QUANTITATION OF THE METABOLIC ACTIVATION OF PHENANTHRENE IN SMOKERS: POTENTIAL USE IN THE ASSESSMENT OF LUNG CANCER SUSCEPTIBILITY

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

JING WANG

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

CHERYL L. ZIMMERMAN, Ph.D., Adviser

May, 2012
Acknowledgements

I would like to thank Dr. Cheryl Zimmerman for her guidance in last 5 years. She is always supportive and encouraging. Working with her is wonderful experience that helped me grow both professionally and personally.

I also want to thank Dr. Stephen Hecht for his support on my thesis project. He is essentially my co-adviser. His expertise on chemistry and cancer research contributed significantly to my thesis work.

I want to thank Dr. Richard Brundage, Dr. Angela Birnbaum, Ghada Ahmed and Akshanth Polepally for their help on the population pharmacokinetic analysis.

I want to thank Dr. William Elmquist and Dr. Carolyn Fairbanks for being on my committee.

I want to thank Dr. Lu Xu who was my manager at Genentech. She trained me how to perform population pharmacokinetic analysis. She is a reliable friend and wonderful mentor. She guided me to the field of biologics. I would like to start my adventure in large molecules in future.

I want to thank my wife Yuan (Ivy) Zhang for editorial assistance.
Abstract

Polycyclic aromatic hydrocarbons (PAHs) in cigarette smoke are among the most likely causes of lung cancer. PAHs require metabolic activation to initiate the carcinogenic process. Phenanthrene (Phe), a non-carcinogenic PAH, was used as a surrogate of benzo[α]pyrene (BaP) and related PAHs to study the metabolic activation of PAHs in smokers. A dose of 10 µg deuterated Phe ([D_{10}]Phe) was administered to 25 healthy smokers in a crossover design, either as an oral solution or by smoking cigarettes containing [D_{10}]Phe. Intensive blood and urine sampling was performed to quantitate the formation of deuterated r-1,t-2,3,c-4-tetrahydroxy-1,2,3,4-tetrahydrophenanthrene ([D_{10}]PheT), a biomarker of the diol epoxide metabolic activation pathway.

In both the oral and smoking arms an average of approximately 6% of the dose was metabolically converted to diol epoxides, with a large inter-subject variability in the formation of [D_{10}]PheT observed. Two diagnostic plots were developed to identify subjects with large systemic exposure and significant lung contribution to metabolic activation, respectively. The combination of the two plots led to the identification of subjects with substantial local exposure. These subjects produced, in one single pass of [D_{10}]Phe through the lung, a [D_{10}]PheT exposure equivalent to the systemic exposure of a typical subject, which may be an indicator of lung cancer susceptibility. Polymorphisms in PAH metabolizing genes of the 25 subjects were also investigated. The integration of phenotyping and genotyping results indicated that GSTM1 null subjects produced approximately 2-fold more [D_{10}]PheT than did GSTM1 positive subjects.

A population pharmacokinetic analysis was also tried as an alternative approach to identify subjects with extensive activation, but the analysis was not completely successful. The preliminary statistical analysis indicated a correlation between renal function and an individual’s capacity to activate PAHs, which may deserve further investigation.
A simulation project based on the pharmacokinetic data of 25 subjects was carried out to simulate the metabolic activation of \([D_{10}]\)Phe in 350 subjects after oral dosing and smoking. The simulation suggested that the collection of a 6-hr urine sample in both treatment arms should be conducted in the future large-scale trials in order to efficiently assess the hypothesized correlation between extensive activation and high lung cancer risk.
Table of Contents

Acknowledgements ................................................................................................................. i
Abstract ..................................................................................................................................... ii
Table of Contents ...................................................................................................................... iv
List of Tables ............................................................................................................................... vi
List of Figures ............................................................................................................................. vii

CHAPTER I  Introduction ............................................................................................................. 1
  I.A  Lung cancer and its prevention .......................................................................................... 1
  I.B  Polycyclic aromatic hydrocarbons .................................................................................... 2
  I.C  Benzo[α]pyrene .................................................................................................................. 3
  I.D  Phenanthrene ..................................................................................................................... 9
  I.E  Genotyping and phenotyping ............................................................................................ 15
  I.F  Research objectives .......................................................................................................... 17

CHAPTER II  Phenanthrene Metabolism in Smokers: Use of a Two-step Diagnostic Plot Approach to Identify Subjects with Extensive Metabolic Activation ........................................... 18
  II.A  Introduction .................................................................................................................... 18
  II.B  Methods .......................................................................................................................... 19
  II.C  Results ............................................................................................................................ 28
  II.D  Discussion ...................................................................................................................... 39

CHAPTER III  Population Pharmacokinetic Analysis of the Metabolic Activation of Phenanthrene in Humans ......................................................................................................................................... 44
  III.A  Introduction .................................................................................................................. 44
  III.B  Methods ....................................................................................................................... 47
  III.C  Results .......................................................................................................................... 50
CHAPTER IV  The Optimization of the Study Design by Simulation ........... 60
IV.A  Introduction ............................................................................. 60
IV.B  Methods .................................................................................... 61
IV.C  Results ...................................................................................... 65
IV.D  Discussion ................................................................................ 69
IV.E  Conclusion ................................................................................ 73
CHAPTER V  Conclusion ..................................................................... 74
Bibliography ..................................................................................... 76
  References Chapter I ....................................................................... 76
  References Chapter II ...................................................................... 89
  References Chapter III .................................................................... 92
  References Chapter IV ...................................................................... 93
Appendix II-A .................................................................................... 94
Appendix III-A .................................................................................. 101
Appendix IV-A .................................................................................. 106
List of Tables

Table II-1. Half-life, clearance and AUC of [D_{10}]PheT after oral dosing and smoking.. 31
Table II-2. $A_{act}$ of oral and smoking arm calculated by two methods.......................... 32
Table II-3. Relative bioavailability by two methods ($F_{AUC}$ and $F_{Xu}$) ......................... 34
Table II-4. Significant lung contribution to metabolic activation in subject 1, 6 and 22 35
Table II-5. Effects of polymorphisms on individual’s capacity to activate PAHs .......... 36
Table III-1. Demographic and PK data of 25 subjects...................................................... 56
Table IV-1. The evaluation of $Xu_{smk}$, $Xu_{oral}$, and $AUC_{PheT}$ in scenario 1 .................. 66
Table IV-2. The evaluation of $Xu_{smk}$, $Xu_{oral}$, and $Xu_{lung}$ in scenario 2 ....................... 66
List of Figures

Figure I-1. Structures of BaP and Phe (adapted from reference ).......................... 4
Figure I-2. Ring opening of the bay region epoxide (BPDE) to form a carbonium ion at C-10 position (adapted from reference 28) ......................................................... 4
Figure I-3. The diol epoxide metabolic activation of BaP (solid arrows) and a subset of other metabolic pathways (dotted arrows) (adapted from reference 29) .... 6
Figure I-4. The diol epoxide metabolic activation of Phe (solid arrows) and a subset of other metabolic pathways (dotted arrows) (adapted from reference 29) .... 12
Figure I-5. The diol epoxide metabolic activation of BaP and Phe (adapted from reference 29) ............................................................................................................. 13
Figure I-6. Tetraols from BaP and Phe bay region diol epoxide pathways (3 and 9) and reverse diol epoxide pathways (6 and 12) (adapted from reference 72) .... 14
Figure II-1. The PK view of metabolic activation of [D_{10}]Phe after administration (detailed description in Appendix II-A) ................................................................. 24
Figure II-2. Diagnostic plot I (top) and II (bottom) ................................................... 26
Figure II-3. Comparison of metabolic activation of [D_{10}]Phe in the oral and smoking arm ................................................................................................................. 33
Figure II-4. Comparison of GSTM1 negative (-) and positive (+) subjects after smoking (a and b) and oral dosing (c and d) (dotted line: median, * p < 0.05, ** p < 0.01) .................................................................................................................. 37
Figure II-5. Correlation between genotyping and phenotyping (triangle: GSTM1 positive; circle: GSTM1 negative) ............................................................................. 38
Figure III-1. Two sources of the formation of [D_{10}]PheT ........................................ 45
Figure III-2. Simulated plasma concentration-time profile of [D_{10}]PheT ............... 46
Figure III-3. Structural model of Pop PK analysis ...................................................... 48
Figure III-4. Plasma [D_{10}]PheT concentration-time profile of 25 subjects after oral dosing ................................................................................................................. 51
Figure III-5. Plasma $[D_{10}]$PheT concentration-time profile of 25 subjects after smoking ................................................................. 52
Figure III-6. Cumulative amount of urinary $[D_{10}]$PheT-time profile of 25 subjects after oral dosing ........................................................................................................ 53
Figure III-7. Cumulative amount of urinary $[D_{10}]$PheT-time profile of 25 subjects after smoking ........................................................................................................ 54
Figure III-8. Correlation between creatinine clearance and $Xu_{smk}$ .......................................................... 55
Figure IV-1. The PK model used in the simulation ......................................................................................... 63
Figure IV-2. The plasma concentration-time profile of $[D_{10}]$PheT of a representative subject from the pilot study (top) and a simulated subject (bottom).......... 67
Figure IV-3. Cumulative amount of $[D_{10}]$PheT in urine-time profile of a representative subject from the pilot study (top) and a simulated subject (bottom)........ 68
Figure IV-4. An example of the effect of BSV of $k_2$ on $Xu_{smk}$................................................................. 70
CHAPTER I    Introduction

I.A    Lung cancer and its prevention

In the United States, lung cancer causes more deaths than any other cancer in both men and women\(^1\). An estimated 160,340 deaths are expected to occur in 2012, accounting for about 28\% of all cancer deaths\(^1\). In the world, the total number of cases that occur annually is estimated to be 1.2 million\(^2\). Unfortunately, the treatment of lung cancer has not been particularly effective as the five-year survival rates have only increased from 12\% in 1975 to 16\% in 2008\(^3\).

While not much progress has been achieved in the treatment of lung cancer, prevention may play an essential role because lung cancer is among the most preventable diseases\(^2\). The major cause of this disease is tobacco smoking, primarily of cigarettes\(^4\). Approximately 90\% of the lung cancer cases were attributed to tobacco smoking\(^2,4\). There are about 47 million cigarette smokers in U.S. and about one billion all over the world. In addition, non-smokers may become victims of involuntary smoking\(^2,5\). Prevention such as smoking cessation could be an effective way to reduce lung cancer risk in both smokers and non-smokers.

Not all smokers are equally susceptible to lung cancer. It is well known that among populations with similar exposure to environmental carcinogens, cancer develops only in a fraction of susceptible individuals\(^6\). In addition, it has been reported that only a small fraction of life-long smokers die from lung cancer, 24.1\% of male smokers and 11.1\% of female smokers\(^2,7,8\), which indicates the existence of subpopulations susceptible to lung cancer.

The identification of susceptible subpopulations is the key to effective prevention. The outcome of smoking cessation depends on the intensity of intervention. The rate of successful smoking cessation at 1 year is 3-5\% when smokers simply try to stop, 7-16\%
when behavioral intervention is provided, and up to 24% when smokers receive both pharmacological treatment and behavioral support. Since intervention resources are often limited, the identification of susceptible individuals could allow concentrating the intensive intervention on those selected individuals. This would improve the outcome of smoking cessation and lead to effective prevention in the more susceptible population.

However, how to identify lung cancer susceptible subjects remains unclear. Based on the concept of multistage carcinogenesis, the end effect depends on critical factors such as the bioavailability of the cigarette carcinogens, the metabolic activation and detoxification of carcinogens, and the relative balance of DNA damage and repair. It is believed that lung cancer susceptibility factors include age, smoking history, an individual’s capacity to activate tobacco carcinogens, an individual’s DNA repair capacity, and pre-existing health impairment (asthma and chronic bronchitis). Body mass index, socioeconomic status and family history of lung cancer are also considered as possible risk factors of lung cancer. In addition, some researchers also argue that ethnicity and gender might have effects on susceptibility.

A few lung cancer risk prediction models based on the retrospective analysis of large datasets, such as the Bach model, the Spitz model, and the Liverpool Lung Project model, have been reported. The use of chest x-ray and computed tomography (CT) to screen high risk individuals have also been investigated. But the clinical utility of these approaches has not been proven to be significant.

