer beneficial beyond 8 hours post-APAP dose [23]. When investigating how thresholds for probable liver injury (based on biomarker concentrations) differed, the effects of providing no antidote versus providing antidote could be clearly seen. This investigation was carried out by supplying different NAC doses at 1-hour-post-APAP dose and determining the resultant APAP dose threshold for probable liver injury. When no NAC was provided, liver injury was predicted to be probable following any APAP dose above approximately 550 mg/kg. Providing a 150 mg/kg dose of NAC increased the toxic APAP threshold to 7,000 mg/kg while providing a 200 mg/kg NAC dose increased the threshold even further to 9,000 mg/kg. Whilst this result supports the benefits of NAC intervention, it must be noted that this framework is based on the concentration of one biomarker (HMGB1). Further validation of the biomarker's predictivity should be carried out. Should cell functionality be the preferred measure of toxicity, much lower APAP dose thresholds were predicted, even following administration of NAC at 1-hour-post-APAP-dose. Although increasing the amount of NAC did not greatly increase the toxic threshold, the distinction was still clear between no antidote, and antidote cases, as can be seen in Figure 4-12; the non-NAC case provided a lower APAP dose threshold than both of the NAC dosed cases.

There has been much discussion around the uncertainty of the optimal combination of both NAC dose amount and time since administration. The current NAC regimen [24] was simulated and compared with a recently proposed shorter NAC protocol (SNAP regimen) [152] to visualise their effects on biomarker time-course profiles,

and also the predicted toxic APAP thresholds (for both probable liver injury and severe loss of cell functionality). As expected, providing NAC (following either regimen) caused large increases in peak NAC and GSH concentrations. The current regimen provided higher peaks which declined at a faster rate, whilst the SNAP regimen provided a lower peak, but with increased concentrations remaining for a longer time period. When using the probability of liver injury (based on HMGB1 concentration) as a toxicity measure, the current regimen had better protective capability than the SNAP regimen; the current regimen was predicted to protect for APAP cases up to 8,000 mg/kg and the SNAP regimen was predicted to protect for APAP cases up to 3,500 mg/kg. However, the lower threshold (3,500 mg/kg) is still a really high APAP dose, and therefore the protective effects of this regimen should not be discounted.

If loss of cell functionality was used as a toxicity measure, both regimens were predicted to provide a similar level of protection (for APAP cases up to approximately 450 mg/kg). Using cell functionality as a toxicity measure, our model predicted severe toxicity occurrence for any APAP dose beyond 400 mg/kg when NAC was not provided. Currently, the clinic suggest that any individual having taken an APAP dose above 250 mg/kg is likely to experience potentially fatal liver toxicity [8]. Our preliminary results suggest that this clinical threshold may have the potential to be increased, possibly leading to huge savings on unnecessary transplantations. However, since there are no repeats of this study, further investigation would be required before such suggestions could be confirmed. The initial results are promising though, and warrant further investigation.

The fact that results differed greatly depending on the toxicity measure used (probability of liver injury based on HMGB1 concentration, or loss of cell functionality based on *in vitro* data) in each scenario when NAC was provided, supports the requirement for better understanding of the relationship between the two measures. Results from our investigation could be used to guide *in vitro* experiments to improve this understanding. Potential experiments could focus on monitoring multidrug resistance proteins (MRPs) over time following different APAP/NAC doses. These transporters are reported to transport GSH and also bind to ATP [153]. Further experimental repeats of the cell viability study would also be required to ensure confidence in the accuracy of the optimised parameters used in our cell functionality model. It must be noted however, that whilst the biomarker in our predictive liver injury framework is heavily dependent on GSH depletion, and although our cell

functionality model is also based on NAPQI accumulation (which is itself dependent on GSH depletion), we have optimised against cell viability data which is representative of intracellular ATP depletion. Our model framework could benefit from optimisation against more diverse cell functionality data such as albumin changes or ability to metabolise MTT. Albumin is the most abundant protein in the blood and decreased levels are widely used to indicate severely ill individuals [11], while MTT is a compound that viable cells with active metabolism can convert into a purple coloured formazan product [154]. When cells die, they lose the ability to convert MTT into formazan and therefore colour formation can serve as a convenient marker for cell viability.

The results from our investigation have provided further insight and confirmation into the benefits of NAC administration in the overdose case. Although the results from our preliminary investigation are promising, the *in vitro* experiments used in the analysis were not repeated due to time and cost limitations. Therefore, before these results can be concluded with confidence, more repeats are highly recommended.

Chapter 5: Clinical paracetamol overdose model

5.1. Background

Ultimately, this research has the aim of making predictions to better inform human safety. The current clinical treatment framework for APAP overdose is known to be highly sub-optimal, mainly due to its dependency on APAP dose amount; and time since overdose, information which is often highly inaccurate. In Chapters 2 and 3, we provide a pre-clinical framework that, once extrapolated to human, could have the potential to improve this current APAP clinical decision-making framework. Although APAP metabolism in mice is similar to that in human [155], humans are known to have much larger inter-individual variability in the metabolism processes in comparison to mice [77]. The modelling approach in the human case must therefore be more extensive than the pre-clinical case, ensuring this inter-individuality is accounted for. In this chapter, we extend our mathematical modelling approaches described in previous chapters to the clinical case.

Through collaboration with Dr. James Dear (Consultant and Honorary Senior Lecturer in Clinical Pharmacology at the Royal Infirmary Edinburgh, Edinburgh University) we obtained access to a database of approximately 3,600 APAP overdose patients, providing data for the optimisation of model parameters. We take a non-linear mixed effects modelling approach to account for both known and unknown errors in model predictions due to physiological differences between humans. The issues with the APAP toxicity framework can be further complicated by these physiological differences. Even though we are able to provide more accurate predictions of overdose amount and time since overdose, there are certain individuals that may respond to treatment differently for matters which are currently not accounted for within our modelling framework. For example, there is an apparent distinct global increase in accidental pediatric overdoses [156], and there have subsequently been many toxicological studies that have tried to understand metabolic differences in younger people versus older people, and whether these potential differences are linked to the increase in pediatric overdoses [157]. Additionally, there have been many cases where APAP toxicity has been apparent although APAP was prescribed and taken appropriately at the recommended therapeutic dosage [158,159], particularly for individuals of low weight (<50 kg) [160,161]. Identifying individuals that are particularly susceptible to APAP-induced liver injury has been heavily debated in the literature (for a review, see Caparrotta et al. [157]). Briefly, individuals are deemed 'high-risk' in the APAP overdose

scenario if they are known to have factors which may impair their ability to synthesise or maintain sufficient GSH levels [10]. Such factors may include pre-existing liver disease, concurrent use of alcohol and/or other liver-metabolised medications, genetic predispositions and acuity/chronicity of APAP use [5]. The debate has been such that the threshold for providing NAC treatment (found by plotting measured APAP concentration against time since overdose) has been lowered by 25% from the original threshold to take into account individuals with potential risk factors and also to account for the possibility of measurement errors [20]. However, this of course bears the risk of treating more patients unnecessarily.

In silico models are being increasingly used to make drug toxicity predictions, with focus on liver toxicity specifically. For a review, see Yang et al. who review *in silico* models for DILI prediction that were published between 2000-2015 [162]. In this chapter, we describe a nonlinear mixed effects modelling approach to pharmacokinetic modelling, which allows us to make population-pharmacokinetic (Pop-PK) predictions for individuals outside the region of the dataset used in optimisation. The APAP Pop-PK model is optimised against clinical data provided by Dr. James Dear at Queens Medical Research Institute Edinburgh and the focus of the modelling at this stage is firstly identifying and investigating any relationships between patient demographics and model parameters. Once the model is formulated and relationships identified, we conduct a model evaluation, emphasising its predictive utility in both acute and multiple dosing scenarios. Additional to this, numerical approaches are used with the aim of investigating of a current clinical problem, APAP toxicity in young, low-weight individuals.

Chapter aims

- Identify relationships between patient demographics and Pop-PK model parameters.
- Investigate the effects of altering administration schedule on therapeutic effects.
- Investigate the effects of age and weight on maximum tolerable dose, with particular focus on young, low-weight individuals.

5.2. Methods

Detailed within this section is the mathematical modelling approach taken to computationally represent APAP metabolism in human. In order to ensure the parameters within the model were representative of the human APAP overdose case, we optimised the model parameters by fitting the model to data from over 3,600 APAP overdose patients. A description of the data is provided within this section. During model development, we tested multiple variations of PK models to identify the most predictive model framework. The model selection criteria used at each stage during the model development are also defined within this section.

5.2.1. Clinical data description

Clinical data was obtained from approximately 3,600 patients from 3 UK hospitals – Newcastle, Edinburgh and London. Data from 1,000 of these patients was published in the Lancet in 2017, within a study carried out by Clarke et al. [136]. The group aimed to determine whether novel biomarkers HMGB1 and K18 could accurately predict the requirement for prolonged hospital treatment, using ALT measurements as their predictive endpoint. This published data did not include all of the patient information. The full raw dataset was provided by our clinical collaborator, Dr. James Dear. For example, for each patient, the following information was obtained: Covariate/demographic information (namely, age, sex, and weight); APAP dose taken; time since APAP dose; and initial APAP concentration measurement (mg/l). The following observations were removed from the data set prior to analysis: patients with missing initial dose information; patients with missing time since dose information; patients with a missing initial APAP measurement; and patients with an initial APAP measurement below the lower limit of detection (LLOD), 10 mg/l [163].

5.2.2. Model selection criteria

Various Pop-PK models were tested throughout the model development process to determine the most predictive. These variations included both one and two compartmental models (Figure 5-2), and models which included different combinations of potential parameter/demographic correlations. Further details of these models can be found in sections 5.2.5 and 5.3.1. In order to determine the most suited model, the Akaike Information Criteria (AIC) and Bayesian Information

Criteria (BIC) were compared at each stage of the development process. Within this analysis, model parameters (within the Pop-PK model structure) were obtained by maximizing their likelihood, in relation to the observational data provided (Maximum likelihood estimation) [164]. The parameter estimation methodology is explained in further detail in section 5.2.4 of this chapter, however we also mention it here briefly. AIC and BIC are defined by:

$$AIC = -2LL(y; \hat{\theta}) + 2P,$$

$$BIC = -2LL(y; \hat{\theta}) + \log(N)P,$$

where LL is the log-likelihood function. Minimising minus two times the log-likelihood, $-2LL$, is equivalent to maximizing the log-likelihood function, LL [82]. Minus two times the log-likelihood, $-2LL$, is a function of y and $\hat{\theta}$ which are representative of the observational data and the estimated parameter, respectively. The total number of parameters to be estimated is represented by P , while N is the number of subjects. Both AIC and BIC measures were investigated at each stage in the model development process, and the model with the lowest combination was chosen to be best suited.

5.2.3. Mixed effects modelling background

If we assume that a sample dataset is quantitatively representative of the true population (for example, genetically homogeneous populations of mice are often bred for medical research purposes, making results of medical trials more uniform [165]), then simple regression models may be sufficient. This would allow the modelling approaches applied in previous chapters to be suitable. These type of models contain leading order parameters that can be assumed to remain constant across individuals (fixed effects). However, if we have reason to believe that certain individuals may respond differently to others (as we do with the human population [166]), this model should be extended to include so-called higher-order random effects. These random effects account for the fact that there may be significant differences between the observations in the subset of data at hand, and the general population. Mixed effects models are necessary for clinical population modelling due to the requirement of both fixed (population averages) and random (individual variation) effects. Fixed effect variables are those that cause differences in observations for reasons of which we are aware, for example, different doses of drug, different time-point measurements of concentration, different weights and

heights of individuals [167]. Random effect variables are necessary to account for the fact that there will be additional differences between individuals in the general population for reasons of which we are unaware. These unknown differences cannot be accounted for solely by the residual error of the dataset that the model is based upon.

Mixed effects modelling generally uses a two-stage hierarchical process [168], firstly creating the model at a population level and secondly at an individual level. This approach formalizes knowledge and assumptions about variation in outcomes and mechanisms both between (stage 1) and within individuals (stage 2) [169]. Once the model describing individual trajectories is embedded within the statistical population model, this then provides a suitable framework for inference of covariate effects (namely, obtaining a greater understanding of how much variance in model output may be attributable an individual's covariate information) [170] .

5.2.4. Model optimisation

Monolix is an advanced computational software programme that provides robust, global solutions for even the most complex pre-clinical and clinical population PKPD models [171]. We used Monolix [172] to estimate parameter values that can represent the clinical human population. Subsequently, parameters were also estimated on an individual level for individuals within this population. Here, I describe how the parameters within our model were estimated, firstly at a population level and secondly at an individual level.

5.2.4.1. Stage 1: Population level model

The parameter component of the population model aims to quantify average values for the whole population:

$$\theta_i = d(\beta, a_i, b_i), \quad i = \{1, 2, \dots, m\}, \quad (5-1)$$

where θ_i is the population parameter estimate; a function of the fixed parameter, β (fixed effects) is given by d ; parameter variation due to systemic association with individual attributes is given by a_i ; and parameter variation that cannot be explained in the population is given by b_i . The total number of population parameters to be estimated is represented by m . In our Pop-PK model (equation (5-23)), we have $m = 5$ parameters to be estimated: $t_{lag}, k_{12}, k_{21}, V, k_{el}$.

5.2.4.2. Parameter estimation – Stochastic Approximation Expectation-Maximisation (SAEM) Algorithm

Estimating the unknown population parameter estimates is a key task in non-linear mixed effect modelling. When faced with an incomplete dataset (clinical data is a perfect example – the data is clearly a subset and does not represent how the whole population behaves), from which unknown parameter estimates are to be estimated, expectation maximisation (EM) methods are particularly useful for parameter optimisation [173]. The key idea that underlies EM methods and differentiates them from other optimisation techniques is the introduction of a latent variable, Z , which is representative of a “complete-data space”. This allows the enhancement of incomplete data by making reasonable guesses of any missing information. EM methods are used for solving maximum likelihood estimation problems (maximising the likelihood function of parameters over a given search space given observational data). EM algorithms are also iterative, meaning they maximize successive local approximations of the likelihood function.

Each iteration k has two phases, an exploratory phase and a smoothing phase.

Exploratory phase – this searches the entire given parameter space to find the neighbourhood of maximum likelihood.

During the exploratory phase, parameters for each individual within the dataset are generated from the conditional distribution first, and then the population parameters are generated based upon these. The conditional distribution is the probability distribution of the individual parameters and is defined for each iteration k as follows:

$$p(\psi_i|y_i, \theta^k), \quad (5-2)$$

where the individual parameter estimate of individual i is represented by ψ_i , the observation or data point for individual i is represented by y_i and the population parameter estimate from the previous iteration is represented by θ^k . For the first iteration, this parameter estimate will be equal to the initial guess.

There is no analytical solution to (5-2). Instead, Kuhn and Lavielle proposed combining the SAEM algorithm with a Markov-Chain Monte Carlo (MCMC) procedure [174]. When an estimated value is required but is difficult to obtain (in our instance individual parameter estimates, ψ_i), simulation approaches such as Metropolis-Hastings can sample from likely distributions to obtain the best estimate possible in the following way [175]. Firstly, the Markov-Chain Monte-Carlo approach

takes a random walk from a user defined initial point within the distribution, to another (more probable) point within the distribution. Then, at each iteration, the probability of the newly proposed point is only dependent on the one point previous, not every point previous. This general definition is applicable in our case, since the probability of the individual parameter, ψ_i , is only dependent on the population parameter from the previous iteration, θ^k , and not all of the parameter estimates from all of the previous iterations. As seen in equation (5-2), the probability of the individual parameter also has a dependency on the data point for that individual, as this helps to pull the search closer to the true maximum likelihood [176]. Multiple iterations then allow the most probable value from the whole distribution to be obtained [177]. The general Metropolis Hastings algorithm described by Robert et al. [175] can therefore be altered to represent the concepts underlying the estimation of our population parameter estimates, as follows.

1. Provide an initial estimate for the individual estimate value, ψ_i .
2. Choose a nearby point within the parameter space to represent the newly proposed individual parameter estimate, ψ_* .
3. Use the Metropolis Hastings algorithm as follows to determine whether the proposed point should be accepted/rejected
 - I. Sample u from a distribution $U(0,1]$.
 - II. If $u < \min \left\{ 1, \frac{p(y_i|\psi_*; \theta^k)}{p(y_i|\psi_i; \theta^k)} \right\}$
 - i. $\psi_{i+1} = \psi_*$
 - else,
 - ii. $\psi_{i+1} = \psi_i$.
4. The process is then repeated, choosing a different nearby point each time, until the most probable value is found.

Once the most probable values are found for each individual, the average of the individual parameter estimates from this iteration is then taken to represent the proposed population parameter estimate, θ_{k+1} , as

$$\theta_{k+1} = \frac{1}{N} \sum_{i=1}^N \psi_i, \quad (5-3)$$

where N is the total number of individuals within the dataset. This whole process is repeated k times until results from iterations begin to converge and a resultant local

neighbourhood is found. This concludes the exploratory phase of this iteration of the EM algorithm.

Following the exploratory phase, is the final (smoothing) stage of the EM algorithm.

Smoothing phase – Once the local neighbourhood is found, this subsection of the parameter space is further explored.

Individual parameter estimates are generated using MCMC as previously described (5-3).

This time, the population parameter estimates are not taken to be the average of the individual parameter estimates at the current iteration, but instead the average across all previous iterations. This can be mathematically described as follows

$$\theta_{k+1} = \frac{1}{k} \left[\frac{1}{N} \sum_{i=1}^n \psi_i^1 + \frac{1}{N} \sum_{i=1}^n \psi_i^2 + \dots + \frac{1}{N} \sum_{i=1}^n \psi_i^k \right], \quad (5-4)$$

Where ψ_i^1 is the individual parameter for individual 1, and ψ_i^2 is the individual parameter 2. This can be equivalently written as,

$$\theta_{k+1} = \frac{1}{k} \left[(k-1)\theta^k + \frac{1}{N} \sum_{i=1}^n \psi_i^k \right], \quad (5-5)$$

which simplifies to,

$$\theta_{k+1} = \theta^k + \frac{1}{k} \left[\frac{1}{N} \sum_{i=1}^n \psi_i^k - \theta^k \right]. \quad (5-6)$$

This process is continued until there is negligible change in the objective function.

5.2.4.3. The simulated annealing SAEM algorithm

SAEM convergence can strongly depend on the initial guess if the likelihood possesses several local maxima [178]. To try to improve convergence towards the global maximum, the simulated annealing version of SAEM is introduced [179]. This simulated annealing version fixes the residual error variance to a large value when initially trying to optimise the model parameters. The large values of the variance allow the estimated values to be less concentrated around its mode. This allows the sequence (θ_k) to escape from the local maxima and converge to a neighbourhood of the global maximum. Once this initial process is complete, the usual MCMC-SAEM algorithm is used [176].

