Modelling the effects of dietary toxicity on the growth of rodents: applications in chemical risk assessment

Thomas Martin

PhD

University of York

Environment & Geography

February 2021

Abstract

Before being approved for use, pesticides undergo ecological risk assessment (ERA) to identify any adverse effects they may have on non-target species. For terrestrial mammals, dietary toxicity studies on laboratory animals, lasting up to two years, are used to assess chronic toxicity. However, such studies are limited in their ability to assess ecological risk, as constant dietary exposure to pesticides over long periods is an unrealistic scenario in the field. *In silico* models have long been suggested as a way of addressing the limitations of laboratory testing, by simulating the processes underlying toxicity. Toxicokinetic-toxicodynamic (TK-TD) models simulate both the uptake of a chemical into an animal's body, and the resulting stress on biological processes which leads to observed effects. This thesis comprises a series of studies in which TK-TD models were developed to simulate the effects of dietary toxicity on the growth of rodents, and applications for ERA and the reduction of animal testing were explored. Unpublished raw data from regulatory toxicokinetics and toxicity studies on several pesticides were provided by Syngenta. First, a protocol for TK-TD modelling with regulatory datasets was developed. This produced TK-TD models that successfully predicted the growth rate of laboratory rats exposed to various pesticides and dietary dose levels. These TK-TD models were then used to facilitate comparisons between *in vitro* and *in vivo* toxicity data. While no predictive relationships were identified, this study demonstrated how *in vitro* and *in silico* methods may be combined to obviate the need for animal testing in future. Finally, a new method for converting observed feeding data into growth model inputs was developed, improving model accuracy and realism. This enhances the ability of models to separate the effects of toxicity and feeding avoidance on body weight, a valuable insight for assessing ecological risk.

List of Contents

List of Tables

List of Figures

Additional Content

Excel file entitled 'Martin_104015910_C3_Parameters_and_Results.xlsx', includes full model parameters and results for every dataset used in Chapter 3.

Acknowledgements

Declaration

I declare that this thesis is a presentation of original work and that I am the sole author. In chapters presenting experiments and data, I have chosen to use the personal pronoun "we" as these chapters were written either for publication, or in preparation for publication, and so acknowledge the involvement of my supervisors and others. The work within this thesis has not previously been presented for an award at this, or any other, University. All sources are acknowledged in the References section. The table below reports the papers arising from this thesis.

Title	Thesis chapter	Status	Journal
Common ground between growth models of rival theories: A useful illustration for beginners	2	Published September 2019	Ecological Modelling
Toxicokinetic–Toxicodynamic Modelling of the Effects of Pesticides on Growth of Rattus norvegicus	3	Published November 2019	Chemical Research in Toxicology

Chapter 1 - Introduction

1.1 Overview

Mechanistic effect models (MEMs) are viewed as potentially important tools for chemical risk assessment (Grimm and Martin, 2013, Ducrot et al., 2016). Such models aim to approximate biological processes with sufficient accuracy to predict responses to scenarios such as exposure to toxins. Mechanistic, rather than empirical, predictions are of greater value for extrapolation from laboratory to field and from individuals to populations and ecological systems. It is therefore argued that modelling methods can provide more relevant proxies for risk assessment than the results of laboratory testing on individuals (Grimm and Martin, 2013) and, if utilised effectively, can improve chemical safety outcomes while reducing the need for animal testing (Jager, 2016).

This project focusses on the development of *in silico* models able to accurately simulate the growth of rodents exposed to pesticides in their diet for periods of up to two years. Prior to this project, there had been very few attempts to develop such models for mammals (Desforges et al., 2017). This task involves modelling the processes of resource assimilation and allocation involved in growth, the exposure and uptake of pesticides, and the stress this places on the organism.

The first aim of this project is to develop a model that can contribute to pesticide risk assessment. For example, as an analytical tool, providing new insight into the results of laboratory toxicity studies, or as a means of extrapolating from laboratory data to untested, ecologically relevant scenarios. Secondly, we will investigate how toxicological models can facilitate comparisons between *in vitro* and *in vivo* data. This could allow the prediction of organism level effects from experiments on cultured cells, potentially reducing the need for animal testing in chemical risk assessment.

This chapter provides a literature review with respect to each of these broad objectives, identifying the knowledge gaps the project aims to address, potential obstacles to overcome, and suitable methods for doing so. The chapter concludes with an overview of the thesis structure, providing a chapter-by-chapter summary of how our objectives were approached over the course of the project.

1.2 Objective 1: Develop TK-TD growth model for rodents to address the limitations of pesticide risk assessment

1.2.1 Limitations of current long-term risk assessment measures

Predicting the long-term effects that a chemical, such as a pesticide, will have on wildlife when used in the environment is a major challenge. To be truly certain of a chemical's ecological impact,

we would have to test it "on every species, include the effects from all other natural and anthropogenic stressors, assess effects on populations and ecosystems in all regions and under all environmental conditions of concern, and use endpoints for quantifying risk that would perfectly indicate whether or not the effect on the environment is acceptable" (Grimm and Martin, 2013). Clearly, this would not be feasible for a number of reasons. In reality, ecological risk assessment (ERA) relies on studies conducted on just a few species, for one chemical at a time and under standard conditions. Such assessment may on occasion fail to highlight the adverse effects of pesticides. For example, pesticide risk assessment in the USA is criticised for its lenience, with 72 pesticides which are banned in the EU - for their impacts on human health or the environment - still approved by the Environmental Protection Agency (Donley, 2019). Conversely, it is also possible that stringent risk assessments may overestimate the risk of a chemical by investigating only the effects of constant exposure, an unlikely scenario in the field (Ducrot et al., 2016). In such cases, chemicals which could benefit food production without damaging the environment may not be approved. Through addressing the gap between laboratory and field, ERA can be made more realistic, avoiding either kind of error occurring.

At a 2005 workshop devoted to long term risk assessment of pesticides to birds and mammals, five main obstacles were identified. The first was relating toxicity endpoints measured in laboratory testing, which are generally reported as summary statistics, in a relevant way to animals living in the field. The second was accurately predicting sensitivity to a chemical from one species to another. Third was estimating the exposure of animals which are free to move in a dynamic environment. Next was accounting for the mismatch in exposure between laboratory and field. The final challenge was evaluating effects at population level when laboratory testing only provides data on individual level effects (Hart and Thompson, 2005).

While these obstacles were identified fifteen years ago, they remain relevant and are worth bearing in mind when considering the current framework for assessing long term toxicity of pesticides in mammals and the ecological risk this poses. The subacute to chronic toxic effects of new active ingredients are assessed with repeated oral dose studies lasting from 28 days to 2 generations. The laboratory bred form of the brown rat, *Rattus norvegicus* (Palm, 1975), is the most common subject although a few other mammalian species may be used (Mineau, 2005). The highest dose should induce toxic effects but not severe suffering or death and generally dose levels decrease as test duration increases. Many endpoints are monitored; these include body weight, reproductive output, reproductive defects, responses to neurological stimuli and organ damage or malformation (OECD, 1998). These data are then used to determine a number of reference doses. The no observed effect level (NOEL) is the highest dose at which there are no statistically significant effects on any endpoint (EFSA, 2006). The no observed adverse effect level

(NOAEL) is the highest dose at which any observed effects are not deemed biologically or ecologically relevant (EFSA, 2009b). The lowest observed adverse effect level (LOAEL) is the lowest dose at which biologically relevant effects are observed (EFSA, 2009a).

Whether the collection of summary statistics is appropriate for ecological risk assessment (ERA) is the subject of much debate among scientists (Murado and Prieto, 2013, Green et al., 2013, Jager, 2012, Jager, 2011, Sanchez-Bayo, 2012). One of the issues with this approach identified by regulators themselves is how to interpret the ecological relevance of effects in tests that were designed to protect human health (EFSA, 2006). Humans' exposure to pesticides typically results from accidental ingestion or inhalation by workers handling them (chronic or acute exposure) or consumption of pesticide residues on food (chronic exposure). Neither of these scenarios are particularly representative of the fluctuating exposure faced by wildlife. Moreover, some of the endpoints monitored in toxicity studies are difficult to relate to ecological impacts. According to guidance it should be considered 'whether the effect could lead to a functional deficit later on' however this still leaves a degree of subjectivity and requires consideration on a "case by case basis" (EFSA, 2009b).

Tier 1 risk assessment aims to determine the risk to species that inhabit agricultural land by first calculating the daily dietary dose (DDD) of 'generic focal species', considered representative of the species at risk. These species should be small in body size and abundant in the relevant crop type in the area of interest. For example, the wood mouse, *Apodemus sylvaticus,* and the field vole, *Microtus agrestis* are commonly considered in European risk assessment (The Danish Environmental Protection Agency, 2014, EFSA, 2009b)

The full calculation of DDD is as follows:

$$\boldsymbol{DDD} = \frac{AR \times RUD \times FIR \times PT \times PD}{BW} \tag{1.1}$$

Where *AR* is application rate ($kg_{A.I.} \times Ha^{-1}$), *RUD* is the mean residue per unit dose resulting from an application rate of 1kg/Ha ($mg \times kg_{food}^{-1} \times kg_{AI}^{-1} \times Ha^{-1}$), *FIR* is food intake rate ($g \times day^{-1}$), *PT* is the proportion of diet from the treated area, *PD* is the proportion of the affected food type in the diet and *BW* is body weight (kg). In practice, the daily dietary dose is initially calculated as

$$DDD = AR \times Shortcut\ Value \times TWA \times MAF_m \tag{1.2}$$

using a shortcut value for the crop/generic species of interest and where *TWA* is time weighted average factor and MAF_m is the multiple application factor. The chronic toxicity exposure ratio (*TER*) is then calculated as:

$$TER = \frac{NOAEL}{DDD} \tag{1.3}$$

The NOAEL value used is the lowest from the 2 generational study (as a daily dose, $mg_{AI} \times kg_{BW}^{-1} \times d^{-1}$). If the *DDD* is greater than one fifth of the NOAEL (TER < 5) then this stage of assessment is failed, and further risk assessment must be conducted. The *RUD* is calculated as the average residue over a 21-day period assuming that the pesticide in question has a 10 day half-life (DT_{50}), that is the time taken for half of the chemical to degrade in the environment. The 21 day and 10 day periods are selected arbitrarily (EFSA, 2009b).

While *TER* based assessment is certainly protective (based on NOAEL derived from two year exposure, far longer than would be expected in the field), there are clear areas for improvement with regard to ecological realism and the obstacles identified by Hart and Thompson (2005). There is a major disparity in the exposure patterns used to derive the values that determine the *TER*. The *DDD* is calculated for exposure over 21 days while the NOAEL is derived from constant exposure over 2 generations. Moreover, the true DT_{50} of a pesticide is not considered, instead it is assumed to be 10 days. Finally, this method treats the animal as a 'black box' considering only the external dose and ignoring the internal processes of absorption, distribution, metabolism and elimination (ADME) despite data on these processes often being available (EFSA, 2009b). Finally, this method also assumes that the species used in laboratory testing is representative of the focal species in the field. This assumption is more difficult to avoid without testing on a wider range of species, an undesirable solution for multiple reasons.

1.2.2 Modelling approaches for ecological risk assessment

In toxicity studies, the toxic effects of a chemical are generally related to the administered dose; this may be an ingested, dermal or inhaled dose for terrestrial species or toxicant concentration in water for aquatic species (OECD, 2020). However, the biologically effective dose is the portion of the administered dose that reaches its target within the organism (Pelkonen and Turpeinen, 2007, Tsaioun et al., 2016). Toxicants may vary greatly in the rate at which they are taken up or eliminated and so there is a clear advantage to relating toxic effects to internal rather than the external concentration.

Toxicokinetics (TK) refers to a chemical's absorption, distribution, metabolism and excretion by a biological system. TK models mathematically describe these processes and so are able to translate an external concentration into an internal concentration over time and vice-versa (Tsaioun et al., 2016). Toxicodynamic (TD) models meanwhile predict how the internal concentration will impact upon biological processes. The effect on these processes is then related to endpoints such as growth, reproduction or survival. When used in combination, the resulting

TK-TD models are able to predict toxic effects of exposure to a toxicant, which may vary through time (Ashauer and Escher, 2010).

TK-TD models work at organism level and are a part of a broader class of predictive models known collectively as mechanistic effect models (MEMs). This term refers to models that simulate the mechanisms by which chemicals affect various levels of biological organisation, from individuals to eco-systems (Grimm and Martin, 2013). These are a valuable tool for extrapolating from laboratory to field scenarios and it is considered that making MEMs an integral part of chemical risk assessment will make the process more comprehensive, ecologically relevant & cost effective while reducing uncertainty (Hommen et al., 2016a).

ERA is generally concerned with higher levels of biological organisation such as populations and communities. However, experimental data collection at these levels is highly impractical and the resulting data are difficult to interpret. In order to make higher level predictions, the level of the individual is vital because it is compatible with experimentation and has direct relevance to population level effects (Jager, 2016). Individual level data are of particular relevance to risk assessment for birds and mammals, as guidelines state that no mortality associated with pesticide use should be accepted (EFSA, 2009b).

The potential of TK-TD modelling to enhance ERA is great. In a workshop organised to investigate how MEMs can be applied in ERA, researchers were provided with data sets on hypothetical pesticides that had failed tier 1 risk assessment (Hommen et al., 2016a). Modelling techniques were then used to address specific protection goals from the EFSA guidance. While the data used in these case studies were artificial, they were also realistic and intended to explore the capacity of modelling to improve ERA. Among the applications identified were using TK modelling to predict internal pesticide concentration in skylarks and wood mice feeding in crops, and to identify the residue levels that would cause mortality (Ducrot et al., 2016). Another study used the general unified threshold for survival (GUTS) model to provide survival rules in individual based population models (IBMs) for three important aquatic arthropod species (Dohmen et al., 2016).

The European Food Safety Authority (EFSA) has long recognised that TK models can be a powerful tool and may serve to reduce the need for animal testing in pesticide risk assessment (EFSA, 2009b). More recently, TK-TD models based on the GUTS framework (Jager et al., 2011) were recommended for use in ERA for aquatic invertebrates (EFSA, 2018). However models based on dynamic energy budget (DEB) theory (Kooijman, 2000) - which are able to predict sublethal effects on graded endpoints such as body weight and reproductive output (Ashauer et

al., 2011) - are currently deemed unsuitable for use in ERA. This is due to the current lack of user-friendly modelling tools (EFSA, 2018).

Developing 'DEBtox' (Kooijman and Bedaux, 1996b) or 'DEB-TKTD' models (Sherborne et al., 2020) suitable for use in ERA is now an active area of research with renewed interest in simplifying the framework (Jager, 2020). Up to now, very few studies have developed DEB-TKTD models for vertebrates (Zimmer et al., 2018, Sadoul et al., 2018). Outside of this project, only one published study applied this approach to a mammalian species, the American mink, (*Mustela vison*), producing promising results despite limited data (Desforges et al., 2017).

1.2.3 Modelling growth in mammals

Bodyweight is an important endpoint in mammalian risk assessment and making predictions on how this will be affected by toxicant exposure over time requires a suitable growth model. The growth of many species follows a sigmoid pattern and three mathematical models which have been frequently applied to mammalian growth are the logistic, Gompertz and Von Bertalanffy models. In order to compare the abilities of these three models to predict mammalian growth Zullinger (1984) fitted all three models to growth data for 49 species from across mammalian taxa. The Von Bertalanffy had the lowest residual sum of squares for the growth curves of 27 of these species, Gompertz for 8 and logistic for 14. The differences in model performance were reduced for larger data sets and it was concluded that in general the models are all reasonably well suited to empirically describing mammal growth.

Of these, only the Von Bertalanffy curve was developed for the purpose of modelling animal growth. In the original form of the equation, growth was defined as the difference between 'anabolism' and 'catabolism'. These processes were poorly defined in his original paper and the suggestion that the anabolism would be proportional to surface area while catabolism would be proportional to volume has also been criticised. With this in mind, all three models can be regarded as purely empirical models (Ricker, 1979) and therefore not suitable for extrapolation to novel scenarios such as restricted feeding or exposure to toxins.

While the biological rationale of Von Bertalanffy was heavily criticised, models have since been developed that make use of its general form while ascribing new biological theory. When modelling toxicity, these give the advantage that toxic effects can be modelled as stress on specific biological processes rather than simply the growth rate. One such theory is the Dynamic Energy Budget (DEB) (Kooijman, 2000). DEB theory considers body mass to be divided into structure (bones, muscles, organs etc.) and reserve (stored lipids, carbohydrates etc.). The theory is based on the principle that certain processes, such as somatic maintenance, are limited by volume while others, such as feeding and assimilation, are limited by surface area (e.g. surface area of feeding

appendages, gut surface area). The theory also relies on the well-known scaling relationships $Volume \propto Length^3$ and $Surface\ Area \propto Volume^{2/3} \propto Length^2$, provided body shape remains the same (isomorphic growth). While the DEB model equations are much more complex and the biological rationale much more thorough, the DEB growth model can be reduced to the form of the Von Bertalanffy curve.

One of the criticisms commonly levelled at DEB is that it contains many parameters, particularly when reserve is considered, which cannot be directly measured or even easily estimated (van der Meer, 2006a). An alternative growth model that claims to address these issues is that of the metabolic theory of ecology (MTE) (Brown and Sibly, 2012). The central equation of MTE states that a $Metabolic\ Rate \propto Body\ Mass^{\alpha}$, with α generally assigned the value ¾ based on Kleiber's law (Kleiber, 1932). MTE provides a biological rationale to Kleiber's law, suggesting that the vascular systems of plants and vertebrates supply resources to cells at a rate which scales with body mass to the power ¾ (Brown and Sibly, 2012, West et al., 2001, West et al., 1997). However, this proposed scaling relationship has faced some strong criticism (Kozlowski and Konarzewski, 2004).

The resulting growth model is considerably simpler than that of DEB, in terms of the number of species-specific parameters. Most parameter values are taxon-specific, based on interspecies trends, so are fixed across related species. However, despite the difference in theory the growth curve is very similar. So similar that, if α is given the value ⅔, the MTE growth model also reduces to the Von Bertalanffy equation. For all the debate around scaling exponents, this change has little effect on predictions (Sibly et al., 2013). In fact, the same data have been employed to demonstrate both a ¾ (Karasov, 2012) and ⅔ (Kearney and White, 2012) scaling relationship between mass and gut surface area of mammals.

While MTE may offer some advantages, it has many of the same drawbacks as DEB and some of its own. While the MTE growth model does have only three basic parameters, and so claims to be more efficient than DEB (Marquet et al., 2014), it is debatable whether their values are truly measurable. All are derived through fitting various model equations to interspecific data (Moses et al., 2008). Moreover, the mathematical and biological basis of MTE has also received significant criticism (Kerkhoff, 2012, Painter, 2005, Kozlowski and Konarzewski, 2004, Van Der Meer, 2006b). The greater number of parameters in DEB is not necessarily a negative for the purpose of TK-TD modelling, as this provides greater scope for hypothesising how the toxicant acts. Additionally, DEBkiss, a simplified version of DEB, was developed more recently (Jager et al., 2013). DEBkiss contains fewer parameters and omits reserve, thereby avoiding the most complex elements of DEB theory (van der Meer, 2006a). The first published study to use DEBkiss to model

toxic effects on growth and reproduction in mammals (Desforges et al., 2017) produced good fits to data. Good fits do not necessarily validate the theory underlying a model (Moses et al., 2008), independent data must be predicted successfully in order to test a theory. Nevertheless, a simpler version of DEB may provide a good starting point when exploring new applications for TK-TD modelling.

1.2.4 Modelling toxicokinetics and toxicodynamics

One of the challenges of modelling toxicokinetics in terrestrial mammals is that the primary route of pesticide uptake is oral consumption and so ingestion rate is a major consideration. This challenge may be approached using body burden models (BBMs) or more complex physiologically based toxicokinetic (PBTK) models. Body burden models treat the whole body as a single compartment, with the gut added as a 'depot' compartment from where the toxicant must be absorbed before it is truly considered inside the body. In this way, internal concentration can be predicted from oral dosing over time. BBMs are relatively simple and provide general suitability for toxicants that distribute evenly throughout the body of an organism without a specific target site (Bednarska et al., 2013a). PBTK models (Nichols et al., 1991, Li et al., 2017, Louisse et al., 2015) divide the body into a number of compartments, corresponding to organs and tissues, and predict the concentrations in those different compartments. Such models are more mechanistic but require a great deal of information about the species physiology, the toxicant's behaviour and how they interact. While PBTK models have been developed for the rat (Li et al., 2017, Louisse et al., 2015), standard toxicity experiments (OECD, 2010) are not designed with model validation in mind and are limited in terms of sample size and dose regimens. This makes it challenging to demonstrate the advantages of PBTK modelling over simpler models. In ecotoxicology, where many species and chemicals are considered, such an approach is often not possible nor necessary (Jager, 2015a). Indeed, for the purpose of wildlife risk assessment, a stronger need is felt for simplicity, generality and flexibility of models (Hunka et al., 2013) with total body burden often deemed sufficient (Bednarska et al., 2013a).

Toxicodynamics describes how a dose metric, for example the internal concentration, of a toxicant is linked to effects on the endpoints of interest. A common assumption in quantifying these processes is that of thresholds (Jager, 2015a). A well-known adage in toxicology is that 'the dose makes the poison', meaning that all substances, even nutrients or water, will become toxic at high enough concentrations within the body (Paracelsus, 1565, Trautmann, 2005). The TD module of DEB-TKTD model assumes that, for any xenobiotic, there exists a 'no effect concentration' (NEC) below which it causes no stress to any biological processes. Unlike the similarly named NOEL or NOAEL, the NEC is time independent and therefore has no relationship with duration of exposure. This assumption allows us to focus on the toxicant only by postulating

that all other xenobiotics within the organism are below their respective NECs. When the toxicant exceeds its NEC, it places 'stress' on one or more biological process. Assuming every molecule of toxicant beyond its threshold has the same effect on a given biological process leads to a 'linear with threshold' relationship, with which stress can be modelled quite simply. This does not mean that the effects on the endpoint will also show a linear relationship, this depends on the role of the stressed parameter in the model of growth, reproduction or other endpoint (Jager, 2015a). The same principle may also be used in survival modelling by assuming an organism's probability to die increases linearly when damage (a function of internal toxicant concentration) passes a threshold (Ashauer et al., 2015). If there is evidence that stress on a biological process does not show a linear relationship with toxicant concentration, then other relationships can be modelled. For example, if stress increases exponentially as toxicant concentration increases beyond its NEC, or if the stressed process has a maximum or minimum rate. Equally, this approach to modelling chemical stress could be applied to models which are not based on DEB.

The threshold approach is not without disadvantages though. Its simplicity is beneficial from a modelling perspective but not entirely representative of responses to toxicants. For example, hormesis describes a bi-phasic response to a chemical whereby low doses bring about beneficial effects while higher doses are toxic (Mattson, 2008). Additionally, thresholds may themselves be altered by chemical exposure through changes to chemical tolerance (Pietrzykowski and Treistman, 2008, Miller, 2001).

1.2.5 Obstacles to long-term Toxicity Modelling

So how well do TK-TD modelling approaches compare to other methods for addressing the five obstacles to long-term pesticide risk in mammals identified earlier (Hart and Thompson, 2005)? The first obstacle identified was the relevance of endpoints collected as part of chronic toxicity testing (Mineau, 2005). The published summary statistics like NOAEL and LOAEL are heavily criticised as they are external doses which are time dependent and so vary depending on the length of exposure (Jager, 2012, Murado and Prieto, 2013). It is also argued that the testing procedures in standardised conditions provide too little information on the role of biotic and abiotic factors that will undoubtedly play a role in the field. Indeed it has been argued that the laboratory testing procedures should be totally rethought in order to make risk assessment more ecologically relevant (Bednarska et al., 2013b) and also to provide data required by models rather than models be designed around the available data (Jager, 2016). TK-TD models can be used to derive more relevant endpoints such as average or peak internal concentration and the no effect concentration (NEC). Published data are generally of limited value to modellers but the raw datasets contain a great deal of useful information recorded on an individual basis. This includes many measured endpoints and often data on recovery after treatment (OECD, 1998, OECD, 2008).

Greater access to existing raw data can therefore be of great benefit to model development. Due to the commercial sensitivity of such data this requires cooperation with industry. If modelling were to become an integral part of ERA (Jager, 2016, Grimm and Martin, 2013, Hommen et al., 2016a, Forbes et al., 2009), then those working in industry would be the users. As such they would always have access to full data sets rather than only the published figures, thus removing this obstacle.

Another difficulty is the extrapolation of toxicity between species. A widely used measure is the species sensitivity distribution (SSD). This involves plotting the sensitivities of different species to a toxicant (in terms of NOAEL or other reference dose) against the proportion of tested species affected. By fitting a curve through the points a dose can be determined at which only X% of species will be affected (Shaw-Allen and Suter, 2016). The data requirements are normally only met for acute toxicity tests on vertebrates since chronic testing more costly, both in terms of animal use and man hours. This method is also not suitable for extrapolating to a specific focal species as it can only make predictions about the proportion of species in a community affected by a particular dose. An alternative method is interspecies correlation estimation (ICE), this involves plotting known reference doses of two species against each other and determining the relationship in sensitivity between them (Raimondo et al., 2007). This approach can be used to make predictions from one species to another, however only if both species have undergone the same testing for a selection of chemicals. This rules out predictions for any wild species not used in animal testing and the required paired data are generally unavailable for chronic toxicity. Cao et al. (2014) developed a method of predicting equivalent doses between species based on scaling of metabolic rate with body size within taxa. These predictions appear reasonably accurate, however testing this method relies on paired data and again this is generally only available for acute toxicity testing. TK modelling offers a more mechanistic approach by incorporating an organism's surface area to volume ratio into model equations. Whether absorption is through the gut, lungs, or skin, it can reasonably be postulated that the rate of absorption is governed by surface area. Once absorbed however, the internal concentration depends on the organism's volume or mass. This approach was used by Gergs et al. (2016) to predict the effects of exposure on survival of the freshwater planktonic crustacean *Daphnia magna* at different life stages. Using the same TD parameters, and so assuming that individuals of different sizes are equally sensitive to the same internal concentrations, survival was predicted reasonably well. While in this case the method was used for intraspecific predictions, using it to make predictions between species would in principle be no different to using allometric relationships. Though, this would rely on the assumption that body size is the only variable and differences in morphology and physiology between species have no impact on sensitivity to a toxicant.

The next major obstacle was how to address the mismatch in exposure between laboratory and field when relating exposure to effects. The approach used when calculating a TER is by using a time weighted average whereby the average pesticide residue over 21 days is determined assuming it has a DT_{50} of 10 days. Naturally, the time over which exposure is averaged has a major effect on this figure and should have some justification (Fischer, 2005) however 21 day averaging time and 10 day DT_{50} are selected arbitrarily (EFSA, 2009b). TK modelling avoids this issue by making dynamic predictions of how internal concentration varies throughout the period of exposure. One approach is to link TK predictions such as maximum internal concentration (C_{max}) or total internal exposure, calculated as the area under the curve (AUC), to time dependent reference doses such as NOAEL (Gajewska et al., 2014). TD modelling takes a different approach, instead linking internal concentrations to the stress on biological processes in real time (Jager, 2015a). TK-TD models can then make more appropriate predictions with regard to toxic effects, either using the same assumed consumption as in the TER calculation or more detailed data.

Quantifying pesticide exposure in field conditions is a very complex task. It has been noted by scientists that the current measures fail to capture the true dynamics of exposure in time and space (Di Guardo and Hermens, 2013). Several criticisms of the daily dietary dose estimate (equation 1.1) have already been discussed, chiefly the lack of ecological relevance of the time periods considered. There is though, a great deal of pertinent information that is overlooked by this calculation, as identified by Crocker (Crocker, 2005). For example, pesticide residue may vary between food types or in space, and animals may show a preference to forage in untreated areas (these are approximated by the parameters *PT* and *PD* in equation 1.1). Another consideration is that, depending on the length of exposure considered and the life stage of the focal species, an individual's body weight may change significantly during exposure. Weight changes during exposure as well as other temporal considerations can be captured by a TK-TD growth model, however describing spatial variation is beyond the capability of this approach. Fully accounting for spatio-temporal variation would require either careful calculation of exposure or the use of an individual based model within which a TK-TD model could be implemented (Liu et al., 2014).

The final hurdle identified in assessing long term pesticide risk to mammals was how to make predictions at the population level (Sibly et al., 2005). Again, this is beyond the scope of TK-TD modelling. However, as has been mentioned, several studies have already begun exploring the potential of TK-TD sub-models to enhance the predictions of IBMs (Gergs et al., 2016, Liu et al., 2014, Ashauer, 2010, Hommen et al., 2016b, Dohmen et al., 2016). As both methods improve, the prospect of using them in tandem to improve our understanding of population level risks of pesticide use is highly promising.

1.2.6 Knowledge gaps and next steps

Many limitations of the current ecological risk assessment of pesticides have been identified and it has already been demonstrated, at least in principle, that modelling approaches can address many of them. TK-TD modelling is capable of accounting for many of the temporal differences between exposure in the laboratory and those expected in the field. It can also relate exposure to toxic effects in a more direct way than summary statistics provided by laboratory tests designed to protect human health. While bodyweight is an important endpoint in mammalian chronic toxicity testing, TK-TD models able to predict toxic effects on mammalian growth are in their infancy. Such capability would significantly aid interpretation of dietary toxicity data and help to relate findings to the exposure of wild mammals in the agricultural landscape. Numerous hurdles remain but these are not insurmountable given recent improvements in modelling methods.

DEBkiss has shown good fits to growth and reproduction of American mink under toxic stress (Desforges et al., 2017). However, due to limited data, many parameters were fitted or assigned default values. Our study on the other hand will use raw data from standard laboratory studies, provided by Syngenta, which will provide individual measurements of dosing, body weight, and feeding rate throughout the test period and data on recovery after exposure ends (OECD, 1998, OECD, 2008). This will present an unparalleled opportunity to test predictions and theories. In particular, detailed feeding data will facilitate the generation of high-resolution inputs representing feeding rate and pesticide ingestion, giving new insight into laboratory data.

This project aims initially to develop a generic TK-TD growth model for rodents that is suitable for use in pesticide ecological risk assessment and able to address some of the limitations of laboratory testing. Using raw data from dietary toxicity studies on 6 pesticides, we will develop a framework for model calibration and validation. We will identify any remaining issues and suggest how they may be addressed in future work.

1.3 Objective 2: Use TK-TD growth model to facilitate *in vitro – in vivo* extrapolation of sublethal toxic effects on rodents

1.3.1 Animal testing and the three Rs

As has been discussed, animal testing plays a major role in the risk assessment of pesticides. This is also true for many other chemicals used in industry or available to the general public, such as clothing dyes, adhesives or cleaning products, which may pose a risk to human health or the environment (ICCA, 2011). Despite its importance to product safety, the use of animal testing has long been the subject of intense scrutiny due to the surrounding ethical concerns. In particular, animal testing of beauty products has been the subject of great opposition and is now banned in the EU. From 2013 even the sale of cosmetic products tested on animals anywhere in the world

was prohibited in the EU (European Commission, 2013a). This is part of a general drive to minimise the practice even where it can be more easily justified.

The UK government is committed to the replacement, reduction and refinement of animal testing (the three R's) (Home Office, 2017) and there is demand for the development of alternative methods from across industries and scientific disciplines. This is because, in addition to ethical issues, animal testing is slow, financially costly and limited in its predictive ability. As an example, the cost of advancing a new pharmaceutical drug to phase 1 human trials may reach $100 million. Failures at this stage are hugely expensive and the differences in kinetics between humans and animal test species account for many of these (Tsaioun et al., 2016). To account for these interspecies differences, PBTK modelling has become a crucial tool in the pharmaceutical industry (Zhuang and Lu, 2016). Meanwhile, the average research and development costs of a new pesticide are $286 million (Phillips McDougall, 2016). The sheer number of animals required is a large part of what makes animal testing so costly and contentious. Across the EU, 28.8 million animals were used in experiments from 2015-2017 (European Commission, 2019) while the most recent figures show 3.4 million animals were used in the UK in 2019 alone (Home Office, 2020). In anticipation of the EU ban on animal tested cosmetics, an expert panel was convened in 2010 by the European Commission to review the capabilities of *in vitro* methods (Adler et al., 2011). The report concluded that the integration of *in vitro* and *in silico* (virtual modelling) methods is essential for quantitative *in vitro* to *in vivo* extrapolation (QIVIVE).

1.3.2 Obstacles to quantitative in vitro to in vivo extrapolation

Even without the use of modelling, toxicity responses *in vivo* can, to a degree, be predicted from *in vitro* results. For example, literature reviews have found good correlations between the reference doses in fish and cultured fish cells (Castano et al., 2003, Schirmer, 2006). Such correlations cannot always be relied upon though, *in vitro* assays generally show low absolute sensitivity and there may be significant differences in behaviour of different assays. One possible reason for this is that in cultures of a single cell type there will be a relatively low number of target sites for toxicants to act upon. Another is that cells will not function in the same way *in vitro* as they would *in vivo*, so transport proteins, enzymes and receptors may be present at different levels (Schirmer, 2006, Groothuis et al., 2015). More complex assays can be constructed to compensate for these differences however they are more complicated to work with and so reduce throughput (Astashkina et al., 2012).

An additional contributor to low sensitivity and inter-assay variation is inappropriate choice of dose metric. Like in toxicity testing with live animals, where administered dose is most commonly linked to effects, the dose metric chosen to predict effects *in vitro* is often not the most suitable.

The commonly used nominal concentration refers to the amount of toxicant added to the medium, divided by the volume of the medium. *In vitro*, a chemical can meet a number of fates including metabolism, binding to proteins in the medium, adsorption to the plastic container, evaporation if volatile or precipitation if insoluble (Tsaioun et al., 2016). The composition of the culture medium differs between cell lines and so toxicant's affinity for proteins in the medium will vary too. Naturally, this can result in contrasting results between assays with the same nominal concentration because the toxicant is present in different concentrations at its target site. The toxicant concentration within the cells would therefore be the most appropriate metric. Even if intracellular measurements are not feasible there are external measures that are preferable to the nominal concentration. For example, the free concentration in the medium, that which is unbound to medium proteins and therefore available for uptake by cells (Groothuis et al., 2015)

1.3.3 *In vitro* **TK modelling**

Once again, TK-TD modelling can be applied to this problem. It has already been discussed how TK-TD modelling is able to relate toxicant exposure to internal concentration and resulting effects at the organism level. This alone however is insufficient to circumvent animal testing. A vital missing component is the ability to make predictions at the cellular level within an *in vitro* environment. In order for QIVIVE to be successful, a thorough understanding of TK *both in vitro* and *in vivo* is vital. Specialised *in vitro* TK models are therefore required to predict the intracellular toxicant concentration that results from a given nominal concentration (Tsaioun et al., 2016, Hamon et al., 2015).

The *in vitro* TK models developed by Wilmes et al. (2013) and Stadnicka-Michalak et al. (2014) both take a similar approach, although the latter is more simply formulated. Both consider the relative concentrations of the toxicant in 3 compartments, namely, the culture medium, the cells and the plastic to which the toxicant may bind (Hamon et al., 2015, Wilmes et al., 2013). In the case of volatile compounds the airspace in the container may be added as a fourth compartment (Stadnicka-Michalak et al., 2014). With a known initial pesticide concentration in the medium, the intracellular concentration at equilibrium can then be modelled.

Due to differences in cell properties and composition of the culture medium, *in vitro* TK models must be calibrated to the specific cell line and the chemical being tested. One approach is to dissolve known quantities of radiolabelled pesticides in the medium and use liquid scintillation counting to quantify the amount of the sample in the medium, absorbed by cells and adsorbed to plastic over time (Stadnicka-Michalak et al., 2014). Model parameters can then be fitted to toxicant concentration in each compartment over time. Repeating this process for enough different chemicals with a range of chemical properties may demonstrate trends in model

parameter values. These values may eventually be reliably predicted from chemical properties, such as the octanol water coefficient, without the need for experimental work.

1.3.4 Reverse Dosing

TK models can be used to predict intracellular toxicant concentration at equilibrium based on initial concentration in the culture medium. This is not necessarily a one-way process though; TK models can also be used in reverse to derive the initial conditions required to result in a desired intracellular concentration. This is known as reverse dosing and is a crucial step toward QIVIVE. Reverse dosing may be performed using *in vivo* or *in vitro* TK models depending on the aim of the study.

Some reverse dosing studies have been conducted with the laboratory rat as their subject, attempting to predict *in vivo* reference doses based on *in vitro* results. For example, Louisse *et al.* (2015) and Li *et al.* (2017) have adopted this approach with increasing success to predict effects of a pharmaceutical and a fungicide respectively. In both studies *in vitro* tests were used to determine the intracellular concentration at which 10% of rat stem cells failed to differentiate. This was taken as a proxy for the *in vivo* benchmark dose at which developmental malformations occur in 10% of individuals (BMD10). A PBTK model for the laboratory rat was then used in reverse, to calculate the oral dose at which the same concentration would be reached in the blood. The predicted doses were roughly one sixth (Louisse et al.) and one third (Li et al.) of the BMD10 published *in vivo* studies. While this approach is certainly relevant to current developmental toxicity testing, the relevance of time dependent reference doses like BMD10 to ecological risk is questionable (Hart and Thompson, 2005, Jager, 2011, Jager, 2012). Regardless, the reference doses used in chronic toxicity studies, such as NOAEL, are not dependent on the proportion of subjects showing a response but depend on the level of response. As such, it would be more beneficial to extrapolate the size of the effect resulting from a given level and duration of exposure.

Stadnicka-Michalak *et al.* (2015) used *in silico* and *in vitro* methods to predict the effects of two fungicides on fish growth. *In vivo* data came from fish early life stage (FELS) tests on two species, the fathead minnow and the rainbow trout. An established PBTK model (Nichols et al., 1991, Nichols et al., 1990, Stadnicka et al., 2012) was used to predict toxicant concentrations in the gills of fish in FELS tests. Reverse dosing was then used to determine the experimental conditions required to achieve the same intracellular concentration in cultured rainbow trout gill cells. Cell population growth under these conditions was then observed over a few days and extrapolated through time using the Von Bertalanffy growth curve. Over the same durations as the FELS tests, the predicted growth inhibition (relative to controls) matched closely with *in vivo* observations.

These results demonstrate the potential of *in vitro* testing and reverse dosing as an alternative to animal testing able to predict toxic effects *in vivo*.

1.3.5 Knowledge gaps and next steps

As yet, the approach used to extrapolate toxic effects on the growth of fish from *in vitro* results (Stadnicka-Michalak et al., 2015) has not been adapted for rodents and this presents a great opportunity to contribute to the three Rs. According to EU figures, around 1.12 million fish were used in experiments in 2017. However, this was only 12% of the 9.39 million animals used in total, rats and mice accounted for 73% (European Commission, 2019). In the UK, over 224,000 rats and mice were used for regulatory testing, with one third of regulatory procedures concerning chemical toxicity and safety (Home Office, 2020). Since long term animal testing uses the most animals and is most costly (Hartung and Rovida, 2009), it is logical to focus on endpoints measured in chronic toxicity testing like bodyweight.

In vitro – in vivo extrapolation of rodent growth under toxic stress will involve addressing several challenges and knowledge gaps. The first is the development of the TK-TD growth model itself, which was discussed in the previous section. Even with a TK-TD model, the procedure developed to extrapolate effects on the growth of fish will not be simple to replicate. For example, relating intracellular *in vitro* and *in vivo* intracellular concentrations to one another presents a greater challenge when considering rodents. A major difference between the toxicity testing frameworks for fish and mammals is the dosing method. While fish in regulatory tests are exposed to constant concentrations of pesticides in their water (OECD, 2014), mammals are dosed via the diet in longer studies (OECD, 1998, OECD, 2008, OECD, 2001). Therefore, the internal pesticide concentration reached depends on feeding rate relative to body size, which fluctuates substantially as animals grow (Laaksonen et al., 2013). Moreover, dosing may impact food consumption which will in turn alter the daily ingested dose and contribute to observed effects on body weight over time. These issues present major challenges for any comparison of *in vitro* and *in vivo* data. A further challenge that has not yet been covered is the identification of a cell line suitable for use as an assay for rodents. The specific gill cell line used by Stadnicka-Michalak et al. (2015) is derived from gill cells of healthy fish. Most mammalian cell lines tend to be transformed and/or cancer derived and so are less representative of the organism.

This project will explore the role that TK-TD modelling can play in facilitating *in vitro – in vivo* extrapolation of toxic effects on rodents and so helping to achieve the three Rs. Our TK-TD growth model will be used to simulate the effects of constant internal exposure to pesticides on rodent growth *in vivo*. This will provide a scenario that is comparable to cultured cells exposed to a constant pesticide concentration via their medium. *In vitro* TK modelling will then be used to

determine the experimental conditions required to match *in vivo* intracellular concentration. The results of these experiments will then be analysed to determine whether they indicate the level of growth inhibition predicted *in vivo*.

1.4 Thesis summary

The rest of this thesis is separated into five chapters. Chapters 2-5 each represent a standalone article or research paper. Chapter 2 presents a side-by-side comparison of the MTE and DEBkiss growth models in their simplest forms, allowing the most fundamental differences in their theory and assumptions to be identified and scrutinised. This exercise was initially conducted in order to determine which modelling framework was most suitable for use in this project. This formed the basis of a short article, providing an entry point for beginners to this subject, which was published in Ecological Modelling.

Chapter 3 is a research paper describing the development of a TK-TD growth model using data from dietary toxicity studies on rats. This was among the very first attempts to model sublethal toxicity in mammals. Additionally, having access to unpublished raw data meant that this study provided a novel implementation of existing modelling approaches, with high resolution model inputs representing pesticide intake and food consumption over time. This was published as a research article in Chemical Research in Toxicology.

In Chapter 4 is a research paper in which the potential of TK-TD models to act as a bridge between *in vitro* and *in vivo* data is investigated. The TK-TD growth model developed in the previous chapter was used to model the effects of constant internal concentrations of five pesticides. *In vitro* TK models were calibrated using experimental data and then used to design cell proliferation experiments. The results were analysed to determine any relationships between cell population growth and *in vivo* growth rate under chemical stress.

Chapter 5 is a methodological study focussing on how feeding data are converted into model inputs. This was conducted to address remaining issues with the growth model developed in Chapter 3. A novel method for deriving feeding inputs was proposed as an alternative to that used in Chapter 3 or the method conventionally used in DEB models. Models using each of the three methods were assessed in terms of accuracy, generality, and biological realism.

Finally, Chapter 6 draws conclusions from the thesis as a whole; reflecting on how successfully project objectives were met and identifying areas for future work.

Chapter 2 - Common ground between growth models of rival theories: a useful illustration for beginners

2.1 Preface

One of the first tasks I faced in this project was the identification of suitable modelling techniques. With respect to growth modelling, I considered methods based on two metabolic theories: the metabolic theory of ecology (MTE), and dynamic energy budget (DEB) theory. These theories represent two different schools of thought within the metabolic modelling community, and several papers have been written debating the merits of one over the other. I researched the growth models of each theory and found that comparing them in their simplest forms highlighted both the considerable overlap and the fundamental differences between them. I felt that such a comparison would be very useful to any newcomers to the subject area but was missing from the literature. This exercise was written up and subsequently published as a discussion article in Ecological Modelling. This article is presented below.

2.2 Abstract

Dynamic energy budget theory (DEB) and the metabolic theory of ecology (MTE) both seek to quantify the processes of resource acquisition and allocation but differ in their underlying mechanisms and assumptions. Some in-depth comparisons of the theories have been conducted in the literature but require a level of knowledge that is likely to be beyond most newcomers to the topic.

We reduce the theories to their simplest forms, their models for growth under optimal conditions, and present a side-by-side comparison of the model equations and key assumptions. This shows considerable overlap in how both theories characterise growth rate while also highlighting fundamental differences, such as MTE's use of taxon specific parameters. Comparing DEB and MTE in this way provides an accessible platform to help beginners gain a better understanding from the existing literature.

2.3 Introduction

DEB (Kooijman, 2000) and MTE (Brown et al., 2004b) are both well-known theories which aim to quantify the processes of acquisition and use of resources, to explain biological patterns. These theories differ fundamentally in their underlying assumptions and mechanisms. DEB assumes that resource assimilation scales with the surface area over which resources are absorbed into the body. This leads to a ⅔ scaling relationship between energetic supply and body mass within a species during isometric growth, in which body shape remains constant (Kooijman, 2000). MTE meanwhile, proposes that fractally branched vascular networks, possessed by plants and many

animals, deliver resources to cells at a rate which scales with body mass to the ¾ power (Brown and Sibly, 2012, Brown et al., 2004b). This scaling relationship and its underlying mechanism faced strong biological and mathematical criticism (Kozlowski and Konarzewski, 2004), however it was contended that this criticism was based on a misunderstanding of the theory's assumptions (Brown et al., 2005). In any case, neither mechanism precludes the other (Maino et al., 2014). It is possible, perhaps likely, that vascular supply networks and the surface area to volume ratio both play a role in limiting energetic supply rates.

The two theories have rarely been directly compared (Maino et al., 2014), leading to a lack of understanding across them and researchers often working only with one or the other (Kearney and White, 2012). This situation is somewhat at odds with the vision of MTE's creators that the theories should be complementary since they have different strengths and applications (Brown et al., 2004a).

Those comparisons that have been carried out focus on theoretical differences and inconsistencies (Marquet et al., 2014, Kearney and White, 2012, Van Der Meer, 2006b), inspiring spirited debate between proponents of the theories (Kearney et al., 2015, Houlahan et al., 2015, Marquet et al., 2015). The conversation, however, is often inaccessible for newcomers to the topic. We suggest that a different approach may be beneficial in this regard, focusing first on the similarity between the theories. The theories are most similar to one another when used to predict growth under constant conditions. We therefore use their most basic growth model as a starting point to clearly highlight the similarities and then isolate the core differences between them.

2.4 Common Ground

Both theories seek to mechanistically model the processes involved in metabolism. Research on DEB began around 1980 (Kooijman, 2000) while the precursor to MTE was published in the late 1990s (West et al., 1997). The growth models of both theories produce a sigmoid curve, as is observed in many species, and follow very similar biological rationale. The central equation of MTE states that metabolic rate (B) scales with total body mass (w) (Brown and Sibly, 2012):

$$B = B_0 w^{\alpha} \tag{2.1}$$

Where B_0 is the metabolic scaling coefficient and α is the exponent (usually ¾) with which metabolic rate scales allometrically (all parameters and dimensions are given in table 2.1). Assuming constant temperature and ad libitum food availability, the resulting growth model is expressed as:

$$\frac{dw}{dt} = \frac{1}{E_m}(B_0 w^\alpha - B_m w)$$
(2.2)

Where B_m is the maintenance rate per unit mass and E_m represents the energetic cost per unit growth (Kerkhoff, 2012). The energetic supply term in this model is the metabolic rate, defined in equation 2.1, since this represents energy that must be supplied for survival and growth.

Unlike that of MTE, the DEB growth model explicitly divides biomass into reserve (stored resources) and structure (functional body mass) (Kooijman, 2000). Therefore DEBkiss (Jager et al., 2013), a simplified version of DEB which does not consider reserve, is a better choice to illustrate the parallels with MTE. Under the same constant conditions, the basic growth model is:

$$\frac{dw}{dt} = y_{VA}(k J^a{}_{AM} w^{\frac{2}{3}} - J^v{}_M w)$$
(2.3)

Where y_{VA} represents conversion efficiency of assimilates to structure, k is the portion of assimilates allocated to structural maintenance and growth, J^a_{Am} is the maximum assimilation rate per unit area and J^V_M the maintenance rate per unit mass. It is clear at this point that both models can be characterised in the same way, illustrated in figure 2.1.
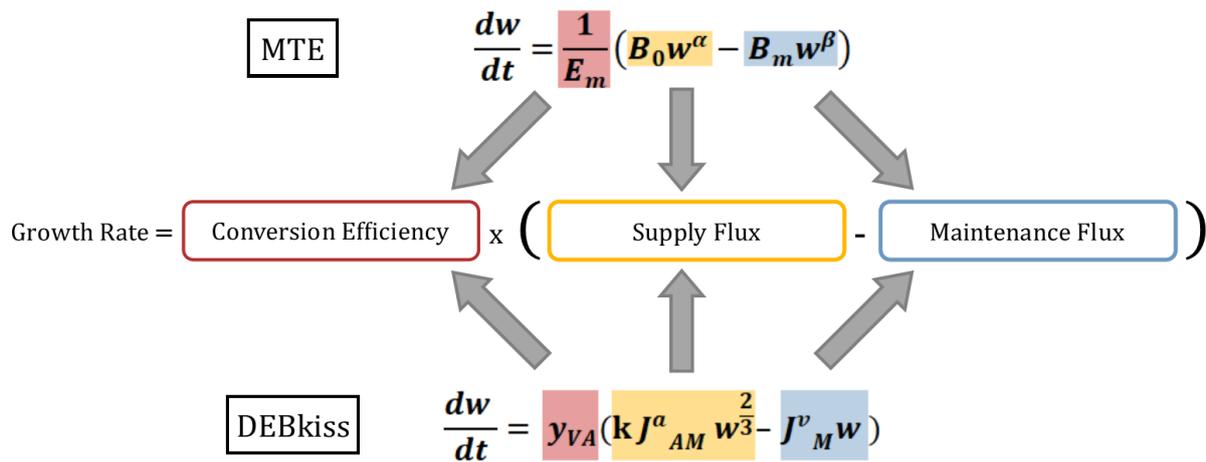
Figure 2.1 *The same general formula for conversion of resources to biomass is followed by both the MTE and DEBkiss growth models.*

Both models describe the conversion of resources into body mass, though their dimensions differ. While MTE considers the energy content of assimilates from food, DEBkiss considers their mass since food consumption is typically measured in terms of mass rather than energy. However, DEBkiss assumes constant composition of structure (and therefore a constant conversion factor between mass and energy) so both theories ultimately consider energy fluxes, regardless of the

units. This difference is somewhat trivial then, but important to note before comparing the model parameters.

A full description of the model parameters can be found in table 2.1. Except for k and α (both dimensionless), parameters relating to resource supply, maintenance or conversion efficiency have equivalent dimensions, differing only in how resources are measured.

Table 2.1 *List of parameters used in each growth model and their dimensions (b.m. = body mass). Thick borders divide parameters relating to (from top to bottom) conversion efficiency, resource supply and maintenance costs.*

Parameter	Theory	Explanation	Dimensions
y_{VA}	DEBkiss	New biomass per unit of resources used	$\text{Mass}_{(b.m.)} \cdot \text{Mass}_{(resources)}^{-1}$
$\frac{1}{E_m}$	MTE	New biomass per unit of energy used	$\text{Mass}_{(b.m.)} \cdot \text{Energy}_{(resources)}^{-1}$
k	DEBkiss	Proportion of resources allocated to growth & maintenance	Dimensionless
J_{Am}^a	DEBkiss	Maximum assimilation rate per unit of surface area	$\text{Mass}_{(resources)} \cdot \text{Mass}_{(b.m.)}^{-2/3} \cdot \text{Time}^{-1}$
α	MTE	Scaling exponent of energy supply with mass	Dimensionless
B_0	MTE	Supply rate per unit of biomass $^\alpha$	$\text{Energy}_{(resources)} \cdot \text{Mass}_{(b.m.)}^{-\alpha} \cdot \text{Time}^{-1}$
J_M^V	DEBkiss	Maintenance rate per unit of biomass	$\text{Mass}_{(resources)} \cdot \text{Mass}_{(b.m.)}^{-1} \cdot \text{Time}^{-1}$
B_m	MTE	Maintenance rate per unit of biomass	$\text{Energy}_{(resources)} \cdot \text{Mass}_{(b.m.)}^{-1} \cdot \text{Time}^{-1}$

Much discourse has concerned whether metabolic rate scales with body size to the power of ⅔ or ¾ (Brown and Sibly, 2012, White and Kearney, 2014, White and Seymour, 2003). While MTE suggests an α value of ¾, this is not fixed (Kerkhoff, 2012, Moses et al., 2008, Brown and Sibly,

2012). It has been noted that setting α to ⅔ actually makes little difference to the growth curve produced (Kerkhoff, 2012) and may be considered preferable in terms of mathematical simplicity (Sibly et al., 2013). This means that the ultimate body mass (W_∞) can be expressed:

$$W_\infty = \left(\frac{B_0}{B_m}\right)^3 \tag{2.4}$$

Substituting this into equation 2.2 means that the growth rate can be expressed in the form of the Von Bertalanffy growth equation:

$$\frac{dw}{dt} = \frac{1}{E_m} B_m w\left[\left(\frac{W_\infty}{w}\right)^{\frac{1}{3}} - 1\right] \tag{2.5}$$

In DEBkiss, the ultimate body mass is calculated as:

$$W_\infty = \left(k\frac{J^a_{AM}}{J^v_M}\right)^3 \tag{2.6}$$

By combining equations 2.3 and 2.6 the DEBkiss growth model can also be expressed in the form:

$$\frac{dw}{dt} = y_{VA} J^v_M w\left[\left(\frac{W_\infty}{w}\right)^{\frac{1}{3}} - 1\right] \tag{2.7}$$

Equations 2.5 and 2.7 are mathematically and conceptually equivalent, further demonstrating the considerable overlap between the two growth models. Parameter values could therefore be transferred between the two models to produce the same growth curve. However, their values would not be consistent with the estimation methods of both theories.

2.5 What Differences Remain?

The growth models may have a great deal in common but, even at this most basic level, important differences remain. The parameters of DEB are generally species-specific and full parameter estimation requires growth and reproduction data, preferably under varying conditions (Kooijman et al., 2008, Lika et al., 2011). A growing, curated database containing DEB parameters for almost 2000 species has been made freely available (Marques et al., 2018, Kooijman et al., 2019) as has software to carry out parameter estimation (Martin et al., 2012). Nevertheless, an understanding of the underlying equations and how parameter values are derived still requires significant investment of time and effort (Jager et al., 2013). One of the goals of MTE was to make this process simpler and less data demanding (Brown and Sibly, 2012), but the devil is in the detail.

It was initially suggested that the cost of growth in MTE, E_m, should be measured as the energy content of tissue (West et al., 2001), however such a simple measure negates any additional

'overhead' energy cost of growth. New estimation methods were subsequently developed by Moses et al. (2008) using the growth curves of animals in early life stages, when it is assumed that maintenance costs are negligible as virtually all energy is devoted to growth. Such data requirements are not dissimilar to those of DEB (Kooijman et al., 2008), so the energetic cost of tissue synthesis remains a property that is not easily measured or estimated. This could potentially be measured *in vitro* using assays measuring energy metabolism (Zhang et al., 2012) in combination with cell proliferation assays. However, to simplify the process, generic values of E_m are recommended for taxa such as mammals, birds and fish, on the basis that alternative methods (Sibly and Calow, 1986b, Moses et al., 2008) have produced similar estimates for these groups (Sibly et al., 2013).

As a taxon specific measure, MTE's metabolic scaling coefficient, B_0, is derived by fitting equation 1 to data on the metabolic rates (measured as respiration) of related species ranging in size (Brown and Sibly, 2012). As such the value of B_0 does not scale with the ultimate size of a species. For some this is an oversimplification as it assumes that variation in ultimate mass between species results solely from interspecific differences in maintenance rate, B_m (equation 2.4). By extension, this means that "the maintenance costs of a lizard are much higher than those of a baby crocodile of the same size" (Van Der Meer, 2006b). Instead, DEB suggests that the maintenance rate per unit of structure is probably very similar among related species. Therefore, differences in the ultimate mass of species are influenced more by interspecific variability in assimilation rate (Van Der Meer, 2006b). In principle, this debate could be settled by measuring the growth and caloric intake of related species. The energy requirements of an adult lizard vs a growing crocodile of the same size would reflect their respective volume specific maintenance costs.

Finally, the kappa rule in DEB states that, for any species, a constant proportion, k (ranging from 0-1), of available resources are allocated to maintenance and growth. The remainder ($1-k$) is invested toward maturation and reproduction (Kooijman, 2000). An equivalent parameter would be redundant in the MTE growth model, as the supply term describes only the energy used by cells for growth and maintenance rather than the total assimilated. As a wider theory MTE does have a similar rule, assuming that resources are allocated in fixed proportions to maintenance, growth and reproduction. However, unlike in DEB, it is also assumed that these proportions are constant across all species (Sibly, 2012), a simplification which has major implications. As a species-specific parameter in DEB, k plays an important role in capturing interspecific variability in ultimate mass (equation 2.6) and reproductive output when this is included model predictions. Estimates of k for different species vary almost across the full range possible, from 0.0388 for the Humboldt squid (Kooijman, 2018) to 0.996 for the frilled anemone (Kooijman, 2020).

2.6 Conclusions

Ultimately all models exist to provide a simplified representation of complex processes and this requires compromise between realism and ease of use. DEB provides a powerful framework for modelling organism life cycles but the data requirements for full parameterisation are substantial, as is the investment needed to fully understand model equations (Jager et al., 2013). MTE aimed to simplify further by utilising interspecific patterns to derive taxon-specific parameters. The only species-specific parameter in MTE's growth model (the maintenance rate B_m) can be calculated from the ultimate weight of a new species of interest. However, subsequent work has detailed how DEB parameters can also be estimated using only the ultimate size of a species if necessary, by scaling the parameter values for a related species (Kooijman et al., 2008). In both cases though, such simplicity comes at a cost in terms of accuracy since parameter values are, in reality, individual specific and vary within and between species.

While our comparison has focused on growth, the purpose of these theories is not simply to mimic observed patterns (for which the Von Bertalanffy equation would suffice) but to provide a mechanistic basis for predictions in novel scenarios and at higher levels of biological organisation. Implementation within individual based models (IBMs) of populations has highlighted the scope for model development in both theories. MTE successfully predicted observed patterns in the mass density of tree seedlings provided there was little competition for nutrients and water but not when these resources were limited (Lin et al., 2014). DEB meanwhile has been used to successfully predict population level patterns observed in *Daphnia magna* but it was found that predictions were just as accurate if reserve was omitted (Martin et al., 2013). This finding inspired the development of the DEBkiss model (Jager et al., 2013) which, as we have demonstrated, offers comparable simplicity to MTE. In addition, the DEBtox framework (Kooijman and Bedaux, 1996b) uses DEB as the basis for predicting the effects of toxic exposure on biological processes. This has now been used to predict observed effects in several invertebrate (Ashauer and Jager, 2018) and vertebrate species (Desforges et al., 2017, Zimmer et al., 2018) demonstrating the ability of DEB to realistically respond to environmental change.

The aim of this article was to compare DEB and MTE in a way that is targeted to beginners and currently missing in the literature. While DEB is indeed a complex theoretical framework, since the creation of MTE it has been developed to offer similar simplicity both in terms of model equations and parameter estimation. At this point, the distinction between the two theories hinges on the underlying mechanisms and assumptions. We hope that, by providing an accessible introduction to this ongoing discussion, this article helps newcomers to the topic to better understand the arguments and feel equipped make their own model choice with confidence.

Chapter 3 – Toxicokinetic-toxicodynamic modelling of the effects of pesticides on growth of *Rattus norvegicus*

3.1 Preface

Syngenta's involvement in this project presented a rare opportunity in terms of data availability. I was provided with the unpublished raw data from toxicokinetics and dietary toxicity studies on seven pesticides. Having identified methods to model growth, toxicokinetics and toxicodynamics, the challenge was then deciding how to best make use of such an abundance of data. I had met with toxicologists at Syngenta's offices and discussed typical patterns that often emerged in dietary toxicology studies. Through these discussions it became clear that feeding rate is a crucial factor influencing the growth rate and the ingested pesticide dose over time. Therefore, properly utilising food consumption data would be absolutely vital to model the effects of dietary toxicity. With such an abundance of data, another question was which data should be used to calibrate model parameters or to test model predictions. The process of model development with regulatory datasets formed a research paper which was published in Chemical Research in Toxicology and is presented below.

3.2 Abstract

Ecological risk assessment is carried out for chemicals such as pesticides before they are released into the environment. Such risk assessment currently relies on summary statistics gathered in standardised laboratory studies. However, these statistics extract only limited information and depend on duration of exposure. Their extrapolation to realistic ecological scenarios is inherently limited. Mechanistic effect models simulate the processes underlying toxicity and so have the potential to overcome these issues. Toxicokinetic-toxicodynamic (TK-TD) models predict the internal concentration of a chemical over time and the stress it places on an individual organism. TK-TD models are particularly suited to addressing the difference in exposure patterns between laboratory (constant) and field (variable) scenarios. Few studies have sought to predict sub-lethal effects of pesticide exposure to mammals in the field, although such effects are of particular interest with respect to longer term exposure. We developed a TK-TD model based on the dynamic energy budget (DEB) theory, which can be parameterised and tested solely using standard regulatory studies on rodents. We demonstrate that this approach is effective in predicting toxic effects on the body weight of rats over time, ≥75% of observations were predicted to within one standard deviation of the mean in 28 of 34 datasets. Model predictions separate the impacts of feeding avoidance and toxic action, highlighting which was the primary driver of effects on growth. Such information is relevant to the ecological risk posed by a compound because in the environment alternative food sources may or may not be available to

focal species. While this study focused on a single endpoint, growth, this approach could be expanded to include reproductive output. The framework developed is simple to use and could be of great utility for ecological and toxicological research as well as to risk assessors in industry and regulatory agencies.

3.3 Introduction

Before chemicals can be registered for use, they undergo ecological risk assessment (ERA), this process is particularly rigorous for agricultural pesticides, which are designed to be toxic to pest species (EFSA, 2009b, van Leeuwen and Vermeire, 2007). It is not practically possible to determine a chemical's ecological impact experimentally. Instead, risk assessment relies on extrapolation from summary statistics such as the 'no observed adverse effect level' (NOAEL) generated for a few species in standardised laboratory studies (Grimm and Martin, 2013). However, such statistics should only be extrapolated with caution as they do not account for the processes that lead to toxic effects and are dependent on duration of exposure (Jager, 2012, Murado and Prieto, 2013). Moreover, they are generated in controlled (supposedly optimal) conditions - regulated temperature and freely available food and water (OECD, 2001) - that are unrealistic in the field. The resulting lack of ecological realism in current standard risk assessment methods is problematic (European Commission, 2014).

Mechanistic effects models (MEMs) aim to simulate the mechanisms by which chemicals affect individuals, populations and communities and therefore enable us to predict how they will respond in untested and more ecologically relevant conditions (Grimm and Martin, 2013). This is an appealing prospect with great potential for use in ERA of pesticides (Forbes et al., 2009, Forbes and Calow, 2012). By focussing on the underlying processes, modelling techniques can add ecological realism to extrapolations and even reduce animal testing requirements (Jager et al., 2006).

Accounting for the mismatch in exposure between laboratory and field (Fischer, 2005) was identified as one of five key obstacles to long term risk assessment of pesticides for mammals (along with selection of suitable toxicity endpoints, extrapolation of toxicity between species, exposure assessment and evaluation of population level effects (Hart and Thompson, 2005)). In chronic toxicity tests rats or mice are exposed to a constant concentration of a pesticide in their diet for periods as long as 2 years (OECD, 2008, OECD, 1998, OECD, 2001). Such constant exposure is unrealistic in the field where pesticides are not applied at a constant rate all year round. This disparity can be addressed through the use of toxicokinetic-toxicodynamic (TK-TD) modelling (Jager et al., 2006). TK-TD models work at the individual level, predicting an internal measure of chemical concentration over time (toxicokinetics) and the stress this places on an

organism (toxicodynamics). As such, the effects on a given endpoint resulting from time varied exposure can be predicted (Jager, 2016, Nyman et al., 2012, Ashauer et al., 2017).

The use of TK-TD modelling has now been recommended for certain regulatory purposes, such as predicting survival of aquatic organisms (EFSA, 2018). However, European protection goals for birds and mammals state that there should be no visible mortality associated with pesticide use (EFSA, 2009b), so sublethal effects are more relevant with respect to realistic exposure. Sublethal effects can be predicted using the 'DEBtox' modelling framework (Kooijman and Bedaux, 1996b, Kooijman and Bedaux, 1996a), combining TK-TD modelling with the Dynamic Energy Budget (DEB) theory (Kooijman, 2000). DEB is an established metabolic theory which has been applied to a range of taxa (Marques et al., 2018), mathematically describing the processes of energy acquisition and allocation that determine the life history of an organism. Using TK-TD modelling to place stress on these processes can produce predictions of effects on sublethal endpoints such as growth and reproduction. Very little research has concerned mammals however, as DEBtox studies have thus far mainly focused on invertebrates (Ashauer and Jager, 2018) and more recently fish (Zimmer et al., 2018, Sadoul et al., 2018).

At present DEBtox is limited to research applications as it is not regarded as user friendly enough for use by regulators (EFSA, 2018). To this end, a simplified version of the theory, 'DEBkiss', was developed in which only structural body mass (bones, muscle, organs etc.) is considered with no reserve storage (Jager et al., 2013). The model retains many DEB principles but with fewer parameters and model equations. It was developed for applications where simplicity and ease of use are important, such as the analysis of toxicity data or for use within individual based population models. The only published study to date in which TK-TD modelling has been used to predict sublethal effects in mammals utilised DEBkiss (Desforges et al., 2017). Although limited data were available, the model accurately simulated observed effects of environmental toxicants on growth and reproduction in the American mink (*Mustela vison*). These results suggest that this simplified framework may be sufficient for practical and regulatory applications.

Here we tested the utility of DEBkiss by working with raw data from repeated dietary dose toxicity tests and modelled the effects of several pesticides on rats. As the first study to use regulatory data for this purpose, we adopted the practice of beginning with the simplest possible methods and identifying areas where more complex techniques may be required. Internal pesticide concentration was modelled with a one compartment model and a single endpoint, body weight, was modelled over time using the DEBkiss growth model. We only considered a single endpoint for simplicity as we aimed to establish a practical procedure for the parameterisation, calibration and validation of DEBtox models using regulatory data and to assess the quality and

utility of predictions. With an established procedure in place, other endpoints can be considered in future studies. Furthermore, we aimed to improve the interpretation of standard toxicity studies by extracting novel, meaningful information using modelling. In this regard, keeping the number of state variables in the model to a minimum provided greater clarity. The data used in this study only came from unmated animals, so reproduction was not omitted from the model, it simply did not occur during the observation period considered.