The current study focuses on one critical event of carcinogenesis: metabolic activation of cigarette carcinogens. Unlike previous studies that used retrospective approaches to identify susceptible individuals, the current study uses a unique prospective approach to investigate the correlation between extensive metabolic activation of cigarette carcinogens and high lung cancer risk.

I.B Polycyclic aromatic hydrocarbons
Polycyclic aromatic hydrocarbons (PAHs) are a major aetiological factor in lung cancer. PAHs are formed during the incomplete combustion of tobacco and are considered to be
strong locally-acting lung carcinogens. The amount of PAHs in the mainstream smoke is about 100-200 ng per cigarette\textsuperscript{2,19}. PAHs are usually very lipophilic and could theoretically accumulate in tissues and reach toxic concentrations\textsuperscript{20}. Thanks to evolution, humans have developed the capacity to detoxify foreign chemicals, including carcinogens, by converting them to less toxic metabolites. Enzymes involved in the detoxification of PAHs include cytochrome P450 1A1 (CYP1A1), the glutathione S-transferases (GSTs), UDP-glucuronosyltransferases (UGTs) and sulfotransferases (SULTs)\textsuperscript{21}.

However, at the cost of solving the acute problem (lethal accumulation of xenobiotics), a small portion of PAHs is transformed to electrophilic intermediates which do not cause lethal hazards in the short term, but may initiate carcinogenesis in the long term. These electrophilic intermediates can attack the nucleophilic macromolecules of the cell, causing gene mutations which can eventually lead to the development of tumors\textsuperscript{5}. The conversion of PAHs to active intermediates is called metabolic activation or bioactivation. Enzymes involved in metabolic activation include CYP1A1, cytochrome P450 1A2 (CYP1A2), 1B1 (CYP1B1) and epoxide hydrolase (EH)\textsuperscript{21}.

Even though only a small fraction of the PAH is activated in humans, it is believed to be responsible for the initiation of carcinogenesis\textsuperscript{36}. Considering the role of bioactivation in lung carcinogenesis, it is reasonable to assume that the more extensive the bioactivation, the larger exposure to active intermediates and the higher susceptibility to lung cancer. Therefore, the major objective of the current project is to quantitate an individual’s capacity to activate PAHs, and identify individuals with extensive bioactivation.

**I.C Benzo[a]pyrene**

The most extensively studied PAH molecule is benzo[a]pyrene (BaP) (Figure I-1), a prototypical PAH which has been classified as carcinogenic to humans\textsuperscript{22}. The amount of BaP is estimated to be 10-25 ng per cigarette\textsuperscript{19,23,24}. The average daily intake of BaP by the general U.S. population is approximately 2.2 μg/day\textsuperscript{23,25}. BaP is a focus of the current study in order to investigate PAH metabolism and lung cancer susceptibility because of its strong carcinogenic effects and its well-characterized mechanism of carcinogenesis.
Like other PAHs, BaP undergoes competing metabolic pathways: activation and detoxification. Metabolic activation of BaP is the prerequisite for the initiation of its carcinogenic effects. One activation center of BaP is at the C-10 position of the bay region (Figure I-1). The formation of an epoxide at C-10 position is closely associated with carcinogenesis. The steric hindrance of the bay region favors oxidation\(^2\). Molecular orbital calculations indicated that epoxides formed at the bay region undergo ring opening more readily than non-bay region epoxides, thereby producing a carbonium ion (Figure I-2) or related species that are highly susceptible to nucleophilic attack, such as that by DNA\(^{23,27,28}\).

**Figure I-1.** Structures of BaP and Phe (adapted from reference 26)

**Figure I-2.** Ring opening of the bay region epoxide (BPDE) to form a carbonium ion at C-10 position (adapted from reference 28)
Benzo[a]pyrene 7,8-diol-9,10-epoxide (BPDE) is a major carcinogenic metabolite of BaP with an epoxide formed at C-10 position of the bay region. A widely accepted scheme for BaP bioactivation is that BPDE is produced during the conversion of BaP to \(r\)-7,\(t\)-8,9,\(c\)-10-tetrahydroxy-7,8,9,10-tetrahydrobenzo[a]pyrene (BaPT) by the collaboration of P450s and EH (Figure I-3) \(^{29}\). The epoxide of BPDE readily undergoes ring opening to form a reactive intermediate with a carbocation (Figure I-2). BPDE binds to DNA through the covalent linkage of carbon of the epoxide (C-10) and the N-2 position of guanine \(^{23,30}\). If the BPDE-DNA adducts persist during DNA replication, then miscoding at crucial genes such as \(P53\) may occur \(^{31,32}\), thereby leading to the loss of normal cellular growth control and eventual development of lung cancer \(^{19}\). Only a small fraction of BPDE binds to DNA and the majority reacts with water to yield the end product BaPT \(^{33}\). After its formation, BaPT is further conjugated to sulfate and glucuronic acid catalyzed by the phase II enzymes SULTs and UGTs \(^{34,36}\).

CYP1A1, 1A2, 1B1 and EH are involved in the activation of BaP to produce BPDE. The activation is initiated by P450s to form BaP-7,8-epoxide followed by the hydrolysis catalyzed by EH to form BP-7,8-diol (Figure I-3) \(^{35-38}\). It has been shown that recombinant human CYP1B1 has 10-fold higher activity than CYP1A1 in catalyzing the conversion of BaP to BaP-7,8-diol \(^{39,40}\). But CYP1A1 and 1B1 catalyze the biotransformation from BP-7,8-diol to the electrophilic intermediate BPDE at similar rates \(^{41,42}\). CYP1A2 also catalyzes the activation of BaP-7,8-diol to reactive intermediates but at much slower rates than CYP1A1 and CYP1B1 \(^{21,43}\). In addition, CYP1B1 is expressed more abundantly than CYP1A1 in most extrahepatic organs except pancreas, small intestine and colon \(^{21,44,45}\). Taken together, it is likely that CYP1B1 may play a more important role in the lung activation of BaP than other P450s.

In addition to the diol epoxide activation pathway, BaP can also be converted to reactive benzo[a]pyrene-7,8-dione by aldo-keto reductase (AKR). Benzo[a]pyrene-7,8-dione is able to interact with DNA and form stable DNA adducts \(^{21,35}\). Another possible activation pathway involves the conversion of BaP to a radical cation by peroxidase but the contribution of this pathway to lung carcinogenesis has not been well established yet \(^{21}\).
Figure I-3. The diol epoxide metabolic activation of BaP (solid arrows) and a subset of other metabolic pathways (dotted arrows) (adapted from reference 29)
Competing against the metabolic activation of BaP are a number of detoxification pathways functioning as protective mechanisms. These detoxification pathways transform BaP to metabolites less toxic than BPDE and reduce the efficiency of the activation pathway. Enzymes involved in detoxification include the P450s, EH, GSTs, UGTs and SULTs. It appears that each step of the diol epoxide metabolic activation has competing detoxification pathways (Figure I-3). The first step of the diol epoxide metabolic activation is the conversion of BaP to BaP-7,8-epoxide. Competing against this step is the P450-catalyzed formation of BaP phenols followed by UGT-catalyzed glucuronidation. 3-HOBaP is the major phenol formed during detoxification and has been used as an indicator of the detoxification pathway. The hydrolysis of BaP-7,8-epoxide to BaP-7,8-diol is the second step of activation. But BaP-7,8-epoxide is likely to be converted to glutathione conjugates catalyzed by GSTs. The third step of activation is the oxidation of BaP-7,8-diol to BPDE. Competing against this step is the glucuronidation of BaP-7,8-diol catalyzed by UGTs. Even after the formation of carcinogenic intermediate BPDE, there is still a detoxification pathway that uses GSH to attack BPDE forming glutathione conjugates.

It remains unclear which enzyme plays the most important role in the detoxification of BaP. Historically, GSTs were considered as important enzymes in detoxification because GSTM1 and GSTP1 were found to be effective catalysts of PAH bay region diol epoxide (such as BPDE) detoxification. However, recent studies indicated that the role of GSTs in the detoxification of diol epoxides might be overestimated as they favor the conjugation of non-carcinogenic (reverse bay region) diol epoxides rather than carcinogenic (bay region) diol epoxides. But it is possible that the primary protective mechanism of GSTs functions through the prevention of the formation of BPDE by catalyzing the conjugation of its precursors such as BaP-7,8-epoxide. The prevention of the formation of BPDE would be more beneficial than the detoxification of the BPDE by conjugation when it is already activated and can potentially attack DNA. It could be speculated that this prevention mechanism of GSTs (i.e. the conjugation of BaP-7,8-epoxide with
glutathione catalyzed by GSTs) gives humans a greater survival advantage over the long history of evolution.

CYP1A1 seems to have a more important role in detoxification than activation. It was reported that CYP1A1 null mice died in 30 days after oral BaP administration of 125 mg/kg but wild-type mice did not show any sign of toxicity by the end of the experiments. Another study reported that after oral BaP administration of 125 mg/kg for 18 days CYP1A1 null mice demonstrated significant toxic effects but CYP1A2 and CYP1B1 null mice did not.

The pharmacokinetics (PK) of BaP was characterized in some preclinical studies. A rat study indicated that after oral administration of 1 mmol/kg (252 mg/kg) BaP, approximately 72% of the total dose was recovered in feces and urine: 63% BaP, 9% 3-HOBaP and 1% other metabolites. The majority of the BaP and its metabolites were recovered from feces, which indicated that fecal excretion was the dominant elimination pathway and urinary excretion was the minor one. This is consistent with the observation from other rat studies that only a very small fraction of the total dose was collected in urine. While most of the PK studies used noncompartmental analysis, one rat study indicated that the PK of BaP could be described by a two-compartment model and that nonlinear PK may exist. The terminal half-life was approximately 30 min for the 2 and 6 mg/kg groups but substantially increased to more than 400 min for the 15 mg/kg group.

The bioavailability of BaP in rats has also been investigated. One rat study reported a bioavailability of approximately 10% after oral administration of 4 nmol (1 µg) and i.v. bolus of 1.7 nmol (0.4 µg). The hepatic and pulmonary extraction ratios were calculated to be 0.4 and 0.1, respectively. Another rat study reported a bioavailability of 40% after oral and i.v. administration of 100 mg/kg. The difference in bioavailability between these two studies might be due to the saturation of metabolic enzymes at the high dose. But further investigation is needed.

Despite numerous studies conducted to understand the activation and detoxification of BaP, there is no convincing conclusion regarding the fraction of BaP converted to active
intermediates in preclinical species and humans. In addition, there are no definitive mass balance studies to quantitatively determine the pathways of elimination of BaP and its metabolites. Although various metabolites such as 3-HOBaP, 9-HOBaP, BaP-4,5-diol, BaP-7,8-diol and BaP-9,10-diol have been identified, these metabolites account for less than 20% of the total dose\textsuperscript{47-51}. It has been reported that a large fraction of BaP (45-65%) is excreted in urine and feces intact\textsuperscript{48,62}. Since these reports were from studies that differed in route of administration, dose, animal species, and study design, it is very difficult to estimate the actual contribution of each elimination pathway from literature data.

In addition, great caution must be exercised when extrapolating preclinical data to the activation and detoxification of BaP in humans. A main reason is that metabolic profiles can be quite different at different dosing levels. The majority of preclinical studies, if not all, focused on toxicity and used doses much higher than the daily exposure of humans. In these studies the majority of BaP was excreted intact probably because of the saturation of metabolic enzymes at extremely high doses. However, enzyme saturation might not be the case in humans. The average human daily exposure to BaP (2.2 µg) is only 0.0003% of the dose (10 mg/kg) commonly used in preclinical studies. It is likely that metabolic enzymes are not saturated at this level of exposure and the majority of the BaP is converted to metabolites. Therefore, preclinical data might not be readily translated to the pharmacokinetic (PK) behavior of BaP in humans.