5.2.4.4. Stage 2: Individual level model

Once average population (fixed effect, and corresponding error) parameters have been estimated, estimates for a new individual from that given population can be estimated:

$$Y_{ij} = f(X_{ij}; \psi_i) + e_{ij}, \quad j = \{1, 2, \dots, n_i\}, \quad (5-7)$$

where the outcome, Y_{ij} (for each individual i at each time point j) is a function of the variables, X_{ij} (for example, time); and fixed regression parameters, ψ_i ; with the addition of some intra-individual error, e_{ij} . The total number of observations for each individual i is represented by n_i .

5.2.4.5. Parameter estimation – conditional distribution

In order to calculate the most probable value of the individual parameter estimates, the conditional parameter distribution is firstly calculated, which is described by (5-2) but detailed here for completeness. In this case, the parameter definition is defined by

$$p(\psi_i | y_i, \hat{\theta}), \quad (5-8)$$

where ψ_i are the individual parameters for individual i , $\hat{\theta}$ are the estimated population parameters and y_i are the observations for individual i . This distribution is representative of the uncertainty in an individual's parameter value taking into account the following:

1. Information available for the individual, which includes:
 - a. Observed data for the individual; and
 - b. Covariate (or demographic) values for them
2. The fact that the individual belongs to a population (for which the parameters have already been estimated).

Within Monolix, there are two possible methods for estimating the most probable value arising from this distribution, either calculating Empirical Bayes Estimates (EBEs) or Markov Chain Monte Carlo (MCMC) sampling from a conditional distribution. I will now summarize both approaches.

5.2.4.6. Empirical Bayes Estimates

Empirical Bayes Estimates (EBEs) are alternatively referred to as the conditional mode, ψ_i^{mode} , which represents the most common individual parameter; they are

found by maximising the aforementioned conditional distribution and therefore represent the most likely value of the individual parameter, ψ_i , given the observations, and a given population distribution:

$$\psi_i^{mode} = \arg_{\psi} \max p(\psi_i | y_i; \hat{\theta}). \quad (5-9)$$

It is therefore necessary to maximize the conditional probability with respect to the individual parameter value, ψ_i . In Monolix, the Nelder-Mead search algorithm is used to find this maximum. Since $p(\psi_i | y_i; \hat{\theta})$ cannot be easily solved for a given ψ_i , this means that the conditional distribution does not have a closed form solution, and therefore Bayes' Law is used to rewrite the conditional distribution in the following way:

$$p(\psi_i | y_i) = \frac{p(y_i | \psi_i) p(\psi_i)}{p(y_i)}. \quad (5-10)$$

If the individual parameter values ψ_i are known, then the conditional density function of the data given these parameter values ($p(y_i | \psi_i)$) is simple to calculate, as is the density function for the individual parameters (ψ_i), since they both have closed form solutions. However, the denominator of (5-10) (the density function of the observations, $p(y_i)$), has no closed form solution. Since the solution of $p(y_i)$ will have no dependency on ψ_i however, this value would remain the same in each calculation of $p(\psi_i | y_i)$, having no direct effect on the maximum value. This element $p(y_i)$ is therefore removed from equation (5-10) for the optimisation procedure, and only $p(y_i | \psi_i) p(\psi_i)$ is optimised.

Once the most probable individual parameter values, ψ_i^{mode} are known, using the population parameters and the covariates, the corresponding individual random effects can then be calculated using

$$\eta_i = \psi_i^{mode} - \psi_{pop} - \beta c_i, \quad (5-11)$$

where η_i is the individual random effect, ψ_i^{mode} is the estimated conditional mode, ψ_{pop} and β are population parameters, and c_i is covariate data for individual i .

For each individual, then, we have a fixed effect and random effect. An individual's parameter value is found by randomly sampling from a normal distribution with the mean value equal to the fixed effect and the standard deviation value equal to the random effect.

5.2.4.7. MCMC sampling of conditional distribution

MCMC methods can be used to obtain samples from the previously predicted conditional distribution, when direct sampling is difficult for individual parameter optimisation. This approach involves a stochastic procedure which yields the most likely value from the probability distribution of interest.

We can use this approach to conclude the most probable value of the individual parameter estimate as the mean value from all of the accepted MCMC draws,

$$\psi_i^{mean} = \frac{1}{\Omega} \sum_{\Omega=1}^{\Omega} \psi_i^{\Omega}, \quad (5-12)$$

where ψ_i^{mean} is the average individual parameter value over all of the samples, ψ_i^{Ω} is the individual parameter value for the current MCMC draw, and Ω is the number of accepted MCMC draws.

5.2.5. Covariate model

An individual's covariate (or demographic) information such as age, gender and weight may affect one or more of that individual's structural parameters. If this is the case, the individual parameter affected should be a function of this continuous covariate. Initially, no covariates were included within model optimisation, but all covariate relationships with individual parameters were tested to determine whether or not inclusion would improve the model fitting criteria. The Pearson correlation coefficient is calculated to determine the relationship between demographic (covariate) information, and the parameter value under investigation. For continuous covariates, the null hypothesis for this test is as follows,

H_0 : The Pearson correlation coefficient between the individual parameter and the covariate (demographic) value is zero. A small p-value indicates that the null hypothesis can be rejected, in which case a correlation is present and the covariate should be included within the model.

Once it is determined that a significant correlation is present, the Wald test is then used to analyse whether including the covariate relationship brings enough change to the model output to validate its inclusion. The null hypothesis of the Wald test is as follows,

H_0 : The approximated beta parameter (from the Pearson correlation test) is equal to zero. A small p-value indicates that the null hypothesis can be rejected, and

therefore the estimated beta parameter is significantly different from zero, subsequently warranting inclusion within the model.

5.2.6. The non-linear mixed effects model

Now that we have described how an individual's fixed effects and random effects can be approximated, we define a general non-linear mixed effects model for continuous outputs:

$$Y_{ij} = f(X_{ij}; \psi_i) + g(X_{ij}; \psi_i; \xi)e_{ij}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq n_i, \quad (5-13)$$

where Y_{ij} is the j th observation of subject i ; the structural model $f(X_{ij}; \psi_i)$ depends on both the regression variables X_{ij} , and a vector of individual parameters ψ_i , the residual error model $g(X_{ij}; \psi_i; \xi)$ depends on both the regression variables and individual parameters but also an additional vector of parameters, ξ , the residual error for individual i is represented by e_{ij} , the number of subjects is represented by N ; and n_i is representative of the number of observations for subject i .

As mentioned, the vector ψ_i is a vector of individual parameters:

$$\psi_i = (\mu, c_i, \eta_i), \quad (5-14)$$

where c_i is a known vector of covariates, μ is an unknown vector of fixed effects, and η_i is an unknown vector of normally distributed random effects.

The residual errors (e_{ij}) are random variables with mean 0 and variance 1, and the residual error model is defined by the function g . This residual error model is further explained in the section below.

5.2.6.1. Residual error model

There are various potential reasons for which a model may not predict the exact observed paracetamol concentration. The structural model used for prediction may only be an approximation of the biology, or the PK responses may have been measured with some level of assay error, for example. The observed difference between actual values and predicted values is known as residual error. Each individual prediction (inclusive of model predictions and residual error predictions) can be described by equation (5-13).

With focus on the error component of the model, $g(X_{ij}; \psi_i; \xi)e_{ij}$, we are reminded that the residual errors are standardized Gaussian random variables with mean 0

and standard deviation 1. It follows then, that $f(X_{ij}; \psi_i)$ is the conditional mean, $E(y_{ij}|\psi_i) = f(X_{ij}; \psi_i)$. It also follows that $g(X_{ij}; \psi_i; \xi)$ is the standard deviation, $sd(y_{ij}|\psi_i) = g(X_{ij}; \psi_i; \xi)$.

In Monolix, the function g is considered to be a function of the structural model, i.e.,

$$g(X_{ij}; \psi_i; \xi) = g(f(X_{ij}; \psi_i); \xi), \quad (5-15)$$

which leads to an expression of the observation model of the form

$$Y_{ij} = f(X_{ij}; \psi_i) + g(f(X_{ij}; \psi_i); \xi)e_{ij}. \quad (5-16)$$

There are many functions that can be chosen to represent the residual error model g . A constant error model assumes that the $f(X_{ij}; \psi_i)$ element of g is eliminated, resulting in the residual error function, g , being constant and the observation model can then be written in the form

$$y = f + \xi e, \quad (5-17)$$

where y is the predicted observation, f is the structural model, ξ is an error related parameter, and e is the residual error. As a reminder, the structural model is dependent upon the model variables, X_{ij} , the individual parameters, ψ_i .

Alternatively, a proportional error model assumes that the residual error function, g , is proportional to the structural model, f , meaning that this element is no longer eliminated from g . The observation model can then be written in the form

$$y = f + \xi f e. \quad (5-18)$$

However, in our analysis, we chose to model the residual error as a linear combination of a constant error term and a term proportional to the structural model. The observational model can therefore be written in the form

$$y = f + (\xi_1 + \xi_2 f)e, \quad (5-19)$$

where y is the predicted observation, f is the structural model and e is the residual error. In this case, there are two error related parameters ξ_1 , a constant error related parameter and ξ_2 , a proportional error related parameter. We chose this combined residual error model to firstly account for any error that may appear due to the structural model being only an approximation of the biology of APAP metabolism in human (proportional error), and secondly to account for any observations near to

the lower limit of detection to ensure that bioassay error is considered (constant error).

Prior to optimisation, the data and parameters were transformed to the log-normal distribution, to ensure the Gaussian assumption is met, and to also ensure that no negative values could be predicted, since negative blood concentrations are not possible. The logit error model was extended to assume that observational concentrations, y , were bounded. Namely, $A < y < B$, with $A = 0$ mg/l being the lowest possible concentration and $B = 50,483$ mg/l being the maximum observed concentration in the dataset.

The transformed model can therefore be represented by

$$T(y_{ij}) = T\left(f(x_{ij}, \psi_i)\right) + g(x_{ij}, \psi_i, \xi)e_{ij}, \quad (5-20)$$

Where T is the transform function, f is the structural model and g is the error model. In order to ensure strictly positive results ranging between A and B , the transform function can be represented by

$$T(y) = \log\left(\frac{y - A}{B - y}\right), \quad (5-21)$$

Applying this transformation to our model, we obtain:

$$y = A + (B - A) \left(\frac{f - a}{f - A + (B - f)e^{-ge}} \right). \quad (5-22)$$

5.3. Results

5.3.1. APAP Pop-PK model formulation

To understand how to effectively model the clinical APAP dynamics, the observed APAP concentrations were transformed into log-space and plotted over time (Figure 5-1).

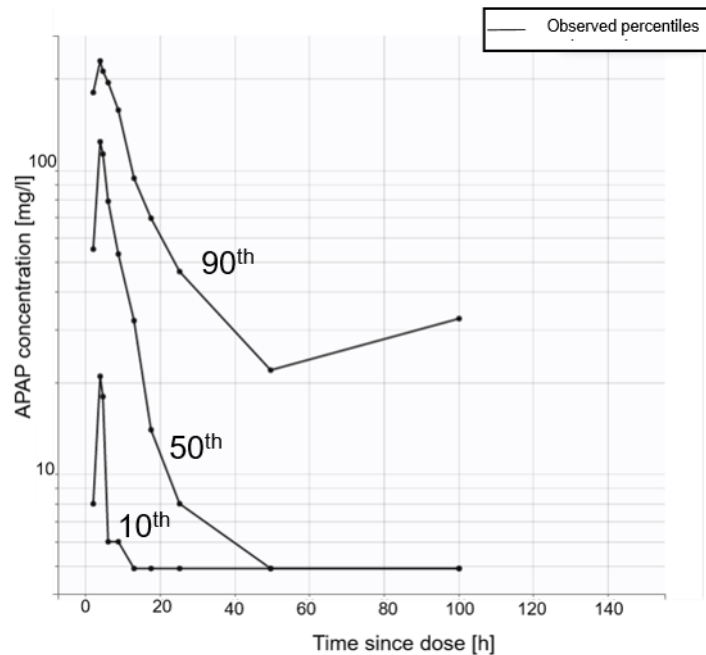


Figure 5-1: Log-transformed paracetamol concentration versus time. The highest trend represents the 90th percentile, meaning that 90% of the observed data falls below this line at these time points. The middle trend is the 50th percentile, meaning that 50% of the observed data falls below this line. The bottom-most line is the 10th percentile, meaning that 10% of the observed data falls below this line.

Since it was not clear whether the distribution/elimination phase of APAP occurs over one or two phases (Figure 5-1), we trialled both one and two-compartmental models. In a one-compartment model, all tissues within the body are assumed to be contained within one compartment [92]. A two-compartmental model assumes that tissues within the body can be separated into two different compartments [92]. The first compartment can be thought of as consisting of well-perfused tissues such as heart, liver and brain. The secondary compartment could represent poorly perfused tissues such as fat and bone. [93]. Schematics of both the one-compartment and two-compartment models can be seen in Figure 5-2. Models representing APAP concentration over time (Cc_t) were built based on the laws of mass action, following a similar methodology as in our pre-clinical APAP model developed in section 2.2.1. Parameters from these models were optimised using the methodology defined in section 5.2.4 and models were selected based on their AIC/BIC values (a description of this criteria is defined in section 5.2.2). Model selection criteria for both the one-compartment and two-compartment models are described in Table 5-1.

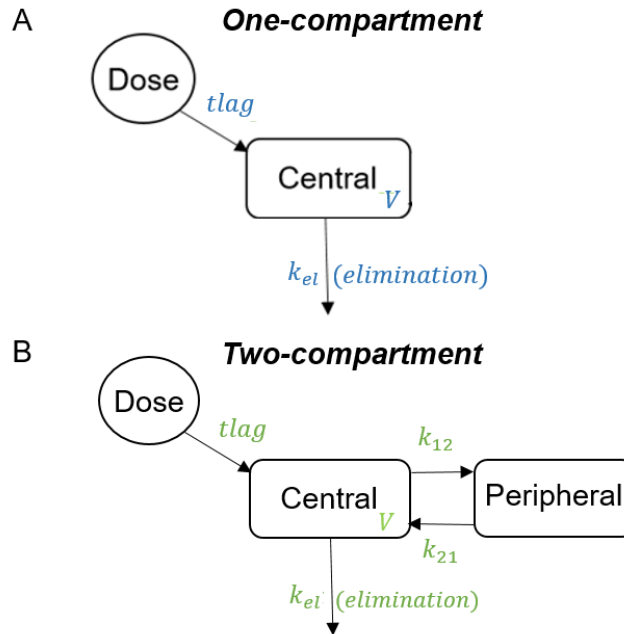


Figure 5-2: Pharmacokinetic model schematic. (A) *One-compartment PK model representing the distribution of APAP. A dose of APAP is transferred to the central compartment but the rate of this transfer is limited by a parameter, t_{lag} . APAP is then eliminated from the central compartment at a rate k_{el} . The central compartment has a theoretical volume, V .* (B) *Two-compartment PK model representing the distribution of APAP. A dose of APAP is transferred to the central compartment (well perfused tissues) but the rate of this transfer is limited by a parameter, t_{lag} . The central compartment has a theoretical volume, V . APAP is secondarily transferred to the peripheral compartment (poorly perfused tissues) at a rate k_{12} , and transferred back from the peripheral compartment to central compartment at a rate k_{21} . APAP is eliminated from the central compartment at a rate k_{el} .*

Compartmental model	Akaike Information Criteria (AIC)	Bayesian Information Criteria (BIC)	-2LL
One	2.67×10^5	2.67×10^5	2.67×10^5
Two	2.64×10^5	2.65×10^5	2.65×10^5

Table 5-1: Model fitting criteria for population pharmacokinetic models. *AIC, BIC, and -2LL are all compared for one- and two-compartment models.*

As can be seen in Table 5-1, the two-compartmental model provided marginally lower values for all three metrics, AIC, BIC and -2LL, and therefore was the PK model taken forward to represent the data. The chosen model is almost identical to the pre-clinical mechanistic APAP model defined in Chapters 2 and 3 (sections 2.2.1 and 3.2.3 respectively), with the difference being that IP administration is not assumed and therefore an absorption rate from the peritoneal cavity is not required. Instead, we assumed an oral dose with absorption from the stomach, with a time

delay representing digestion/lag time. The corresponding model for APAP concentration over time in the central compartment (Cc_t) can be written as follows:

$$Cc_t = \begin{cases} 0, & t < tlag, \\ D(Ae^{-\alpha\omega} + Be^{-\beta\omega} - (A+B)e^{-k_a\omega}), & t \geq tlag, \end{cases} \quad (5-23)$$

where

$$\alpha = \beta \left(\frac{k_{21}}{k_{el}} \right),$$

$$\beta = 0.5(k_{12} + k_{21} + k_{el}) + \sqrt{(k_{12} + k_{21} + k_{el})^2 - (4k_{21}k_{el})},$$

$$\omega = t - tlag, \quad A = \frac{k_a}{V} \left(\frac{k_{21} - \alpha}{k_a - \alpha} \right) (\beta - \alpha), \quad B = \frac{k_a}{V} \left(\frac{k_{21} - \beta}{k_a - \beta} \right) (\alpha - \beta).$$

The dose amount of APAP (mg) is represented by D , the parameter that limits transfer of the oral dose to the central compartment due to, for example, digestion is represented by $tlag$. The theoretical volume of the central compartment (l/kg) is represented by V , the rate at which APAP is transferred to the peripheral compartment (poorly perfused tissues) (h^{-1}) is represented by k_{12} , the rate at which APAP transferred from the peripheral compartment to central compartment (h^{-1}) is represented by k_{21} , the rate at which APAP is eliminated from the central compartment (h^{-1}) is represented by k_{el} and t represents the current time point (h).

5.3.2. Identifying relationships between patient demographics and model parameters

Multiple variations of the population PK model were fitted to the clinical dataset with the aim to optimise the APAP PK parameters such that computational simulations could mimic the clinical dataset provided. During initial model development, no demographic/parameter relationships were included in the model, although all possible combinations were statistically tested to determine whether their inclusion could improve the model fitting criteria. Results from the Pearson and Wald tests suggested that the final correlations to be included in the model should be between $tlag$ and age (negative correlation); k_{12} and age (positive correlation); and k_{21} and sex (positive correlation), where 1=male and 2=female. Literature searches were carried out to try and understand/confirm these potential relationships. Mian et al. reviewed 27 PK studies [180]. Of those defined within the review (that investigated a possible link between $tlag$ and age), none found a significant result. It must be

noted, however, that only a small number of studies investigated this correlation (n=3) and none of these studies incorporated inter individual variability.

5.3.3. Model parameterisation

The mixed-effects Pop-PK model resulted in optimised parameter values representative of the population, a random individual, and the level of correlation between an individual subject's parameter estimate and their demographic information. Optimised population parameter values are provided in Table 5-2.