3.4 Methods

3.4.1 Data

All data used here were made available from existing regulatory studies (Syngenta, unpublished) carried out according to 94/79/EC (European Commission, 1994), investigating toxicokinetics and chronic toxicity of acibenzolar-S-methyl (benzothiadiazole; fungicide, insecticide and plant activator (PPDB, 2019a)), azoxystrobin (strobilurin; fungicide (PPDB, 2019b)), fenpropidin (unclassified; fungicide (PPDB, 2019c)), fludioxonil (phenylpyrrole; fungicide (PPDB, 2019d)), mandipropamid (mandelamide; fungicide (PPDB, 2019e)) and prosulfuron (sulfonylurea; herbicide (PPDB, 2019f)) and thiamethoxam (neonicotinoid; insecticide (PPDB, 2019g)) in laboratory rats (*Rattus norvegicus*).

The toxicokinetics studies followed OECD test no. 417 (2010) guidelines. Animals were treated with a single oral (gavage) dose of a ^{14}C radiolabelled pesticide with total radioactivity found in various tissues and excreta monitored over a period of days. The animals were allowed free access to a certified standard diet.

Data sets differed between pesticides but followed a common framework. At least two dose levels were studied with typically three male and three female animals in each treatment group. The reports include data detailing the proportion of the initial dose excreted in faeces, urine and bile over ~48 hours after a single oral dose, providing an average percentage of the dose which was absorbed into the body. Pesticide concentration in the blood of animals was measured over ~48 hours following a single high or low dose. Pesticide concentration was also measured in different body tissues from animals terminated at ~4 time points following a single high or low dose. Details of dosing, including exact dose ($mg_{(AI)}$), body weight (g) at the start (and in some cases the end) of testing and achieved dose ($mg_{(AI)} \times kg_{(BW)}^{-1}$) were provided for each individual animal.

Chronic toxicity studies lasting 28 days (OECD test no. 407 (2008)), 90 days (OECD test no. 408 (1998)) or 2 years (OECD test no. 416 (2001), 451 (2018a) or 452 (2018b)) were carried out according to OECD guidelines. Animals of around 5-7 weeks in age were provided with a diet containing pesticide and multiple toxicological endpoints monitored over the study period.

Animals were kept in standard conditions with food and water available ad libitum. Each study provides individual weekly observations of body weight (g) and food consumption ($g_{(diet)} \times day^{-1}$). Sample size was typically 5 animals per sex per treatment in 28 day studies, 10 animals per sex in 90 day studies and 50-80 animals per sex in 2 year studies. The measured concentration of pesticide in the diet of each treatment was also reported. Each study comprised a control group and at least 3 treatments fed diets containing different concentrations of pesticide. As study duration was increased dietary doses were generally decreased.

3.4.2 Toxicokinetic model

For each toxicant (denoted 'AI' for active ingredient), the internal concentration was modelled using a one compartment TK model with first order kinetics. The internal concentration here refers only to toxicant present in body tissues at a given time and excludes any in the gut which has not yet been absorbed. As terrestrial mammals are primarily exposed to pesticides via the diet, the toxicant concentration in the gut was also modelled as an intermediate 'depot' compartment. This was a very similar approach to Bednarska et $al.$ (2013a) but we also account for change in body size by including dilution by growth and changes to surface area to volume ratio as per Gergs et $al.$ (2016). The model equations are shown below.

$$\Delta C_{Gut} = I - {^{L_\infty}}\!/_{L}\, k_a C_{Gut} F - C_{Gut}({^{\Delta W}}\!/_{W}) \qquad\qquad (3.1)$$

$$\Delta C_{Int} = {^{L_\infty}}\!/_{L}\, k_a C_{Gut} F - {^{L_\infty}}\!/_{L}\, k_e C_{Int} - C_{Int}({^{\Delta W}}\!/_{W}) \qquad\qquad (3.2)$$

Where ΔC indicates change in the body weight (denoted 'BW') normalized dose, C, of toxicant over time ($Mass_{(AI)} \times Mass_{(BW)}^{-1} \times t^{-1}$) and subscripts Gut and Int denote gut and internal respectively; I is toxicant ingestion rate ($Mass_{(AI)} \times Mass_{(BW)}^{-1} \times t^{-1}$); F represents bioavailability (dimensionless); k_a and k_e represent the rate constants of toxicant absorption from the gut and toxicant elimination from the system respectively (t^{-1}); W denotes body weight and ΔW is change in weight over time; L is volumetric length (the cube root of body volume) and L_∞ is the ultimate volumetric length of the test species.

DEB states that volume, V, surface area, a, and volumetric length, L, scale such that $V = L^3$, $a = L^2$ and $a = V^{2/3}$. Thus, the surface area to volume ratio can be calculated as $V/a = L$, so L_∞/L gives an animal's surface area relative to that of a fully grown adult. Uptake and elimination are area mediated processes while internal concentration is determined relative to volume or weight. Multiplying the uptake and elimination terms by L_∞/L therefore accounts for changes to these rates that occur as animals grow. The final term of equations 3.1 and 3.2 accounts for dilution by growth, that is the change to concentration that occurs if there is no uptake or elimination but the

compartment changes in size. Egestion is not explicitly considered in this model. Instead, fitted elimination rate constants are assumed to account for elimination by egestion.

3.4.3 Growth model

Growth was modelled with the DEBkiss (Jager et al., 2013) growth equation:

$$\Delta W = y_{VA}(kf\,J^a{}_{AM}\,W^{2/3} - J^v{}_M W)\qquad\qquad\qquad\text{(3.3)}$$

Where ΔW represents the change in total body weight (W) over time, y_{VA} represents the efficiency with which assimilates are converted to structural mass (Mass$_{(BW)}$ × Mass$_{(Assimilates)}^{-1}$), k is the proportion of assimilates allocated to the soma, f is the scaled feeding rate (unitless), J^a_{Am} is the maximum surface area specific assimilation rate (Mass$_{(Assimilates)}$ × Mass$_{(BW)}^{-2/3}$ × t^{-1}) and J^V_M the mass specific maintenance rate (Mass$_{(Assimilates)}$ × Mass$_{(BW)}^{-1}$ × t^{-1}). The DEBkiss model is represented graphically in figure 3.1.

Endotherms are also subject to surface area specific maintenance costs, accounting for heat loss to the environment. However as long as the ambient temperature is within the thermoneutral zone of a species (Kingma et al., 2014) these are assumed to be zero as heating costs are at their minimum so are simply part of volume specific maintenance (Lika et al., 2011). Laboratory guidelines require rodents to be kept at 22±3°C, as this was considered to be within the thermoneutral zone of the rat (Poole and Stephenson, 1977). More recent research has suggested that this temperature range is below the thermoneutral zone of the rat (Le and Brown, 2008) but for simplicity it was assumed that heat loss could be omitted. Omitting heating costs would have minimal impact on model predictions themselves but would impact on the relative values of J^a_{Am} and J^V_M. However, any costs near the thermoneutral zone would be minimal.
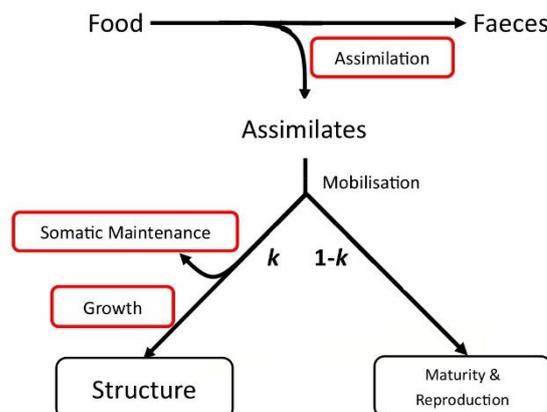
Figure 3.1 *A graphical representation of the DEBkiss model. The value of k determines the proportion of resources assimilated from food allocated to maintenance and growth or maturity and reproduction. Processes outlined in red are those that can be subjected to stress.*

The parameter *k* represents the proportion of assimilates allocated to maintenance and growth with the remainder (1-*k*) going towards maturation and reproduction. DEBkiss theory states that up to puberty, these resources are used up as animals develop and are stored as 'reproduction buffer' during puberty and adulthood (Jager et al., 2013). This buffer is generally used to represent mass for egg production which is not applicable to mammals. Additionally, as reproduction was not modelled in this study, the reproduction buffer would serve no purpose other than as a reserve for use under starvation. As there were no data for body length, there could be no distinction between growth and weight gain as fat, so all body weight was simply modelled as structure. This necessitated the assumption that 1-*k* branch continues to be used up as the animals develop into sexually mature adults (at 70- 90 days of age (Tacutu et al., 2018)) and then to maintain maturity, a process which can also be included in DEBkiss (Jager et al., 2013). Only un-mated animals were included in this study, so body mass was not impacted upon by pregnancy.

If food intake is reduced such that $kf\,J^a_{AM}\,W^{2/3} < J^v_M W$ but the total assimilation rate $f\,J^a_{AM}\,W^{2/3} \geq J^v_M W$ (i.e. in a situation where food intake is sufficient to maintain homeostasis but not to grow), then $\Delta W = 0$ as available resources are diverted from the 1-*k* branch to meet maintenance costs. If the total assimilation rate is insufficient to meet maintenance costs, that is $f\,J^a_{AM}\,W^{2/3} < J^v_M W$, then the growth rate becomes negative as tissue is metabolised to meet maintenance requirements.

$$\Delta W = (f\,J^a_{Am}W^{2/3} - J^V_M W)/y_{AV} \tag{3.4}$$

Where y_{AV} is conversion efficiency of structure to assimilates. The value of *k* therefore determines the point at which the feeding rate becomes insufficient for growth but does not impact the onset of weight loss. While it would be possible to model the effects of starvation on survival using the GUTS framework (Jager et al., 2011), there is a lack of data on the topic as the experiments required would be unethical. Any treatment that induced drastic reductions in feeding would be abandoned as guidelines state that dosing should not cause 'death or severe suffering' (OECD, 1998, OECD, 2008).

3.4.4 Toxicodynamic model

Finally, the DEBtox toxicodynamic model (Kooijman and Bedaux, 1996b, Kooijman and Bedaux, 1996a, Jager, 2015b) was used to link internal toxicant concentration to stress on growth. It is

assumed that for any xenobiotic, there exists a 'no effect concentration' (*NEC*) below which it causes no stress to any biological processes. The *NEC* is a time independent threshold and therefore has no relationship with duration of exposure. Assuming that every molecule of toxicant beyond its *NEC* creates the same amount of 'stress' leads to a 'linear with threshold' relationship, which can be modelled quite simply (figure 3.2):
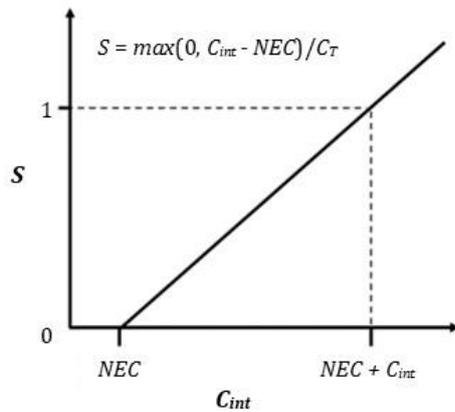
Figure 3.2 *Stress increases with internal toxicant concentration beyond a threshold. Where S is dimensionless stress and C_T is the 'tolerance concentration' ($Mass_{(AI)} \times Mass_{(BW)}^{-1}$). Here the NEC determines the point at which stress exceeds zero while C_T is the increase in C_{Int} corresponding to an increase in S of 1. This means the gradient of S is $1/C_T$ when C_{Int} exceeds the NEC.*

In our approach, stress can be applied to one of three growth parameters ('physiological modes of action' (Kooijman and Bedaux, 1996a, Alvarez et al., 2006)) each of which respond differently to stress, these are the maximum assimilation rate, J_{Am}^a, the maintenance rate, J_M^V, or the conversion efficiency, y_{VA} (figure 3.3). The proportion of resources allocated to the soma, k, could theoretically be affected by a toxicant, but data on reproduction would be required to distinguish this from effects on assimilation and such effects are not well documented (Jager, 2015b). Details of all TK-TD and growth model parameters are included in table 3.1.
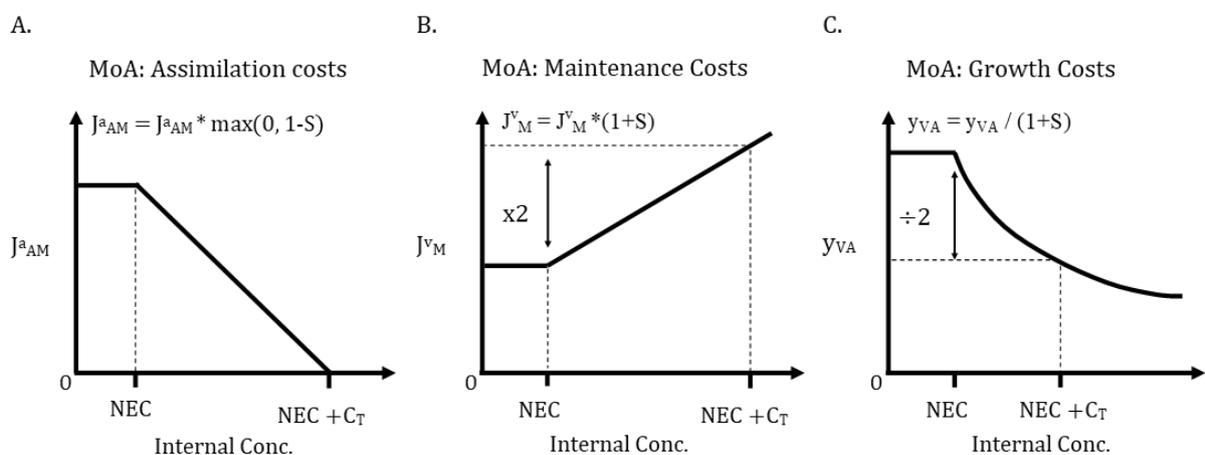
Figure 3.3 *Plots showing how growth model parameters respond to internal toxicant concentration when stress is applied. A. The maximum assimilation rate, J^a_{Am}, decreases linearly with stress until it reaches zero when S=1. B. The maintenance rate, J^V_M, increases linearly with stress and is doubled when S=1. C. The costs per unit of tissue synthesis increase linearly with stress, doubling when S=1. At this point conversion efficiency y_{VA} is halved as it approaches zero asymptotically.*

Table 3.1 *contains a full list of model parameters. Parameter values marked with a * are default values suggested by Jager, Martin & Zimmer (2013).*

Name	Explanation	Value	Dimensions
Growth			
k	Fraction of assimilates for growth and maintenance	0.8 *	-
J^a_{Am}	Maximum assimilation rate per unit of surface area	Fitted to data	$g_{(Assimilates)} \times g_{(BW)}^{-2/3} \times t^{-1}$
J^V_M	Maintenance rate per unit of biomass	Fitted to data	$g_{(Assimilates)} \times g_{(BW)}^{-1} \times t^{-1}$
y_{VA}	New biomass per unit assimilates	0.45 (as per Sibly and Calow (1986a))	$g_{(BW)} \times g_{(Assimilates)}^{-1}$
y_{AV}	Yield of assimilates per unit biomass	0.8 *	$g_{(Assimilates)} \times g_{(BW)}^{-1}$
Toxicokinetics			
F	Proportion absorbed from gut	Calculated from data	-
k_a	Absorption rate constant	Fitted to data	t^{-1}
k_e	Elimination rate constant	Fitted to data	t^{-1}
Toxicodynamics			
NEC	No effect concentration	Fitted to data	$mg_{(AI)} \times kg_{(BW)}^{-1}$
C_T	Tolerance Concentration	Fitted to data	$mg_{(AI)} \times kg_{(BW)}^{-1}$
Other Parameters			
f	Scaled feeding rate	Calculated from data	-
I	Pesticide ingestion rate	Calculated from data	$mg_{(AI)} \times kg_{(BW)}^{-1} \times t^{-1}$
L	Volumetric Length	$W^{1/3}$	cm
L_∞	Ultimate volumetric Length	$W_\infty^{1/3}$ or $kf\frac{J^a_{AM}}{J^v_M}$	cm
L_m	Maximum volumetric Length	$W_m^{1/3}$ or $k\frac{J^a_{AM}}{J^v_M}$	cm
W_∞	Ultimate structural body mass	782 (as per Hubert et al. (2000)) or L_∞^3	g
W_m	Maximum structural body mass	L_m^3	g
State Variables			

C_{Gut}	Pesticide concentration in gut (ΔC_{Gut})	$I - {L_\infty}/{L}\, k_a C_{Gut} F - C_{Gut}({\Delta W}/{W})$	$mg_{(AI)} \times kg_{(BW)}^{-1} \times t^{-1}$
C_{Int}	Internal pesticide concentration	$\int {L_\infty}/{L}\, k_a C_{Gut} F - {L_\infty}/{L}\, k_e C_{Int} - C_{Int}({\Delta W}/{W})$	$mg_{(AI)} \times kg_{(BW)}^{-1} \times t^{-1}$
W	Structural body mass (ΔW)	$y_{VA}(kf\, J^a_{AM}\, W^{2/3} - J^v_M W\,)$	$g_{(BW)} \times t^{-1}$

3.4.5 Model implementation

All models were implemented in Matlab (ver. R2016b). TK and TK-TD growth models were developed with the BYOM (Jager, 2019) flexible model platform (ver. 4.1), several additional functions and scripts were also developed as part of this study. All fitted parameter values were derived using the Nelder Mead simplex algorithm to maximise the likelihood function, given the observed data (Pan and Fang, 2002). Likelihood profiling was also used to check that initial fits were not local optima (Kreutz et al., 2013). TK model parameters were fitted to mean internal pesticide concentration over time, while growth and TD model parameters were fitted to mean body weight over time.

Toxicodynamic modelling also required selection of a physiological mode of action. The best fitting physiological mode of action was determined using the Akaike information criterion (Burthe et al., 2010). As the alternative models were not nested, the likelihood ratio test would be inappropriate however the AIC has no such requirement (Murtaugh, 2014).

3.4.6 TK Modelling

3.4.6.1 Parameterisation

The percentage of the dose which was absorbed into the body was reported in excretion studies. This data provided the value of the parameter F in the TK model (eqn. 1 & 2). If the percentage absorbed was very high (>90%) then F was assigned the value 1 as this represents the worst-case scenario. For single dose studies the value of I was zero while the starting gut concentration was the average achieved dose for each treatment. If body weight was recorded at the beginning and end of testing then ΔW was calculated as the linear growth rate observed in each treatment. If not, it was assumed that $\Delta W = 0$ as dilution by growth has a minimal effect on the model over the short testing period, typically 48 hours.

Ultimate length was calculated as the cube root of 782cm³ which is the average ultimate volume of male Sprague Dawley rats with ad libitum food availability (Hubert et al., 2000), assuming that average wet tissue density is equal to that of water (Lika et al., 2011).

3.4.6.2 Calibration

Having determined the other model inputs directly from experimental data, two free TK model parameters were left to be fitted to data, the absorption and elimination rate constants k_a and k_e. The best time course data for internal toxicant concentration (highest number of time points) came from the blood as it can be sampled without terminating animals. In order to ascertain whether pesticide concentration in the blood was a suitable proxy for overall body burden, the Pearson correlation coefficient was used to determine if it was significantly correlated with the concentration in all other sampled tissues. If so, then blood concentration data (in $\mu g \times g_{(blood)}^{-1}$) were deemed representative of C_{Int} (in $mg \times kg_{(BW)}^{-1}$) and could be used to fit ΔC_{Int} (eqn. 2). Where available, whole carcass concentration could be used as an alternative however, fewer data points were available.

Data were first inspected to determine whether there were obvious differences in kinetics between males and females. Properties considered were the mean peak concentration (C_{max}), the time after dosing at which it was reached (T_{max}) and the time taken to eliminate the pesticide from the blood. If clear differences were evident then models were calibrated separately for males and females.

For some pesticides, the same individual animals were sampled for the whole observation period after dosing, for others only partial time course data were available for each individual. Where complete time course data were available, the model was first fitted to data for each individual. A multi-way ANOVA was conducted to determine whether there were significant differences in fitted values of k_a and k_e associated with sex, dose level (both discrete) or weight (continuous). If more than one radiolabel was used then this was also included as a factor. This was to determine whether data needed to be separated for calibration or the dataset could be used as a whole, which would be preferable in terms of sample size.

If sex was the only factor to have a significant effect, then the model was calibrated to males and females separately. If both sex and weight were shown to have significant effects, then a Mann Whitney U test was used to determine whether there was a significant difference in weight between males and females. If so, then the ANOVA was repeated with the data separated by sex. If no effects were found within each sex then the model was simply fitted to males and females separately. If the rate constants were significantly affected by factors other than sex then this was noted along with the full TK results.

Finally, models were calibrated to mean (at each time point) blood concentration observed in the high and low dose groups simultaneously. Where appropriate, this was carried out separately for

males and females. Parameters were not fitted individually and then averaged as this was not possible for all compounds due to data availability.

3.4.6.3 Validation

As part of the toxicokinetics studies, pesticide concentration was measured different body tissues of animals terminated at different time points following a single oral dose. These tissues included the blood. Blood samples from these animals were arranged to provide an independent time course data set of internal toxicant concentration. These were then used to assess the model performance by comparing to the internal concentration predicted by the model with fixed parameters.

Generally, these data covered the same two dose levels as the calibration data set so could not provide validation per se, since the model inputs (i.e. dose levels) were virtually the same. Any differences in achieved doses were generally very small so differences between observed toxicokinetics were primarily due to individual variability. This did however provide an indication of how well the true average response was represented by the calibration data set and therefore the calibrated model.

In most cases blood concentration data sets included only three samples for each time point and responses could be highly variable between individuals. We did not carry out any quantitative assessment of predictions because this would be misleading due to the low sample size combined with strong inter-individual variability

3.4.7 TK-TD Growth Modelling

3.4.7.1 Parameterisation

The full TK-TD growth model comprises all the model equations (eqns. 3.1, 3.2 & 3.3) and simultaneously predicts toxicant concentration in the gut (mg$_{(AI)}$ × kg$_{(BW)}$$^{-1}$), internal toxicant concentration (mg$_{(AI)}$×kg$_{(BW)}$$^{-1}$) and body weight (g).

The growth parameter k describes the proportion of resources allocated to maintenance and growth. When only modelling growth, its precise value is not crucial (only its product with the fitted parameter J_{Am}^a contributes to the model) so this was fixed at its default value of 0.8 (Jager et al., 2013) which estimates suggest is reasonable for the species (Kooijman, 2015, Rakel and Gergs, 2018). Physiological studies suggest assimilated energy is converted to new tissue by homeotherms with an efficiency between 0.4 and 0.5 (Sibly and Calow, 1986a) so the parameter y_{VA} was fixed at 0.45. If strong evidence were provided suggesting different values for either of these parameters, J_{Am}^a and J_M^V would simply need to be adjusted by the appropriate correction

factor. Stress functions would continue to have the same impact and so would not need adjustment.

In repeated dose toxicity tests each animal's weight and food consumption were recorded at least weekly for some or all of the study period (OECD, 2001, OECD, 1998, OECD, 2008). In some cases, food consumption was recorded per cage, so the values provided were an average per animal but each weight measurement had a corresponding measurement of food consumption. The achieved toxicant concentration in the diet of each treatment group was also measured. This allowed the growth parameter f (scaled feeding rate) and the TK parameter I (toxicant ingestion rate) to be calculated directly from the data.

The maximum feeding rate at a given food density is assumed to be proportional to surface area. The scaled feeding rate, f, is equal to an individual's actual feeding rate at a given food density divided by the maximum feeding rate for its size and therefore ranges from 0 to 1 (Jager et al., 2013).

Measured food consumption per day was converted into surface area specific feeding rate by dividing by the associated body weight raised to the power 2/3. Dividing these values by the maximum feeding rate recorded in the study group (separated by sex) provided scaled f values between 0 and 1 for each individual in each week of the study period. A matrix was then produced containing average weekly feeding rates for each treatment group, these provided the value of f for each treatment in each weekly interval.

Multiplying the achieved toxicant concentration in the diet ($mg_{(AI)} \times kg_{(diet)}^{-1}$) by the mass specific daily feeding rate ($kg_{(diet)} \times kg_{(BW)}^{-1} \times d^{-1}$) provides the ingested dose ($mg_{(AI)} \times kg_{(BW)}^{-1} \times d^{-1}$). Again, a matrix was produced, this time containing the weekly averages of daily ingested dose for each treatment. This provided the values of the toxicant ingestion rate I which fluctuates with feeding rate throughout the study period. All other TK parameters remain fixed at the values determined during TK model calibration (k_a and k_e were multiplied by 24 to convert them from hourly to daily rates).

3.4.7.2 Calibration

The data from 90 day toxicity studies (OECD 408) were intermediate in terms of sample size and dietary dose levels. Thus, these data were more representative than the 28 day studies (OECD 407) while the observed effects on growth were generally larger than in 2 year studies (OECD 416, 451 or 452). For this reason, the 90 day studies were used to calibrate the TK-TD model. Calibration of the growth and TD parameters was conducted separately for males and females.

The two free growth model parameters, J_{Am}^a and J_M^V, were fitted to the growth data from the control group and then fixed. Next, the model was run for all treatments but with no stress applied. This step was used to identify the lowest dose group in which observed growth was lower than predicted by the growth model based on feeding rate alone (i.e. the lowest dose group in which chemical stress occurred). The initial value for the estimation of the TD parameter *NEC* was then set to half the average internal concentration predicted by the TK model for that treatment. Only two treatments (the lowest affected treatment and the top dose) were used for fitting as this allowed more of the data to be used for testing predictions while still providing a wide range of internal concentration predictions.

The TD parameters *NEC* and C_T were then fitted to the lowest affected treatment and the top dose group simultaneously, this was repeated for each physiological mode of action. The fit which produced the lowest AIC value was selected. The physiological mode of action and resulting TD parameter values were then fixed.

3.4.7.3 Validation

For verification, the resulting model was then run for all treatments in the 90 day study producing interpolations to the intermediate dose groups. This was to show how the model could be validated with data from only one study. Then, the model was used to predict growth in the 28 day and two year studies. Only the first 12-14 weeks of growth data from two year studies were used to test model predictions. The reasons for this cut-off point are addressed in detail in the discussion.

Matrices containing weekly averages of feeding rate, *f*, and toxicant ingestion rate, *I*, were generated from the data as described previously. While the TK and TD parameters remained fixed, it was necessary to repeat the fitting of growth parameters (J_{Am}^a and J_M^V) to the control group data. This was important so that the effects of feeding rate on growth were predicted relative to the control group of each study rather than to that of a separate study in which conditions (laboratory rat strain, feed, average temperature) may have differed.

All model parameters were then fixed and the model was used to predict effects on growth in all treatments. Predictions were compared to observed data for each treatment at each observed time point. Predictions were considered in terms of animal weight (g) or the proportional effect on body weight relative to the control group (mean weight$_{\text{(treatment)}}$ × mean weight$_{\text{(control)}}^{-1}$). Predictions were deemed accurate if they were within one standard deviation of the mean observed value at each time point as this measure takes into account the individual variability within the data. The percentage of predictions that were accurate was reported and any

exceptions were noted. Exceptions were used to determine the limitations of the model and to infer the underlying reasons why they arise.

3.5 Results

Results are summarised in this section however, due to the number of studies used to calibrate and test models, it was not practical to include all tables and figures here. These can be found in the Appendix A.

3.5.1 Toxicokinetics

For all the pesticides the concentration reached in the blood was significantly correlated with that in all other tissues sampled ($p<0.01$) except for the concentration of fenpropidin in fat which was not significant ($p = 0.0574$). However when one outlier (residual > 3 s.d. from mean) was removed from the analysis, the correlation was highly significant ($p<0.001$).

All the compounds exhibited first order kinetics, producing blood concentration time curves which could be reproduced by fitting of a one compartment TK model. In some cases adjustments to the modelling procedure were required which will be described in turn.

For thiamethoxam (fig. A6), global model fits at both high (R^2 = 0.83 for males and 0.74 for females) and low dose levels (R^2 = 0.68 for males and 0.71 for females) closely matched the observed data. When model predictions were tested against independent blood concentration data, the predicted curves again closely emulated the observed data.

For four of the compounds, acibenzolar-S-methyl, azoxystrobin, fludioxonil and fenpropidin (figs. A1-A4), the global model fits better represented blood kinetics in the high dose group ($0.65<R^2 \leq 0.89$), with modelled curves not reaching the peak blood concentration (C_{max}) observed in the low dose groups ($-0.48<R^2 \leq 0.37$). In all these cases this same pattern was observed when model predictions were tested against an independent blood concentration data set. The likely explanation for this phenomenon is that, as pesticide concentration in the gut is increased, absorption rate becomes saturated and reaches a maximum (Sjovall et al., 1985). However, with only two dose levels tested in most toxicokinetics studies, generally differing by a factor of at least 100, it is not possible to estimate the point at which this occurs or to determine whether it is a gradual process or happens suddenly. Lower model accuracy at low internal concentrations has little impact upon eventual predictions of effects, and none at all if below the *NEC*. As the high dose levels were more relevant to the dietary ingestion rates associated with effects on body weight, these parameter values were accepted.

Male rats administered a high dose ($100mg \times kg_{(BW)}^{-1}$) of fenpropidin appeared to exhibit a double peak in the concentration reached in the blood. An initial peak was reached 1 hour after dosing

with a second, lower peak after around 8 hours. Double peaks have been attributed to variable absorption in different regions of the gut, enterohepatic recirculation or delayed gastric emptying (Godfrey et al., 2011). Were the dose delivered at a more constant rate in the diet rather than as a single large dose one would expect the overall rates of absorption and elimination to reach equilibrium and so any of these mechanisms would have less impact on blood kinetics. As the intended application of the model was to predict the effects of dietary dosing, it was decided to fit the model for males with the data collected 2h, 3h, 4h and 6h after dosing excluded in the high dose group. This allowed the model to fit a single peak in which the observed C_{max}, the time at which it was reached (T_{max}) and time for total elimination of the dose were matched closely by the model (fig. A3).

Following administration of a high dose, peak concentration of prosulfuron in the blood ($\mu g_{(AI)} \times g_{(Blood)}^{-1}$) exceeded the body weight normalised dose ($mg_{(AI)} \times kg_{(BW)}^{-1}$). This could not be modelled by our TK model (equations 3.1 & 3.2) which uses blood concentration as a proxy for overall internal concentration. However, blood only accounts for around 7% of body mass (Lindstedt and Schaeffer, 2002) so in reality only >7% of the dose needs to be present in the blood at one time for this to occur. Nevertheless, this phenomenon was unusual among the chemicals included in the study.

In order to address this, the relationship between prosulfuron concentration in the blood (C_{Blood}) and overall carcass concentration of terminated animals was investigated. Only male animals were used in the tissue sampling experiments with three animals sampled at each of four time points following a high or low dose. Blood and carcass concentration were strongly and significantly correlated (Pearson's correlation coefficient, r = 0.99, n=24, p <0.0001) . The line of best fit, intercepting the y axis at zero, was derived by finding the least squares solution to the equation $C_{Blood} = XC_{Int}$. The gradient, $X=2.4337$, was determined as the concentration factor by which prosulfuron concentration in the blood exceeds that in the body as a whole.

A third equation, incorporating that concentration factor but otherwise identical to equation 3.2, was then added to the TK model to describe blood concentration over time as:

$$\Delta C_{Blood} = 2.4337 \, {}^{L_\infty}/_L \, k_a C_{Gut} F - {}^{L_\infty}/_L \, k_e C_{Int} - C_{Int}(\Delta W/_W) \qquad \textbf{(3.5)}$$

This determines that, $C_{Blood} = 2.4337 \times C_{Int}$ at any given time point, with ΔC_{Int} modelled by equation 3.2. Blood and whole body internal concentration could then be modelled simultaneously. The fitted model produced curves matching the data well for both variables at the high dose ($R^2 = 0.95$ for C_{Blood} and 0.97 for C_{int}, fig. A5).

In contrast to other compounds the prosulfuron model predicted higher than observed internal concentrations at the low dose level. This was the case for concentration in the whole body as well as in the blood so the concentration factor was not the cause of the discrepancy. In fact, for the low dose group alone the concentration factor was higher than the overall figure. A possible explanation in this case is that the elimination rate becomes saturated beyond a certain internal concentration (Lesko, 1979, Caccia et al., 1990), this would be consistent with the unusually high internal concentrations measured in the high dose groups. The high dose in the prosulfuron toxicokinetics experiments was around $450\text{mg}_{(AI)} \times \text{kg}_{(BW)}^{-1}$, several times higher than the high dose used for other chemicals included in this study ($100 \text{ mg}_{(AI)} \times \text{kg}_{(BW)}^{-1}$). Had the other chemicals been tested at such high doses it is possible that a similar pattern would have been evident. This meant that, of the compounds in this study, only thiamethoxam showed no dose dependence in uptake. Yet again, only two dose levels (in this case differing by a factor of around 900) provide insufficient data to determine the maximum elimination rate and the internal concentration at which it is reached.

With fixed parameters and independent data, the model again predicted higher prosulfuron concentration in the blood than was observed at the low dose. The data were predicted well at the high dose however. For both sexes, C_{max} and T_{max} were predicted with reasonable accuracy. Elimination of the compound was slower than predicted in females, but the parameters were deemed acceptable and the model was not fitted to males and females separately.

3.5.2 Growth

The fitted DEBkiss growth curve was able to accurately model growth of rats aged around 6-20 weeks. In total, the model was fitted to 34 control group datasets comprising weekly observations of body weight and food consumption rate. Modelled body weight was within 1 standard deviation of the observed mean at all time points in 30 out of 34 cases and at >90% of time points in 32 out of 34 cases (figs. A7 – A40). The deviations were most pronounced in two data sets. For the female control group in the 28 day toxicity study of fenpropidin, the modelled body weight was lower than the observed mean by more than one standard deviation at week one only. As a result, only 75% of the modelled weights were within one standard deviation. For the male control group in the 90 day toxicity study of azoxystrobin, the modelled body weight was lower than the mean by more than one standard deviation in weeks two and four. This resulted in only 84.6% of the modelled weights being within one standard deviation of the observed mean. For every data set, all predictions of body weight in the control group were within 10% of the mean at all time points.

3.5.3 Toxicodynamics

With the growth and toxicokinetic parameter values fixed, the toxicodynamic parameters, NEC and C_T, were fitted to selected data as described in the methods section. Predictions were interpolated to other treatments in the 90 day studies and extrapolated to 28 day and 2 year studies.

In terms of body weight ≥75% of predictions were within 1 standard deviation of observed means for 28 out of the 34 study groups. In terms of effect on body weight, ≥75% of predictions were within 1 standard deviation of observed means for 30 out of the 34 study groups. Model accuracy was low for male and female rats given acibenzolar-S-methyl over 28 days and males given fenpropidin over 28 days. A summary of results is shown in table 3.2. For female rats administered thiamethoxam or fludioxonil, TD parameters were fitted to a single treatment group. This was because, in studies of both compounds, body weight reductions beyond those predicted based on feeding rate were only evident in the top dose group. In the case of thiamethoxam, significant body weight reductions were only observed in females dosed with 10,000mg/kg$_{(diet)}$ of thiamethoxam over 28 days (fig. A39). Females administered fludioxonil in their diet did show significant body weight reductions. However, these were predicted entirely based on reduced feeding rate in all but one treatment, those dosed with 20,000mg/kg$_{(diet)}$ over 90 days (fig. A27). The consequence was that while the *NEC* was consistent with data from several treatments, C_T was determined using data from only one treatment. However, since toxicant ingestion rate was dynamic, C_T was fitted to a range of internal concentrations even within one treatment.

Table 3.2 *Toxicodynamic parameters used to model the effects of each compound on male and female rats. The percentage of predictions (in terms of absolute body weight and effect on body weight relative to the control group at each time point) within one standard deviation of the observed mean, are shown. Percentages ≥75% are highlighted in green, those of ≥50% and <75% are shown in blue while those <50% are highlighted in orange. Those marked with a '*' were fitted to only one treatment group. pMoA: best fitting physiological Mode of Action.*

Compound	Acibenzolar-S-methyl						Azoxystrobin						Fenpropidin					
Sex	Male			Female			Male			Female			Male			Female		
pMoA	Maintenance			Growth Efficiency			Assimilation			Growth Efficiency			Maintenance			Assimilation		
NEC (mg × kg$_{(BW)}^{-1}$)	1.08E-06			7.93E-07			1.20E-04			4.62E-08			1.40E-07			0.21		
C_T (mg × kg$_{(BW)}^{-1}$)	507.4			79.99			424.77			270			24.02			31.86		
Study	90d	28d	2y	90d	28d	2y	90d	28d	2y	90d	28d	2y	90d	28d	2y	90d	28d	2y
% weight predictions ±1s.d. of mean	92%	42%	92%	75%	25%	73%	87%	94%	100%	92%	100%	100%	77%	25%	70%	100%	63%	80%
% effects predictions ±1s.d. of mean	100%	42%	90%	85%	33%	75%	90%	94%	98%	100%	94%	100%	80%	25%	70%	100%	75%	83%

Compound	Fludioxonil						Prosulfuron						Thiamethoxam					
Sex	Male			Female*			Male			Female			Male			Female*		
pMoA	Maintenance			Maintenance			Maintenance			Growth Efficiency			Maintenance			Maintenance		
NEC (mg × kg$_{(BW)}^{-1}$)	7.32E-11			38.09			1.01E-06			3.26E-07			1.34E-08			50.64		
C_T (mg × kg$_{(BW)}^{-1}$)	251.7			49.33			2676			561.5			656.3			154.6		
Study	90d	28d	2y	90d	28d	2y	90d	28d	2y	90d	28d	2y	90d	28d	2y	90d	28d	2y
% weight predictions ±1s.d. of mean	100%	75%	-	98%	100%	-	100%	95%	96%	94%	100%	100%	86%	75%	90%	91%	100%	96%
% effects predictions ±1s.d. of mean	100%	75%	-	100%	100%	-	100%	100%	100%	96%	100%	100%	100%	75%	100%	98%	100%	100%

Model predictions were also used to investigate the extent to which reductions in body weight could be attributed to reduced feeding or direct toxic action. This was done by comparing experimental data to model simulations in which no stress was applied, producing growth curves predicted based solely on feeding rate. Comparisons between expected growth modelled with actual feeding rates and observed growth were conducted separately for every treatment group. The observed and predicted body weights in each treatment were converted to proportions of the control body weight (observed and predicted respectively) at each time point. The proportion of any observed body weight reductions (relative to controls) that were predicted based on feeding rate alone could then be calculated. The remainder was attributed to toxic action. If body weight predicted based purely on actual feeding data was below that observed at a given timepoint then any observed weight reduction (relative to controls) was attributed entirely to reduced feeding rate. Likewise, if body weight predicted based on actual feeding rate in a given treatment group was higher than controls, then any observed weight reduction was attributed entirely to toxic action. An example of this process is shown in figure 3.4; selected results for all compounds are shown in figure 3.5.
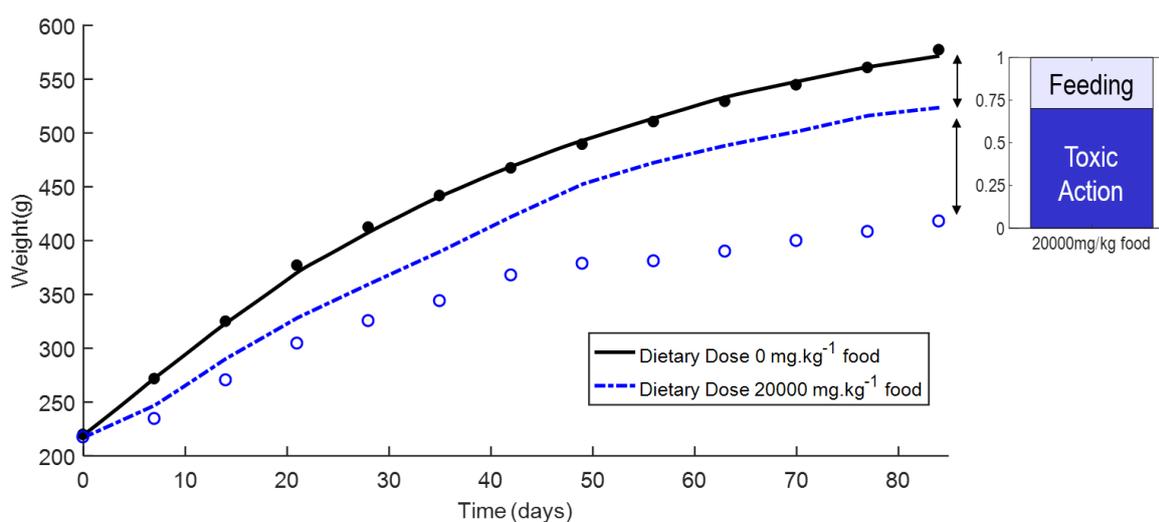
Figure 3.4 *Growth modelled based on feeding rate only (lines) and observed growth (circles) of male rats. The control group and those dosed with 20,000 mg × $kg_{(diet)}^{-1}$ fludioxonil are shown. The proportional breakdown of the observed reduction in body weight of treated rats vs controls at the end of testing is represented in a bar chart.*

For azoxystrobin, prosulfuron, thiamethoxam and fludioxonil there appeared to be a pattern across the sexes, with reduced growth being driven more by feeding rate in females and by toxicity in males (figure 3.5). This was most evident in the case of fludioxonil, which was associated with significant body weight reductions in both sexes. While reductions in male body weight were attributed largely to toxicity, the reductions observed in females were predicted based entirely on reduced feeding rate in all but the highest dose group across two studies. A

similar pattern was seen for thiamethoxam. Once again toxic effects were only predicted to impact upon female bodyweight in the highest dose group. In this case however, reductions in female body weight were not observed in most treatments as feeding rate was not affected either.
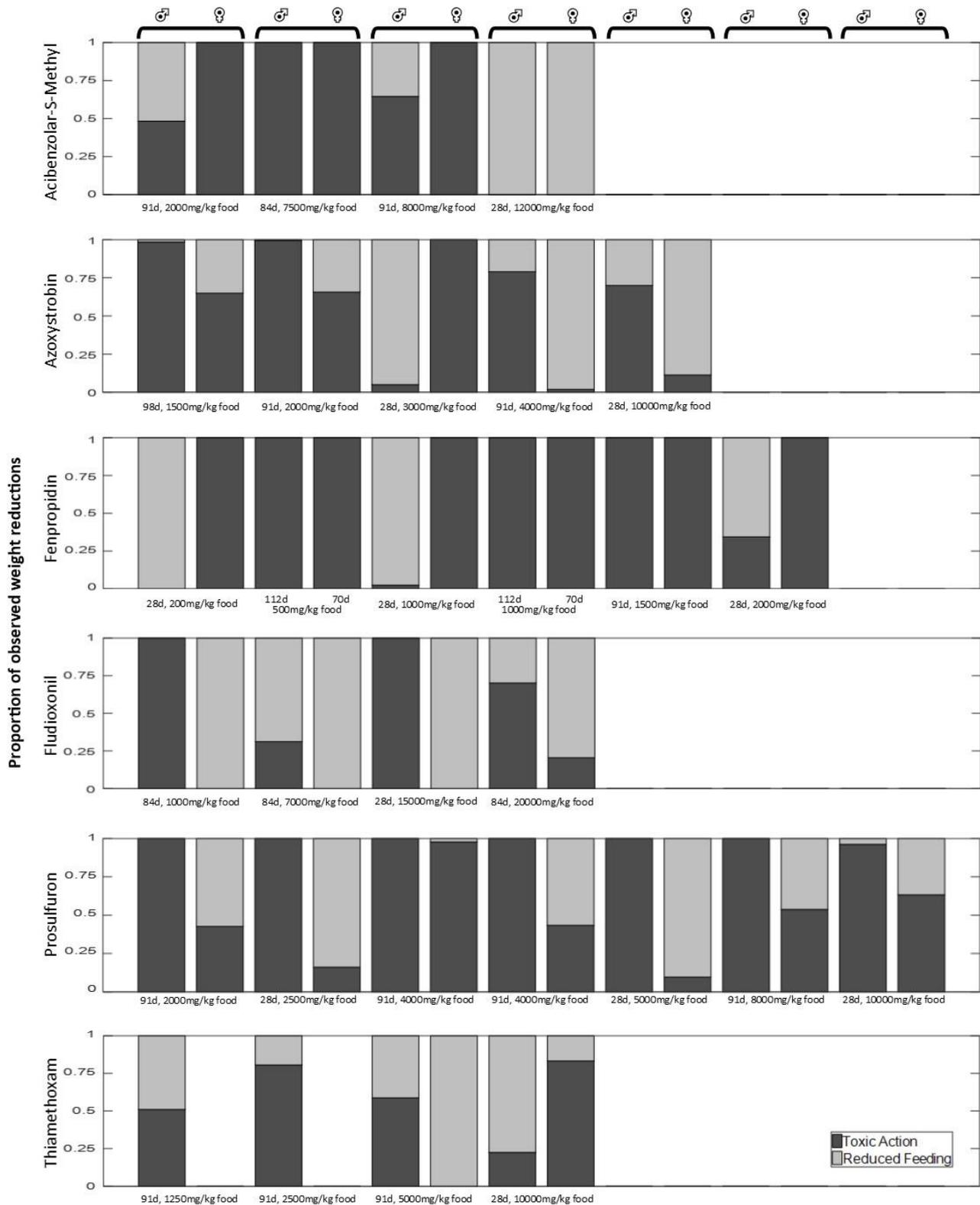
Figure 3.5 *Bar charts showing the proportion of observed weight reductions relative to the control group attributed to reduced feeding rate and/or toxic stress by the growth model. All treatments in which a weight reduction was evident at the end of the analysed period are included. X-axis labels denote the observation date and dietary dose, in some cases treatments were duplicated between*

studies. No bar is displayed where there was no reduction in weight. Note that bars are the same size regardless of the magnitude of the observed effect.

3.6 Discussion

3.6.1 Predicting Growth Under Chemical Stress

Raw data from chronic toxicity studies were used to test DEBtox predictions of sublethal toxic effects in mammals. Weekly measurements of body weight and food consumption as well as precisely measured dietary concentration provide good quality data with which to calibrate models. As the regulatory framework requires several such studies, abundant data are available to test predictions (OECD, 1998, OECD, 2008, OECD, 2001). As has been noted previously, the focus on novelty means that funding for academic studies is rarely allocated to corroboration studies. Moreover, any such studies which are funded are also unlikely to be published in the scientific journals for the same reason (Jager and Ashauer, 2018).

Our findings showed good agreement between predictions and data. With minimal model fitting (only selected treatments from 90 day study), observed effects on body weight were predicted reliably (≥75% of predictions accurate to within one standard deviation of the observed mean in all studies for which data were available) in males and females for four of the six chemicals modelled. This suggests TK predictions were at least proportional to actual internal concentration over time and that the 'linear past threshold' TD model (figure 3.2) is based on reasonable assumptions.

3.6.2 Feeding Rate vs Toxicant Ingestion

An obstacle when analysing the effects of dietary toxicant exposure is that the ingested dose depends as much on the feeding rate of the study animals as it does on the concentration in the diet. While a high feeding rate will have a positive effect on growth, the corresponding toxicant ingestion will place stress on growth parameters. Understanding of this trade-off is important (Thompson, 2007) as ingestion is considered the primary exposure route for terrestrial mammals to pesticides in the field (Bednarska et al., 2013a) although it is argued that other routes should receive greater attention (Mineau, 2011, EFSA, 2009b). In the modelling approach used here, both the scaled feeding rate and pesticide ingestion rate were calculated as dynamic model inputs, directly from the data. This allowed model predictions to separate the competing effects of feeding rate and toxic action on growth. For example, male rats given 2500 mg/kg$_{(diet)}$ thiamethoxam over 90 days grew larger than those given 1250 mg/kg$_{(diet)}$. This result was correctly predicted by the model as the higher feeding rate of the 2500 mg/kg$_{(diet)}$ group partially counteracted the chemical stress (fig A35).

Model predictions can therefore provide new insight into the observed data by comparing data to model simulations in which no stress is applied. Such predictions only require the calibrated growth model and data on feeding rate. Thus, even if toxic effects cannot be reliably predicted, growth model predictions can indicate the degree to which observed reductions in body weight were driven by toxicity or reduced feeding relative to controls. Such information is valuable for assessing the risk that a chemical poses to terrestrial mammals in the field. If strong avoidance is observed, then this may increase or decrease the risk posed depending on whether animals would have a choice of food items in the field scenario (EFSA, 2009b, Thompson, 2007). The degree to which reduced feeding can be regarded a toxic effect is an interesting question and depends on the underlying cause. Reduced feeding due to nausea or suppressed appetite could be considered a toxic effect but reduced feeding due to palatability is simply a behavioural response.

For several compounds there appeared to be a pattern across the sexes, with reduced growth being driven more by feeding rate in females but by toxicity in males. This pattern was strongest for fludioxonil and thiamethoxam and in both cases was reflected by a large difference in the values assigned to the NEC for each sex. This would suggest that while females exhibit a higher tolerance for these compounds, at least with respect to growth, they show stronger feeding avoidance. Such inconsistency in the toxicodynamic parameters of males and females may seem surprising however, large differences between reference doses for each sex have long been documented in rats (Calabrese, 1986). Moreover, the results were unequivocal with respect to thiamethoxam, as females were unaffected at several dose levels which affected males. In several cases, differences have been noted in the sensitivity of the liver and kidneys to toxicity (Calabrese, 1986, Seralini et al., 2007), possibly related to variable enzyme production between the sexes (Moser et al., 1998). The liver and kidneys were identified as the target organs of these pesticides in mammals so these results would appear consistent with previous findings (Syngenta, unpublished). Given that such differences in chemical sensitivity can occur between the sexes it should not be surprising either that the models suggested different modes of action for several of the compounds in males and females (table 3.2). Mode of action in DEBtox refers to abstract processes rather than specific chemical pathways so, this simply implies that the effects on the growth curve differed between the sexes (Jager, 2015b).

3.6.3 Model Limitations

A fundamental limitation to any model of a complex system is the trade-off between realism and simplicity. As user friendliness is a significant consideration for regulatory use (EFSA, 2018), DEBkiss was selected as the simplest possible approach to investigate how raw lab data should be utilised to parameterise and calibrate models with data from dietary toxicity studies. Another reason for prioritising simplicity was so that potential issues could be clearly identified at this

early stage for future models to develop, with scope for further elaboration if necessary. Several such issues were highlighted by our results which are discussed in this section.

While our results demonstrate that DEBtox is a useful framework, predictions of growth under chemical stress were not reliable in all cases. The effects of fenpropidin and acibenzolar-S-methyl were predicted accurately in several treatments, but overall model accuracy was lower than for the other chemicals. When considering *in vivo* effects, any observed deviations from predictions are unlikely to result from measurement error, more often, individual variability is the cause. This represents a hurdle to predictive modelling, even for genetically similar laboratory strains kept in controlled conditions. Indeed, individual variability in growth is still evident in studies using genetically identical springtail (*Folsomia candida*) clones kept individually and provided unlimited food (Jager, 2013). As such, we cannot know for sure how treated animals would have grown without a toxicant in their diet, so growth parameters fitted to the control group must be fixed across treatments. Consequently, models provide predictions of mean body weight over time under specified conditions but do not account for individual variability in parameter values.

Another consequence of individual variability is that it can be difficult to identify the underlying causes of model inaccuracy. For example, several treatment groups of female rats fed lower dietary doses of acibenzolar-S-methyl grew larger than controls despite feeding at a lower rate. This is clearly a result the model would not predict and could simply be the result of variability in average growth parameters between treatment groups. Alternatively, this could be interpreted as evidence of hormesis (Mattson, 2008), the phenomenon by which lower doses of a chemical have the opposite effect of higher doses on a given endpoint. If this were the case, then the stress function rather than the growth parameters would require alteration, but we cannot be sure which. In a few treatments the opposite issue arose when the animals fed at a relatively high rate but did not grow as expected. Since feeding rate was calculated based on actual body size rather than predicted size, modelled growth continued at a higher than observed rate. For these treatments, model predictions far exceeded observed growth. This issue could be somewhat resolved by calculating feeding rate relative to the predicted body size over time. However, it is likely that variability in growth parameters also played a role.

Furthermore, growth is not the only modelled property subject to individual variability. Tolerance to a toxicant (Barata et al., 2002) or the rates at which it is taken up and eliminated (Bednarska et al., 2013a) may be highly variable among individuals in a population. Inter-treatment differences in any one or more of these properties could result in observed effects not being uniformly correlated with internal concentration and are therefore difficult to predict without further knowledge. This was evident for mandipropamid; results were not reported as

the data were not suitable to test model predictions. In the 90 day study, males in the highest dose group grew larger than those at a lower dose despite feeding at roughly the same rate. Females meanwhile grew larger than predicted based on their feeding rate in all dose groups. The highest dose groups in the 28 day study were terminated early due to unacceptable reductions to feeding rate, while reduced growth was not observed in the 2 year study so toxicodynamic predictions could not possibly be validated.

Poor model accuracy (<50% predictions accurate to within one standard deviation) only occurred when predicting the effects of 28 day dietary exposure. This is not entirely surprising as there are several factors making these data sets more challenging to model. In 28 day studies sample size was lowest, with only 5 individuals per sex and only four time points observed so naturally, individual and temporal variability would be expected to have a larger impact. Moreover, in the early weeks of dietary studies feeding rate can be highly variable between treatments and over time, as animals react behaviourally to a novel ingredient added to their diet. In all 3 data sets for which predictions were poor, growth predictions in the highest dose groups were substantially lower than observed. This may have resulted from internal dose being overpredicted as dose-dependent uptake was apparent for these compounds. Another possibility is that the default value of k played a role here. If this value were too low this would stop growth when it could still occur, though this would be almost entirely compensated for in the fitting of other growth parameters. Including the reproduction buffer could have delayed the need to metabolise structure (Jager et al., 2013), but using up the buffer would still correspond to a reduction in overall body weight. Moreover, the model only predicted starvation to occur in the early weeks of testing. At this point, the rats were around the onset of puberty (Rakel and Gergs, 2018) when the buffer begins to accumulate (Jager et al., 2013), so any buffer amassed by this point would be almost negligible.

It is quite possible that the dual stresses of reduced feeding and toxicity elicit compensatory physiological or behavioural responses not predicted by the model. Reduced body temperature has been documented as a response to starvation in rats (Sakurada et al., 2000) meanwhile chemical stress has been shown to induce reductions in body temperature and activity (Buwalda et al., 2001). Such responses would likely correspond to a reduction in the maintenance rate, J_M^V, and should be considered in future models of physical and chemical stress.

Individual variability was also evident in the toxicokinetics data. Individual responses varied with regard to the toxicant concentration reached in the blood and the speed with which it was absorbed and eliminated. The low sample size of three individuals per treatment meant that mean observations for each time point, to which the models were fitted, could be heavily influenced by

variability on either axis. As an example, even if three individuals exhibited a single peak with very little variability in C_{max}, the average data could show a double peak if just one individual peaked later than the other two. Moreover, both peaks would likely be lower than the C_{max} of any individual. Even simple quantitative assessment of model fits, such as whether model predictions were in the observed range for toxicokinetic statistics such as C_{max}, T_{max} or area under the curve, were therefore problematic. Besides, which of these statistics would be more relevant for predicting the internal concentration resulting from prolonged dietary exposure is debatable and none of them they are used by TK-TD models to predict effects. Nevertheless, TK model predictions provide plausible estimates of internal concentration resulting from the recorded time varied ingestion rates, with parameter values fitted to the best available data. It is likely that the predictions are at least proportional to the true values and therefore form a credible basis for the fitting of TD parameters.

While it is not possible to separate variability in TD and growth parameters, future model iterations could incorporate stochasticity in growth parameters by utilising the wealth of control data available from studies on *R. norvegicus* (Jager, 2013). This would also be possible for TK parameters although low sample size would present a challenge. An alternative would be to fit the TK model individually (where appropriate data are available) and use the average parameter values, rather than fitting to average data. However, suitable individual data were not available for all the compounds in this study.

3.6.4 Implications for DEB theory

In this study, precise food consumption data were available, rather than simply food availability. This allowed the mean surface area specific feeding rate to be calculated on a weekly basis for each treatment in a study, before being scaled as a proportion of the maximum value in each data set. Consequently, the growth parameters were fitted to controls in each data set to account for variability in feeding rate within and between studies. Previously, it has been assumed that the scaled feeding rate, f, is equal to 1 when food is available *ad libitum* (Desforges et al., 2017). For certain purposes this is a reasonable and necessary assumption. When modelling growth in the field for example, detailed data are unlikely to be available and so feeding rate must be estimated as a fraction of *ad libitum* feeding in laboratory studies. However, this was not satisfactory in this study as variability in feeding rate over time and between treatments was an important driver of effects. Moreover, several treatment groups fed at a higher rate than controls, so it was important not to assume the maximum value of *f* as the default.

Growth ceases at the ultimate weight, W_∞, as this is the point at which the maximum assimilation rate can only match the maintenance requirements of the organism (Jager et al., 2013). Assuming

f = 1 means that W_∞ is equal to the theoretical maximum weight, W_m, when food is available *ad libitum* (see table 3.1). However, a marked trend in the data was that, even with unrestricted access to food, feeding rate relative to surface area declined as the animals grew. This presents a clear issue in that it places an additional limit on the assimilation rate and ensures that the growth curve reaches a plateau before W_m is reached. In many cases, by week 13-14 of observation, the value of f was well below the maximum and so the theoretical W_m based on fitted parameter values was unrealistically high.

Recent DEBtox studies of toxic effects on rainbow trout have also sought to account for variability in f, however this has focused on differences between study groups rather than within treatment changes over time (Zimmer et al., 2018, Sadoul et al., 2018). Much effort has gone into deriving standard DEB parameters for species of interest (Rakel and Gergs, 2018, Marques et al., 2018, Kooijman, 2015). However, lifetime variation in f must be considered for these to be compatible with time varied feeding data rather than the constant food density. For laboratory strains of *R.norvegicus* this could certainly be addressed; due to their extensive use in regulatory testing there exists a vast database of growth and food consumption in control conditions. If the relationship between feeding rate and body size were described mathematically, this could be utilised for scaling observed feeding rate such that the maximum feeding rate decreases with size and the resulting value of f remains roughly constant over the lifetime.

Although longer term data were available, predictions of toxic effects on body weight were not reported beyond around 12-14 weeks of dietary exposure for the two year studies. Beyond this point the intervals between body weight and food consumption observations increased. Since observed feeding rate was used as a model input, this reduction in data resolution would be expected to adversely affect predictions. There were also more fundamental reasons behind this cut-off. While sigmoidal curves such as the Von Bertalanffy growth model (to which DEBkiss simplifies (Jager et al., 2013)) can approximate the growth curve of rats, there are distinct stages where observed growth deviates from such a model. It has previously been reported that the Gompertz function – also sigmoid – matches data closely when fitted to the first 70-105 days of rat growth but that longer term predictions are problematic (Pahl, 1969). When fitted to the full two years of control data the DEBkiss curve also showed systematic errors. For both males and female rats, predicted body weight was lower than observed for roughly the first three months, higher than observed until around month 14 and then lower than observed for the final 10 months.

A possible explanation is that k (the allocation to soma and reproduction), does not remain constant throughout the rat's lifespan. A logical suggestion is that a greater proportion of energy

may be allocated to growth earlier in life with more energy used for sexual maturation during puberty. Indeed, it has been postulated that k may change in humans at puberty (Kooijman, 2000). It is also likely that reserve dynamics become more important in adult rats, as continued ad libitum food availability allows animals to develop significant reserve stores after structural growth has ceased. The full DEB model which, unlike DEBkiss, models reserve as well as structure may therefore be better able to address this issue.

Another issue with long term predictions is that there is insufficient knowledge regarding recovery after inhibition of growth in rats. By default, after a stressor is removed, modelled body weight may theoretically reach W_m if feeding rate is high enough. Based on our analysis, this assumption appeared sound for rats up to around 23 weeks of age. Model predictions agreed well with data in cases where 90 day studies included four weeks of recovery for the highest dose group. However, skeletal growth is known to stop in rats at around 26 weeks of age, the underlying processes are complex but appear to be related to age rather than size (Roach et al., 2003). Logically, if growth has been suppressed up to a critical age then a full recovery, relative to controls, will not be possible as growth will cease before W_∞, let alone W_m, can be reached. We hypothesise that this occurs because, beyond a certain age, energy is allocated to processes other than growth, such as maturation and reproduction. This would correspond to a reduction in the parameter k, resulting in a reduced growth rate and, crucially, a lower W_∞, for animals that had experienced stress (Jager et al., 2013, Kooijman, 2000).

Realistic constraint of recovery is essential for long term predictions to be of use. Otherwise, to match observed data, TD parameters must continually stress growth even when it is no longer possible, and so exaggerate the toxicity of a chemical. Recovery may be limited as a function of the (structural) weight reached by a critical age. However, determining rules by which to accurately decrease the value of k thereafter would likely require significant experimental work. Removing a given stressor at different timepoints may identify the age at which a full recovery becomes impossible. However, subsequent experiments would still be required to examine how potential for recovery is affected by the level of stress as well as the duration. Such experimental work was beyond the scope of this study. In any case, for regulatory purposes such long term predictions are of limited relevance as pesticides are not applied at constant rates for years at a time. The 12-14 week exposure periods modelled in study go well beyond the 21day exposure considered in ecological risk assessment (EFSA, 2009b)so should be sufficient for extrapolation to realistic exposure scenarios

3.7 Conclusions

This study shows that DEBtox modelling with DEBkiss can provide an effective and simple to use tool for predicting toxic effects on growth in rats. We show how time varying model inputs for feeding rate and pesticide ingestion rate can be calculated directly from data generated in standard chronic toxicity studies, providing additional insight into data by indicating to what extent body weight is impacted by feeding rate or toxic effects over time.

We also identify several difficulties which future models should aim to overcome. Individual variability presents a significant obstacle to assessing model accuracy. Our models simulate toxicokinetics, toxicodynamics and growth, all of which may be subject to variability. In most cases predictions were accurate to within one standard deviation of the observed mean and so provide useful estimates of the mean but not exact projections including variability.

Our findings support DEB theory as an effective basis for predictions of sub-lethal toxic effects in mammals. However, some issues became apparent regarding its compatibility with chronic toxicity data. Given the extensive use of *R. norvegicus* in laboratory testing and the resulting wealth of control data, these complications can be addressed. Further analysis of lifetime variation in feeding rate and energy allocation to the soma may improve model accuracy and realistically constrain recovery. Such adjustments would broaden the range of applications for which DEBtox may be used.

Chapter 4 - Can TK-TD modelling bridge the gap between in vitro and in vivo mammalian toxicity data?

4.1 Preface

The aim of this chapter was to build on research by Stadnicka-Michalak et al. (2015), which used TK-TD modelling and *in vitro* assays to predict toxic effects on the growth rate of fish. As an experimental study, this chapter is very different to the others in the thesis, but the models developed in Chapter 3 were an integral part of the methodology being tested.

Having built TK-TD models to predict the growth of rats under toxic stress, I began to consider how they might be used to develop *in vitro – in vivo* extrapolation methods for rodents. The dietary dosing route, used in studies on terrestrial species, means that ingested dose and the resulting internal concentration can be highly variable over time. Therefore, any direct comparison between *in vitro* and *in vivo* data is not meaningful, so model predictions of growth under constant pesticide exposure must serve as a proxy for *in vivo* observations. With this in mind, I formulated an *in vitro/in silico* approach and planned experiments to determine whether toxic effects on *in vivo* growth could be predicted. These were carried out in labs at the University Hertfordshire. Chapter 4 is a research paper describing this study.

4.2 Abstract

Repeated dietary dose testing is used to assess longer term toxicity of chemicals, such as pesticides, to mammals. However, the internal pesticide concentration varies significantly as feeding rate relative to body size fluctuates over time. Toxicokinetic-toxicodynamic (TK-TD) models estimate internal toxicant concentration over time and link this directly to observed effects on growth rate of laboratory rats. Using TK-TD models it is therefore possible to predict the effects that would result from a constant internal concentration of a pesticide. This presents the possibility of comparison with data from *in vitro* experiments, potentially facilitating quantitative *in vitro* to *in vivo* extrapolation (QIVIVE). We used *in vivo* TK-TD models to identify relevant internal concentrations and then estimated the experimental conditions required to replicate these in cultured cells, using *in vitro* TK models. Cell population growth was measured, with a view to extrapolating through time and comparing effect sizes with *in vivo* predictions. However, observed cell proliferation was not significantly affected (p>0.05) by the tested concentrations of any of the five pesticides in this study and so extrapolation was not possible. In light of this negative result, we highlight areas for future work toward QIVIVE of graded sublethal effects in mammals. The most pressing objective is to improve the accuracy of *in vivo* TK predictions. The inclusion of a dietary dose group in regulatory TK studies would facilitate this by providing calibration data better suited to predicting internal concentration in dietary toxicity

studies. If the current issues can be addressed, TK-TD modelling may yet serve as a vital bridge between *in vitro* and *in vivo* data, enabling accurate QIVIVE methods for mammals. This would represent a major step toward the reduction and replacement of animal testing.

4.3 Introduction

Though it is an important part of chemical safety testing, the use of animal testing has long been the subject of intense scrutiny due to ethical concerns. The European Union (European Parliament, 2010) and many national governments including those of the UK (Home Office et al., 2014), and US (EPA Press Office, 2019) are committed to the replacement, reduction and refinement of animal testing (the 3 R's) through the development of alternative methods. However, the push towards new methods is not driven solely by policy makers, there is also demand from across industries and scientific disciplines. This is because, in addition to ethical issues, animal testing is time consuming, financially costly and limited in its predictive ability for human safety. In the development of a new pesticide for example, an average of $29 million is spent on toxicology assessment (Phillips McDougall, 2016). Meanwhile, the cost of advancing a new pharmaceutical to phase 1 human trials, may reach $100 million and failures at this stage are often due to differences in drug kinetics between humans and animal test species (Tsaioun et al., 2016).