1.D Phenanthrene

Since BaP is generally used as the model compound to study the carcinogenesis of PAHs, the characterization of the metabolic activation of BaP in humans would definitely provide valuable information on the identification of susceptible subpopulations. Unfortunately, a human PK study of BaP is impossible because of its strong carcinogenic effects. Then the question becomes the following: how can an individual's capacity to activate BaP be quantitated without giving BaP to human subjects? One answer to the question is the use of phenanthrene (Phe) as the surrogate of BaP (Figure I-1)\textsuperscript{29}.
Phe is a non-carcinogenic PAH ubiquitous in the environment. The PK of Phe has been investigated in various preclinical studies in laboratory animals. Major metabolic pathways are shown in Figure I-4. One study reported that the majority of the dose was recovered in 5 days after oral administration of 10 mg/kg \( ^{14} \text{C} \)-labelled Phe to rats. Of the dose recovered, 90% of the radioactivity was found in urine and 10% in feces. In urine only 5% of the radioactive content was from the parent compound and 95% of the radioactivity was from metabolites. Primary metabolites included Phe-9,10-diol (45%), Phe-3,4-diol (12%) and Phe-1,2-diol (10%). Most of the metabolites were present as conjugates such as glucuronides, sulfates and cysteinylglycines. The oral administration of 10 mg/kg \( ^{14} \text{C} \)-labelled Phe to guinea pigs showed similar metabolic profiles with minor differences. In a mouse study with an oral dose of 4.5 mg/kg, the half-life and oral clearance were determined to be 0.32 hr and 2.18 L/hr·kg, respectively. It should be noted that the doses used in these preclinical trials were much higher than the average daily human exposure to Phe (48-180 ng) and total PAHs (1.8-16 µg). Therefore, literature data may not be readily used to predict the PK of Phe in humans.

The diol epoxide metabolic activation of Phe mimics that of BaP including enzymes involved and intermediates formed. As shown in Figure I-5, there are three steps in the activation of BaP and Phe: 1) oxidation of the parent compound to produce epoxides; 2) hydrolysis of epoxides to form diols; 3) further oxidation of diols to produce diol epoxides. In the first two steps, P450s coordinate with EH to convert the parent compound to diols. CYP1A1, CYP1B1 and CYP1A2 are the most active P450s in the catalysis of the formation of both BaP-7,8-diol and Phe-1,2-diol according to in vitro studies with recombinant human P450s. In the third step, the oxidation of Phe-1,2-diol produces anti-1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydrophenanthrene (PheDE), which is analogous to the oxidation of BaP-7,8 diol to form BPDE. In other words, PheDE is a biomarker of active intermediate BPDE. However, PheDE is an unstable intermediate like BPDE and direct quantitation of PheDE is very difficult. Since the majority of PheDE readily reacts with water and produces \( r^{-1},t-2,3,c^{-4}\)-tetrahydroxy-1,2,3,4-tetrahydrophenanthrene (PheT), the quantitation of PheT would give a reasonable
estimate of the amount of PheDE formed during activation. Collectively, considering that the diol epoxide metabolic activation of Phe is similar to BaP and that Phe is safe to be used in a clinical study, a human PK study of Phe will shed light on the metabolic activation of BaP and likely other PAHs. Subjects with extensive activation of Phe would be assumed to have high efficiency in the activation of BaP.

It should be noted that there are also differences in the diol epoxide activation of BaP and Phe. As shown in Figure I-6, enantiomers 3 and 9 are the end products of the bay region diol epoxide pathway. Enantiomers 6 and 12 are the end products of the reverse bay region diol epoxide pathway. A recent study of urine samples of smokers indicated that 96% of the PheT was enantiomer 12 and 78% of the BaPT was enantiomer 3. Obviously, reverse bay region diol epoxide formation is dominant in Phe metabolism while bay region diol epoxide formation is dominant in BaP metabolism. Since the reverse bay region diol epoxides are generally believed to be less carcinogenic than the bay region diol epoxides, this may raise the question as to the utility of racemic PheT as the biomarker of the bay region diol epoxide activation. But this issue has been addressed by a recent study which demonstrated strong correlations between the levels of the racemic PheT (enantiomers 9 + 12), enantiomer 12 of PheT and enantiomer 3 of BaPT. Therefore, the quantitation of the racemic PheT without enantiomeric resolution could provide a reasonable estimate of the efficiency of the bay region diol epoxide pathway.
Figure I-4. The diol epoxide metabolic activation of Phe (solid arrows) and a subset of other metabolic pathways (dotted arrows) (adapted from reference 29)
Figure I-5. The diol epoxide metabolic activation of BaP and Phe (adapted from reference 29)
Figure I-6. Tetraols from BaP and Phe bay region diol epoxide pathways (3 and 9) and reverse diol epoxide pathways (6 and 12) (adapted from reference 72)
I.E Genotyping and phenotyping

It is well known that there are large inter-individual variations in the metabolism of PAHs, leading to the hypothesis that individual differences in PAH metabolism influence cancer risk. Numerous epidemiological studies have been done to correlate cancer susceptibility with certain genetic polymorphisms or the combination of several variants of metabolic enzymes. These studies undoubtedly provided valuable insight into the complex mechanism of lung cancer. However, results from epidemiological studies are often mixed and sometimes even contradictory. For example, studies in Japan indicated a strong correlation between CYP1A1*2A (CYP1A1 MspI or m1 mutation) and lung cancer risk, which was confirmed by two studies in the U.S. However, other studies in China, Norway and U.S. did not support the correlation. While the statistical power of one single study might be insufficient to detect the correlation, the integration of data from multiple studies has also been tried. A number of papers using meta-analysis seemed to come to a similar conclusion that CYP1A1*2A is associated with a slight increase in lung cancer risk, but other meta-analysis papers failed to observe a significant correlation.

Another commonly used approach in epidemiological studies is to correlate the combination of several polymorphisms with lung cancer risk. For example, it has been proposed that gene-gene interactions between CYP1A1 and glutathione S-transferase M1 (GSTM1) may result in a greater-than-additive risk for lung cancer because the deletion of GSTM1 has been reported to be associated with the increase of the inducibility of CYP1A1. But results from a number of meta-analysis studies are mixed.

The inconsistency of genotyping studies might be due to a couple of reasons. First, the enzyme-environment interaction was often ignored. It is well established that cigarette smoking as well as other environmental exposure to PAHs can increase the activity of CYP1A1, 1B1 and 1A2 through the interactions of PAHs, and possibly other cigarette smoke constituents, with the aryl hydrocarbon receptor. Lack of consideration of enzyme induction, especially the inter-individual variability in the extent of enzyme induction, may underestimate the variability of enzyme activity in the population. As a
result, sample size in these studies might not be large enough to raise sufficient statistical power and identify the effect of a specific polymorphism. Second, local (lung) exposure to cigarette carcinogens has not yet been quantitated. For example, two individuals with the same genotype may have quite different expression levels of metabolic enzymes at the lung which could lead to a significant difference in local exposure. As a result, the inter-individual variability in local exposure will contribute to the variability of lung cancer susceptibility, thereby requiring a much larger sample size to detect the effect of a specific polymorphism. Last but not least, the metabolism of PAHs is very complex because 1) a great many elimination pathways are involved; 2) each elimination pathway is catalyzed by multiple metabolic enzymes; 3) these enzymes are often involved in both activation and detoxification pathways. Therefore, the genotype of metabolic enzymes may not be readily related to an individual’s capacity to activate PAHs.

The current PK study of Phe is a phenotyping approach that can serve as an alternative to the genotyping approach. In the PK study, deuterated Phe ([D\text{10}]Phe) was used in order to avoid the interference from ubiquitous environmental Phe. The end product of the activation pathway, deuterated PheT ([D\text{10}]PheT), was quantitated to estimate the amount of “active” metabolite, deuterated PheDE ([D\text{10}]PheDE), formed during activation. The actual measurement of the amount of [D\text{10}]PheT can capture the gene-environment interaction because subjects with more extensive induction of activation enzymes would produce larger amounts of [D\text{10}]PheT. In addition, our study could also reveal crucial information regarding local pulmonary exposure. In the study, a dose of 10 µg (53.2 nmol) [D\text{10}]Phe was administered to 25 subjects, either as an oral solution or by smoking cigarettes doped with [D\text{10}]Phe. Since the dose (10 µg) is similar to the daily exposure to PAHs (1.8-16 µg), the metabolic profiles obtained in our study would be more relevant to lung carcinogenesis than those reported in animal studies with extremely high doses. Furthermore, the integration of data from two routes of administration would provide the direct comparison of the first-pass activation in lung and liver, respectively. This information could help to identify subjects with substantial local exposure and presumably higher lung cancer risk.
I.F  Research objectives
Collectively, the hypothesis is that an individual’s capacity to activate PAHs is one
determinant of lung cancer susceptibility. Subjects with extensive metabolic activation
are at a higher lung cancer risk than other subjects. In order to test the hypothesis, two
research projects have been proposed:

1) To develop a PK method to identify subjects with extensive metabolic activation
   and presumably high lung cancer risk
2) To determine the correlation between extensive metabolic activation and high
   lung cancer risk

The first project was the focus of this thesis. The specific aims of the first project were as
follows:

1) To characterize the metabolic activation of Phe after the administration of
   \([D_{10}]\)Phe to 25 subjects via two routes of administration
2) To identify subjects with extensive activation and substantial local exposure to
   activated PAHs
3) To estimate the population mean and between-subject variability of PK
   parameters that describe an individual’s capacity to activate PAHs
4) To investigate the effects of weight, gender, race, renal function, smoking status,
   genotype on an individual’s capacity to activate PAHs
5) To optimize the study protocol of the future large-scale clinical trial

The second project is not within the scope of the current study. It involves the
participation of approximately 350 subjects and the performance of bronchoscopy
screening. Subjects with bronchoepithelial metaplasia or dysplasia would be
hypothesized to have higher levels of \([D_{10}]\)PheT in the urine than subjects without
metaplasia or dysplasia. This project will be initiated in the near future.
CHAPTER II  Phenanthrene Metabolism in Smokers: Use of a Two-step Diagnostic Plot Approach to Identify Subjects with Extensive Metabolic Activation

II.A  Introduction

Cigarette smoking is the leading cause of lung cancer, accounting for 90% of cases of the disease. The polycyclic aromatic hydrocarbons (PAHs) in cigarettes are among the most likely causes of lung cancer in smokers. PAHs usually undergo two competing metabolic pathways: activation and detoxification. The metabolic activation pathway produces intermediates that bind to DNA and form adducts, thereby initiating carcinogenic effects. On the other hand, the detoxification pathway converts PAHs to less toxic metabolites and reduces the efficiency of metabolic activation. Due to the crucial role of metabolic activation in the carcinogenesis of PAHs, it has been hypothesized that subjects with extensive activation have higher lung cancer risk.

Benzo[α]pyrene (BaP) is found in cigarette smoke and is a prototypic and widely studied compound for the investigation of carcinogenesis by PAHs. A major bioactivation pathway of BaP (Figure 1-5) is its conversion to r-7,r-8,9,10-tetrahydroxy-7,8,9,10-tetrahydrobenzo[α]pyrene (BaPT) via the formation of the “bay region diol epoxide” benzo[α]pyrene 7,8-diol-9,10-epoxide (BPDE), which reacts readily with DNA and is mutagenic and carcinogenic. BPDE is thought to be one electrophilic reactive intermediate responsible for BaP mutagenesis and carcinogenesis, and related bay region diol epoxides are considered to be major ultimate carcinogens of a number of other PAHs. As such, the characterization of the diol epoxide metabolic activation of BaP in humans would provide valuable information on the identification of susceptible subjects because theoretically subjects with extensive activation would have higher lung cancer risk.
However, most PAHs like BaP are carcinogenic and cannot be administered to humans and there have not been any detailed pharmacokinetic (PK) studies of PAHs in humans. To address this issue, a novel biomarker approach has been proposed: the use of phenanthrene (Phe) as a surrogate of BaP\textsuperscript{12}. Phe is a non-carcinogenic PAH ubiquitous in the environment and can be safely administered to human subjects because all humans are exposed to Phe. The conversion of Phe to \textit{r}-1,\textit{r}-2,3,\textit{r}-4-tetrahydroxy-1,2,3,4-tetrahydrophenanthrene (PheT) mimics the diol epoxide metabolic activation of BaP including the intermediates formed and enzymes involved (Figure I-5)\textsuperscript{13,14}, although there are some differences\textsuperscript{15}.

In the present study, deuterated phenanthrene ([D\textsubscript{10}]Phe) was administered to 25 subjects and the metabolism of [D\textsubscript{10}]Phe to deuterated \textit{r}-1,\textit{r}-2,3,\textit{r}-4-tetrahydroxy-1,2,3,4-tetrahydrophenanthrene ([D\textsubscript{10}]PheT) was characterized in order to evaluate each individual’s capacity to carry out the diol epoxide metabolism pathway. We used [D\textsubscript{10}]Phe to avoid interference from ubiquitous exposure to environmental Phe. An intermediate formed during the activation of [D\textsubscript{10}]Phe was deuterated \textit{anti}-1,2-dihydroxy-3,4-epoxy-1,2,3,4-tetrahydrophenanthrene ([D\textsubscript{10}]PheDE), a biomarker of the diol epoxide pathway leading to BPDE from BaP (Figure I-5). The primary objective of our study was to quantify the amount of [D\textsubscript{10}]PheDE formed in the metabolism of [D\textsubscript{10}]Phe. [D\textsubscript{10}]PheDE cannot be quantified directly because it rapidly reacts with H\textsubscript{2}O producing [D\textsubscript{10}]PheT, which was measured in this study\textsuperscript{14}.