Parameter	Description	Value
$tlag_{\theta}$	Parameter limiting the rate of transferring an APAP dose to the central compartment due to, for example, digestion	3.16
$k_{a_{\theta}}$	APAP absorption rate	0.381
V_{θ}	Theoretical volume of central compartment	0.0213
$k_{12_{\theta}}$	Rate of APAP transfer from central to peripheral compartment	0.929
$k_{21_{\theta}}$	Rate of APAP transfer from peripheral to central compartment	1.2
$k_{el_{\theta}}$	APAP elimination rate from central compartment	0.000479

Table 5-2: Optimised population parameter estimates for clinical APAP Pop-PK model.

Individual values, η_i , were randomly sampled for each individual and were assumed to be normally distributed, with mean of 0 and variance of ω^2 . The variance for each parameter is provided in Table 5-3.

Parameter	Description	Value
$tlag_{\omega}$	Variance of the parameter limiting the rate of transferring an APAP dose to the central compartment due to, for example, digestion	1
$k_{a_{\omega}}$	Variance of the APAP absorption rate	1.24
V_{ω}	Variance of the theoretical volume of central compartment	1.75
$k_{12_{\omega}}$	Variance of the rate of APAP transfer from central to peripheral compartment	1.28
$k_{21_{\omega}}$	Variance of the rate of APAP transfer from peripheral to central compartment	1.02
$k_{el_{\omega}}$	Variance of the APAP elimination rate from central compartment	4.17

Table 5-3: Optimised variance of parameter values for each individual Pop-PK parameter.

As detailed in Section 5.3.2, some parameter estimates were correlated with an individual's demographic information. Coefficients for these correlations are detailed in Table 5-4.

Parameter	Description	Value
β_tlag_{age}	Coefficient of the correlation between the APAP lag time in absorption into the central compartment, and an individual's age	-0.00585
β_k12_{age}	Coefficient of the correlation between the rate of APAP transfer from central to peripheral compartment, and an individual's age	0.0163
β_k21_{sex}	Coefficient of the correlation between the rate of APAP transfer from peripheral to central compartment, and an individual's gender	0.5

Table 5-4: Optimised coefficients of significant correlations between Pop-PK parameters and an individual's demographic information.

Assuming a log-normal distribution, to allow positive values only for predicted APAP concentration, the individual and population parameter values were combined to provide a parameter value for the i^{th} subject as follows:

$$k_{a_i} = k_{a\theta} e^{k_{a\eta}},$$

$$V_i = V_{\theta} e^{V_{\eta}},$$

$$k_{el_i} = k_{el\theta} e^{k_{el\eta}}.$$

For those parameters with significant correlations, the parameter values for the i^{th} subject were as follows:

$$tlag_i = tlag_{\theta} e^{tlag_{\eta}} + \beta_tlag_{age}(AGE),$$

$$k12_i = k12_{\theta} e^{k12_{\eta}} + \beta_k12_{age}(AGE),$$

$$k21_i = k21_{\theta} e^{k21_{\eta}} + \beta_k21_{sex}(SEX),$$

where *AGE* and *SEX* were the actual demographic values for the i^{th} subject.

5.3.4. Model evaluation

We tested the predictivity of the non-linear mixed effects model by plotting simulated output against the observed clinical data (Figure 5-3). For each individual observed data point, the corresponding individual model output is plotted as a blue circle. As a point of reference, in a model that was 100% predictive, these blue circles would fall along the black solid diagonal line. If observations fall above the diagonal line, this would be an example of model under-prediction, since the observed value would be higher than the predicted value. Conversely, if observations fall below the line,

this would be an example of model over-prediction, since the observed value would be lower than the predicted value.

Additionally, individual weighted residuals (IWRES) were plotted versus both predicted time and APAP concentrations (Figure 5-4). In order for the model to be deemed predictive, predictions are required to remain within approximately ± 1.96 standard deviations [181]. As can be seen in Figure 5-4, for the most part, this requirement is met, with only very few predictions falling outside of this range (approximately 1.05% in the IWRES versus time case and approximately 0.5% in the IWRES versus APAP concentration case).

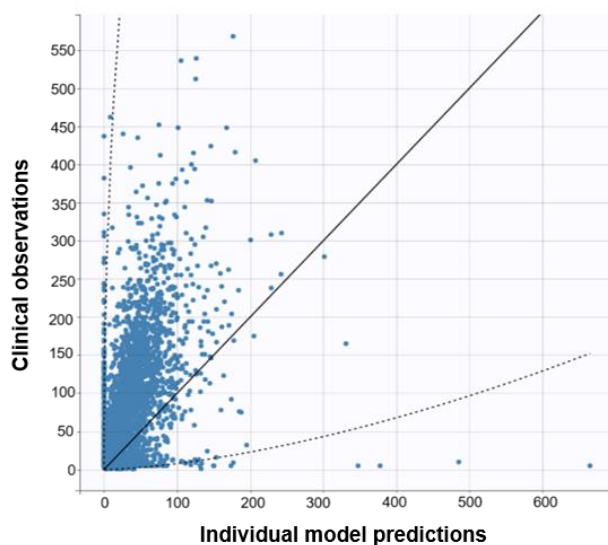


Figure 5-3: Observational data versus individual predictions. The black solid line is the reference $y = x$ line. The black dotted lines define the region containing 95% of the data.

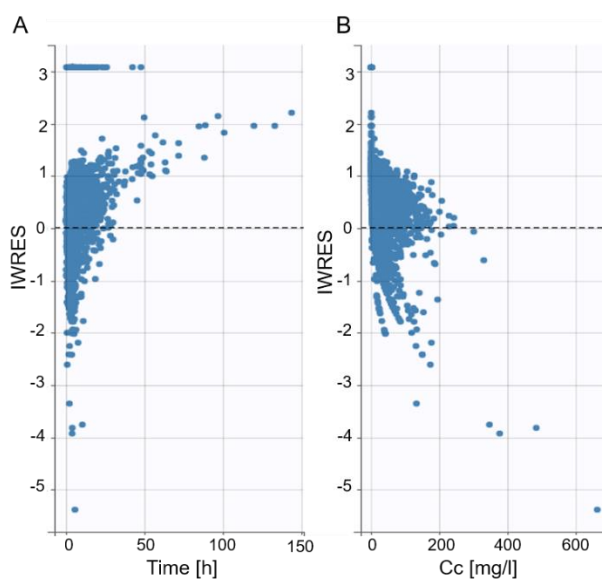


Figure 5-4: Individual weighted residuals (IWRES) comparison. (a) IWRES versus time and (b) IWRES versus APAP concentration (Cc). The black dashed line is a reference line at $IWRES=0$. Observations should ideally fall between ± 1.96 standard deviations.

In Figure 5-5, we provide a Visual Predictive Check (VPC) of the predictivity of our population pharmacokinetic model. If our model was completely predictive of the real clinical environment, we would expect the predicted percentiles (--) to directly overlap with the empirical percentiles (-). Additionally, the prediction intervals (shaded regions) would ideally encase the empirical percentiles. The upmost percentile and prediction intervals are representative of the 90th percentile, meaning that 90% of the data is either observed (clinically) below this line (-) or are predicted by the model to fall below this line (--). The upmost prediction interval (blue shaded area) means that there is a 95% chance that the 90% predicted percentile will fall within this range. Following these descriptions then, we can see that although our model slightly over predicted the peak APAP concentration, the predicted outputs were reasonably similar to the observational outputs.

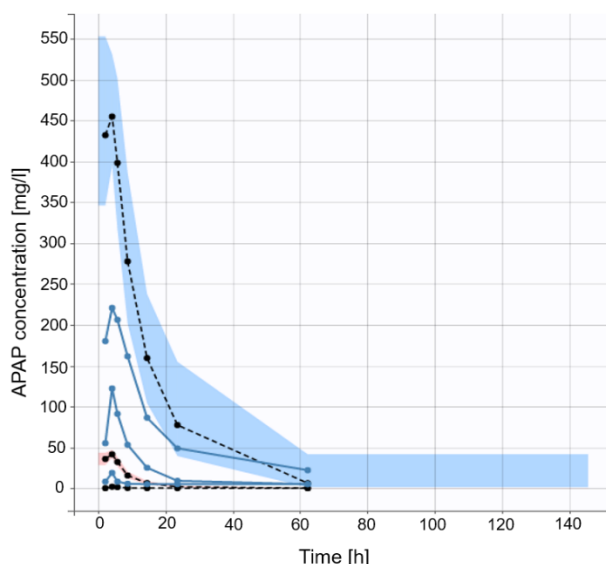


Figure 5-5: Visual Predictive Check (VPC). Empirical percentiles are represented by blue solid lines. Predicted percentiles are represented by black dashed lines. The highest trend is the 90th percentile, central trend is the 50th percentile and the bottom-most trend is the 10th percentile. The blue shaded region around the 90th percentile is the 95% prediction interval, meaning there is a 95% chance the simulated prediction interval will fall within this range. The 95% prediction interval for the 50th percentile is indicated by a red shaded region.

5.3.5. Model simulation

5.3.5.1. Acute dosing

Individual simulations

The population-pharmacokinetic model was firstly used to simulate an individual taking acute doses of paracetamol. In the first instance, paracetamol concentration profiles were explored for a 25-year-old, 45kg female taking two different doses,

1,000 mg (2 tablets) and 336,000 mg (672 tablets, which was the maximum amount defined in the clinical data used within this investigation). Results of the simulations can be seen in Figure 5-6.

Figure 5-6(A) shows the predicted APAP concentration profile following a therapeutically recommended dose of APAP (1,000 mg). Considering the therapeutic window for the maximum APAP concentration (C_{max}) is predicted to be between 5-20 mg/l, and APAP is recommended to be taken every four hours, the predicted profile for an acute dose was as expected; almost reaching the therapeutic threshold a few hours-post-dose, but indicating that an additional dose may be required for full therapeutic effects. If a dose much higher than the recommended dose was taken acutely by the same individual, the predicted difference in profiles was clear (Figure 5-6(B)). The threshold for toxicity is thought to be any APAP C_{max} above 25 mg/l. As expected, for a dose 336 times the recommended amount, the predicted peak concentration far exceeded this, at almost 3,000 mg/l.

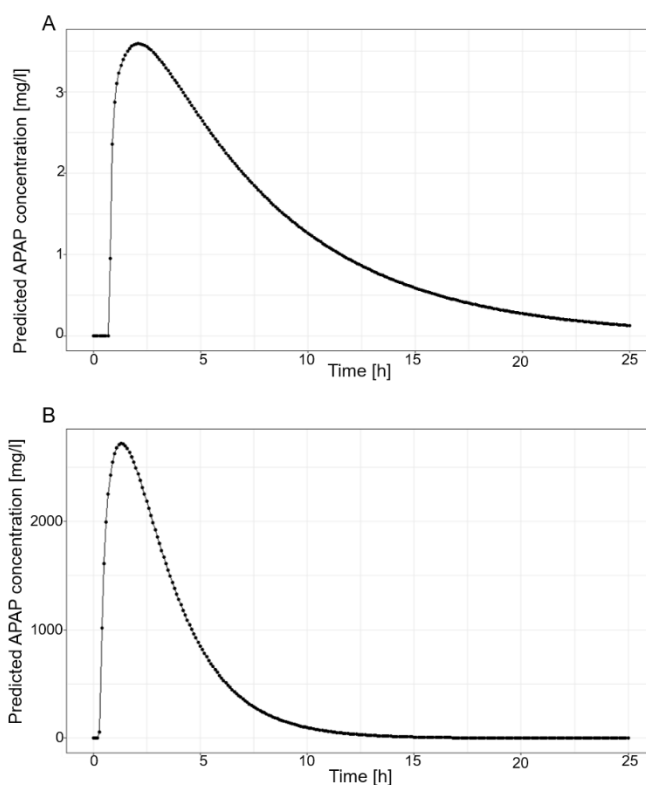


Figure 5-6: Simulated concentration over time for an individual taking two different acute APAP doses. Simulated APAP concentration over time for a 25-year-old female weighing 45kg following an APAP dose of 1,000 mg (A) and 336,000 mg (B).

Multiple individual simulations

Although a patient may have similar demographics, e.g. weight, age and gender, their physiological parameters (e.g. rate of APAP absorption, rate of APAP elimination) will differ due to inter-individual variability. Taking this inter-individual variability into account, the model was used to simulate three individuals with the same demographic information (gender - female, age - 25 weight - 45kg) receiving a high APAP dose (336,000 mg) to see how their paracetamol concentration profiles differed.

The clear differences between simulated individuals (with the same demographic information) receiving the same highly toxic APAP dose, can be seen in Figure 5-7. Although, as expected, all simulated individuals had predicted concentrations that far exceeded the toxic threshold (25 mg/l), the predicted profile shapes appeared to be very different for the three simulations. While two of the simulations showed fairly similar absorption and clearance of APAP, the third simulation had a much more rapid absorption of APAP, leading to a much higher maximum APAP concentration (C_{max}), although the clearance was also much quicker, resulting in the APAP being cleared by around the same time as the other two simulations. The visual differences in profiles for only three simulations, shows the importance of accounting for this inter-individual variability when making population predictions.

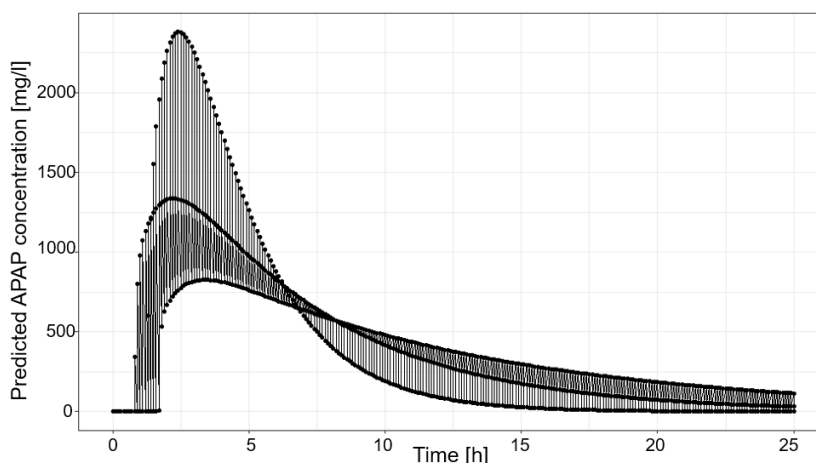


Figure 5-7: Simulation of APAP concentration over time for multiple individuals with the same demographic information taking the same 336,000 mg dose. Portrayal of the variation in simulated outputs when inter-individual variability of parameters is taken into account.

Population simulations

Since differences in APAP concentration profiles could be clearly seen in individuals with the same demographics, it was clear that simulations of a high number of

individuals would be necessary to make predictions for a more representative population. For this reason, acute dosing was simulated for 1,000 individuals, each having the same covariate/demographic information (gender, age, weight) but taking into account inter-individual variability in their physiological parameters. From these 1,000 simulations, predicted population profiles were calculated (Figure 5-8).

Predicted population profiles were visually as expected. For example, in Figure 5-8(A) we see that for an acute therapeutic dose of 1,000 mg, 95% of females who weigh 45kg were predicted to have a C_{max} that was within the therapeutic range (5-20 mg/l). Figure 5-8(B) shows that by increasing this dose to 4,000 mg (a daily recommended dose in one acute instance), 95% of females were predicted to have a C_{max} above the toxic threshold (25 mg/l). If this dose was increased even further, to an amount chosen to represent the maximum dose recorded in the clinical dataset (336,000 mg/672 tablets), visually, we see in Figure 5-8(C), that 95% of the population were predicted to far exceed the toxic threshold.

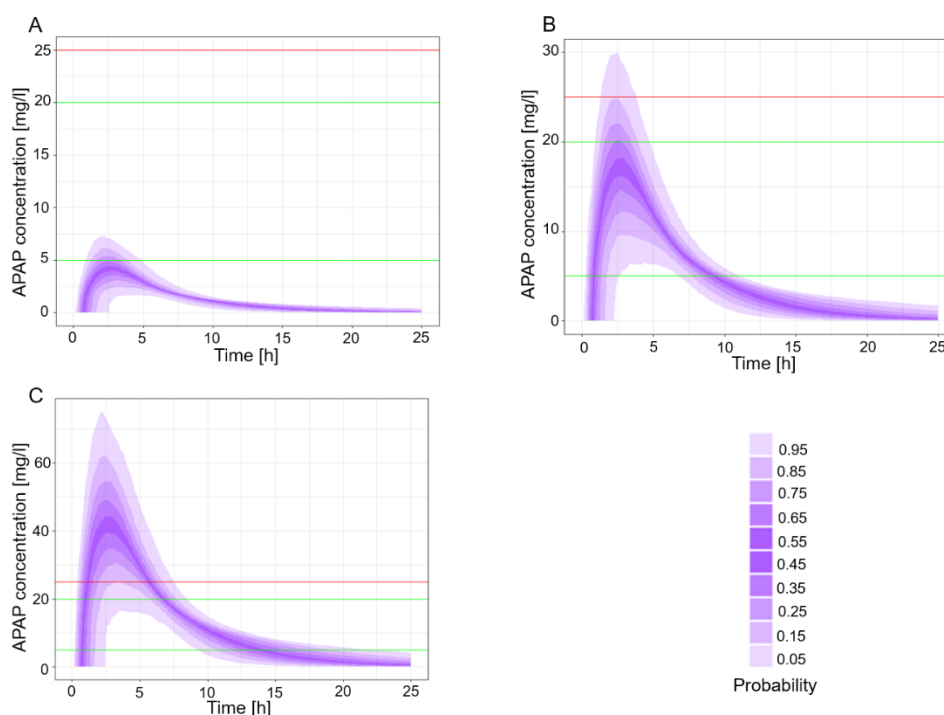


Figure 5-8: Predicted population APAP concentration profiles for varying doses. For 25-year-old women weighing 45kg, there is a 95% chance that their profile will be within the upper and lower bound (palest purple bounds). There is a 55% chance that their profile will be within the inner-most band (darkest purple). The therapeutic index for the maximum of the concentration profile (C_{max}) is bounded by the green horizontal lines (5-20 mg/l). The threshold for toxicity is indicated by the red horizontal line (25 mg/l). If the C_{max} is above this threshold, toxicity is predicted by the model. (A) shows the population predicted profiles for individuals of this demographic taking an acute dose of 1,000 mg APAP. (B) shows the population predicted

profiles following a 4,000 mg dose. (C) shows the population predicted profiles following a 36,000 mg dose.

5.3.5.2. Multiple dosing

Although the simulated output was as expected for the acute dosing situation, paracetamol is therapeutically advised to follow a multiple dosing regimen. Prior to using the model for further investigation, the model was exposed to a multiple dosing regimen for validation.