The need for high throughput alternatives to animal testing is made more pressing by the EU's REACH legislation, requiring many more chemical products to undergo risk assessment (Rovida and Hartung, 2009) and the US EPA's recent pledge to end funding for tests on mammals by 2035 (EPA Press Office, 2019). To this end, the adverse outcome pathway (AOP) framework has been developed (Willett, 2019), seeking to mechanistically link processes that lead from molecular interactions to adverse outcomes at higher levels of biological organisation. In this framework, non-animal methods such as quantitative structure-activity relationships (QSARS) and *in vitro* assays may be used to identify potential hazards associated with a compound. This can reduce the need for *in vivo* assays through targeted testing. However, the risks posed at organ, organism and population level cannot yet be quantified without *in vivo* data (Sewell et al., 2018).

Quantitative *in vitro-in vivo* extrapolation (QIVIVE) could eventually obviate the need for testing on animals. However, there are many obstacles to overcome before this can be achieved (Blaauboer, 2010). Although correlations have been observed between reference doses determined *in vitro* and *in vivo* (Castano et al., 2003, Schirmer, 2006), such simple relationships cannot be relied upon. Cultured cells generally show low absolute sensitivity and there may be significant differences in behaviour of different cell lines. One possible reason for this is that in cultures of a single cell type there will be a relatively low number of target sites for toxicants to

act upon. Another issue is differences in cell function between different cell lines and between *in vitro* or *in vivo* assays, specifically with regards to metabolic clearance and toxic metabolite formation (Schirmer, 2006, Groothuis et al., 2015).

Another essential consideration is the dose metric. In *in vitro* studies, the nominal concentration (simply the amount of test compound divided by the volume of the medium) is often linked to effects. However, this measure is inappropriate for extrapolation (Groothuis et al., 2015). For example, the same nominal concentration is not necessarily even an equivalent dose between two cell lines requiring different culture media due to chemical partitioning. The test compound's affinity for proteins and lipids in each medium will likely vary, leading to different free concentrations available for absorption by cells in each assay (Fischer et al., 2017). Additionally, differences in absorption and clearance rates of the cells themselves will result in further disparity in the toxicant's concentration at its target site between the two assays. The intracellular concentration is therefore a more appropriate dose metric for comparison of responses between different cultured cell lines and between *in vitro* and *in vivo* assays but, is more complex to measure.

A further complication is that the risk posed to an organism by a compound depends not just on the dose level but also the duration and route of exposure (Groothuis et al., 2015, Jager, 2011). This is a particularly important issue in respect to chronic toxicity, generally leading to graded, sublethal effects such as reduced growth rather than binary endpoints such as mortality (Ashauer et al., 2011). This is a key hurdle to the development of alternative methods for the ecological risk assessment (ERA) of agricultural pesticides. Due to the nature of these products, it is inevitable that wildlife will be exposed and so it is vital to identify application rates at which they are effective against their targets without causing adverse effects to wildlife. For birds and mammals, European guidelines state that no mortality should occur as a result of pesticide application (EFSA, 2009b) so chronic toxicity data is essential for decision making. The benefits of alternatives to long-term *in vivo* studies, lasting up to two years (OECD, 2001), are clear as these are the costliest in terms of animal use, man hours and resources.

A potential solution to the issues outlined above is to combine *in vitro* assays with *in silico* modelling of toxic effects (Stadnicka-Michalak et al., 2015, Adler et al., 2011, Li et al., 2017, Louisse et al., 2015). Toxicokinetic-toxicodynamic (TK-TD) models predict an internal measure of toxicant concentration over time based on exposure rate (toxicokinetics) and use this as the basis for prediction of effects (toxicodynamics) (Jager and Zimmer, 2012, Ashauer et al., 2011). This method accounts for exposure duration and profile, allowing for prediction of effects resulting from prolonged, constant or fluctuating exposure (Nyman et al., 2012, Martin et al., 2019, Ashauer

et al., 2017). In recent years models have been developed for the prediction of graded sublethal effects in vertebrate species (Zimmer et al., 2018, Sadoul et al., 2018, Martin et al., 2019, Desforges et al., 2017). Crucially, the use of internal concentration as a dose metric means that TK-TD models can play an important role in linking *in vitro* and *in vivo* data to facilitate quantitative predictions of *in vivo* effects based on *in vitro* observations.

Using *in vivo* and *in vitro* TK modelling, Stadnicka-Michalak *et al.* (2015) designed experiments in which predicted pesticide concentration in the gills of fish in regulatory toxicity studies could be replicated in cultured gill cells. The effects of two pesticides on *in vivo* growth rate could be accurately predicted from observed cell population growth over a few days. While many fish are used in animal testing, they are not the most common test subjects. In the UK, mice and rats were the subjects in 69% of all experimental procedures using animals in 2018 (Home Office, 2019). A means of predicting effects on rodents from *in vitro* data could be a major step towards the reduction and replacement of animal testing. However, differences in exposure route between fish and rodents present an additional hurdle. While fish in regulatory tests are exposed to pesticides at a constant rate in water (OECD, 2014), mammals are dosed via the diet in longer studies (OECD, 1998, OECD, 2008, OECD, 2001). Therefore, internal pesticide concentration depends on feeding rate relative to body size, which fluctuates substantially over time (Martin et al., 2019). Moreover, any effects on feeding rate associated with dosing, such as avoidance of feeding, also contribute to observed effects on body weight over time. These issues present major challenges for any comparison of *in vitro* and *in vivo* data.

In this study we investigated the potential of QIVIVE methods to predict the effects of pesticides on growth in rats. We employed similar methods to Stadnicka-Michalak *et al.* (2015). However, adaptations to the approach were required to account for dietary dosing in mammalian toxicity studies. An important additional step was the use of *in vivo* TK-TD modelling to simulate a constant exposure scenario comparable to *in vitro* dosing. This step also allowed the prediction of effects at internal concentrations relevant to regulatory *in vivo* test conditions, without interference from effects on feeding rate. We aimed to determine whether the intracellular pesticide concentrations predicted *in vivo* would also inhibit cell population growth *in vitro*. Any consistent relationships between *in vitro* and *in vivo* effects could form the basis for predictions across the two scales. This would represent a promising step towards QIVIVE and, more immediately, a useful early screen for toxic effects which could be applicable as an alternative to dose setting studies.

4.4 Materials and Methods

4.4.1 Pre-Existing Models and Data

This study made use of toxicokinetic-toxicodynamic (TK-TD) models developed by Martin *et al.* (2019) and existing data made available from regulatory testing of several pesticides (Syngenta, unpublished). These studies investigated toxicokinetics, sub-chronic and chronic dietary toxicity in laboratory rats (*Rattus norvegicus*) according to 94/79/EC (European Commission, 1994).

4.4.2 Test Compounds

The pesticides included in this study were acibenzolar-S-methyl (benzothiadiazole; fungicide, insecticide and plant activator (PPDB, 2019a)), azoxystrobin (strobilurin; fungicide (PPDB, 2019b)), fludioxonil (phenylpyrrole; fungicide (PPDB, 2019d)), prosulfuron (sulfonylurea; herbicide (PPDB, 2019f)) and thiamethoxam (neonicotinoid; insecticide (PPDB, 2019g)). These compounds covered a range of octanol: water partition coefficient, *P*, with log(*P*) values from -0.13 to 4.12 and all were of low volatility (Henry's law constant ≤1.3 x 10^{-7} atm x m^3 x mol^{-1} at 25°C). Analytical grade samples were purchased from Sigma Aldrich while radiolabelled (^{14}C) samples of each pesticide compound were provided by Syngenta. Pesticides were dissolved in dimethyl sulfoxide (DMSO) before being added to culture medium such that the medium DMSO content of the experimental media was always 0.3%.

4.4.3 Cell Culture

The L6 (ATCC® CRL-1458™) rat skeletal muscle cell line was selected for this study. Rat skeletal muscle cells (L6 cells) were grown in Nunclon™ T25 and T75 flasks (Thermo Scientific) and used between passage numbers 2 and 12 from purchase. L6 cells were maintained in a humidified atmosphere at 37 °C with 5% v/v CO_2 and cultured in high glucose Dulbecco's Modified Eagle's Medium (DMEM), containing 10% v/v heat inactivated fetal bovine serum (FBS), 100 IU/mL penicillin-100 µg/mL streptomycin solution and 2 mM L-glutamine (Sigma Aldrich). Cells were routinely sub-cultured when 70-80% confluent and seeded onto 96 well plates.

4.4.4 Model Implementation

All models were implemented in Matlab (ver. R2016b) and were developed with the BYOM (Jager, 2019) flexible model platform (ver. 4.1). TK model rate constants were fitted to mean pesticide concentration in medium, cells and on plastic over time. The fitting procedure was the Nelder Mead simplex algorithm to maximise the likelihood function, given the observed data (Pan and Fang, 2002). Likelihood profiling was also used to check that initial fits were not local optima (Kreutz et al., 2013) and to calculate parameter confidence intervals.

4.4.5 Modelling effects of constant internal exposure *in vivo*

Dietary toxicity studies (OECD, 1998, OECD, 2008, OECD, 2001), do not provide constant exposure to test compounds. While test diets contain a constant concentration of test compound, body weight specific feeding rate and resulting ingested dose vary significantly over time. As well as temporal variability, inter-treatment effects on feeding rate are common, particularly in the early stages of testing as animals react to a novel ingredient in the diet (Martin et al., 2019). This presents a major obstacle for any comparison between *in vivo* and *in vitro* data.

Models were used to remove the inherent noise from experimental data sets. We used the TK-TD models developed by Martin *et al.* (2019) to simulate a scenario in which all treatments fed at the same rate as the control group and internal pesticide concentration was constant for each treatment. For each treatment group, the average internal concentration predicted under test conditions was calculated (as the mean of all values predicted by the TK model over the test period) and used as a constant input to the TD model. This produced growth predictions within the range observed (figure 4.1) which were suitable for comparison with *in vitro* data.

In Martin *et al.* (2019), the data used to calibrate TK models were pesticide concentration in the blood. For all the pesticides in this and the aforementioned study, pesticide concentration in the blood was significantly correlated with that in all other tissues, so was deemed a suitable proxy for overall internal concentration or 'body burden'. As the L6 muscle cell line was chosen for our experiments, it was necessary to estimate the corresponding pesticide concentration in the muscle. Correlations between pesticide concentration in the blood and muscle were highly significant ($p<0.0001$), and linear relationships were assumed and fitted to these data (table B1 and figure B1). This enabled estimation of target intracellular muscle concentrations to be replicated *in vitro*.
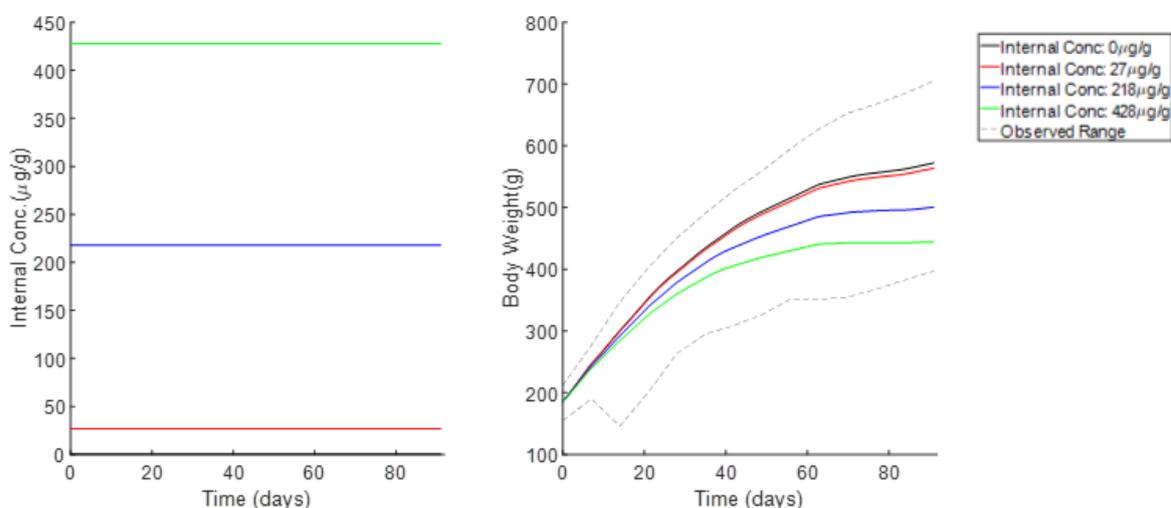
Figure 4.1 *Plots showing predicted effects on growth in response to various internal concentrations of prosulfuron. These predictions are based on growth and toxicodynamics parameters estimated using experimental data in Martin et al. (2019).*

4.4.6 *In Vitro* Toxicokinetic experiments

To assess the *in vitro* kinetics of the pesticides in this study, liquid scintillation counting was used to determine the quantity of radiolabelled pesticide in the medium, in cells and adsorbed to plastic over time. Samples were added to 8 mL HDPE Hinge Cap Vials™ (Perkin Elmer) along with 4 ml Ultima Gold™ scintillation cocktail (Perkin Elmer). The total radioactivity in each vial was then measured in decays per minute by a scintillation counter. The radioactivity could then be converted to quantity, and ultimately concentration, of pesticide based on the known specific activity of the radiolabelled samples.

Cells were seeded in 96 well plates in complete culture medium at a density of 1.56×10^5 cells×cm^{-2} in 100 uL complete cell culture medium and incubated at 37 °C, 5% v/v CO_2. After 24 hours, cell culture medium was removed and replaced with 100 μL cell culture medium containing 2.5% v/v FBS, 0.3 % v/v DMSO and radiolabelled pesticide at a designated high or low concentration (these concentrations were different for each test compound). FBS content was reduced from 10% in order to maintain cell viability but limit proliferation. This was to ensure a roughly constant number of cells over the 48-hour observation period (supporting figure B2, appendix B). All solutions were filtered through a 0.22μm filter syringe to ensure sterility prior to addition to cells. After 3, 24 and 48 hours, 100 μL supernatant was removed and added to 4 mL scintillation fluid for analysis. Cells were then incubated with 50 μL trypsin-EDTA (0.1 % w/v) for 10 mins and then scraped with a pipette tip to ensure detachment. The trypsin solution and detached cells were transferred to 4 mL scintillation fluid for analysis. Finally, 200 μl DMSO was added to each well and incubated at room temperature to dissolve any pesticide adsorbed to the plastic. After 15 mins the DMSO was transferred to 4 mL scintillation fluid for analysis. Each pesticide was assessed using four individual samples for each treatment, at each timepoint. The experiment was repeated a second time with a different cell passage number. Throughout, the CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega) was performed on duplicate plates – containing unlabelled pesticides – to confirm cell population size remained relatively constant.

No toxicokinetic studies for these pesticides have been performed in mammalian cell culture in vitro previously, therefore the concentrations of pesticide used for this study were deduced from previous studies in fish. Relationships between pesticide log(P) and uptake by fish gill cells have been identified by Stadnicka-Michalak *et al.* (2014). These relationships were used to estimate appropriate nominal concentrations for these studies, albeit provisionally as this study used a mammalian cell line, requiring different conditions (culture medium, incubation temperature).

For each test compound, the aim was to estimate the nominal concentration that would result in the highest intracellular concentration predicted in the muscle *in vivo* and use this as the high concentration treatment. Preliminary experiments had indicated that the dose levels in these experiments were not sufficient to result in cytotoxicity.

4.4.7 *In Vitro* Toxicokinetics Model and Reverse Dosing

To model *in vitro* toxicokinetics, a three-compartment model based on Stadnicka-Michalak *et al.* (2014) was used, modelling pesticide concentration in culture medium, in cells and on plastic over time. The model comprised the following equations:

$$\frac{dC_{medium}}{dt} = (V_{cell}/V_{medium})\,k_{out\_cell}C_{cell} + (A_{plastic}/V_{medium})\,k_{out\_plastic}C_{plastic} \tag{4.1}$$

$$- (V_{cell}/V_{medium})\,k_{in\_cell}C_{medium} - k_{in\_plastic}C_{medium}$$

$$\frac{dC_{cell}}{dt} = k_{in\_cell}C_{medium} - k_{out\_cell}C_{cell} \tag{4.2}$$

$$\frac{dC_{plastic}}{dt} = (V_{medium}/A_{plastic})\,k_{in\_plastic}C_{medium} - k_{out\_plastic}C_{plastic} \tag{4.3}$$

Where C_x is concentration, with the subscript denoting the compartment (mass×cm^{-3} C_{medium} & C_{cell} and mass×cm^{-2} for $C_{plastic}$), t is time, k_{in_x} and k_{out_x} are the rate constants of uptake and elimination in each compartment (time^{-1}), V_x is volume of a compartment (cm^3) and A_x is surface area of a compartment (cm^2).

As pesticide moves into the medium from the cells, or vice versa, the change to the C_{medium} is equal to the corresponding change to C_{cell} multiplied by V_{cell}/V_{medium} (dimensionless ratio). This accounts for the different volumes of the compartments relative to the amount of pesticide exchanged. Similarly, in order to model uptake from medium on to the plastic surface, the change in $C_{plastic}$ is multiplied by $V_{med}/A_{plastic}$ (cm^3/cm^2 = cm). This accounts for the relative sizes of the two compartments and performs the necessary unit conversion from mass×cm^{-3}×t^{-1} to mass×cm^{-2}×t^{-1}. Likewise, as the test compound is eliminated from the plastic into the medium the change in C_{medium} is multiplied by $A_{plastic}/V_{med}$ (cm^2/cm^3 = cm^{-1}) which performs the opposite conversion. The volume or surface area of the respective compartments were known, as were the starting concentrations. The rate constants were fitted to the first LSC dataset. As there was some variation in starting concentration between repeats, the data from the repeated LSC experiments were used to test model predictions with fixed parameters.

Using the calibrated model parameters, it was then possible to calculate the initial value of C_{medium} to estimate the nominal concentration required to result in a desired intracellular concentration (C_{cell}) at equilibrium. For each pesticide the nominal concentration required to match the highest intracellular concentration predicted *in vivo* was estimated as the reference dose to investigate the effects on cell population growth.

4.4.8 Cell proliferation and cytotoxicity experiments

Cells were seeded at 3.1×10^4 cells×cm^{-2} (10^4 cells per well) in 100 μL cell culture medium and incubated at 37 °C, 5% v/v CO_2. After 24 hours, medium was removed and replaced with complete cell culture medium containing 10 % v/v FBS and 0.3 % v/v DMSO, and 0.5, 1, 2.5, 5 or 10x the reference dose for each pesticide, identified by reverse dosing. Cell proliferation and percentage mortality were assessed after 24, 48 and 72-hour incubation times with pesticide solutions (n=5).

Cell proliferation was assessed using the CellTiter 96® AQueous One Solution Cell Proliferation Assay. In this colorimetric method, the tetrazolium compound (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium, or MTS) is metabolised by viable cells to a coloured formazan product. The amount of formazan produced in a given time period is quantified as absorbance at 490nm with a microplate reader. This absorbance is often assumed to increase linearly with viable cell count. However, the metabolic rate of cells can be affected by test compounds, so no such relationship was assumed here. Instead, the assay was used to identify proportional effects on overall mitochondrial activity between treatments, as an indicator of effects on cell proliferation. The assay was conducted in accordance with the manufacturer's protocol. Controls in place to identify anomalous background absorbance were medium without cells and untreated cells lysed with 0.5% v/v Triton X-100 (n ≥ 3).

Significant effects of pesticide treatment on mitochondrial activity - relative to the vehicle control treatment - were identified by analysis of variance (ANOVA) and pairwise comparison (Tukey's honest significance test) at each timepoint. Linear modelling was used to identify significant relationships between initial pesticide concentration in the medium and mitochondrial activity.

Cytotoxicity was monitored using the The CytoTox-ONE™ Homogeneous Membrane Integrity Assay (Promega). This fluorometric assay quantifies the proportion of non-viable cells in the well by measuring the release of lactate dehydrogenase (LDH) from cells with compromised membrane integrity. The enzymatic reaction results in the conversion of resazurin into resorufin which can be measured by a microplate reader as fluorescence with an excitation wavelength of 560nm and an emission wavelength of 590nm. The assay was carried out in line with the manufacturer's protocol. As an additional control measure, for each experimental treatment, cells in three wells were lysed with 0.5% v/v Triton X- 100 to establish fluorescence corresponding to

100% cell mortality while accounting for differences in cell number between treatments (Smith et al., 2011). This allowed cell mortality to be calculated as a percentage using equation 4.4.

$$\% \; Mortality \; = 100 \; \times \left[\frac{(Experimental \; Fluroesence - Medium \; Background)}{(100\% \; Mortality \; - \; Medium \; Background)} \right] \tag{4.4}$$

Results were analysed using the bootstrapping method of Smith et al. (2011). For each treatment, at each timepoint, percentage mortality was computed by entering every combination of experimental (n=5), medium background (n=3) and 100% mortality wells (n=3) into equation 4.4, producing a sample of 45 values. Mean percentage mortality was calculated as the mean of these 45 values. Next, the sample was used to generate 200 random bootstrap samples (n=5, the number of experimental wells per treatment), from which 95% confidence intervals were derived. If the lower confidence interval of an experimental treatment was both positive and higher than the upper confidence interval of the controls, this was deemed a significant effect on cell mortality.

4.5 Results

4.5.1 *In vitro* toxicokinetics

Initial pesticide concentration in the medium was determined as the mean total recovered at the 3h timepoint - rather than from samples taken immediately after medium preparation - to control for sample loss due to additional pipetting. Generally, pesticide concentration in medium, cells and on plastic quickly reached equilibrium and modelled uptake matched well with observations (figure 4.2).

Due to low sample size (n=4 per treatment per timepoint), quantitative analysis of the model predictions by treatment would be of limited value. By normalising data as a percentage of the initial dose (the mean total recovered at 3h in each treatment) recovered from each compartment at each time point, all treatments could be aggregated into a larger dataset. This was a suitable measure by which to assess model performance, as the model parameters - which were constant across treatments - simply predict proportional uptake at equilibrium (determined by ratio of rate constants) and speed of uptake (determined by magnitude of rate constants). The purpose of the TK model in this study was to determine the initial conditions required to achieve a given intracellular concentration at equilibrium. As such, the 24h and 48h timepoints (at which all models had reached equilibrium) were of greatest relevance.

Predictions at equilibrium were compared to the mean recovery of each compound by compartment across treatments at 24h and 48h (table 4.1). Uptake by cells was predicted to within one standard deviation of the overall mean at both timepoints for thiamethoxam, prosulfuron & acibenzolar-S-methyl and at 24h for azoxystrobin and fludioxonil. For both fludioxonil and azoxystrobin, uptake by cells was underpredicted at 48h, suggesting that these compounds exhibited slower kinetics and had not fully equilibrated by this point. Due to practical constraints, it was not possible to verify this with a longer observation period, however, in both cases it appeared that uptake had slowed and that substantial increases in intracellular concentration were unlikely.

Mean percentage sorption to plastic was predicted to within one standard deviation at both timepoints for thiamethoxam, prosulfuron and fludioxonil, at 24h for acibenzolar-S-methyl and at neither timepoint for azoxystrobin. Sorption of azoxystrobin was predicted to within 0.6% of the observed mean at both timepoints but, due to low variability in the data, still differed by more than 1 standard deviation. Recovery of acibenzolar-S-methyl from plastic decreased with time (figure 4.2 & table 4.1) suggesting that soaking in DMSO may have been insufficient to remove sample adsorbed to the plastic after longer periods.
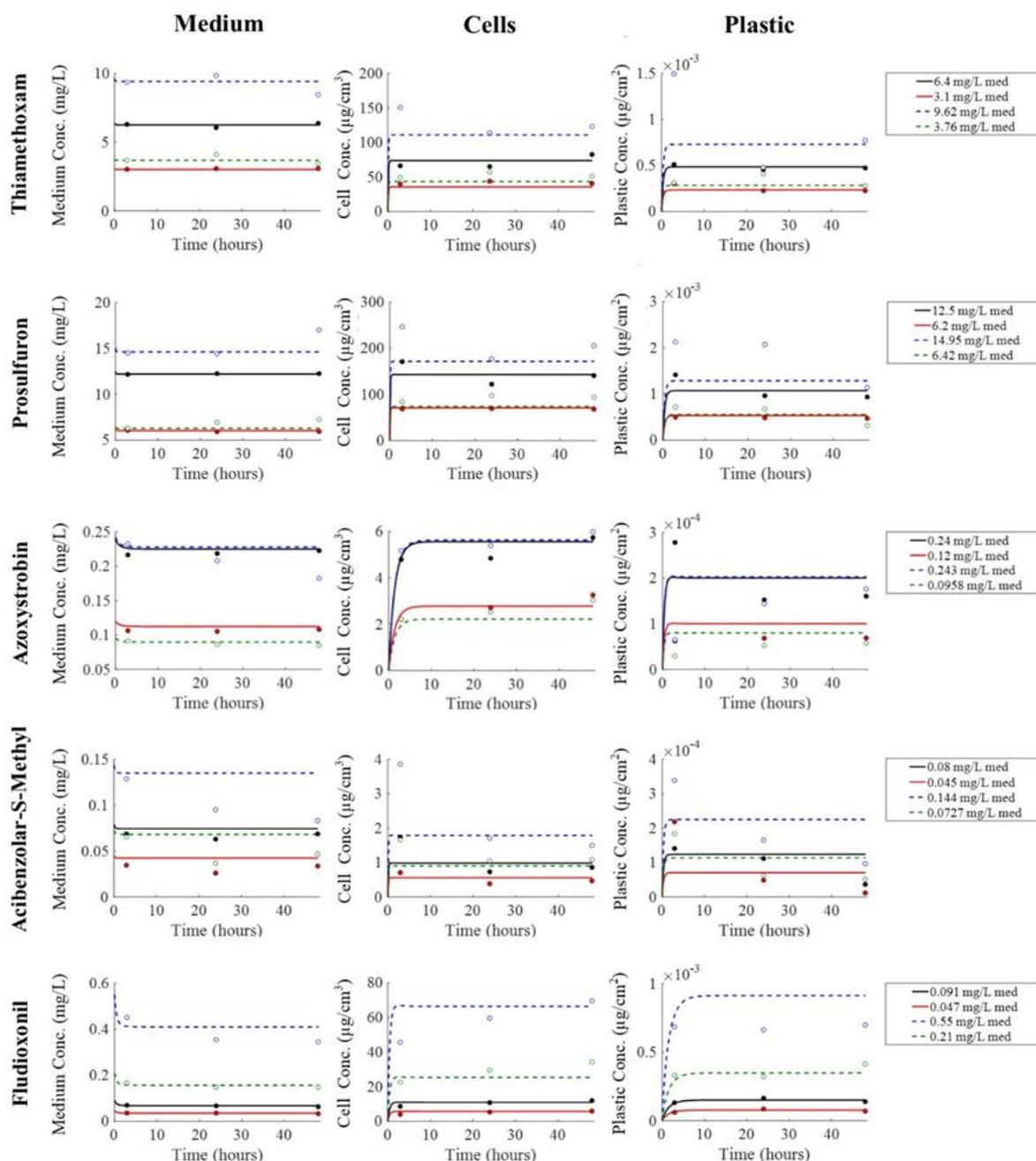
Figure 4.2 *Model outputs showing observed (circles) and predicted (lines) concentration in medium, cells and on plastic over time. Solid lines and circles denote treatments used for model calibration while dashed lines and hollow circles denote treatments used for validation.*

Pesticide retention in the medium was predicted to within one standard deviation of the overall mean at 24h and 48h for all compounds except acibenzolar-S-methyl. Acibenzolar-S-methyl was the only compound for which losses of over 5% (as evaporation or otherwise unrecovered) were observed. This was the most volatile of the compounds tested and was also the second most lipophilic. In all treatments, medium concentration dropped below model predictions, as the model assumes all of the initial dose remains in the system. However, losses were smaller in the

treatments used for model calibration, and predictions of uptake by cells were not adversely affected.

Table 4.1 *A summary of observed and predicted recovery of each compound at the 24h and 48h timepoints, normalised as a percentage of the mean total recovered at 3h in each treatment. The overall mean is based on all data points from all treatments (n=16), also shown are the lowest and highest mean uptake in a single treatment (n=4). Colour coding denotes whether model predictions were within one standard deviation of the overall mean at both (green), one (yellow) or neither (red) of the timepoints.*

| | | Percentage Recovery by Compartment | | | | | | |
| | | 24hr | | | 48hr | | | Model Prediction at Equilibrium |
Chemical	Compartment	Overall Mean ± s.d.	Lowest Group Mean	Highest Group Mean	Overall Mean ± s.d.	Lowest Group Mean	Highest Group Mean	
Thiamethoxam	Medium	100.99±8.93	94.31	108.63	94.15±13.01	87.47	99.36	97.81
	Cells	2.26±0.48	1.79	2.67	2.30±0.55	2.25	2.38	2.04
	Plastic	0.19±0.089	0.12	0.27	0.19±0.082	0.18	0.20	0.19
Prosulfuron	Medium	99.28±8.06	95.54	107.30	105.03±10.42	95.79	113.62	97.75
	Cells	2.12±0.53	1.73	2.68	2.22±0.48	1.93	2.57	2.03
	Plastic	0.25±0.15	0.19	0.35	0.17±0.11	0.12	0.19	0.22
Azoxystrobin	Medium	90.87±8.67	85.57	94.20	89.03±11.13	74.98	96.70	93.75
	Cells	4.14±0.67	3.69	4.64	4.86±0.67	4.35	5.58	4.10
	Plastic	1.52±0.35	1.37	1.66	1.66±0.22	1.51	1.83	2.11
Acibenzolar-S-Methyl	Medium	65.60±15.55	50.22	83.58	73.43±15.87	57.61	91.18	93.85
	Cells	1.99±0.54	1.62	2.54	2.11±0.47	1.83	2.62	2.20
	Plastic	2.93±1.40	2.13	3.74	1.33±0.66	0.67	1.77	3.96
Fludioxonil	Medium	72.17±7.39	64.39	80.39	69.25±6.33	62.71	73.82	74.50
	Cells	21.61±2.75	19.18	24.32	24.59±2.69	22.39	28.19	21.34
	Plastic	4.13±1.00	3.06	4.93	4.03±0.96	3.22	4.88	4.20

The percentage uptake by cells and plastic increased with hydrophobicity (figure 4.3). Both were strongly (r = 0.98 and 0.87 respectively) and significantly (p<0.0001) correlated with octanol: water partition coefficient, *P*. A more comprehensive data set may yield a predictive relationship between *P* and TK parameters. However, with only five compounds in total and substantial losses of one of those, any relationship derived from this data set would be questionable.
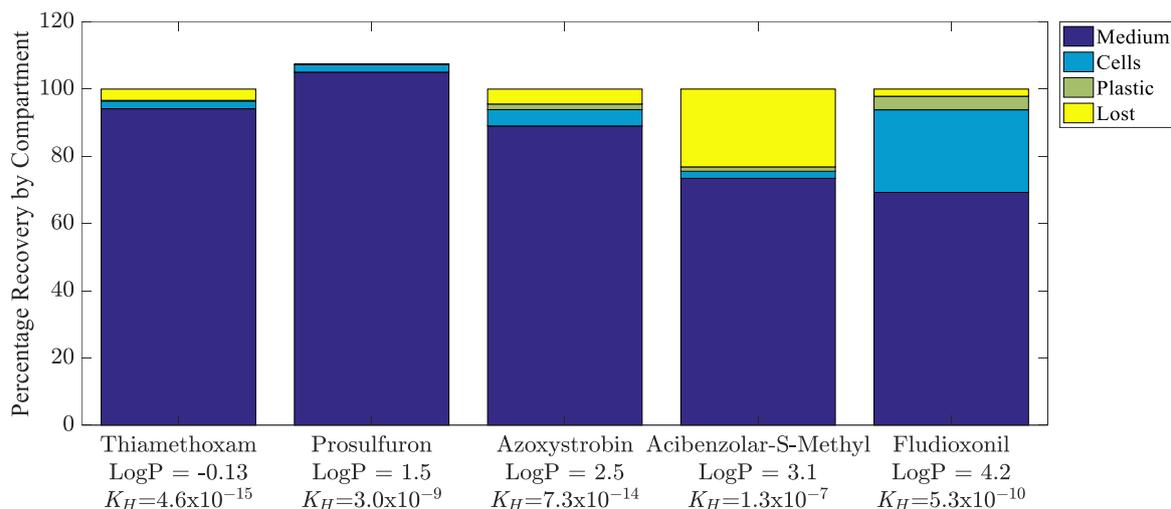
Figure 4.3 *Percentage recovery by compartment for each compound at 48h. Results shown are an average from all treatments normalised as percentages of the total recovered at 3h in each treatment. In one case the total recovered at 48h exceeded that recovered at 3h. Henry's law constant, K_H, is given in atm x m³ x mol⁻¹ at 25°C.*

4.5.2 Reverse dosing

The calibrated models were then used to determine the nominal concentration required to result in a desired intracellular concentration. For each pesticide, the target concentration was the average predicted in the highest dose group in *in vivo* dietary toxicity testing, using the models developed in Martin *et al.* (2019). The model parameters and experimental conditions derived from reverse dosing are shown in table 4.2.

Table 4.2 *Summary of reverse dosing. The calibrated absorption and elimination rate constants for each compound are represented in the form k_{in_x} and k_{out_x} respectively, with x denoting the relevant model compartment. The target intracellular concentration and the medium concentration required to achieve this, as calculated by the in vitro TK model, are also shown.*

Compound	k_{in_cell} (h⁻¹)	k_{out_cell} (h⁻¹)	$k_{in\_plastic}$ (h⁻¹)	$k_{out\_plastic}$ (h⁻¹)	Target Intracellular Conc. (µg/g)	Required nominal medium concentration (mg/L)
Thiamethoxam	113.9	9.676	0.005532	2.827	62.70	5.443
Prosulfuron	128.0	10.91	0.005374	2.419	213.7	18.63
Azoxystrobin	16.71	0.6770	0.05167	2.295	45.57	1.998
Acibenzolar-S-Methyl	111.6	8.428	0.1277	3.027	4.583	0.3702
Fludioxonil	277.5	1.715	0.03200	0.5675	26.02	0.2155

4.5.3 Effects on cell proliferation

Linear modelling showed a significant negative relationship between concentration of acibenzolar-S-methyl and mean mitochondrial activity of cells after 24 hours exposure (p =

0.0013). However, this pattern did not persist beyond this timepoint. The only other significant relationships between chemical concentration and mitochondrial activity were positive. These were thiamethoxam (p = 0.012) and fludioxonil (p = 0.049) at 48 hours exposure. ANOVA and pairwise comparison showed that after 72 hours exposure, mitochondrial activity was not significantly reduced, relative to the vehicle control treatment, by any dose of any of the compounds tested. As such, these data were not suitable for extrapolating effects on growth *in vivo* (see 'Discussion' for more details).
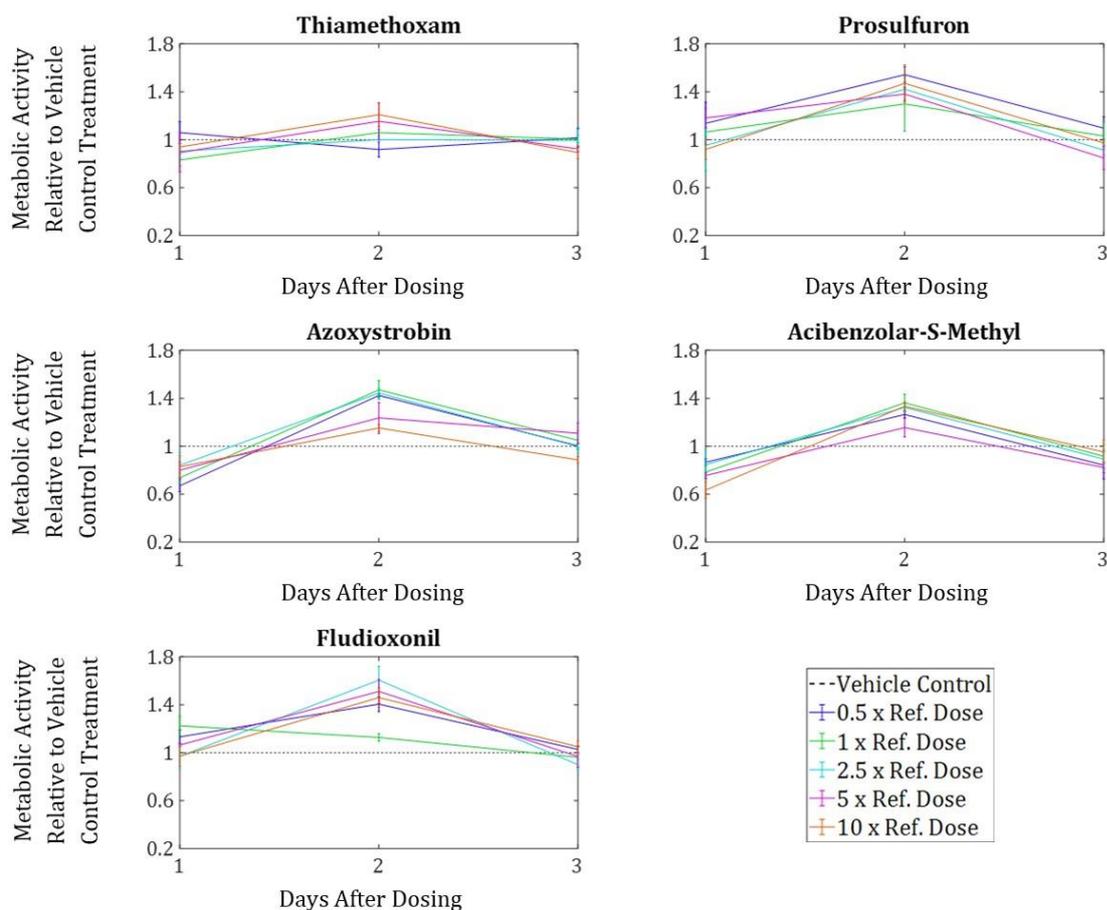
Figure 4.4 *Mitochondrial activity (determined by CellTiter 96® AQueous One Solution Cell Proliferation Assay) normalised against vehicle control treatment at each time point, over 72 hours exposure to pesticides. Lines represent the mean for each treatment (n=5), with error bars showing 95% confidence intervals.*

4.5.4 Effects on cell mortality

After 24 hours exposure, anomalous background fluorescence (untreated medium, no cells) exceeded that of several experimental treatments. This resulted in equation 4.4 producing theoretically impossible negative values of percentage mortality for most treatments, including controls. Additionally, the fluorescence of maximum LDH release wells also was low relative to background at this timepoint (due to cell number still being low shortly after seeding), meaning that effects were amplified. Estimated mortality at this time point ranged from −24.4% to 24.8%,

with −23.6% estimated in the vehicle control treatment. For this reason, the results at this timepoint were not considered reliable and were excluded from further analyses. Significant increases in mortality were recorded in two azoxystrobin treatments at 48 hours (2.5x and 10x the reference dose). Mortality was also significantly increased in the highest prosulfuron treatment and two fludioxonil treatments (2.5x and 5x the reference dose) at 72 hours only. The highest average mortality was recorded at 72 hours exposure in the highest prosulfuron dose group, with a mean value of 13.5%.

4.6 Discussion

4.6.1 *In Vitro* TK Modelling

Achieving specific pesticide concentrations in the medium was challenging due to unpredictable loss of pesticides during sterile filtering. Samples were measured before and after filtration and losses were quite different between experimental repeats. The second experimental run was intended to assess the repeatability of results and increase sample size but in most cases the achieved nominal concentrations were substantially different from the first run. However, as concentration was measured in all cases, this was not a problem for the aims of the study. In fact, it increased the range over which model predictions could be tested.

For all compounds, the three-compartment model used was able to predict uptake by cells with reasonable accuracy. However, in the cases of azoxystrobin and fludioxonil the intracellular concentration at equilibrium may have been underestimated as the data suggested that these compounds had not fully equilibrated after 48 hours. Sorption to plastic and retention in the medium were also generally well predicted by the model. The most notable exception was acibenzolar-S-methyl for which inaccurate predictions appeared to be caused by loss of the test compound, most likely due to evaporation. The TK model could be adapted to consider evaporation (Stadnicka-Michalak et al., 2014) but, as intracellular concentration was accurately predicted across treatments, this was not deemed necessary for the purposes of this study.

Intracellular concentration at equilibrium was the endpoint of greatest interest in this study. Model predictions of this endpoint result from the ratio of absorption and elimination rate constants. However, it is the absolute values of these model parameters that determine the time taken for equilibrium to be reached. Azoxystrobin and fludioxonil showed slower kinetics but for the other three compounds, equilibrium appeared to have been reached by the first data point at three hours after exposure. Due to practical constraints, it was not possible to collect data at multiple observation times shortly after dosing or beyond 48 hours after exposure. Both would have been desirable in order to improve the accuracy, particularly in terms of absolute value, of fitted parameter values and this should be considered in the design of future kinetics studies.

Another consideration for future studies is the role of FBS concentration. The concentration of FBS in the medium was reduced to 2.5% during TK experiments in order to minimise cell proliferation, however, it has been shown that increased FBS concentration can reduce the free concentration of test compounds and, therefore, uptake by cells (Fischer et al., 2018). In fact, this process of serum mediated passive dosing stabilizes cell dose and compensates for chemical depletion as a result of cell proliferation. It would have been preferable then to have kept FBS concentration constant at 10% for all experiments. As a result of changing the FBS concentration,

intracellular concentration was likely overpredicted in the cell proliferation experiments (at 10% v/v FBS). However, none of the relationships between cellular uptake and FBS concentration reported by Fischer et al. (2018) were linear, a twentyfold (0.5% to 10% v/v) increase in FBS concentration led to a maximum 4.5-fold reduction in cellular uptake. Therefore, at the very least, it is likely that >2.5 times the intracellular concentration predicted *in vivo* was achieved in our cell proliferation experiments.

The compounds in this study were not selected for their range of chemical properties but because *in vivo* experimental data for these pesticides had previously been made available for the development of a TK-TD model (Martin et al., 2019). However, as was expected based on previous studies (Stadnicka-Michalak et al., 2014), cellular uptake and sorption to plastic were correlated with octanol: water partition coefficient, *P.* This suggests it may be possible to estimate model parameters from chemical properties in future. A similar study including more compounds and a broader spectrum of chemical properties could uncover significant patterns from which parameter values could be accurately predicted without further experimental work. As well as hydrophobicity, volatility should also be considered. A previous study found this posed no issues for compounds with low volatility (K_H < 1 x 10^{-6} atm x m^3 x mol^{-1}) in culture with fish cells at 25°C (Stadnicka-Michalak et al., 2014) however, evaporation of acibenzolar-S-methyl (K_H =1.3 x 10^{-7} atm x m^3 x mol^{-1} at 25°C, (PPDB, 2019a)) did occur within our test system where mammalian cells are cultured at 37°C. This may place limits on the estimation of parameter values from chemical properties and highlights the influence of temperature on *in vitro* toxicokinetics.

Finally, because culture medium needed to be re-sterilised after the radiolabelled samples were added, it was difficult to achieve a specific nominal concentration. To avoid this issue in future, it would be preferable to have a sterile workspace designated for dispensation of radioactive stocks. However, as previously mentioned, this was not a major problem because concentration was still measured with LSC in this study.

4.6.2 Cell proliferation and mortality

Cells were cultured for 72 hours in medium containing up to 10x the reference dose – identified by reverse dosing – of each pesticide. Due to culture medium containing 10 % v/v FBS in these experiments, the intracellular concentrations achieved were probably lower than expected (Fischer et al., 2018). However, in the highest dose groups, intracellular concentration was likely >2.5x the highest concentrations predicted *in vivo* in dietary toxicity studies (associated with body weight reductions of >10% (Martin et al., 2019)). At only one time point, for one compound (acibenzolar-S-methyl at 24 hours), was a significant negative relationship between pesticide concentration and mean mitochondrial activity found. After 72 hours exposure, mitochondrial

activity was not significantly inhibited by any pesticide treatment, indicating no effects on cell proliferation.

In some treatments, significantly increased cell mortality was observed. For data to be suitable for extrapolation, any observed effects must be significant, consistent over time and exhibit a dose related response so that they can be attributed to the presence of the test compound. Additionally, doses should be below the threshold for effects on mortality so that cytotoxicity can be discounted from effects on cell population growth (Stadnicka-Michalak et al., 2015). No chemical in this study met these criteria and as such it was not possible to use these data to predict effects on *in vivo* growth.

4.6.3 Obstacles to QIVIVE

There are several possible factors that may underlie the negative results of the cell proliferation experiments. First is simply that the sensitivity of the chosen cell line to the compounds in this study did not represent that of the whole animal. The rationale for selecting of the L6 cell line was twofold. As muscle accounts for 35.3% of rat body mass, the largest proportion of any tissue (Lindstedt and Schaeffer, 2002), it was deemed the most appropriate tissue to investigate chemical effects on growth. Additionally, with a lack of information as to the role of metabolites in the toxic mode of action of each pesticide, it was assumed that effects are caused by parent compounds. As such, a cell line with limited propensity for biotransformation was desirable in order to test this assumption. In past studies using the L6 cell line, concentration dependent responses to known toxins have been demonstrated using the same assays employed in this study (Kalam et al., 2011, Cybulski et al., 2015). However, these studies were investigating large effects on cell viability rather than simply reduced population growth. This would agree with previous observations of low absolute sensitivity in cultured cells, possibly the result of few target sites being present in a single cell type relative to an organism (Schirmer, 2006, Groothuis et al., 2015). The sensitivity of L6 cells relative to other cell lines varies in the literature, depending on the test compound (Kalam et al., 2011, Cybulski et al., 2015). However, unlike other cell lines, L6 did show higher sensitivity to membrane damage by certain chemicals than reduced mitochondrial activity (Cybulski et al., 2015) which could present an obstacle for separating effects on cell proliferation and cell death. Many mammalian cell lines are available and naturally this choice will impact on the observed responses to chemicals. As was noted by Stadnicka et al. (2015), one advantage of the RT-gill-W1 is that it is derived from healthy gill cells whereas mammalian cell lines, such as L6 (Yaffe, 1968), are often established by exposure to carcinogens. The ES-D3 cell line, derived from embryos of *Mus musculus,* could be a suitable candidate for future studies. This cell line was not cancer derived (Doetschman et al., 1985) and has produced promising results in the past (Li et al., 2017, Louisse et al., 2015). While biological rationale can be applied in the selection of a cell

line, any issues will not become apparent before experiments begin and finding a cell line with the desired attributes could be a time-consuming task involving a degree of trial and error.

A related issue is that metabolites of the compounds may contribute significantly to the toxicity of the compounds *in vivo* (Pirmohamed et al., 1994) and these metabolic pathways are not necessarily replicated *in vitro* (Coecke et al., 2006). For example, thiamethoxam is mostly excreted untransformed but its main metabolite, clothianidin (FAO and WHO, 2011), is more toxic to mammals than the parent compound (PPDB, 2019g, PPDB, 2021). For the fitting of TK model parameters, total radioactivity was converted to concentrations on the assumption it corresponded to the presence of the parent compound, without accounting for biotransformation. One reason that a muscle cell line was selected over tissues involved in biotransformation, such as the liver, was in order to meet this assumption. Previous studies have also made this assumption (Stadnicka-Michalak et al., 2015, Li et al., 2017) yet found close relationships between dose-response predictions based on *in vitro* data and observed *in vivo* data. The assumption may be necessary because proposed metabolic pathways would be impossible to replicate *in vitro*. For example, of the compounds in this study, prosulfuron, azoxystrobin and fludioxonil are broken down into at least 14, 15 and 20 metabolites respectively, with significant differences observed between the sexes (European Commission, 2013c, European Commission, 2009, European Commission, 2006). Models of *in vivo* toxicokinetics would need to predict the uptake of the parent compound, biotransformation into toxic metabolites, and elimination the metabolites and parent compound. If impacts on growth are not predictable without accurately accounting for biotransformation *in vitro,* this method may not be a viable alternative to animal testing due to requirements for detailed data on *in vivo* metabolism and because of the number of chemicals with complex metabolic pathways.

It has been demonstrated that observed effects on fish growth can be predicted by replicating the intracellular toxicant concentration and using *in vitro* data to calibrate a growth model before extrapolating through time (Stadnicka-Michalak et al., 2015). However, it may not be possible to directly predict mammalian growth in this way due to the relative size of observed effects in toxicity testing using fish versus mammals. The highest *in vivo* doses of cyproconazole and propiconazole caused average body weight reductions in fish of 49% and 77% respectively, relative to controls (Stadnicka-Michalak et al., 2015). In mammalian chronic toxicity studies, body weight reductions relative to controls rarely reach 20% by the end of testing (Wang et al., 2019). Logically, prediction of smaller effects *in vivo* requires the detection of smaller effects *in vitro,* which can be challenging due to variability in data. It may instead be necessary to induce larger effects *in vitro* before extrapolating to realistic intracellular concentrations or effect sizes and investigating the relationships between predictions and *in vivo* data. This was part of the rationale

for aiming to achieve 10 times the intracellular concentration predicted *in vivo* however, this was not high enough to induce significant effects. Another consideration is that cells are neither male nor female yet chemical effects on endpoints such as body weight can vary greatly between the sexes in animals (Calabrese, 1986, Martin et al., 2019). Such differences will be extremely challenging to predict from *in vitro* data so it will be necessary to account for this uncertainty when interpreting any future extrapolations from *in vitro* data.

Finally, and perhaps most crucially, it may be that the *in vivo* TK predictions, from which target concentrations were derived, were not sufficiently accurate. Although the *in vivo* model had been used to produce accurate TK-TD predictions, the requirements for QIVIVE are different. A TK-TD model only requires that predicted internal concentration is proportional to the actual value at the target site (e.g. 'scaled internal concentration' in the most simplified TK-TD models), as the fitting of TD parameters will compensate for the difference (Jager and Zimmer, 2012). To compare *in vitro* and *in vivo* effects however, it is important that the intracellular concentrations actually match (Stadnicka-Michalak et al., 2015). Predictions of chemical concentration in muscle tissue were derived from simple observed correlations with blood concentration. This may appear a simplistic method but, at equilibrium, more complex physiologically based TK models also predict that chemical concentrations in different tissues are simply multiples of one another (Li et al., 2017, Louisse et al., 2015). No matter how simple or realistic the TK model, predictions are only as good as the data used for calibration and this is an area that could be improved. Animals in regulatory toxicokinetics studies are generally given a single gavage dose (OECD, 2010). Whilst this does represent an oral exposure, it is not representative of the dietary dosing used in longer term toxicity studies (OECD, 2001, OECD, 1998, OECD, 2008). Toxicokinetics may vary substantially between gavage and dietary dose groups (Foster et al., 2015, Vandenberg et al., 2014, Kapetanovic et al., 2006). Additionally, equilibrium is never reached in single dose treatments because the intake is not sustained. A major advantage of the testing framework for fish is that the route of exposure is consistent throughout testing with animals exposed to a constant toxicant concentration in their water (except in the case of hydrophobic compounds for which dietary exposure is more relevant) (OECD, 2012, OECD, 2013, OECD, 1984, OECD, 2019, OECD, 2014). The inclusion of dietary exposure in mammalian toxicokinetics studies would be the surest way of improving the accuracy of TK modelling, which will be crucial if the methodology is to help replace animal testing in future.

4.6.4 Future Work

While it was not possible to extrapolate *in vivo* effects on growth from *in vitro* data, this study demonstrates how, in principle, TK-TD models can act as a bridge between *in vitro* and *in vivo* data from mammalian dietary toxicity studies.

Before further attempts are made to develop QIVIVE methods using TK-TD models it is worth evaluating the place of such methods within wider efforts to find alternatives to animal testing. The establishment of AOPs, which link molecular interactions to 'key events' which in turn cause adverse effects at higher levels of biological organisation, has received a great deal of focus at policy level (Willett, 2019). At lower levels of organisation, predictive QSARs and *in vitro* methods are used to highlight a compound's propensity to initiate key events. However, at higher levels, *in vivo* assays are still required. At present, the information gained from AOPs is generally qualitative and more suited to identify hazards presented by a compound rather than quantifying the environmental or health risk it poses (Sewell et al., 2018). A need for quantifying thresholds for progression from one key event to the next has been identified as well as the need to account for level, duration and route of exposure in quantifying risk. It is for these purposes that TK-TD modelling is particularly well suited because TK-TD model parameters have a biological or chemical interpretation. Further, linking toxic responses at cellular level to responses at the organism level can be aided by models that allow parameterisation from bioassays at either level (Stadnicka-Michalak et al., 2015, Ashauer and Jager, 2018).

Previous studies using mammal cells have shown that benchmark doses, or BMDs, (i.e. the dose required to induce a given effect) can be predicted to within threefold of observed values using *in vitro* and *in silico* methods (Louisse et al., 2015, Li et al., 2017). However, BMDs are time dependent measures and their value to assessing risk has been questioned (Jager, 2011, Jager, 2012). As such, a method of predicting the degree to which graded endpoints, such as growth rate, are affected by a given level and duration of exposure would be an extremely valuable tool. Promising results with fish cells (Stadnicka-Michalak et al., 2015) suggest this may be possible but there is still some way to go in order to replicate that success for mammals.

Reliable *in vivo* TK-TD models can be developed using existing datasets from regulatory toxicity studies on mammals but these do not require precise TK predictions. TK models that can accurately predict internal concentration will be required. For this to be achieved, more relevant data with which to calibrate and validate models must be collected. While this would represent an increase in the number of animals used in toxicokinetics studies, such studies have small sample sizes (OECD, 2010) and the resulting data could facilitate major progress towards alternative methods in future. Models have been developed which can be parameterised using without *in vivo* data (Li et al., 2017, Louisse et al., 2015). However, their validation still requires animal data. Developing the capability to model kinetics following dietary exposure in rodents could fulfil a crucial role in improving the utility of non-animal methods in the AOP framework going forward.

Chapter 5 - Modelling effects of variability in feeding rate on growth – a vital step for DEB-TKTD modelling

5.1 Preface

One of the aims stated at the start of the project was to explore applications of TK-TD models in ecological risk assessment, such as extrapolating the effects of realistic pesticide exposure on wild rodents. However, despite some success in separating the effects of toxicity and feeding rate on the growth of rats, Chapter 3 highlighted important issues around the way model inputs were derived from feeding data. Without accurately modelling growth based on feeding rate, any inhibition that results from chemical toxicity cannot be differentiated from the effects of reduced feeding. Therefore toxicodynamic parameters cannot accurately reflect the toxicity of a chemical, limiting their value for extrapolation. Therefore, I felt that the final research chapter should return to the growth model and address the issues identified around feeding inputs.

In Chapter 3.6, an alternative method of deriving feeding inputs was briefly outlined. Chapter 5 presents a research paper in which this new approach was fully developed, and compared to the previously employed methods, both in terms of model accuracy and theoretical consistency.

5.2 Abstract

A major limitation of dietary toxicity studies on rodents is that food consumption often differs between treatments. The control treatment serves as a reference of how animals would have grown if not for the toxicant in their diet, but this comparison unavoidably conflates the effects of toxicity and feeding rate on body weight over time. A key advantage of toxicity models based on dynamic energy budget theory (DEB) is that chemical stress and food consumption are separate model inputs, so their effects on growth rate can be separated.

To reduce data requirements, DEB convention is to derive a simplified feeding input, f, from food availability; its value ranges from zero (starvation) to one (food available *ad libitum*). Data show that the feeding rate of rats does not scale with body size, contradicting DEB assumptions regarding. Relatively little work has focused on addressing this mismatch, but accurately modelling the effects of food intake on growth rate is essential for the effects of toxicity to be isolated. This can provide greater insight into the results of chronic toxicity studies, allowing accurate extrapolation of toxic effects from laboratory data.

Here we trial a new method for calculating f, based on the observed relationships between food consumption and body size in laboratory rats. We compare model results with those of the conventional DEB method and another method from a previous study. Our results showed that the new method improved model accuracy ($R^2 \geq 0.97$ when calibrated) while modelled reserve

dynamics closely followed observed body fat percentage over time (closest to independent data at six out of seven timepoints). The new method assumes that digestive efficiency increases with body size. Verifying this relationship through data collection would strengthen the basis of DEB theory and support the case for its use in ecological risk assessment.

5.3 Introduction

Mechanistic effects models (MEMs) aim to simulate the mechanisms by which chemicals affect individuals, populations and communities (Grimm and Martin, 2013). This is an appealing prospect with great potential for use in ecological risk assessment (ERA) of chemicals such as pesticides (Forbes et al., 2009, Forbes and Calow, 2012). Simulating underlying processes confers several advantages over traditional analysis of data from laboratory-based toxicity studies and extrapolations to field scenarios based on summary statistics. Mechanistic modelling enables the prediction of toxic effects in untested, ecologically relevant conditions. This can add ecological realism to extrapolations and potentially even reduce animal testing requirements (Jager et al., 2006).

Accounting for the mismatch in exposure between laboratory and field is a key obstacles to long term risk assessment of pesticides for mammals (Fischer, 2005). For example, in chronic toxicity testing of pesticides, rats are exposed to a constant concentration of test compound in their diet for up to two years (OECD, 2018b, OECD, 2018a, OECD, 2001). Such constant exposure is unrealistic in the field as pesticides are not applied at a constant rate all year round. This disparity can be addressed using toxicokinetic-toxicodynamic (TK-TD) models (Jager et al., 2006). These are a class of MEMs that work at the individual level, predicting an internal measure of chemical concentration over time (toxicokinetics) and the stress this places on an organism (toxicodynamics). As such, the effects on a given endpoint resulting from realistic, time varied exposure can be predicted (Nyman et al., 2012).

The use of TK-TD modelling has now been recommended for certain regulatory purposes, such as predicting survival of aquatic organisms (EFSA, 2018). However, for birds and mammals sublethal effects are most relevant at realistic exposure levels, as no mortality associated with pesticide use is accepted under European regulations (EFSA, 2009b). The 'DEBtox' or 'DEB-TKTD' modelling framework (Kooijman and Bedaux, 1996b, Kooijman and Bedaux, 1996a, Sherborne et al., 2020), combining TK-TD modelling with the Dynamic Energy Budget (DEB) theory (Kooijman, 2000) provides a means of predicting sublethal toxic effects. DEB is an established metabolic theory, mathematically describing the processes of energy acquisition and allocation to predict endpoints such as body size and reproductive output. DEB has been applied to a wide range of taxa, with parameters available in the Add My Pet (AmP) library (Marques et al., 2018). The

majority of DEB-TKTD studies thus far have focused on invertebrates (Ashauer and Jager, 2018) and more recently fish (Zimmer et al., 2018, Sadoul et al., 2018), with very few studies concerning terrestrial vertebrates (Martin et al., 2019, Desforges et al., 2017).

A particular advantage of DEB-TKTD modelling is the ability to separate the effects of feeding rate and toxic action on growth rate. This is particularly important in dietary toxicity studies, where ingested dose is directly related to feeding rate. This property is relatively unexplored but, in fact, it is crucial for TD models to accurately reflect toxicity and therefore to be of use for extrapolation to novel scenarios. Temporal and inter-treatment variability in feeding rate is a crucial driver of observed growth so any observed effects on body weight cannot simply be attributed to toxic action. Moreover, the extent to which a compound induces feeding avoidance may increase or decrease the risk posed to wildlife, depending on whether animals would have a choice of food items in the field scenario (Thompson, 2007). A previous study (Martin et al., 2019) developed methods to account for variability in feeding while modelling the effects of dietary toxicity on growth of domestic laboratory rats (*Rattus norvegicus*). However, some important issues with these methods were identified as areas for improvement in future.

In DEB theory, feeding rate is assumed to be limited by surface area (e.g. area of feeding appendages in filter feeders), which is proportional to body mass to the power 2/3. Where data are available, observed area specific feeding rate is divided by a maximum value so that it can be entered into the model as a dimensionless parameter ranging from zero to one (Jager et al., 2013, Kooijman, 2000). In Martin et al. (2019), we generated model inputs by scaling weekly area specific feeding rate in each treatment relative to the maximum observed rate within each dataset. While this was a logical approach, two major issues became apparent.

The first issue was that, in rats, area specific feeding rate decreases as animals grow (Laaksonen et al., 2013, Martin et al., 2019). As such, the scaled feeding rate entered into the model dropped well below one in the latter stages of growth. This meant that, according to model equations, animals could have grown to many times their maximum observed body weight if they had continued to feed at the maximum observed rate throughout their lifetime. While this was a theoretical rather than a practical issue, it must be addressed for models to realistically represent the processes involved in growth.

The second issue arose because area specific feeding rate was calculated relative to observed (rather than predicted) body size. This meant that when predictions differed from observed data, this could result in a positive feedback loop or 'snowball effect' (figure 5.1). For example, if a rat with area of 40cm^2 ate 20g food/day at time t this would be a feeding rate 0.5 g/cm^2/day. However, if predicted surface area at time t were larger than that observed, say 50cm^2, then 0.5

g/cm²/day would equate to 25g/day. Therefore, the modelled growth rate in the next time step would correspond to 25% higher food consumption than was observed, exacerbating the problem with each time step.
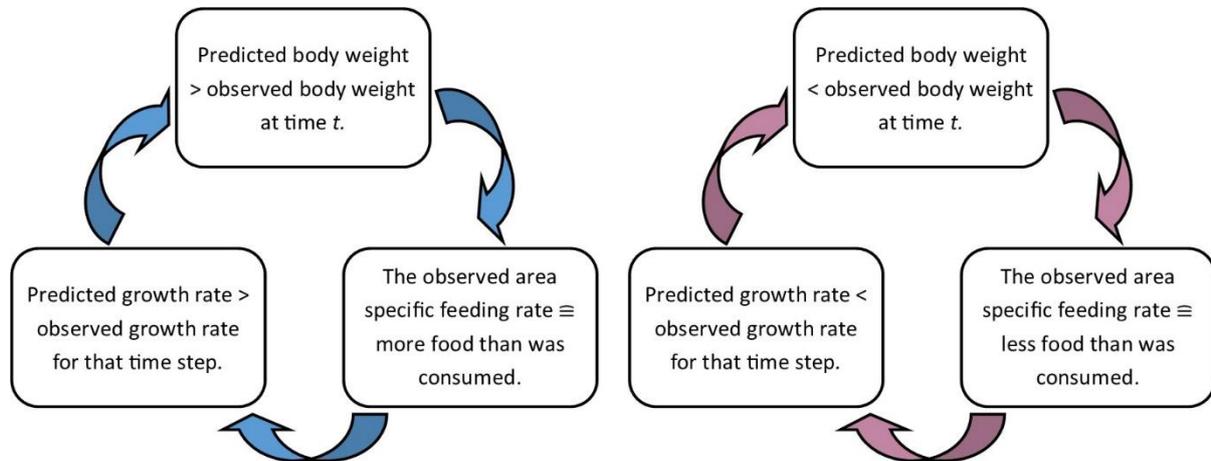
Figure 5.1 *Flowchart illustrating the positive feedback loop ('snowball effect') that can occur when area specific is calculated directly from data as in Martin et al. (2019).*

Here we investigate the potential of new methods to solve these issues and the implications for DEB theory. As suggested in Martin et al. (2019), we look to mathematically describe the relationship between feeding rate and body size in rats over the entire growth period and use this as a reference for scaling observed area specific feeding rate. We assess the resulting models from three standpoints: accuracy - how closely fitted models agreed with observed growth curves; generality – how well independent data are predicted without additional fitting; biological realism – how realistically the models simulate the processes underlying growth. We use a model based on the DEBkiss modelling framework (Jager et al., 2013) - a simplified version of DEB, following the same fundamental principles but with fewer parameters. It was desirable to prioritise model simplicity in this study. Firstly, because eliminating complex reserve dynamics from the model meant that the effects of different feeding inputs on model predictions could be more easily analysed. Reproduction was also omitted for this reason. Additionally, the lack of user-friendly modelling tools was recently identified as a barrier preventing the use of DEB-TKTD models by regulators (EFSA, 2018) which has prompted renewed interest in DEBkiss (Jager, 2020).

5.4 Methods

5.4.1 Data

All data used here were made available from existing regulatory studies (Syngenta, unpublished) required under 94/79/EC (European Commission, 1994), investigating chronic toxicity of acibenzolar-S-methyl, prosulfuron and thiamethoxam in Sprague Dawley laboratory rats (*Rattus norvegicus*) (Palm, 1975).

Chronic toxicity studies lasting two years were carried out according to OECD guidelines (OECD, 2018b, OECD, 2018a, OECD, 2001). Animals were kept in standard conditions with food and water available *ad libitum*. Each study comprised a control group and at least three dose groups with individual observations of body weight (g) initially at weekly intervals (later observations were up to five weeks apart). Food consumption ($g_{(food)} \times day^{-1}$) was recorded alongside body weight either individually or per cage (2-5 individuals), providing the average per animal per day. Sample size was initially 80 animals per treatment per sex and only data for unmated animals were included in this study.

5.4.1.1. Calibration dataset

The control group from the two-year dietary toxicity study of acibenzolar-S-methyl was selected as the calibration data in this study, as it was intermediate in terms of total food consumption for both sexes. This dataset, henceforth referred to as group A, comprised observations of an initial 80 animals of each sex at 37 timepoints over 104 weeks (a total of 2659 observations for males and 2678 for females). Animals in this study were fed a diet of Nafag 890 pelleted food.

5.4.1.2 Independent datasets

Independent datasets B and C were the control groups from the two-year dietary toxicity studies of prosulfuron and thiamethoxam respectively. Initial sample size in both datasets was 80 animals of each sex. Animals in group B were fed a diet of Rodent Chow #5002 pellets and observations took place at 37 timepoints over 104 weeks. Animals in group C were fed a diet of Nafag 890 pelleted food and observations took place at 36 timepoints over 103 weeks. Nafag 890 and Rodent Chow #5002 are similar in protein (18-20%), fat (3-4.5%) and energy content (12-14kJ/g) although Nafag 890 is substantially higher in fibre (Ruhlen et al., 2011, Leonhardt and Langhans, 2002, Silberbauer et al., 2000).

5.4.2 Theoretical basis of the bioenergetic model

To simulate rat growth, we used a slightly modified version of the DEBkiss modelling framework (Jager et al., 2013). The model in this study employs the most basic rules for starvation and the storage of assimilates. The reason for this choice was to determine how accurately growth can be

predicted using simple equations, if feeding data are used to produce accurate and high-resolution model inputs.