Preliminary reports from a subset of subjects demonstrated the rapid formation of diol epoxides and potential immediate negative health consequences of smoking\textsuperscript{16}, and examined the metabolism of [D\textsubscript{10}]Phe administered to 16 smokers either in a cigarette or orally as a biomarker of the activation pathway\textsuperscript{17}. With the completion of the clinical trial, a comprehensive analysis of the bioactivation of Phe is now reported.

\textbf{II.B Methods}

\textbf{Clinical Study Design}. The study was approved by the U.S. Food and Drug Administration and the University of Minnesota Institutional Review Board. Details on
the recruitment of subjects and dosing protocols were previously reported\textsuperscript{16,17}. Subjects were recruited using advertisements on the radio, television or metropolitan and campus newspapers. Volunteers interested in the study called the University of Minnesota Transdisciplinary Tobacco Use Research Center (TTURC) and were informed about the study. The preliminary screening was performed over the phone to select subjects meeting the following specific inclusion criteria: the smoking of at least 10 cigarettes daily for the past year and good physical and mental health. Female subjects who were pregnant or nursing were excluded. Eligible subjects were further invited to the clinic site for an orientation visit to fill out a detailed questionnaire regarding the smoking and medical history. Pregnancy tests were also conducted. Subject recruitment incentives were used and an average of $500 was paid to each subject for the completion of the study.

A total of twenty-five eligible subjects, eight male, were recruited between April 2008 and September 2010. The age of these subjects was between 23 and 54 years with a mean age (± standard deviation) of 36.2 ± 10.4 years. Their weights ranged from 61 to 113 kg with a mean weight (± standard deviation) of 87.4 ± 15.4 kg. Twelve subjects were Caucasian, eight were African-American and five subjects reported being multi-racial.

A dose of 10 µg (53.2 nmol) [D\textsubscript{10}]Phe was administered to 25 subjects, either as an oral solution or by smoking cigarettes to which [D\textsubscript{10}]Phe had been added. [D\textsubscript{10}]Phe (98%, containing 2% nondeuterated Phe) was purchased from Cambridge Isotope Laboratories, and then repurified in the University of Minnesota Molecular and Cellular Therapeutics GMP facility. The study had a randomized, open-label, single-dose, crossover design. The order of administration was randomized and each dose was separated by a washout period of at least one week. For the oral dose, each subject was given 10 µg [D\textsubscript{10}]Phe (5 mL of 20% ethanol-80% water solution). The dosing bottle was rinsed twice with water to ensure accurate dosing.

Before the administration of [D\textsubscript{10}]Phe to subjects by smoking, [D\textsubscript{10}]Phe was dissolved in 20 µL ethanol and added to each of 20 Marlboro cigarettes at each of the following levels.
(µg): 0, 25, 50, 100, and 200. Ten cigarettes at each level were smoked on a machine under International Organization for Standardization (ISO) conditions (puff volume, 35 mL; puff interval, 60 s; puff duration, 2 s; ventilation holes not blocked) and Health Canada intense conditions (puff volume, 55 mL; puff interval, 30 s; puff duration, 2 s; ventilation holes blocked), respectively.\textsuperscript{16} [D\textsubscript{10}]Phe in the mainstream smoke of three cigarettes at each level was quantitated. Levels of [D\textsubscript{10}]Phe in the cigarettes were plotted against levels of [D\textsubscript{10}]Phe detected in mainstream smoke. It was determined that adding 80 µg [D\textsubscript{10}]Phe in the cigarettes followed by the smoking of cigarettes under Health Canada intense conditions would lead to the delivery of 10 µg [D\textsubscript{10}]Phe. As such, in the smoking arm, a total of 80 µg [D\textsubscript{10}]Phe was added to cigarettes and subjects were instructed to follow Health Canada intense conditions while smoking the cigarettes containing [D\textsubscript{10}]Phe in order to receive a dose of 10 µg [D\textsubscript{10}]Phe in the mainstream smoke.\textsuperscript{16} The administration of [D\textsubscript{10}]Phe by smoking was performed in a specially ventilated room at TTURC. Subjects underwent an adaptation trial prior to smoking the cigarettes containing [D\textsubscript{10}]Phe. Subjects smoked the cigarettes through the smoking topography device which recorded the puff volume, puff duration and puff number. A Marlboro cigarette was used in the practice session, after which subjects were then allowed to smoke the cigarette containing [D\textsubscript{10}]Phe. The smoking process was also observed by the clinician to ensure good compliance. The smoking session usually lasted about 4-5 min.

Blood samples of 10 mL each were taken prior to dosing and 15, 30, 45, 60, 90, 120, 150, 240, 360, 540 or 720, and 1440 min after the completion of administration. Blood samples were centrifuged to obtain plasma which was stored at -20 °C until analysis. Urine samples were obtained pre-dosing and at the following post-dosing intervals: 0-30, 30-60, 60-120, 120-360, 360-720, 720-1440, 1440-2880 min. The volume of each urine collection was measured and an aliquot of 50 mL was stored at -20°C until analysis.

**Analysis of urine.** The method was similar to that described previously.\textsuperscript{16,17} A mixture of 0.5 mL urine sample, 0.75 mL pH 5 sodium acetate buffer, 18 µL β-glucuronidase plus arylsulfatase (recombinant from *Escherichia coli* BL21, ~100,000 units per mL, Roche
Applied Sciences, Indianapolis, IN) and 5 µL 100 pg/µL [¹³C₆]PheT (internal standard) was incubated in a 15-mL disposable glass centrifuge tube at 37°C overnight and then applied to a preconditioned Strata-X polymeric solid-phase extraction cartridge (200mg/6mL, Phenomenex, Torrance, CA). The cartridge was washed by 5 mL of 1% NH₄OH-10% MeOH in H₂O and the eluent was discarded. The PheT-containing fraction was eluted with 5 mL of 50% CH₃OH in H₂O, concentrated to dryness and reconstituted in 0.5 mL H₂O. The solution was further applied to a preconditioned phenyl boronic acid (PBA) column (100 mg/1 mL, Agilent Technologies, Palo Alto, CA). The PBA column was washed by 0.1 mL H₂O and placed in the vacuum pump overnight to completely remove water. The column was further rinsed by 2 mL acetone and the eluent was discarded. The PheT-containing fraction was eluted with 1 mL CH₃OH, concentrated to dryness, reconstituted in 180 µL CH₃OH and transferred to a glass insert vial (CP-0952-03, Chrom Tech Inc, Apple Valley, MN). The vial was placed in the vacuum pump until dry and 15 µL Bis(trimethylsilyl)trifluoroacetamide (BSTFA, Regis Technologies, Morton Grove, IL) with 2 fmol BaPT was added as the injection standard to correct the between-injection variability during the sample analysis by mass spectrometry. The samples were incubated at 60°C for 1 hr and 2 µL were injected on the gas chromatography-negative ion chemical ionization-tandem mass spectrometry (GC-NICI-MS/MS).

The GC-NICI-MS/MS analysis was carried out with a TSQ Quantum instrument (Thermo Scientific, San Jose, CA) and a DB-17MS column (30 m x 0.25 mm x 15µm, Agilent Technologies, Palo Alto, CA). The analytes were detected as their trimethylsilyl derivatives. The oven temperature program after the injection of samples was set as follows: 80°C for 1 min, then increased from 80 to 190°C at 30°C/min, then increased from 190 to 210 at 3°C/min, then increased from 210 to 320°C at 30°C/min, and held for 2 min. The temperature of the injection port and MS transfer line temperature were set at 250 and 280°C, respectively. The NICI-MS/MS conditions were set as follows: CI gas, methane at 1.5 mL/min; source temperature, 200°C; emission current, 350 µA; collision energy 12 eV and electron energy -150eV. [D₁₀]PheT, [¹³C₆]PheT and PheT were
detected at m/z 382 → m/z 220, m/z 378 → m/z 216, and m/z 372 → m/z 210, respectively. The total run time was 20 min and the retention time for [D_{10}]PheT, [^{13}C_6]PheT and PheT was about 10.9 min.

**Analysis of plasma.** The preparation of plasma samples was similar as urine samples except the replacement of Strata-X polymeric solid-phase extraction cartridge by Oasis MCX Mixed-mode cation exchange solid-phase extraction cartridge (60mg/1mL, Waters, Milford, MA). The MCX column was rinsed with 3 mL 1M HCl and 5 mL 1% CH$_3$OH. The PheT-containing fraction was eluted with 3 mL 40% CH$_3$OH, concentrated to dryness and reconstituted in 0.5 mL H$_2$O. The samples were then applied to PBA columns for further purification. The GC-NICI-MS/MS analysis of plasma samples was the same as urine samples.

It should be noted that since approximately 90% of the PheT in human urine exists as sulfate and glucuronide conjugates$^{12}$, plasma and urine samples were incubated with β-glucuronidase and arylsulfatase before quantitation. Hence the reported level of [D$_{10}$]PheT in both plasma and urine is the sum of the free [D$_{10}$]PheT and its conjugates. In addition, the [D$_{10}$]PheT quantitated in the current study was racemic, consisting of both enantiomer 9 and 12 as shown in Figure I-6.

**Pharmacokinetic Analysis.** Noncompartmental analysis was performed with the use of Phoenix WinNonlin$^{\text{TM}}$ (v6.1, PharSight, Cary, NC) to calculate the area under the [D$_{10}$]PheT plasma concentration-time curve ($AUC_{(PheT)}$) and the elimination half-life ($t_{1/2}$) of [D$_{10}$]PheT.

The clearance of [D$_{10}$]PheT ($CL_7$, Figure II-1) was estimated from the slope of the urinary excretion rate vs. mid-point plasma concentration curve:

$$\frac{\Delta X}{\Delta t} = CL_7 \times CP_{mid} \quad \text{Equation II-1}$$

where $CP_{mid}$ was the plasma concentration of [D$_{10}$]PheT at the mid-point of the urine collection interval. The urinary excretion rate ($\Delta X/\Delta t$) was determined from the amount
of [D_{10}]PheT excreted in urine at each interval (ΔX) divided by the length of the collection interval (Δt). Urine collections at 720-1440 and 1440-2880 min were not included in the estimate of CL7 because very low concentrations were present in urine in those intervals, and each of those intervals was > 1 half-life of [D_{10}]PheT^{18}.

The total amount of [D_{10}]PheDE formed during activation (A_{act}) was estimated by two methods (equations II-2 and II-3; see Appendix II-A for detailed derivation of the equations). The first method was using the product of AUC_{(PheT)} and CL7:

\[ A_{act,plasma} = AUC_{(PheT)} \times CL_7 \quad \text{Equation II-2} \]
The second method was to calculate the amount of $[\text{D}_{10}\text{PheT}]$ collected in urine up to 48 hr ($X_{u(PheT),t=48}$):

$$A_{act,urine} = X_{u(PheT),t=48} \quad \text{Equation II-3}$$

The percentage of dose activated ($f_{act}$) was calculated as:

$$f_{act} = \frac{A_{act}}{D} \times 100\% \quad \text{Equation II-4}$$

where $D$ is the administered dose of $[\text{D}_{10}\text{Phe}]$ (10 µg or 53.2 nmol).

$A_{act,plasma}$ and $A_{act,urine}$ are two indicators of systemic exposure to $[\text{D}_{10}\text{PheT}]$ and thus formed the basis of diagnostic plot I (Figure II-2, top).

Differences in $A_{act}$ between smoking and oral dosing were calculated as:

$$A_{act (lung)} = A_{act,smk} - A_{act,oral} \quad \text{Equation II-5}$$

where $A_{act,smk}$ and $A_{act,oral}$ are $A_{act}$ after smoking and oral dosing, respectively. $A_{act (lung)}$ is a measure of lung contribution to the formation of $[\text{D}_{10}\text{PheT}]$. A positive value of $A_{act (lung)}$ indicates that the lung contributed more to the formation of $[\text{D}_{10}\text{PheT}]$ than liver, and hence played a major role in metabolic activation. In addition, the larger the value of $A_{act (lung)}$, the larger local (lung) exposure to $[\text{D}_{10}\text{PheT}]$.