Test population introduction

To test the predictive capacity of the model in the multiple dosing scenario, three test subjects were exposed to the recommended APAP dose (1,000 mg) 4 times daily. Test subjects were created by providing demographic information on three known individuals – namely myself, Chantelle Mason (gender – female, age – 25 years, weight 47 kg), and my supervisory team, Dr Steven Webb (gender – male, age – 44 years, weight 92 kg), and Dr Joseph Leedale (gender – male, age – 32 years, weight 86 kg).

Recommended dosing

The recommended dose (1,000 mg, 4 x daily) was firstly assumed to be taken at times typical of breakfast (7am), lunch (12pm), dinner (5pm) and bedtime (9pm). Resultant paracetamol concentration profiles can be seen in Figure 5-9 for each test individual.

In Figure 5-9 we see that for all test individuals, concentrations did not appear to reach the therapeutic window following the first dose. However, following a second dose, the therapeutic window was reached. Profiles appeared to be generally similar for all three individuals, the main difference being Chantelle (orange) had a greater lag time (parameter limiting the transfer of the oral dose to the central compartment) than both Steven (purple) and Joseph (green). While it took longer for Chantelle's oral dose to enter the central circulation, Chantelle also took the longest to completely clear the APAP. If this one-day approach to dosing was taken, therapeutic effects were predicted to diminish while each of the subjects were sleeping (beyond around 1am). If the targeted pain was short-lived (e.g. a headache), this may have been a sufficient schedule. However, if post-operative pain relief was required, for example, it may have been necessary for the therapeutic effects to remain for longer than a one-day period. For this reason, the model was

used to simulate a repeated administration schedule, the outputs of which can be seen in Figure 5-10.

In this simulation scenario, the recommended dose (1,000 mg 4 x daily) was provided to all three test individuals, this time for two consecutive days. Figure 5-10 shows that predicted concentration profiles for all three test individuals predominantly remained within the therapeutic window as expected. Similar to the one-day-dosing simulation scenario, for all test individuals, taking the first dose of APAP did not result in predicted therapeutic effects. However, taking the first dose of APAP on the second day resulted in almost immediate therapeutic effects. This is likely due to the non-cleared APAP from the previous day having an accumulating effect on the total concentration. Consecutive doses (2nd, 3rd and 4th) on day 2 resulted in similar peak concentrations to the previous day for all three test individuals.

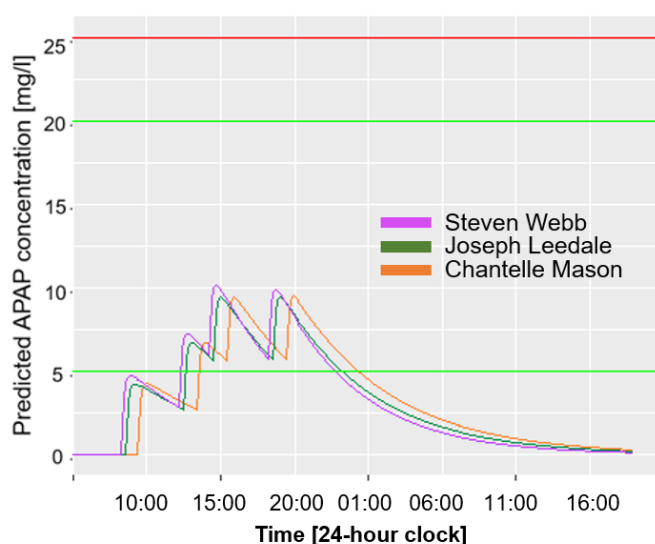


Figure 5-9. Predicted APAP concentration profiles for three test individuals following the recommended APAP dosing schedule – 1,000 mg administered at 7am,12pm,5pm,9pm. Purple - Concentration profile for test individual Steven Webb (44-year-old male weighing 92kg). Green - Concentration profile for test individual Joseph Leedale (32-year-old male weighing 86kg). Orange – Concentration profile for test individual Chantelle Mason (25-year-old female weighing 47kg). Green horizontal lines bound the therapeutic window (a peak APAP concentration within this window would be assumed to represent an efficacious case). The red horizontal line defines the toxic threshold (a peak APAP concentration reaching this level would be assumed to represent a toxic case).

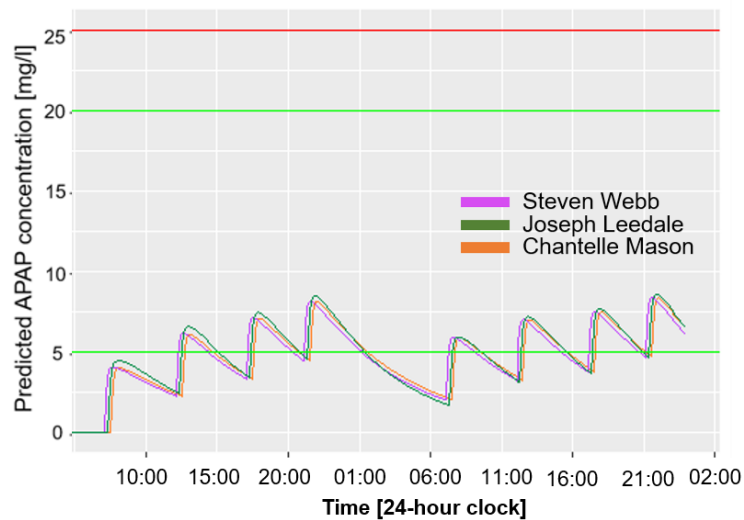


Figure 5-10: Predicted APAP concentration profiles for three test individuals following the recommended APAP dosing schedule, repeated for 2 days – 1,000 mg administered at 7am, 12pm, 5pm and 9pm. Purple - Concentration profile for test individual Steven Webb (44-year-old male weighing 92kg). Green - Concentration profile for test individual Joseph Leedale (32-year-old male weighing 86kg). Orange – Concentration profile for test individual Chantelle Mason (25-year-old female weighing 47kg). Green horizontal lines bound the therapeutic window (a peak APAP concentration within this window would be assumed to represent an efficacious case). The red horizontal line defines the toxic threshold (a peak APAP concentration reaching this level would be assumed to represent a toxic case).

An interesting result in both the one-day and two-day scenarios, was the time-period of drug holiday (time periods in which no therapeutic effects are caused by the drug). That is, with reference to the one-day dosing scenario for Chantelle (Figure 5-9), if the recommended dose was taken at 7am, 12pm, 5pm and 9pm, there are 8 hours throughout this period where no pain relief was achieved. Since no effective therapy was seen following the first dose (reflected by peak concentration resulting in a value lower than the therapeutic window requirement, 5-20 mg/l), this resulted in no therapeutic effects between 7am-12pm. Additionally, a drug holiday was seen between 2.30pm-5pm and finally between 8.30pm-9pm.

Differing demographic combination responses to varying doses of APAP

One important current clinical question is whether responses to the same APAP dose are different in younger, lighter individuals when compared to either older, average-weight or heavy-weight individuals. Once concentration profiles were confirmed to be as expected by our clinical collaborator (Dr. James Dear), for both the acute and multiple recommended dosing scenarios, we used the model to investigate differences in responses for the demographic groups of interest.

For this investigation a 15 years old was chosen to represent a “young” individual since “young” overdose cases are usually classed as those occurring in individuals aged between 15-24 [182]. 90-years-old was chosen to represent an “old” individual since this is beyond the general elderly age threshold of 75 [183]. Low weights were chosen to be below the average for the age group of interest [184] and heavy weights were chosen to be representative of obese/class 3 obese [185]. With this approach, a young, light weight individual was assumed to be aged 15, weighing 35 kg. A young, heavy individual was assumed to be aged 15, weighing 90 kg. This weight (90 kg) was chosen since this is the threshold for obesity for an individual of average height [185]. An old, light individual was assumed to be aged 90, weighing 40 kg. For elderly individuals, a weight beyond the class 3 obesity threshold was used [185]. An old, heavy individual was assumed to be aged 90, weighing 135 kg. For each of these demographic combination groups, an acute dose of 1,000 mg was provided, and simulated concentration-time profiles were compared (Figure 5-11).

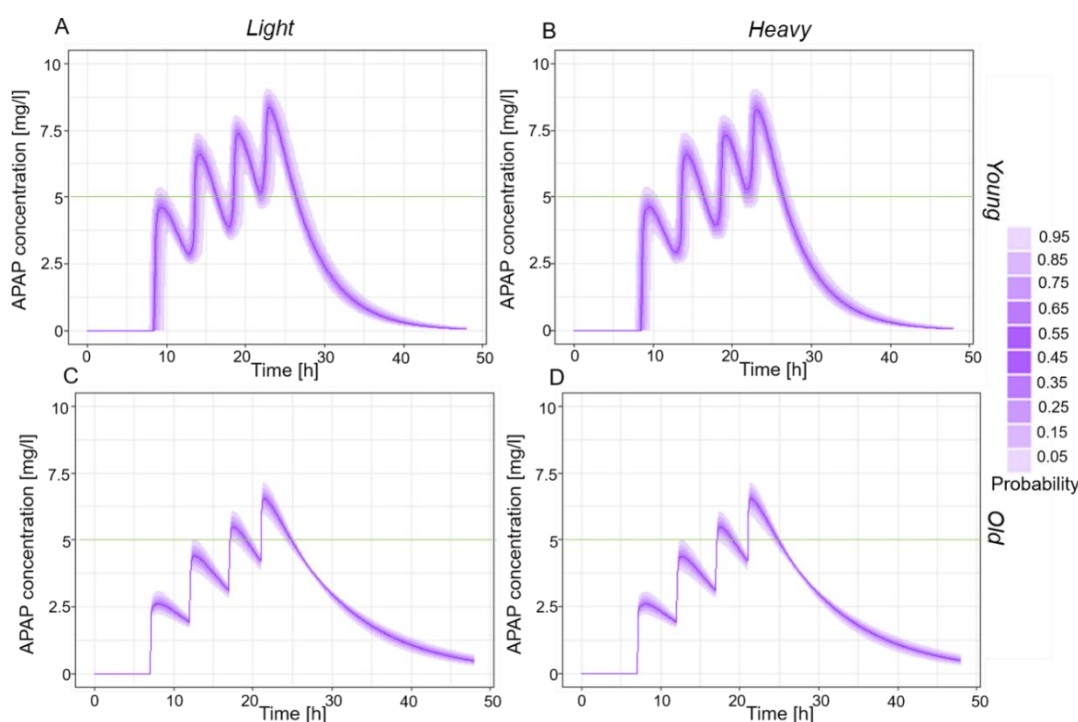


Figure 5-11: Predicted population APAP concentration profiles for various demographic groups following administration of a 1,000 mg APAP dose at 7am, 12pm, 5pm and 9pm. For each demographic group, there is a 95% chance that their concentration profile will be within the upper and lower bound (palest purple bounds). There is a 55% chance that their profile will be within the inner-most band (darkest purple). (A) predicted population APAP concentration profiles of young, low weight individuals. (B) predicted population APAP concentration profiles of young, heavy weight individuals. (C) predicted population APAP concentration profiles of old, light weight individuals. (D) predicted population APAP concentration profiles of old, heavy weight individuals.

For the recommended APAP dose, Figure 5-11 shows that all demographic groups were exposed to therapeutic effects, as expected. Interestingly though, individuals in the older groups (Figure 5-11(C-D)) were predicted to only see therapeutic effects following a third dose of APAP, whilst younger individuals (Figure 5-11(A-B)) were predicted to see therapeutic effects following the first dose. The group with the highest variance in their profile were the young, light group of individuals (Figure 5-11(A)), meaning that there was less certainty in predicting exact results for these individuals. This is a key result. For the younger group, if the weight was increased (Figure 5-11(B)), this seemed to improve confidence in the predictions, although any uncertainty in predictions was reduced further in the older population (Figure 5-11(C-D)). Generally, there was a greater variance in model predictions for lower weight individuals. As both age and weight increased, there was an improvement in the confidence of predictions.

The dose was then increased from the recommended 1,000 mg to 2,000 mg for an investigation into a higher dose in the same demographic groups. Resultant simulated concentration profiles are shown in Figure 5-12.

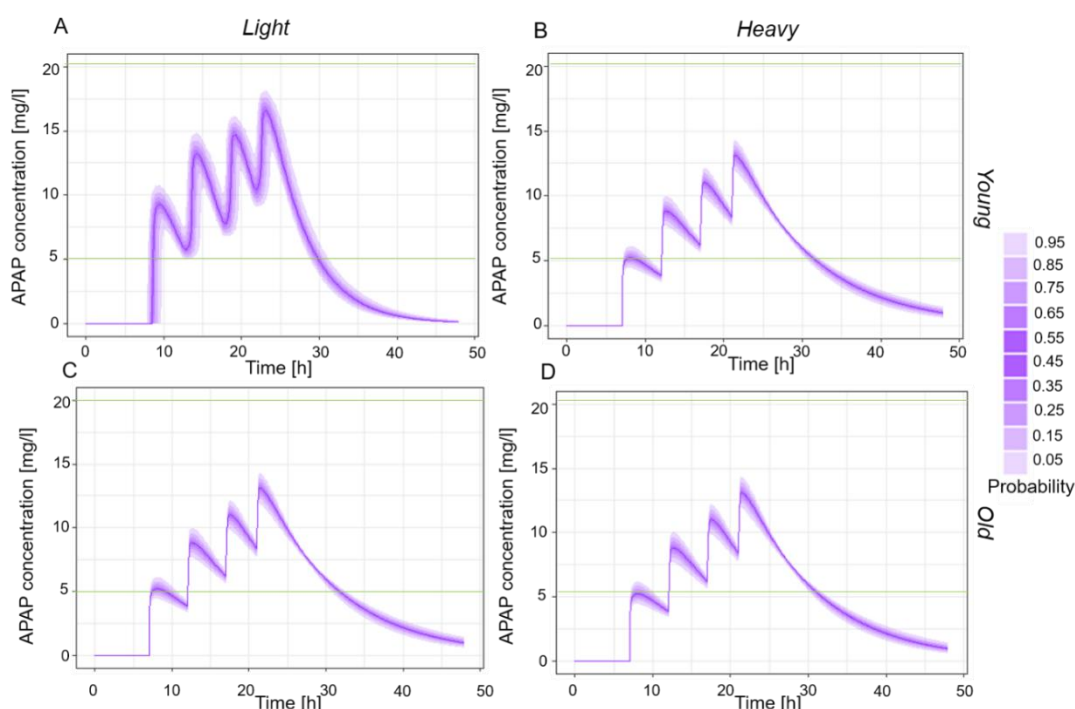


Figure 5-12: Predicted population APAP concentration profiles for various demographic groups following administration of a 2,000 mg APAP dose at 7am, 12pm, 5pm and 9pm. For each demographic group, there is a 95% chance that their concentration profile will be within the upper and lower bound (palest purple bounds). There is a 55% chance that their profile will be within the inner-most band (darkest purple). (A) predicted population APAP concentration profiles of young, low weight individuals. (B) predicted population APAP concentration profiles of young, heavy weight individuals. (C) predicted population APAP concentration profiles of old,

light weight individuals. (D) predicted population APAP concentration profiles of old, heavy weight individuals.

The main difference in the concentration profiles when increasing the dose, was the therapeutic effects being seen earlier for the older populations (Figure 5-12(C-D)). The younger populations (Figure 5-12(A-B)) were predicted to have therapeutic concentrations for the whole duration of the dosing schedule. For the older populations (Figure 5-12(C-D)), although there was a small drug holiday window between around 10h post-dose and 12h post-dose, the situation seemed to have improved for this higher dose, with therapeutic effects predicted following the first 2,000 mg dose, rather than the third 1,000 mg dose.

Differences in toxic responses for the demographic groups were investigated by simulating two highly toxic doses, 5,000 mg, and 10,000 mg.

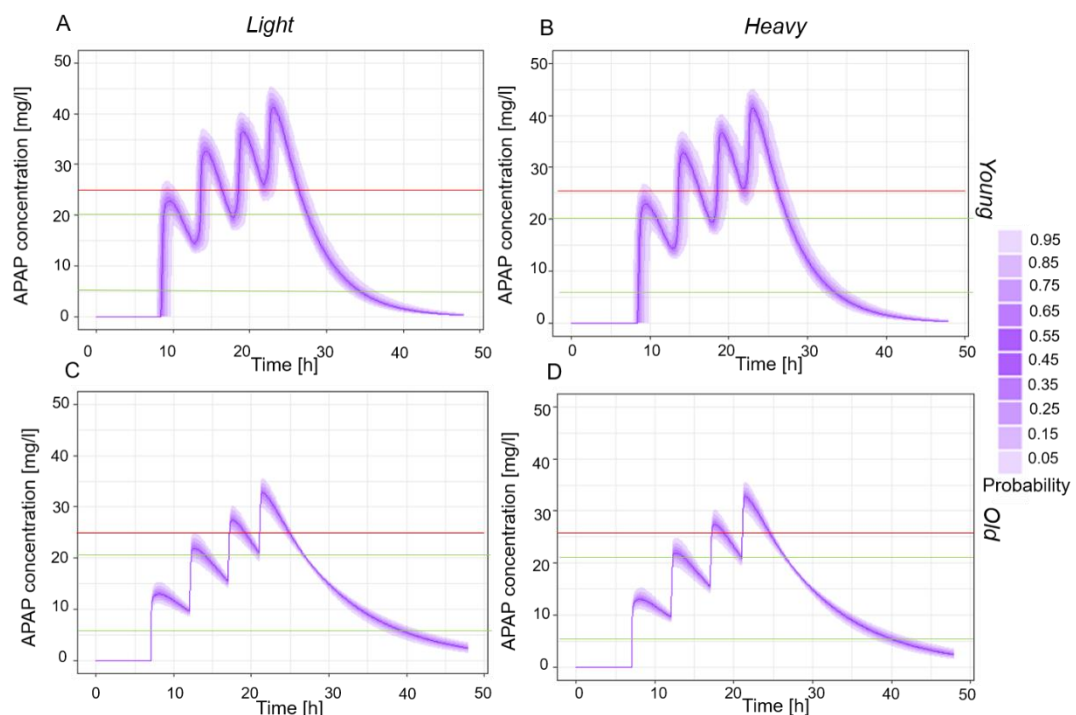


Figure 5-13: Predicted population APAP concentration profiles for various demographic groups following administration of a 5,000 mg APAP dose at 7am, 12pm, 5pm and 9pm. For each demographic group, there is a 95% chance that their concentration profile will be within the upper and lower bound (palest purple bounds). There is a 55% chance that their profile will be within the inner-most band (darkest purple). (A) predicted population APAP concentration profiles of young, low weight individuals. (B) predicted population APAP concentration profiles of young, heavy weight individuals. (C) predicted population APAP concentration profiles of old, light weight individuals. (D) predicted population APAP concentration profiles of old, heavy weight individuals.

The predicted concentration-time profiles of individuals following 4x 5,000 mg doses can be seen in Figure 5-13 and the predicted concentration over time profiles of individuals following 4 x10,000 mg doses can be seen in Figure 5-14.