All DEB (Kooijman, 2000) models are based on the principle that certain processes are limited by volume or surface area and that an animal's length, surface area and volume scale such that $Volume \propto Length^3$ and $Surface\ Area \propto Volume^{2/3} \propto Length^2$, provided body shape remains the same (isomorphic growth). In DEBkiss (Jager et al., 2013), an animal's total wet weight, W_w, is divided into structural (bones, muscle, organs etc.) weight, W_V, and stored assimilates known as the reproduction buffer, W_R. DEBkiss was developed with invertebrates in mind and so the reproduction buffer is generally meant to provide mass for egg production. This function does not apply to viviparous mammals, for which the main costs of reproduction – foetal development and milk production- only occur post-fertilisation. Since reproduction is not modelled in this study, this buffer is used only for its secondary function, as a reserve to maintain structure in times of starvation. As such, it is more intuitive to think of W_R simply as mass of 'reserve' and it will be referred to as such throughout. In this model implementation, the term reserve simply refers to stored assimilates and so its definition differs from that given in full DEB models (Kooijman, 2000). DEBkiss assumes that juvenile animals allocate all available resources to growth and maturation, and so $W_R=0$ until the onset of puberty. From this point on a portion of assimilates are stored for reproductive investment (Jager et al., 2013). We make the same assumption, with the distinction that assimilates are stored to cover the costs of reproduction or maintenance as needed.

Wet weight is more practical to measure than structural volume, V. Assuming that average wet tissue density, d_w (g × cm^{-3}), is equal to that of water, that is $d_w = 1$ (Lika et al., 2011), means that in juvenile animals $V = W_w/1$ g × cm^{-3}. Rather than any specific measure of length, such as nose to tail, the volumetric length, L, is defined as $V^{1/3}$ and surface area, a, is equal to L^2 or $V^{2/3}$. It is also helpful to estimate the density of structure, d_V (g × cm^{-3}), allowing conversion between dry weight and volume such that $W_v = V \times d_V$ and $W_w = W_V/d_V + W_R$. Multiple studies have estimated average tissue water content of *R. norvegicus* as between 64% and 74% (Reinoso et al., 1997) suggesting that 0.3 is a realistic value of d_V for this species.

5.4.3 Model notation

J - Flux or rate

y - Yield or efficiency

d - Density

X - Food

A – Assimilates

M – Maintenance

R - Reserve

W – Weight or mass

V - Structural volume

a - Surface area

L - Volumetric length

w - Wet tissue

m – Maximum

5.4.4 Growth model

Assimilation of nutrients from food into the body occurs across membranes and so this process is assumed to be mediated by surface area. Assimilation flux, *JA,* is defined as

$$JA = fJ_{Am}^a V^{\frac{2}{3}} \tag{5.1}$$

where J_{Am}^a is the maximum surface area specific assimilation rate (g(assimilates) × cm$_{(L)}$$^{-2}$ × d^{-1}) and V is volume. The parameter *f* is 'scaled functional response' to food availability (Jager et al., 2013, Kooijman et al., 2008) or 'scaled feeding rate' depending on how it is calculated. The distinction between these two terms is detailed later.

Maintenance flux, *JM,* is given as

$$JM = J_M^V V \tag{5.2}$$

Where J_M^V is the mass specific maintenance rate (g(assimilates) × cm$_{(L)}$$^{-3}$ × d^{-1}). Endotherms are also subject to surface area specific maintenance costs, accounting for heat loss to the environment. However as long as the ambient temperature is within the thermoneutral zone of a species (Kingma et al., 2014) these are assumed to be zero (Lika et al., 2011). Laboratory guidelines require rodents to be kept at 22±3°C, as this was considered to be within the thermoneutral zone of the rat (Poole and Stephenson, 1977). More recent research has suggested that this

temperature range is too low (Le and Brown, 2008) but for simplicity we assumed that heat loss could be omitted.
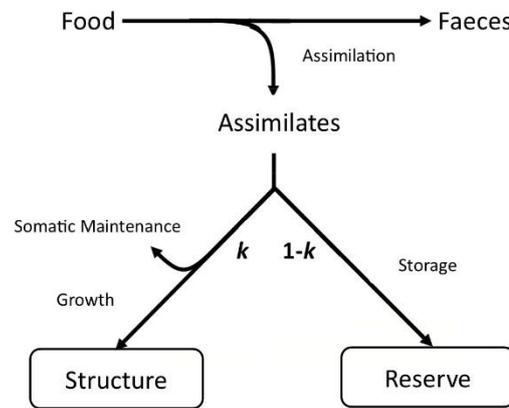
It is assumed that a certain proportion of assimilates are allocated to structural maintenance and growth and this is denoted k (dimensionless).

If $kJA > JM$, that is, assimilation is sufficient for growth and reserve storage

$$\Delta W_V = y_{VA}(kJA - JM) \tag{5.3}$$

$$\Delta W_R = (1 - k)JA \tag{5.4}$$

Where y_{VA} (g$_{(structure)}$ × g$_{(assimilates)}$$^{-1}$) is the yield of structure over assimilates, (*i.e.* the efficiency with which assimilates can be converted into structure). Puberty is estimated to begin in rats at 5-7 weeks of age (Rakel and Gergs, 2018) which is also the age of the study animals at the start of observation. As such, we assumed that $W_R = 0$ initially and begins to accumulate immediately. Like the full DEB model, our model implementation divides wet weight into structure and reserve (although reserve is more narrowly defined in this case). However, the model equations used are unaltered from DEBkiss and follow the simple assumption that any assimilates not required for maintenance, or allocated to growth, are stored. This system is represented in figure 5.2.

Figure 5.2 *A graphical representation of the DEBkiss model used when assimilation is sufficient for growth. The value of k determines the proportion of resources assimilated from food allocated to maintenance and growth or stored for reproductive investment.*

At any constant value of f, growth ceases when $J_A = J_M$. This is the point at which the ultimate structural volume, V_∞, is reached, which can be calculated as $(kfJ_{Am}^a/J_M^v)^3$. The theoretical maximum structural volume, V_m, is reached when $J_A = J_M$ and f=1 such that $V_m = (kJ_{Am}^a/J_M^v)^3$. At all times $\Delta W_w = \Delta W_V/d_V + \Delta W_R$ but ΔW_V and ΔW_R depend on the value of f.

If $kJA < JM < JA$, that is, overall assimilation flux JA is sufficient for homeostasis but not growth then

$$\Delta W_V = 0 \tag{5.5}$$

$$\Delta W_R = JA - JM \tag{5.6}$$

Maintenance is prioritised above growth, with the 1-k branch utilised to pay maintenance costs and any remainder stored as reserve. Equations 5.5 and 5.6 also describe change in body mass when $JA < JM$ and $W_R > 0$. In this scenario, the animal is starving, and W_V is maintained by utilising reserve. Both ΔW_R and ΔW_W become negative as the reserve decreases.

If $JA < JM$ and $W_R = 0$, that is, reserve has been used up and assimilation is insufficient to meet maintenance costs

$$\Delta W_V = (JA - JM)/y_{AV} \tag{5.7}$$

$$\Delta W_R = 0 \tag{5.8}$$

where y_{AV} (g$_{(assimilates)}$ × mass$_{(structure)}^{-1}$) is the yield of assimilates over structure (*i.e.* the efficiency with which assimilates can be extracted from structure). Therefore, structural weight is lost until it can be sustained by feeding.

The values of k and y_{VA}, 0.9472 and 0.7988 respectively, were taken from the most recent AmP entry for *R.* norvegicus (Rakel and Gergs, 2018), while y_{AV} was assigned its default value of 0.8 (Jager et al., 2013). The maximum surface area specific assimilation rate, J_{Am}^a and volume specific maintenance rate, J_M^V, were fitted to data. The value of f was calculated from data or food availability. Various approaches to this calculation, and their theoretical implications, are now summarised.

5.4.5 Methods for calculating f

5.4.5.1 Method 1: f = scaled functional response to food availability

The approach used most commonly in DEB literature is to calculate the value of f based on food availability because, in most cases, detailed feeding data are unavailable. This approach uses the Holling Type II functional response

$$f = X/(X + H) \tag{5.9}$$

where X denotes the density of food in the environment (g$_{(food)}$ × m^{-2}) and H (g$_{(food)}$ × m^{-2}) is the half-saturation food density at which food consumption rate, J_X (g$_{(food)}$ × day^{-1}) is half of its

maximum. When food is available *ad libitum, X =* ∞ and therefore $f = 1$ (van der Meer, 2006a, Kooijman et al., 2008, Jager et al., 2013). Using food availability as a proxy for observations of food consumption in this way relies on the assumption that, when provided with as much food as they can eat, animals eat as much as they can.

In other studies, *f* has been fixed to one during calibration and then estimated for independent data (Sadoul et al., 2018). This approach was not followed in this study as it is based on observed growth rather than feeding data, so the mechanistic basis is unclear. An alternative would be to compare overall average area specific feeding rate in the independent dataset to that of the calibration dataset and adjust *f* accordingly. However, this would not test whether food availability is a suitable proxy for feeding observations. Since feeding availability was always *ad libitum* in all datasets included in this study, *f* was fixed at its maximum value of 1 for all datasets in this study.

5.4.5.2 Methods 2 & 3: From scaled functional response to scaled feeding rate

The rate at which an animal can consume food depends on body size and so J_X has no fixed upper limit. Instead, it is assumed that feeding rate is limited by surface area so

$$J_X = fJ_{Xm}^a L^2 \tag{5.10}$$

where J_{Xm}^a (g*(food)* × cm*(L)*$^{-2}$ × day^{-1}) is the maximum area specific feeding rate for a species and L is the animal's volumetric length (cm) (Jager et al., 2013).

$$J_{Xm}^a = J_{Am}^a / y_{AX} \tag{5.11}$$

where y_{AX} is the yield of assimilates from food or digestive efficiency (g*(assimilates)* × g*(food)*$^{-1}$) and J_{Am}^a is the maximum surface area specific assimilation rate (g*(assimilates)* × cm*(L)*$^{-2}$ × d^{-1}). Since $y_{AX} \leq 1$, J_{Xm}^a provides the upper limit when fitting J_{Am}^a.

Dividing equation 5.10 by L^2 gives

$$J_X^a = fJ_{Xm}^a \tag{5.12}$$

where J_X^a is area specific feeding rate (g*(food)* × cm*(L)*$^{-2}$ × day^{-1}). Solving for *f* gives

$$f = J_X^a / J_{Xm}^a \tag{5.13}$$

So, where J_X^a can be calculated from observed data, it is more appropriate to calculate *f* using equation 5.13 and define it as 'scaled feeding rate' rather than scaled functional response. In Methods 2 and 3, J_X^a is calculated for each observation interval by dividing observed of daily food consumption by the associated observation of wet weight, W_w, raised to the power of 2/3. Strictly

speaking, calculations of J_X^a should be based on structural surface area, a or $V^{2/3}$. However, since V is not quantifiable from observed data, W_w was used instead. While both Methods 2 and 3 use equation 5.13, they differ is in how J_{Xm}^a is calculated.

5.4.5.3 Method 2: J_{Xm}^a = maximum observed area specific feeding rate

In this approach, previously employed in Martin et al. (2019), J_{Xm}^a is defined as the maximum individual observed area specific feeding rate within a dataset (separated by sex). In group A, J_{Xm}^a was 0.822g × cm^{-2} × day^{-1} for males and 0.715g × cm^{-2} × day^{-1} for females. These values were used for all datasets. Identifying J_{Xm}^a in this way guarantees that scaled f values do not exceed one for the calibration data set. It is possible, though unlikely, for f to exceed one when using (mean) independent data. This would require the average feeding in the independent dataset to exceed the highest individual observation in the calibration dataset.

5.4.5.4 Method 3: J_{Xm}^a = predicted J_X^a at a given body size, maximum food availability

In this method, observed daily food consumption, J_X (g$_{(food)}$ × day^{-1}), and area specific feeding rate, J_X^a (g$_{(food)}$ × cm$_{(L)}^{-2}$ × day^{-1}) with food available *ad libitum*, were described empirically as functions of surface area (calculated as $W_w^{2/3}$). Visual inspection showed that, rather than continually increasing as animals grew, J_X roughly followed a sigmoid pattern when plotted against surface area. The generalised logistic function (Richards, 1959) was selected as a flexible sigmoid curve which could meet the necessary conditions to model J_X as a function of body size. It was specified that the curve must pass through the origin, as an animal with zero mass would be unable to consume any food.

One expression of the generalised logistic formula to describe J_X in terms of surface area, a, is

$$J_X = G + \frac{U-G}{1+e^{-B(a-M)}} \tag{5.14}$$

Where G is the lower asymptote (g$_{(food)}$ × day^{-1}), U is the upper asymptote (g$_{(food)}$ × day^{-1}), M is inflection point (cm^2) and B is the growth rate (cm^{-2}). The simplest way (minimum number of parameters) in which this can be adjusted to pass through the origin is as a symmetrical curve with its inflection point at (0,0). This can be done by stipulating that $G = -U$ and $M = 0$ such that
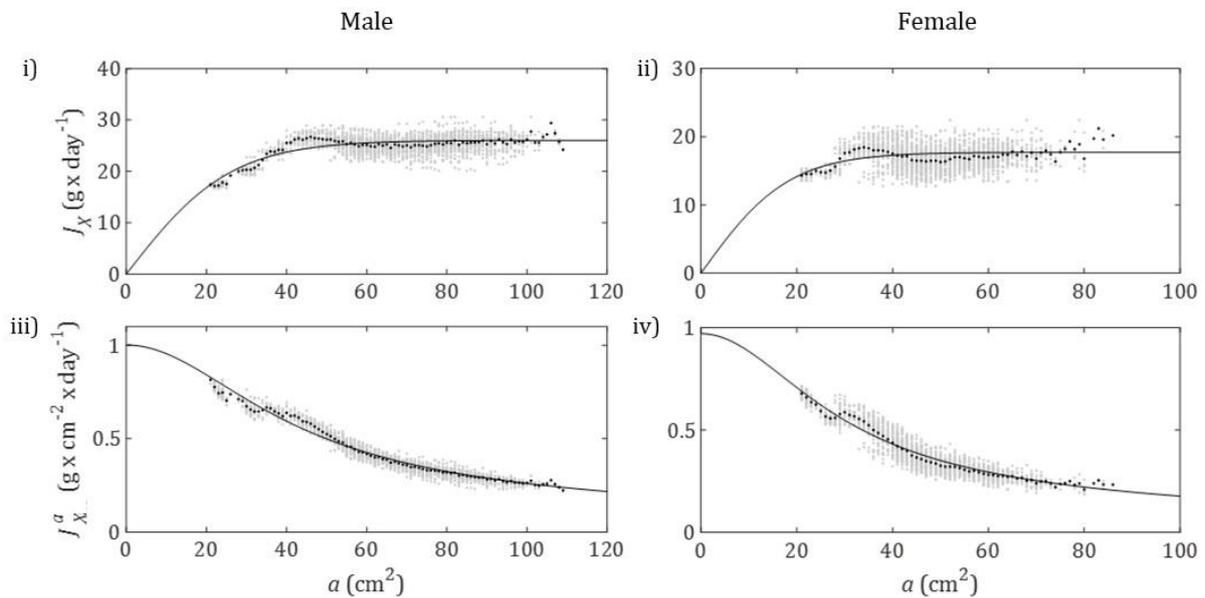
$$J_X = \frac{2U}{(1+e^{-B(a)})} - U \tag{5.15}$$

With only two free parameters, U and B, this function was then fitted to mean J_X (at each unique value of a) in the calibration dataset (figure 5.3). The coefficient of determination, R^2, was then calculated. This showed that, for males (U=26.03 and B=0.07693), 82% of variation in in mean J_X

was explained by the fitted function of surface area. For females (U=17.72 and B=0.1096), this figure was 38%, as data were more variable particularly at extreme body sizes with fewer observations (figure 5.3 ii). In order to eliminate values of J_X that were insufficient to meet maintenance costs, any data collected after mean body size had peaked (day 539 for males, day 686 for females) were excluded. Area specific feeding rate, J_X^a, was then simply modelled as

$$J_X^a = J_X/a \tag{5.16}$$

This explained 98% of variability in mean area specific feeding rate for males and females (figure 5.3 iii-iv). Modelled J_X^a, as a function of a, was then used as the reference for scaling observed J_X^a, meaning that J_{Xm}^a was redefined as predicted area specific feeding rate at a given body size, at maximum food availability (equation 5.16).

Figure 5.3 *Plots i & ii show observed (circles) and modelled (line) daily food consumption J_X vs surface area, a, of males and females respectively. Raw data are plotted in light grey while mean values are plotted in black. Method 3 uses equation 5.15, fitted to mean data to model food consumption per day, J_X. Plots iii & iv show observed (circles) and modelled (line) area specific feeding rate J_X^a vs surface area, a, of males and females respectively. Raw data are plotted in light grey while mean values are plotted in black. Models plot the fitted formula for J_X divided by surface area, a.*

Redefining J_{Xm}^a has several important theoretical implications. Since J_{Xm}^a is no longer a true maximum, the scaled feeding rate, f, may exceed one. Because $J_A^a = fJ_{Am}^a$ this means J_{Am}^a no longer represents the maximum area specific assimilation rate and is redefined as the area specific assimilation rate at maximum food availability. Crucially, as per equation 5.11, if J_{Am}^a remains fixed but J_{Xm}^a decreases as animals grow, then digestive efficiency, y_{AX}, must increase with body size. Furthermore, $y_{AX} \leq 1$ $\therefore$ $J_{Am}^a \leq J_{Xm}^a$, so the lowest predicted J_{Xm}^a within the observed range of body size provides an upper limit for J_{Am}^a.

5.4.5.5 Addressing the 'snowball effect' in Methods 2 and 3

To avoid the feedback loop described in the introduction (Martin et al., 2019), J_X^a was calculated in real time from observed food consumption, J_X (g × day^{-1}), and modelled surface area, a (cm²). Surface area, a, at time t was defined as modelled $W_w^{2/3}$ (in order to be consistent with how J_X^a was calculated from data). Mean observed food consumption, J_X, at time t was then divided by a, to yield J_X^a for the next time step. In Method 2, J_X^a was simply divided by the fixed value of J_{Xm}^a to give a value of f at time t. In Method 3, J_{Xm}^a was calculated by entering modelled a at time t into equations 5.15 and 5.16 before using equation 5.13 to yield f.

This meant that growth rate was modelled based on observed food consumption rather than observed area specific feeding rate over time.

5.4.5 Model assessment

Initially, the growth model was fitted to wet weight data from the calibration dataset A using each method of f calculation. To reduce the impact of heteroscedasticity, the square root transformation was used during fitting. The accuracy of the model fits was then assessed in a variety of ways. Overall goodness of fit was measured with the coefficient of determination, R^2, and the root mean square error, RMSE. Additionally, the proportion of observations predicted to within one standard deviation of the mean was calculated.

Next, the models were used to predict independent datasets B and C, without recalibration, to assess the generality of the model parameters derived using each method. Again, predictions were assessed using R^2, RMSE and the proportion of observations predicted to within one standard deviation of the mean.

The biological realism of each method was then assessed through comparison of the theoretical maximum volume, V_m, and modelled reserve, W_R, to relevant literature data. Finally, the impact of real time f calculation to avoid positive feedback was assessed with a worked example.

5.4.6 Model implementation

All models were implemented in Matlab (ver. R2016b). Growth models were developed with the BYOM (Jager, 2019) flexible model platform (ver. 4.1). All fitted parameter values were derived using the Nelder Mead simplex algorithm to maximise the likelihood function, given the observed data (Pan and Fang, 2002). Likelihood profiling was also used to check that initial fits were not local optima (Kreutz et al., 2013).

5.5 Results

5.5.1 Calibrated Growth Curves

The growth curve was fitted to mean wet weight, W_w, observed over two years in group A, using each of the three methods for determining the scaled feeding rate, f (figure 5.4). Fitted parameter values as well as various measures to assess goodness of fit are given in table 5.1. Goodness of fit measures only relate to total body weight, W_w, as this was the only model variable monitored in the dietary toxicity studies which provided data for this investigation. For illustrative purposes, the breakdown of modelled W_w into reserve and structure is shown on plots. Modelled reserve dynamics are assessed with respect to literature data in the Biological Realism subsection. Full size plots of calibration results can be found Appendix C (figures C1-C6 and C19-C24).
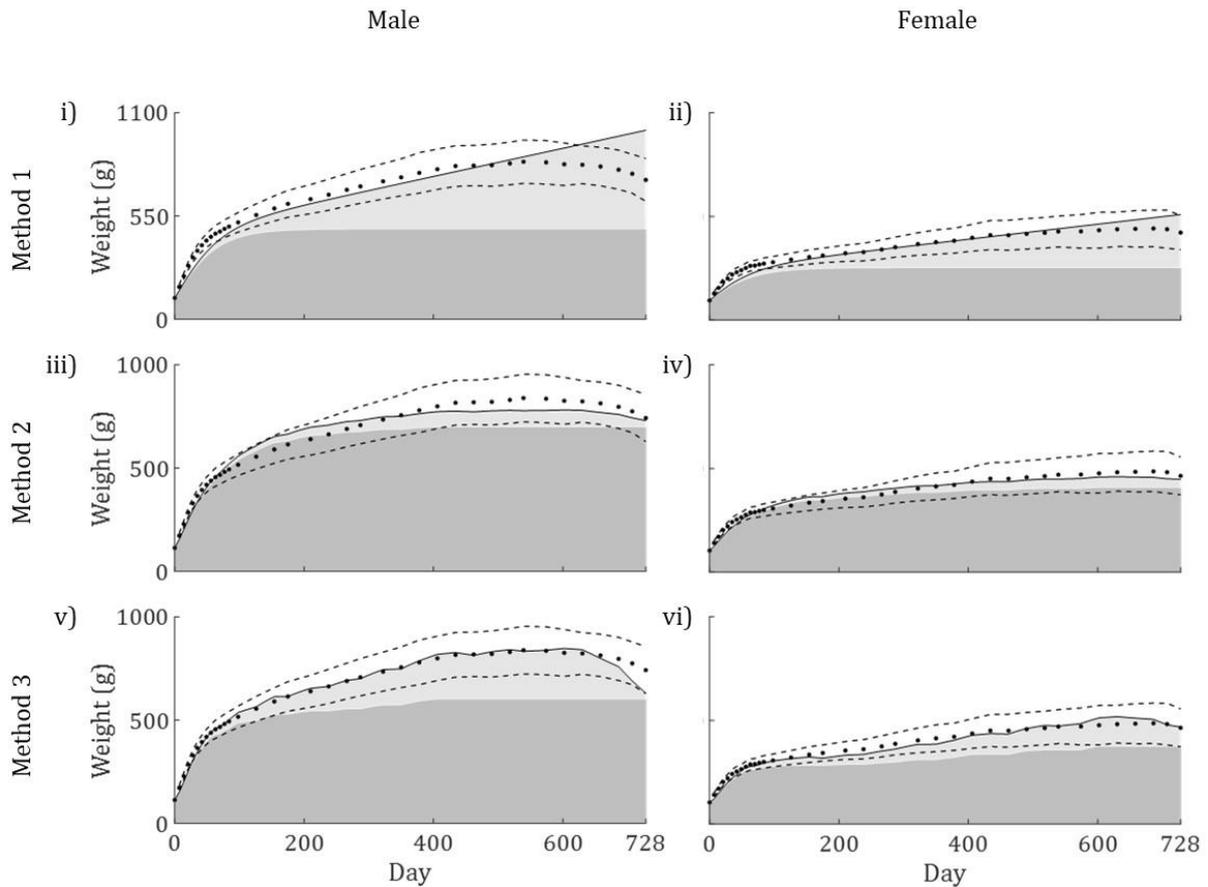
Table 5.1 *Fitted parameter values, selected observed and modelled endpoints, and goodness of fit measures for each method of calculating the scaled feeding rate, f.*

Sex	Males			Females		
Observed max. W_w (g)	839.0			486.2		
Method	1	2	3	1	2	3
J_{AM}^a (g × cm⁻² × d⁻¹)	0.2312	0.1196	0.1921	0.1775	0.07966	0.1313
J_M^V (g × cm⁻³ × d⁻¹)	0.02794	0.004829	0.02183	0.02575	0.004295	0.01754
Modelled max. W_w (g)	1006.6	781.4	847.3	560.0	460.4	518.0
R^2	0.857	0.970	0.983	0.900	0.968	0.971
RMSE	79.55	36.35	27.53	33.67	19.13	18.18
% W_w observations modelled to within 1 SD	61.11	86.11	91.67	66.67	86.11	88.89

In Method 1, f = 1 for the duration of the study. This meant that a smooth curve was produced and stored reserve, W_R, rose continuously. Consequently, this method produced the weakest fits to mean body weight, W_w, over time. For both sexes, this method produced the lowest R^2, the highest RMSE and modelled the fewest observations to within 1 s.d. of the mean.

Method 2 defined J_{Xm}^a as the highest observed area specific feeding rate in group A. Good fits were calculated for mean W_w over time (R^2 > 0.96). Modelled growth rate fluctuated in response to variation in food intake over time and became negative toward the end of the study period, matching observations. The overall shape of the curve was similar for males and females, showing signs of systematic error. Modelled growth rate lagged behind that observed until modelled body weight overtook observations after around 9-11 weeks. This persisted until modelled body weight fell below observations once again after 50-58 weeks (figure 5.4 iii-iv).

In Method 3, J_{Xm}^a was calculated as a function of surface area, a. The calculated fits to mean data were slightly better, for all measures, than those of Method 2. Modelled growth rate was highly responsive to fluctuations in f, becoming negative as area specific feeding rate dropped in the late stages. For both sexes, modelled body weights were very close to observed data for most of the observation period, with significant deviations only occurring late in the study. Maximum modelled W_w was only 1% higher than the maximum observed body weight in males and 6.5% higher for females.

Figure 5.4 *Plots showing models (solid line) fitted to observed mean body weight of group A male and female rats over 2 years (circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey) while dashed lines represent observed mean ± SD. The results of Method 1 are shown in plots i-ii, Method 2 in plots iii-iv, and Method 3 in plot v-vi.*

5.5.2 Summary analysis of food consumption and body weight data

Based on mean observed body weight and food consumption at each timepoint, summary analyses were conducted to highlight broad differences between the data sets (table 5.2). For both males and females, total food consumption was highest in group B, intermediate in group A and lowest in group C. Males and females in group A had the lowest starting weight but were intermediate in terms of maximum body weight and final body weight, with the highest weight

gain (final weight minus initial weight) over two years. Weight gain, maximum weight and final weight were lowest for males and females in group C.

Table 5.2 *Summary data based on mean observed body weight and food consumption at each timepoint.*
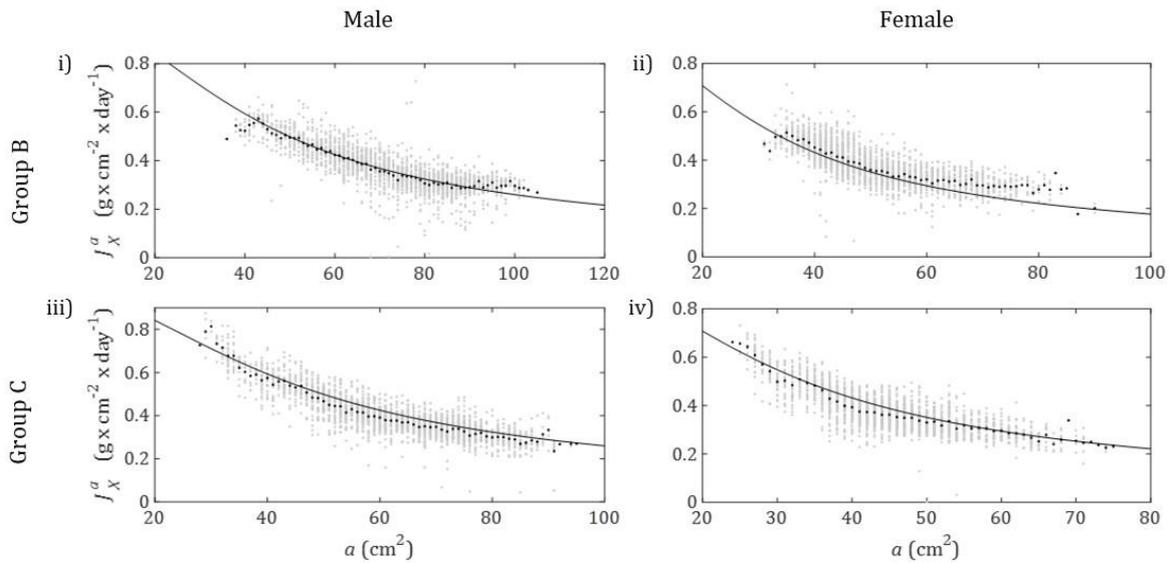
Sex	Male			Female		
Dataset	A (Calibration)	B	C	A (Calibration)	B	C
Total Food Consumption (kg)	17.97	18.36	17.32	12.24	13.76	12.16
Initial Weight (g)	113.37	209.38	190.35	104.14	189.22	144.69
Max Weight (g)	839.04	859.78	695.08	486.21	559.79	433.77
Final Weight (g)	742.35	784.36	672.11	464.31	549.12	430.20
Weight gain (g)	628.99	574.98	481.76	360.17	359.90	285.51

5.5.3 Feeding Rate Predictions

As part of Method 3, the generalised logistic curve was fitted to mean observed daily feeding rate, J_X (g × day^{-1}), as a function of surface area, a, (cm^2) of male and female rats in group A. This produced R^2 values of 0.82 and 0.37 respectively. Dividing fitted J_X by a to predict, mean area specific feeding rate at maximum food availability, J_X^a, produced R^2 values of 0.98 for males and females.

To assess the uniformity of the relationship between J_X^a and a across study groups, the predictions of the calibrated curves were compared to independent datasets B and C (figure 5.5). For males, variation in mean J_X^a was well predicted by surface area with R^2 values of 0.90 for group B and 0.95 for group C. Observed J_X^a in group B agreed closely with predictions at medium body sizes but exceeded predictions at large sizes and showed a decrease at low body size that was not evident in the calibration data. Observed J_X^a in group C showed a similar shape to the predicted curve but was generally slightly lower. The relationship was less consistent for females though, R^2 was 0.68 for group B and 0.95 for group C. Observed J_X^a in group B was higher than predicted, particularly at larger body sizes. As was the case for males, J_X^a in group C was slightly lower than predicted for most body sizes.

The relationship between J_X^a and a appears less uniform among female rats. However, deviations from predicted J_X^a may be reflected by a predictable increase or decrease in growth rate.

Figure 5.5 *Plots comparing independent data (circles) to predictions (lines) of area specific feeding rate, J_X^a vs surface area, a. Raw data are plotted in light grey while mean values are plotted in black. Data for males and females in group B are shown in plots i & ii respectively while data for males and females in group C are shown in plots iii & iv respectively.*
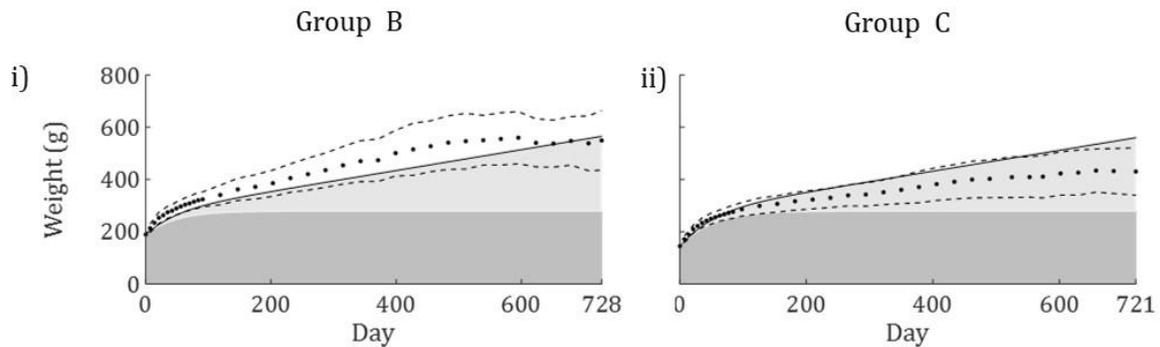
5.5.4 Growth curve validation

The calibrated growth models were used to predict independent datasets B and C. The accuracy of the predictions produced by each method was assessed by calculating R^2, RMSE and the percentage of observations predicted to within one standard deviation (table 5.3). Full size plots of all model predictions of independent data are included in Appendix C (figures C7-C18 and C25-C36).

Table 5.3 *Selected measures of the accuracy of each method when used to predict independent data.*
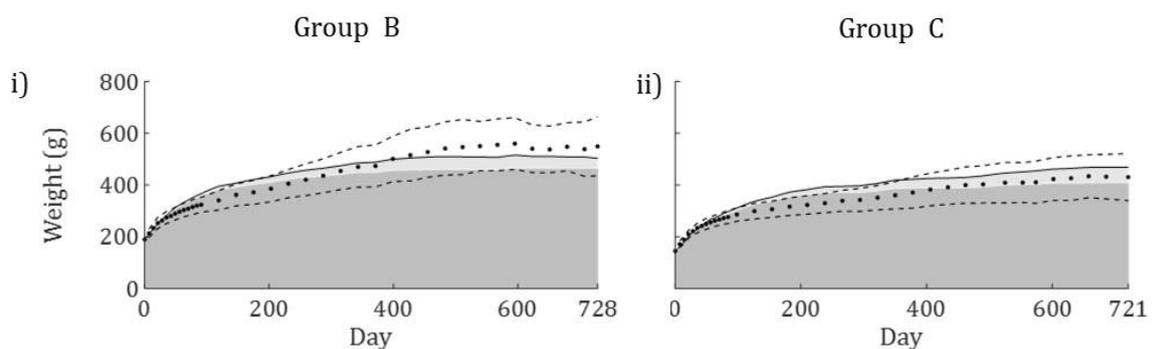
Sex	Male					
Dataset	B			C		
Method	1	2	3	1	2	3
R^2	0.7749	0.9641	0.9434	0.0889	0.7311	0.8624
RMSE	86.25	34.44	43.23	137.7	74.82	53.52
% observations predicted to ±1 s.d.	36.11	86.11	80.56	34.29	54.29	80.00
Sex	Female					
Dataset	B			C		
Method	1	2	3	1	2	3
R^2	0.8831	0.9263	0.2609	0.5513	0.7970	0.6098
RMSE	40.25	31.95	101.2	54.55	36.69	50.87
% observations predicted to ±1 s.d.	83.33	75.00	16.67	71.43	74.29	71.43

Method 1 produced virtually identical curves for all datasets. This is because initial weight was the only model input that differed from the calibration data. This method produced the poorest predictions of mean body weight over time for males in both independent datasets and for females in group C. For females in group B, Method 1 produced the highest proportion (83.33%) of predictions within one standard deviation of the observed mean (figure 5.6). However, had the model continued to run, modelled body weight would have continued to increase, as reserve accumulated indefinitely.

Figure 5.6 *Plots showing predicted (black lines) and mean observed body weight of female rats over 2 years (circles), using Method 1 to calculate the scaled feeding rate, f. The shaded area under the model curves shows structure (dark grey) and reserve (light grey) while dashed lines represent observed mean ± SD. Results for group B and C are shown in plots i & ii respectively.*
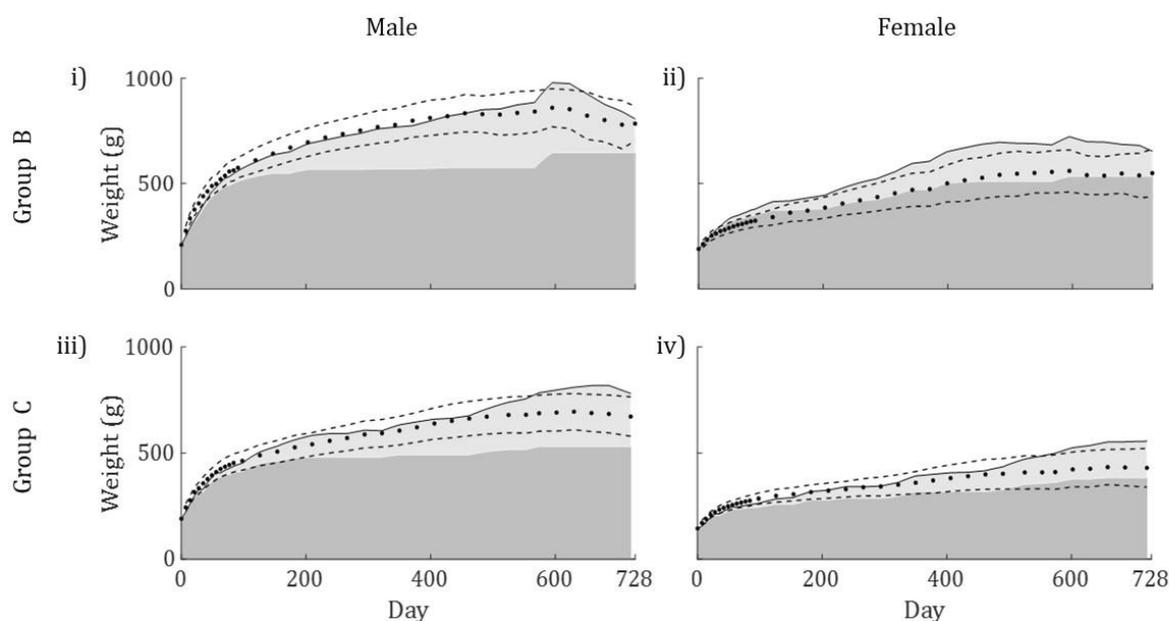
Method 2 produced the most accurate predictions (highest R^2 and lowest MRSE) of mean growth rate for males and females in group B. Body weight of both sexes in group C was overpredicted for all but the early stages of observation. Despite this, Method 2 did produce the most accurate predictions for females in this dataset (figure 5.7).

Figure 5.7 *Plots showing predicted (black lines) and mean observed body weight of female rats over 2 years (circles), using Method 2 to calculate the scaled feeding rate, f. The shaded area under the model curves shows structure (dark grey) and reserve (light grey) while dashed lines represent observed mean ± SD. Results for group B and C are shown in plots i & ii respectively.*

Method 3 predicted growth of males in group B slightly less accurately than Method 2, and was the most accurate for males in group C. In both cases, model predictions closely followed the

observations until the late stages of observation. This was also the case for females in group C, for which this method produced the second most accurate predictions. However, growth of females in group B was poorly predicted. Modelled body weight was well above that observed for almost all of the observation period. These predictions are shown in figure 5.8.
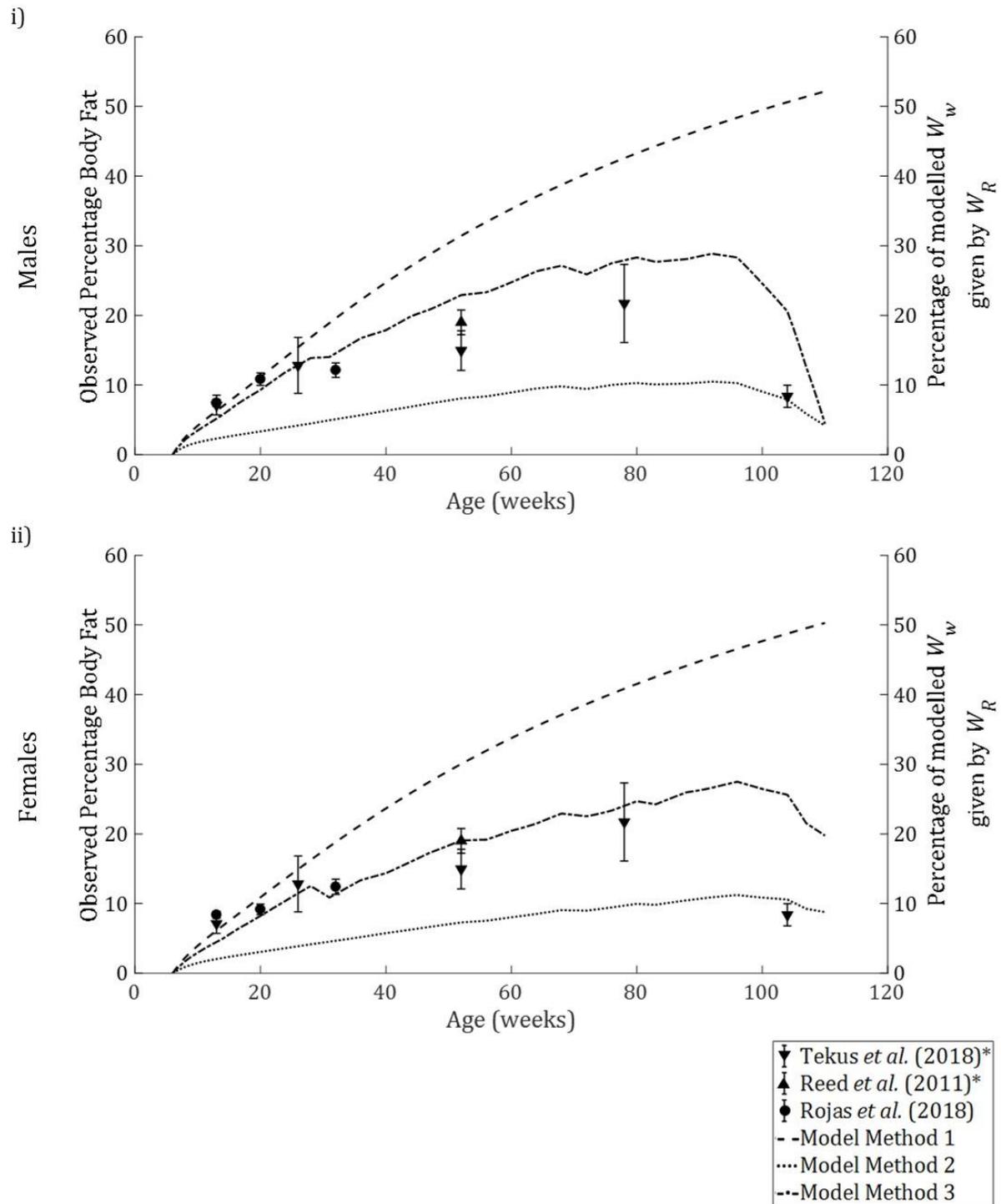

Figure 5.8 *Plots showing predicted (solid line) and mean observed body weight of rats over 2 years (circles), using Method 3 to calculate the scaled feeding rate, f. The shaded area under the model curves shows structure (dark grey) and reserve (light grey) while dashed lines represent observed mean ± SD. The results for males and females in group B are shown in plots i-ii respectively while the results for males and females in group C are shown in plots iii-iv respectively.*

5.5.5 Biological Realism

Although goodness of fit to observed W_w (the only model endpoint measured in toxicity studies) quantifies model accuracy, it gives no information as to the biological realism of the model itself. In order to address this question, literature data were utilised to assess other model variables. No data are available for W_R, as this represents stored assimilates from food. This would include not only stored lipids but also carbohydrates stored as glycogen, and fat-soluble vitamins. Nevertheless, observed body fat percentage of *ad libitum* fed rats is a useful, if not ideal, comparator, as it would be expected to follow very similar temporal patterns.

Data from Tekus *et al.* (2018) provide reference values of body fat percentage of rats at various ages up to two years. While the study used only male Wistar (rather than Sprague Dawley) rats, other studies indicate that body fat percentage is similar across the two strains (Reed et al., 2011) and between male and female Sprague Dawley rats (Rojas et al., 2018). Figure 5.9 shows

literature data plotted against calibrated model simulations (group A) of W_R as a percentage of modelled W_w.

Figure 5.9 *Plots showing mean ± SE body fat percentage (left hand axis) recorded in rats of various ages by Tekus et al. (2018), Reed et al. (2011) and Rojas et al. (2018), and calibrated model simulations of W_R as a percentage of W_w over time (right hand axis) for male (plot i) and female rats (plot ii). * denotes that data were available for male animals only.*

Method 1 assumes that $f = 1$ at all times where food is available *ad libitum*, leading to constant accumulation of reserve. For both sexes, modelled W_R was around 50% of modelled W_w after 2 years, over five times the value reported at that age and more than double the maximum reported percentage body fat. Using Method 2, the observed decline in body fat in the late life stages was reflected by model simulations. W_R peaked at 10.48% of modelled W_w of males and 11.22% for females, only about half of the value reported by Tekus *et al.* (2018).

With Method 3, model simulations were relatively consistent with observations. For both sexes, W_R as a percentage of W_w matched observed body fat percentage at 6 months of age before reaching a peak between the ages of 18 months and two years and declining thereafter. Peak W_R percentage was 28.85% for males and 27.49%, slightly exceeding the highest mean body fat percentage + standard error (27.33%) reported by Tekus *et al.* (2018).

The maximum volume of structure, V_m, is a theoretical maximum calculated from model parameters as $(kJ_{Am}^a/J_M^v)^3$. Multiplying V_m by the density of wet tissue, d_w, (assumed to be 1 g × cm^{-3} (Lika et al., 2011)) gives the maximum wet weight of structure, W_{Vwm}. Peak lean weight would serve as a sensible proxy for comparison but is not measured in toxicity studies. Instead, we can assume that the weights of structure and reserve peak simultaneously, meaning that W_{Vwm} can be estimated from data as

$$W_{Vwm} = max.\,observed\,body\,weight \times (1 - max.\,recorded\,proportion\,body\,fat) \qquad (17)$$

Using the relevant values from group A (used in calibration) and Tekus *et al.*, gives 839.04g × 0.783 = 656.97g for males and 486.21g × 0.783 = 380.70g for females (table 5.4).

Table 5.4 *A comparison of maximum wet weight of structure, W_{Vmw}, estimated from data and calculated from model parameters.*

Sex	Male			Female		
Estimated W_{Vmw} (g)	656.97			380.7		
Method	1	2	3	1	2	3
Modelled W_{Vmw} (g)	481.5	12910	579.1	278.3	5422	356.5
Model/ Estimate	0.733	19.7	0.882	0.731	14.2	0.936

Modelled W_{Vmw} was lowest when using Method 1, for both sexes its value was 73% of that estimated from data. Method 2 meanwhile produced very high values of W_{Vmw}, almost 20 times the estimated value for males and over 14 times the estimate for females. The values of W_{Vwm} given by Method 3 were closest to the estimates at only 12% and 6% lower for males and females respectively. This was also the only method for which W_{Vmw} was not a strict maximum, as f could

exceed one. The highest modelled wet weights of structure using this method were closer still at 602.8g for males and 375.6g for females.

5.5.6 Impact of real time f calculations

For Methods 2 and 3, the scaled feeding rate f was calculated in real time as the growth model ran. This was not applicable to Method 1 as f was not derived from feeding data in that approach. This was in order to ensure that the models reflected growth based on the quantity of food consumed, rather than observed area specific feeding rate which could lead to a 'snowball effect' where predictions deviate from data (figure 5.1). Hypothetical data were used to illustrate the impact this had on modelled ΔW_w (using the fitted parameter values given in table 5.1). A male rat was considered, weighing 300g and consuming 24g food per day, with Methods 2 and 3 returning f values 0.65 and 0.98 respectively (table 5.5).

Table 5.5 *Calculations of scaled feeding rate, f, using Methods 2 and 3 for the same hypothetical data.*

Method	J_X, $g_{food} \times day^{-1}$	W_w, g	J_X^a, $g_{food} \times cm^{-2} \times day^{-1}$	J_{Xm}^a (Method 2), $g_{food} \times cm^{-2} \times day^{-1}$	Predicted J_{Xm}^a at 300g (Method 3), $g_{food} \times cm^{-2} \times day^{-1}$	f
2	24	300	0.54	0.82	N/A	0.65
3	24	300	0.54	N/A	0.55	0.98

We then supposed that modelled W_w was either accurate (300g) or 50g above or below that observed at time t (in all cases W_w was broken down into 88% structure and 12% reserve). We then calculated ΔW_w for the next time step. Using Method 2 and the f value (table 5.5) derived directly from data, as in Martin et al. (2019), showed that f always corresponds to the same area specific feeding rate, J_X^a, equating to a higher J_X in larger animals and vice versa (table 5.6). Therefore, modelled growth rate does not reflect the quantity of food consumed, this leads to positive feedback between modelled W_w and ΔW_w. As a result, once predictions of W_w deviate from data, they become less accurate with each time step. Using Method 3, the value of J_X^a corresponding to the given value of f decreases with body size (table 5.6). In this case, negative feedback occurs between modelled W_w and ΔW_w at a given value of f. However, a positive relationship still exists between modelled W_w and J_X, so ΔW_w does not exactly reflect the quantity of food consumed.

Table 5.6 *Food consumption, J_X, and growth rate, ΔW_w, as calculated from predicted body weight, W_w, and observed scaled feeding rate, f, using Methods 2 and 3. This approach was not used in this study due to positive feedback.*

Method	Inputted f value	Predicted W_w at time t, g	Corresponding J_X^a at time t, g$_{(food)}$ × cm^{-2} × day^{-1}	Corresponding J_X at time t, g$_{(food)}$ × day^{-1}	Predicted ΔW_w, g × day^{-1}
Method 2	0.65	350	0.54	26.60	1.68
	0.65	300	0.54	24.00	1.57
	0.65	250	0.54	21.25	1.44
Method 3	0.98	350	0.49	24.48	1.60
	0.98	300	0.54	24.00	1.68
	0.98	250	0.59	23.27	1.72

The exercise was then repeated, this time using the approach employed throughout this study. In this case, f was instead derived from observed food consumption, J_X, and modelled W_w (table 5.7). When the calculations are performed in this order, J_X is fixed so J_X^a and f decrease as modelled body size increases, and vice versa. For Method 2, this curtails the 'snowball effect', instead leading to negative feedback between modelled W_w and ΔW_w at a given value of J_X. For Method 3, negative feedback between W_w and ΔW_w is strengthened (a difference of 0.44 g × day^{-1} between ΔW_w at 250g and 350g body weight) as predicted growth rate accurately reflects observed daily food consumption.

Table 5.7 *Scaled feeding rate, f, and growth rate, ΔW_w, calculated from predicted body weight, W_w, and observed food consumption, J_X, using Methods 2 and 3. This is the approach used in this study.*

Method	Observed J_X at time t, g$_{(food)}$ × day^{-1}	Predicted W_w at time t, (g)	Corresponding J_X^a at time t, g$_{(food)}$ × cm^{-2} × day^{-1}	Inputted f value	Predicted ΔW_w, g × day^{-1}
2	24	350	0.48	0.59	1.40
	24	300	0.54	0.65	1.57
	24	250	0.60	0.74	1.74
3	24	350	0.48	0.96	1.46
	24	300	0.54	0.98	1.68
	24	250	0.60	1.01	1.90

5.6 Discussion

DEB models are designed to function without the need for detailed feeding data (Kooijman, 2000, Jager et al., 2013). However, this presents the question of what to do with such data when they are available. The conventional approach to deriving feeding inputs in DEB models does not reflect temporal or intertreatment variability in feeding rate, only food availability. In a previous study (Martin et al., 2019), we developed a method to derive feeding inputs directly from feeding data, but this approach had problems of its own. In this study, a novel method was developed, with the aim of addressing all the issues previously identified. We used a simple model based on DEBkiss to assess three approaches, for their impact on model accuracy, generality, and realism.

5.6.1 Accuracy and generality

Model accuracy was assessed by fitting the models to growth data for Group A. Method 3 produced the most accurate fits to calibration data for both males and females. Method 2 was only slightly less accurate. However, errors appeared more systematic in nature when using Method 2, following a similar pattern over time for both males and females. Method 1 was the least accurate, producing a smooth curve which did not respond to temporal variability in area specific feeding rate.

Model generality was then assessed by using the models to predict independent growth data (Groups B and C) without recalibration. Method 2 performed best in terms of model generality, despite systematic errors still being apparent. Without recalibration, body weight over time was predicted most accurately using this method for females in both independent datasets and for males in group B. Method 1 was again least accurate for all but one dataset. Using this method, the only model input to change between datasets was initial body weight which has minimal effects on predictions. The resulting model outputs for Method 1 were essentially different sections of the same curve for all datasets. While body weight of group B females was predicted relatively well, this result was coincidental as this dataset was quite different to that used in the calibration (table 5.2).

Predictions using Method 3 were most accurate for group C males and a close second for males in group B. Results were mixed for females however, this method was second most accurate for group C females but the least accurate for those in group B. It is notable that observed feeding patterns in this dataset were most different from predictions of J_{Xm}^a (figure 5.5) and that animals were fed on a different diet to those in groups A and C. While predictions of total body weight of group B females were poor, it is notable that predicted structural weight followed observed wet weight very closely (figure 5.8). This suggests that the profile of the scaled feeding rate, f, over

time was accurate, if not the values themselves. This could potentially be remedied by fitting the logistic curve to feeding data for each dataset. However, in this study no additional fitting was performed in order to ensure a fair comparison of all three methods.

5.6.2 Addressing the 'snowball effect'

We showed that real time calculation of f works effectively to stop the 'snowball effect' that occurred in a previous study. The large impact this had on ΔW_w predicted by Method 2, calls into question the findings of Martin et al. (2019) regarding the relative contributions of feeding toxicity to observed effects on body weight over time. In most cases, growth rate without chemical stress would have been overestimated due to positive feedback, leading to the effects of feeding avoidance being understated. Method 3 mitigated this problem even without this additional step; however, the extent of this mitigation would vary depending on fitted parameter values. It should be considered good practice in future studies to calculate f as the model runs, at least when predicting independent data.

5.6.3 Biological realism

Method 1, that is suggested by DEB literature (Kooijman et al., 2008, van der Meer, 2006a, Jager et al., 2013), assumes that $f = 1$ when food is freely available. In our growth model, this meant that W_{Vmw} was approached quickly with ΔW_R becoming linear *ad infinitum*. Modelled W_R reached over 50% of W_w after 2 years, more than twice the maximum reported percentage body fat in the literature.

It should be mentioned that our model did not include the maturity maintenance parameter, this represents the costs of maintaining sexual maturity and is taken from the 1- k branch of the model. However, these costs are assumed to be proportional to structural weight at puberty and so remain constant as animals grow larger (Jager et al., 2013). Maturity maintenance would have, therefore, only slightly mitigated the issue of constant reserve accumulation. A potential solution would be to implement the full DEB model which uses a more complex equation to model reserve dynamics. This includes an additional parameter, the maximum reserve density ($g_{(reserve)}$ × $g_{(structure)}^{-1}$), which provides a limit on reserve accumulation. However, despite using the same simple equations, constant reserve accumulation was not an issue with Methods 2 and 3. It may then be the case that, rather than representing a real biological processes or limits, maturity maintenance and maximum reserve density simply serve to compensate for the oversimplicity of the feeding input.

However, the major issue with Method 1 is that rats provided with food *ad libitum* do not feed at the maximum area specific rate, as is assumed (figure 5.10). Rats moderate their feeding significantly as they grow (Laaksonen et al., 2013, Martin et al., 2019). In fact, data in this study and others (Hubert et al., 2000, Tekus et al., 2018) show that weight loss occurs as rats approach two years of age. These patterns cannot possibly be modelled based on constant *ad libitum* food availability (equation 5.9). Therefore, $f \neq 1$ and/or digestive efficiency, y_{AX}, and maximum area specific feeding rate, J^a_{Xm}, are not fixed values.
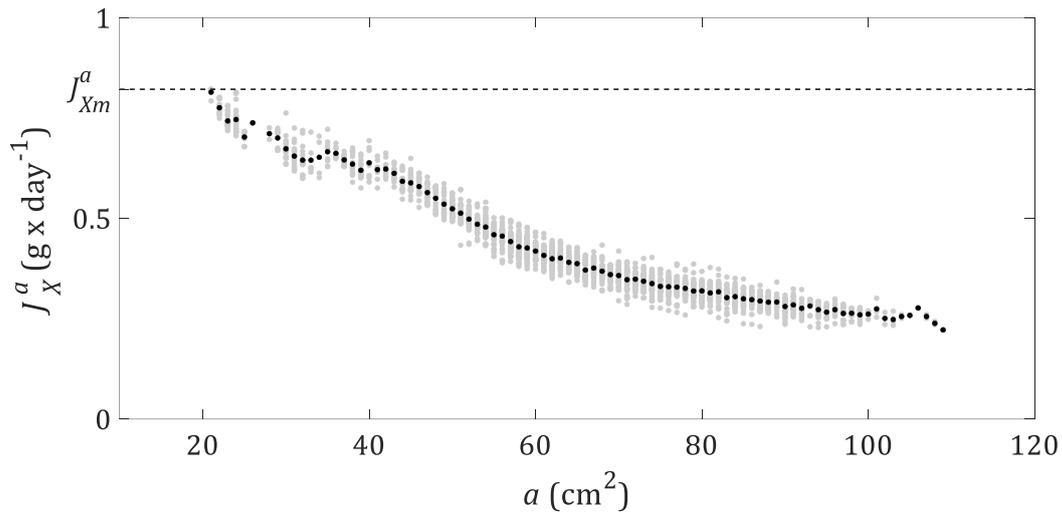
Figure 5.10 *Observed area specific feeding rate, J^a_X, of male rats in group A plotted against surface area, a. Raw data are plotted in light grey while mean values are plotted in black. The dashed line shows the relationship required for scaled feeding rate, f = 1 if digestive efficiency, y_{AX}, is fixed as a primary parameter. Maximum area specific feeding rate, J^a_{Xm}, is assigned the value of the highest observation in the dataset.*

In Method 2, previously employed in Martin et al. (2019), *f* is a dynamic input calculated based on observed area specific feeding rate over time. However, as before, the values of y_{AX} and J^a_{Xm} remain fixed as animals grow. Using this method, W_R relative to W_w was substantially lower than body fat percentages reported in the literature (Reed et al., 2011, Rojas et al., 2018, Tekus et al., 2018) although observed patterns in fat storage over time were reflected by the model (figure 5.9). As previously noted, the observed negative relationship between body size and area specific feeding rate leads to *f* values that decline to well below one. Fitted parameter values must compensate for this and as a result, V_m is extremely high. The parameter values in this study meant that, in theory, male rats could grow to almost 13kg if they fed at a sufficiently high rate. This suggests that this method is fundamentally flawed, as such sizes are far beyond the highest observations in the literature (Hubert et al., 2000, Rojas et al., 2018). The strong performance of Method 2, both in terms of model fits and predictions of independent data, is therefore an example

of a model being right for the wrong reasons. Based on accuracy and generality, Method 2 would appear to perform best overall, but assessing biological realism reveals serious faults.

Method 3 attempts to address issues with both the previous methods by positing that the values of y_{AX} and J_{Xm}^a vary as functions of body size and allowing f to fluctuate and exceed one. This attempt seems to have been largely successful. Using this method, modelled percentage body weight given by W_R was closest to observed body fat percentage for both sexes at all but one timepoint. Moreover, the modelled maximum wet weight of structure, W_{Vwm}, was only slightly lower than estimated peak lean weight in the calibration dataset.

5.6.4 What issues remain?

Conventionally, digestive efficiency, y_{AX}, is treated as a primary (fixed) parameter (Jager et al., 2013, Kooijman et al., 2008) but Method 3 changes this, such that this value increases with body size. This is highly plausible; several literature studies report increases in digestive efficiency associated with body size in a range of species (Smith, 1995, Illius and Gordon, 1992, Hansson and Jaarola, 1989, Demment and Vansoest, 1985). This occurs because increased gut capacity of larger animals allows the same volume of food to attain a greater surface area, while increased gut length leads to increased retention time for the extraction of nutrients.

In lieu of digestive efficiency data for rats, the generalised logistic model was fitted to food consumption data. This relationship relies on the assumption that growing rats, supplied with food *ad libitum*, consume enough food for area specific assimilation to equal J_{Am}^a and for structural volume to reach V_m. This appeared to be most true of male animals, with growth predictions being more accurate than for females. This would be consistent with the behavioural ecology of the species. Whereas females do not compete for mates and tend not to migrate, heavier males fare better in competition for dominance with unfamiliar individuals so there is selective pressure to grow as large as possible (Macdonald et al., 1999).

However, predictions with Method 3 were not always accurate for males either. Generally, predictions matched data well up until around day 500 but substantial deviations from data did occur thereafter. One possible explanation is that this occurred because the model allowed structural growth in older animals despite skeletal growth in rats generally ceasing after around 6 months; a process that appears related to age rather than body size (Roach et al., 2003). This certainly contributed to higher assimilation and reserve accumulation late on. Another possibility though, is that predicted J_{Xm}^a was too low at large body sizes making even small deviations from predictions proportionally larger than they should have been. This would exaggerate fluctuations in f and therefore ΔW_w in the later stages of growth.

A relatively minor issue is our lack of knowledge around weight loss and starvation in rats. Based on the data in this study and the literature, it appears to be typical for rats to reduce feeding and lose weight as body fat as they approach two years of age (Hubert et al., 2000, Tekus et al., 2018). This weight loss was overestimated by the model for males in group A. A possible reason is that reduced feeding elicits compensatory physiological or behavioural responses not included in the model's starvation rules. For example reduced body temperature has been documented as a response to short-term starvation in rats (Sakurada et al., 2000), which would correspond to a reduction in maintenance rate, J_M^V. Finding the data needed to refine the starvation rules proposed by DEBkiss represents a challenge though. While some studies have restricted food availability (Hubert et al., 2000), enforcing longer term starvation leading to weight loss would be unethical due to the suffering this would cause.

In order to address the remaining issues, the clear solution is to simply measure digestive efficiency of standard laboratory diets (Batzli and Cole, 1979, Veloso and Bozinovic, 1993) alongside food consumption and body weight in growing rats. This would allow the relationship between J_{Xm}^a and body size to be determined mechanistically and for equation 5.11 to be solved, providing the value of the maximum assimilation rate J_{Am}^a. Inevitably, empirical relationships can only provide an imperfect representation of reality. Indeed, at extreme body sizes (>1.577kg for males and >1.567kg for females) our parameters mean that $J_{Xm}^a < J_{Am}^a$ and therefore $y_{AX} > 1$. This is a physical impossibility as assimilates from food cannot exceed the mass of the food itself. The strong performance of Method 2 in predicting independent data suggests that the reality may sit between Methods 2 and 3. It appears likely that J_{Xm}^a does decrease as animals grow, though perhaps less dramatically than Method 3 predicts. Likewise, f likely does fall as animals grow, but less markedly than suggested by Method 2.

5.7 Conclusions

DEBkiss (Jager et al., 2013) inevitably made some compromises in order to simplify the DEB framework. However, our results suggest that it is a property common to all versions of DEB, the calculation used to derive feeding inputs (Kooijman et al., 2008), which represents an over-simplification. This was designed to circumvent the need for detailed feeding data, which are rarely available (Kooijman, 2000, van der Meer, 2006a). However, observed patterns between the feeding rate and surface area of rats clearly contradict model assumptions and so changes are required.

We have developed methods which extract more information from feeding data in order to broaden the applicability of models based on DEBkiss. With this approach we have produced

accurate and biologically sound models that use simple equations to model growth and reserve dynamics. This removes the assumption of first order dynamics of reserve density, which is the most difficult aspect of the full DEB growth model (van der Meer, 2006a). Where feeding data are unavailable, conventional methods by which constant or simple f inputs are assumed, may still be most suitable. However, we suggest that y_{AX} and J_{Xm}^a are dynamic variables that vary with surface area and that, even if these relationships cannot be quantified for most species, DEB theory should reflect this.

While the new method is a significant step in the right direction, relying on empirical relationships is not ideal and several issues remain that could be addressed by data collection. Models able to accurately predict how animals in dietary toxicity studies would have grown if fed a control diet are now within reach. Such models are a prerequisite for DEB-TKTD models that accurately reflect a compound's toxicity. Equally though, they represent an exciting new tool with which to analyse toxicological data, avoiding the conflation of effects due to toxicity and differences in feeding rate. This will allow assessment of how feeding avoidance impacts upon the ecological risk posed by a chemical in a way that was not previously possible.

Chapter 6 – Thesis Conclusions

This chapter presents conclusions from the entire thesis; each section relates to one of the broad challenges approached during the project. The final section presents suggestions for continued research towards this project's objectives and identifies the challenges that will need to be overcome.

6.1 Identifying Suitable Methods

When this project began, the potential of TK-TD modelling to refine vertebrate ecological risk assessment (ERA) had been recognised for some time. Regulators (EFSA, 2009b) and researchers (Ducrot et al., 2016) had identified TK-TD modelling as a relevant approach for risk assessment refinement, able to bring more realism to the risk assessment by simulating realistic exposure scenarios. However, few attempts had been made to apply these methods to vertebrate species and those efforts had focused on acute toxicity and risks to survival. Up until relatively recently, it was considered that no available TK-TD modelling techniques were suitable to simulate vertebrate growth under chemical stress (Ashauer et al., 2011). This was a clear knowledge gap as protection goals state that no mortality of birds and mammals associated with the use of pesticides must occur (EFSA, 2009b). Therefore, sublethal impacts should be the only risk relevant to realistic pesticide exposure to terrestrial mammals.

The primary aim of the project was to develop a TK-TD growth model for rodents using data from dietary toxicity studies using laboratory rats (*Rattus norvegicus*), with the secondary objectives of exploring model applications for ERA and quantitative *in vitro-in vivo* extrapolation (QIVIVE). For any endpoint of interest, whether survival, reproduction or growth, the process model is the fundamental component of a TK-TD model. The initial task was therefore to identify suitable methods to model growth in rats. Almost coinciding with the start of this project, Desforges et al. (2017) published promising results having used TK-TD models based on dynamic energy budget theory (DEB, Kooijman (2000)) to model effects of polychlorinated biphenyls on growth and reproduction of the American mink (*Mustela vison*). This study used the simplified DEB framework, DEBkiss (Jager et al., 2013), suggesting that this method may be applicable to mammals, despite having been developed primarily for invertebrates.

A common criticism of DEB models is that they contain too many species-specific parameters, which must be fitted to data as they do not represent directly measurable processes (Marquet et al., 2014). In chapter 2 we conducted a side by side comparison of the DEBkiss growth model with that of the metabolic theory of ecology (MTE, Brown et al. (2004b)), which claims to remedy this issue. By reducing the models to the simplest form, growth under constant conditions, we showed that they share many common features, offer comparable simplicity, and are similarly reliant on

fitted parameters. The meaningful difference between the two methods is in their underlying mechanisms and assumptions. DEB is founded upon the undisputed scaling relationships $Surface\ Area \propto Length^2$, $Volume \propto Length^3$ and $Mass \propto Volume$, with certain processes assumed to be area or mass specific (Kooijman, 2000). MTE, meanwhile, proposes that the fractally branched vascular networks of many plants and animals, deliver resources to cells at a rate which scales with body mass to the 3/4 power (Brown et al., 2004b). This proposed relationship has faced criticism from mathematicians (Kozlowski and Konarzewski, 2004) and so we found that, at this point in time, there is no compelling reason to reject DEB in favour of MTE for use in TK-TD modelling. Additionally, the DEBtox framework (Kooijman and Bedaux, 1996a) provides an established method for applying chemical stress to the model parameters.