The relative bioavailability of $[\text{D}_{10}\text{PheT}]$ after oral dosing compared to the smoking administration was calculated from the ratio of oral/smoking $AUC_{(PheT)}$ values ($F_{AUC}$) and the ratio of oral/smoking $X_{u(PheT),t=48}$ values ($F_{Xu}$). $F_{AUC}$ and $F_{Xu}$ were used as two indicators of lung contribution to activation, and formed the basis for diagnostic plot II (Figure II-2, bottom). The combination of plots I and II allowed the identification of subjects with substantial local exposure.
Figure II-2. Diagnostic plot I (top) and II (bottom)
According to equation II-A23 in Appendix II-A, the amount of \([D_{10}]\text{PheDE}\) formed after oral dosing and smoking could be calculated as

\[
A_{act} = [(f_a \times f_m) + f_{act}] \times \text{Dose} \quad \text{Equation II-6}
\]

where \(f_a\), \(f_{act}\) and \(f_m\) are the fraction of the dose absorbed, the fraction of the dose converted to \([D_{10}]\text{PheDE}\) during first-pass activation and the fraction of \([D_{10}]\text{Phe}\) converted to \([D_{10}]\text{PheDE}\) in the systemic circulation, respectively. Since changes in the route of administration would only affect \(f_a\) and \(f_{act}\), but \(f_m\) would remain the same, equation II-6 could be rewritten as II-6-1 and II-6-2 for the oral and smoking arm, respectively.

\[
A_{act \ (oral)} = [(f_a \ (oral) \times f_m) + f_{act \ (oral)}] \times \text{Dose} \quad \text{Equation II-6-1}
\]

\[
A_{act \ (smk)} = [(f_a \ (smk) \times f_m) + f_{act \ (smk)}] \times \text{Dose} \quad \text{Equation II-6-2}
\]

\(f_a \ (oral)\) and \(f_a \ (smk)\) are the fraction of \([D_{10}]\text{Phe}\) dose absorbed after oral dosing and smoking, respectively. \(f_{act \ (oral)}\) and \(f_{act \ (smk)}\) are the fraction of \([D_{10}]\text{Phe}\) converted to active intermediates during first-pass metabolism after oral dosing and smoking, respectively.

**Genotyping.** Twelve polymorphisms of metabolizing enzymes of Phe were determined by the BioMedical Genomics Center at the University of Minnesota. Genotyping was performed using the iPLEX® Gold method (Sequenom, Inc, San Diego, CA). Similar methods have been previously described\(^{19}\). Briefly, the method is based on the primer-extension reaction that generates allele-specific products with distinct masses detected by Matrix-Assisted Laser Desorption / Ionization Time-Of-Flight (MALDI-TOF) mass spectrometry. Briefly, the method started with polymerase chain reaction (PCR) amplification followed by shrimp alkaline phosphatase (SAP) treatment to remove unincorporated dNTPs. Single-base extension (SBE) was carried out by the addition of SBE primers, iPLEX® enzymes and buffers. SBE products were measured with the MassARRAY® system (Sequenom, Inc, San Diego, CA), and mass spectra were analyzed.
with TYPER® software (Sequenom, Inc, San Diego, CA). iPLEX reagents and protocols for multiplex PCR, SBE and generation of mass spectra were based on the manufacturer’s instructions.

The polymorphisms of CYP1A1 and CYP1B1 investigated in the study were CYP1A1MspI, CYP1A1I462V, CYP1B1R48G, CYP1B119S, CYP1B1L432V and CYP1B1N453S. In addition, EPHX1Y113H and EPHX1H139R, two polymorphisms of microsomal epoxide hydrolase 1 (EPHX1), were also measured. The polymorphisms of the detoxification enzymes glutathione S-transferase P1 (GSTP1), T1 (GSTT1), and M1 (GSTM1) measured in the study were GSTP1A114V, GSTP1I105V, GSTT1 null and GSTM1 null.

**Statistical Analysis.** A paired t-test was used to compare $t_{1/2}$, $CL_7$ and $AUC_{(PheT)}$ between the oral and smoking arm. The measurement of relative bioavailability by two methods ($F_{AUC}$ vs. $F_{X0}$) was compared by using a paired t-test after the logarithmic transformation of the original data. Two-way ANOVA was employed to investigate the influence of data source (plasma vs. urinary data) and route of administration (oral vs. smoking) on the estimates of systemic exposure to [D$_{10}$]PheT. One-way ANOVA was used to identify polymorphisms that might have effects on an individual’s capacity to activate PAHs. A $p$-value of < 0.05 was considered to be significant.

**II.C Results**

Table II-1 reports the half-life, clearance ($CL_7$) and $AUC_{(PheT)}$ of [D$_{10}$]PheT after oral dosing and smoking administration of [D$_{10}$]Phe. No significant difference was observed in $t_{1/2}$, $CL_7$, or $AUC_{(PheT)}$ between the oral and smoking arms of the study, consistent with the results reported previously for 16 subjects$^{17}$.

$A_{act}$ is an estimate of the systemic exposure to [D$_{10}$]PheT and was calculated by two methods: use of plasma data ($A_{act,plasma}$) and use of urinary data ($A_{act,urine}$). As shown in Table II-2, $A_{act,plasma}$ and $A_{act,urine}$ were 4.23 ± 3.94 and 3.06 ± 1.91 nmol in the smoking arm, respectively. The percentage of the [D$_{10}$]Phe dose activated ($f_{act}$) was calculated as 7.96 ± 7.41 and 5.75 ± 3.60 based on plasma and urinary data, respectively.
Similar results were obtained from the oral arm. Neither route of administration (smoking vs. oral) nor source of data (plasma vs. urine) had a significant impact on the estimate of $A_{act}$.

Although at the group level no difference was observed in the means of $A_{act}$, a large inter-subject variability (> 20 fold) in $A_{act}$ was observed in both the smoking and oral arms (Table II-2). One purpose of this study was to identify people with a potentially increased susceptibility to carcinogenesis by PAHs because of their ability to carry out the bay region diol epoxide pathway. Since the route of administration of [D$_{10}$]Phe in the smoking arm mimicked the uptake of carcinogens by cigarette smokers, plasma and urinary data of the smoking arm were used to identify subjects with a large systemic exposure to [D$_{10}$]PheT, as shown in diagnostic plot I (Figure II-2, top). Six subjects whose estimates of systemic exposure ($A_{act}$) were in the top 30% of the population as measured by plasma data ($\geq 4.68$ nmol by equation II-2) and urine data ($\geq 4.29$ nmol by equation II-3) fell into zone A (large systemic exposure zone): subjects 1, 6, 14, 22, 23 and 24. Similarly, 5 subjects whose estimates of systemic exposure were in the lowest 30% of the population as measured by plasma data ($\leq 1.96$ nmol by equation II-2) and urine data ($\leq 1.58$ nmol by equation II-3), fell into zone B (low systemic exposure zone): subjects 2, 7, 12, 18 and 21.

The relative bioavailability of [D$_{10}$]PheT after oral dosing compared to the smoking administration was calculated from the ratio of oral/smoking $AUC_{(PheT)}$ values ($F_{AUC}$) and the ratio of oral/smoking $Xu_{(PheT),t=48}$ values ($F_{Xu}$). $F_{AUC}$ and $F_{Xu}$ were used as two indicators of lung contribution to metabolic activation. As shown in Figure II-3, after the administration of [D$_{10}$]Phe by smoking, the parent molecule [D$_{10}$]Phe passes through the lung before it reaches the systemic circulation. After oral dosing the parent molecule passes through the liver before it reaches the systemic circulation. Therefore, if $F_{AUC}$ and $F_{Xu}$ are less than 1, the lung has a greater contribution to the formation of [D$_{10}$]PheT. As shown in Table II-3, the relative bioavailability measured by $F_{AUC}$ and $F_{Xu}$ was 1.35 ± 1.33 and 1.35 ± 0.96, respectively. No significant difference was observed in relative bioavailability measured by the two methods. At the group level no significant difference
was observed in the disposition of parent molecule by lung and liver. However, a large inter-subject variability was observed (Table II-3), illustrating a greater than a 9-fold range in both $F_{AUC}$ and $F_{Xu}$. Diagnostic plot II was then developed to identify subjects with significant lung contribution to metabolic activation. Nine subjects with both $F_{AUC}$ and $F_{Xu}$ values of less than 1 were identified in zone C (Figure II-2, bottom): subjects 1, 3, 4, 5, 6, 8, 13, 22 and 25.
**Table II-1.** Half-life, clearance and AUC of [D\(_{10}\)]PheT after oral dosing and smoking

<table>
<thead>
<tr>
<th>ID</th>
<th>(t_{1/2}) (hr)</th>
<th>Oral</th>
<th>SMK</th>
<th>Oral</th>
<th>SMK</th>
<th>Oral</th>
<th>SMK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.2</td>
<td>7.8</td>
<td>13.2</td>
<td>1.07</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>17.7</td>
<td>8.1</td>
<td>7.3</td>
<td>0.13</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12.6</td>
<td>5.8</td>
<td>6.4</td>
<td>0.97</td>
<td>1.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11.0</td>
<td>4.9</td>
<td>6.4</td>
<td>0.37</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.9</td>
<td>5.5</td>
<td>4.8</td>
<td>0.24</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>6.3</td>
<td>4.3</td>
<td>7.3</td>
<td>1.10</td>
<td>1.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12.6</td>
<td>4.5</td>
<td>13.0</td>
<td>0.97</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5.5</td>
<td>5.4</td>
<td>4.6</td>
<td>0.46</td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6.4</td>
<td>3.2</td>
<td>1.3</td>
<td>1.04</td>
<td>0.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>6.9</td>
<td>5.1</td>
<td>7.1</td>
<td>0.55</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>4.9</td>
<td>7.5</td>
<td>3.1</td>
<td>0.85</td>
<td>0.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>4.4</td>
<td>3.2</td>
<td>3.6</td>
<td>0.43</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>7.9</td>
<td>6.0</td>
<td>7.0</td>
<td>0.58</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>7.4</td>
<td>4.0</td>
<td>4.5</td>
<td>1.70</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>8.8</td>
<td>4.6</td>
<td>4.0</td>
<td>0.94</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>5.1</td>
<td>4.4</td>
<td>4.5</td>
<td>0.93</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>9.3</td>
<td>1.9</td>
<td>1.9</td>
<td>1.98</td>
<td>1.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>6.1</td>
<td>2.2</td>
<td>2.2</td>
<td>0.80</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>5.4</td>
<td>N/A</td>
<td>1.8</td>
<td>N/A</td>
<td>0.90</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5.2</td>
<td>3.7</td>
<td>4.1</td>
<td>1.04</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>7.4</td>
<td>3.2</td>
<td>2.9</td>
<td>1.45</td>
<td>0.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>5.8</td>
<td>2.8</td>
<td>4.7</td>
<td>0.82</td>
<td>1.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>7.0</td>
<td>3.0</td>
<td>5.1</td>
<td>2.08</td>
<td>1.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4.8</td>
<td>5.3</td>
<td>3.0</td>
<td>1.95</td>
<td>1.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>6.3</td>
<td>2.4</td>
<td>2.9</td>
<td>0.65</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7.4</td>
<td>6.9</td>
<td>2.4</td>
<td>2.9</td>
<td>1.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>3.2</td>
<td>2.0</td>
<td>1.6</td>
<td>3.0</td>
<td>0.49</td>
<td>0.44</td>
<td></td>
</tr>
</tbody>
</table>

\(p\) value \(^b\) N.S. N.S. N.S.