For a toxic acute APAP dose of 5,000 mg, as shown in Figure 5-13(A-B), toxicity (represented by concentrations higher than 25 mg/l) was predicted to be apparent almost immediately in the younger individuals. For the older individuals (Figure 5-13(C-D)), following administration of one 5,000 mg dose, therapeutic concentrations were predicted, but following the third consecutive dose, toxic concentrations were predicted. As in previous results, the younger individuals (Figure 5-13(A-B)), particularly those who were also light-weight (Figure 5-13(A)), had the highest variance and least certainty in their predictions.

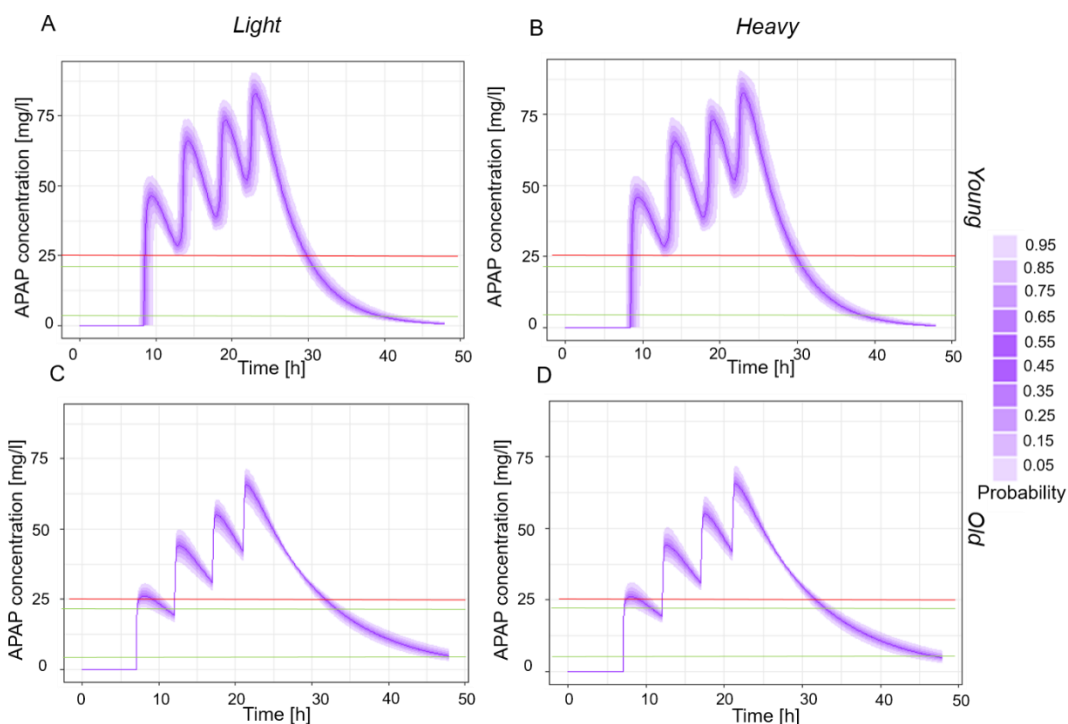


Figure 5-14: Predicted population APAP concentration profiles for various demographic groups following administration of a 10,000 mg APAP dose at 7am, 12pm, 5pm and 9pm. For each demographic group, there is a 95% chance that their concentration profile will be within the upper and lower bound (palest purple bounds). There is a 55% chance that their profile will be within the inner-most band (darkest purple). (A) predicted population APAP concentration profiles of young, low weight individuals. (B) predicted population APAP concentration profiles of young, heavy weight individuals. (C) predicted population APAP concentration profiles of old, light weight individuals. (D) predicted population APAP concentration profiles of old, heavy weight individuals.

For the highest dose of 10,000 mg (Figure 5-14), as expected, toxic concentrations were predicted in all demographic groups. There did appear to be a difference however, with regard to the extent of toxicity when comparing the younger (Figure 5-14(A-B)) and older individuals (Figure 5-14(C-D)). For the younger individuals, (both light and heavy in weight) the predicted concentrations far exceeded the toxic threshold. For the older population (both light and heavy in weight), the toxic threshold was only just exceeded after the first dose.

5.3.5.3. Investigating the effects of demographic information on APAP tolerance

Influence of age and weight

Our clinical collaborator (Dr. James Dear) requested an investigation into the effects of a patient's age and weight on APAP tolerance. This is of particular interest currently due to an increasing amount of toxicity cases being identified in young, low-weight individuals. These demographic factors were investigated in terms of their effect on maximum tolerated dose. For each age/weight combination for an individual, doses were increased until the toxic APAP C_{max} threshold was reached. This process was repeated for 10 different virtual humans per age/weight class, with the average and standard deviation recorded. The results can be seen in Figure 5-15. The figure represents results for simulated females. The approach was repeated for males, however the visual difference in results was negligible and therefore to avoid repetition, the figures for males are not provided.

On average, younger lighter people were predicted to have a lower tolerance to APAP (Figure 5-15(A)). This is indicated by the blue region, in the bottom left-hand corner of the plot, mainly accounting for individuals below age 40 and weight 80 kg (approximately). On average, individuals aged between 40-60 years and weighing between 80-100 kg had a slightly higher predicted tolerance to APAP dose compared to the younger, light-weight individuals. Individuals aged between 60-75 years and weighing between 100-120 kg had an even higher tolerance, whilst individuals above 75 years old and also weighing above 120 kg were predicted to, on average, have the highest tolerance to APAP dose.

The results from Figure 5-15(B) suggest that if a younger person has a higher weight, there is less variance in their predictions. This means that for these age/weight classes, we could predict their maximum tolerance with more certainty. This is indicated by the dark blue region in the bottom right-hand corner of the figure. This figure also shows that for any individual weighing less than 80 kg, the predicted maximum tolerable dose had a high potential to vary, with lower weight individuals over the age of 60 years having the highest variance in predictions. Visualising the data in terms of relative standard deviation (Figure 5-15(C)) however, we see that regardless of age, there was predicted to be a higher variance in tolerance for low weight individuals. Both Figure 5-15(B and C) suggest that for older people (above

approximately 65 years), an increased weight does not necessarily reduce the variability in maximum dose tolerance.

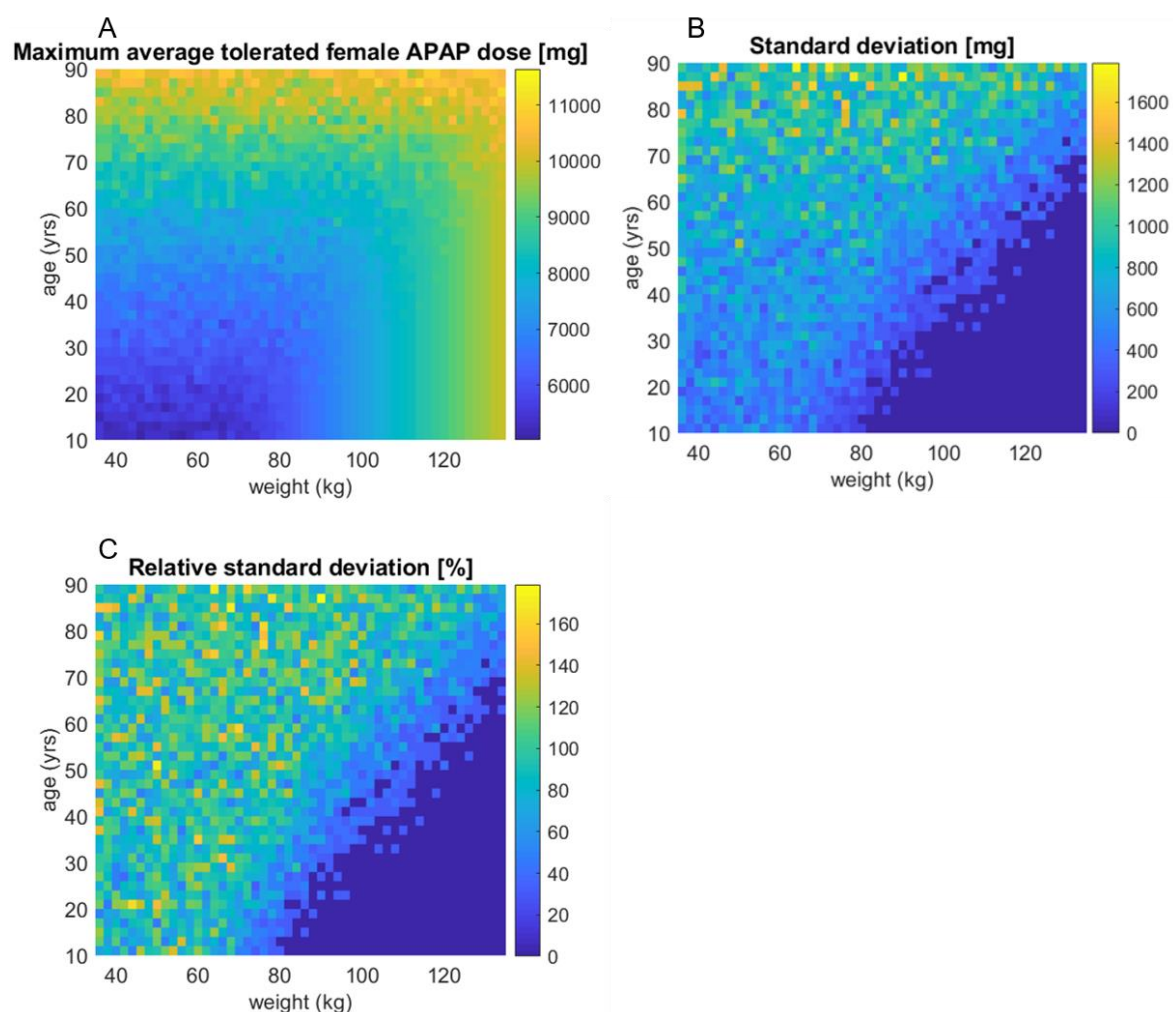


Figure 5-15: Maximum tolerated dose (the dose at which a toxic APAP C_{max} is reached) for females with various age and weight combinations. *Lowest values are indicated by a dark blue colour; as the value increases, the colour becomes green and eventually yellow at the highest values. (A) Average predicted maximum tolerated dose. (B) Standard deviation of predicted maximum tolerated doses. (C) Relative standard deviation of predicted maximum tolerated doses.*

Age differences

We investigated age-related differences in maximum tolerated APAP dose with particular focus on young individuals, regardless of weight. Whilst previously, we analyse APAP concentration (mg/l), upon arrival at the clinic, often an APAP dose (mg/kg) is the initial information provided to the clinician to aid in their decision of whether or not intervention is necessary. For this reason, our analysis is now conducted in terms of APAP dose (mg/kg) rather than APAP concentration (mg/l). We compared the predicted maximum tolerated doses with current thresholds used in the clinic which may deem a patient “unnecessary for intervention”, “potentially toxic” or “requiring transplant”. The aim of this was to determine whether or not “one-size fits all” thresholds are predictive, or whether further accountability for age should be considered. Results of which can be seen in Figure 5-16.

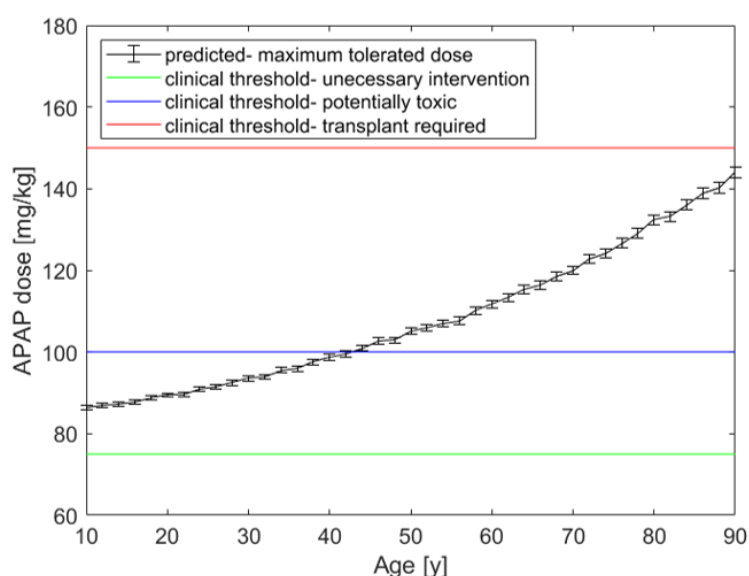


Figure 5-16: Predicted age effect on maximum tolerated APAP dose. *Black lines indicate the average maximum tolerated dose predicted by the model, with error bars representing standard error mean in predictions. The green horizontal line represents a currently used clinical threshold (75 mg/kg) for which any patient reporting an ingested dose below this would be deemed unnecessary for intervention. The blue horizontal line represents a currently used clinical threshold (100 mg/kg) for which any patient reporting an ingested dose above this would be deemed potentially toxic and therefore would require treatment. The red horizontal line represents a currently used clinical threshold (150 mg/kg) for which any patient reported to have ingested a dose above this would be thought to require a liver transplant.*

The results from Figure 5-16 show that although the currently used potentially toxic dose threshold may be suitable for individuals above 40 years of age, keeping the threshold this high may mean that toxicity could potentially be missed for individuals below this age. For the youngest individuals in our investigation, the average dose at which toxic concentrations were present was 85 mg/kg. Should a young individual ingest a dose of approximately 85 mg/kg, clinically, they would currently be assumed

to be a non-toxic case. However, our predictions suggest that toxic concentrations may in fact be present in this scenario. Conversely, if a 75-year-old individual reports an ingested dose of 100 mg/kg, currently, they would be treated with N'Acetylcysteine. However, our results suggest that individuals of this age could tolerate a maximum dose of 120 mg/kg, and therefore this individual could potentially be unnecessarily treated under the current recommendations. The threshold currently used to advise transplant requirement (reported dose above 150 mg/kg) lies above the average maximum tolerated dose predicted for all age ranges, and therefore our results suggest this threshold may be sufficient for this purpose. The threshold used to deem a patient unnecessary for intervention is 75 mg/kg. For all age ranges, the average maximum tolerated dose was predicted to be higher than this, and therefore this threshold may also be considered suitable.

5.4. Chapter discussion

The current clinical framework for treating paracetamol overdose is known to be sub optimal [10]. In previous chapters, specifically Chapters 2 and 3, we proposed a framework to improve the accuracy of predictions for initial dose and time since ingestion, and subsequently, the probability of liver injury. The mathematical approach we detailed is a proof-of-concept pre-clinical framework, which shows potential for translation to human and subsequent clinical application. Within this chapter, we provide a predictive framework applicable to humans. However, since biomarker data was unavailable at the time of investigation, within this chapter we detailed initial investigations on the PK aspect of the model only. The interesting results from these initial investigations provide additional support for further developing the model to include the pharmacodynamic aspect and strengthen the utility of the framework.

Large differences in inter-individual susceptibility to APAP overdose have been reported, but the specific reasons for this variability remain unknown [77]. To account for this variability in responses we employed a mixed effect modelling approach. This modelling approach takes into consideration parameters that are representative of the average population, but additionally considers how much an individual may potentially deviate from this average, providing a range to sample for individual *in silico* investigations. The parameters within our mixed effects model were optimised by fitting to data from approximately 3,600 APAP overdose patients. The optimised parameters (based on the clinical data) that varied greatest between

individuals were the APAP absorption rate (k_a) and the theoretical volume of distribution (V). There are many reasons which may justify variability in the absorption of a drug such as an individual's nutritional state, gastrointestinal health, or necessity for drug co-administration [186]. Since the theoretical volume of distribution is associated with the weight of an individual, it is no surprise that this varies greatly within the population. The PK parameters predicted to vary the greatest were therefore those as expected.

Simulating the model for three test individuals with different demographic information (gender, age and weight) allowed for the visualisation of differing concentration versus time profiles. There is persisting confusion regarding whether or not the weight of an individual should influence their oral APAP dose requirement. This has resulted in many independent clinics choosing to reduce the oral dose for low-weight individuals as a precautionary measure [157]. For just three test individuals, differences were identified in their predicted APAP concentration profiles, suggesting that profiles may vary greatly for a large amount of individuals with differing demographic information. It must be recollected however, that the model parameters used for simulation were optimised based on overdose data, and further optimisation with non-overdose data would be necessary to make therapeutic predictions with improved confidence.

With the potential effects of demographic information in mind, we investigated further the relationship between age and weight for large ranges which were not beyond the scope of the clinical dataset used for optimisation (10-90 years, 35-135 kg), with the aim of determining how this information affected the maximum tolerated dose. Results from this more extensive analysis suggest that, on average, younger lighter people are predicted to have a lower tolerance to APAP. It must be noted, however, that there was also predicted to be a high level of variability in the maximum tolerated dose for the young, low-weight group of individuals. For heavier, younger individuals, the variability in profiles was predicted to be less, suggesting that we can have more confidence that a younger person can tolerate less than an older person, specifically if they are of a higher weight.

Results suggest that regardless of age, there is much more variability in predictions for low-weight individuals, supporting the idea that some clinics take precautionary measures for these individuals. Since our findings thus far suggested that age was an important factor in differing responses to APAP, we used our model to compare predicted toxic thresholds for individuals ranging from 10-90 years with currently

used clinical thresholds. The key prediction from this analysis was that individuals below the age of 40 may not be investigated for potential toxicity under current clinical guidance (where 100 mg/kg is the considered toxic threshold). Results from our analysis suggest that individuals younger than 40 years old could have toxic APAP concentrations following a dose of 85 mg/kg. When simulating various four times daily doses and comparing how the concentration profiles differed for varying demographic groups, in every dosing scenario ([1,000; 2,000; 5,000; 10,000] mg) the group with the highest predicted variability in profiles were young, light weight individuals. Again, it must be noted that this model has been optimised against overdose data and therefore extrapolating predictions for therapeutic purposes may require additional optimisation against therapeutic data.

Clinically, there has been much discussion around this demographic group with regards to their differing toxicity when compared with other demographic groups, although there is currently little evidence to guide clinicians on how to appropriately treat this group. Our study highlights the necessity for further investigation into this demographic group. Better understanding should be sought in terms of why this specific demographic group has such variability in APAP response, which should in turn be used to try and better guide clinicians to amend treatment.