6.2 Modelling growth with abundant data – an opportunity and a challenge

In general, experiments are not conducted with model parameterisation in mind, rather modellers work with the data available from published studies (Jager, 2020). An added obstacle is that multiple datasets are required to calibrate models and then validate their predictions. Due to funding structures in science, corroboration studies are rarely conducted in academia or published in the scientific journals as they are not considered novel (Jager and Ashauer, 2018). By contrast, ecological risk assessment requires multiple animal studies for each new pesticide, varying in duration and dose level. However, only summary data from regulatory studies are normally made public. In this project we had access to raw data from dietary toxicity studies on rats for several pesticides. Such access to unpublished data represented a major opportunity but also a challenge, which was first approached in chapter 3.

The challenge arose because DEB models are designed to be compatible with the data generally available to modellers, or lack thereof. Field studies and summaries of regulatory studies do not provide the detailed data necessary to determine primary DEB parameters concerning feeding (Jager et al., 2013). To deal with this, DEB uses a simplified feeding input, the scaled functional response, f, which is dimensionless and ranges from zero to one according to food availability. It is generally assumed that f = 1 when food is available *ad libitum* (Jager et al., 2013, Kooijman et al., 2008). However, the data made available for this project did include observations of animals' food consumption, recorded on a regular basis alongside body weight (OECD, 2001, OECD, 2008, OECD, 1998). Simply assuming that f =1 because animals in all treatments are provided with food *ad libitum* would disregard these data, possibly conflating the effects of toxicity and feeding rate on growth rate over time.

DEB assumes that feeding rate is limited by surface area (e.g. area of feeding appendages in filter feeders), which is proportional to body mass to the power 2/3 (Kooijman, 2000). In DEBkiss, the

scaled functional response or scaled feeding rate, f, is the observed area specific feeding rate, J_X^a ($g_{(food)} \times cm_{(length)}^{-2} \times day^{-1}$) scaled as a proportion of the maximum area specific feeding rate, J_{Xm}^a ($g_{(food)} \times cm_{(length)}^{-2} \times day^{-1}$), for a species (Jager et al. (2013) & equation 5.13). The challenge was therefore assigning a value to J_{Xm}^a, which is not simple to determine. In chapter 3, the highest observed area specific feeding rate in a dataset was used, producing feeding inputs that reflected temporal and intertreatment variability in feeding rate. For the first time, this allowed the effects of feeding rate and chemical stress on growth to be modelled separately at high resolution. This capability represents a key advantage of DEB based models over traditional data analysis. We showed how it was possible to isolate the toxic component of observed effects on body weight over time, using only a growth model with no TK-TD model required.

However, the findings presented in chapter 3 should be regarded as a proof of concept rather than accurate breakdowns of observed effects. This is because of significant issues with the way feeding data were entered into the growth model. These issues were noted in chapter 3, but their true extent was shown when they were investigated fully in chapter 5. The first issue was a practical one. The value of f for each weekly interval was calculated - based on the mean area specific feeding rate observed in each treatment - before the model ran. This led to a positive feedback loop or 'snowball effect' when predictions differed from data, as a given f value would correspond to a different mass of food at the predicted body weight than was consumed. This issue is explained in detail in chapter 5 and visualised in figure 5.1. We showed that this problem could be dealt with effectively by instead calculating f as the model runs. This meant that predicted growth rate reflected the mass of food consumed rather than the observed area specific feeding rate. This practice should be adopted in future models of growth which make use of feeding data, particularly when those models are used to predict independent data or extrapolate to novel scenarios.

The second issue was theoretical, concerning DEB assumptions regarding feeding rate. Data showed that area specific feeding rate of rats falls dramatically as they grow (figure 5.10). This contradicts the assumption that animals provided with food *ad libitum* feed at the maximum rate possible for their body size, meaning that the value of f cannot be fixed at 1. The explanatory mechanism we propose is that digestive efficiency increases with body size as increased gut length leads to longer gut transit time as well as greater surface area. If this is indeed the case, then it is likely true for many other species. In chapter 3, f was not fixed but the maximum area specific feeding rate, J_{Xm}^a, was. The result was a modelled growth curve that plateaued not because assimilation from food could no longer exceed maintenance costs, but because animals apparently opted to limit their growth substantially by regulating their feeding. The upper limits

of body weight calculated from model parameters in that study were well above the highest observations for the species, suggesting a major flaw in this method of calculating *f*.

Again, this issue was addressed in chapter 5. We proposed a new method for calculating *f*. Rather than being assigned a fixed value, J_{Xm}^a was calculated as a function of surface area, fitted to observed area specific feeding rates under *ad libitum* food availability. The new approach resulted in improved model fits to mean body weight over time, relative to the methods employed in previous studies. The new method also appeared to increase model realism. Modelled reserve dynamics followed similar patterns to observed body fat percentage as rats age while the upper limit of the growth model was closest to the peak lean weight estimated from observed data.

This new method makes a fundamental adjustment to DEB assumptions. Whereas DEB literature states that the efficiency of digestion is a fixed primary parameter (Kooijman et al., 2008, Jager et al., 2013, van der Meer, 2006a), it is instead assumed that digestive efficiency increases with body size. This would be consistent with observed interspecific patterns, which are related to increases in gut surface area and gut transit time (Demment and Vansoest, 1985, Hansson and Jaarola, 1989, Illius and Gordon, 1992, Smith, 1995). Were the project to continue, collecting intraspecific data to verify or refute this relationship in the laboratory rat would clearly be the next step. Regardless of the outcome, collecting such data for a species commonly used in laboratory studies would be extremely valuable to inform DEB theory and increase model accuracy and applicability.

6.3 Experimental design limits toxicokinetics models

Modelling toxicokinetics in terrestrial vertebrates is a challenge as the primary exposure route for animals in the field is through dietary consumption. Therefore, toxicant concentration inside the body is not simply a function of the concentration of a toxicant in an animal's diet but is also dependent on the rate at which food is consumed. Models developed by Bednarska et al. (2013a) provided a simple method by which to simulate this, with the toxicant being ingested into a depot compartment representing the gut, before being absorbed into, and eliminated from, a single internal compartment according to first order kinetics. In chapter 3, those model equations were adapted to account for dilution by growth, which is essential for predicting internal toxicant concentration in growing animals (Gergs et al., 2016).

We used data from regulatory toxicokinetics studies (OECD, 2010) to calibrate and test the TK models. These data were the best available for the pesticides included in this study but were far from ideal for model development. Firstly, the experiments are designed to assess risk to humans rather than to wildlife exposed to pesticides. For humans working with pesticides, any ingestion would be accidental so toxicokinetics studies generally only consider a single oral dose delivered by gavage rather than in food. Repeated dosing may be included to investigate bioaccumulation,

but this is not a requirement. Sample size is typically only 3-4 individuals of each per treatment, with only two dose levels used throughout.

Using pesticide concentration in the blood as a proxy for overall body burden, the simple TK model based on Bednarska et al. (2013a) was fitted to observed kinetics data for the selected pesticides. However, the models could not be validated *per se* as all experiments used the same two dose levels. Therefore, independent data simply reflected individual variability rather than a different set of conditions. The accuracy of model fits to calibration data was the only way to assess models and these were satisfactory with the simple TK model. While more physiologically accurate TK models have been developed for *R.*norvegicus (Li et al., 2017, Louisse et al., 2015), this limitation of the data meant there was no way to determine whether added model complexity conferred a meaningful advantage. A TK-TD model only requires that predicted internal concentration is proportional to the true value (Jager and Zimmer, 2012). So long as this requirement is met, the fitting of TD parameters can compensate for the mismatch.

6.4 Modelling growth under chemical stress – a promising first attempt

Much like toxicokinetics studies, dietary toxicity studies are designed to assess the risk of chronic pesticide exposure to humans, primarily through residues left on food (EFSA, 2006). Accordingly, the summary statistics derived from results, such as the no observed adverse effect level (NOAEL), relate to the concentration in food resulting, or not, in a specified effect. Measures like the NOAEL are just as dependent on duration of exposure as toxicant concentration (Jager, 2011, Jager, 2012). Moreover, in dietary toxicity studies, dose is also dependent on the rate of food consumption. This this leads to a lack of ecological realism when using the results to determine ecological risk posed by a pesticide(EFSA, 2006). TK-TD models have been advocated as a means of dealing with these issues by providing a more detailed assessment of exposure profiles and intrinsic toxicity of pesticides (Ashauer et al., 2013, Ducrot et al., 2016).

In Chapter 3, our simple TK model, DEBkiss growth model, and DEBtox stress functions were combined as a TK-TD growth model and assessed against 34 dietary toxicity datasets for six pesticides. TK predictions could not be verified, as internal pesticide concentration is not monitored during dietary toxicity studies. Likewise, modelled 'stress' (figs. 3.2 & 3.3) could not be compared to data as this is an abstract concept for which data cannot be collected. However, the data provided - pesticide concentration in diet, food consumption per day and body weight – could be combined to produce accurate records of mean ingested dose per treatment ($mg_{(pesticide)}$ $\times kg_{(body\ weight)}^{-1} \times day^{-1}$), which varied substantially over time. These served as highly detailed TK model inputs. While the model components could not be assessed individually, performance of the full TK-TD growth model could be assessed through comparison to observed growth under

chemical stress. For 28 out of 34 datasets, ≥75% of mean body weight observations were predicted to within one standard deviation.

Again, these results represent proof of concept rather than the finished product. It is unlikely that the values of the fitted TD parameters truly reflected the intrinsic toxicity of the compounds. There were no data to verify the predictions of internal concentration and the TK models were calibrated to data from studies which used a different dosing method. Without the inclusion of reproduction data, the mode of action can only be provisionally identified (Desforges et al., 2017, Jager et al., 2006) and there were the issues with the calculation of the scaled feeding rate, discussed above. Regardless of all these issues though, the effects of dietary toxicity on growth were generally predicted accurately - across datasets differing in sample size, dose level, duration, and food consumption rates - with fixed TK-TD parameters. The results suggest that predictions of internal pesticide concentrations were at least proportional to their actual values and that average uptake and elimination rates were consistent between groups. Likewise, the stress placed on growth parameters by the pesticides also appeared consistent and predictable, lending support to the assumptions of the DEBtox stress functions. These findings are extremely encouraging, especially considering this was the first attempt at using regulatory data to model the effects of dietary toxicity on growth of a mammalian species. Moreover, an emphasis was placed on using the simplest available methods and, though some issues remain, they appeared sufficient for the task.

6.5 Mismatch in dose route hinders *in vitro – in vivo* extrapolation of growth effects

Perhaps the most ambitious application of TK-TD modelling considered in this study was quantitative *in vitro-in vivo* extrapolation of growth effects in rats. This capability has been demonstrated previously for effects on growth in fish (Stadnicka-Michalak et al., 2015). In Chapter 4, we investigated whether this was also possible for rodents, which make up around 76% of the animals used in toxicology studies in the EU (European Commission, 2013b). We found that several additional obstacles made QIVIVE more complex for mammals and found no relationships between effects of intracellular pesticide concentration on *in vitro* popultaion and *in vivo* growth rate.

The principal issue throughout was the difference in dose route between studies. Chronic toxicity in fish is monitored by exposing fish to a constant concentration of a pesticide in their water (OECD, 2013), this dose route is also used in toxicokinetics studies on fish (OECD, 2012). This is analogous to cultured cells being exposed to a constant concentration of a compound in their medium, meaning that *in vitro* and *in vivo* data can be readily compared. For mammals in dietary toxicity studies, the situation is more complex. While pesticide concentration in the diet is

constant (OECD, 2008, OECD, 1998, OECD, 2001), the ingested dose ($mg_{(pesticide)} \times kg_{(body\ weight)}^{-1} \times day^{-1}$) varies substantially over the study period. Therefore, direct comparison between *in vitro* and *in vivo* data is not possible. In this regard, TK-TD models have great potential to act as a bridge, by predicting the effects of a truly constant exposure scenario (constant internal concentration) on growth, eliminating noise from observed data. However, this requires great confidence in the precision of TK-TD model predictions, and at this point the required precision cannot be achieved.

Again, the problem is primarily related to dose route. Rats in toxicokinetics studies are dosed orally but this is delivered by gavage rather than in the diet (OECD, 2010). This leaves uncertainty as to how bioavailability changes between toxicokinetics studies and dietary toxicity studies. As discussed above, this mismatch is not a major issue for an *in vivo* TK-TD growth model (Jager and Zimmer, 2012). However, for QIVIVE it is vital that internal concentration is predicted accurately, so that it can be replicated *in vitro* (Stadnicka-Michalak et al., 2015). In Chapter 4, experiments found no consistent effects of pesticide concentration on cell population growth. Even if effects had been shown, without confidence in the relationship between intracellular concentration achieved *in vitro* and *in vivo*, it would still not have been possible to accurately extrapolate effects between the two.

While there are other obstacles - such as identifying the optimal cell line and assays - the inconsistency in dose route between datasets is the most fundamental barrier to QIVIVE of effects on graded sublethal endpoints, such as growth, in rodents. Without the inclusion of dietary dosing in toxicokinetics studies it is not possible to predict internal pesticide concentration *in vivo* with the necessary precision to inform *in vitro* experimental design. This gap in the data must be addressed before any further attempts to use the methods outlined in Chapter 4 can be successful.

6.6 Final thoughts

This project has shown that the DEBtox modelling framework is suitable to predict the effects of chemical stress on growth rate in one of the world's most commonly used study animals (European Commission, 2013b), the laboratory rat. Moreover, this has been achieved using models based on the DEBkiss framework throughout, helping to address issues around accessibility for regulators who are not DEB specialists (EFSA, 2018). However, further work is needed to develop models suitable for use in ERA.

Ultimately, it is not enough for a TK-TD model to simply predict observed effects accurately. The modelling procedure must derive time independent TD parameters that precisely reflect the intrinsic toxicity of a compound. This allows extrapolation of toxic effects to realistic exposure scenarios, conferring a clear advantage over time dependent summary statistics such as the

NOAEL (Ducrot et al., 2016), which are more suited to assessing risk to human health than ecological risk. The task is not a simple one and, while the existing modelling frameworks themselves are fit for purpose, data requirements remain the key barrier to overcome.

In this project, specific data gaps have been identified that, if filled, would greatly improve the prospect of DEBtox models meeting the requirements for use in ERA. Data concerning the relationship between digestive efficiency and body size would validate, or more likely refine, the methods proposed in Chapter 5. Not only would this improve the accuracy of TD parameters by minimising conflation of effects due to toxicity and feeding rate, but this would also make the growth model alone an extremely valuable tool with which to analyse the results of dietary toxicity studies. The inclusion of dietary dosing in toxicokinetics studies would greatly increase the accuracy of TK models used to predict internal concentration in dietary toxicity studies. This would in turn increase the accuracy of TD parameters and remove a significant obstacle to QIVIVE of graded sublethal effects linked to chronic dietary exposure to pesticides.

TK-TD models and other *in silico* methods have great potential to complement and even replace animal testing in risk assessment. However, animal data are required for that potential to be realised and the window for these data to be collected appears limited. The US Environmental Protection Agency has committed to a 30% reduction in its requests for, and funding of, mammal studies by 2025 and the elimination of requests and funding by 2035 (EPA Press Office, 2019). The EU meanwhile is committed to the refinement, reduction and replacement of animal testing (European Commission, 2013a) and has faced public pressure to strengthen existing legislation limiting the use of animal testing (Peter, 2015). I echo calls for a break from the status quo (Jager, 2020). Rather than modellers making do with the data made available to them, a more cyclical process is required, with animal experiments designed to consider model parameterisation. Without this shift in approach occurring soon, we risk a situation where animal testing is phased out before *in silico* methods are fit to act as replacements.

Appendix A – Supporting Figures for Chapter 3

Acibenzolar-S-Methyl Toxicokinetics
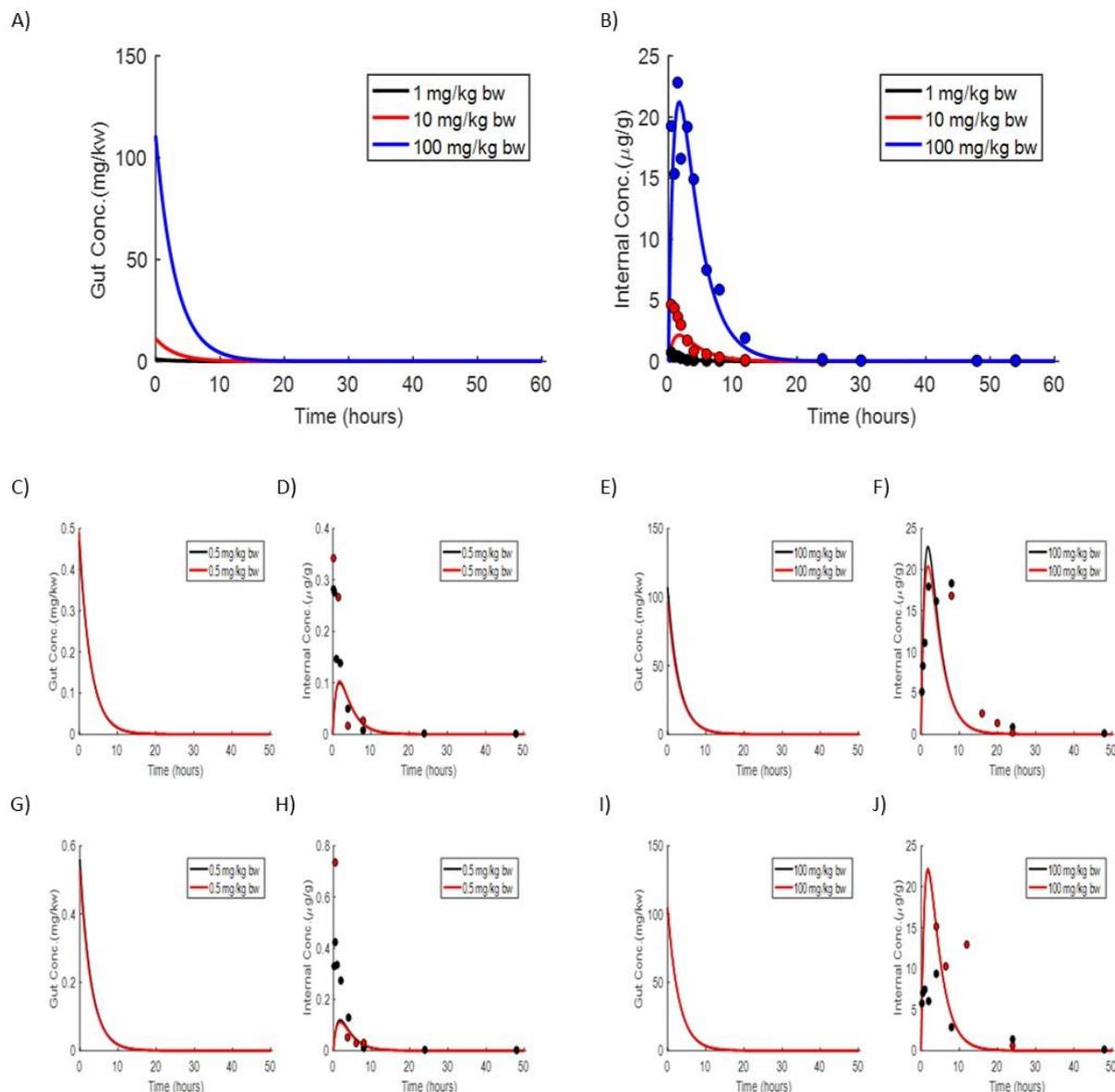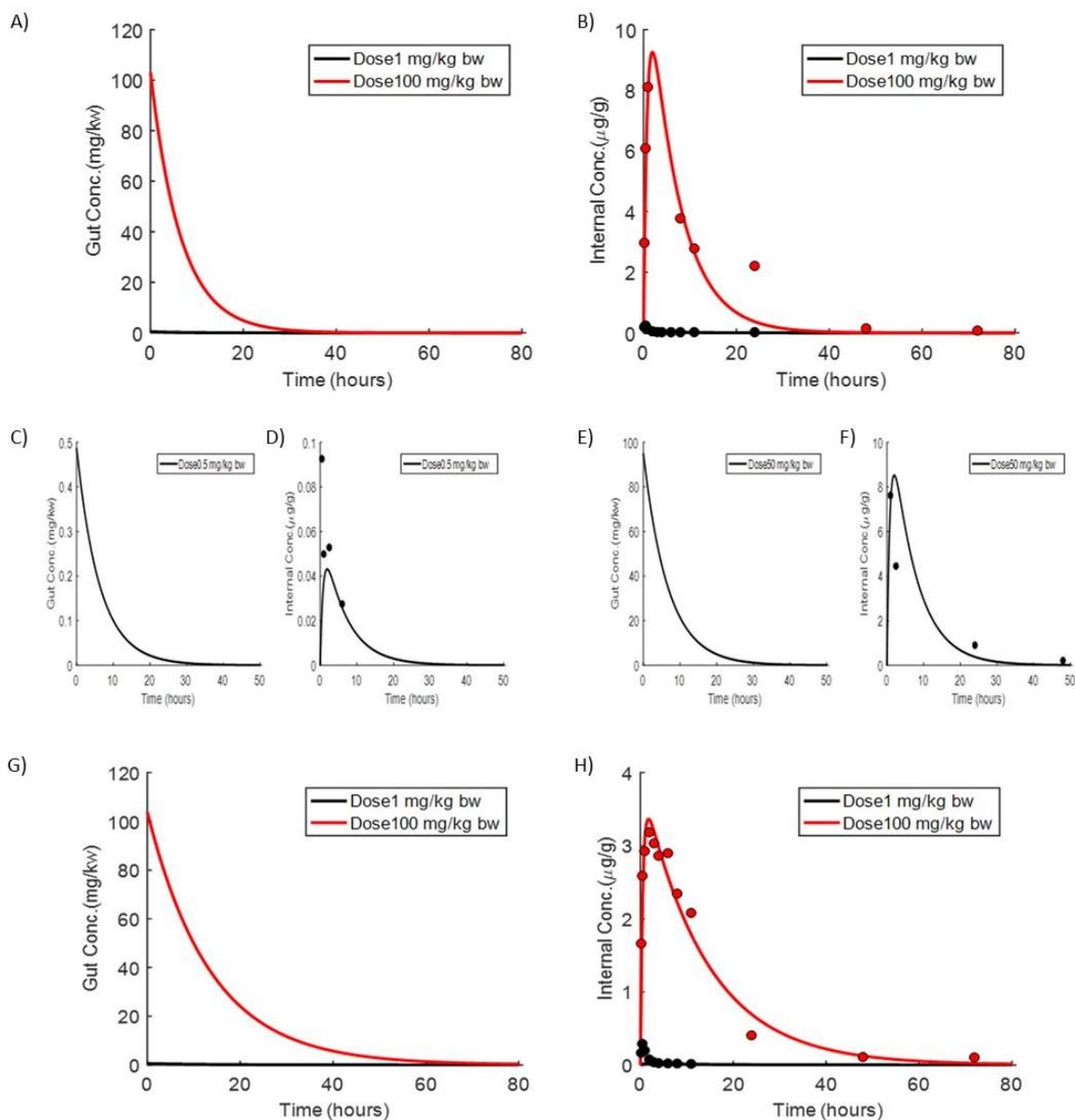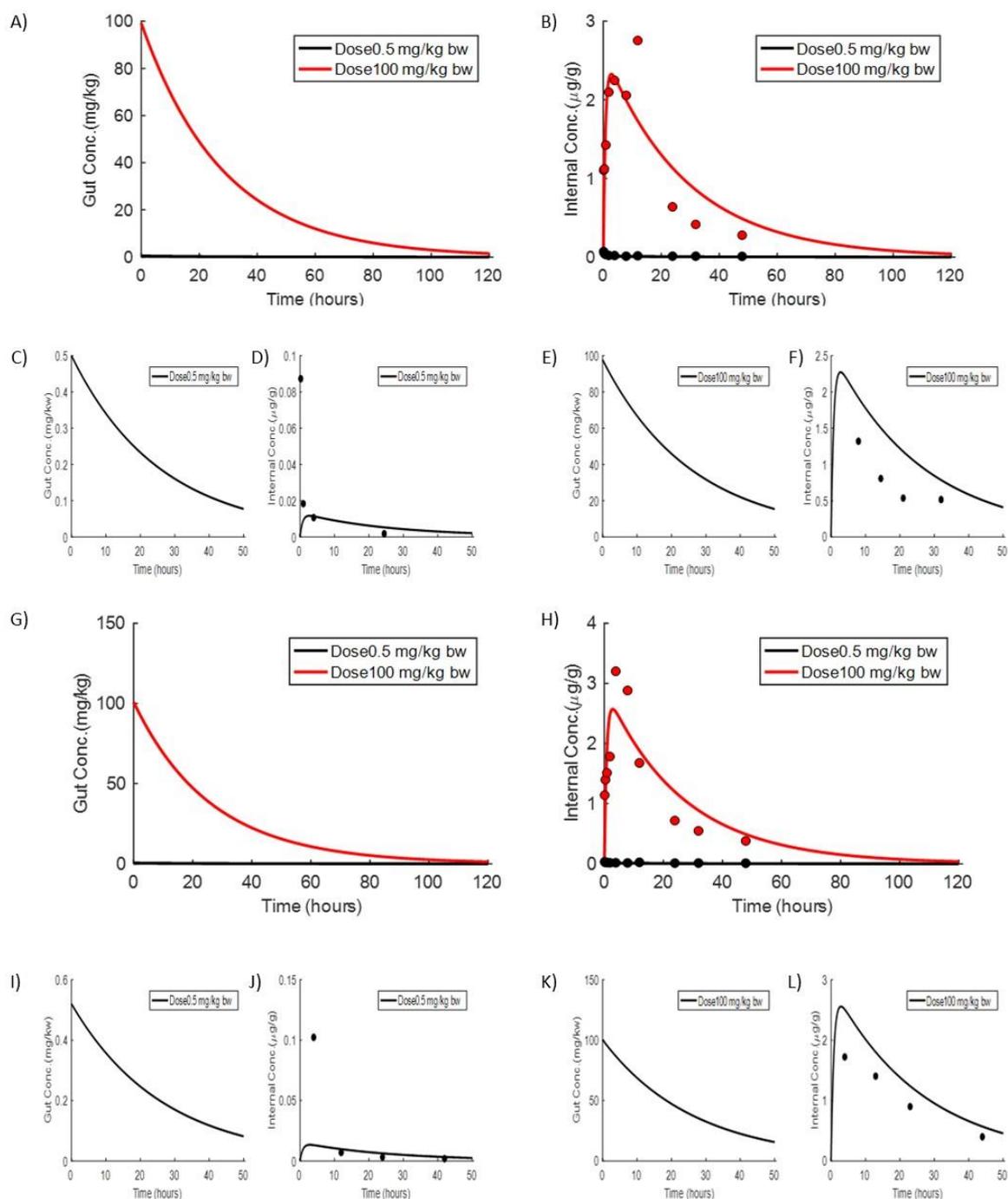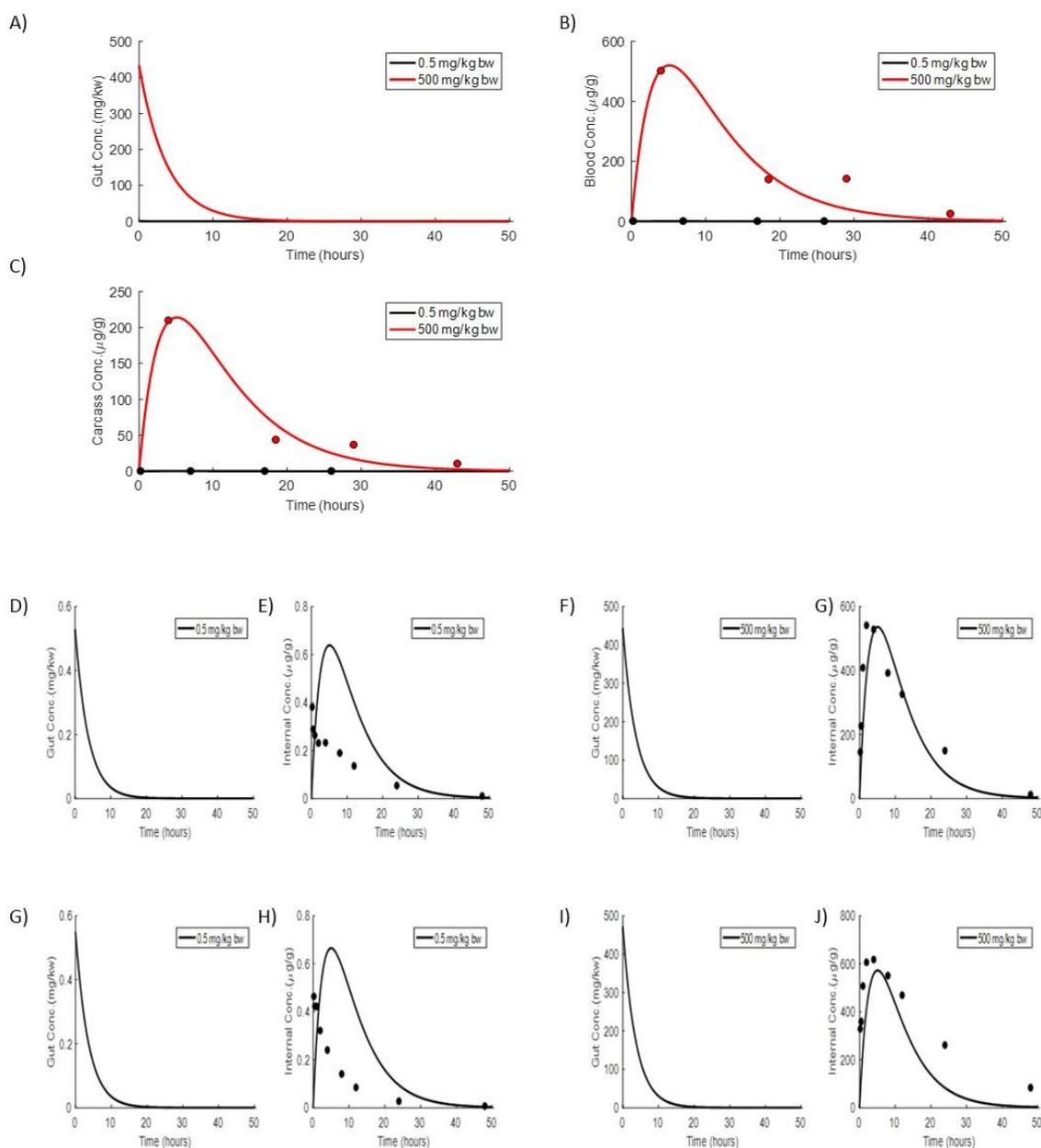

Figure A1 Modelled gut (A) and internal (B) concentrations over time when fitted to mean blood concentration data (circles) from female rats at three dose levels.

Plots C-F show model predictions with fixed parameters, data points (circles) are from two different groups of male rats dosed at 0.5mg/kg body weight (C&D) and 100mg/kg body weight (E&F). Slightly different curves are produced owing to differences in the average achieved dose for each group.

Plots I-J show model predictions with fixed parameters, data points (circles) are from two different groups of female rats dosed at 0.5mg/kg body weight (G&H) and 100mg/kg body weight (I&J). Slightly different curves are produced owing to differences in the average achieved dose for each group.

Azoxystrobin Toxicokinetics
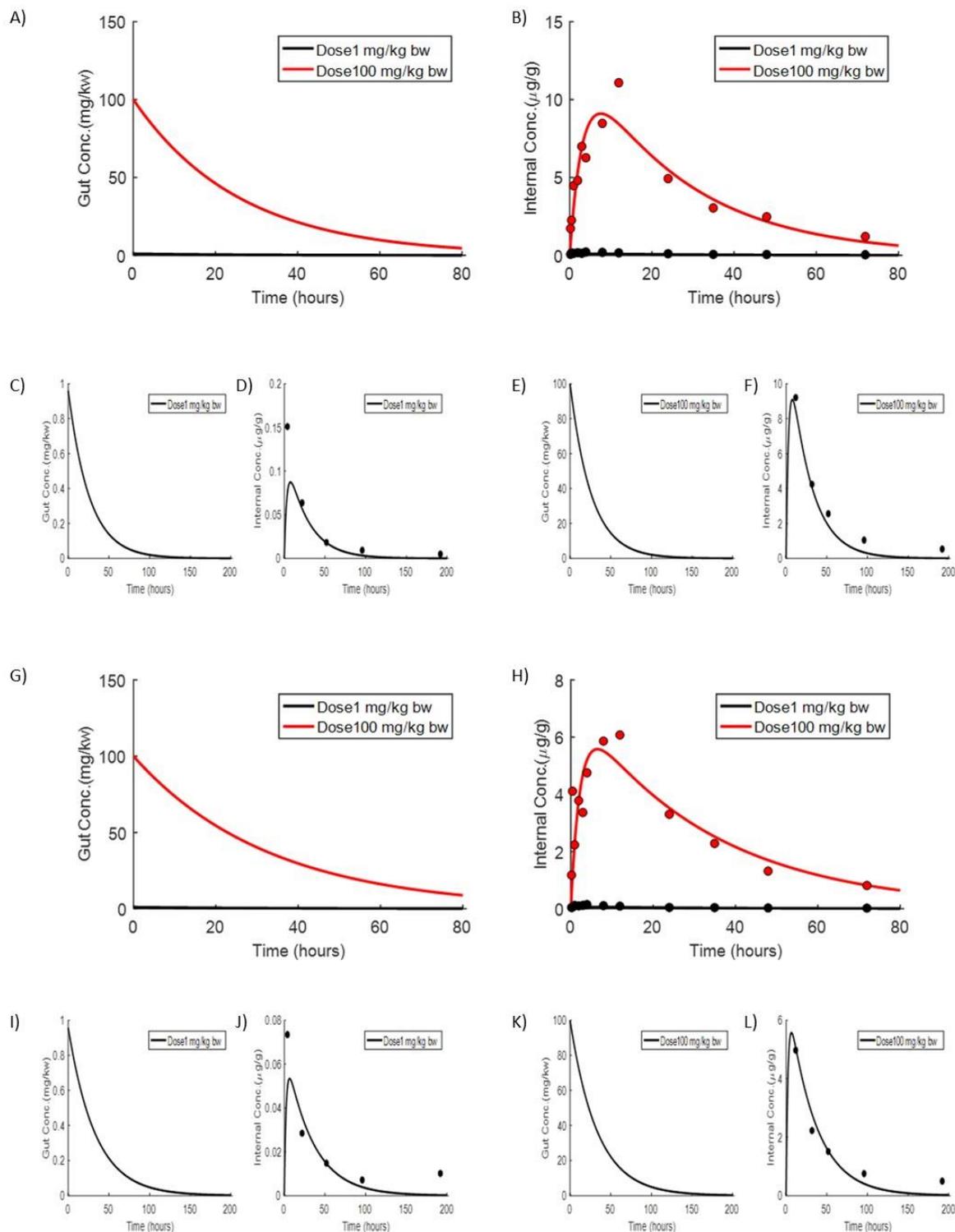
Figure A2 Modelled gut (A) and internal (B) concentrations over time when fitted to mean blood concentration data (circles) from male rats at two dose levels. Plots C-F show model predictions with fixed parameters. Data points (circles) are from different groups of male rats also dosed at 1mg/kg body weight (C&D) and 100mg/kg body weight (E&F).

Modelled gut (G) and internal (H) concentrations over time when fitted to blood concentration data (circles) from female rats at two dose levels. Plots I-L show model predictions with fixed parameters. Data points (circles) are from different groups of female rats also dosed at 1mg/kg body weight (I&J) and 100mg/kg body weight (K&L).

Fenpropidin Toxicokinetics

Figure A3 Modelled gut (A) and internal (B) concentrations over time when fitted to mean blood concentration data (circles) from male rats at two dose levels. Plots C-F show model predictions with fixed parameters. Data points (circles) are from different groups of male rats also dosed at 1mg/kg body weight (C&D) and 100mg/kg body weight (E&F).

Modelled gut (G) and internal (H) concentrations over time when fitted to mean blood concentration data (circles) from female rats at two dose levels. No alternative data set was available for females to compare to predictions with fixed parameters.

Fludioxonil Toxicokinetics

Figure A4 Modelled gut (A) and internal (B) concentrations over time when fitted to mean blood concentration data (circles) from male rats at two dose levels. Plots C-F show model predictions with fixed parameters. Data points (circles) are from different groups of male rats also dosed at 0.5mg/kg body weight (C&D) and 100mg/kg body weight (E&F).

Modelled gut (G) and internal (H) concentrations over time when fitted to mean blood concentration data (circles) from female rats at two dose levels. Plots I-L show model predictions with fixed parameters. Data points (circles) are from different groups of female rats also dosed at 0.5mg/kg body weight (I&J) and 100mg/kg body weight (K&L).

Prosulfuron Toxicokinetics

Figure A5 Modelled gut (A), blood (B) and carcass (C) toxicant concentrations over time when fitted to mean blood and carcass concentration data (circles) from male rats at two dose levels.

Plots D-G show model predictions with fixed parameters. Data points (circles) are from different groups of male rats also dosed at 0.5mg/kg body weight (D&E) and ~500mg/kg body weight (F&G).

Plots G-J show model predictions with fixed parameters. Data points (circles) are from different groups of female also dosed at 0.5mg/kg body weight (G&H) and ~500mg/kg body weight (I&J).

Thiamethoxam Toxicokinetics
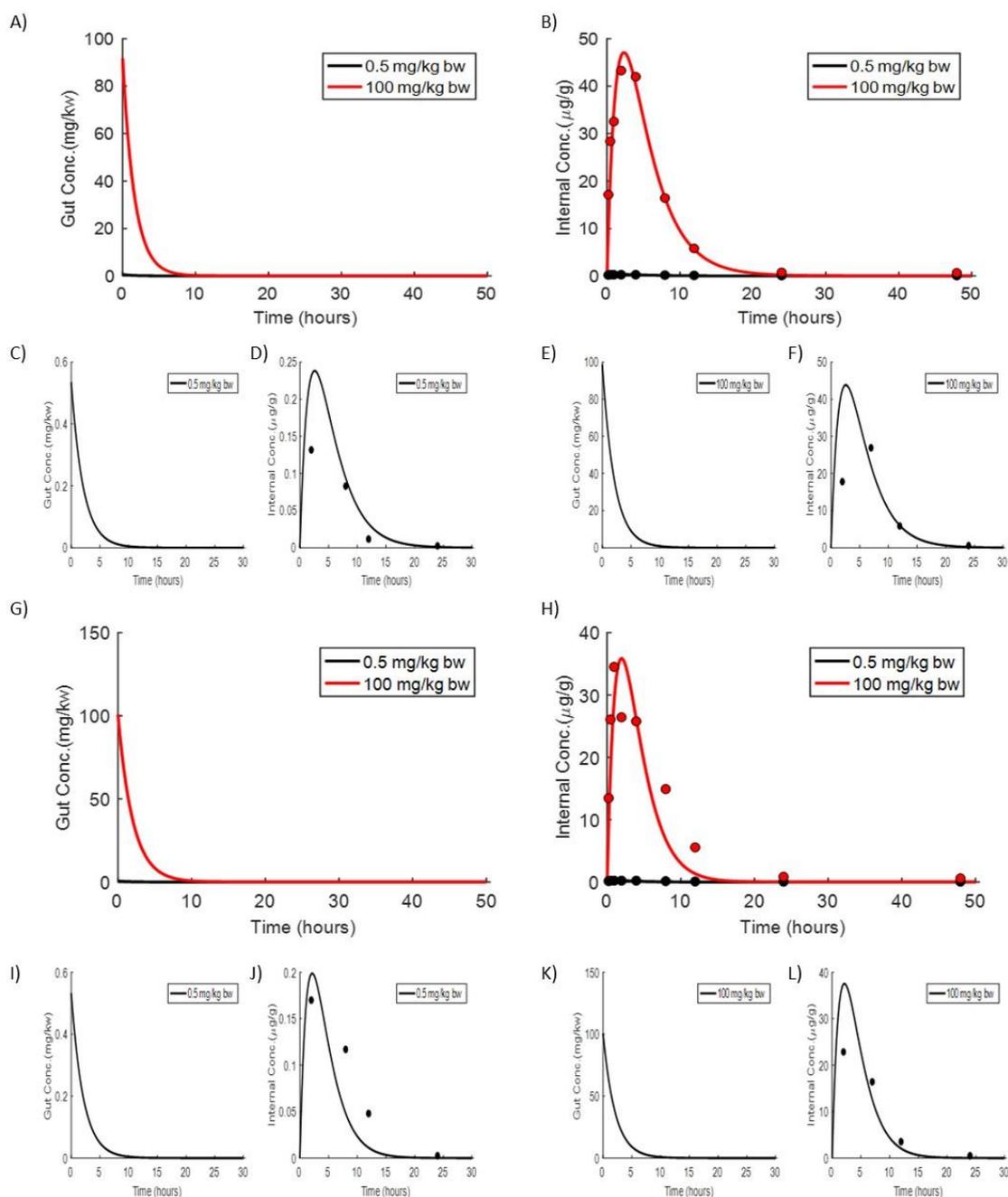
Figure A6 Modelled gut (A) and internal (B) concentrations over time when fitted to mean blood concentration data (circles) from male rats at two dose levels. Plots C-F show model predictions with fixed parameters. Data points (circles) are from different groups of male rats also dosed at 0.5mg/kg body weight (C&D) and 100mg/kg body weight (E&F).

Modelled gut (G) and internal (H) concentrations over time when fitted to mean blood concentration data (circles) from female rats at two dose levels. Plots I-L show model predictions with fixed parameters. Data points (circles) are from different groups of female rats also dosed at 0.5mg/kg body weight (I&J) and 100mg/kg body weight (K&L).

TK-TD model outputs for male rats dosed with Acibenzolar-S-Methyl

90 Day Study

A)

B)

C)
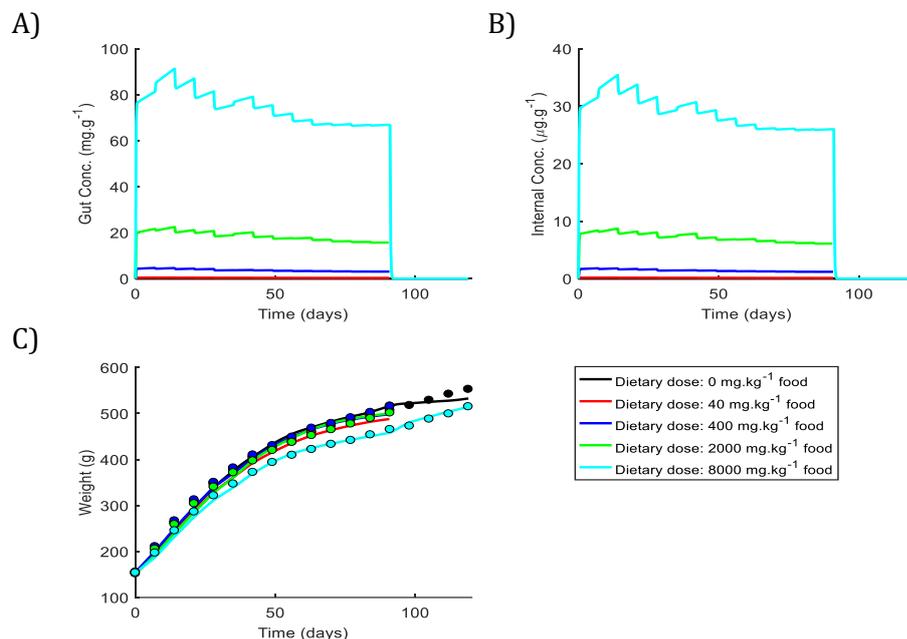
Figure A7: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles). NOTE: The control group and the highest dose group were observed for an additional 4 weeks with no pesticide in the diet to investigate recovery.

28 Day Study

A)

B)

C)
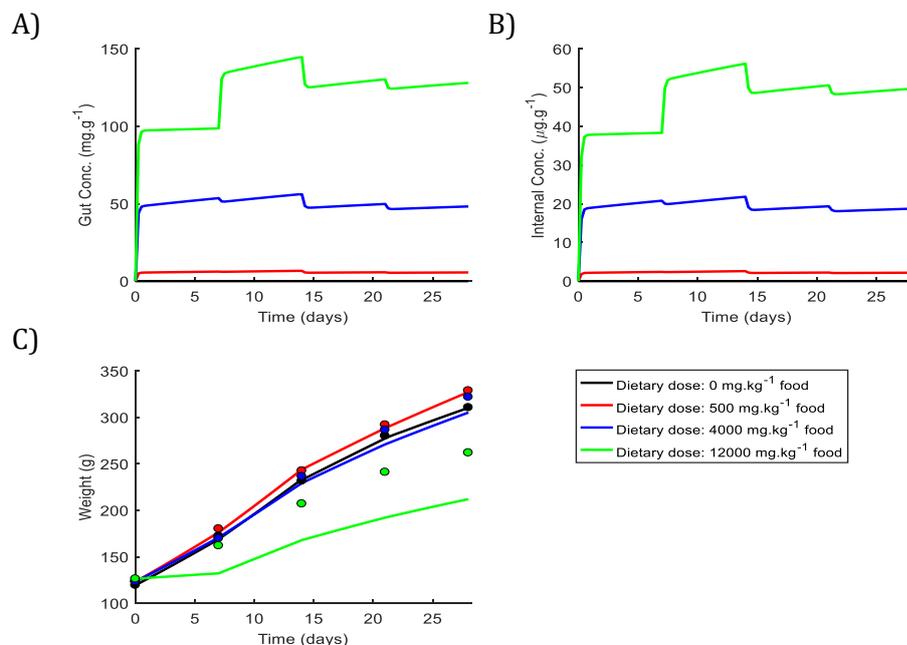
Figure A8: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
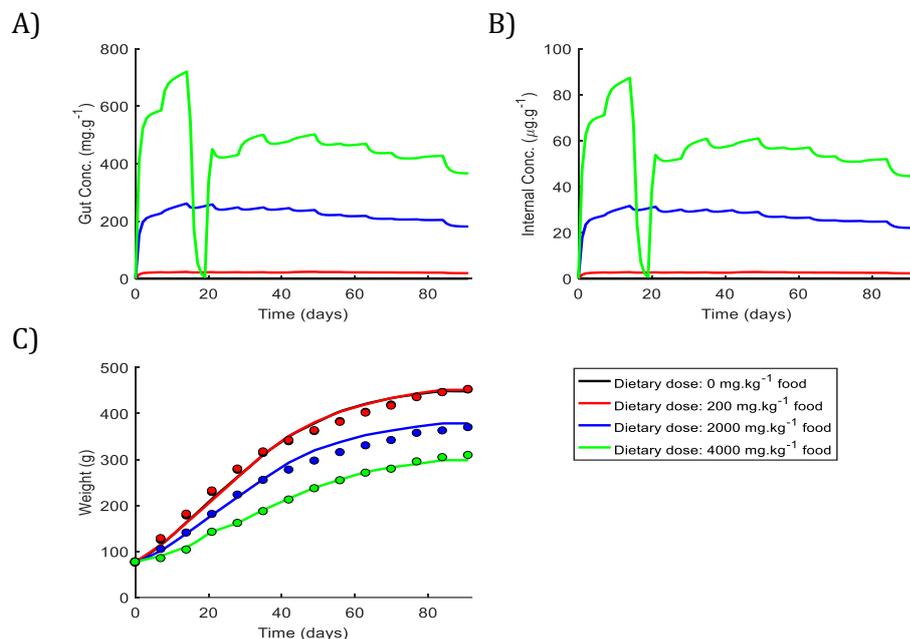
Two Year Study

A)

B)

C)
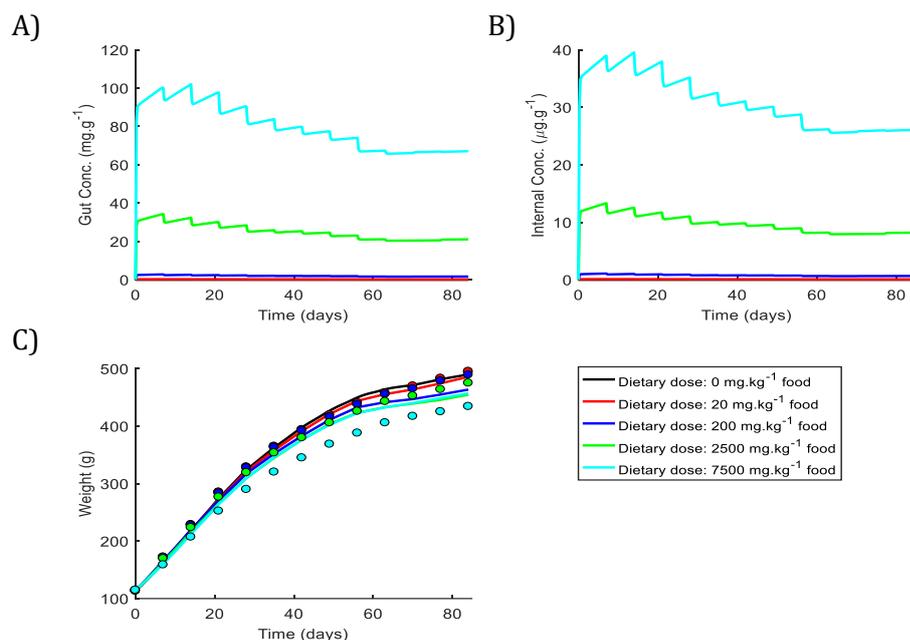
Figure A9: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
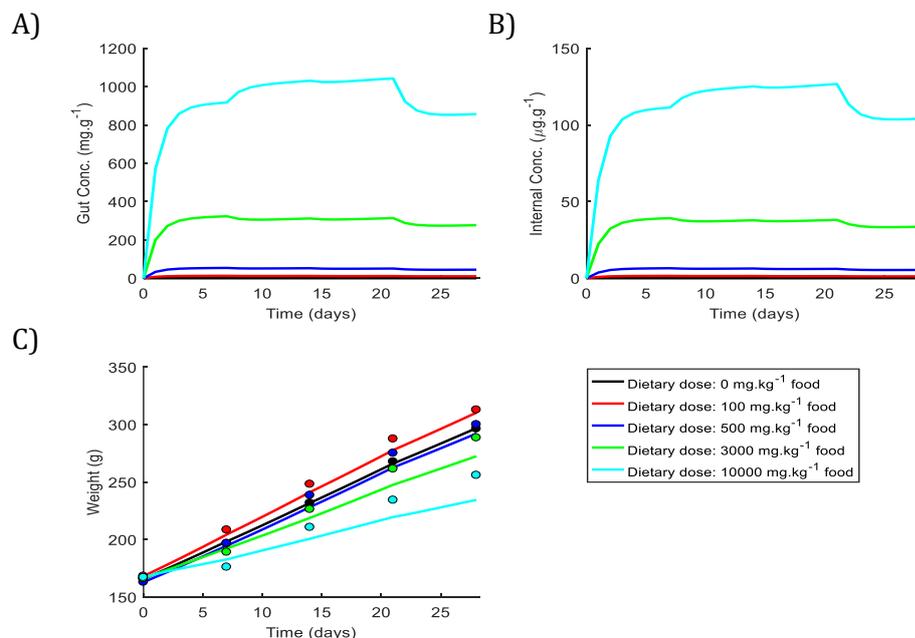
TK-TD model outputs for female rats dosed with Acibenzolar-S-Methyl

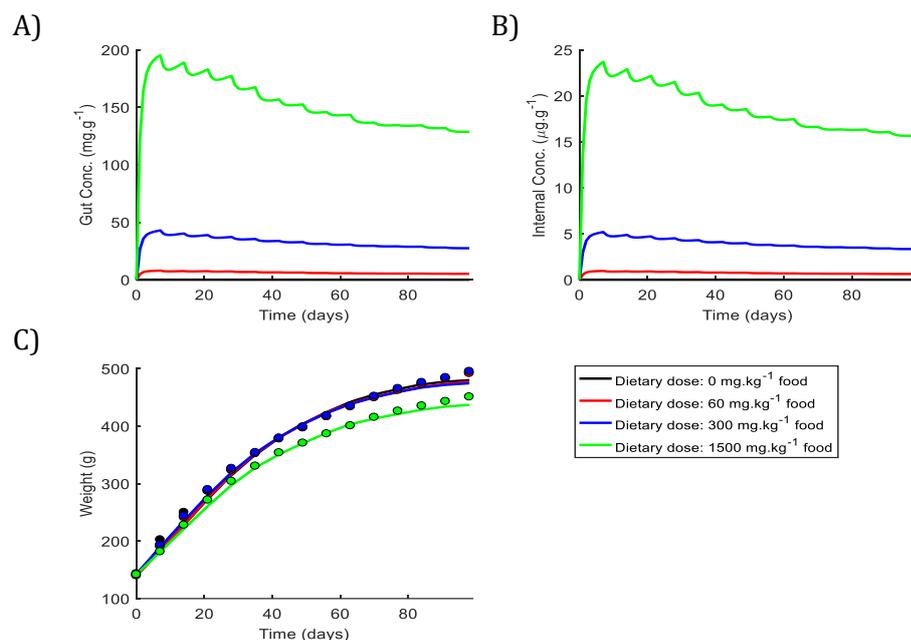**90 Day Study**

A)

B)

C)
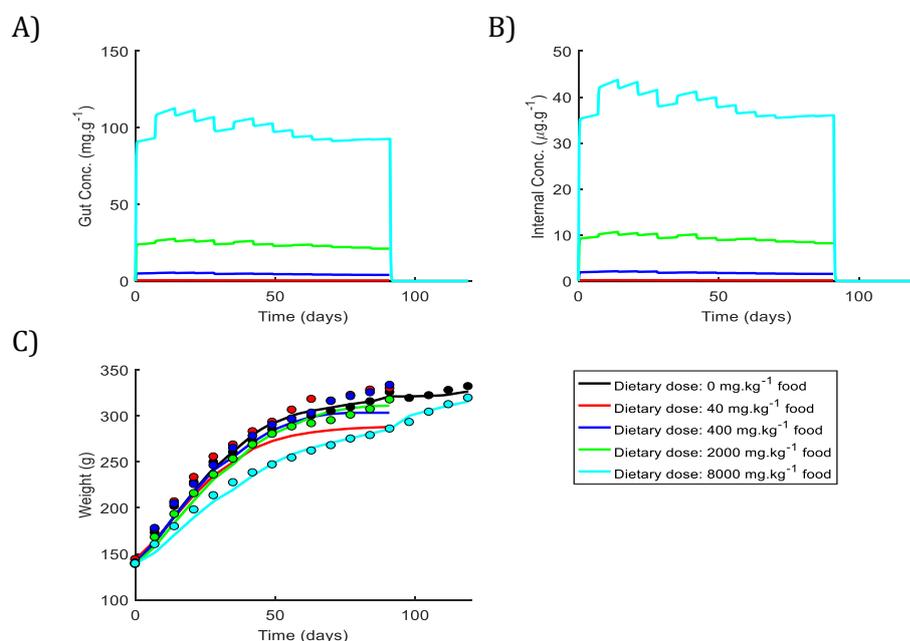
Figure A10: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles). NOTE: The control group and the highest dose group were observed for an additional 4 weeks with no pesticide in the diet to investigate recovery.

28 Day Study

A)

B)

C)
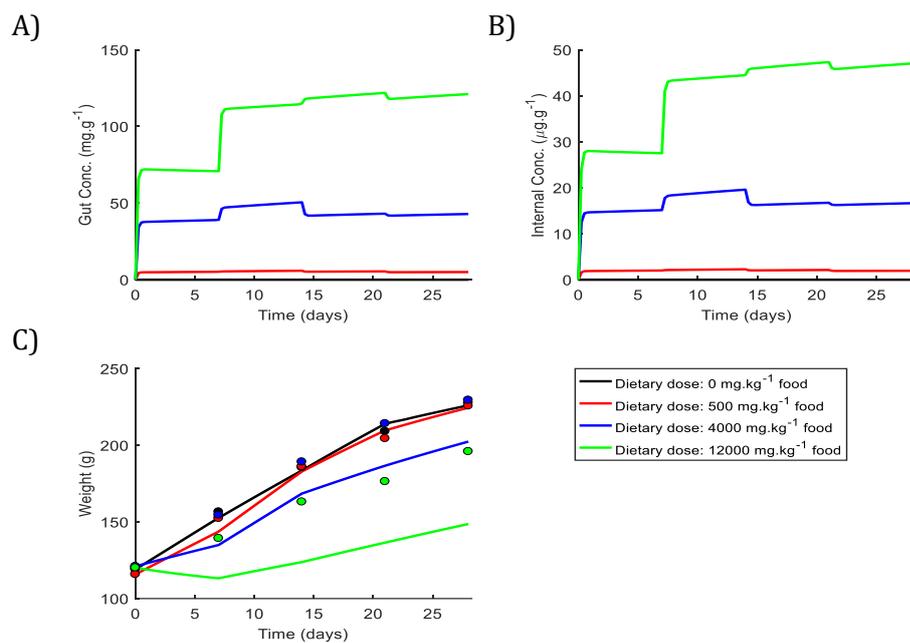
Figure A11: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

Two Year Study

A)

B)

C)
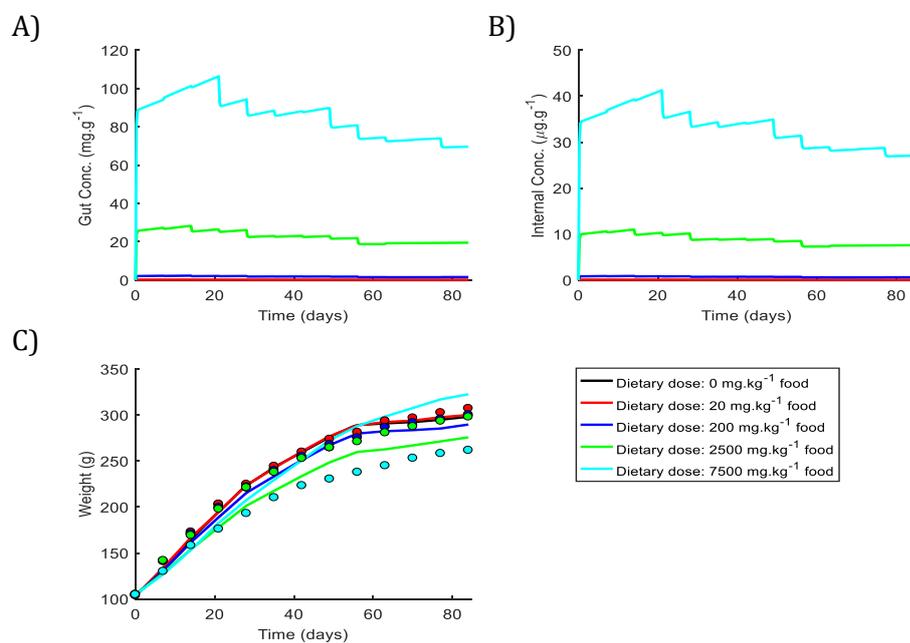
Figure A12: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
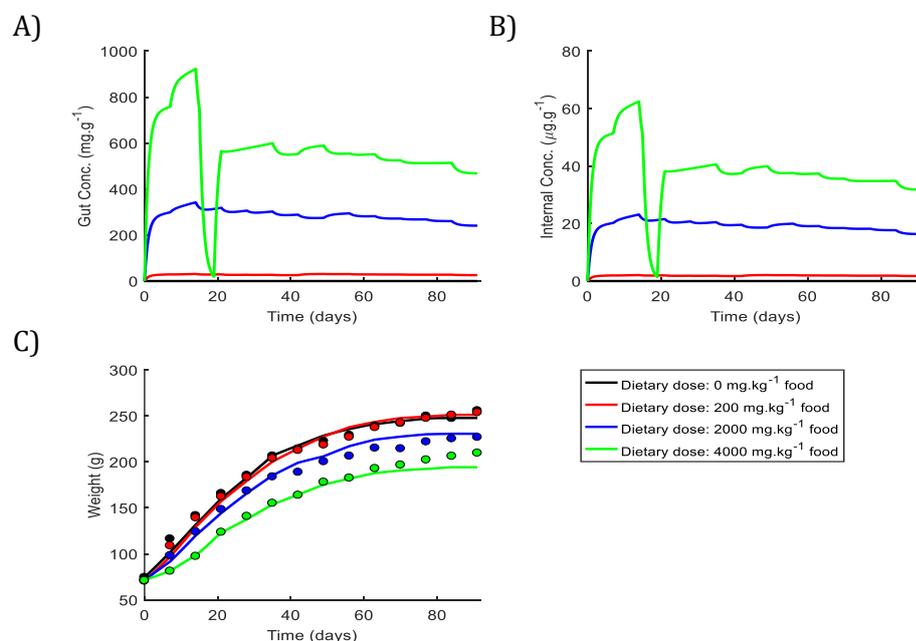
TK-TD model outputs for male rats dosed with Azoxystrobin

90 Day Study

Figure A13: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of bod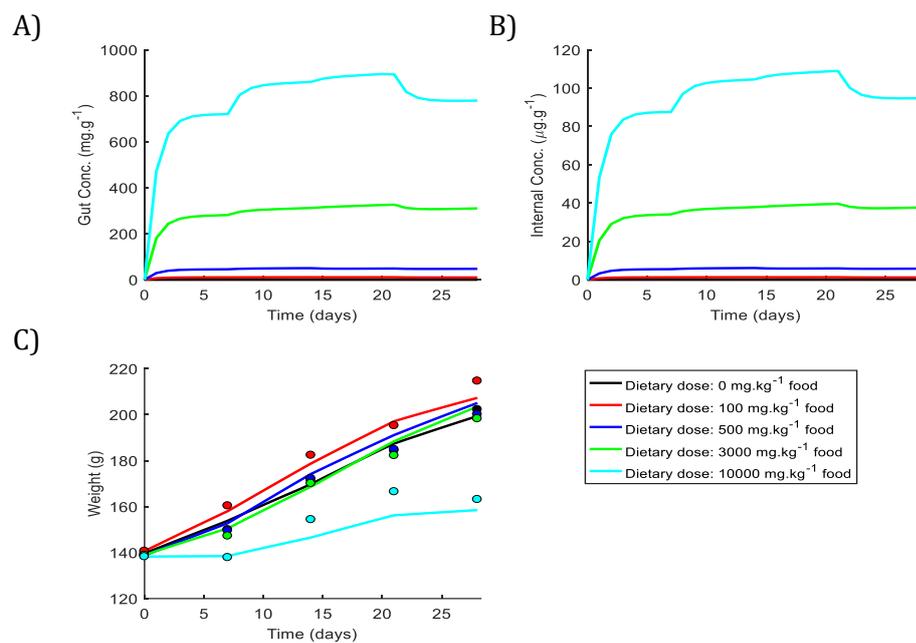y weight are also shown (circles). NOTE: The highest dose group was initially 6000 mg.kg-1 food but was reduced after 2 weeks to 4000 mg.kg-1 food following 5 days of feeding on a control diet.
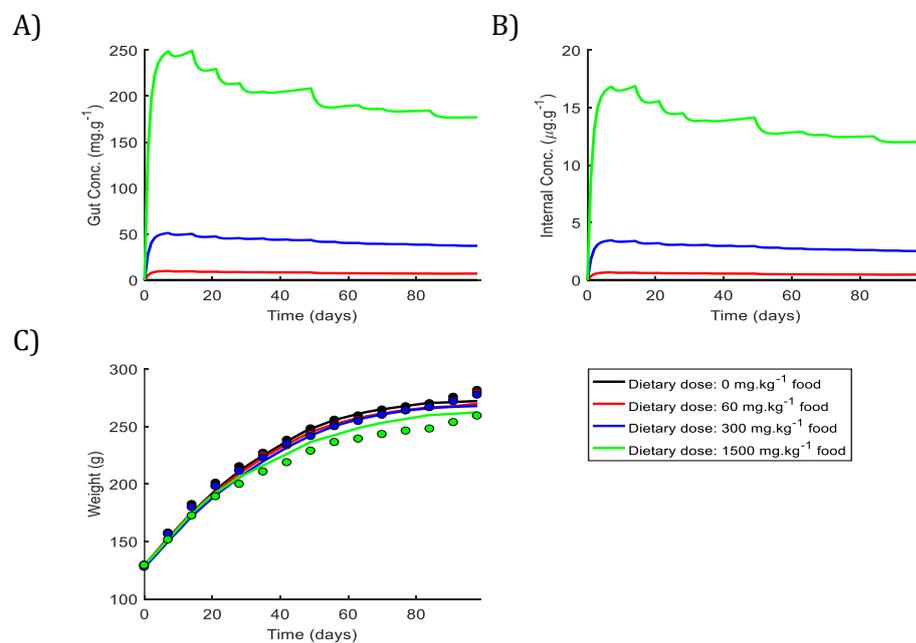
28 Day Study

Figure A14: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

Two Year Study

Figure A15: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for female rats dosed with Azoxystrobin

90 Day Study

Figure A16: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles). NOTE: The highest dose group was initially 6000 mg.kg^{-1} food but was reduced after 2 weeks to 4000 mg.kg^{-1} food following 5 days of feeding on a control diet.