\(^a\): Not available because plasma concentration of [D\(_{10}\)]PheT of subject 19 after smoking was below the limit of quantitation; \(^b\): Paired t-test was performed to compare the effect of route of administration; N.S. not significant, \(p > 0.05\)
Table II-2. $A_{act}$ of oral and smoking arm calculated by two methods

<table>
<thead>
<tr>
<th>ID</th>
<th>Oral arm</th>
<th>Smoking arm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_{act,plasma}$</td>
<td>$A_{act,urine}$</td>
</tr>
<tr>
<td></td>
<td>nmol</td>
<td>% of dose</td>
</tr>
<tr>
<td>1</td>
<td>8.25</td>
<td>15.51</td>
</tr>
<tr>
<td>2</td>
<td>0.41</td>
<td>0.77</td>
</tr>
<tr>
<td>3</td>
<td>5.57</td>
<td>10.47</td>
</tr>
<tr>
<td>4</td>
<td>1.79</td>
<td>3.36</td>
</tr>
<tr>
<td>5</td>
<td>1.32</td>
<td>2.48</td>
</tr>
<tr>
<td>6</td>
<td>4.62</td>
<td>8.68</td>
</tr>
<tr>
<td>7</td>
<td>4.36</td>
<td>8.20</td>
</tr>
<tr>
<td>8</td>
<td>2.45</td>
<td>4.61</td>
</tr>
<tr>
<td>9</td>
<td>3.24</td>
<td>6.09</td>
</tr>
<tr>
<td>10</td>
<td>2.83</td>
<td>5.32</td>
</tr>
<tr>
<td>11</td>
<td>6.31</td>
<td>11.86</td>
</tr>
<tr>
<td>12</td>
<td>1.36</td>
<td>2.56</td>
</tr>
<tr>
<td>13</td>
<td>3.43</td>
<td>6.45</td>
</tr>
<tr>
<td>14</td>
<td>6.75</td>
<td>12.69</td>
</tr>
<tr>
<td>15</td>
<td>4.3</td>
<td>8.08</td>
</tr>
<tr>
<td>16</td>
<td>4.08</td>
<td>7.67</td>
</tr>
<tr>
<td>17</td>
<td>3.72</td>
<td>6.99</td>
</tr>
<tr>
<td>18</td>
<td>1.74</td>
<td>3.27</td>
</tr>
<tr>
<td>19</td>
<td>1.59</td>
<td>2.99</td>
</tr>
<tr>
<td>20</td>
<td>3.83</td>
<td>7.20</td>
</tr>
<tr>
<td>21</td>
<td>4.6</td>
<td>8.65</td>
</tr>
<tr>
<td>22</td>
<td>2.29</td>
<td>4.30</td>
</tr>
<tr>
<td>23</td>
<td>6.24</td>
<td>11.73</td>
</tr>
<tr>
<td>24</td>
<td>6.77</td>
<td>12.73</td>
</tr>
<tr>
<td>25</td>
<td>1.57</td>
<td>2.95</td>
</tr>
<tr>
<td>Mean</td>
<td>3.74</td>
<td>7.02</td>
</tr>
<tr>
<td>SD</td>
<td>2.06</td>
<td>3.87</td>
</tr>
</tbody>
</table>

$p$ value $^b$ N.S.

a: Not available because plasma concentration of [D$_{10}$]PheT after smoking was below the limit of quantitation; b: Two-way ANOVA was performed to evaluate the influence of route of administration and data source (plasma, urine); N.S. not significant, $p > 0.05$
The combination of zones A (diagnostic plot I) and C (diagnostic plot II) led to the identification of three subjects with both a large systemic exposure and a significant lung contribution to activation: subjects 1, 6 and 22. These three subjects formed much more $[D_{10}]\text{PheT}$ after smoking than oral dosing (Table II-4). In addition, $A_{a(e) \text{ (lung)}}$ of these 3 subjects was between 0.66 and 10.56 nmol. This amount of $[D_{10}]\text{PheT}$ indicated substantial local exposure, especially considering that the total exposure to $[D_{10}]\text{PheT}$ in a typical subject after smoking ranged from 3.06 to 4.23 nmol.

**Figure II-3.** Comparison of metabolic activation of $[D_{10}]\text{Phe}$ in the oral and smoking arm
Table II-3. Relative bioavailability by two methods ($F_{AUC}$ and $F_{Xu}$)

<table>
<thead>
<tr>
<th>ID</th>
<th>Relative Bioavailability</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F_{AUC}$</td>
<td>$F_{Xu}$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.20</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.92</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.98</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.50</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.73</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7.32</td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.91</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1.29</td>
<td>1.41</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1.14</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1.21</td>
<td>1.90</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.88</td>
<td>1.09</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.76</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1.53</td>
<td>1.33</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1.80</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1.06</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1.56</td>
<td>1.67</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>1.22</td>
<td>2.19</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>N/A</td>
<td>5.01</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1.02</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2.11</td>
<td>1.87</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>0.53</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>1.43</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>0.82</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.83</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.35</td>
<td>1.35</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.33</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>90% CI b</td>
<td>(0.93 -1.35)</td>
<td>(0.95 -1.38)</td>
<td></td>
</tr>
<tr>
<td>$p$ value c</td>
<td>N.S.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a: Not available because plasma concentration of [D$_{10}$]PheT after smoking was below the limit of quantitation; b: 90% CI of geometric means; c: Paired t test was performed to compare two methods of calculating relative bioavailability; N.S. not significant, $p > 0.05$
Table II-4. Significant lung contribution to metabolic activation in subject 1, 6 and 22

<table>
<thead>
<tr>
<th>ID</th>
<th>$A_{act, plasma}$ (nmol)</th>
<th>$A_{act(lung)}$ b (nmol)</th>
<th>$A_{act, urine}$</th>
<th>$A_{act(lung)}$ c (nmol)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Smoking</td>
<td>Oral</td>
<td>Smoking</td>
<td>Oral</td>
</tr>
<tr>
<td>1</td>
<td>18.81</td>
<td>8.25</td>
<td>10.56</td>
<td>8.00</td>
</tr>
<tr>
<td>6</td>
<td>10.77</td>
<td>4.62</td>
<td>6.15</td>
<td>7.03</td>
</tr>
<tr>
<td>22</td>
<td>7.14</td>
<td>2.29</td>
<td>4.85</td>
<td>4.29</td>
</tr>
<tr>
<td>A typical subject$^a$</td>
<td>4.23</td>
<td>3.74</td>
<td>-</td>
<td>3.06</td>
</tr>
</tbody>
</table>

a: Data of a typical subject was based on the group mean (n = 25); b: The difference of $A_{act, plasma}$ between smoking and oral dosing; c: The difference of $A_{act, urine}$ between smoking and oral dosing

Table II-5 shows the effects of metabolic enzyme polymorphisms on an individual’s capacity to activate PAHs as measured by estimates of $A_{act}$ (equations II-2 and II-3) of both the oral and smoking arms. Among 12 polymorphisms tested only the $GSTM1$ polymorphism was associated with a difference in systemic exposure. Figure II-4a (plasma data) and b (urine data) show the maximum, minimum and median of $A_{act}$ in $GSTM1$ negative and positive subjects after smoking. A more than two-fold difference was observed in $A_{act}$ between $GSTM1$ negative subjects ($A_{act, plasma}$: 5.87 ± 4.85 nmol; $A_{act, plasma}$: 4.16 ± 2.00 nmol; n = 12) and $GSTM1$ positive subjects ($A_{act, plasma}$: 2.60 ± 1.76 nmol; $A_{act, plasma}$: 2.05 ± 1.16 nmol; n=13) after smoking (p < 0.05). To further confirm the effects of $GSTM1$ on the metabolic activation, estimates of $A_{act}$ from the oral arm were investigated (Figure II-4c and d) and similar results were observed.
**Table II-5.** Effects of polymorphisms on individual’s capacity to activate PAHs

<table>
<thead>
<tr>
<th>Gene Polymorphism</th>
<th>% of Occurrence</th>
<th>p values&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Oral</th>
<th>Smoking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Heterozytes</td>
<td>Homozygotes</td>
<td>$A_{act,plasma}$</td>
</tr>
<tr>
<td><strong>CYP1A1 MspI</strong></td>
<td>64</td>
<td>36</td>
<td>0</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>CYP1A1 I462V</strong></td>
<td>88</td>
<td>12</td>
<td>0</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>CYP1B1 R48G</strong></td>
<td>32</td>
<td>60</td>
<td>8</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>CYP1B1 A119S</strong></td>
<td>8</td>
<td>92</td>
<td>0</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>CYP1B1 L432V</strong></td>
<td>20</td>
<td>52</td>
<td>28</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>CYP1B1 N453S</strong></td>
<td>84</td>
<td>16</td>
<td>0</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>GSTP1 I105V</strong></td>
<td>36</td>
<td>52</td>
<td>12</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>GSTP1 A114V</strong></td>
<td>84</td>
<td>16</td>
<td>0</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>EPHX1 Y113H</strong></td>
<td>32</td>
<td>68</td>
<td>0</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>EPHX1 H139R</strong></td>
<td>64</td>
<td>28</td>
<td>8</td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>GSTT1</strong></td>
<td>12 (presence)</td>
<td>88 (null)</td>
<td></td>
<td>N.S.</td>
</tr>
<tr>
<td><strong>GSTM1</strong></td>
<td>52 (presence)</td>
<td>48 (null)</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

<sup>a</sup>: One way ANOVA was performed to determine the influence of genotype on $A_{act}$; N.S. not significant $p > 0.05$, * $p < 0.05$, ** $p < 0.01$
Figure II-4. Comparison of GSTM1 negative (-) and positive (+) subjects after smoking (a and b) and oral dosing (c and d) (dotted line: median, * $p < 0.05$, ** $p < 0.01$)
Figure II-5. Correlation between genotyping and phenotyping (triangle: *GSTM1* positive; circle: *GSTM1* negative)
Figure II-5 integrates diagnostic plot I and GSTM1 genotype information, and reveals an interesting correlation between genotyping and phenotyping results. The occurrence of the GSTM1 negative genotype was 100% (n = 6) in zone A compared with 48% in the group as a whole (n = 25). In contrast, the occurrence of the GSTM1 negative genotype was only 20% (n = 5) in zone B. High occurrence of the GSTM1 negative genotype in subjects with a large systemic exposure to [D$_{10}$]PheT, and low occurrence of the GSTM1 negative genotype in subjects with a low systemic exposure to [D$_{10}$]PheT clearly suggests a correlation between GSTM1 genotype and an individual’s capacity to activate PAHs. It is of interest to note that the three subjects (1, 6 and 22) with both high systemic exposure and significant lung contribution to activation were all GSTM1 negative and African American.

II.D Discussion
Since lung cancer treatment is not particularly effective (five-year survival rate < 16%), prevention is an important alternative, especially considering that tobacco smoking accounts for 90% of cases of the disease in the United States$^{1-20}$. The outcome of smoking cessation, a major prevention strategy, depends on the intensity of intervention. The rate of successful smoking cessation at 1 year is 3-5% when smokers simply try to stop, 7-16% when behavioral intervention is provided, and up to 24% when smokers receive both pharmacological treatment and behavioral support$^{21}$. Therefore, the successful identification of susceptible individuals could lead to increasing the intensity of intervention for these individuals. This would potentially improve the outcome of smoking cessation interventions. The current study aimed to quantitate an individual’s capacity to metabolically activate PAHs (one group of carcinogens involved in lung cancer), and identify subjects with extensive activation and presumably higher lung cancer risk.

The large inter-subject variability (> 20 fold) in the capacity to activate [D$_{10}$]Phe was consistent with the large inter-subject variability in the activation of PAHs reported in the literature$^{22}$, and further justified the use of a PK approach to identify highly susceptible individuals. In our PK study of 25 subjects, the fraction of [D$_{10}$]Phe converted to
[D$_{10}$]PheT ranged from 0.49% to 15.04% ($A_{act,urine}$) in the smoking arm. Even though only a relatively small fraction of a PAH, as represented by [D$_{10}$]Phe, is metabolically activated in humans, it is believed to be critical in the initiation of carcinogenesis$^{8-11,23,24}$. Metabolites not quantified in this study include phenols and dihydrodiols as well as unidentified material.