Chapter 6: Discussion

6.1. Conclusions

Paracetamol (APAP) overdose is a concerning global health issue [5]. Considered safe at therapeutic doses, this drug is also the most commonly used in overdose situations, accounting for 40% of all drug poisoning cases in the UK [3]. APAP overdose is the leading cause of acute liver failure (ALF) in the western world [4], with ALF being responsible for approximately 56,000 emergency room visits, 2,600 hospitalisations and 500 deaths per year in the USA [6]. Considering the severity level of this public health concern, the current clinical treatment framework for APAP overdose is criticised for being sub-optimal. The two main issues with the current framework are: decisions to administer NAC are currently heavily dependent on the patients' knowledge of time elapsed since overdose [10]; and, decisions to continue NAC treatment are heavily based on biomarkers ALT and INR, which are criticised for lacking sensitivity and specificity [11,17]. Further complexities arise as some individuals are more susceptible to toxicity from APAP overdose than others, particularly those with a weakened ability to synthesise or maintain sufficient glutathione (GSH) levels [10]. Such high-risk individuals may be of a particular age, suffer from pre-existing liver disease, concurrently use alcohol, or be malnourished, for example [5].

While the decision of whether or not to administer NAC is currently sub-optimal, so too is the ideal treatment regimen (i.e. NAC dose amount and administration time), resulting in recent trials of newly proposed treatment protocols [136]. As well as known risk-factors that cause individuals to have a higher susceptibility to toxicity following overdose, there are also variations in the physiology of humans which may cause increased susceptibility [77], and more work is needed to try and quantify these effects in order to improve the treatment framework. Whilst pre-clinical animal testing remains an important element of toxicity identification methodology, quantitative systems toxicology (QST) modelling is a recently well-employed tool in modern toxicity testing, providing useful predictions efficiently and reducing the dependency on animal tests [43]. Within this thesis, we describe a QST modelling approach with the aim of improving the current APAP clinical treatment framework.

Chapter 2 describes a pre-clinical proof-of concept framework representing APAP metabolism and corresponding toxicity biomarkers in mice. The biomarkers used within the investigation include the conventional (ALT), as well as the novel biomarkers (HMGB1, full and fragmented K18) that are thought to have improved

sensitivity over those currently used in the clinic [87]. We identified pharmacokinetic (PK) parameters for APAP in mice, which were previously undocumented within the literature. These values can now form a basis for other researchers within the field aiming to undertake similar pre-clinical APAP toxicity modelling investigations. All biomarkers were tested for their ability to predict potential liver injury against *in vivo* animal histology data via binomial logistic regression analysis. Although the biomarkers have been previously investigated individually [39], the biomarkers were tested in combinations within this study to better understand their relationship in the APAP toxicity setting. Although the current clinical combination of APAP and ALT turned out to be predictive for liver injury, HMGB1 alone was found to be the most significant predictor. This result formed the basis of our proposed liver injury identification framework.

Whilst Remien et al. [73] proposed a predictive liver injury framework based on a single measurement of biomarker concentrations, their study was focused upon conventional biomarkers e.g. alanine aminotransferase (ALT) and international normalised ratio (INR). Our study combined conventional and novel biomarkers with the aim of predicting the initial APAP dose, and how much time had elapsed since the dose was taken. The initial dose and time since administration predictions were found by combining PKPD modelling and simulation with statistical (visualisation and classification) techniques. The PKPD model was used to simulate a virtual population, with each individual receiving a random dose of APAP. Conventional and novel biomarker concentrations were extracted at a random time point to mimic the clinical situation. Measuring biomarker concentrations and combining with our proposed statistical framework allows the following to be predicted for each individual: initial APAP dose amount; time since overdose; and most importantly, a quantitative probability of liver injury (based solely on HMGB1 concentration).

Currently, the clinical treatment framework is heavily dependent on knowledge of time since administration [10], information which is often unknown or highly inaccurate. We have developed a framework that can predict the time elapsed since overdose, based on biological information rather than verbal. The decision to continue NAC treatment is currently based on monitoring levels of biomarkers such as international normalised ratio (INR) and alanine aminotransferase (ALT), both of which are criticised for their lack of sensitivity and specificity [11,17]. An additional issue is the requirement of multiple invasive biomarker measurements, which has the potential to lead to incomplete results. Our proposed framework predicts the

probability of liver injury based on HMGB1 concentration (a biomarker defined to have improved sensitivity over those currently used [39]). Additionally, our framework requires only one measurement of biomarkers, eradicating the requirement for multiple invasive measures.

Although QST models are being increasingly employed in the field of drug toxicity, skepticism around their utility still exists due to the necessity of underlying assumptions within the models. Therefore, any uncertainties that exist within the model should be explicitly defined to ensure the model may be used to its highest potential [46]. In Chapter 3, an identifiability analysis technique was defined with the aim to identify and quantify parameter uncertainty within the model. The technique was developed following a two-week award training placement at Freiburg Institute for Advanced Studies (FRIAS). In order to quantify any uncertainties in the model framework described in Chapter 2, we applied an identifiability analysis to the PKPD model which resulted in the identification of several model parameters that were unidentifiable. Structural unidentifiabilities (associated with the mechanistic structure of the model) were dominant in the HMGB1 and fragmented K18 components of the model. Practical unidentifiabilities (associated with a lack of experimental data used during parameter optimisation) were dominant in the GSH component of the model, with additional practical unidentifiabilities in the full and fragmented K18 components.

Chapter 3 describes the refinement of the model described in Chapter 2, adjusting the structure of the mechanistic framework and subsequently fitting the model to additional data to optimise the parameters. Since one of the complexities in APAP overdose treatment is that some individuals are more susceptible (“high-risk”) with regard to toxicity than others [10], we chose to extend our model framework to also predict toxicity in these “high-risk” individuals. We achieved this by modifying our pre-clinical PKPD model to represent both a fed and fasted state (fasted mice are known to have depleted GSH stores) and optimised parameter values by fitting to additional fed and fasted mouse data.

The extended model can be used to make toxicity predictions for either healthy, high-risk or mixed populations. Prediction accuracy for time since APAP administration and initial APAP dose were similar for the healthy and mixed populations, with a 3-4 h error for time since administration and a 65-75 mg/kg error for initial dose. The initial dose was much harder to predict in the high-risk population (184.1 mg/kg error). This result is unsurprising since a much larger range of doses

may have a toxic effect if the liver is already impaired. Time since administration, however, could be predicted in the high-risk group more accurately than in a healthy/mixed population (2.118 h error). It has already been found that amending treatment thresholds to account for high-risk individuals can better protect those with greater liver injury susceptibility [5].

Results from this pre-clinical study further endorse the idea that there are likely very different outcomes with respect to liver injury potential for healthy and high-risk individuals. This was concluded from the visualisation of *in silico* individuals from both “healthy” and “high-risk” populations, with each observation being discriminated by their probability of liver injury. For the healthy population, most of the observations were predicted to have less than around 35% chance of DILI progression. In stark contrast, however, most of the observations in the high-risk population were predicted to have almost 100% chance of liver injury progression. More informed decisions could therefore be made regarding optimal treatment if clinicians can identify those who are more susceptible to overdose. This would significantly improve patient outcomes while reducing the cost and burden of unnecessary antidote treatment.

In order to create a framework that has the potential to improve current APAP toxicity treatment in the clinic, it is important to include the antidote (N’acetylcysteine, NAC) and its effects on toxicity potential. In Chapter 4, we included *in vitro* experimental data showing the effects of APAP and NAC on cell functionality obtained through a laboratory experimental training placement at the University of Liverpool. The results of a multiple linear regression analysis concluded that, if biomarker measurements are taken throughout treatment, ALT, full and fragmented K18 could significantly predict cell viability. This insight was made possible by modelling full and fragmented K18 as separate necrotic and apoptotic versions of the same biomarker, suggesting its critical role in quantifying cell functionality.

There has been much discussion around the uncertainty of the optimal combination of NAC dose amount and time of administration. In Chapter 4, we compared two NAC treatment regimens [24,136], the first, being the regimen currently used in the clinic, and the second being a proposed regimen that administers NAC over a shorter time period (SNAP) regimen. We assessed the regimens’ effects on the amount of APAP exposure allowed before toxicity was predicted. We used the thresholds defined by our framework: probable liver injury; and severe loss of cell

functionality. We found that when using probable liver injury as a marker of toxicity, the current NAC regimen was predicted to have better protective effects than the proposed SNAP regimen (protecting against APAP doses of 8,000 mg/kg compared to the SNAP regimen which protects against 3,500 mg/kg). When using loss of cell functionality as a measure, however, both regimens were predicted to provide similar protective effects. If cell functionality is used as a toxicity measure, our framework predicted toxicity occurring for any APAP dose beyond 400 mg/kg. Currently, the clinical expected fatal toxicity dose is thought to occur at APAP doses beyond 250 mg/kg [8]. Our preliminary results suggest that this clinical threshold may have the potential to be increased, possibly leading to huge savings on unnecessary interventions. However, further investigations would be required before such predictions could be confirmed.

In Chapters 2 to 4, we have provided a proof-of-concept framework that is based on pre-clinical *in vivo/in vitro* scenarios. Chapter 5 describes a population-pharmacokinetic (Pop-PK) model that can be applied to the human clinical case. The framework has parameters that were optimised by fitting the Pop-PK model to a clinical dataset of 3,616 patients (provided by our clinical collaborator, Dr. James Dear). The APAP overdose patients were from 3 UK hospitals – Newcastle, Edinburgh and London. For each patient, we had the following information: Covariate (demographic) information (age, sex, weight); APAP dose taken (mg); time since APAP dose (h); and APAP measurement upon hospital admission (mg/l). Once confident that the simulations from our Pop-PK framework were representative of the clinical data, we used the modelling framework to investigate some concerns which are currently of clinical interest.

Firstly, large differences between inter-individual susceptibility to APAP overdose have been reported. However, specific reasons for this variability remain unknown [77]. We used modelling and simulation approaches to obtain a better quantification of this variability. Namely, the parameters that varied greatest between individuals were the APAP absorption rate, k_a , and the theoretical volume of distribution, V . There are many reasons which may explain differences in the absorption of a drug, such as an individual's nutritional state, gastrointestinal health or polypharmacy [186].

Secondly, within the clinic, there has been much discussion around a specific demographic group (young, lower weight individuals) with regards to their differing toxicity when compared to other demographic groups. At present, there is little

quantitative evidence to guide clinicians on how to appropriately treat this specific group. Through various simulations of our model, we were able to compare APAP concentration profiles for individuals in different demographic classes. We compared predicted population profiles from four different classes of individuals; young, light weight; young, heavy weight; old, light weight; and old, heavy weight. A consistent result, throughout testing various APAP doses, was that the concentration profiles for the young, light weight individuals were predicted to have much greater variance compared to the other demographic groups, suggesting less certainty in their predicted APAP profiles. We also simulated over 2,000 different age and weight combinations with the aim of investigating any differences in the APAP dose threshold predicted to induce toxicity. The results from this analysis clearly suggest that on average, younger, light weight individuals may have a lower tolerance to APAP. Something important to note, however, is that there was a high level of variability in the predicted maximum dose for this demographic group. As the weight of the young individuals increased, the relative variability in their predictions was much lower. This suggests that we can have more confidence that young individuals may have a lower APAP tolerance than an older person, if they are of a higher weight.

One of the key highlights from our study suggests that under the current clinical guidelines, some patients who are deemed unnecessary for investigation, may in fact require treatment for APAP toxicity. An APAP dose of 100 mg/kg is thought to be the toxic APAP threshold at present. Results from our analysis suggest that individuals below the age of 40 years old could have toxic APAP concentrations following a dose of 85 mg/kg. Missing potential toxicity cases upon arrival at the hospital can have several detrimental downstream effects. For example, a nationwide study of over 12,000 subjects conducted by Huang et al. found that APAP poisoning is associated with increased long-term mortality and early referral for intensive after-care is essential [187].

In summary, this thesis aims to show the utility of mathematical modelling in toxicity predictions, with the particular objective of making improvements to the clinical APAP overdose treatment framework. We provide a proof-of-concept framework that, from a single measurement of conventional and novel biomarker concentrations, predictions about the amount of APAP taken and how long it has been since the overdose took place can be made. We also have the capability of predicting the probability of liver injury for each individual, and the severity of the

loss of cell functionality. We investigated the effects of NAC treatment and propose potential adjustments to the current framework which may improve the clinical situation. Finally, we provide a fully-optimised Pop-PK clinical model that has provided a better understanding of which physiological parameters could have the greatest influence on inter-individual susceptibility to APAP overdose. Whilst our clinical framework in its Pop-PK form has provided insight to the APAP overdose problem, once extended to include the biomarker dynamics (such as those investigated in Chapters 2-4), our hope is that our framework would have the utility to make toxicity predictions in the human case based on biomarkers that are more sensitive and specific than those currently used.

There have been some mathematical approaches to improving the APAP overdose problem to date. For better understanding of APAP metabolism, mechanistic models often take the form of systems of ordinary differential equations (ODEs). These kind of mathematical models are useful for representing and making predictions of biological processes over time. Examples of such studies are defined by the DILI-sim-initiative and Geenen et al. [69,83]. We have taken a similar ODE modelling approach to represent APAP metabolism and the corresponding toxicity biomarkers. However, considering the limitation of data dependence, we define an identifiability analysis technique, which following application to the model defined in Chapter 2, enabled us to determine areas of uncertainty within our model structure that required refinement/additional data during optimisation. We were subsequently able to define the uncertainty within our predictions, and improve confidence in our framework.

The key novelty of this study is the combination of mathematical and statistical modelling approaches. Simulations of the mechanistic models are combined with statistical visualisation and classification techniques to provide key mechanistic information that can be utilised by experimentalists, whilst additionally providing results from statistical analysis that can be easily interpreted and used by a clinician.

6.2. Future work

While the results of this research are promising, with additional time and resources, the utility of the model predictions could clearly be improved. With regard to the proof-of-concept framework defined in Chapters 2 and 3, although the results from the T-SNE method for visualisation showed clear separation, particularly with regards to the probability of liver injury, there was a slight overlap in the time-since-

administration and dose plots. This result supports the possibility of defining further classes through unsupervised methodologies in future investigations. Also, the classification techniques used provided incredibly high accuracy levels considering the nature of the problem. However, a further investigation of interest could be the rate of misclassification between the classes with regard to critical errors at the edges of the variable ranges.

Better understanding of the relationship between probability of liver injury (based on HMGB1 concentration) and loss of cell functionality is required. This is supported by our APAP tolerability results in Chapter 4 varying greatly depending on the toxicity measure used. Results from our investigation could be used to guide *in vitro* experiments to improve this understanding. For example, potential experiments could focus on monitoring multidrug resistance proteins (MRPs) over time following different APAP/NAC doses; these transporters are reported to transport GSH (which will affect HMGB1 concentration) and also bind to ATP (which will affect hepatocyte viability) [153]. Further experimental repeats of our cell viability study would also be necessary to ensure confidence in the accuracy of the optimised parameters used in our cell functionality model, and therefore the cell functionality metric.

For QST models to be useful, relevant human biomarkers are essential [188]. During the time this research was carried out, the novel biomarkers used in our pre-clinical framework were unavailable within the clinical setting. Since this time, human biomarker data has been made available for HMGB1, full and fragmented K18 [136]. The next stage in this research will be to extend the currently developed APAP Pop-PK model (Chapter 5) to include the pharmacodynamic components of the pre-clinical novel biomarkers. A crucial investigation will be determining whether the predictions made from our framework defined in Chapters 2 and 3 are translatable to the clinical setting. Results from our pre-clinical analysis suggest that biomarkers ALT and full K18 are particularly sensitive to perturbations in parameters, and therefore sufficiently dense data will be required when translating the work to the clinical context. Dr. James Dear has expressed his interest in continuing our collaborative research. He and his team at the Queens Medical Research Institute, Edinburgh, predict that if translated to the clinic, our framework has the potential to save approximately 20,000 hospital bed days per year in the UK. This equates to an approximate saving of £8 million. Dr. Dear suggests that if translatability potential is deemed successful, their team could employ the framework within their clinic within a matter of months.

References

1. Chiew A, Fountain J, Graudins A, Geoffrey KI, Reith D, Buckley NA. Guidelines for the management of paracetamol poisoning in Australia and New Zealand. *Med J Aust*. 2015;
2. Gunnell D, Murray V, Hawton K. Use of paracetamol (acetaminophen) for suicide and nonfatal poisoning: worldwide patterns of use and misuse. *Suicide Life Threat*. 2000;30(0363–0234):313–26.
3. Doctors.net. Paracetamol (Acetaminophen) Poisoning Evidence Based Review [Internet]. 2017 [cited 2017 May 30]. Available from: https://www.doctors.net.uk/_datastore/ecme/mod1145/paracetamol_poisoning_EBR_ch_v01.pdf
4. Larson AM, Polson J, Fontana RJ, Davern TJ, Lalani E, Hynan LS, et al. Acetaminophen-induced acute liver failure: Results of a United States multicenter, prospective study. *Hepatology*. 2005;42(6):1364–1372.
5. Yoon E, Babar A, Choudhary M, Kutner M, Pysopoulos N. Acetaminophen-Induced Hepatotoxicity: a Comprehensive Update. *J Clin Transl Hepatol*. 2016;4(2):131–42.
6. Lee WM. Acetaminophen and the U.S. acute liver failure study group: Lowering the risks of hepatic failure. *Hepatology*. 2004;40(1):6–9.
7. ONS. Deaths related to drug poisoning in England and Wales: 2016 registrations [Internet]. statistical bulletin. 2016 [cited 2018 Sep 11]. Available from: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsrelatedtodrugpoisoninginenglandandwales/2016>

6registrations

8. Buckley N, Eddleston M. Paracetamol (acetaminophen) poisoning. Clin Evid (Online). 2007;(March):1–14.
9. Rumack B, Peterson R, Koch G, Amara I. Acetaminophen overdose. 662 cases with evaluation of oral acetylcysteine treatment. Arch Int Med. 1981;141(3):380–5.
10. Ferner R, Dear WJ, Bateman N. Management of paracetamol poisoning. BMJ. 2011;342:1–19.
11. Merlot AM, Kalinowski DS, Kovacevic Z, Jansson PJ. Making a case for albumin - a highly promising drug delivery system. Med Chem (Los Angeles). 2015;7:553–6.
12. Williams M. Alanine aminotransferase (serum, plasma). Assoc Clin Biochem. 2012;3–7.
13. McGill MR. The past and present of serum aminotransferases and the future of liver injury biomarkers. EXCLI J. 2016;15:817–28.
14. Brace LD. Current Status of the International Normalized Ratio. Lab Med. 2003;32(7):390–2.
15. Kopec AK, Luyendyk JP. Coagulation in liver toxicity and disease: Role of hepatocyte tissue factor. Thromb Res. 2014;133(0 1):1–7.
16. BritishLiverTrust. Liver Function Tests [Internet]. 2017 [cited 2017 May 5]. Available from: <https://www.britishlivertrust.org.uk/liver-information/tests-and-screening/liver-function-tests/>
17. Clarke J, Dear J, Antoine DJ. Recent advances in biomarkers and therapeutic interventions for hepatic drug safety - false dawn or new horizon? Expert Opin