28 Day Study

A)

B)

C)

Figure A17: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
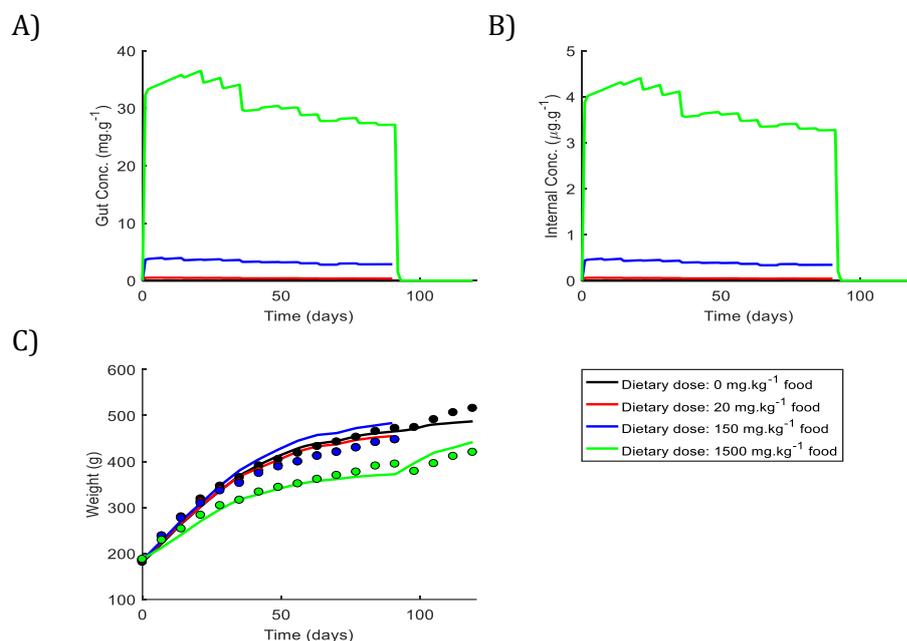
Two Year Study

A)

B)

C)

Figure A18: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
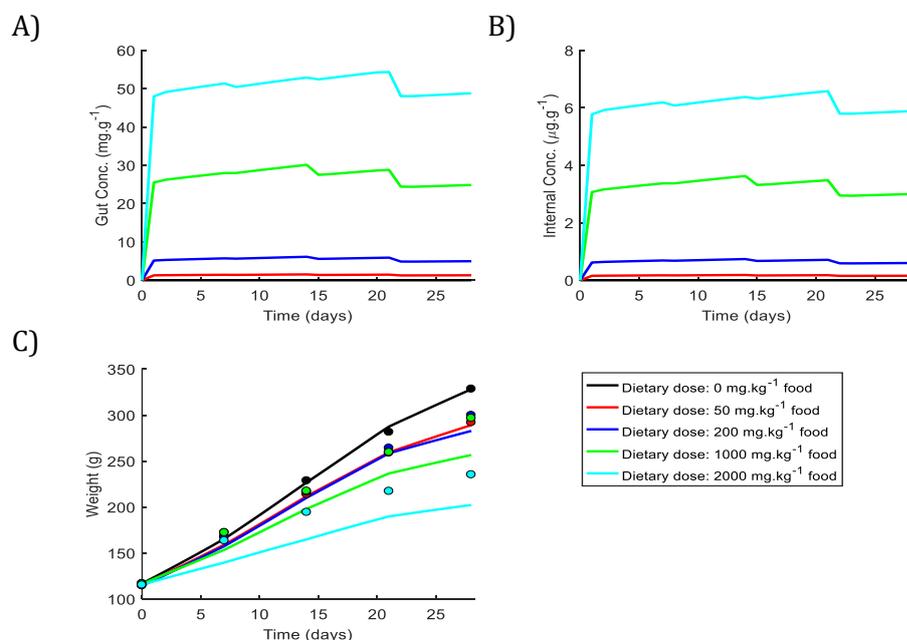
TK-TD model outputs for male rats dosed with Fenpropidin
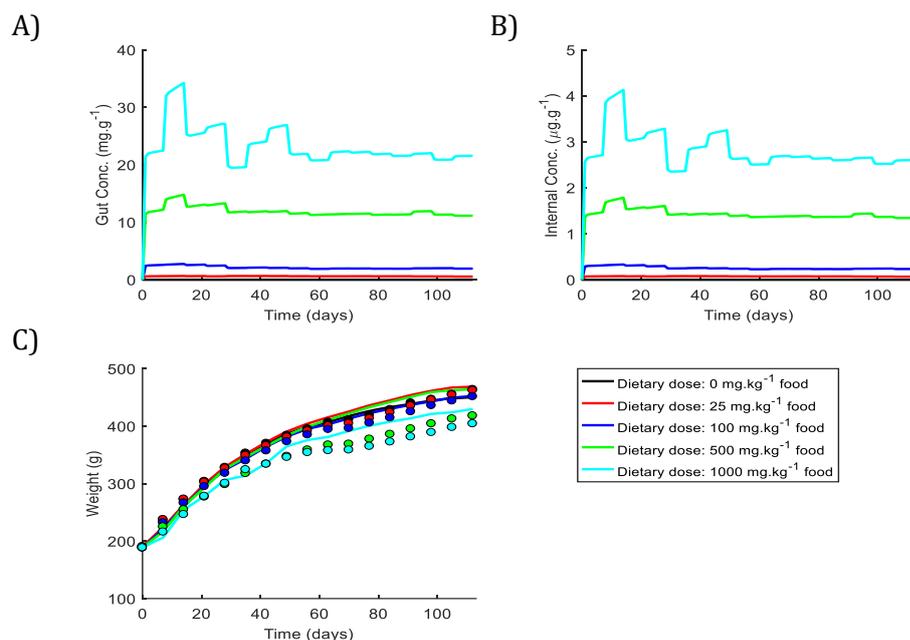
90 Day Study

Figure A19: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles). NOTE: The control group and the highest dose group were observed for an additional 4 weeks with no pesticide in the diet to investigate recovery.

28 Day Study

Figure A20: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

Two Year Study

A)

B)

C)

Figure A21: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
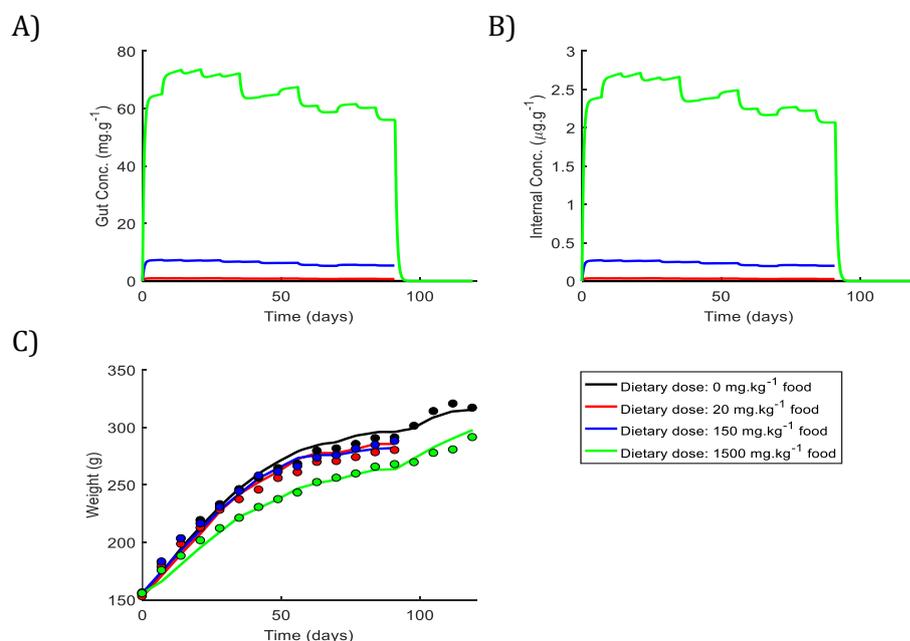
TK-TD model outputs for female rats dosed with Fenpropidin

90 Day Study

A)

B)

C)

Figure A22: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body 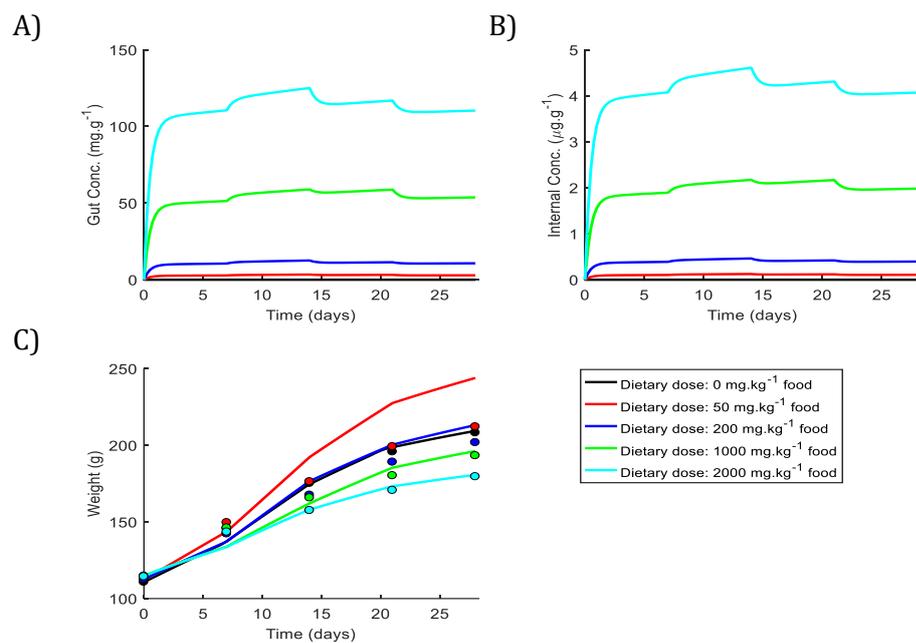weight are also shown (circles). NOTE: The control group and the highest dose group were observed for an additional 4 weeks with no pesticide in the diet to investigate recovery.

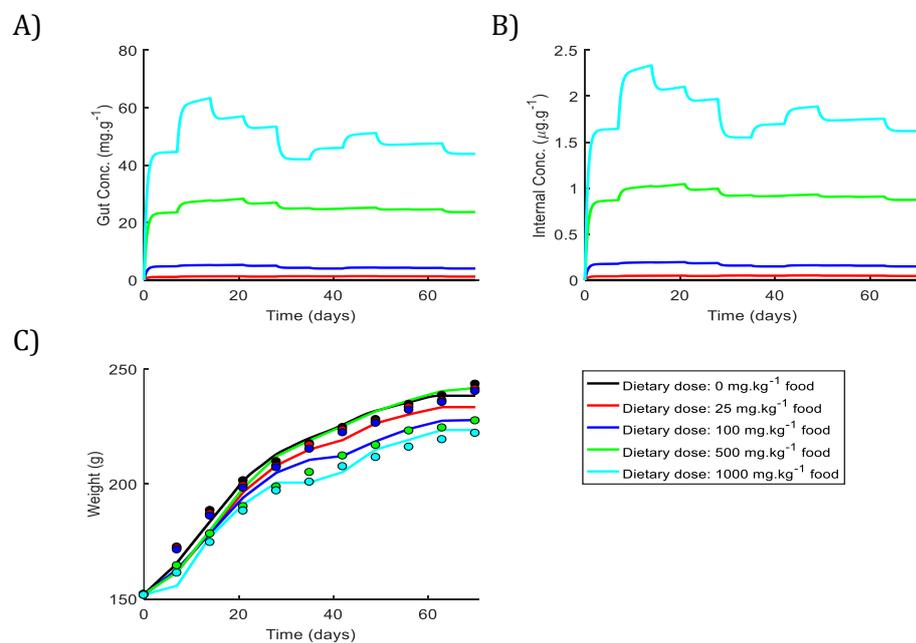**28 Day Study**

A)

B)

C)

Figure A23: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

Two Year Study

A)

B)

C)

Figure A24: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for male rats dosed with Fludioxonil

90 Day Study

Figure A25: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
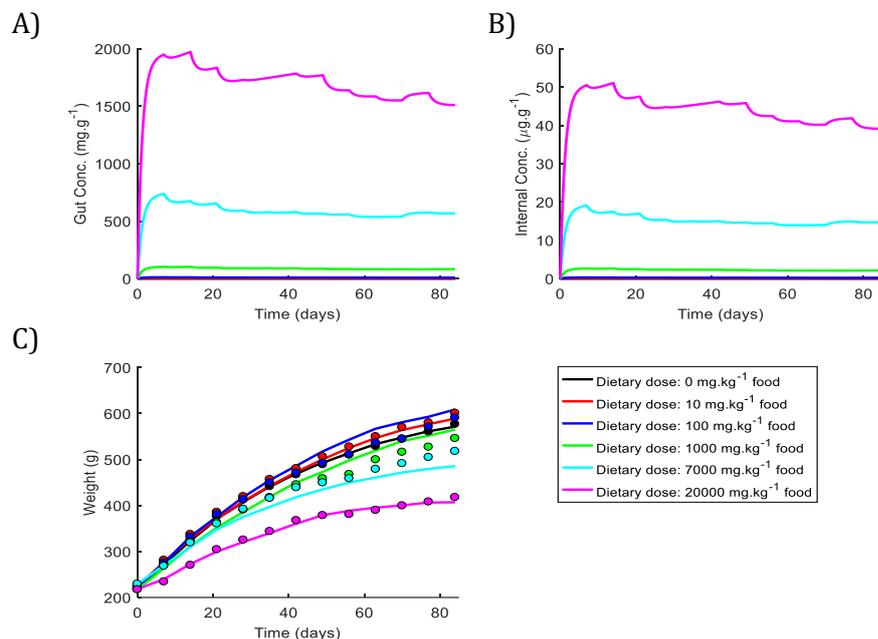
28 Day Study

Figure A26: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for female rats dosed with Fludioxonil

90 Day Study

Figure A27: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
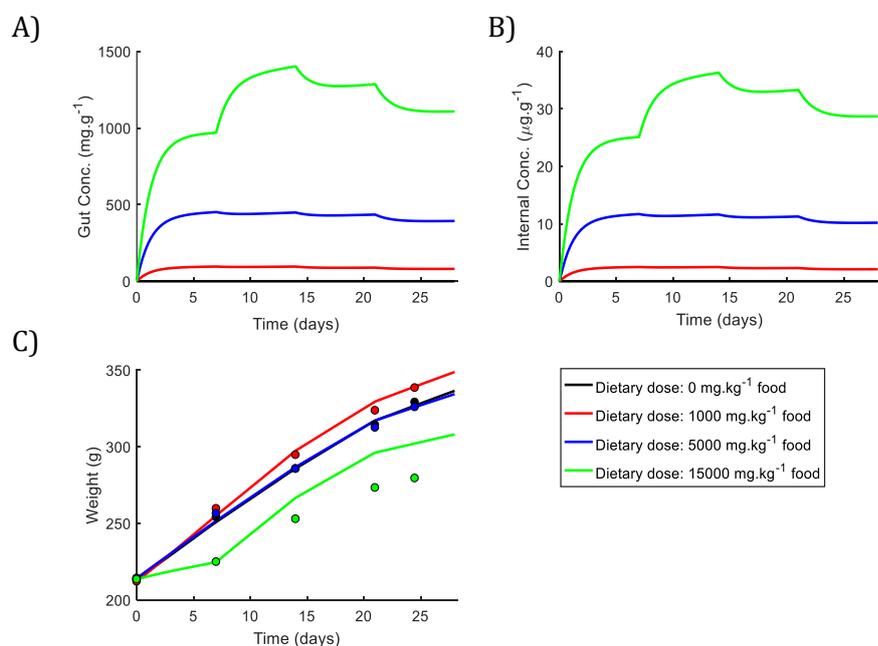
28 Day Study

Figure A28: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for male rats dosed with Prosulfuron

90 Day Study

Figure A29: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
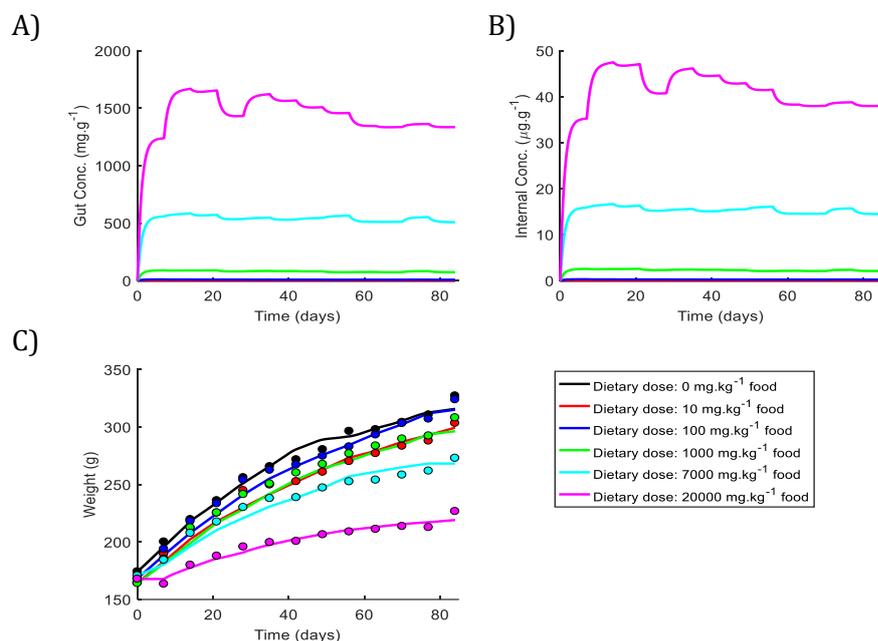
28 Day Study

Figure A30: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
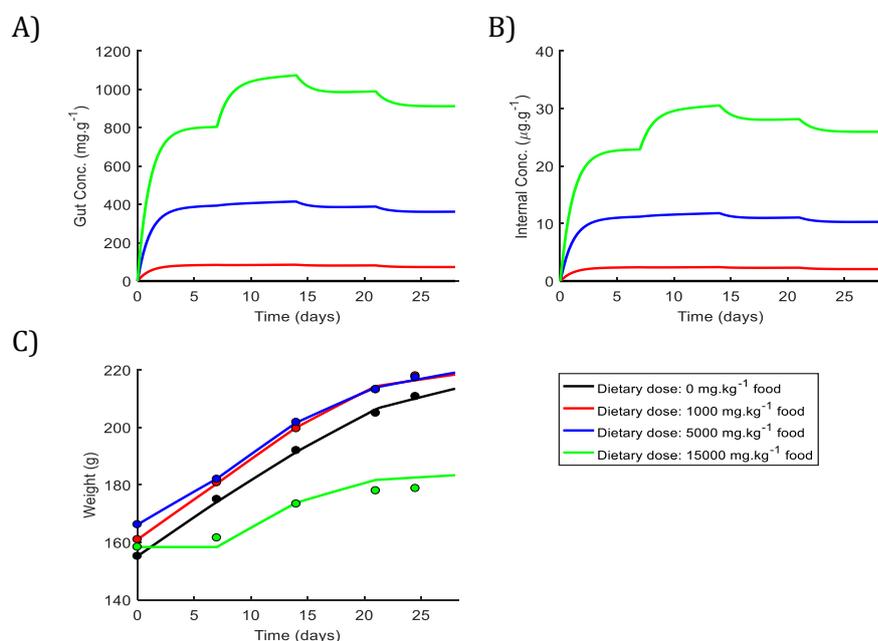
Two Year Study

A)

B)

C)

Figure A31: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for female rats dosed with Prosulfuron

90 Day Study

A)

B)

C)

Figure A32: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
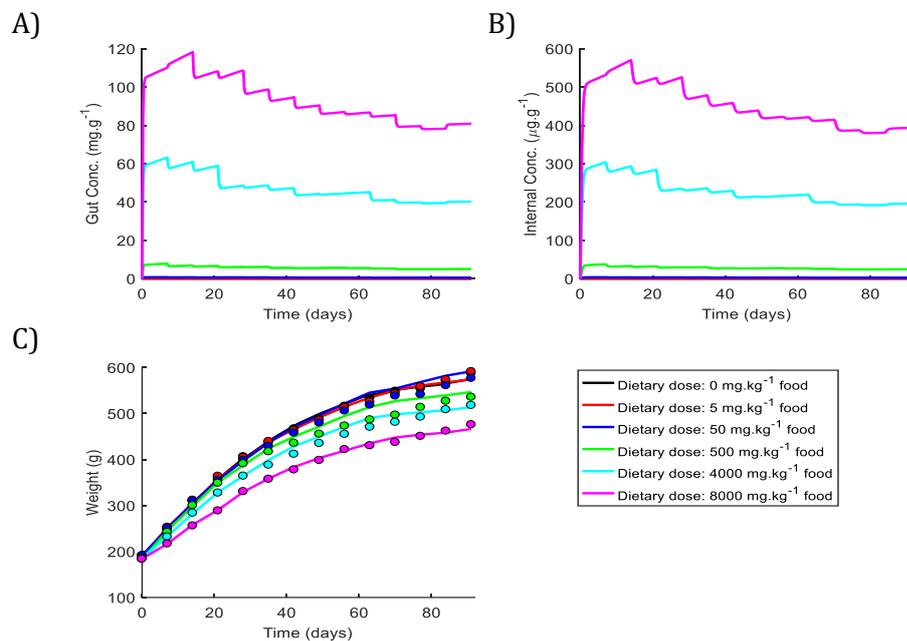
28 Day Study

A)

B)

C)

Figure A33: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
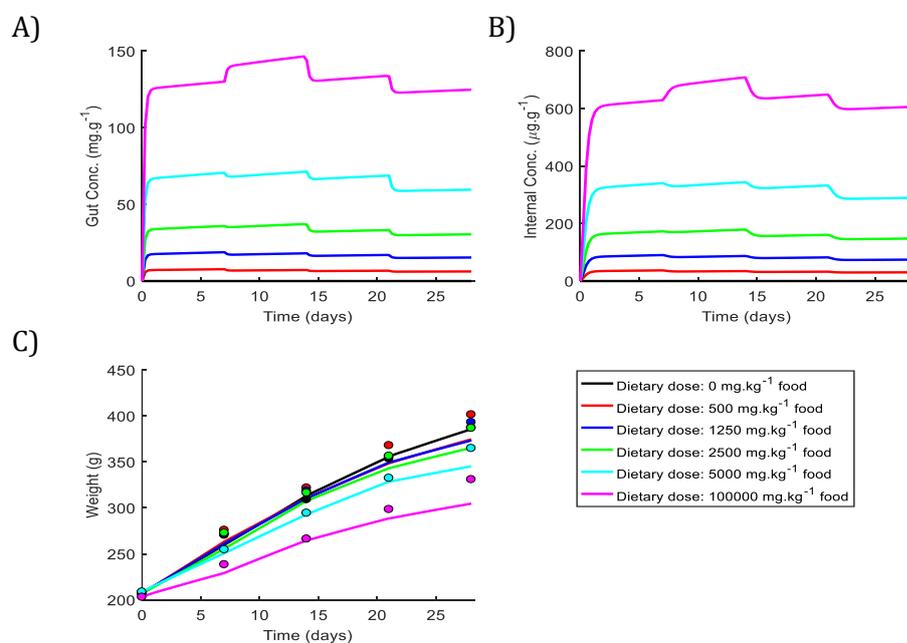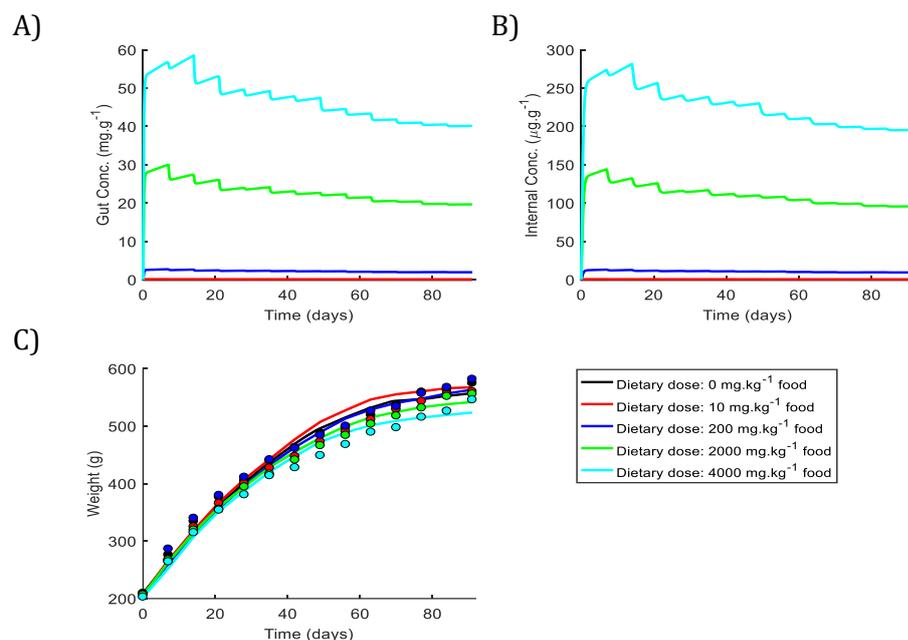
Two Year Study

A)

B)

C)

Figure A34: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for male rats dosed with Thiamethoxam

90 Day Study

A)

B)

C)
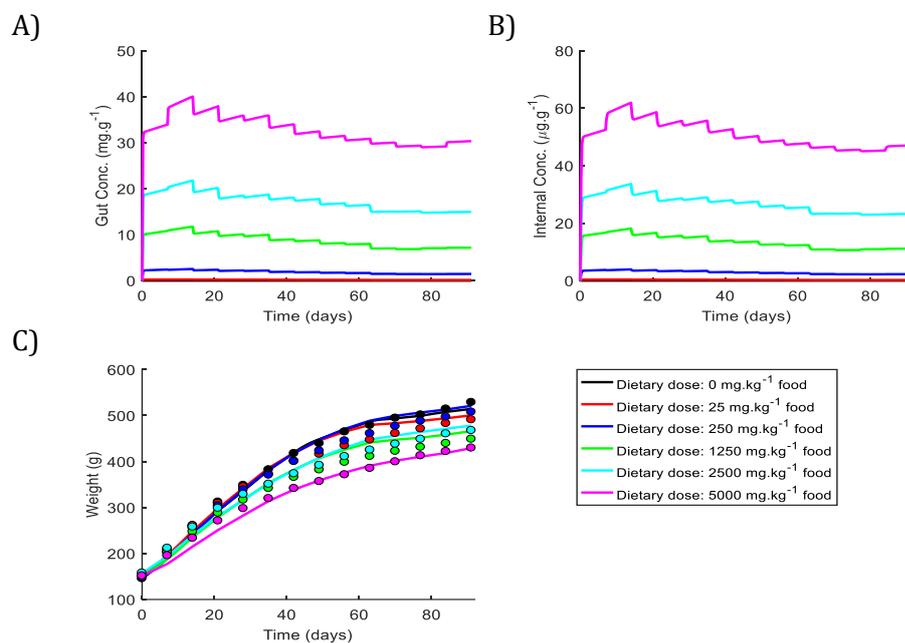
Figure A35: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

28 Day Study

A)

B)

C)
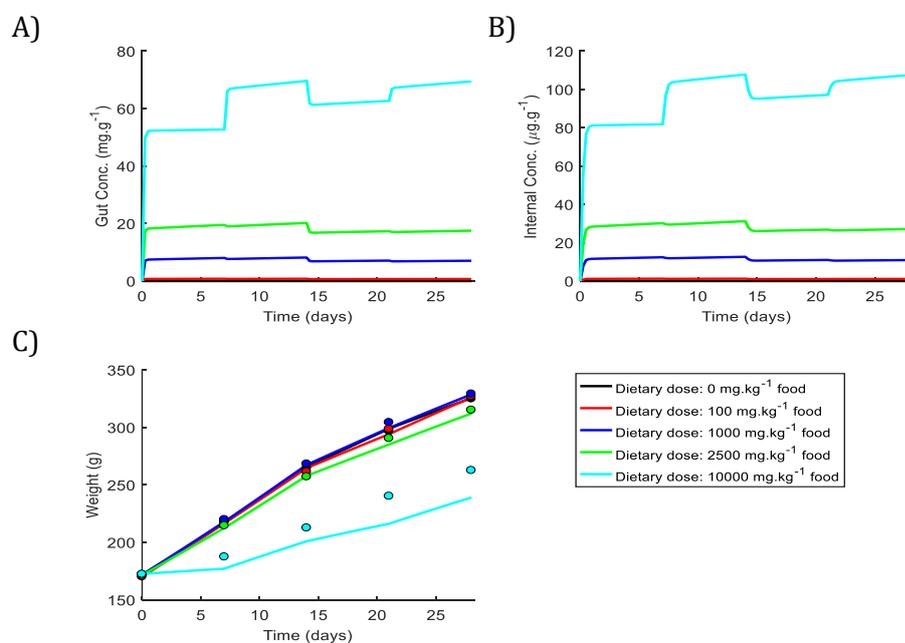
Figure A36: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
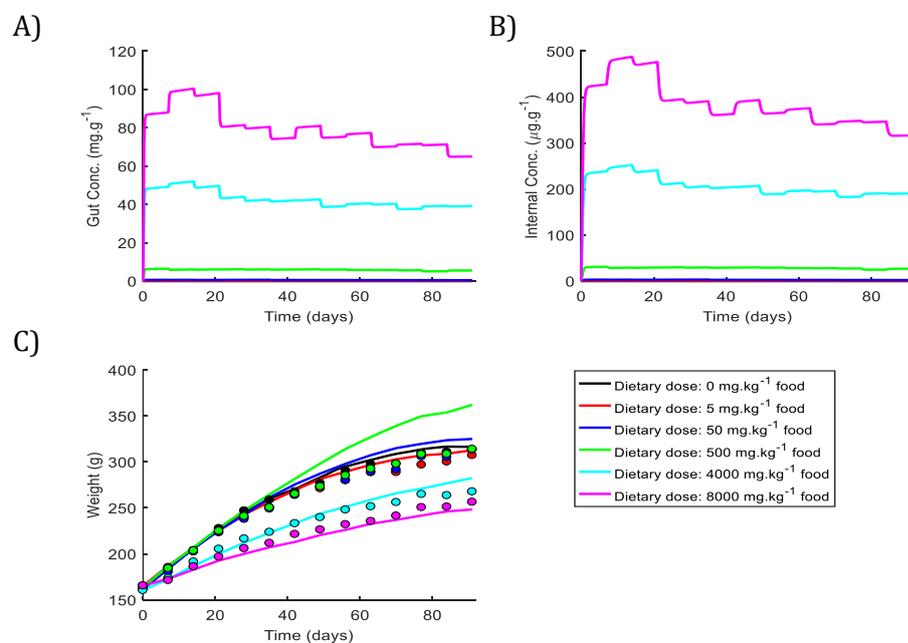
Two Year Study

A)

B)

C)
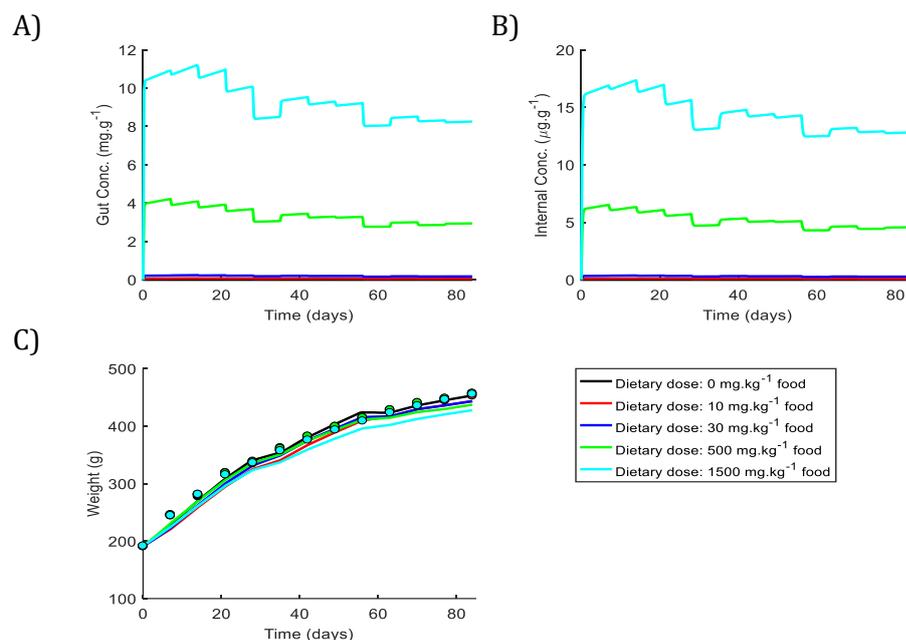
Figure A37: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

TK-TD model outputs for female rats dosed with Thiamethoxam

90 Day Study

A)

B)

C)
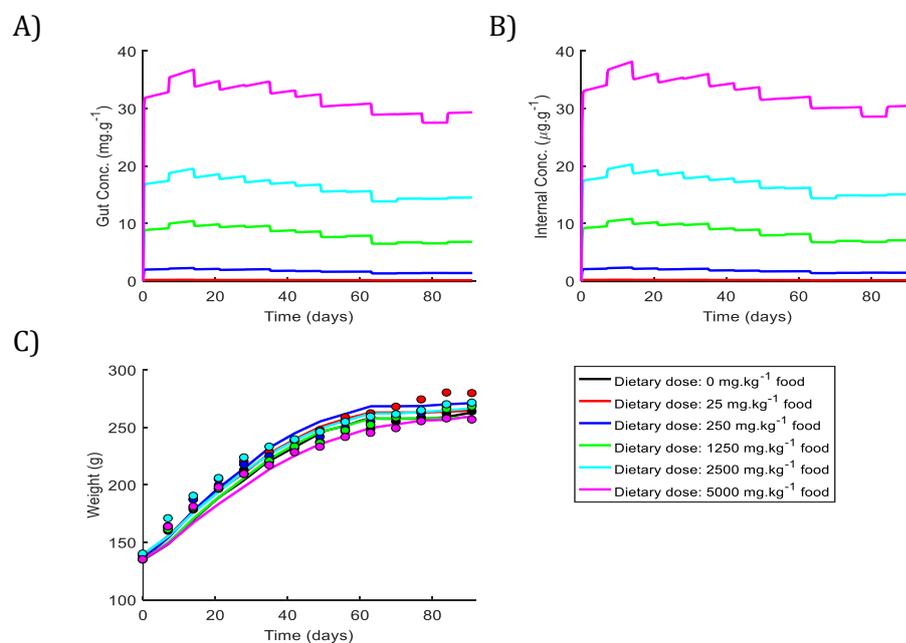
Figure A38: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
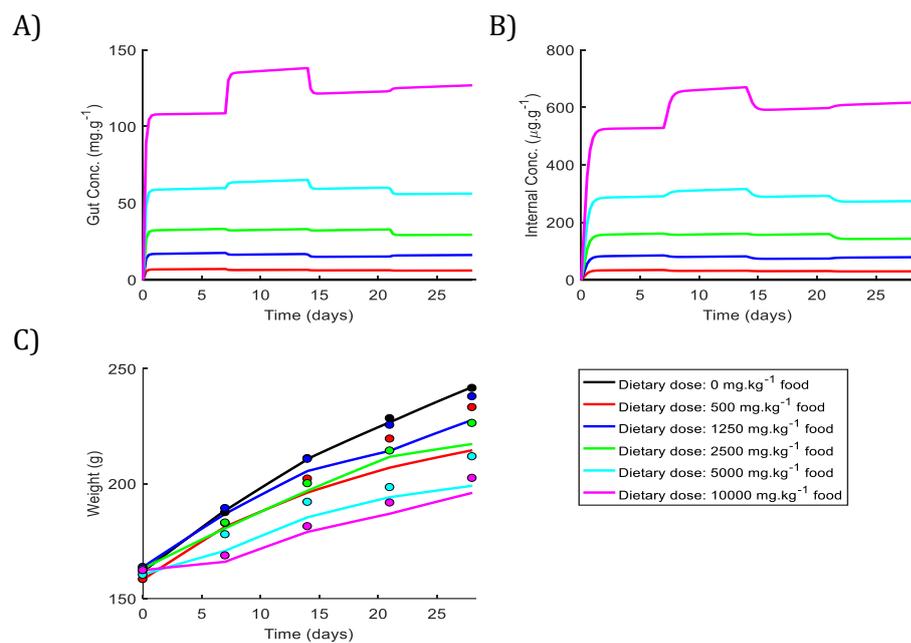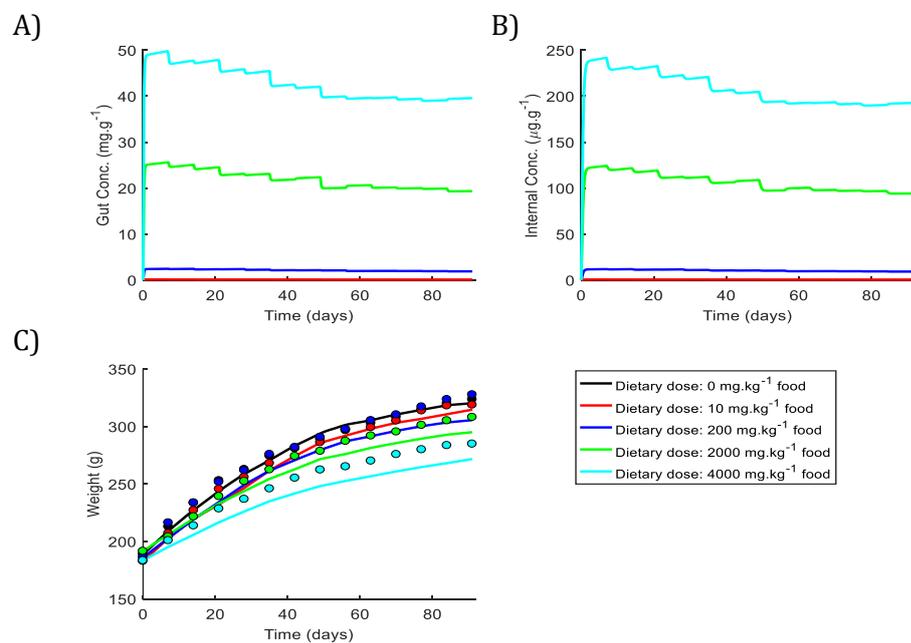
28 Day Study

A)

B)

C)

Figure A39: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).

Two Year Study

A)

B)

C)

Figure A40: Model predictions (lines) of toxicant concentration in the gut (A), internal toxicant concentration (B) and body weight (C) over time. Weekly observations of body weight are also shown (circles).
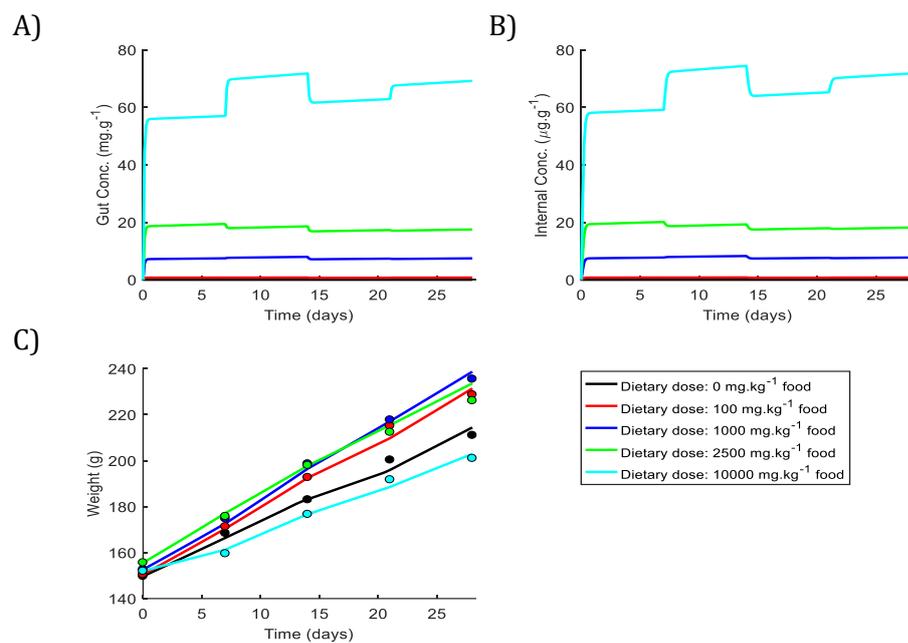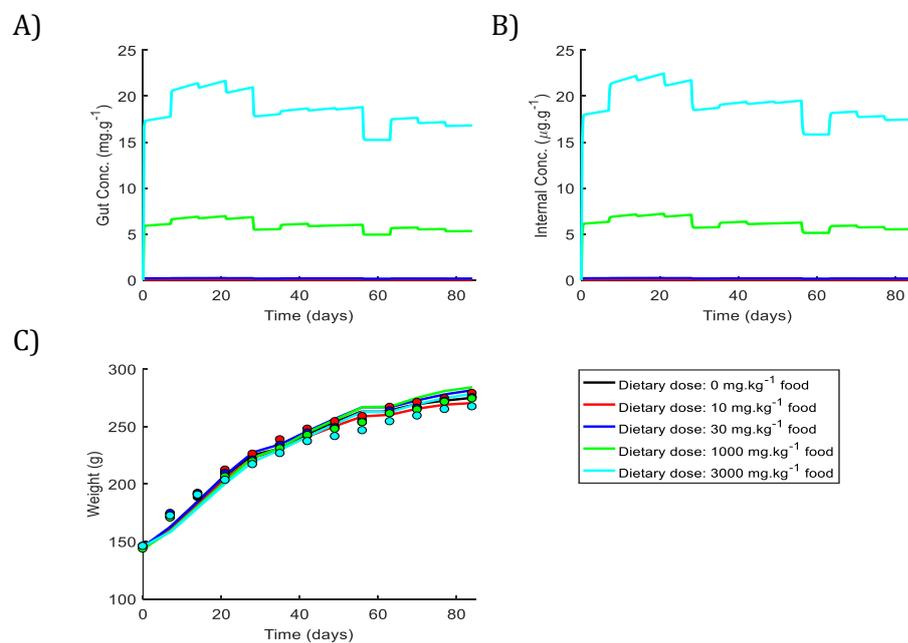
Appendix B - Supporting Information for Chapter 4

Identifying target concentration in muscle

The TK model used in Martin et al. (2019) was calibrated using blood concentration data. Concentration in blood was used as a proxy for overall internal concentration or 'body burden'. Predictions of internal concentration over the duration of repeated dietary dose toxicity tests were then used to calculate mean values over the testing period in different treatments. The highest average internal concentration in any treatment was identified as a 'reference concentration in blood' for each pesticide. However, the tissue chosen for use in cell culture was skeletal muscle. For all the pesticides in the study, concentration in the blood was significantly ($p<0.0001$) correlated with that in muscle ($n \geq 23$). These relationships were determined (Table B1 & figure S1) and used to estimate the concentration in muscle corresponding to reference concentration in the blood for each pesticide. Concentrations in muscle were estimated for male and female rats and the higher of the two was used as the reference for *in vitro* experiments.

Table B1 shows the details of the relationship between pesticide concentration in blood and in muscle. For prosulfuron, data were only available to determine this relationship in males. These relationships were used to estimate the highest average muscle concentration reached in repeated dose dietary toxicity testing.

Pesticide	Sex	Correlation coefficient	Muscle Conc. / Blood Conc.	95% CI	Reference conc. in blood	Reference conc. in muscle
Thiamethoxam	Male	0.994	1.05	1.01 - 1.08	49.36	51.58
	Female	0.999	1.03	1.02 - 1.04	60.80	62.70
Prosulfuron	Male	0.972	0.40	0.36 - 0.44	536.47	213.67
	Female	''	''	''	499.04	198.77
Azoxystrobin	Male	0.936	0.33	0.29 - 0.37	93.94	31.01
	Female	0.936	0.56	0.49 - 0.62	82.07	45.57
Acibenzolar-S-Methyl	Male	0.999	0.13	0.129 -0.134	34.96	4.58
	Female	0.994	0.11	0.106 -0.116	37.66	4.17
Fludioxonil	Males	0.930	0.36	0.31 - 0.41	43.17	15.36
	Female	0.893	0.64	0.52 - 0.77	40.35	26.02
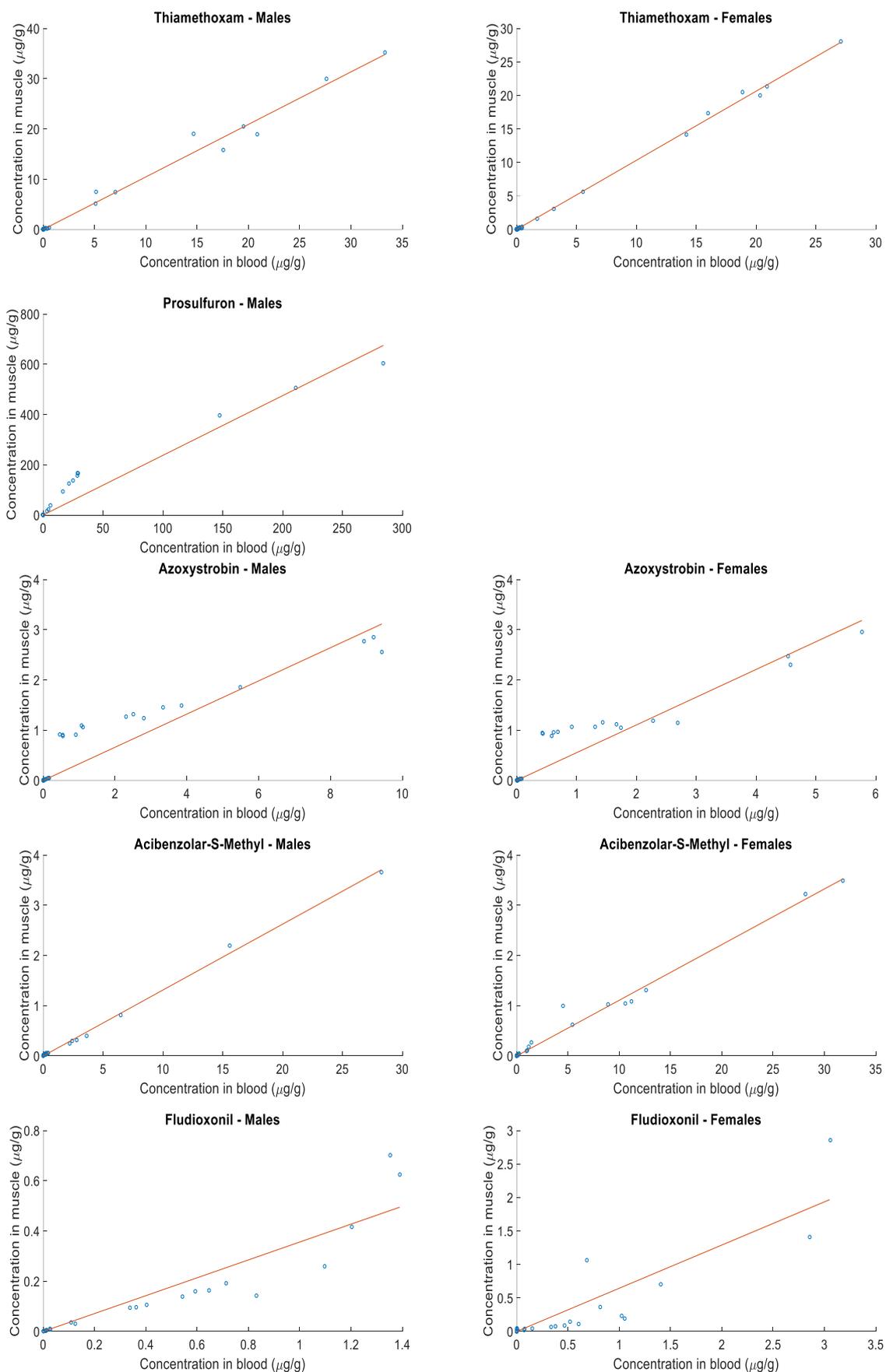
Figure B1 Scatterplots showing observed pesticide concentration in blood and in muscle (blue circles) for five pesticides with line of best fit shown in red.

Effects of FBS concentration on cell proliferation

Cell proliferation was monitored over two days in culture medium contain five different FBS concentrations. This was to identify the FBS concentration sufficient to maintain cell population while inhibiting proliferation. Mitochondrial activity after 48 hours was similar to that recorded at day 0 in medium containing 1-5% w/v FBS. As the intermediated value, 2.5% was selected for use in the TK experiments.
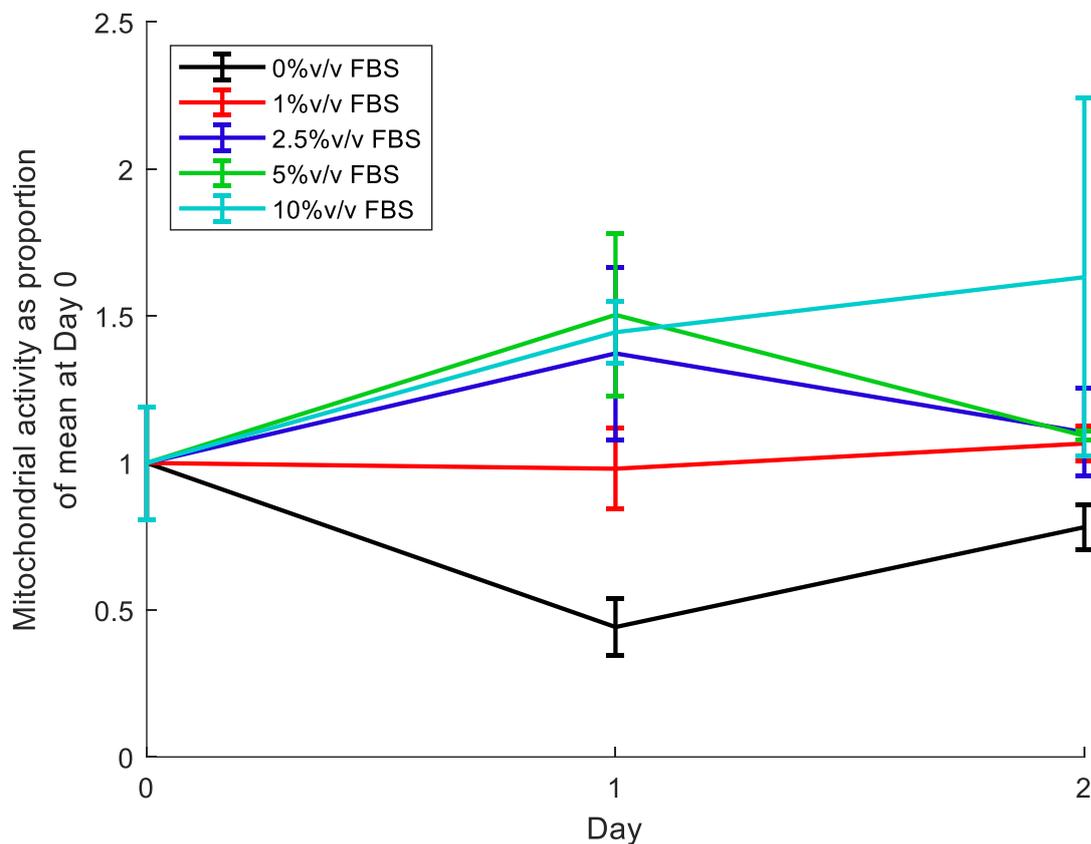
Figure B2 Mean mitochondrial activity over two days in medium containing various concentrations of FBS, plotted as the proportion of mean mitochondrial activity recorded at day 0. Error bars represent 95% confidence intervals.

Calculation of concentration from measured radioactivity

In the *in vitro* toxicokinetics experiments, the pesticide concentration in different compartments was determined by measuring radioactivity though liquid scintillation counting. Radioactivity, measured in DPM, was first converted to $mg_{pesticide} \times well^{-1}$. DPM was divided by 60 to give $Bq \times well^{-1}$ and then by 10^6 to give $MBq \times well^{-1}$. This was then divided by the specific activity of the test compound ($MBq \times mg^{-1}$), giving to $mg_{pesticide} \times well^{-1}$.

These values were then divided by the size of each compartment to convert to concentrations. Each well contained 0.1ml medium, so $mg_{pesticide} \times well^{-1}$ recovered from the medium was divided by $0.0001L_{medium} \times well^{-1}$ to give concentration in $mg \times L^{-1}_{medium}$. The internal surface area of each well was calculated as 2.53cm², so $mg_{pesticide} \times well^{-1}$ recovered from the plastic was divided by $2.53cm^2_{plastic} \times well^{-1}$ and then multiplied by 1000 to give concentration in $\mu g \times cm^{-2}_{plastic}$.

Finally, the volume of cells $\times well^{-1}$ was estimated by first using Image J to measure the diameter of detached cells photographed on a haemocytometer, the average value (n=24) was 15.97μm. Assuming detached cells to be spherical, the average cell volume was calculated to be 2135.25 μm³. This was multiplied by 85,000 (the average number of cells, to the nearest 5000), to give a value of $0.00018cm^3_{cells} \times well^{-1}$. Pesticide recovered from cells ($mg_{pesticide} \times well^{-1}$) was then divided by $0.00018cm^3_{cells} \times well^{-1}$ and then multiplied by 1000 to give concentration in $\mu g \times cm^{-3}_{cells}$.

***In vitro* TK results**

Full results for of the *in vitro* toxicokinetics experiments are given in tables S2-S21.

Thiamethoxam

Table B2 shows the mean and standard deviation of thiamethoxam concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 6.40 mg/L.

Time (h)	Initial Medium Concentration: 6.40 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	6.27	0.10	65.78	10.20	5.07E-04	4.03E-05
24	6.04	0.39	64.71	6.37	4.53E-04	1.34E-04
48	6.34	0.15	82.17	13.54	4.70E-04	2.61E-04

Table B3 shows the mean and standard deviation of thiamethoxam concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 3.10 mg/L.

Time (h)	Initial Medium Concentration: 3.10 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	3.01	0.19	39.47	4.35	2.99E-04	7.01E-05
24	3.06	0.14	43.26	5.79	2.23E-04	1.47E-04
48	3.07	0.10	40.19	19.63	2.23E-04	1.26E-04

Table B4 shows the mean and standard deviation of thiamethoxam concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 9.76 mg/L.

Time (h)	Initial Medium Concentration: 9.76 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	9.34	2.06	150.30	57.82	1.49E-03	9.68E-04
24	9.82	1.08	113.66	22.44	4.75E-04	2.07E-04
48	8.43	2.21	122.67	8.34	7.71E-04	3.17E-04

Table B5 shows the mean and standard deviation of thiamethoxam concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 3.76 mg/L.

Time (h)	Initial Medium Concentration: 3.76 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	3.67	0.27	48.88	10.82	3.11E-04	1.70E-04
24	4.09	0.32	56.87	9.76	4.05E-04	9.93E-05
48	3.42	0.48	50.64	6.15	2.77E-04	1.06E-04

Prosulfuron

Table B6 shows the mean and standard deviation of prosulfuron concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 12.50 mg/L.

	Initial Medium Concentration: 12.50 mg/L					
	Medium (mg/L)		Cells (µg/cm³)		Plastic (µg/cm²)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	12.13	0.47	170.33	57.23	1.41E-03	6.29E-04
24	12.22	0.75	121.58	11.02	9.55E-04	1.26E-04
48	12.22	0.53	140.01	23.16	9.26E-04	1.49E-04

Table B7 shows the mean and standard deviation of prosulfuron concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 6.20 mg/L.

	Initial Medium Concentration: 6.20 mg/L					
	Medium (mg/L)		Cells (µg/cm³)		Plastic (µg/cm²)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	6.03	0.38	68.05	2.66	4.90E-04	1.00E-04
24	5.89	0.10	68.71	21.51	4.75E-04	8.00E-05
48	5.90	0.18	67.20	3.61	4.63E-04	2.01E-04

Table B8 shows the mean and standard deviation of prosulfuron concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 14.95 mg/L.

	Initial Medium Concentration: 14.95 mg/L					
	Medium (mg/L)		Cells (µg/cm³)		Plastic (µg/cm²)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	14.46	0.40	245.91	130.65	2.12E-03	2.15E-04
24	14.38	1.13	176.17	17.80	2.06E-03	1.75E-03
48	16.98	0.49	204.72	115.00	1.13E-03	1.22E-03

Table B9 shows the mean and standard deviation of prosulfuron concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 6.24 mg/L.

	Initial Medium Concentration: 6.42 mg/L					
	Medium (mg/L)		Cells (µg/cm³)		Plastic (µg/cm²)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	6.25	0.71	83.56	11.90	7.08E-04	1.59E-04
24	6.89	0.67	97.33	20.28	6.64E-04	1.86E-04
48	7.23	0.79	93.25	5.25	3.07E-04	9.46E-05

Azoxystrobin

Table B10 shows the mean and standard deviation of azoxystrobin concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.24 mg/L.

	Initial Medium Concentration: 0.24 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.22	0.01	4.78	0.91	2.78E-04	2.01E-04
24	0.22	0.01	4.83	0.77	1.52E-04	1.70E-05
48	0.22	0.01	5.72	0.70	1.60E-04	1.12E-05

Table B11 shows the mean and standard deviation of azoxystrobin concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.12 mg/L.

	Initial Medium Concentration: 0.12 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.11	0.004	2.20	0.20	6.18E-05	1.03E-05
24	0.10	0.004	2.71	0.06	6.78E-05	8.25E-06
48	0.11	0.004	3.24	0.17	6.84E-05	1.16E-05

Table B12 shows the mean and standard deviation of azoxystrobin concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.243 mg/L.

	Initial Medium Concentration: 0.243 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.23	0.05	5.16	0.96	6.47E-05	1.78E-05
24	0.21	0.03	5.38	1.33	1.44E-04	5.26E-05
48	0.18	0.01	5.98	0.73	1.76E-04	1.02E-05

Table B13 shows the mean and standard deviation of azoxystrobin concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.096 mg/L.

	Initial Medium Concentration: 0.096 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.09	0.01	2.21	0.48	2.86E-05	1.50E-05
24	0.09	0.01	2.51	0.27	5.21E-05	1.60E-05
48	0.08	0.01	3.02	0.21	5.73E-05	7.68E-06

Acibenzolar-S-Methyl

Table B14 shows the mean and standard deviation of acibenzolar-S-methyl concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.08 mg/L.

	Initial Medium Concentration: 0.08 mg/L					
	Medium (mg/L)		Cells ($\mu g/cm^3$)		Plastic ($\mu g/cm^2$)	
Time (h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.07	0.004	1.74	0.24	1.40E-04	1.40E-05
24	0.06	0.006	0.72	0.09	1.11E-04	4.84E-05
48	0.07	0.0003	0.85	0.04	3.57E-05	9.85E-06

Table B15 shows the mean and standard deviation of acibenzolar-S-methyl concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.045 mg/L.

	Initial Medium Concentration: 0.045 mg/L					
	Medium (mg/L)		Cells ($\mu g/cm^3$)		Plastic ($\mu g/cm^2$)	
Time (h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.03	0.00	0.70	0.13	2.18E-04	2.21E-04
24	0.03	0.00	0.38	0.07	4.81E-05	3.37E-05
48	0.03	0.00	0.46	0.12	1.09E-05	9.58E-06

Table B16 shows the mean and standard deviation of acibenzolar-S-methyl concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.144 mg/L.

	Initial Medium Concentration: 0.144 mg/L					
	Medium (mg/L)		Cells ($\mu g/cm^3$)		Plastic ($\mu g/cm^2$)	
Time (h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.13	0.01	3.85	0.81	3.38E-04	2.56E-05
24	0.09	0.02	1.70	0.22	1.65E-04	4.27E-05
48	0.08	0.02	1.49	0.33	9.57E-05	1.18E-05

Table B17 shows the mean and standard deviation of acibenzolar-S-methyl concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.073 mg/L.

	Initial Medium Concentration: 0.073 mg/L					
	Medium (mg/L)		Cells ($\mu g/cm^3$)		Plastic ($\mu g/cm^2$)	
Time (h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.07	0.01	1.65	0.27	1.83E-04	2.44E-05
24	0.04	0.01	1.04	0.31	6.12E-05	1.95E-05
48	0.05	0.01	1.08	0.18	5.10E-05	2.36E-05

Fludioxonil

Table B18 shows the mean and standard deviation of fludioxonil concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.091 mg/L.

	Initial Medium Concentration: 0.091 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.07	0.003	8.49	1.11	1.30E-04	1.34E-05
24	0.07	0.003	10.60	1.01	1.63E-04	1.73E-05
48	0.06	0.003	11.81	0.35	1.37E-04	8.67E-06

Table B19 shows the mean and standard deviation of fludioxonil concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.047 mg/L.

	Initial Medium Concentration: 0.047 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.03	0.002	3.86	0.27	5.94E-05	8.72E-06
24	0.03	0.001	5.18	0.43	8.38E-05	9.70E-06
48	0.03	0.001	5.76	0.49	6.84E-05	2.28E-05

Table B20 shows the mean and standard deviation of fludioxonil concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.55 mg/L.

	Initial Medium Concentration: 0.55 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.45	0.02	45.45	4.61	6.83E-04	1.30E-04
24	0.35	0.03	59.40	1.80	6.62E-04	1.28E-04
48	0.34	0.02	69.33	5.83	6.97E-04	9.27E-05

Table B21 shows the mean and standard deviation of fludioxonil concentration detected in culture, medium and on plastic at 3, 24 and 48 hours after dosing at 0.21 mg/L.

	Initial Medium Concentration: 0.21 mg/L					
	Medium (mg/L)		Cells (μg/cm^3)		Plastic (μg/cm^2)	
Time(h)	Mean	s.d.	Mean	s.d.	Mean	s.d.
3	0.17	0.01	22.38	2.41	3.31E-04	8.96E-05
24	0.15	0.01	29.36	4.22	3.20E-04	8.81E-05
48	0.15	0.02	34.03	2.01	4.12E-04	6.96E-05

Cell Proliferation

Results of MTS assay measuring cell proliferation in the presence of each pesticide are given in tables S22-S26. Results were analysed by ANOVA or Kruskall Wallis if assumptions of normality and homoscedasticity were not met and pairwise comparisons were made using Tukey's honest significant difference criterion.

Table B22 shows mean and standard deviation of mitochondrial activity relative to the vehicle control treatment measured 24, 48 and 72 hours after dosing with thiamethoxam. Any mean values marked with an asterisk are significantly different the vehicle control at that timepoint.

Thiamethoxam	Mitochondrial activity relative to the vehicle control treatment					
	24h		48h		72h	
Treatment (mg/L$_{medium}$)	Mean	SD	Mean	SD	Mean	SD
Untreated	1.04	0.23	1.13	0.23	1.05	0.21
Vehicle Control	1.00	0.22	1.00	0.21	1.00	0.12
2.72	1.08	0.27	0.89	0.18	1.02	0.25
5.44	0.77	0.25	1.08	0.22	1.01	0.17
13.6	0.87	0.15	1.00	0.15	0.99	0.21
27.2	0.85	0.36	1.21	0.38	0.90	0.14
54.40	0.92	0.09	1.29	0.30	0.86	0.11

Table B23 shows mean and standard deviation of mitochondrial activity relative to the vehicle control treatment measured 24, 48 and 72 hours after dosing with prosulfuron. Any mean values marked with an asterisk are significantly different the vehicle control at that timepoint.

Prosulfuron	Mitochondrial activity relative to the vehicle control treatment					
	24h		48h		72h	
Treatment (mg/L$_{medium}$)	Mean	SD	Mean	SD	Mean	SD
Untreated	1.04	0.23	1.13	0.23	1.05	0.21
Vehicle Control	1.00	0.22	1.00	0.21	1.00	0.12
9.3	1.19	0.51	1.75*	0.19	1.12	0.36
18.6	1.09	0.32	1.41	0.53	1.04	0.25
46.5	0.94	0.50	1.58	0.11	0.89	0.09
93	1.25	0.23	1.52	0.52	0.80	0.25
186	0.89	0.20	1.65	0.41	0.97	0.13

Table B24 shows mean and standard deviation of mitochondrial activity relative to the vehicle control treatment measured 24, 48 and 72 hours after dosing with azoxystrobin. Any mean values marked with an asterisk are significantly different the vehicle control at that timepoint.

Azoxystrobin	Mitochondrial activity relative to the vehicle control treatment					
	24h		48h		72h	
Treatment (mg/L$_{medium}$)	Mean	SD	Mean	SD	Mean	SD
Untreated	1.04	0.23	1.13	0.23	1.05	0.21
Vehicle Control	1.00	0.22	1.00	0.21	1.00	0.12
0.995	0.55*	0.17	1.58*	0.11	1.00	0.14
1.99	0.64	0.14	1.65*	0.21	1.06	0.33
4.98	0.78	0.28	1.61*	0.18	1.00	0.24
9.95	0.73	0.16	1.33	0.52	1.14	0.20
19.9	0.77	0.23	1.21	0.25	0.85	0.07

Table B25 shows mean and standard deviation of mitochondrial activity relative to the vehicle control treatment measured 24, 48 and 72 hours after dosing with acibenzolar-S-Methyl. Any mean values marked with an asterisk are significantly different the vehicle control at that timepoint.

Acibenzolar-S-Methyl	Mitochondrial activity relative to the vehicle control treatment					
	24h		48h		72h	
Treatment (mg/L$_{medium}$)	Mean	SD	Mean	SD	Mean	SD
Untreated	1.04	0.23	1.13	0.23	1.05	0.21
Vehicle Control	1.00	0.22	1.00	0.21	1.00	0.12
0.185	0.82	0.06	1.37	0.34	0.80	0.25
0.37	0.70	0.24	1.50	0.22	0.89	0.36
0.925	0.79	0.32	1.45	0.32	0.86	0.18
1.85	0.67	0.09	1.21	0.38	0.77	0.10
3.7	0.50*	0.16	1.46	0.20	0.94	0.20

Table B26 shows mean and standard deviation of mitochondrial activity relative to the vehicle control treatment measured 24, 48 and 72 hours after dosing with fludioxonil. Any mean values marked with an asterisk are significantly different the vehicle control at that timepoint.

Fludioxonil	Mitochondrial activity relative to the vehicle control treatment					
	24h		48h		72h	
Treatment (mg/L$_{medium}$)	Mean	SD	Mean	SD	Mean	SD
Untreated	1.04	0.23	1.13	0.23	1.05	0.21
Vehicle Control	1.00	0.22	1.00	0.21	1.00	0.12
0.11	1.18	0.23	1.56	0.37	1.04	0.14
0.22	1.31	0.26	1.18	0.15	0.95	0.19
0.55	0.96	0.27	1.83*	0.38	0.87	0.11
1.11	1.09	0.17	1.70*	0.51	0.96	0.21
2.22	0.96	0.32	1.63	0.26	1.06	0.28

Cell mortality

Percentage mortality was estimated alongside cell proliferation by quantifying the amount lactate dehydrogenase (LDH) released by dead cells at 24, 48 and 72 hours after dosing. Mortality data were analysed by generating bootstrap confidence intervals. Significant effects were determined when confidence intervals did not overlap with those of either control treatment.