The quantitation of an individual’s capacity to activate PAHs was the primary objective of the current study. The unique PK approach used here quantitated both systemic and local exposure to [D$_{10}$]PheT. There are no other published studies on PAH PK in humans. While the measurement of systemic exposure is important in the identification of susceptible subjects, local (lung) exposure is a physiologically more relevant measurement for two reasons: 1) it captures the magnitude of exposure to carcinogenic intermediates at the specific site where carcinogenesis occurs; 2) [D$_{10}$]Phe was employed as a surrogate of BaP and other related PAHs which generally have their strongest carcinogenic effects at the site of application, although there are exceptions$^{3,7,13}$. For locally acting carcinogens, lung exposure is much more relevant to tobacco smoke carcinogenesis of lung cancer than systemic exposure. Despite its crucial role in carcinogenesis by PAHs, local exposure is usually very difficult to measure in clinical trials. In the present study the integration of systemic exposure and relative bioavailability revealed important information regarding local exposure. The concept is that subjects with both large systemic exposure (zone A of plot I, Figure II-2) and significant lung contribution to activation (zone C of plot II, Figure II-2) had substantial local (lung) exposure. As shown in Figure II-3, for a given individual, the difference in systemic exposure after oral dosing and smoking is due to the different pathways that [D$_{10}$]Phe takes before reaching the systemic circulation, i.e., one single pass of parent molecule through the liver or lung. In other words, if more [D$_{10}$]Phe were activated after smoking than oral dosing ($F_{AUC} < 1$ or $F_{Xu} < 1$), then this subject’s lung contributed more to the activation than did liver. Therefore, subjects with relative bioavailability less than 1 as well as subjects with large systemic exposure were of interest.
If relative bioavailability is less than 1, then by definition $A_{act(oral)} < A_{act(smk)}$.

According to equations II-6-1 and II-6-2, there are three scenarios when $A_{act(oral)} < A_{act(smk)}$:

1) $f_a(oral) < f_a(smk)$ and $f_{act}(oral) < f_{act}(smk)$

2) $f_a(oral) > f_a(smk)$ and $f_{act}(oral) < f_{act}(smk)$

3) $f_a(oral) < f_a(smk)$ and $f_{act}(oral) > f_{act}(smk)$

In the first case, since $f_a(oral) < f_a(smk)$, the lung allowed a larger fraction of [D$_{10}$]Phe to enter systemic circulation than did the liver, indicating a larger systemic exposure after smoking than oral dosing. In addition, since $f_{act}(oral) < f_{act}(smk)$, more active intermediates would be produced in lung than liver, indicating substantial local exposure. Therefore, case 1 was “large systemic exposure, large local exposure”. Similarly, case 2 and 3 would be considered “low systemic exposure, large local exposure”, and “large systemic exposure, low local exposure”, respectively. Although in case 3 lung may receive lower exposure to the active intermediates than does the liver, given the substantial amount of active intermediates in the circulation, the attack of electrophilic intermediates against lung DNA would still probably be substantial. Therefore, for all the cases related to a relative bioavailability of less than 1, potentially negative health consequences due to substantial local exposure, may be likely. As such, diagnostic plot II was developed to identify subjects with a significant lung contribution to metabolic activation.

As shown in plot I (Figure II-2, top), systemic exposure was calculated by two methods (equations 2 and 3). Similarly, in plot II (Figure II-2, bottom) the relative bioavailability was also estimated by two methods ($F_{AUC}$ and $F_{Xa}$) to confirm the identification of subjects with significant lung contribution to activation. The combination of plot I and plot II led to the identification of 3 subjects. In these 3 subjects one single pass through lung produced the same level of [D$_{10}$]PheT as total exposure in a typical subject. This
suggests that these subjects may have higher susceptibility than the rest of the study group.

There was a statistically significant difference in systemic exposure to [D\textsubscript{10}]PheT between \textit{GSTM1} negative and positive subjects. Since \textit{GSTM1} is a detoxification enzyme that catalyzes the conjugation of electrophilic intermediates with glutathione, the deletion of the \textit{GSTM1} gene reduces the efficiency of the detoxification pathways. As a result, more [D\textsubscript{10}]Phe may go through the activation pathway and more end product of the activation pathway was observed in \textit{GSTM1} negative subjects. Furthermore, about a two-fold difference in activation caused by \textit{GSTM1} might lead to clinically different outcomes considering the important role of PAHs in lung carcinogenesis and decades of uptake of PAHs from cigarette smoking. It has been reported that \textit{GSTM1} negative subjects have higher PAH-DNA adducts in the lung than \textit{GSTM1} positive subjects\textsuperscript{25,26}. It is worth noting that the three subjects identified by diagnostic plots I and II were all \textit{GSTM1} negative. In addition, all three subjects were African American. The development of a population PK model is in progress which will estimate the population mean and inter-subject variability of the efficiency of the activation pathway. Age, gender, weight, race, renal function, genotype, and smoking history will be incorporated in the population PK analysis to examine their influence on lung cancer susceptibility.

In other studies, the analysis of the urine samples of smokers indicated that conjugation of Phe diol epoxides by glutathione is a relatively minor pathway, at least based on the amount of the appropriate N-acetylcysteine conjugate found in urine\textsuperscript{27}. The amounts appear to be too low to account for the decreased levels of [D\textsubscript{10}]PheT in the \textit{GSTM1} competent vs. \textit{GSTM1} null individuals observed here. It is possible that glutathione conjugation occurs earlier in the pathway that produces [D\textsubscript{10}]PheT, for example by reaction with Phe-1,2-epoxide (Figure I-4).

Although no effects of other polymorphisms on systemic exposure were observed, it is likely that multiple genes are involved in the metabolic activation of PAHs\textsuperscript{28}. The low frequency of minor alleles and relatively small number of subjects in the present study
may explain the failure to detect an association of these polymorphisms with our measures of exposure.

One limitation of the present study was the lack of a PK profile of the parent compound [D_{10}]Phe due to its extremely low concentrations in plasma (< 25 pg/mL) and urine (< 5 pg/mL) following the 10 µg dose of [D_{10}]Phe^{17}. Animal PK studies have reported using a dose of 4.5-10 mg/kg which was at least 30,000-fold higher than the dose used in this study^{29,30}. The half-life and oral clearance of phenanthrene in a mouse study with an oral dose of 4.5 mg/kg were reported to be 0.32 hr and 2.18 L/hr·kg, respectively. Another study in rat and guinea pig indicated that more than 90% of the parent compound was converted to metabolites after oral administration of 10 mg/kg 14C-labelled Phe. The distribution of Phe to tissues was not well characterized, but Phe is probably extensively distributed to tissues because of its lipophilicity. Collectively, the low dose of [D_{10}]Phe in human study, extensive metabolism and rapid distribution are likely to contribute to the low concentration of [D_{10}]Phe in plasma and urine.

In summary, in the present study the metabolic activation of Phe in humans has been investigated. Plasma and urinary data of the smoking arm indicated a more than 20-fold difference in an individual’s capacity to activate PAHs, and formed the basis of diagnostic plot I to identify subjects with large systemic exposure. The relative bioavailability between oral dosing and smoking also showed a large inter-subject variability (> 9 fold), and formed the basis of diagnostic plot II to identify subjects with significant lung contribution to activation. The combination of plots I and II led to the identification of subjects with substantial lung exposure. This approach may have significant potential in the prediction of lung cancer risk. The integration of phenotyping and genotyping results indicated that GSTM1 played an important role in the detoxification of Phe.
CHAPTER III  Population Pharmacokinetic Analysis of the
Metabolic Activation of Phenanthrene in Humans

III.A Introduction

[D_{10}]PheT is a biomarker of the diol epoxide metabolic activation of [D_{10}]Phe. The more
[D_{10}]PheT formed, the more extensive the metabolic activation. But the location of the
activation is no less important than the extent of the activation. While the extent of the
activation contributes to systemic exposure, the location of the activation determines the
magnitude of local exposure such as lung exposure, which is crucial in lung
carcinogenesis.

One can consider a general model for [D_{10}]PheT appearance after dosing [D_{10}]Phe by any
extravascular route. Two pathways will contribute to the formation of [D_{10}]PheT as
shown in Figure III-1. One pathway (Pathway I) is first-pass activation. The other
(Pathway II) is the absorption of the parent compound followed by the activation in the
systemic circulation. Two assumptions regarding the plasma concentration-time profiles
of [D_{10}]PheT from these two pathways can be made. As shown in Figure III-2, it can be
assumed that [D_{10}]PheT from Pathway I demonstrated a PK profile similar to an i.v.
bolus because of its rapid formation and immediate appearance in the systemic
circulation. The decline of plasma concentration is due to the elimination constant for
[D_{10}]PheT. [D_{10}]PheT appearing from Pathway II is assumed to have a biexponential PK
profile. Three rate processes are involved in the determination of the shape of the PK
profile of [D_{10}]PheT: 1) absorption of the parent compound [D_{10}]Phe; 2) the elimination
of the parent compound; 3) the excretion of [D_{10}]PheT. Because [D_{10}]Phe cannot be
measured in the plasma, processes 1 and 2 cannot be distinguished from one another, and
the rate of increase of [D_{10}]PheT is controlled by the slower process of the two. Therefore,
it is assumed that the plasma concentration of [D_{10}]PheT from Pathway II will have a
biexponential shape. The observed plasma concentration-time profile of $[D_{10}]$PheT is a combination of Pathways I and II.

The relative contribution of each pathway is of particular interest. The fraction of the total dose (10 µg or 53.2 nmol) activated through Pathways I and II are defined as $f_1$ and $f_2$, respectively. In the smoking arm, if $f_1 >> f_2$, this would indicate that the lung is the primary site of activation and that local exposure is substantial. Theoretically subjects with substantial local exposure could be identified by smoking data alone if the two assumptions about the PK profiles of $[D_{10}]$PheT from Pathways I and II are reasonable. Similarly, the relative magnitude of $f_1$ and $f_2$ in the oral arm might also provide information about the role of liver and intestine in the metabolic activation.

Figure III-1. Two sources of the formation of $[D_{10}]$PheT
As such, a population (Pop) PK approach was used to estimate the typical value of $f_1$ and $f_2$ in the oral and smoking arms, respectively. But the smoking arm was the focus of the Pop PK analysis because the administration of $[D_{10}]Phe$ by smoking mimicked the uptake of carcinogens by cigarette smokers. After smoking cigarettes containing $[D_{10}]Phe$, subjects with large $f_1$ estimated by the Pop PK model were believed to have substantial lung exposure. The use of the Pop PK approach to identify subjects with substantial local exposure could be validated by the two-step diagnostic plot approach discussed in Chapter II. If the majority of the subjects with substantial local exposure identified by these two approaches are the same, then the validation would considered to be successful.
III.B Methods

Model development

The 24-hr plasma and 48-hr urine samples of 25 subjects after the administration of 10 µg (53.2 nmol) [D_{10}]Phe by smoking and oral doing were collected and analyzed by GC-MS as described in Chapter II. The statistical software R 2.14.1 (open source software) was used to organize the data. Pop PK analysis was performed with NONMEM® version 7.2.0, implemented with PDx-Pop® version 5 (ICON Development Solutions, Ellicott City, MD). The first order conditional estimation (FOCE) and the subroutine ADVAN6 were used to estimate Pop PK parameters. Model evaluation was based on the maximum likelihood and model diagnostic plots. Scatter plots of dependent variable (DV) vs. population-predicted concentration (PRED), DV vs. individual predicted concentration (IPRED), conditional weighted residual (CWRES) vs. PRED, CWRES vs. time and CWRES vs. ID were used as model diagnostic plots.

The structural model is illustrated in Figure III-3. Compartments 1, 2, 3 and 4 were abbreviated as CMT1, CMT2, CMT3 and CMT4. CMT1 and CMT2 were two dosing compartments and represented the first-pass activation pathway (Pathways I) and Pathway II, which both contributed to the formation of [D_{10}]PheT, respectively. The input rate constant \( k_1 \) was fixed to 1000 hr\(^{-1}\) to describe the immediate formation of [D_{10}]PheT by the first-pass activation. The input rate constant \( k_2 \) is a PK parameter estimated by the model. The transformation process from CMT2 to CMT3 described by \( k_2 \) is the hybrid of absorption and activation (Pathway II) shown in Figure III-1. Other PK parameters included the clearance (\( CL \)) and volume of distribution (\( V \)) of [D_{10}]PheT. But two most important PK parameters were the bioavailability fraction parameters of CMT1 and CMT2: F1 and F2. The bioavailability fraction parameters F1 and F2 were defined as the fraction of the dose introduced to CMT1 and CMT2 that was transformed to CMT3\(^1\). Since the doses in CMT1 and CMT2 were both set to be 10 µg (53.2 nmol), the amount of [D_{10}]PheT contributed by Pathways I and II were 53.2 × F1 and 53.2 × F2 nmol, respectively. The total amount of [D_{10}]PheT formed after the administration of [D_{10}]Phe
was 53.2 × (F1 + F2) nmol. Therefore, by definition F1 and F2 in the NONMEM® model were \( f_1 \) and \( f_2 \), respectively. A large value of \( F1 \) (\( f_1 \)) estimated by the PK model of the smoking arm would indicate that the subject has substantial local exposure.