- Drug Saf. 2016;
18. EBMedicine. Make Calculated Decisions. 2018;
 19. Achterbergh R, Lammers LA, Kuijsten L, Klümpen HJ, Mathôt RAA, Romijn JA. Effects of nutritional status on acetaminophen measurement and exposure. *Clin Toxicol.* 2019;57(1):42–9.
 20. GlaxoSmithKline. Guidelines for the management of paracetamol overdose [Internet]. 2007 [cited 2017 Apr 10]. Available from: <http://www.asem.org.au/document.php/njxudmy/Paracetamol+Overdose+Treatment+Nomogram.pdf>
 21. NHS. Statistics on alcohol, England 2017 [Internet]. 2017 [cited 2018 Sep 12]. Available from: <http://webarchive.nationalarchives.gov.uk/20180328130416/http://digital.nhs.uk/catalogue/PUB23940>
 22. Beat. Eating disorder statistics [Internet]. 2018 [cited 2018 Sep 11]. Available from: <https://www.beateatingdisorders.org.uk/media-centre/eating-disorder-statistics>
 23. Algren A. Review of N-Acetylcysteine for The Treatment of Acetaminophen (Paracetamol) Toxicity in Pediatrics. Expert Committee on The Selection and Use of Essential Medicines; 2008.
 24. Bateman DN, Carroll R, Pettie J, Yamamoto T, Elamin MEMO, Peart L, et al. Effect of the UK's revised paracetamol poisoning management guidelines on admissions, adverse reactions and costs of treatment. *Br J Clin Pharmacol.* 2014;78(3):610–8.
 25. Hayward KL, Powell EE, Irvine KM, Martin JH. Can paracetamol (acetaminophen) be administered to patients with liver impairment? *Br J Clin*

- Pharmacol. 2016;81(2):210–22.
26. Smith HS, Royal MA. Acetaminophen. First Edit. Current Therapy in Pain. Elsevier Inc.; 2009. 429–436 p.
 27. Hundt M, John S. Bile Secretion. In: Physiology [Internet]. 2019. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK470209/>
 28. Mazaleuskaya LL, Sangkuhl K, Thorn CF, Fitzgerald GA, Altman RB, Klein TE. PharmGKB summary: Pathways of acetaminophen metabolism at the therapeutic versus toxic doses. Pharmacogenet Genomics. 2015;25(8):416–26.
 29. Alwood AJ. Acetaminophen. In: J Toxicol Clin Toxicol. 1999. p. 6–9.
 30. Antoine DJ, Dear JW. How to treat paracetamol overdose and when to do it. Expert Rev Clin Pharmacol. 2016;9(5):633–5.
 31. Malhi H, Gores GJ, Lemasters JJ. Apoptosis and necrosis in the liver: A tale of two deaths? Hepatology. 2006;43(2 SUPPL. 1):31–44.
 32. Manning F, Zuzel K. Comparison of types of cell death: Apoptosis and necrosis. J Biol Educ. 2003;37(3):141–5.
 33. Schulza-Osthoff K, Bantel H. Necrosis Versus Apoptosis in Acetaminophen-Induced Hepatotoxicity. Hepatology. 2011;1070–1.
 34. Malhi H, Gores GJ. Cellular and Molecular Mechanisms of Liver Injury. Gastroenterology. 2008;134(6):1641–54.
 35. Antoine DJ, Williams DP, Kipar A, Laverty H, Park BK. Diet Restriction Inhibits Apoptosis and HMGB1 Oxidation and Promotes Inflammatory Cell Recruitment during Acetaminophen Hepatotoxicity. 2010;16(December):479–90.

36. Nunnari JRF and J. Mitochondrial form and function Jonathan. *Nano*. 2008;6(9):2166–71.
37. Robles-Díaz M, Medina-Caliz I, Stephens C, Andrade RJ, Lucena MI. Biomarkers in DILI: One more step forward. *Front Pharmacol*. 2016;7(AUG):1–7.
38. Schomaker S, Warner R, Bock J, Johnson K, Potter D, Winkle J Van, et al. Assessment of Emerging Biomarkers of Liver Injury in Human Subjects. 2013;132(2):276–83.
39. Antoine DJ, Williams DP, Jenkins AK, Regan SL, Sathish JG, Kitteringham NR, et al. High-mobility group box-1 protein and keratin-18, circulating serum proteins informative of acetaminophen-induced necrosis and apoptosis in vivo. *Toxicol Sci*. 2009;112(2):521–31.
40. David P. Bartel. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*. 2004;116(2):281–97.
41. Raies AB, Bajic VB. In silico toxicology: computational methods for the prediction of chemical toxicity. *Wiley Interdiscip Rev Comput Mol Sci*. 2016;6:147–72.
42. Ballentine C. Sulfanilamide Disaster. *FDA consumer magazine*. 1981;(June).
43. Bloomingdale P, Housand C, Apgar JF, Millard BL, Mager DE, Burke JM, et al. Quantitative systems toxicology. *Curr Opin Toxicol*. 2017;4:79–87.
44. Bai JP., Fontana RJ, Price ND, Sangar V. Systems pharmacology modeling : an approach to improving drug safety. *Biopharm Drug Dispos*. 2014;35:1–14.
45. Wagner BK, Kitami T, Gilbert TJ, Peck D, Ramanathan A, Schreiber SL, et al. Large-scale chemical dissection of mitochondrial function. *Nat Biotechnol*.

- 2008;26(3):343–51.
46. Vanlier J. *Uncertainty Analysis in Systems Biology*. Eindhoven: Technische Universiteit Eindhoven; 2014.
 47. Tylutki Z, Polak S, Wi B. Top-down , Bottom-up and Middle-out Strategies for Drug Cardiac Safety Assessment via Modeling and Simulations. 2016;171–7.
 48. Cherkasov A, Muratov E, Fourches D, Varnek A, Baskin I, Cronin M, et al. QSAR Modeling: Where have you been? Where are you going to? *Med Chem*. 2014;57(12):4977–5010.
 49. Patlewicz G, Dimitrov SD, Low LK, Kern PS, Dimitrova GD, Comber MIH, et al. TIMES-SS-A promising tool for the assessment of skin sensitization hazard. A characterization with respect to the OECD validation principles for (Q)SARs and an external evaluation for predictivity. *Regul Toxicol Pharmacol*. 2007;48(2):225–39.
 50. Benigni R, Bossa C. Structure alerts for carcinogenicity, and the Salmonella assay system: A novel insight through the chemical relational databases technology. *Mutat Res - Rev Mutat Res*. 2008;659(3):248–61.
 51. Cheng A, Dixon SL. In silico models for the prediction of dose-dependent human hepatotoxicity. *J Comput Aided Mol Des*. 2003;17(12):811–23.
 52. Li X, Chen Y, Song X, Zhang Y, Li H, Zhao Y. The development and application of: In silico models for drug induced liver injury. *RSC Adv*. 2018;8(15):8101–11.
 53. Liebert MA. Predicting CNS Permeability of Drug Molecules : 2002;9(6):849–64.
 54. Dearden JC, Cronin MTD, Kaiser KLE. How not to develop a quantitative

- structure-activity or structure-property relationship (QSAR/QSPR). SAR QSAR Environ Res. 2009;20(3–4):241–66.
55. Hedaya M. Basic Pharmacokinetics. 2012.
 56. Rosenbaum S. Basic pharmacokinetics and pharmacodynamics. Wiley; 2011.
 57. Gopu D, Gomathi P. Pharmacokinetic/Pharmacodynamic (PK/PD) modeling: An investigational tool for drug development. Int J Pharm Pharm Sci. 2012;4(3):30–7.
 58. Charles B. Population pharmacokinetics: An overview. Aust Prescr. 2014;37(6):210–3.
 59. Watkins P. The DILI-sim Initiative: Insights into Hepatotoxicity Mechanisms and Biomarker Interpretation. Clin Transl Sci. 2019;122–9.
 60. Shoda LKM, Battista C, Siler SQ, Pisetsky DS, Watkins PB, Howell BA. Mechanistic Modelling of Drug-Induced Liver Injury : Investigating the Role of Innate Immune Responses. 2017;
 61. Howell BA, Siler SQ, Barton HA, Joshi EM, Cabal A, Eichenbaum G, et al. Development of quantitative systems pharmacology and toxicology models within consortia: experiences and lessons learned through DILIsym development. Drug Discov Today Dis Model. 2016;22:5–13.
 62. Woodhead JL, Paech F, Maurer M, Engelhardt M, Schmitt-Hoffmann AH, Spickermann J, et al. Prediction of Safety Margin and Optimization of Dosing Protocol for a Novel Antibiotic using Quantitative Systems Pharmacology Modeling. Clin Transl Sci. 2018;11(5):498–505.
 63. Longo DM, Yang Y, Watkins PB, Howell BA, Siler SQ. Elucidating differences in the hepatotoxic potential of tolcapone and entacapone with DILIsym®, a

- mechanistic model of drug-induced liver injury. *CPT Pharmacometrics Syst Pharmacol*. 2016;5(1):31–9.
64. Woodhead JL, Howell BA, Yang Y, Harrill AH, Clewell HJ, Andersen ME, et al. An Analysis of N-Acetylcysteine Treatment for Acetaminophen Overdose Using a Systems Model of Drug-Induced Liver Injury. *J Pharmacol Exp Ther*. 2012;342(2):529–40.
65. Will Y, McDuffie EJ, Olaharski AJ, Jeffy BD. Systems pharmacology and DILI. In: *Drug Discovery Toxicology: From Target Assessment to Translational Biomarkers*. 2016. p. 118–20.
66. Reith D, Medlicott NJ, Kumara De Silva R, Yang L, Hickling J, Zacharias M. Simultaneous modelling of the michaelis-menten kinetics of paracetamol sulphation and glucuronidation. *Clin Exp Pharmacol Physiol*. 2009;36(1):35–42.
67. Diaz Ochoa JG, Bucher J, Péry ARR, Zaldivar Comenges JM, Niklas J, Mauch K. A multi-scale modeling framework for individualized, spatiotemporal prediction of drug effects and toxicological risk. *Front Pharmacol*. 2013;3(204):1–11.
68. Krauss M, Schaller S, Borchers S, Findeisen R, Lippert J, Kuepfer L. Integrating Cellular Metabolism into a Multiscale Whole-Body Model. *PLoS Comput Biol*. 2012;8(10).
69. Geenen S, Yates JWT, Kenna JG, Bois FY, Wilson ID, Westerhoff H V. Multiscale modelling approach combining a kinetic model of glutathione metabolism with PBPK models of paracetamol and the potential glutathione-depletion biomarkers ophthalmic acid and 5-oxoproline in humans and rats. *Integr Biol (United Kingdom)*. 2013;5(6):877–88.

70. Geenen S, Du Preez FB, Snoep JL, Foster AJ, Sarda S, Kenna JG, et al. Glutathione metabolism modeling: A mechanism for liver drug-robustness and a new biomarker strategy. *Biochim Biophys Acta - Gen Subj*. 2013;1830(10):4943–59.
71. Stahl SH, Yates JW, Nicholls AW, Kenna JG, Coen M, Ortega F, et al. Systems toxicology: Modelling biomarkers of glutathione homeostasis and paracetamol metabolism. *Drug Discov Today Technol* [Internet]. 2015;15:9–14. Available from: <http://dx.doi.org/10.1016/j.ddtec.2015.06.003>
72. Leclerc E, Hamon J, Claude I, Jellali R, Naudot M, Bois F. Investigation of acetaminophen toxicity in HepG2/C3a microscale cultures using a system biology model of glutathione depletion. *Cell Biol Toxicol*. 2015;31(3):173–85.
73. Remien CH, Adler FR, Waddoups L, Box TD, Sussman NL. Mathematical modeling of liver injury and dysfunction after acetaminophen overdose: Early discrimination between survival and death. *Hepatology*. 2012;56(2):727–34.
74. Remien CH, Sussman NL, Adler FR. Mathematical modelling of chronic acetaminophen metabolism and liver injury. *Math Med Biol*. 2014;31(3):302–17.
75. Ben-Shachar R, Chen Y, Luo S, Hartman C, Reed M, Nijhout HF. The biochemistry of acetaminophen hepatotoxicity and rescue: a mathematical model. *Theor Biol Med Model*. 2012;9:55.
76. Reddyhoff D, Ward J, Williams D, Regan S, Webb S. Timescale analysis of a mathematical model of acetaminophen metabolism and toxicity. *J Theor Biol*. 2015;386:132–46.
77. Ruiz-Aracama A, Claessen SM, Peijnenburg AACM, van Herwijnen MHM, Lommen A, Kleinjans JCS, et al. Interindividual variation in gene expression

- responses and metabolite formation in acetaminophen-exposed primary human hepatocytes. *Arch Toxicol.* 2015;90(5):1103–15.
78. Zhu AZX. Quantitative translational modeling to facilitate preclinical to clinical efficacy & toxicity translation in oncology. 2018;4.
79. Péry ARR, Brochot C, Zeman FA, Mombelli E, Desmots S, Pavan M, et al. Prediction of dose-hepatotoxic response in humans based on toxicokinetic/toxicodynamic modeling with or without in vivo data: A case study with acetaminophen. *Toxicol Lett.* 2013;220(1):26–34.
80. Zurlinden TJ, Heard K, Reifeld B. A novel approach for estimating ingested dose associated with paracetamol overdose. *Br J Clin Pharmacol.* 2016;81(4):634–45.
81. Navid A, Ng DM, Stewart BJ, Wong SE, Lightstone FC. Quantitative In Silico analysis of transient metabolism of acetaminophen and associated causes of hepatotoxicity in humans. *Silico Pharmacol.* 2013;1(1):14.
82. Kreutz C, Raue A, Kaschek D, Timmer J. Profile likelihood in systems biology. 2013;280:2564–71.
83. DILI-simInitiative. DILIsym [Internet]. 2018. Available from: <https://www.simulations-plus.com/software/dilisym/>
84. Juurlink D. Drug-Induced Hepatotoxicity. *N Engl J Med.* 2003;349:1974–6.
85. LabTestsOnline. Prothrombin Time and International Normalized Ratio (PT/INR) [Internet]. 2019 [cited 2019 Apr 14]. Available from: <https://labtestsonline.org/tests/prothrombin-time-and-international-normalized-ratio-ptinr>
86. Giannini EG, Testa R, Savarino V. Liver enzyme alteration: A guide for

- clinicians. *Cmaj*. 2005;172(3):367–79.
87. Antoine DJ, Dear JW, Lewis PS, Platt V, Coyle J, Masson M, et al. Mechanistic Biomarkers Provide Early and Sensitive Detection of Acetaminophen-Induced Acute Liver Injury at First Presentation to Hospital. *Hepatology*. 2013;58(2):777–87.
 88. Beger RD, Bhattacharyya S, Yang X, Gill PS, Schnackenberg LK, Sun J, et al. Translational biomarkers of acetaminophen-induced acute liver injury. *Arch Toxicol*. 2015;89(9):1497–522.
 89. NC3Rs. The 3Rs [Internet]. 2017 [cited 2017 Apr 28]. Available from: <https://www.nc3rs.org.uk/the-3rs>
 90. Jaeschke H, Xie Y, McGill MR. Review Article Acetaminophen-induced Liver Injury : from Animal Models to Humans. 2014;2:153–61.
 91. Coen M, Ruepp SU, Lindon JC, Nicholson JK, Pognan F, Lenz EM, et al. Integrated application of transcriptomics and metabolomics yields new insight into the toxicity due to paracetamol in the mouse. *J Pharm Biomed Anal*. 2004;35(1):93–105.
 92. Certara. What are compartmental models? [Internet]. 2011 [cited 2019 Apr 15]. Available from: <https://www.certara.com/2011/01/05/what-are-compartmental-models/>
 93. Boroujerdi M. Pharmacokinetics and toxicokinetics. Taylor & Francis; 2012.
 94. Metzler CM. Open of the Two-Compartment Usefulness Model in Pharmacokinetics. *J Am Stat Assoc*. 2012;66(333):49–53.
 95. Wright MH. Nelder, Mead, and the Other Simplex Method. 2010 [cited 2017 May 12]; Available from: <http://www.math.uiuc.edu/documenta/vol->

96. Mathworks. Matlab. Natick, Massachusetts: The MathWorks Inc.; 2016.
97. Hinson JA, Roberts DW, James LP. Mechanisms of Acetaminophen-Induced Liver Necrosis. *Pharmacology*. 2010;196(196):1–34.
98. Prescott L. Kinetics and metabolism of paracetamol and phenacetin. *Br J Clin Pharmacol*. 1980;10(2):291–8.
99. Teuscher N. What are direct and indirect pharmacodynamic models? 2011.
100. Sharma A, Jusko WJ. Characteristics of indirect pharmacodynamic models and applications to clinical drug responses. *Br J Clin Pharmacol*. 1998;45(3):229–39.
101. Mathworks. Simbiology Toolbox: User's guide (R2017b). Natick, Massachusetts: The MathWorks Inc.; 2016.
102. Martins J, Kroo I, Alonso J. An automated method for sensitivity analysis using complex variables. *AIAA Pap*. 2000;
103. Saranya C, Guimarães M. A Study on Normalization Techniques for Privacy Preserving Data Mining. *Int J Eng Technol*. 2015;5:2701–4.
104. Maaten L Van Der. Accelerating t-SNE using tree-based algorithms. 2014;15:3221–45.
105. Thorne J. Introduction to Principal Components and FactorAnalysis. 2018.
106. Smith LI. A tutorial on Principal Components Analysis Introduction. *Statistics (Ber)* [Internet]. 2002;51:52. Available from: <http://www.mendeley.com/research/computational-genome-analysis-an-introduction-statistics-for-biology-and-health/>

107. Dubey A. The mathematics behind principal component analysis [Internet]. 2018 [cited 2019 Apr 12]. Available from: <https://towardsdatascience.com/the-mathematics-behind-principal-component-analysis-fff2d7f4b643>
108. Pathak M. Introduction to TSNE [Internet]. 2018. Available from: <https://www.datacamp.com/community/tutorials/introduction-t-sne>
109. Laurens van der Maaten, Hinton G. Visualizing Data using t-SNE Laurens. J Mach Learn Res. 2008;9:2579–605.
110. Starkweather J, Moske AK. Multinomial logistic regression. Multinomial Logist Regres. 2011;51(6):404–10.
111. Welling M. Fisher Linear Discriminant Analysis. Science (80-). 2009;1(2):1–3.
112. Leung KM. Naive bayesian classifier. Polytech Univ Dep Comput Sci Risk Eng. 2007;
113. Kuhkan M. A Method to Improve the Accuracy of K-Nearest Neighbor Algorithm. Int J Comput Eng Inf Technol (IJCEIT), Vol 8, Issue 6, June 2016. 2016;8(6):90–5.
114. Samworth RJ. Optimal Weighted Nearest Neighbour Classifiers. Ann Stat. 2012;40(5):2733–63.
115. Hair J., Anderson R., Tatham R., Black W. Multivariate analysis. 3rd ed. New York: Macmillan; 1995.
116. Zelen M. Multinomial response models. Comput Stat Data Anal. 1991;12(2):249–54.
117. Tharwat A, Gaber T, Ibrahim A, Hassanien AE. Linear discriminant analysis: A detailed tutorial. 2017;

118. Thatcher T. Discriminant Analysis Documentation. 2016;
119. Freitas AT. Naive Bayes Classifier. 2016;1–12.
120. Akella, Alan, Downey A, Gorm A, Lincoln A, Arauzo, et al. Naive Bayes classifier. 2010;1–9.
121. Bejar J. K Nearest Neighbours. Expo Times. 1985;96(7):209–10.
122. Hechenbichler K, Schliep K, Wilson A. Weighted k-Nearest-Neighbor Techniques and Ordinal Classification. Discuss Pap. 2004;399.
123. Bewick V, Ball Jonathan. Statistics Review 14: Logistic Regression. Crit Care Febr. 2005;9(1).
124. Brannick M. Logistic Regression [Internet]. 2016 [cited 2017 Apr 28]. Available from: <http://faculty.cas.usf.edu/mbrannick/regression/Logistic.html>
125. Antoine DJ. Chemical and molecular markers of hepatic drug bioactivation, apoptosis and necrosis. University of Liverpool; 2009.
126. Lea JD, Clarke JI, Mcguire N, Antoine DJ. Redox-Dependent HMGB1 Isoforms as Pivotal Co-Ordinators of Drug-Induced Liver Injury: Mechanistic Biomarkers and Therapeutic Targets. Antioxidants Redox Cycl. 2016;24(12):1–45.
127. Investopedia. Sensitivity Analysis [Internet]. Investopedia. 2017 [cited 2017 Jul 29]. Available from: <http://www.investopedia.com/terms/s/sensitivityanalysis.asp>
128. Hammonds JS, Hoffman FO, Bartell SM. An Introductory Guide to Uncertainty Analysis in Environmental and Health Risk Assessment. 1994.
129. Butler T. Bayesian Push-forward based Inference A Consistent Bayesian Formulation for Stochastic Inverse Problems Based on Push-forward

Measures. 2017;1–25.

130. Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, et al. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*. 2009;25(15):1923–9.
131. Mason CL, Leedale J, Tasoulis S, Jarman I, Antoine DJ, Webb SD. Systems Toxicology Approach to Identifying Paracetamol Overdose. *CPT Pharmacometrics Syst Pharmacol*. 2018;7(6):394–403.
132. Raue A, Becker V, Klingmüller U, Timmer J, Raue A, Becker V, et al. Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos*. 2010;20.
133. Antoine DJ, Jenkins RE, Dear JW, Williams DP, Mitchell R, Sharpe MR, et al. Molecular forms of HMGB1 and Keratin-18 as mechanistic biomarkers for mode of cell death and prognosis during clinical acetaminophen hepatotoxicity. *Hepatology*. 2012;56(5):1070–9.
134. Longo DM, Generaux GT, Howell BA, Siler SQ, Antoine DJ, Button D, et al. Refining Liver Safety Risk Assessment : Application of Mechanistic Modeling and Serum Biomarkers to Cimaglermin Alfa (GGF2) Clinical Trials. 2017;102(6):961–9.
135. Craig DGN, Bates CM, Davidson JS, Martin KG, Hayes PC, Simpson KJ. Overdose pattern and outcome in paracetamol- induced acute severe hepatotoxicity. 2011;
136. Dear JW, Clarke JI, Francis B, Allen L, Wraight J, Shen J, et al. Risk stratification after paracetamol overdose using mechanistic biomarkers : results from two prospective cohort studies. *Lancet Gastroenterol Hepatol*.

- 2017;1253(17):1–10.
137. Gujral JS, Knight TR, Farhood A, Bajt ML, Jaeschke H. Mode of cell death after acetaminophen overdose in mice: Apoptosis or oncotic necrosis? *Toxicol Sci.* 2002;67(2):322–8.
138. Ozer J, Ratner M, Shaw M, Bailey W, Schomaker S. The current state of serum biomarkers of hepatotoxicity. *Toxicology.* 2008;245(3):194–205.
139. Nicotera P, Leist M, Ferrando-May E. Intracellular ATP, a switch in the decision between apoptosis and necrosis. *Toxicol Lett.* 1998;102–103(8):139–42.
140. Cole LA. Adenosine Triphosphate Energetics. *Biol Life.* 2016;(Chapter 5):65–77.
141. Nature. Western blot [Internet]. Nature education. 2014 [cited 2019 Mar 23]. Available from: <https://www.nature.com/scitable/definition/western-blot-288>
142. PrimerDesign. PrimerDesigns Beginners Guide To Real-Time PCR. PrimerDesign. 2015. p. 1–8.
143. Brown M, Wittwer C. Flow cytometry: Principles and clinical applications in hematology. *Clin Chem.* 2000;46(8):1221–9.
144. Antibodies-Online. What is flow cytometry (FACS analysis)? [Internet]. 2013 [cited 2019 Mar 23]. Available from: <https://www.antibodies-online.com/resources/17/1247/what-is-flow-cytometry-facs-analysis/>
145. Crawley M. *Statistics: An introduction using R.* 2017;1–9.
146. Durbin J, Watson G. Testing for correlation in least squares regression. II. *Biometrika.* 1951;38:159–77.
147. Brown M, Bjorksten AEA, Medved AEI. Pharmacokinetics of intravenous N -

- acetylcysteine in men at rest and during exercise. 2004;717–23.
148. Shen F, Coulter C V, Isbister GK, Duffull SB, Shen F, Coulter C V, et al. A dosing regimen for immediate N-acetylcysteine treatment for acute paracetamol overdose A dosing regimen for immediate N-acetylcysteine treatment for acute paracetamol overdose. 2017;3650:5–10.
 149. Saito C, Zwingmann C, Jaeschke H. Novel mechanisms of protection against acetaminophen hepatotoxicity in mice by glutathione and N-acetylcysteine. *Hepatology*. 2010;51(1):246–54.
 150. Gove CD, Hughes RD. Liver regeneration in relationship to acute liver failure. *Gut*. 2007;32:S92–6.
 151. Malhi H, Guicciardi ME, Gores GJ. Hepatocyte death: A clear and present danger. *Gastroenterology*. 2010;90(3):1165–94.
 152. Bateman DN, Dear JW, Thanacoody HKR, Thomas SHL, Eddleston M, Sandilands EA, et al. Reduction of adverse effects from intravenous acetylcysteine treatment for paracetamol poisoning : a randomised controlled trial. *Lancet*. 2012;383(9918):697–704.
 153. Sodani K, Patel A, Kathawala RJ, Chen ZS. Multidrug resistance associated proteins in multidrug resistance. *Chin J Cancer*. :58–73.
 154. Riss TL, Moravec RA, Niles AL, Benink HA, Worzella TJ. *Cell Viability Assays*. 2016;1–31.
 155. McGill MR, Williams D, Xie Y, Ramachandran A, Jaeschke H. Acetaminophen-induced Liver Injury in Rats and Mice: Comparison of Protein Adducts, Mitochondrial Dysfunction, and Oxidative Stress in the Mechanism of Toxicity. *Toxicol Appl Pharmacol* Novemb. 2012;1(2643):387–94.

156. Mund ME, Quarcoo D, Gyo C, Brüggmann D, Groneberg DA. Paracetamol as a toxic substance for children : aspects of legislation in selected countries. *J Occup Med Toxicol.* 2015;1–7.
157. Caparrotta TM, Antoine DJ, Dear JW. Are some people at increased risk of paracetamol-induced liver injury? A critical review of the literature. *Eur J Clin Pharmacol.* 2018;74(2):147–60.
158. Ging P, Mikulich O, Reilly KMA. Unexpected paracetamol (acetaminophen) hepatotoxicity at standard dosage in two older patients : time to rethink 1 g four times daily ? 2016;566–7.
159. Yaghi C, Assaf A. Acetaminophen Toxicity at Therapeutic Doses. 2017;3(11):1–13.
160. Claridge L, Eksteen B, Smith A, Shah T, Holt A. Acute liver failure after administration of paracetamol at the maximum recommended daily dose in adults. *BMJ.* 2010;341.
161. Pearce B, Grant IS. Acute liver failure following therapeutic paracetamol administration in patients with muscular dystrophies. *Anaesthesia.* 2008;63(1):89–91.
162. Yang H, Sun L, Li W, Liu G, Tang Y. In Silico Prediction of Chemical Toxicity for Drug Design Using Machine Learning Methods and Structural Alerts. *Front Chem.* 2018;6(February):1–12.
163. Polson J, Wians FH, Orsulak P, Fuller D, Murray NG, Koff JM, et al. False positive acetaminophen concentrations in patients with liver injury. *Clin Chim Acta* [Internet]. 2008;391(1–2):24–30. Available from: <http://dx.doi.org/10.1016/j.cca.2008.01.018>
164. StanfordUniversity. Maximum Likelihood Estimation. 2011. p. 1–3.

165. Melina R. Why do medical researchers use mice? [Internet]. 2010 [cited 2019 Apr 28]. Available from: <https://www.livescience.com/32860-why-do-medical-researchers-use-mice.html>
166. NuffieldBioEthics. Animal use in toxicity studies. 2005;
167. Taylor J. Introduction to regression and analysis of variance: Fixed and random effects. 2018. p. 1–19.
168. Lavielle M. Mixed effects models for the population approach: Models, tasks, methods and tools. *Mix Eff Model Popul Approach Model Tasks, Methods Tools*. 2014;(January):1–361.
169. Carter AA, Rosenbaum SE, Dudley MN. Review of methods in population pharmacokinetics. *Clin Res Regul Aff*. 1995;12(1):1–21.
170. Gastonguay MR. Full Covariate Models as an Alternative to Methods Relying on Statistical Significance for Inferences about Covariate Effects: A Review of Methodology and 42 Case Studies. *Page 2011*. 2011;(Abstr 2299):A16.
171. Lixoft. Lixoft - Monolix. Built for model based drug development computing [Internet]. 2019 [cited 2019 Apr 28]. Available from: <http://lixoft.com/products/monolix/>
172. Lixoft. Monolix. 2018.
173. Do C, Batzoglou S. What is the expectation maximisation algorithm? *J Oper Res Soc*. 2006;46(4):543.
174. Kuhn E, Lavielle M. Coupling a stochastic approximation version of EM with an MCMC Procedure. *ReCALL*. 2004;8(January):12–24.
175. Robert CP. The Metropolis – Hastings algorithm. 2016;(Mcmc):1–15.
176. Monolix. Population parameter estimation using SAEM - Simulated Annealing

- [Internet]. 2018 [cited 2019 Apr 16]. Available from: <http://monolix.lixoft.com/tasks/population-parameter-estimation-using-saem/simulated-annealing/>
177. Earl DJ, Deem MW. Monte Carlo Simulations. 2008;443:25–37.
178. Popix. The SAEM algorithm for estimating population parameters [Internet]. 2016. Available from: http://wiki.webpopix.org/index.php/The_SAEM_algorithm_for_estimating_population_parameters
179. Henderson D, Johnson AW. The Theory and Practice of Simulated Annealing. *Handb Metaheuristics*. 2003;287–319.
180. Mian P, Allegaert K, Spriet I, Tibboel D, Petrovic M. Paracetamol overdose in older people – towards evidence based dosing. *Drugs and Aging* [Internet]. 2018;(35):603–24. Available from: <https://doi.org/10.1007/s40266-018-0559-x>
181. Nguyen THT, Mouksassi MS, Holford N, Al-Huniti N, Freedman I, Hooker AC, et al. Model evaluation of continuous data pharmacometric models: Metrics and graphics. *CPT Pharmacometrics Syst Pharmacol*. 2017;6(2):87–109.
182. NIDABT. The National Instructure on Drug Abuse Blog Team. Overdoses in Youth [Internet]. 2019 [cited 2019 Apr 17]. Available from: <https://teens.drugabuse.gov/drug-facts/drug-overdoses-youth>
183. Orimo H. Reviewing the definition of elderly. *Nippon Ronen Igakkai Zasshi Japanese J Geriatr*. 2012;43(1):27–34.
184. OnAverage. Average child weight [Internet]. 2019 [cited 2019 Apr 17]. Available from: <https://www.onaverage.co.uk/body-averages/average-child-weight>

185. CDCP. Centers for Disease Control and Prevention. Defining overweight and obesity [Internet]. 2017 [cited 2019 Apr 17]. Available from: <https://www.cdc.gov/obesity/adult/defining.html>
186. Welling PG, Tse FLS. Factors contributing to variability in drug pharmacokinetics. I. Absorption. *J Clin Hosp Pharm.* 1984;9(3):163–79.
187. Huang HS, Ho CH, Weng SF, Hsu CC, Wang JJ, Su S Bin, et al. Long-term mortality of acetaminophen poisoning: A nationwide population-based cohort study with 10-year follow-up in Taiwan. *Scand J Trauma Resusc Emerg Med.* 2018;26(1):1–7.
188. Cook D, Brown D, Alexander R, March R, Morgan P, Satterthwaite G, et al. Lessons learned from the fate of AstraZeneca's drug pipeline: A five-dimensional framework. *Nat Rev Drug Discov.* 2014;13(6):419–31.

Appendices

Model equations – Dynamic equations for the full pre-clinical systems toxicology model: APAP, corresponding toxicity biomarkers, NAC and hepatocyte functionality. Models for toxicity biomarkers are provided for the fed and fasted_f cases (note that we use subscript f to denote parameters and variables in the fasted case). Note that for Full and Fragmented K18, $H(x)$ is the Heaviside function where $H(x) = 1$ when $x \geq 0$ and $H(x) = 0$ when $x \leq 0$.

PK

$$\text{APAP central } [C_c]: \quad \frac{d[C_c]}{dt} = \frac{k_a D_0 e^{-k_a t}}{V_c} + k_{21} [C_p] \frac{V_p}{V_c} - k_{12} [C_c] - k_{el} [C_c] \quad (\text{A-1})$$

$$\begin{aligned} \text{APAP peripheral} \\ [C_p]: \end{aligned} \quad \frac{d[C_p]}{dt} = k_{12} [C_c] \frac{V_c}{V_p} - k_{21} [C_p] \quad (\text{A-2})$$

GSH

$$\text{Fed GSH } [gsh]: \quad \frac{d[gsh]}{dt} = k_o (gsh_0 - [gsh]) + \zeta [N_A] - \frac{\xi k_{el} C_c [gsh]}{[gsh] + k_{pr}}, \quad (\text{A-3})$$

$$\begin{aligned} \text{Fasted GSH} \\ [gsh_f]: \end{aligned} \quad \frac{d[gsh_f]}{dt} = \delta k_o (gsh_0 - [gsh_f]) + \zeta [N_A] - \frac{\xi k_{el} C_c [gsh_f]}{[gsh_f] + k_{pr}}, \quad \delta < 1 \quad (\text{A-4})$$

ALT and HMGB1, (fed and fasted)

$$\begin{aligned} r = [alt] \text{ or} \\ [hmgb1]: \end{aligned} \quad \frac{d[r]}{dt} = r_0 k_{out} \left(\frac{R_{50}^n + gsh_0^n}{R_{50}^n} \right) \left(1 - \frac{g_{max} [gsh]^n}{R_{50}^n + [gsh]^n} \right) - k_{out} [r], \quad (\text{A-5})$$

In the fed case, $g_{max} = 1$, in the fasted case, $g_{max} < 1$.

Full K18 (fed and fasted)

$$\begin{aligned} \text{Full K18} = \\ [K18]: \end{aligned} \quad \frac{d[K18]}{dt} = r_0^{18} k_{out}^{18} + k_{max} k_{in}^{18} H(gsh_{\theta_2} - [gsh]) - k_{out}^{18} [k18], \quad (\text{A-6})$$

In the fed case, $k_{max} = 1$, in the fasted case, $k_{max} < 1$.

Fragmented K18

$$\frac{d[fK18]}{dt} = r_0^{f18} k_{out}^{f18} + k_{in}^{f18} \left(H([gsh] - gsh_{\theta_1}) - H([gsh] - gsh_{\theta_2}) \right) - k_{out}^{f18} [fk18], \quad (\text{A-1})$$

$$\frac{d[fK18_f]}{dt} = r_0^{f18} k_{out}^{f18} + k_{in}^{f18} \left(H([gsh_f] - gsh_{\theta_1} - \varepsilon) - H([gsh_f] - gsh_{\theta_2}) \right) - k_{out}^{f18} [fk18_f], \quad (\text{A-2})$$

$$\text{Fed fragmented K18} = [fK18], \text{ Fasted fragmented K18} = [fK18_f], \varepsilon > 0. \quad (\text{A-3})$$

NAC

$$\frac{d[N_A]}{dt} = \frac{NDose}{V_{NA}} + k_{N21}[N_B] \left(\frac{V_{NB}}{V_{NA}} \right) + k_{N31}[N_C] \left(\frac{V_{NC}}{V_{NA}} \right) - (k_{N12} + k_{N13} + k_{Nel})[N_A], \quad (\text{A-10})$$

$$\frac{d[N_B]}{dt} = k_{N12}[N_A] \left(\frac{V_{NA}}{V_{NB}} \right) - k_{N21}[N_B], \quad (\text{A-11})$$

$$\frac{d[N_C]}{dt} = k_{N13}[N_A] \left(\frac{V_{NA}}{V_{NC}} \right) - k_{N31}[N_C], \quad (\text{A-12})$$

Hepatocyte functionality

$$\frac{dH}{dt} = rH \left(1 - \frac{H+Z}{H_{max}} \right) - \eta[N]H, \quad (\text{A-13})$$

$$\frac{dZ}{dt} = \eta[N]H - \delta_z Z, \quad (\text{A-14})$$

Model initial conditions

Variable	Initial condition (fed case)	Initial condition (fasted case)
$[C_c]$	0 $\mu\text{mol/l}$	0 $\mu\text{mol/l}$
$[C_p]$	0 $\mu\text{mol/l}$	0 $\mu\text{mol/l}$
$[gsh]$	559.47497 $\mu\text{mol/l}$	374.0909 $\mu\text{mol/l}$
$[alt]$	0.7626 $\mu\text{mol/l}$	0.9528 $\mu\text{mol/l}$
$[hmgbl]$	0.0005 $\mu\text{mol/l}$	0.0007 $\mu\text{mol/l}$
$[fullk18]$	0.0088 $\mu\text{mol/l}$	0.0113 $\mu\text{mol/l}$
$[fragk18]$	0.0977 $\mu\text{mol/l}$	0.1634 $\mu\text{mol/l}$
$[N_A]$	0 $\mu\text{mol/l}$	
$[N_B]$	0 $\mu\text{mol/l}$	
$[N_C]$	0 $\mu\text{mol/l}$	
$[H]$	1	
$[Z]$	0	

Table A1: Initial conditions for each variable within the dynamic APAP systems toxicology pre-clinical framework.