Table B27 shows estimated mortality rates 24 hours after dosing. These results were not analysed as unusually high background readings led to the estimation of theoretically impossible large negative mortality rates.

| Compound	Treatment (mg/L$_{medium}$)	Mean Mortality (%)	Lower CI	Upper CI
None	Untreated	-18.89	-43.21	-0.22
	Vehicle Control	-23.65	-55.71	-4.86
Thiamethoxam	2.72	-6.45	-28.60	9.18
	5.44	-8.26	-30.51	11.79
	13.6	-15.98	-38.38	-0.06
	27.2	-22.68	-56.10	-3.21
	54.40	-24.35	-57.50	-3.82
Prosulfuron	9.3	-6.16	-11.30	-1.32
	18.6	-8.11	-14.87	-2.26
	46.5	6.27	-11.08	23.52
	93	-12.41	-23.59	-3.55
	186	-16.00	-27.77	-6.78
Azoxystrobin	0.995	8.03	2.53	13.44
	1.99	11.38	4.14	18.22
	4.98	9.27	1.72	17.00
	9.95	17.01	1.59	29.17
	19.9	3.49	-3.54	11.57
Acibenzolar-S-Methyl	0.185	15.10	-12.50	50.75
	0.37	9.54	3.27	27.74
	0.925	20.07	1.03	34.36
	1.85	16.23	6.03	14.32
	3.7	24.76	-6.20	27.86
Fludioxonil	0.11	-6.43	-17.21	5.36
	0.22	-21.34	-34.50	-10.02
	0.55	-12.80	-34.37	9.80
	1.11	-10.02	-19.72	-0.29
	2.22	-9.43	-21.81	4.28

Table B28 shows estimated mortality rates 48 hours after dosing. Any mean values marked with an asterisk are significantly different from the vehicle control treatment.

Compound	Treatment	Mean Mortality (%)	Lower CI	Upper CI
None	Untreated	1.14	-2.14	4.01
	Vehicle Control	-1.09	-4.01	2.12
Thiamethoxam	2.72	-2.10	-3.64	-0.23
	5.44	-2.17	-6.05	0.61
	13.6	-3.63	-6.08	-1.45
	27.2	-4.46	-6.24	-2.78
	54.40	-5.27	-7.31	-2.78
Prosulfuron	9.3	1.56	0.44	3.15
	18.6	1.77	0.72	3.07
	46.5	1.86	0.69	3.48
	93	1.33	-0.36	3.04
	186	3.18	1.16	5.47
Azoxystrobin	0.995	5.68	3.47	7.34
	1.99	4.97	3.23	6.98
	4.98	8.57 *	4.37	13.00
	9.95	6.66	3.94	9.14
	19.9	10.55 *	4.90	15.06
Acibenzolar-S-Methyl	0.185	4.18	-1.22	4.40
	0.37	-0.30	-0.32	3.23
	0.925	0.64	-1.19	2.06
	1.85	1.63	-1.25	0.58
	3.7	2.25	0.56	7.99
Fludioxonil	0.11	0.63	-0.42	1.68
	0.22	0.12	-0.76	1.26
	0.55	0.11	-0.74	1.42
	1.11	1.02	-0.65	3.04
	2.22	2.47	-0.14	4.64

Table B29 shows estimated mortality rates 72 hours after dosing. Any mean values marked with an asterisk are significantly different from the vehicle control treatment.

Compound	Treatment	Mean Mortality (%)	Lower CI	Upper CI
None	Untreated	0.13	-2.75	2.85
	Vehicle Control	1.84	-1.11	5.72
Thiamethoxam	2.72	-0.65	-4.52	2.86
	5.44	-2.49	-6.18	1.20
	13.6	-0.60	-5.43	3.83
	27.2	0.72	-3.67	5.49
	54.40	1.29	-3.85	6.39
Prosulfuron	9.3	4.78	2.11	7.54
	18.6	6.12	4.20	8.22
	46.5	5.47	2.84	8.58
	93	8.22	5.60	11.01
	186	13.49 *	9.38	17.27
Azoxystrobin	0.995	1.91	0.72	3.21
	1.99	2.06	-2.83	7.02
	4.98	2.36	-0.70	6.34
	9.95	1.69	-2.24	5.31
	19.9	2.07	-0.72	4.56
Acibenzolar-S-Methyl	0.185	-1.17	-2.38	0.52
	0.37	-2.03	-3.60	3.69
	0.925	-1.35	-2.72	0.07
	1.85	0.93	-3.43	-0.37
	3.7	-1.09	-2.85	0.59
Fludioxonil	0.11	9.16	5.39	13.44
	0.22	8.30	2.81	12.23
	0.55	9.90 *	7.19	12.65
	1.11	8.41 *	6.11	10.94
	2.22	7.96	4.55	13.50

Appendix C – Supporting Figures for Chapter 5

Modelled Wet Weight – Calibration Data
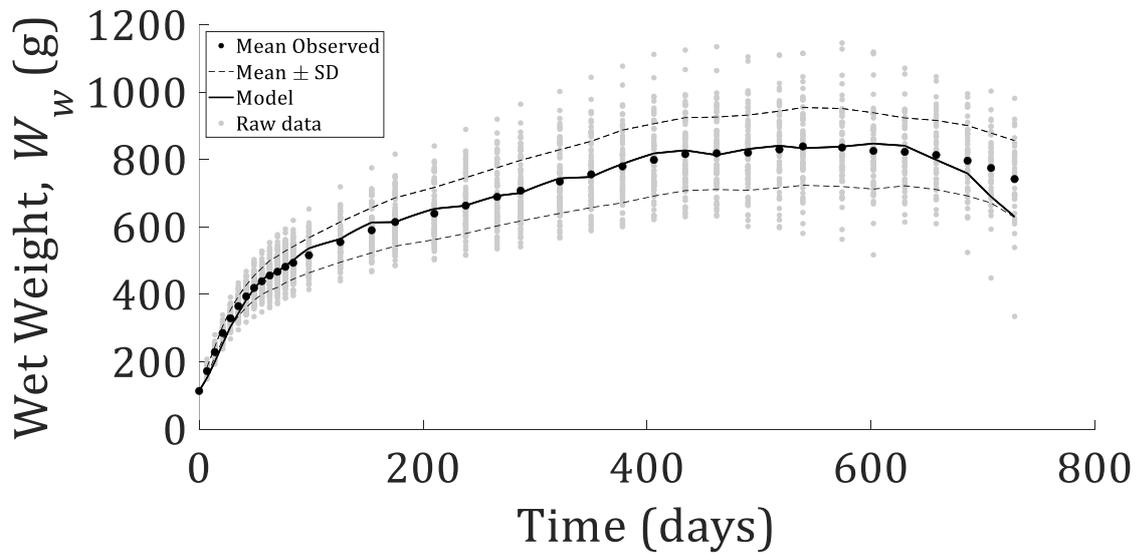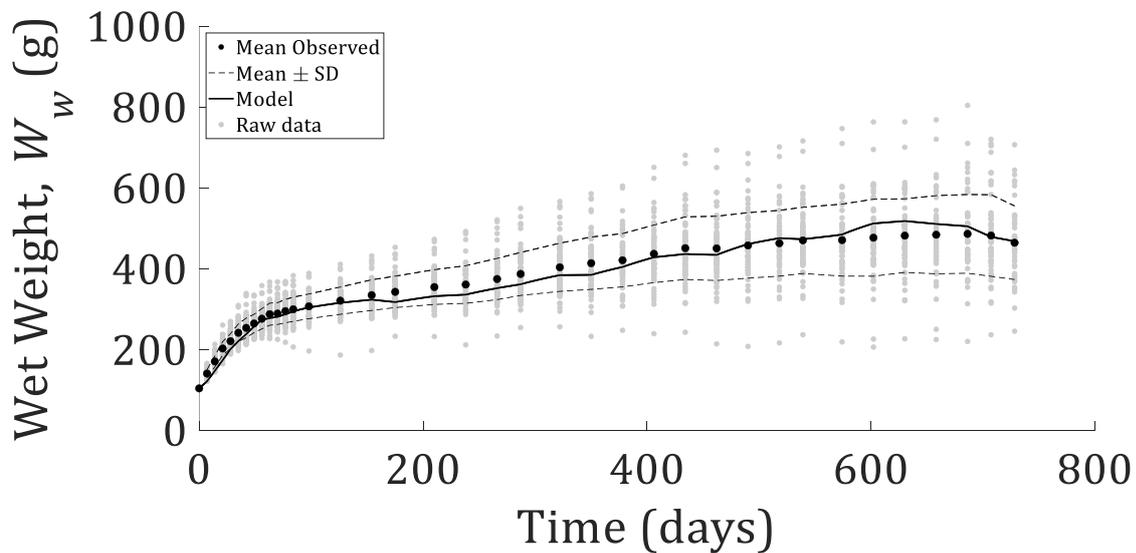
Method 1 – Dataset A
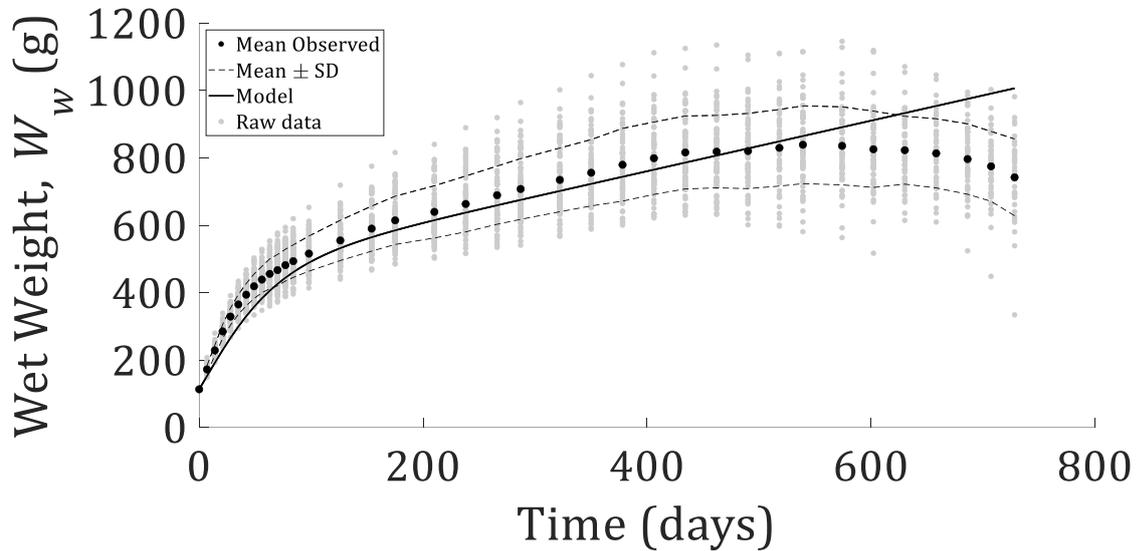
Figure C1 Wet weight of male rats in group A over two years, modelled using Method 1 (solid line), fitted to mean observations (black circles). Grey circles represent raw data while dashed lines show the observed mean ±1 SD.
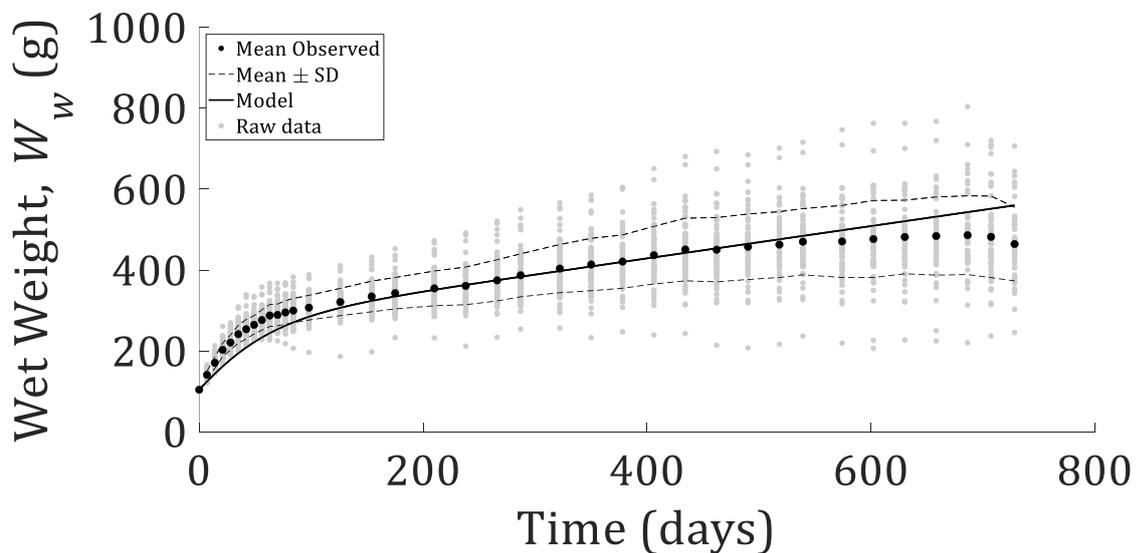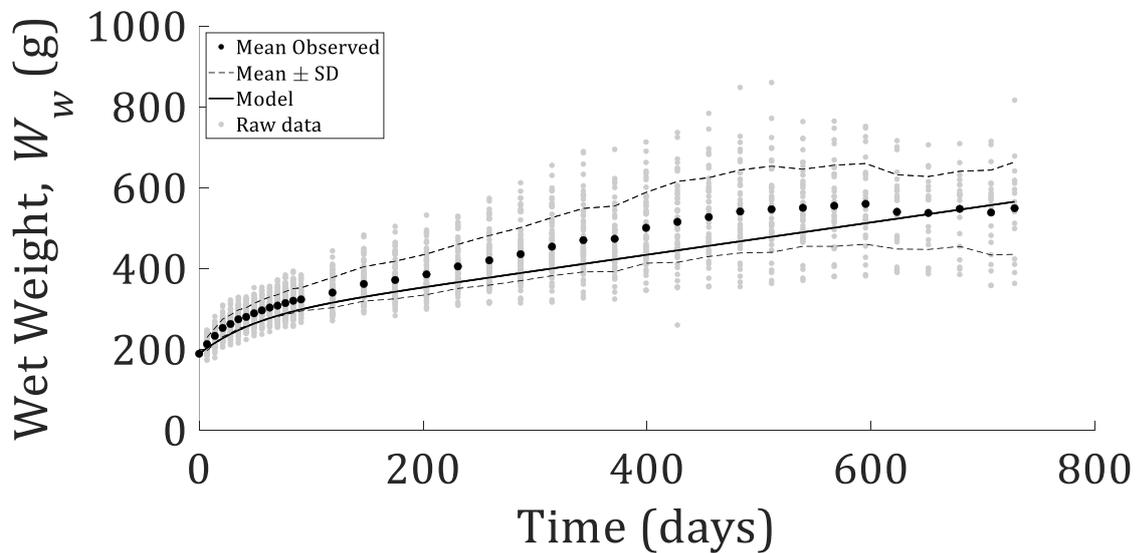

Figure C2 Wet weight of female rats in group A over two years, modelled using Method 1 (solid line), fitted to mean observations (black circles). Grey circles represent raw data while dashed lines show the observed mean ±1 SD.

Method 2 - Dataset A
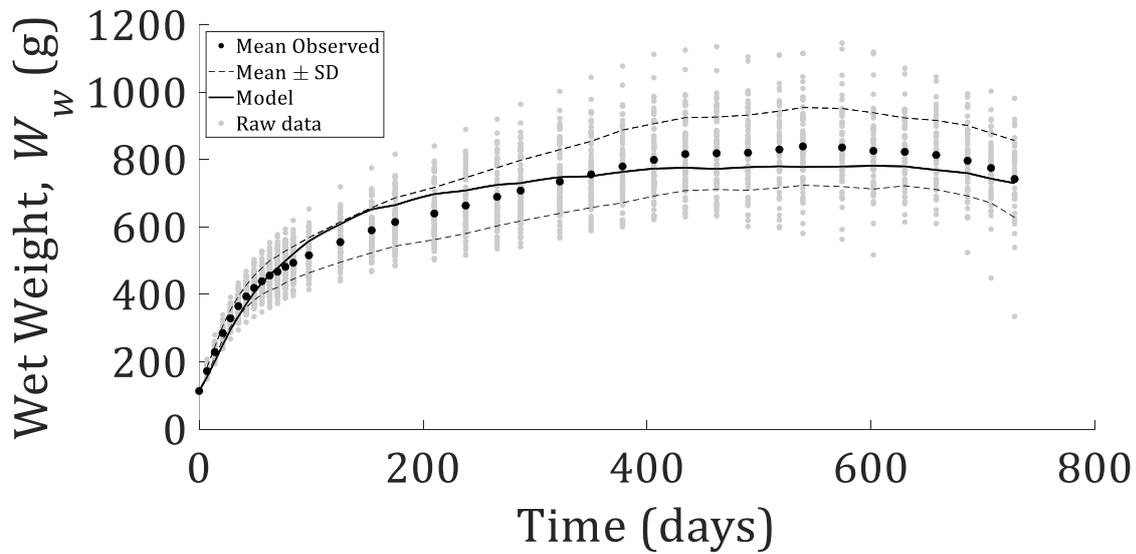
Figure C3 Wet weight of male rats in group A over two years, modelled using Method 2 (solid line), fitted to mean observations (black circles). Grey circles represent raw data while dashed lines show the observed mean ±1 SD.
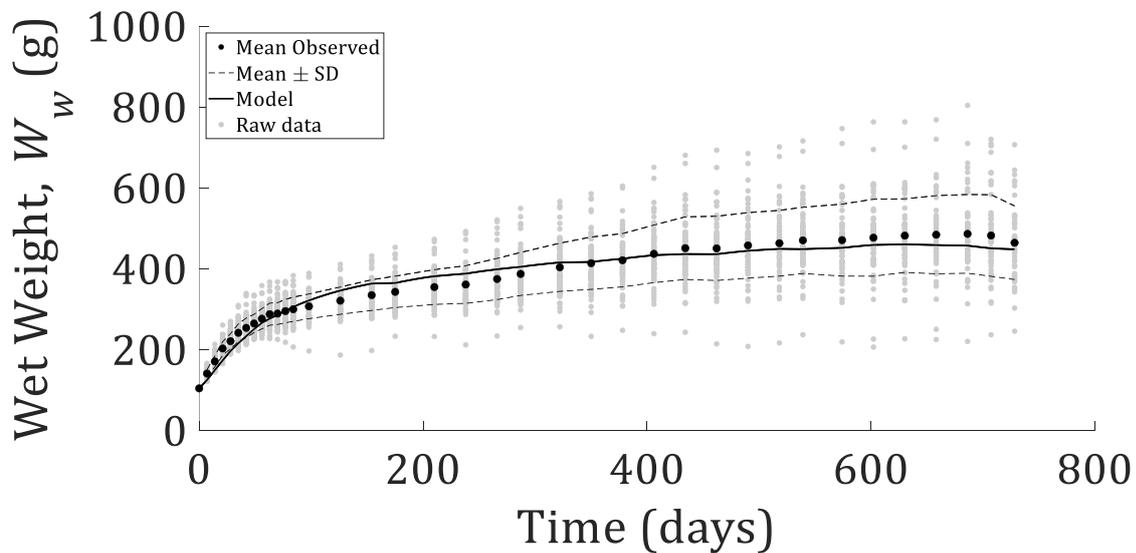
Figure C4 Wet weight of female rats in group A over two years, modelled using Method 2 (solid line), fitted to mean observations (black circles). Grey circles represent raw data while dashed lines show the observed mean ±1 SD.

Method 3 – Dataset A

Figure C5 Wet weight of male rats in group A over two years, modelled using Method 3 (solid line), fitted to mean observations (black circles). Grey circles represent raw data while dashed lines show the observed mean ±1 SD.

Figure C6 Wet weight of female rats in group A over two years, modelled using Method 3 (solid line), fitted to mean observations (black circles). Grey circles represent raw data while dashed lines show the observed mean ±1 SD.

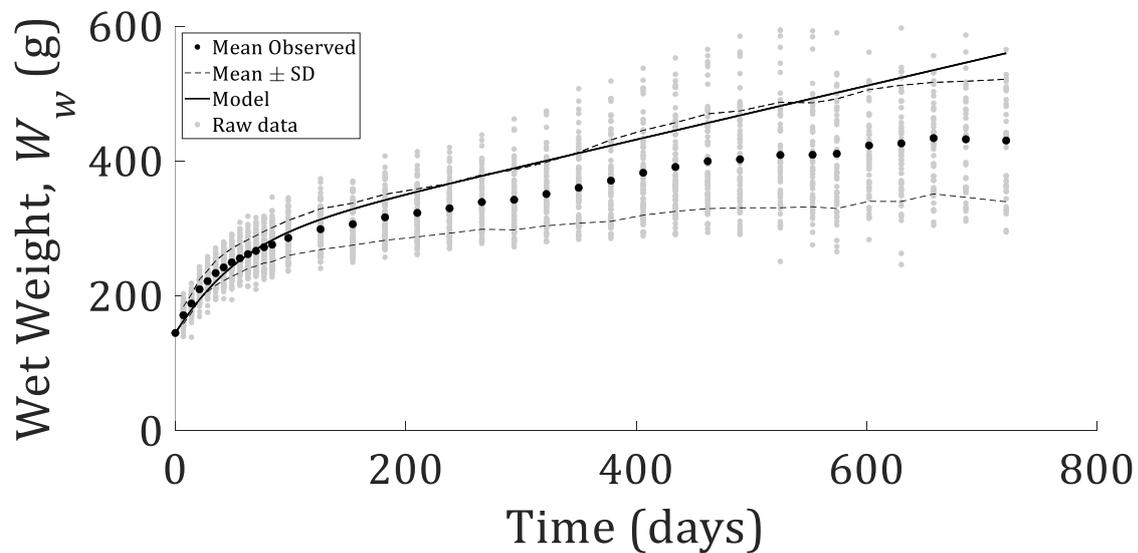**Predicted Wet Weight – Independent Data**
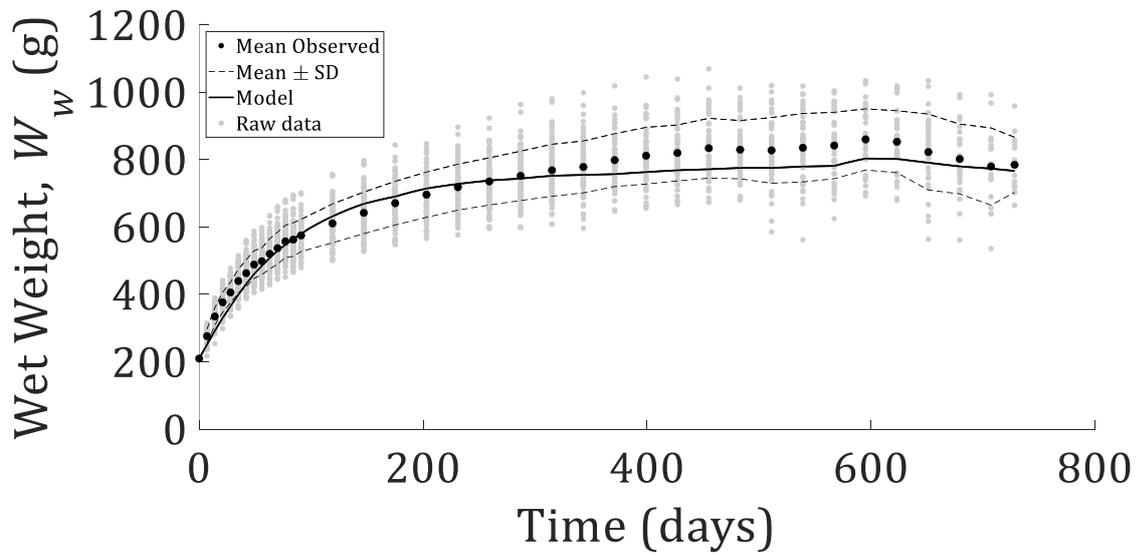
Method 1 – Dataset B
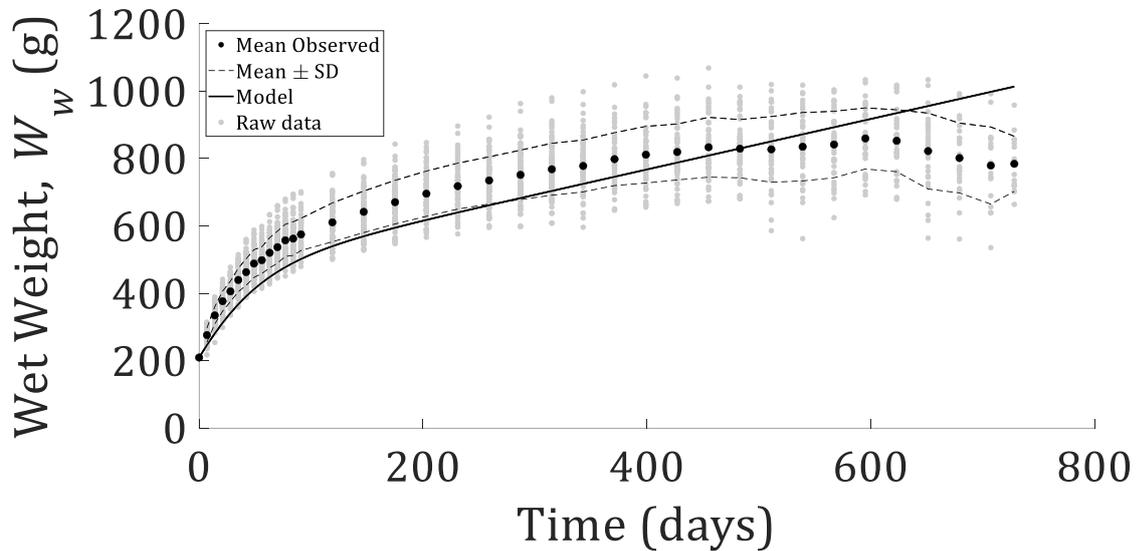
Figure C7 Observed (circles) and predicted (solid line) wet weight of male rats in group B over two years, using Method 1. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

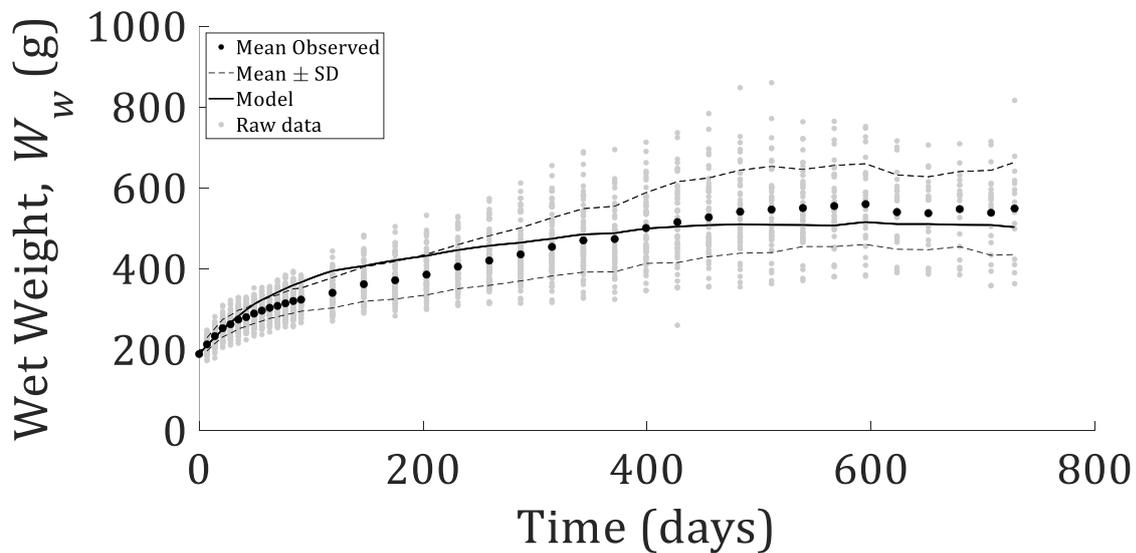
Figure C8 Observed (circles) and predicted (solid line) wet weight of female rats in group B over two years, using Method 1. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

Method 1 – Dataset C
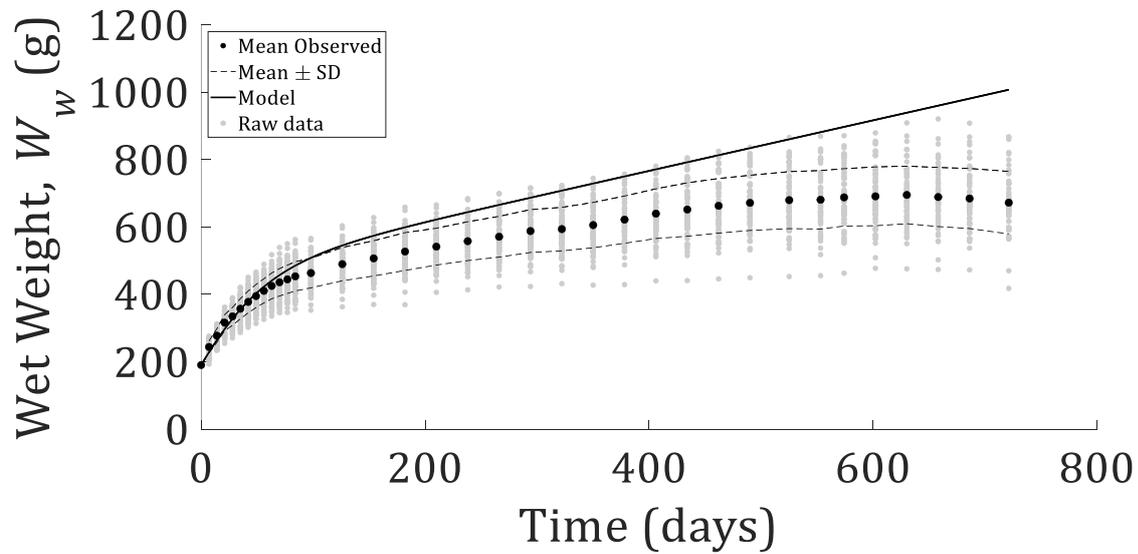
Figure C9 Observed (circles) and predicted (solid line) wet weight of male rats in group C over two years, using Method 1. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.
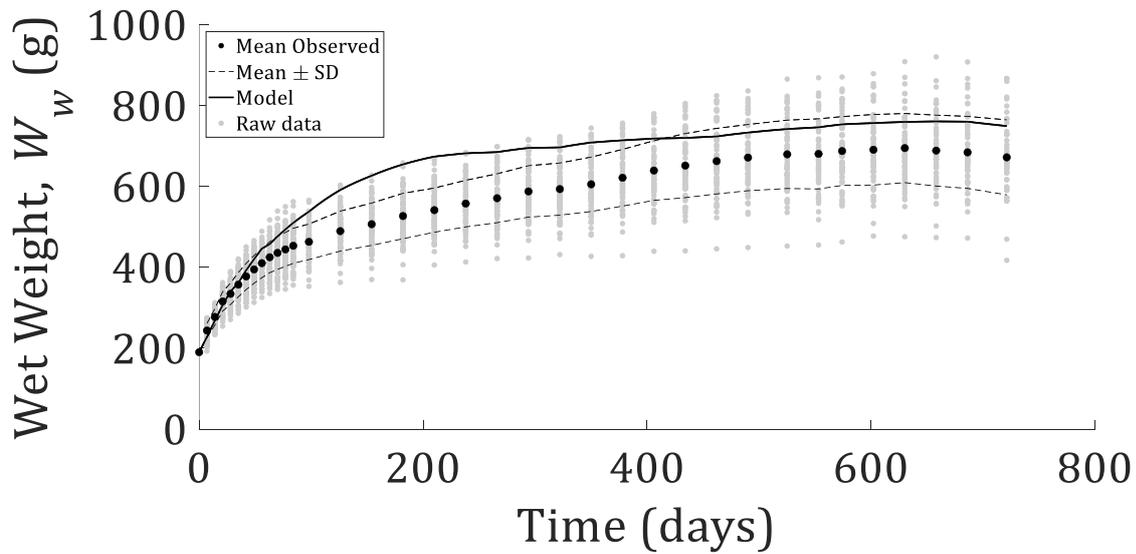

Figure C10 Observed (circles) and predicted (solid line) wet weight of female rats in group C over two years, using Method 1. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

Method 2 - Dataset B

Figure C11 Observed (circles) and predicted (solid line) wet weight of male rats in group B over two years, using Method 2. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.
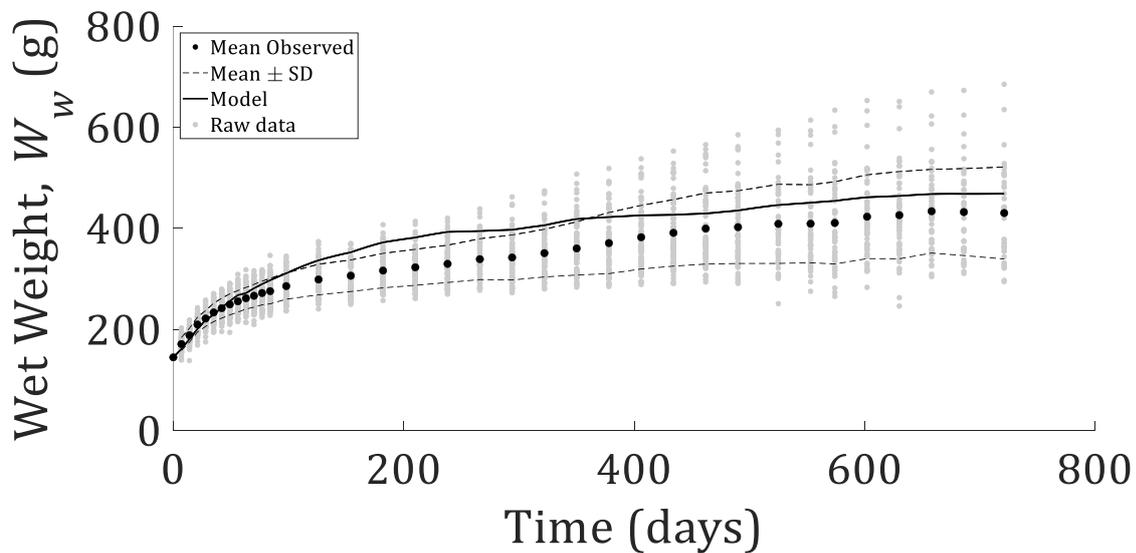

Figure C12 Observed (circles) and predicted (solid line) wet weight of female rats in group B over two years, using Method 2. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.
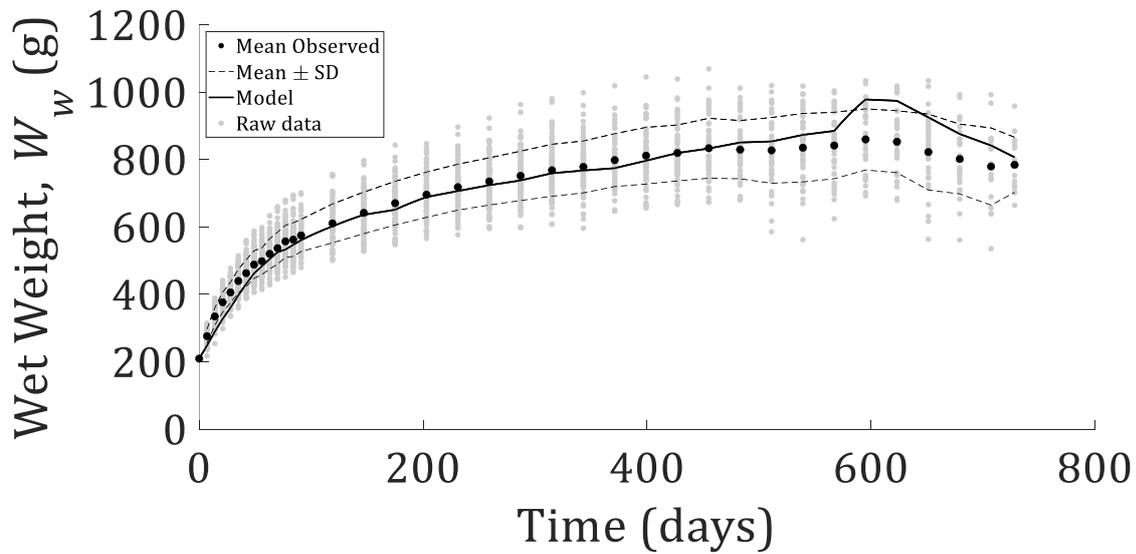
Method 2 - Dataset C

Figure C13 Observed (circles) and predicted (solid line) wet weight of male rats in group C over two years, using Method 2. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

Figure C14 Observed (circles) and predicted (solid line) wet weight of female rats in group C over two years, using Method 2. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

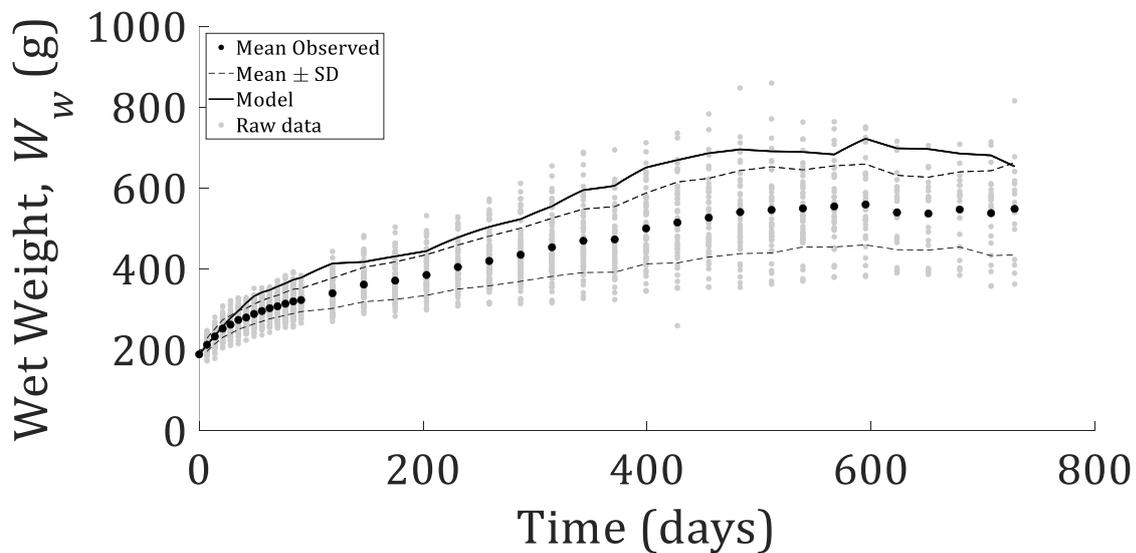**Method 3 – Dataset C**

Figure C15 Observed (circles) and predicted (solid line) wet weight of male rats in group B over two years, using Method 3. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.
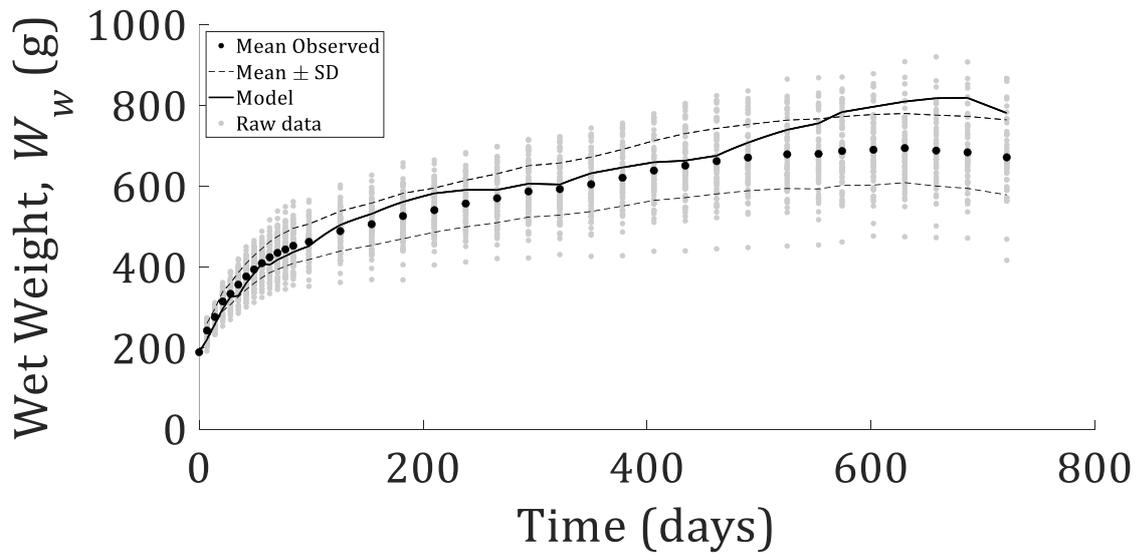

Figure C16 Observed (circles) and predicted (solid line) wet weight of female rats in group B over two years, using Method 3. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

Method 3 – Dataset C

Figure C17 Observed (circles) and predicted (solid line) wet weight of male rats in group C over two years, using Method 3. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.
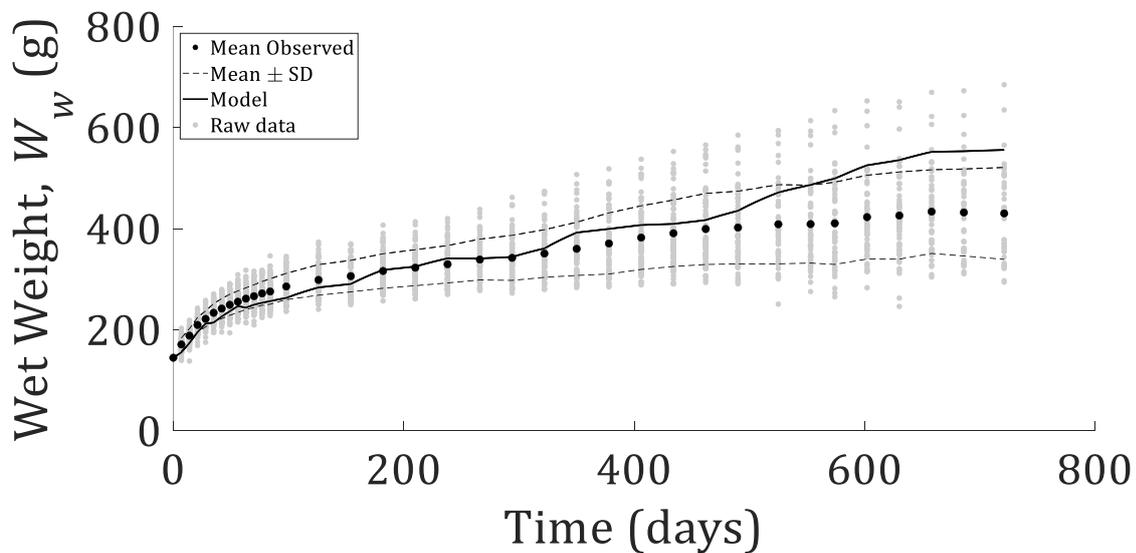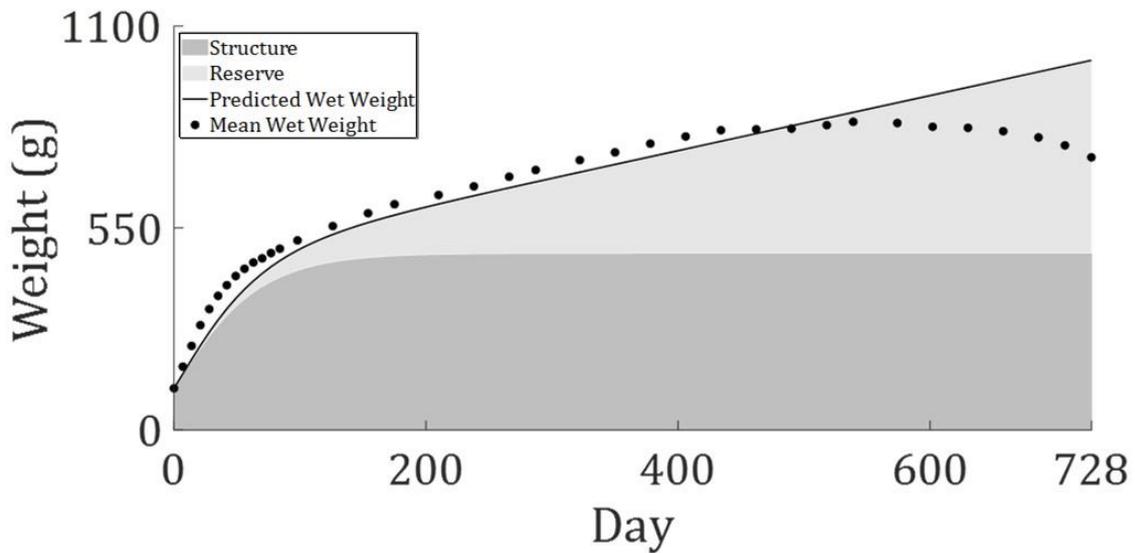

Figure C18 Observed (circles) and predicted (solid line) wet weight of female rats in group C over two years, using Method 3. Mean observations are shown in black while grey circles represent raw data. Dashed lines show the observed mean ±1 SD.

Modelled Structure, Reserve and Wet Weight– Calibration Data

Method 1 – Dataset A

Figure C19 Wet weight of male rats in group A over two years, modelled using Method 1 (solid line), fitted to mean observations (black circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
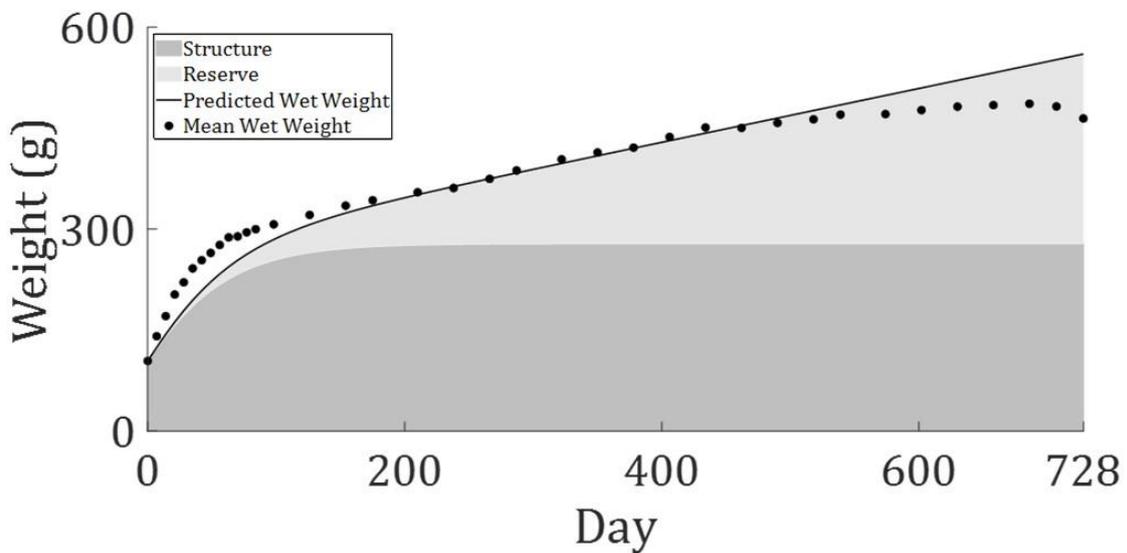
Figure C20 Wet weight of female rats in group A over two years, modelled using Method 1 (solid line), fitted to mean observations (black circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Method 2- Dataset A

Figure C21 Wet weight of male rats in group A over two years, modelled using Method 2 (solid line), fitted to mean observations (black circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
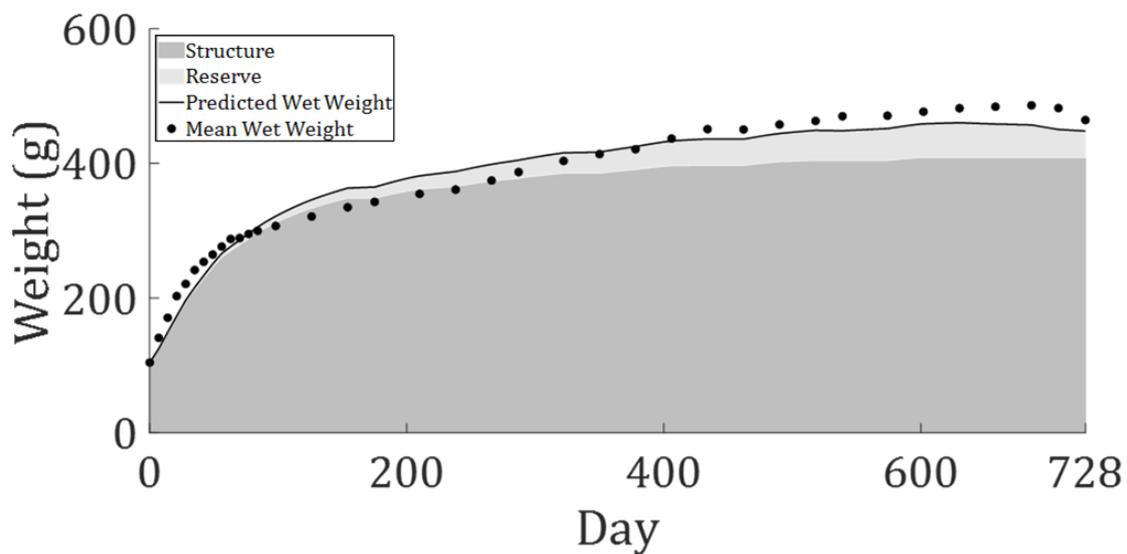
Figure C22 Wet weight of female rats in group A over two years, modelled using Method 2 (solid line), fitted to mean observations (black circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
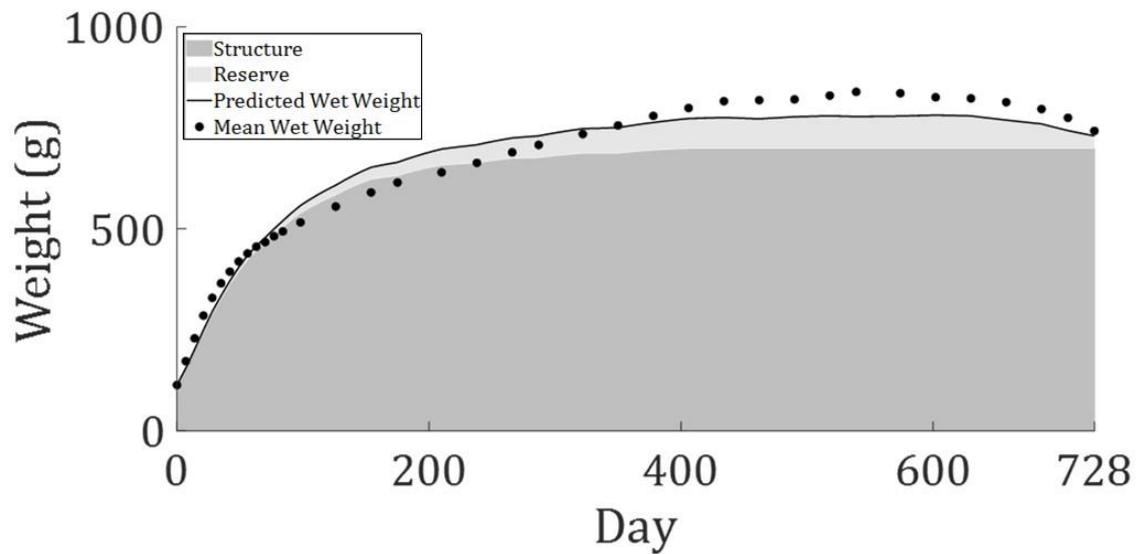
Method 3 – Dataset A

Figure C23 Wet weight of male rats in group A over two years, modelled using Method 3 (solid line), fitted to mean observations (black circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
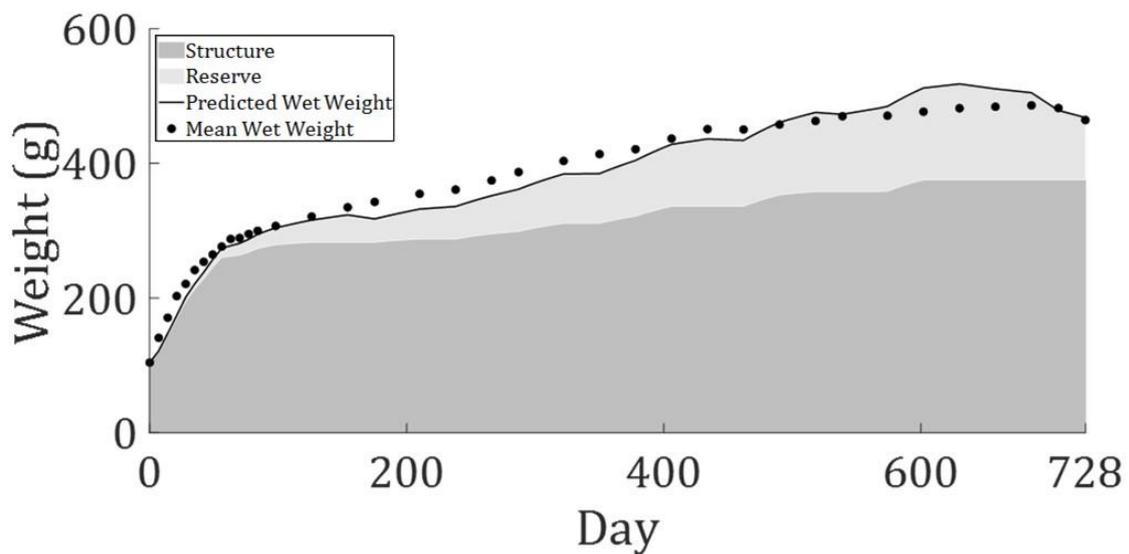
Figure C24 Wet weight of female rats in group A over two years, modelled using Method 3 (solid line), fitted to mean observations (black circles). The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
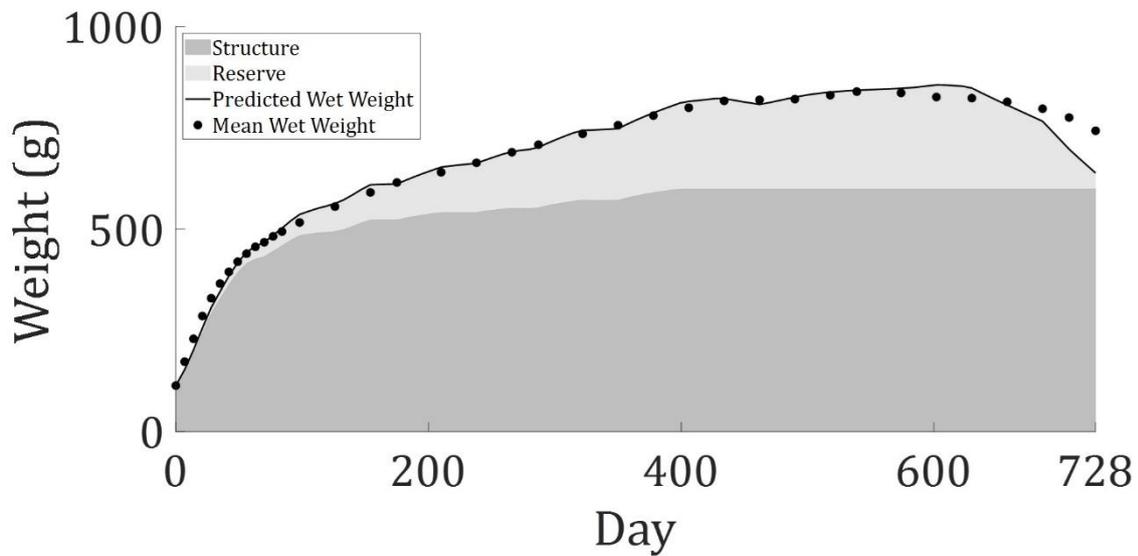
Predicted Structure, Reserve and Wet Weight – Independent Data

Method 1 – Dataset B

Figure C25 Observed (circles) and predicted (solid line) wet weight of male rats in group B over two years, using Method 1. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
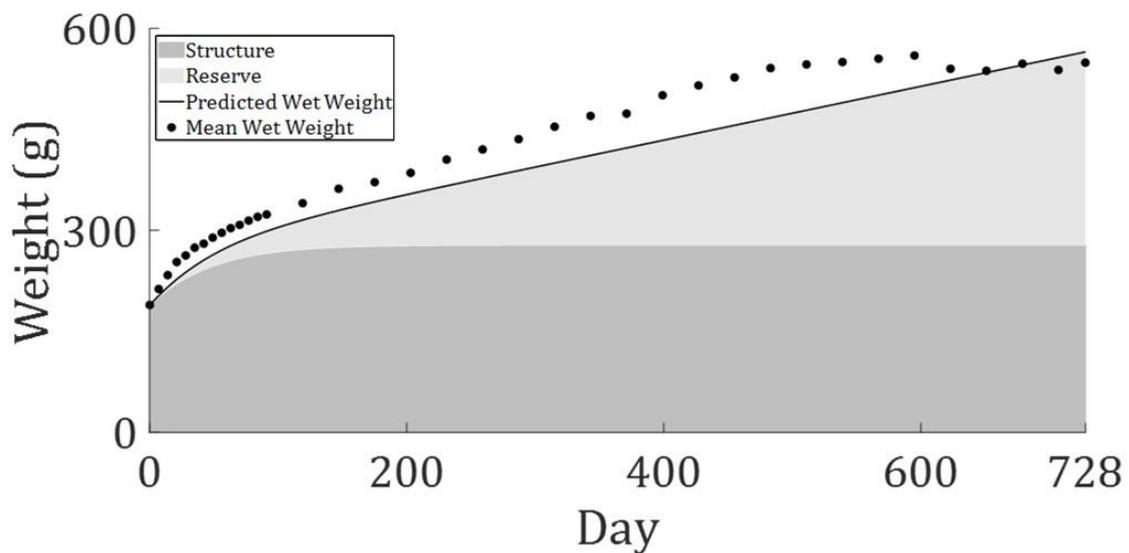
Figure C26 Observed (circles) and predicted (solid line) wet weight of female rats in group B over two years, using Method 1. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Method 1 – Dataset C

Figure C27 Observed (circles) and predicted (solid line) wet weight of male rats in group C over two years, using Method 1. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Figure C28 Observed (circles) and predicted (solid line) wet weight of female rats in group C over two years, using Method 1. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
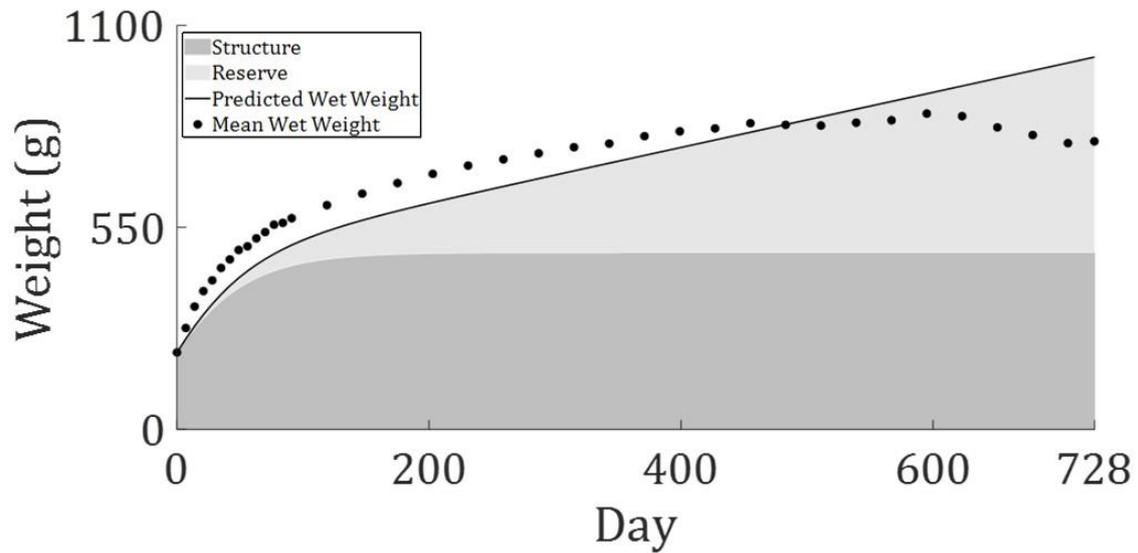
Method 2 - Dataset B

Figure C29 Observed (circles) and predicted (solid line) wet weight of male rats in group B over two years, using Method 2. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Figure C30 Observed (circles) and predicted (solid line) wet weight of female rats in group B over two years, using Method 2. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Method 2 - Dataset C

Figure C31 Observed (circles) and predicted (solid line) wet weight of male rats in group C over two years, using Method 2. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Figure C32 Observed (circles) and predicted (solid line) wet weight of female rats in group C over two years, using Method 2. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
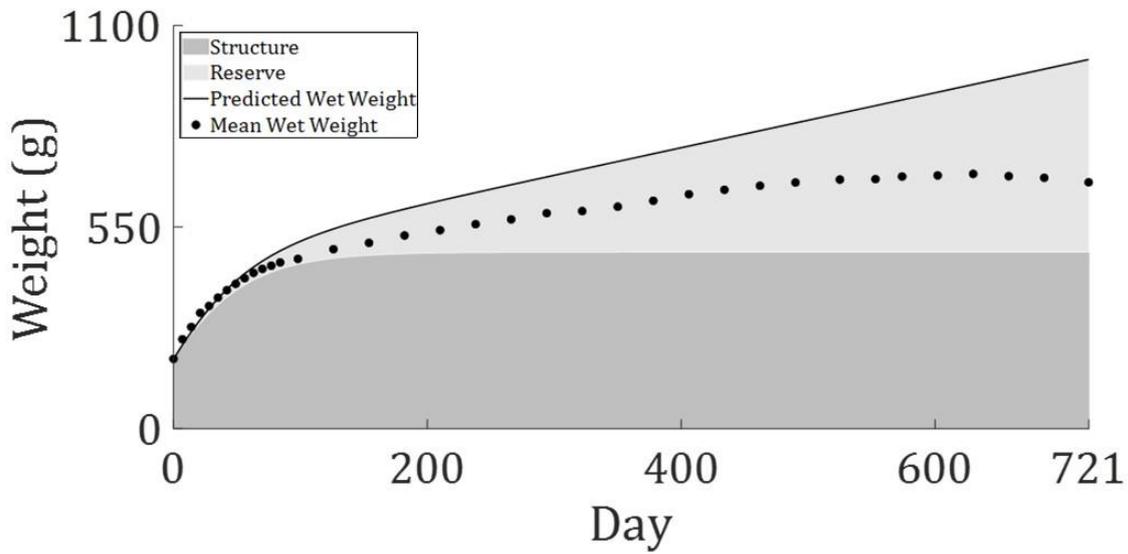
Method 3 – Dataset C

Figure C33 Observed (circles) and predicted (solid line) wet weight of male rats in group B over two years, using Method 3. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
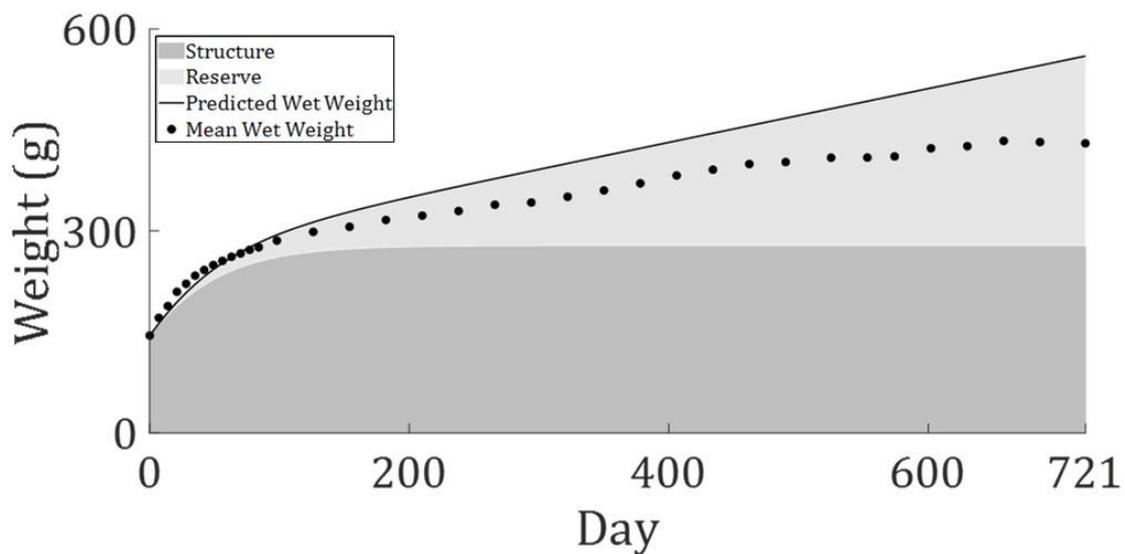

Figure C34 Observed (circles) and predicted (solid line) wet weight of female rats in group B over two years, using Method 3. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

Method 3 – Dataset C

Figure C35 Observed (circles) and predicted (solid line) wet weight of male rats in group C over two years, using Method 3. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).
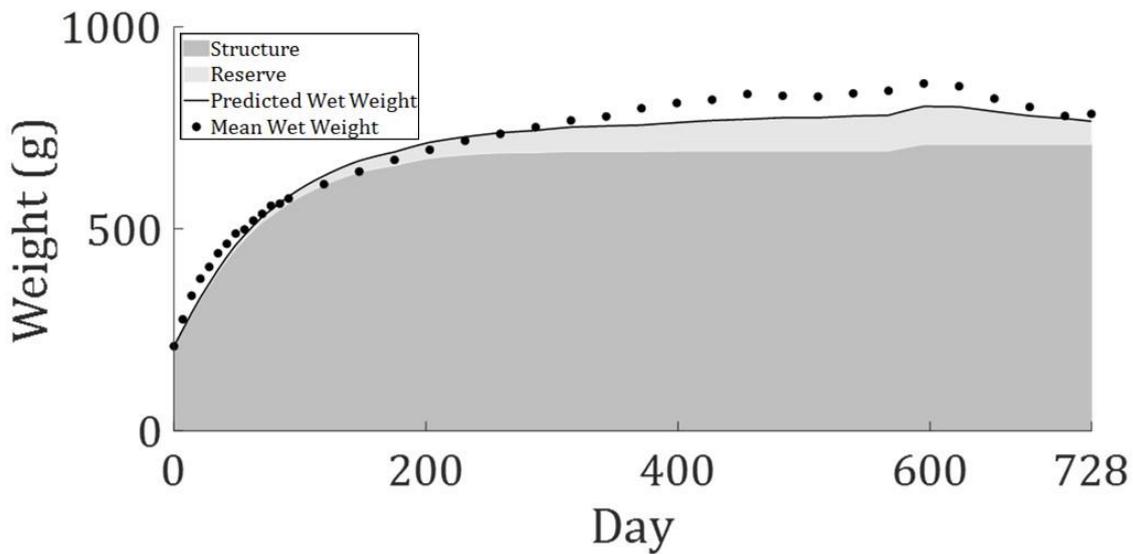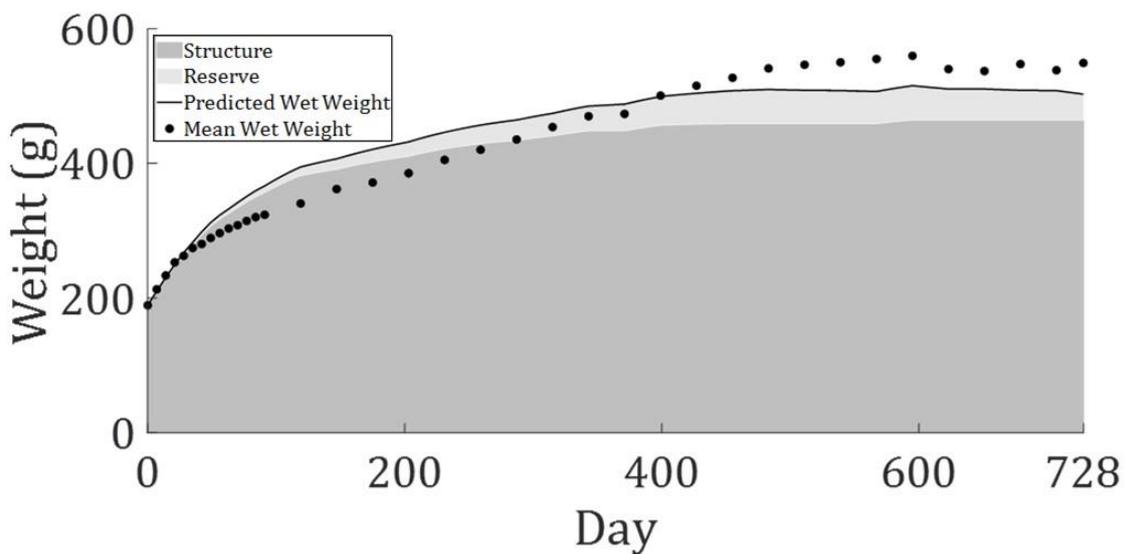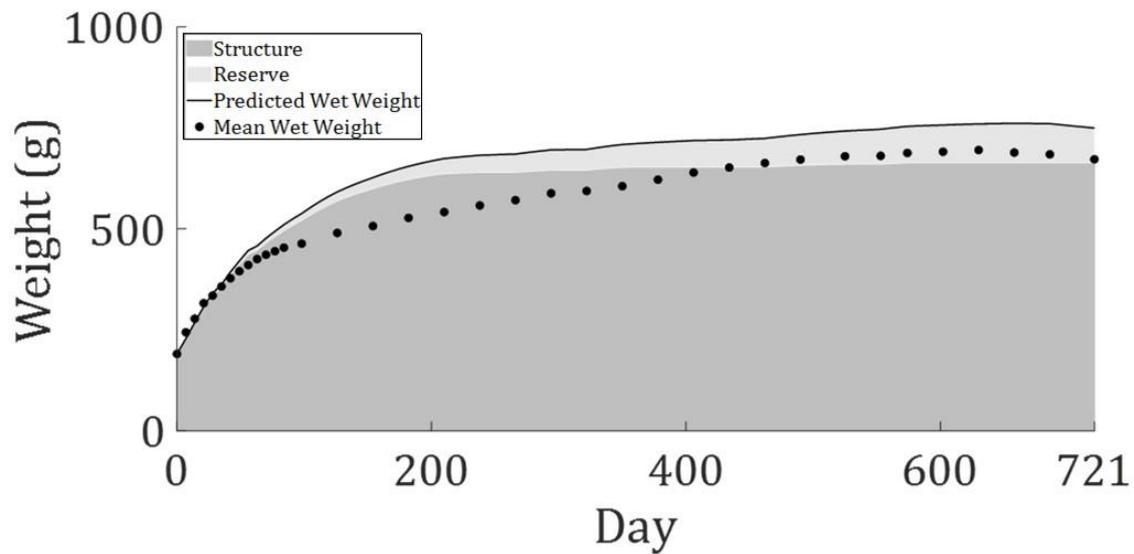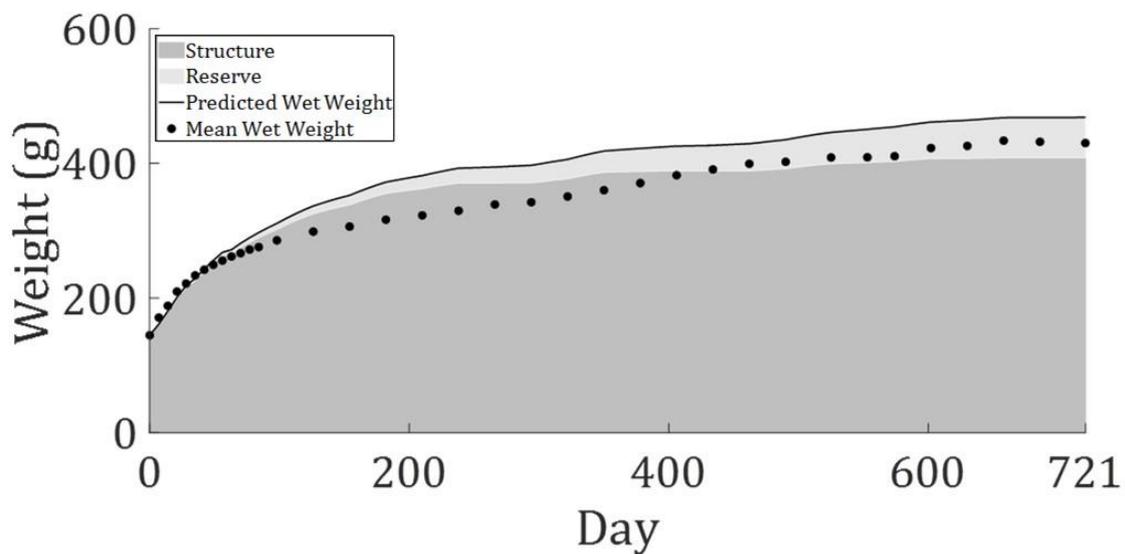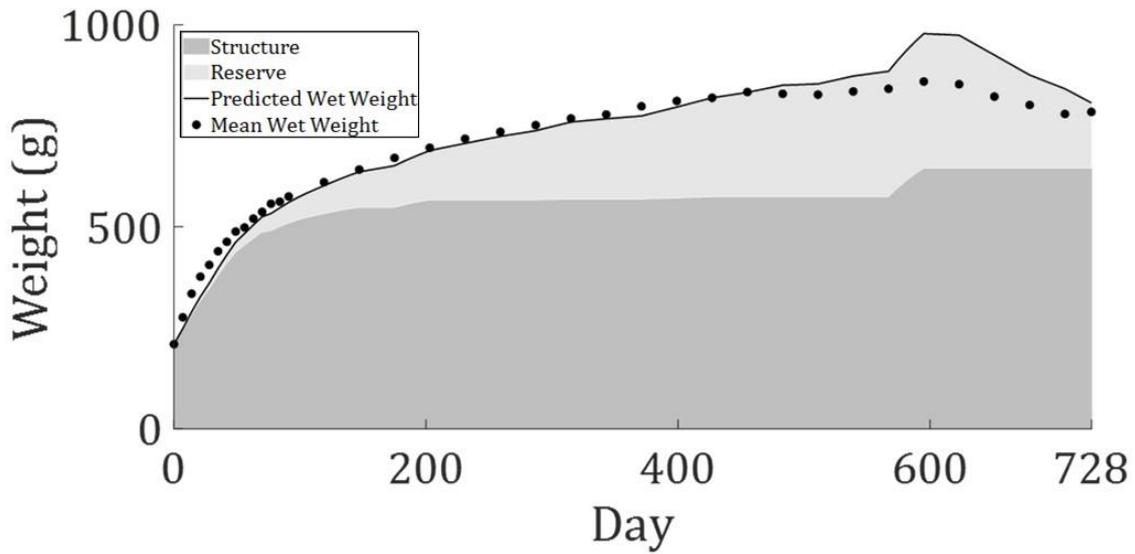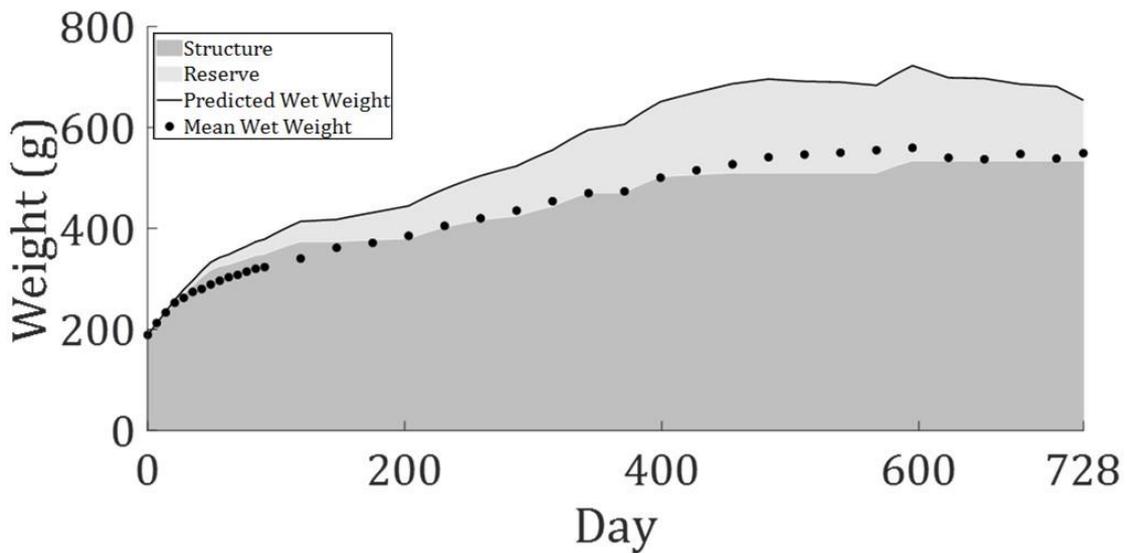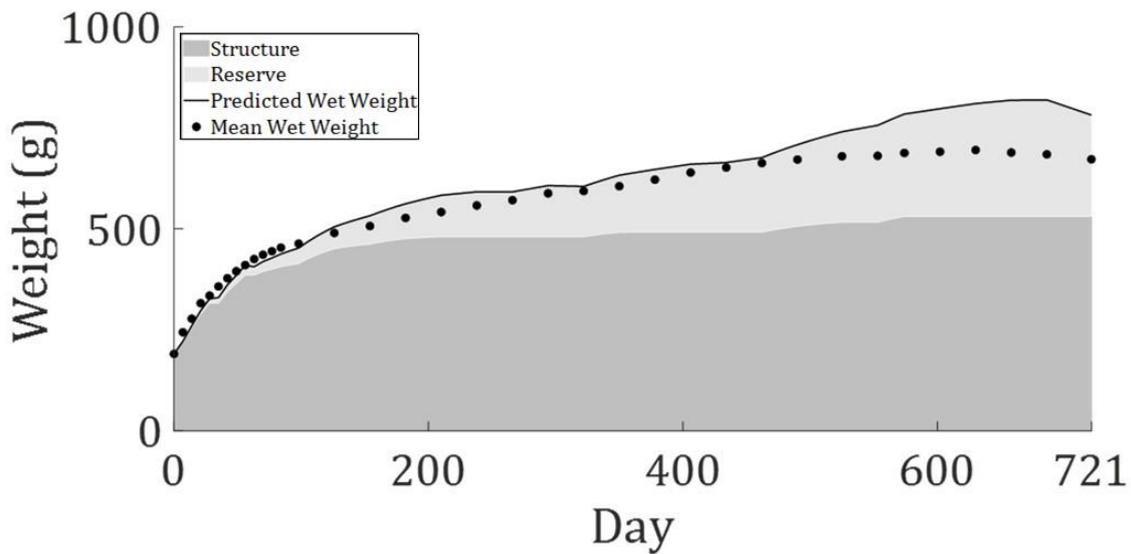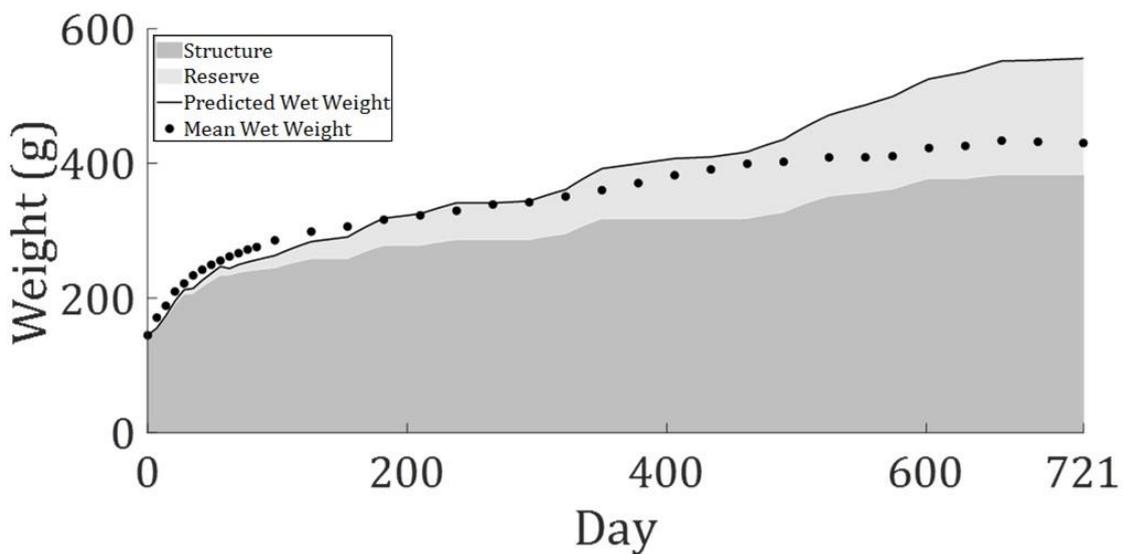

Figure C36 Observed (circles) and predicted (solid line) wet weight of female rats in group c over two years, using Method 3. The shaded area under the model curves shows structure (dark grey) and reserve (light grey).

References

ADLER, S., BASKETTER, D., CRETON, S., PELKONEN, O., VAN BENTHEM, J., ZUANG, V., ANDERSEN, K. E., ANGERS-LOUSTAU, A., APTULA, A., BAL-PRICE, A., BENFENATI, E., BERNAUER, U., BESSEMS, J., BOIS, F. Y., BOOBIS, A., BRANDON, E., BREMER, S., BROSCHARD, T., CASATI, S., COECKE, S., CORVI, R., CRONIN, M., DASTON, G., DEKANT, W., FELTER, S., GRIGNARD, E., GUNDERT-REMY, U., HEINONEN, T., KIMBER, I., KLEINJANS, J., KOMULAINEN, H., KREILING, R., KREYSA, J., LEITE, S. B., LOIZOU, G., MAXWELL, G., MAZZATORTA, P., MUNN, S., PFUHLER, S., PHRAKONKHAM, P., PIERSMA, A., POTH, A., PRIETO, P., REPETTO, G., ROGIERS, V., SCHOETERS, G., SCHWARZ, M., SERAFIMOVA, R., TAHTI, H., TESTAI, E., VAN DELFT, J., VAN LOVEREN, H., VINKEN, M., WORTH, A. & ZALDIVAR, J. M. 2011. Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology,* 85**,** 367-485.

ALVAREZ, O. A., JAGER, T., REDONDO, E. M. & KAMMENGA, J. E. 2006. Physiological modes of action of toxic chemicals in the nematode Acrobeloides nanus. *Environmental Toxicology and Chemistry,* 25**,** 3230-3237.

ASHAUER, R. 2010. Toxicokinetic-toxicodynamic modelling in an individual based context-Consequences of parameter variability. *Ecological Modelling,* 221**,** 1325-1328.

ASHAUER, R., AGATZ, A., ALBERT, C., DUCROT, V., GALIC, N., HENDRIKS, J., JAGER, T., KRETSCHMANN, A., O'CONNOR, I., RUBACH, M. N., NYMAN, A. M., SCHMITT, W., STADNICKA, J., VAN DEN BRINK, P. J. & PREUSS, T. G. 2011. Toxicokinetic-toxicodynamic modeling of quantal and graded sublethal endpoints: A brief discussion of concepts. *Environmental Toxicology and Chemistry,* 30**,** 2519-2524.

ASHAUER, R. & ESCHER, B. I. 2010. Advantages of toxicokinetic and toxicodynamic modelling in aquatic ecotoxicology and risk assessment. *Journal of Environmental Monitoring,* 12**,** 2056-2061.

ASHAUER, R. & JAGER, T. 2018. Physiological modes of action across species and toxicants: the key to predictive ecotoxicology. *Environmental Science-Processes & Impacts,* 20**,** 48-57.

ASHAUER, R., O'CONNOR, I. & ESCHER, B. I. 2017. Toxic Mixtures in Time-The Sequence Makes the Poison. *Environmental Science & Technology,* 51**,** 3084-3092.

ASHAUER, R., O'CONNOR, I., HINTERMEISTER, A. & ESCHER, B. I. 2015. Death Dilemma and Organism Recovery in Ecotoxicology. *Environmental Science & Technology,* 49**,** 10136-10146.

ASHAUER, R., THORBEK, P., WARINTON, J. S., WHEELER, J. R. & MAUND, S. 2013. A method to predict and understand fish survival under dynamic chemical stress using standard ecotoxicity data. *Environmental Toxicology and Chemistry,* 32**,** 954-965.

ASTASHKINA, A., MANN, B. & GRAINGER, D. W. 2012. A critical evaluation of in vitro cell culture models for high-throughput drug screening and toxicity. *Pharmacology & Therapeutics,* 134**,** 82-106.

BARATA, C., BAIRD, D. J., MITCHELL, S. E. & SOARES, A. 2002. Among- and within-population variability in tolerance to cadmium stress in natural populations of Daphnia magna: Implications for ecological risk assessment. *Environmental Toxicology and Chemistry,* 21**,** 1058-1064.

BATZLI, G. O. & COLE, F. R. 1979. Nutritional ecology of microtine rodents - digestibility of forage. *Journal of Mammalogy,* 60**,** 740-750.

BEDNARSKA, A. J., EDWARDS, P., SIBLY, R. & THORBEK, P. 2013a. A toxicokinetic model for thiamethoxam in rats: implications for higher-tier risk assessment. *Ecotoxicology,* 22**,** 548-557.

BEDNARSKA, A. J., JEVTIC, D. M. & LASKOWSKI, R. 2013b. More ecological ERA: Incorporating natural environmental factors and animal behavior. *Integrated Environmental Assessment and Management,* 9**,** E39-E46.

BLAAUBOER, B. J. 2010. Biokinetic modeling and in vitro-in vivo extrapolations. *Journal of Toxicology and Environmental Health-Part B-Critical Reviews,* 13**,** 242-252.

BROWN, J. H., GILLOOLY, J. F., ALLEN, A. P., SAVAGE, V. M. & WEST, G. B. 2004a. Response to forum commentary on "toward a metabolic theory of ecology". *Ecology,* 85**,** 1818-1821.

BROWN, J. H., GILLOOLY, J. F., ALLEN, A. P., SAVAGE, V. M. & WEST, G. B. 2004b. Toward a metabolic theory of ecology. *Ecology,* 85**,** 1771-1789.

BROWN, J. H. & SIBLY, R. M. 2012. The Metabolic Theory of Ecology and Its Central Equation. *Metabolic Ecology.* John Wiley & Sons, Ltd.

BROWN, J. H., WEST, G. B. & ENQUIST, B. J. 2005. Yes, West, Brown and Enquist's model of allometric scaling is both mathematically correct and biologically relevant. *Functional Ecology,* 19**,** 735-738.

BURTHE, S. J., LAMBIN, X., TELFER, S., DOUGLAS, A., BELDOMENICO, P., SMITH, A. & BEGON, M. 2010. Individual growth rates in natural field vole, Microtus agrestis, populations exhibiting cyclic population dynamics. *Oecologia,* 162**,** 653-661.

BUWALDA, B., BLOM, W. A. M., KOOLHAAS, J. M. & VAN DIJK, G. 2001. Behavioral and physiological responses to stress are affected by high-fat feeding in male rats. *Physiology & Behavior,* 73**,** 371-377.

CACCIA, S., CAPPI, M., FRACASSO, C. & GARATTINI, S. 1990. Influence of dose and route of exposure on the kinetics of fluoxetine and its metabolite norfluoxetine in the rat. *Psychopharmacology,* 100**,** 509-514.

CALABRESE, E. J. 1986. Sex differences in susceptibility to toxic industrial chemicals. *British Journal of Industrial Medicine,* 43**,** 577-579.

CAO, Q. M., YU, J. & CONNELL, D. 2014. New Allometric Scaling Relationships and Applications for Dose and Toxicity Extrapolation. *International Journal of Toxicology,* 33**,** 482-489.

CASTANO, A., BOLS, N., BRAUNBECK, T., DIERICKX, P., HALDER, M., ISOMAA, B., KAWAHARA, K., LEE, L. E. J., MOTHERSILL, C., PART, P., REPETTO, G., SINTES, J. R., RUFLI, H., SMITH, R., WOOD, C. & SEGNER, H. 2003. The use of fish cells in ecotoxicology - The report and recommendations of ECVAM workshop 47. *Atla-Alternatives to Laboratory Animals,* 31**,** 317-351.

COECKE, S., AHR, H., BLAAUBOER, B. J., BREMER, S., CASATI, S., CASTELL, J., COMBES, R., CORVI, R., CRESPI, C. L., CUNNINGHAM, M. L., ELAUT, G., ELETTI, B., FREIDIG, A., GENNARI, A., GHERSI-EGEA, J.-F., GUILLOUZO, A., HARTUNG, T., HOET, P., INGELMAN-SUNDBERG, M., MUNN, S., JANSSENS, W., LADSTETTER, B., LEAHY, D., LONG, A., MENEGUZ, A., MONSHOUWER, M., MORATH, S., NAGELKERKE, F., PELKONEN, O., PONTI, J., PRIETO, P., RICHERT, L., SABBIONI, E., SCHAACK, B., STEILING, W., TESTAI, E., VERICAT, J.-A. & WORTH, A. 2006. Metabolism: A Bottleneck in In Vitro Toxicological Test Development: The Report and Recommendations of ECVAM Workshop 541. *Alternatives to Laboratory Animals,* 34**,** 49-84.

CROCKER, D. R. 2005. Estimating the exposure of birds and mammals to pesticides in long-term risk assessments. *Ecotoxicology,* 14**,** 833-851.

CYBULSKI, W., RADKO, L. & RZESKI, W. 2015. Cytotoxicity of monensin, narasin and salinomycin and their interaction with silybin in HepG2, LMH and L6 cell cultures. *Toxicology in Vitro,* 29**,** 337-344.

DEMMENT, M. W. & VANSOEST, P. J. 1985. A nutritional explanation for body-size patterns of ruminant and nonruminant herbivores. *American Naturalist,* 125**,** 641-672.

DESFORGES, J. P. W., SONNE, C. & DIETZ, R. 2017. Using energy budgets to combine ecology and toxicology in a mammalian sentinel species. *Scientific Reports,* 7.

DI GUARDO, A. & HERMENS, J. L. M. 2013. Challenges for exposure prediction in ecological risk assessment. *Integrated Environmental Assessment and Management,* 9**,** E4-E14.

DOETSCHMAN, T. C., EISTETTER, H., KATZ, M., SCHMIDT, W. & KEMLER, R. 1985. The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of Embryology and Experimental Morphology,* 87**,** 27.

DOHMEN, G. P., PREUSS, T. G., HAMER, M., GALIC, N., STRAUSS, T., VAN DEN BRINK, P. J., DE LAENDER, F. & BOPP, S. 2016. Population-Level Effects and Recovery of Aquatic Invertebrates after Multiple Applications of an Insecticide. *Integrated Environmental Assessment and Management,* 12**,** 67-81.

DONLEY, N. 2019. The USA lags behind other agricultural nations in banning harmful pesticides. *Environmental Health,* 18**,** 44.

DUCROT, V., ASHAUER, R., BEDNARSKA, A. J., HINAREJOS, S., THORBEK, P. & WEYMAN, G. 2016. Using Toxicokinetic-Toxicodynamic Modeling as an Acute Risk Assessment Refinement Approach in Vertebrate Ecological Risk Assessment. *Integrated Environmental Assessment and Management,* 12**,** 32-45.

EFSA 2006. Opinion of the Scientific Panel on Plant protection products and their residues (PPR) related to the choice of endpoints to assess the long term risk to mammals. *EFSA Journal,* 4**,** 344-n/a.

EFSA 2009a. Guidance of the Scientific Committee on Use of the benchmark dose approach in risk assessment. *EFSA Journal,* 7**,** 1150-n/a.

EFSA 2009b. Risk Assessment for Birds and Mammals. *EFSA Journal,* 7**,** 1438.

EFSA 2018. Scientific Opinion on the state of the art of Toxicokinetic/Toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. *EFSA Journal,* 16.

EPA PRESS OFFICE 2019. Administrator Wheeler Signs Memo to Reduce Animal Testing, Awards $4.25 Million to Advance Research on Alternative Methods to Animal Testing.

EUROPEAN COMMISSION 1994. Commission Directive 94/79/EC of 21 December 1994 amending Council Directive 91/414/EEC concerning the placing of plant protection products on the market.

EUROPEAN COMMISSION 2006. Draft Renewal Assessment Report - Fludioxonil Volume 3, Annex B.6.

EUROPEAN COMMISSION 2009. Draft Renewal Assessment Report - Azoxystrobin Volume 3, Annex B.6.

EUROPEAN COMMISSION. 2013a. *Ban on Animal Testing* [Online]. Available: https://ec.europa.eu/growth/sectors/cosmetics/animal-testing_en [Accessed 15/03/2017 2017].

EUROPEAN COMMISSION 2013b. Commission Staff Working Document accompanying the Seventh Report on the Statistics on the Number of Animals used for Experimental and other Scientific Purposes in the Member States of the European Union. Brussels.

EUROPEAN COMMISSION 2013c. Draft Renewal Assessment Report - Prosulfuron Volume 3, Annex B.6.

EUROPEAN COMMISSION 2014. Addressing the new challenges for risk assessment. *In:* DIRECTORATE-GENERAL FOR HEALTH AND CONSUMERS (ed.).

EUROPEAN COMMISSION 2019. Report from the commission to the European parliament and the council 2019 report on the statistics on the use of animals for scientific purposes in the Member States of the European Union in 2015-2017. Brussels: European Union.

EUROPEAN PARLIAMENT 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. The Official Journal of The European Union, 276.

FAO & WHO 2011. Joint Meeting on Pesticide Residues Report 2010.

FISCHER, D. L. 2005. Accounting for differing exposure patterns between laboratory tests and the field in the assessment of long-term risks of pesticides to terrestrial vertebrates. *Ecotoxicology,* 14**,** 853-862.

FISCHER, F. C., ABELE, C., DROGE, S. T. J., HENNEBERGER, L., KONIG, M., SCHLICHTING, R., SCHOLZ, S. & ESCHER, B. I. 2018. Cellular Uptake Kinetics of Neutral and Charged Chemicals in in Vitro Assays Measured by Fluorescence Microscopy. *Chemical Research in Toxicology,* 31**,** 646-657.

FISCHER, F. C., HENNEBERGER, L., KÖNIG, M., BITTERMANN, K., LINDEN, L., GOSS, K.-U. & ESCHER, B. I. 2017. Modeling Exposure in the Tox21 in Vitro Bioassays. *Chemical Research in Toxicology,* 30**,** 1197-1208.

FORBES, V. E. & CALOW, P. 2012. Promises and problems for the new paradigm for risk assessment and an alternative approach involving predictive systems models. *Environmental Toxicology and Chemistry,* 31**,** 2663-2671.

FORBES, V. E., HOMMEN, U., THORBEK, P., HEIMBACH, F., VAN DEN BRINK, P. J., WOGRAM, J., THULKE, H.-H. & GRIMM, V. 2009. Ecological models in support of regulatory risk assessments of

pesticides: developing a strategy for the future. *Integrated Environmental Assessment and Management,* 5**,** 167-172.

FOSTER, M. L., BARTNIKAS, T. B., JOHNSON, L. C., HERRERA, C., PETTIGLIO, M. A., KEENE, A. M., TAYLOR, M. D. & DORMAN, D. C. 2015. Pharmacokinetic Evaluation of the Equivalency of Gavage, Dietary, and Drinking Water Exposure to Manganese in F344 Rats. *Toxicological Sciences,* 145**,** 244-251.

GAJEWSKA, M., WORTH, A., URANI, C., BRIESEN, H. & SCHRAMM, K. W. 2014. Application of physiologically-based toxicokinetic modelling in oral-to-dermal extrapolation of threshold doses of cosmetic ingredients. *Toxicology Letters,* 227**,** 189-202.

GERGS, A., GABSI, F., ZENKER, A. & PREUSS, T. G. 2016. Demographic Toxicokinetic-Toxicodynamic Modeling of Lethal Effects. *Environmental Science & Technology,* 50**,** 6017-6024.

GODFREY, K. R., ARUNDEL, P. A., DONG, Z. M. & BRYANT, R. 2011. Modelling the Double Peak Phenomenon in pharmacokinetics. *Computer Methods and Programs in Biomedicine,* 104**,** 62-69.

GREEN, J. W., SPRINGER, T. A. & STAVELEY, J. P. 2013. The drive to ban the NOEC/LOEC in favor of ECx is misguided and misinformed. *Integrated Environmental Assessment and Management,* 9**,** 12-16.

GRIMM, V. & MARTIN, B. T. 2013. Mechanistic effect modeling for ecological risk assessment: Where to go from here? *Integrated Environmental Assessment and Management,* 9**,** E58-E63.

GROOTHUIS, F. A., HERINGA, M. B., NICOL, B., HERMENS, J. L. M., BLAAUBOER, B. J. & KRAMER, N. I. 2015. Dose metric considerations in in vitro assays to improve quantitative in vitro-in vivo dose extrapolations. *Toxicology,* 332**,** 30-40.

HAMON, J., RENNER, M., JAMEI, M., LUKAS, A., KOPP-SCHNEIDER, A. & BOIS, F. Y. 2015. Quantitative in vitro to in vivo extrapolation of tissues toxicity. *Toxicology in Vitro,* 30**,** 203-216.

HANSSON, L. & JAAROLA, M. 1989. Body size related to cyclicity in microtines - dominance behavior or digestive efficiency. *Oikos,* 55**,** 356-364.

HART, A. D. M. & THOMPSON, H. M. 2005. Improved approaches to assessing long-term risks to birds and mammals. *Ecotoxicology,* 14**,** 771-773.

HARTUNG, T. & ROVIDA, C. 2009. Chemical regulators have overreached. *Nature,* 460**,** 1080-1081.

HOME OFFICE. 2017. *Research and Testing Using Animals* [Online]. Available: https://www.gov.uk/guidance/research-and-testing-using-animals [Accessed 28/02/2017 2017].

HOME OFFICE 2019. Annual Statistics of Scientific Procedures on Living   Animals Great Britain 2018.

HOME OFFICE 2020. Annual Statistics of Scientific Procedures on Living Animals, Great Britain 2019.

HOME OFFICE, DEPARTMENT FOR BUSINESS INNOVATION & SKILLS & DEPARTMENT OF HEALTH AND SOCIAL CARE 2014. Working to reduce the use of animals in research: delivery plan.

HOMMEN, U., FORBES, V., GRIMM, V., PREUSS, T. G., THORBEK, P. & DUCROT, V. 2016a. How to Use Mechanistic Effect Models in Environmental Risk Assessment of Pesticides: Case Studies and Recommendations from the SETAC Workshop MODELINK. *Integrated Environmental Assessment and Management,* 12**,** 21-31.

HOMMEN, U., SCHMITT, W., HEINE, S., BROCK, T. C. M., DUQUESNE, S., MANSON, P., MEREGALLI, G., OCHOA-ACUNA, H., VAN VLIET, P. & ARTS, G. 2016b. How TK-TD and Population Models for Aquatic Macrophytes Could Support the Risk Assessment for Plant Protection Products. *Integrated Environmental Assessment and Management,* 12**,** 82-95.

HOULAHAN, J. E., MCKINNEY, S. T. & ROCHETTE, R. 2015. On Theory in Ecology: Another Perspective. *Bioscience,* 65**,** 341-342.

HUBERT, M. F., LAROQUE, P., GILLET, J. P. & KEENAN, K. P. 2000. The effects of diet, ad libitum feeding, and moderate and severe dietary restriction on body weight, survival, clinical pathology parameters, and cause of death in control Sprague-Dawley rats. *Toxicological Sciences,* 58**,** 195-207.

HUNKA, A. D., MELI, M., THIT, A., PALMQVIST, A., THORBEK, P. & FORBES, V. E. 2013. Stakeholders' Perspective on Ecological Modeling in Environmental Risk Assessment of Pesticides: Challenges and Opportunities. *Risk Analysis,* 33**,** 68-79.

ICCA 2011. ICCA Guidance on Chemical Risk Assessment. 2nd Edition ed.: International Control of Chemical Associations.

ILLIUS, A. W. & GORDON, I. J. 1992. Modeling the nutritional ecology of ungulate herbivores - evolution of body size and competitive interactions. *Oecologia,* 89**,** 428-434.

JAGER, T. 2011. Some Good Reasons to Ban ECx and Related Concepts in Ecotoxicology. *Environmental Science & Technology,* 45**,** 8180-8181.

JAGER, T. 2012. Bad habits die hard: The NOEC's persistence reflects poorly on ecotoxicology. *Environmental Toxicology and Chemistry,* 31**,** 228-229.

JAGER, T. 2013. All Individuals Are Not Created Equal; Accounting for Interindividual Variation in Fitting Life-History Responses to Toxicants. *Environmental Science & Technology,* 47**,** 1664-1669.

JAGER, T. 2015a. Making sense of chemical stress. *Application of Dynamic Energy Budget theory in ecotoxicology and stress ecology.* Leanpub.

JAGER, T. 2015b. Making sense of chemical stress: Application of Dynamic Energy Budget theory in ecotoxicology and stress ecology. Leanpub.

JAGER, T. 2016. Predicting environmental risk: A road map for the future. *Journal of Toxicology and Environmental Health-Part a-Current Issues,* 79**,** 572-584.

JAGER, T. 2019. *BYOM Flexible Model Platform* [Online]. http://www.debtox.info/byom.html. Available: http://www.debtox.info/byom.html [Accessed 04/10/2019].

JAGER, T. 2020. Revisiting simplified DEBtox models for analysing ecotoxicity data. *Ecological Modelling,* 416**,** 108904.

JAGER, T., ALBERT, C., PREUSS, T. G. & ASHAUER, R. 2011. General Unified Threshold Model of Survival - a Toxicokinetic-Toxicodynamic Framework for Ecotoxicology. *Environmental Science & Technology,* 45**,** 2529-2540.

JAGER, T. & ASHAUER, R. 2018. How to Evaluate the Quality of ToxicokineticToxicodynamic Models in the Context of Environmental Risk Assessment. *Integrated Environmental Assessment and Management,* 14**,** 604-614.

JAGER, T., HEUGENS, E. H. W. & KOOIJMAN, S. A. L. M. 2006. Making sense of ecotoxicological test results: Towards application of process-based models. *Ecotoxicology,* 15**,** 305-314.

JAGER, T., MARTIN, B. T. & ZIMMER, E. I. 2013. DEBkiss or the quest for the simplest generic model of animal life history. *Journal of Theoretical Biology,* 328**,** 9-18.

JAGER, T. & ZIMMER, E. I. 2012. Simplified Dynamic Energy Budget model for analysing ecotoxicity data. *Ecological Modelling,* 225**,** 74-81.

KALAM, Y., ISBISTER, G. K., MIRTSCHIN, P., HODGSON, W. C. & KONSTANTAKOPOULOS, N. 2011. Validation of a cell-based assay to differentiate between the cytotoxic effects of elapid snake venoms. *Journal of Pharmacological and Toxicological Methods,* 63**,** 137-142.

KAPETANOVIC, I. M., KRISHNARAJ, R., MARTIN-JIMENEZ, T., YUAN, L., VAN BREEMEN, R. B. & LYUBIMOV, A. 2006. Effects of oral dosing paradigms (gavage versus diet) on pharmacokinetics and pharmacodynamics. *Chemico-Biological Interactions,* 164**,** 68-75.

KARASOV, W. 2012. Terrestrial Vertebrates. *Metabolic Ecology.* John Wiley & Sons, Ltd.

KEARNEY, M. R., DOMINGOS, T. & NISBET, R. 2015. Dynamic Energy Budget Theory: An Efficient and General Theory for Ecology. *Bioscience,* 65**,** 341-341.

KEARNEY, M. R. & WHITE, C. R. 2012. Testing Metabolic Theories. *American Naturalist,* 180**,** 546-565.

KERKHOFF, A. J. 2012. Modeling Metazoan Growth and Ontogeny. *Metabolic Ecology.* John Wiley & Sons, Ltd.

KINGMA, B. R. M., FRIJNS, A. J. H., SCHELLEN, L. & VAN MARKEN LICHTENBELT, W. D. 2014. Beyond the classic thermoneutral zone: Including thermal comfort. *Temperature: Multidisciplinary Biomedical Journal,* 1**,** 142-149.

KLEIBER, M. 1932. Body size and metabolism. *Hilgardia,* 6**,** 315-353.

KOOIJMAN, S. 2000. Dynamic Energy and Mass Budgets in Biological Systems. Cambridge: Cambridge University Press.

KOOIJMAN, S. 2018. *AmP Dosidicus gigas, version 2018/12/15* [Online]. http://www.bio.vu.nl/thb/deb/deblab/add_my_pet. Available: http://www.bio.vu.nl/thb/deb/deblab/add_my_pet/entries_web/Dosidicus_gigas/Dosidicus_gigas_res.html [Accessed].

KOOIJMAN, S. 2020. *AmP Metridium senile, version 2020/02/01* [Online]. http://www.bio.vu.nl/thb/deb/deblab/add_my_pet. Available: http://www.bio.vu.nl/thb/deb/deblab/add_my_pet/entries_web/Metridium_senile/Metridium_senile_res.html [Accessed].

KOOIJMAN, S. & BEDAUX, J. J. M. 1996a. *The Analysis of Aquatic Toxicity Data*, VU University Press.

KOOIJMAN, S., LIKA, D., AUGUSTINE, S., KEARNEY, M. & MARN, N. 2019. *Add-my-Pet Species List* [Online]. Available: https://www.bio.vu.nl/thb/deb/deblab/add_my_pet/species_list.html [Accessed June 16 2019].

KOOIJMAN, S., SOUSA, T., PECQUERIE, L., VAN DER MEER, J. & JAGER, T. 2008. From food-dependent statistics to metabolic parameters, a practical guide to the use of dynamic energy budget theory. *Biological Reviews,* 83**,** 533-552.

KOOIJMAN, S. A. L. M. 2015. *AmP Rattus norvegicus version 2015/11/06* [Online]. https://www.bio.vu.nl/thb/deb/deblab/add_my_pet/entries_web/Rattus_norvegicus/Rattus_norvegicus_res.html. Available: https://www.bio.vu.nl/thb/deb/deblab/add_my_pet/entries_web/Rattus_norvegicus/Rattus_norvegicus_res.html [Accessed 19/09/2019 2019].

KOOIJMAN, S. A. L. M. & BEDAUX, J. J. M. 1996b. Analysis of toxicity tests on Daphnia survival and reproduction. *Water Research,* 30**,** 1711-1723.

KOZLOWSKI, J. & KONARZEWSKI, M. 2004. Is West, Brown and Enquist's model of allometric scaling mathematically correct and biologically relevant? *Functional Ecology,* 18**,** 283-289.

KREUTZ, C., RAUE, A., KASCHEK, D. & TIMMER, J. 2013. Profile likelihood in systems biology. *Febs Journal,* 280**,** 2564-2571.

LAAKSONEN, K. S., NEVALAINEN, T. O., HAASIO, K., KASANEN, I. H. E., NIEMINEN, P. A. & VOIPIO, H. M. 2013. Food and water intake, growth, and adiposity of Sprague-Dawley rats with diet board for 24 months. *Laboratory Animals,* 47**,** 245-256.

LE, N. M. P. & BROWN, J. W. 2008. Characterization of the Thermoneutral Zone of the Laboratory Rat. *Faseb Journal,* 22.

LEONHARDT, M. & LANGHANS, W. 2002. Hydroxycitrate has long-term effects on feeding behavior, body weight regain and metabolism after body weight loss in male rats. *Journal of Nutrition,* 132**,** 1977-1982.

LESKO, L. J. 1979. Dose-dependent elimination kinetics of theophylline. *Clinical Pharmacokinetics,* 4**,** 449-459.

LI, H., ZHANG, M., VERVOORT, J., RIETJENS, I. M. C. M., VAN RAVENZWAAY, B. & LOUISSE, J. 2017. Use of physiologically based kinetic modeling-facilitated reverse dosimetry of in vitro toxicity data for prediction of in vivo developmental toxicity of tebuconazole in rats. *Toxicology Letters,* 266**,** 85-93.

LIKA, K., KEARNEY, M. R., FREITAS, V., VAN DER VEER, H. W., VAN DER MEER, J., WIJSMAN, J. W. M., PECQUERIE, L. & KOOIJMAN, S. 2011. The "covariation method" for estimating the parameters of the standard Dynamic Energy Budget model I: Philosophy and approach. *Journal of Sea Research,* 66**,** 270-277.

LIN, Y., HUTH, F., BERGER, U. & GRIMM, V. 2014. The role of belowground competition and plastic biomass allocation in altering plant mass-density relationships. *Oikos,* 123**,** 248-256.

LINDSTEDT, S. L. & SCHAEFFER, P. J. 2002. Use of allometry in predicting anatomical and physiological parameters of mammals. *Laboratory Animals,* 36**,** 1-19.

LIU, C., BEDNARSKA, A. J., SIBLY, R. M., MURFITT, R. C., EDWARDS, P. & THORBEK, P. 2014. Incorporating toxicokinetics into an individual-based model for more realistic pesticide exposure estimates: A case study of the wood mouse. *Ecological Modelling,* 280**,** 30-39.

LOUISSE, J., BOSGRA, S., BLAAUBOER, B. J., RIETJENS, I. & VERWEI, M. 2015. Prediction of in vivo developmental toxicity of all-trans-retinoic acid based on in vitro toxicity data and in silico physiologically based kinetic modeling. *Archives of Toxicology,* 89**,** 1135-1148.

MACDONALD, D. W., MATHEWS, F. & BERDOY, M. 1999. The Behaviour and Ecology of Rattus norvegicus : from Opportunism to Kamikaze Tendencies. *In:* SINGLETON, G. R., HINDS, L. A., LEIRS, H. & ZHANG, Z. (eds.) *Ecologically-based Management of Rodent Pests.* Canberra: Australian Centre for International Agricultural Research.

MAINO, J. L., KEARNEY, M. R., NISBET, R. M. & KOOIJMAN, S. 2014. Reconciling theories for metabolic scaling. *Journal of Animal Ecology,* 83**,** 20-29.

MARQUES, G. M., AUGUSTINE, S., LIKA, K., PECQUERIE, L., DOMINGOS, T. & KOOIJMAN, S. 2018. The AmP project: Comparing species on the basis of dynamic energy budget parameters. *Plos Computational Biology,* 14.

MARQUET, P. A., ALLEN, A. P., BROWN, J. H., DUNNE, J. A., ENQUIST, B. J., GILLOOLY, J. F., GOWATY, P. A., GREEN, J. L., HARTE, J., HUBBELL, S. P., O'DWYER, J., OKIE, J. G., OSTLING, A., RITCHIE, M., STORCH, D. & WEST, G. B. 2014. On Theory in Ecology. *Bioscience,* 64**,** 701-710.

MARQUET, P. A., ALLEN, A. P., BROWN, J. H., DUNNE, J. A., ENQUIST, B. J., GILLOOLY, J. F., GOWATY, P. A., HARTE, J., HUBBELL, S. P., OKIE, J. G., OSTLING, A., RITCHIE, M., STORCH, D. & WEST, G. B. 2015. On the Importance of First Principles in Ecological Theory Development. *Bioscience,* 65**,** 342-343.

MARTIN, B. T., JAGER, T., NISBET, R. M., PREUSS, T. G. & GRIMM, V. 2013. Predicting Population Dynamics from the Properties of Individuals: A Cross-Level Test of Dynamic Energy Budget Theory. *American Naturalist,* 181**,** 506-519.

MARTIN, B. T., ZIMMER, E. I., GRIMM, V. & JAGER, T. 2012. Dynamic Energy Budget theory meets individual-based modelling: a generic and accessible implementation. *Methods in Ecology and Evolution,* 3**,** 445-449.

MARTIN, T., THOMPSON, H., THORBEK, P. & ASHAUER, R. 2019. Toxicokinetic–Toxicodynamic Modeling of the Effects of Pesticides on Growth of Rattus norvegicus. *Chemical Research in Toxicology,* 32**,** 2281-2294.

MATTSON, M. P. 2008. Hormesis defined. *Ageing Research Reviews,* 7**,** 1-7.

MILLER, C. S. 2001. The Compelling Anomaly of Chemical Intolerance. *Annals of the New York Academy of Sciences,* 933**,** 1-23.

MINEAU, P. 2005. A review and analysis of study endpoints relevant to the assessment of "long term" pesticide toxicity in avian and mammalian wildlife. *Ecotoxicology,* 14**,** 775-799.

MINEAU, P. 2011. Barking up the wrong perch: Why we should stop ignoring non dietary routes of pesticide exposure in birds. *Integrated Environmental Assessment and Management,* 7**,** 297-299.

MOSER, V. C., CHANDA, S. M., MORTENSEN, S. R. & PADILLA, S. 1998. Age- and gender-related differences in sensitivity to chlorpyrifos in the rat reflect developmental profiles of esterase activities. *Toxicological Sciences,* 46**,** 211-222.

MOSES, MELANIE E., HOU, C., WOODRUFF, WILLIAM H., WEST, GEOFFREY B., NEKOLA, JEFFERY C., ZUO, W. & BROWN, JAMES H. 2008. Revisiting a Model of Ontogenetic Growth: Estimating Model Parameters from Theory and Data. *The American Naturalist,* 171**,** 632-645.

MURADO, M. A. & PRIETO, M. A. 2013. NOEC and LOEC as merely concessive expedients: Two unambiguous alternatives and some criteria to maximize the efficiency of dose-response experimental designs. *Science of the Total Environment,* 461**,** 576-586.

MURTAUGH, P. A. 2014. In defense of P values. *Ecology,* 95**,** 611-617.

NICHOLS, J. W., MCKIM, J. M., ANDERSEN, M. E., GARGAS, M. L., CLEWELL, H. J. & ERICKSON, R. J. 1990. A physiologically based toxicokinetic model for the uptake and disposition of waterborne organic-chemicals in fish. *Toxicology and Applied Pharmacology,* 106**,** 433-447.

NICHOLS, J. W., MCKIM, J. M., LIEN, G. J., HOFFMAN, A. D. & BERTELSEN, S. L. 1991. Physiologically based toxicokinetic modeling of 3 waterborne chloroethanes in rainbow-trout (*Oncorhynchus-mykiss*). *Toxicology and Applied Pharmacology,* 110**,** 374-389.

NYMAN, A. M., SCHIRMER, K. & ASHAUER, R. 2012. Toxicokinetic-toxicodynamic modelling of survival of Gammarus pulex in multiple pulse exposures to propiconazole: model assumptions, calibration data requirements and predictive power. *Ecotoxicology,* 21**,** 1828-1840.

OECD 1984. *Test No. 204: Fish, Prolonged Toxicity Test: 14-Day Study*.

OECD 1998. Test No. 408: Repeated Dose 90-Day Oral Toxicity Study in Rodents.

OECD 2001. Test No. 416: Two-Generation Reproduction Toxicity.

OECD 2008. Test No. 407: Repeated Dose 28-day Oral Toxicity Study in Rodents.

OECD 2010. Test No. 417: Toxicokinetics.

OECD 2012. *Test No. 305: Bioaccumulation in Fish: Aqueous and Dietary Exposure*.

OECD 2013. Test No. 210: Fish Early Life Stage Toxicity Test.

OECD 2014. *Fish Toxicity Testing Framework*.

OECD 2018a. Test No. 451: Carcinogenicity Studies.

OECD 2018b. Test No. 452: Chronic Toxicity Studies.

OECD 2019. *Test No. 203: Fish, Acute Toxicity Test*.

OECD. 2020. *OECD guidelines for testing of chemicals: Full list of test guidelines June 2020* [Online]. Available: https://www.oecd.org/chemicalsafety/testing/2020_TG_List_ENG_June_2020.pdf [Accessed 2021].

PAHL, P. J. 1969. Growth Curves for Body Weight of the Laboratory Rat. *Australian Journal of Biological Sciences,* 22 1077-1080.

PAINTER, P. R. 2005. Supply-demand balance in outward-directed networks and Kleiber's law. *Theoretical Biology and Medical Modelling,* 2.

PALM, J. E. 1975. The Laboratory Rat, Rattus norvegicus. *In:* KING, R. C. (ed.) *Handbook of Genetics: Volume 4 Vertebrates of Genetic Interest.* Boston, MA: Springer US.

PAN, J. X. & FANG, K. T. 2002. Maximum Likelihood Estimation. *Growth Curve Models and Statistical Diagnostics.* New York, NY: Springer.

PARACELSUS 1565. The third defenson for writing the new prescriptions. *The Book Paragranum -* Frankfurt: Christian Egenolff.

PELKONEN, O. & TURPEINEN, M. 2007. In vitro-in vivo extrapolation of hepatic clearance: Biological tools, scaling factors, model assumptions and correct concentrations. *Xenobiotica,* 37**,** 1066-1089.

PETER, L. 2015. *EU rejects bid to ban animal testing in European labs* [Online]. BBC News. Available: https://www.bbc.co.uk/news/world-europe-33015460 [Accessed 02/12/2020 2020].

PHILLIPS MCDOUGALL 2016. Agrochemical Research and development: A Consultancy Study for Crop Life America and the European Crop Protection Association.

PIETRZYKOWSKI, A. Z. & TREISTMAN, S. N. 2008. The molecular basis of tolerance. *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism,* 31**,** 298-309.

PIRMOHAMED, M., KITTERINGHAM, N. R. & KEVIN PARK, B. 1994. The Role of Active Metabolites in Drug Toxicity. *Drug Safety,* 11**,** 114-144.

POOLE, S. & STEPHENSON, J. D. 1977. Body temperature regulation and thermoneutrality in rats. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences,* 62**,** 143-149.

PPDB. 2019a. *Pesticide Properties Database - Acibenzolar-S-Methyl* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/13.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/13.htm [Accessed 19/09/2019 2019].

PPDB. 2019b. *Pesticide Properties Database - Azoxystrobin* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/54.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/54.htm [Accessed 19/09/2019 2019].

PPDB. 2019c. *Pesticide Properties Database - Fenpropidin* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/307.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/307.htm [Accessed 19/09/2019 2019].

PPDB. 2019d. *Pesticide Properties Database - Fludioxonil* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/330.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/330.htm [Accessed 19/09/2019 2019].

PPDB. 2019e. *Pesticide Properties Database - Mandipropamid* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/425.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/425.htm [Accessed 19/09/2019 2019].

PPDB. 2019f. *Pesticide Properties Database - Prosulfuron* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/558.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/558.htm [Accessed 19/09/2019 2019].

PPDB. 2019g. *Pesticide Properties Database - Thiamethoxam* [Online]. https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/631.htm. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/631.htm [Accessed 19/09/2019 2019].

PPDB. 2021. *Pesticide Properties Database - Clothianidin* [Online]. http://sitem.herts.ac.uk/aeru/ppdb/en/Reports/171.htm. Available: http://sitem.herts.ac.uk/aeru/ppdb/en/Reports/171.htm [Accessed 2021].

RAIMONDO, S., MINEAU, P. & BARRON, M. G. 2007. Estimation of chemical toxicity to wildlife species using interspecies correlation models. *Environmental Science & Technology,* 41**,** 5888-5894.

RAKEL, K. & GERGS, A. 2018. *AmP Rattus norvegicus Wistar version 2018/06/18* [Online]. https://www.bio.vu.nl/thb/deb/deblab/add_my_pet/entries_web/Rattus_norvegicus_Wistar/Rattus_norvegicus_Wistar_res.html. Available: https://www.bio.vu.nl/thb/deb/deblab/add_my_pet/ [Accessed 17/10/2018].

REED, D. R., DUKE, F. F., ELLIS, H. K., ROSAZZA, M. R., LAWLER, M. P., ALARCON, L. K. & TORDOFF, M. G. 2011. Body fat distribution and organ weights of 14 common strains and a 22-strain consomic panel of rats. *Physiology & Behavior,* 103**,** 523-529.

REINOSO, R. F., TELFER, B. A. & ROWLAND, M. 1997. Tissue water content in rats measured by desiccation. *Journal of Pharmacological and Toxicological Methods,* 38**,** 87-92.

RICHARDS, F. J. 1959. A flexible growth function for empirical use. *Journal of Experimental Botany,* 10**,** 290-300.

RICKER, W. E. 1979. Growth Rates and Models. *In:* HOAR, W. S., RANDALL, D. J. & CONTE, F. P. (eds.) *Fish Physiology : Bioenergetics and growth.* Academic Press.

ROACH, H. I., MEHTA, G., OREFFO, R. O. C., CLARKE, N. M. P. & COOPER, C. 2003. Temporal analysis of rat growth plates: Cessation of growth with age despite presence of a physis. *Journal of Histochemistry & Cytochemistry,* 51**,** 373-383.

ROJAS, J. M., BOLZE, F., THORUP, I., NOWAK, J., DALSGAARD, C. M., SKYDSGAARD, M., BERTHELSEN, L. O., KEANE, K. A., SOEBORG, H., SJOGREN, I., JENSEN, J. T., FELS, J. J., OFFENBERG, H. K., ANDERSEN, L. W. & DALGAARD, M. 2018. The Effect of Diet-induced Obesity on Toxicological Parameters in the Polygenic Sprague-Dawley Rat Model. *Toxicologic Pathology,* 46**,** 777-798.

ROVIDA, C. & HARTUNG, T. 2009. Re-Evaluation of Animal Numbers and Costs for In Vivo Tests to Accomplish REACH Legislation Requirements for Chemicals - a Report by the Transatlantic Think Tank for Toxicology (t(4)). *Altex-Alternativen Zu Tierexperimenten,* 26**,** 187-208.

RUHLEN, R. L., TAYLOR, J. A., MAO, J., KIRKPATRICK, J., WELSHONS, W. V. & VOM SAAL, F. S. 2011. Choice of animal feed can alter fetal steroid levels and mask developmental effects of endocrine disrupting chemicals. *Journal of Developmental Origins of Health and Disease,* 2**,** 36-48.

SADOUL, B., AUGUSTINE, S., ZIMMER, E., BÉGOUT, M. L. & VIJAYAN, M. M. 2018. Prediction of long-term variation in offspring metabolism due to BPA in eggs in rainbow trout using the DEB model. *Journal of Sea Research*.

SAKURADA, S., SHIDO, O., SUGIMOTO, N., HIRATSUKA, Y., YODA, T. & KANOSUE, K. 2000. Autonomic and behavioural thermoregulation in starved rats. *Journal of Physiology-London,* 526**,** 417-424.

SANCHEZ-BAYO, F. 2012. Should we forget NOECs? *Integrated Environmental Assessment and Management,* 8**,** 564-565.

SCHIRMER, K. 2006. Proposal to improve vertebrate cell cultures to establish them as substitutes for the regulatory testing of chemicals and effluents using fish. *Toxicology,* 224**,** 163-183.

SERALINI, G. E., CELLIER, D. & DE VENDOMOIS, J. S. 2007. New analysis of a rat feeding study with a genetically modified maize reveals signs of hepatorenal toxicity. *Archives of Environmental Contamination and Toxicology,* 52**,** 596-602.

SEWELL, F., GELLATLY, N., BEAUMONT, M., BURDEN, N., CURRIE, R., DE HAAN, L., HUTCHINSON, T. H., JACOBS, M., MAHONY, C., MALCOMBER, I., MEHTA, J., WHALE, G. & KIMBER, I. 2018. The future trajectory of adverse outcome pathways: a commentary. *Archives of Toxicology,* 92**,** 1657-1661.

SHAW-ALLEN, P. & SUTER, G. W. 2016. *Species Sensitivity Distributions (SSDs)* [Online]. United States Environmental Protection Agency. Available: https://www3.epa.gov/caddis/da_advanced_2.html [Accessed 20/04/2017 2017].

SHERBORNE, N., GALIC, N. & ASHAUER, R. 2020. Sublethal effect modelling for environmental risk assessment of chemicals: Problem definition, model variants, application and challenges. *Science of The Total Environment,* 745**,** 141027.

SIBLY, R. & CALOW, P. 1986a. *Physiological ecology of animals : an evolutionary approach*, Blackwell Scientific Publications.

SIBLY, R. & CALOW, P. 1986b. *Physiological ecology of animals : an evolutionary approach / R. M. Sibly, P. Calow*, Blackwell Scientific Publications.

SIBLY, R. M. 2012. Life History. *Metabolic Ecology.* John Wiley & Sons, Ltd.

SIBLY, R. M., AKCAKAYA, H. R., TOPPING, C. J. & O'CONNOR, R. J. 2005. Population-level assessment of risks of pesticides to birds and mammals in the UK. *Ecotoxicology,* 14**,** 863-876.

SIBLY, R. M., GRIMM, V., MARTIN, B. T., JOHNSTON, A. S. A., KULAKOWSKA, K., TOPPING, C. J., CALOW, P., NABE-NIELSEN, J., THORBEK, P. & DEANGELIS, D. L. 2013. Representing the acquisition and use of energy by individuals in agent-based models of animal populations. *Methods in Ecology and Evolution,* 4**,** 151-161.

SILBERBAUER, C. J., SURINA-BAUMGARTNER, D. M., ARNOLD, M. & LANGHANS, W. 2000. Prandial lactate infusion inhibits spontaneous feeding in rats. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology,* 278**,** R646-R653.

SJOVALL, J., ALVAN, G. & WESTERLUND, D. 1985. Dose-dependent absorption of amoxycillin and bacampicillin. *Clinical Pharmacology & Therapeutics,* 38**,** 241-250.

SMITH, F. A. 1995. Scaling of digestive efficiency with body-mass in neotoma. *Functional Ecology,* 9**,** 299-305.

SMITH, S. M., WUNDER, M. B., NORRIS, D. A. & SHELLMAN, Y. G. 2011. A simple protocol for using a LDH-based cytotoxicity assay to assess the effects of death and growth inhibition at the same time. *PloS one,* 6**,** e26908-e26908.

STADNICKA-MICHALAK, J., SCHIRMER, K. & ASHAUER, R. 2015. Toxicology across scales: Cell population growth in vitro predicts reduced fish growth. *Science Advances,* 1.

STADNICKA-MICHALAK, J., TANNEBERGER, K., SCHIRMER, K. & ASHAUER, R. 2014. Measured and Modeled Toxicokinetics in Cultured Fish Cells and Application to In Vitro - In Vivo Toxicity Extrapolation. *Plos One,* 9.

STADNICKA, J., SCHIRMER, K. & ASHAUER, R. 2012. Predicting Concentrations of Organic Chemicals in Fish by Using Toxicokinetic Models. *Environmental Science & Technology,* 46**,** 3273-3280.

TACUTU, R., THORNTON, D., JOHNSON, E., BUDOVSKY, A., BARARDO, D., CRAIG, T., DIANA, E., LEHMANN, G., TOREN, D., WANG, J. W., FRAIFELD, V. E. & DE MAGALHAES, J. P. 2018. Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Research,* 46**,** D1083-D1090.

TEKUS, E., MIKO, A., FUREDI, N., ROSTAS, I., TENK, J., KISS, T., SZITTER, I., BALASKO, M., HELYES, Z., WILHELM, M. & PETERVARI, E. 2018. Body fat of rats of different age groups and nutritional states: assessment by micro-CT and skinfold thickness. *Journal of Applied Physiology,* 124**,** 268-275.

THE DANISH ENVIRONMENTAL PROTECTION AGENCY 2014. Pesticide Risk Assessment For Birds And Mammals. *In:* AAGAARD, A. (ed.) 1.1 ed.: The Danish Environmental Protection Agency.

THOMPSON, H. M. 2007. Addressing issues in pesticide risk assessment for birds and mammals. *Outlooks on Pest Management,* 18**,** 23-27.

TRAUTMANN, N. 2005. *The Dose Makes the Poison - Or Does It?* [Online]. www.actionbioscience.org. Available: http://www.actionbioscience.org/environment/trautmann.html [Accessed 18/04/2017 2017].

TSAIOUN, K., BLAAUBOER, B. J. & HARTUNG, T. 2016. Evidence-Based Absorption, Distribution, Metabolism, Excretion (ADME) and its Interplay with Alternative Toxicity Methods. *Altex-Alternatives to Animal Experimentation,* 33**,** 343-358.

VAN DER MEER, J. 2006a. An introduction to Dynamic Energy Budget (DEB) models with special emphasis on parameter estimation. *Journal of Sea Research,* 56**,** 85-102.

VAN DER MEER, J. 2006b. Metabolic theories in ecology. *Trends in Ecology & Evolution,* 21**,** 136-140.

VAN LEEUWEN, C. J. & VERMEIRE, T. G. 2007. *Risk Assessment of Chemicals: An Introduction*, Springer Netherlands.

VANDENBERG, L. N., WELSHONS, W. V., VOM SAAL, F. S., TOUTAIN, P.-L. & MYERS, J. P. 2014. Should oral gavage be abandoned in toxicity testing of endocrine disruptors? *Environmental Health,* 13**,** 46.

VELOSO, C. & BOZINOVIC, F. 1993. Dietary and digestive constraints on basal energy-metabolism in a small herbivorous rodent. *Ecology,* 74**,** 2003-2010.

WANG, M., GUCKLAND, A., MURFITT, R., EBELING, M., SPRENGER, D., FOUDOULAKIS, M. & KOUTSAFTIS, A. 2019. Relationship between magnitude of body weight effects and exposure duration in mammalian toxicology studies and implications for ecotoxicological risk assessment. *Environmental Sciences Europe,* 31**,** 38.

WEST, G. B., BROWN, J. H. & ENQUIST, B. J. 1997. A general model for the origin of allometric scaling laws in biology. *Science,* 276**,** 122-126.

WEST, G. B., BROWN, J. H. & ENQUIST, B. J. 2001. A general model for ontogenetic growth. *Nature,* 413**,** 628-631.

WHITE, C. R. & KEARNEY, M. R. 2014. Metabolic Scaling in Animals: Methods, Empirical Results, and Theoretical Explanations. *Comprehensive Physiology,* 4**,** 231-256.

WHITE, C. R. & SEYMOUR, R. S. 2003. Mammalian basal metabolic rate is proportional to body mass(2/3). *Proceedings of the National Academy of Sciences of the United States of America,* 100**,** 4046-4049.

WILLETT, C. The Use of Adverse Outcome Pathways (AOPs) to Support Chemical Safety Decisions Within the Context of Integrated Approaches to Testing and Assessment (IATA). *In:* KOJIMA, H., SEIDLE, T. & SPIELMANN, H., eds. Alternatives to Animal Testing, 2019// 2019 Singapore. Springer Singapore, 83-90.

WILMES, A., LIMONCIEL, A., ASCHAUER, L., MOENKS, K., BIELOW, C., LEONARD, M. O., HAMON, J., CARPI, D., RUZEK, S., HANDLER, A., SCHMAL, O., HERRGEN, K., BELLWON, P., BUREK, C., TRUISI, G. L., HEWITT, P., DI CONSIGLIO, E., TESTAI, E., BLAAUBOER, B. J., GUILLOU, C., HUBER, C. G., LUKAS, A., PFALLER, W., MUELLER, S. O., BOIS, F. Y., DEKANT, W. & JENNINGS, P. 2013. Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. *Journal of Proteomics,* 79**,** 180-194.

YAFFE, D. 1968. Retention of differentiation potentialities during prolonged cultivation of myogenic cells. *Proceedings of the National Academy of Sciences of the United States of America,* 61**,** 477-483.

ZHANG, J., NUEBEL, E., WISIDAGAMA, D. R. R., SETOGUCHI, K., HONG, J. S., VAN HORN, C. M., IMAM, S. S., VERGNES, L., MALONE, C. S., KOEHLER, C. M. & TEITELL, M. A. 2012. Measuring energy metabolism in cultured cells, including human pluripotent stem cells and differentiated cells. *Nature protocols,* 7**,** 1068-1085.

ZHUANG, X. & LU, C. 2016. PBPK modeling and simulation in drug research and development. *Acta Pharmaceutica Sinica B,* 6**,** 430-440.

ZIMMER, E. I., PREUSS, T. G., NORMAN, S., MINTEN, B. & DUCROT, V. 2018. Modelling effects of time-variable exposure to the pyrethroid beta-cyfluthrin on rainbow trout early life stages. *Environmental Sciences Europe,* 30.

ZULLINGER, E. M., RICKLEFS, R. E., REDFORD, K. H. & MACE, G. M. 1984. Fitting sigmoidal equations to mammalian growth-curves. *Journal of Mammalogy,* 65**,** 607-636.