Figure III-3. Structural model of Pop PK analysis

The differential equations of the model were listed as the following:

\[
\frac{dA_1}{dt} = -f_1 \times k_1 \times A_1 \quad \text{Equation III-1}
\]

\[
\frac{dA_2}{dt} = -f_2 \times k_2 \times A_2 \quad \text{Equation III-2}
\]

\[
\frac{dA_3}{dt} = (f_1 \times k_1 \times A_1) + (f_2 \times k_2 \times A_2) - (CL \times C_{PheT}) \quad \text{Equation III-3}
\]

\[
\frac{dA_4}{dt} = CL \times C_{PheT} \quad \text{Equation III-4}
\]

\( A_1 \) and \( A_2 \) were the amounts of parent compound in CTM1 and CMT2, respectively. \( C_{PheT} \) was the plasma concentration of [D\textsubscript{10}]PheT. \( A_3 \) and \( A_4 \) were the amounts of [D\textsubscript{10}]PheT in CMT3 and CMT4, respectively.

The inter-individual variability in the PK parameters were modeled assuming a log-normal distribution\textsuperscript{2} described as

\[
P_{ij} = P_i \times \exp(\eta_{ij}) \quad \text{Equation III-5}
\]
\( P_{ij} \) was the individual estimate of the \( ith \) PK parameter of subject \( j \) and \( P_i \) was the population estimate of the \( ith \) PK parameter. \( \eta_{ij} \) was the random variable normally distributed with a mean of zero and a variance of \( \omega_i^2 \).

The intra-subject variability of the plasma data was modeled using a combined additive and exponential error model\(^2\) described as
\[
Y_{obs} = Y_{pred} \times \exp(\epsilon_i) + \epsilon_2 \quad \text{Equation III-6}
\]
\( Y_{obs} \) was the observed plasma concentration and \( Y_{pred} \) was the plasma concentration predicted by the model. \( \epsilon_1 \) and \( \epsilon_2 \) were normally distributed error terms with means of zero and variances of \( \sigma_1^2 \) and \( \sigma_2^2 \), respectively.

Similarly, the intra-subject variability of the urine data were described as
\[
Y_{obs} = Y_{pred} \times \exp(\epsilon_i) + \epsilon_4 \quad \text{Equation III-7}
\]
\( Y_{obs} \) was the observed urine concentration and \( Y_{pred} \) was the urine concentration predicted by the model. \( \epsilon_3 \) and \( \epsilon_4 \) were normally distributed error terms with means of zero and variances of \( \sigma_3^2 \) and \( \sigma_4^2 \), respectively.

The foregoing equations could be used to describe the metabolic activation of [D\(_{10}\)Phe in both oral and smoking arms. As such, the data of both oral and smoking arms were fitted by the PK model simultaneously. The \( CL \) and \( V \) were set to be the same in two arms. In addition, the input rate constant \( k_1 \) was set to be 1000 hr\(^{-1}\) in both arms to describe the immediate appearance of [D\(_{10}\)PheT contributed by the first-pass activation. But two separate sets of \( f_1 \), \( f_2 \) and \( k_2 \) were used in the oral and smoking arms to explain the differences in the extent and rate of the formation of [D\(_{10}\)PheT caused by different first-pass organs.

**Statistical analysis**

One-way analysis of variance (ANOVA) was used to evaluate the effects of age, weight (WGT), gender, race, renal function, clearance of [D\(_{10}\)PheT, number of cigarettes per
day (CPD), years of smoking, and total exposure to cigarette carcinogens on an individual’s capacity to activate PAHs. Creatinine clearance ($CL_{cr}$) was used as the indicator of renal function and was calculated by the ratio of the excretion rate of creatinine obtained from urine data to the plasma concentration of creatinine. Total exposure to cigarette carcinogens was calculated by the multiplication of CPD and years of smoking. The amount of $[D_{10}]PheT$ collected in urine over 48 hr after smoking the cigarette containing $[D_{10}]Phe$ ($X_{u_{smk}}$) was used as the indicator of an individual’s capacity to activate PAHs. A $p$-value of $< 0.05$ was considered to be significant.

### III.C Results

The plasma data of the oral and smoking arms are shown in Figure III-4 and Figure III-5, respectively. The urine data of the oral and smoking arms are shown in Figure III-6 and Figure III-7, respectively. Examples of a NONMEM® dataset and control file are listed in the Appendix III-A. The demographic information has been summarized in Chapter II and is not discussed in the current chapter.

The minimization procedure searching for the final estimates of the PK parameters was terminated “due to rounding errors (error = 134)”. Although various approaches have been tried to revise the minimization process, the global minimum has not been located yet and therefore the model was not able to provide final estimates for the PK parameters.

Since the development of the structural model failed, covariate analysis based on the Pop PK model was not performed. Instead, one-way ANOVA was used to identify potential covariates of $X_{u_{smk}}$. The dataset used for statistical analysis is shown in Table III-1. The preliminary statistical analysis showed no significant effects of age, weight, gender, race, clearance of $[D_{10}]PheT$, CPD, years of smoking, and total exposure to cigarette carcinogens on $X_{u_{smk}}$. However, there was a significant correlation between renal function and $X_{u_{smk}}$. Subjects with higher creatinine clearance seemed to have larger $X_{u_{smk}}$ (Figure III-8).
Figure III-4. Plasma $[D_{10}]$PheT concentration-time profile of 25 subjects after oral dosing.
Figure III-5. Plasma [D_{10}]PheT concentration-time profile of 25 subjects after smoking
Figure III-6. Cumulative amount of urinary $[D_{10}]$PheT-time profile of 25 subjects after oral dosing
Figure III-7. Cumulative amount of urinary [D_{10}]PheT-time profile of 25 subjects after smoking
Figure III-8. Correlation between creatinine clearance and $Xu_{smk}$
Table III-1. Demographic and PK data of 25 subjects

<table>
<thead>
<tr>
<th>ID</th>
<th>$X_{t_{\text{smk}}}$ (nmol)</th>
<th>Age (years)</th>
<th>WGT (kg)</th>
<th>Gender</th>
<th>Race</th>
<th>CLcr (ml/min)</th>
<th>CPD</th>
<th>Years of smoking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>40</td>
<td>97.6</td>
<td>F</td>
<td>AA</td>
<td>161.3</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>0.43</td>
<td>47</td>
<td>79.5</td>
<td>M</td>
<td>AA</td>
<td>NA</td>
<td>23</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>3.45</td>
<td>47</td>
<td>79.5</td>
<td>M</td>
<td>AA</td>
<td>NA</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>2.06</td>
<td>44</td>
<td>72.4</td>
<td>F</td>
<td>AA</td>
<td>NA</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>2.27</td>
<td>38</td>
<td>76.3</td>
<td>F</td>
<td>CA</td>
<td>NA</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>7.03</td>
<td>54</td>
<td>94.8</td>
<td>M</td>
<td>AA</td>
<td>149.9</td>
<td>25</td>
<td>37</td>
</tr>
<tr>
<td>7</td>
<td>1.08</td>
<td>31</td>
<td>101.9</td>
<td>F</td>
<td>CA</td>
<td>NA</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>2.46</td>
<td>25</td>
<td>60.8</td>
<td>F</td>
<td>CA</td>
<td>110.7</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>3.76</td>
<td>52</td>
<td>100.5</td>
<td>F</td>
<td>CA</td>
<td>101.5</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>10</td>
<td>2.22</td>
<td>45</td>
<td>62.5</td>
<td>M</td>
<td>AA</td>
<td>66.3</td>
<td>14</td>
<td>25</td>
</tr>
<tr>
<td>11</td>
<td>3.22</td>
<td>39</td>
<td>99.3</td>
<td>F</td>
<td>CA</td>
<td>141.7</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>12</td>
<td>1.58</td>
<td>28</td>
<td>71.7</td>
<td>F</td>
<td>CA</td>
<td>93.2</td>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>2.96</td>
<td>24</td>
<td>91.9</td>
<td>M</td>
<td>MR</td>
<td>146.9</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>14</td>
<td>4.4</td>
<td>25</td>
<td>68</td>
<td>F</td>
<td>CA</td>
<td>92.1</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>15</td>
<td>3.29</td>
<td>41</td>
<td>112.7</td>
<td>M</td>
<td>CA</td>
<td>159.5</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>16</td>
<td>4.13</td>
<td>26</td>
<td>61.3</td>
<td>F</td>
<td>CA</td>
<td>118.4</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>17</td>
<td>2.33</td>
<td>25</td>
<td>63.5</td>
<td>F</td>
<td>CA</td>
<td>60.1</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>18</td>
<td>0.77</td>
<td>51</td>
<td>65.8</td>
<td>F</td>
<td>AA</td>
<td>51.9</td>
<td>12</td>
<td>34</td>
</tr>
<tr>
<td>19</td>
<td>0.26</td>
<td>33</td>
<td>94.3</td>
<td>F</td>
<td>MR</td>
<td>107.7</td>
<td>30</td>
<td>18</td>
</tr>
<tr>
<td>20</td>
<td>2.86</td>
<td>25</td>
<td>64.7</td>
<td>F</td>
<td>MR</td>
<td>97.1</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>21</td>
<td>1.42</td>
<td>31</td>
<td>97.4</td>
<td>M</td>
<td>MR</td>
<td>138.7</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>22</td>
<td>4.29</td>
<td>36</td>
<td>87.3</td>
<td>M</td>
<td>AA</td>
<td>102.8</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>23</td>
<td>5.26</td>
<td>24</td>
<td>73.1</td>
<td>F</td>
<td>MR</td>
<td>154</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>24</td>
<td>5.13</td>
<td>23</td>
<td>82.6</td>
<td>F</td>
<td>CA</td>
<td>130.5</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>25</td>
<td>1.87</td>
<td>50</td>
<td>72.2</td>
<td>F</td>
<td>CA</td>
<td>81</td>
<td>20</td>
<td>33</td>
</tr>
</tbody>
</table>

Gender: F for female and M for male

Race: AA for African American, CA for Caucasian, and MR for more than one races

$CL_{cr}$: creatinine clearance

CPD: cigarettes per day

NA: not available
III.D Discussion

The primary objective of this chapter was to develop an alternative approach to identify subjects with substantial local exposure. If the identification of subjects with substantial local exposure by the Pop PK approach was consistent with the two-step diagnostic approach, then the Pop PK approach would be preferred because only one dosing arm (smoking) would be required. In addition, the $\omega^2$ of $f_1$ ($f_1$) obtained from the Pop PK analysis of the smoking arm would also reveal the between-subject variability of an individual’s capacity to activate PAHs in lung.

However, the development of the Pop PK model has not been completed yet. The error message “rounding errors (error = 134)” is very common in the development of NONMEM® models\(^1\). The following approaches are often used to address the issue caused by “rounding errors (error = 134)”:

1. Try different initial estimates
2. Change the significant digits to a lower value
3. Simplify the OMEGA matrix
4. Try other estimation algorithms

In addition, a model-specific approach has also been tried to facilitate the minimization process: 1) use the cumulative amount of [D$_{10}$]PheT in the urine at 48 hr as the dose for CMT1 and CMT2 instead of 10 µg; 2) revise the definition of $f_1$ and $f_2$ to be the fraction of [D$_{10}$]PheT contributed by Pathways I and II, respectively; 3) express $f_2$ as 1 - $f_1$ to reduce the number of PK parameters. In addition, the $CL$ and $V$ have been fixed to reduce the number of PK parameters of the model. Oral and smoking data have also been fitted separately by the PK model to reduce the number of PK parameters. But none of these approaches resulted in successful minimization.

One possible explanation might be the lack of sufficient data to support the PK model. Although intensive sampling was performed, the information about the rising part of the plasma concentration-time curve may still be insufficient to support the estimation of $k_2$.  

\(^1\)
For example, the $C_{\text{max}}$ of the smoking arm was usually reached around 15-30 min. There were at most 2 sampling points (15 and 30 min) before the $C_{\text{max}}$ which might not be enough to estimate $k_2$. One suggestion for future studies might be to collect blood samples during the smoking administration to try to capture the rate of absorption.

Another possible explanation might be extensive activation in the lung. This is plausible because the plasma concentration-time profiles of most subjects in the smoking arm seemed to be monoexponential curves (Figure III-5). In other words, in the smoking arm Pathway I might be dominant with $f_2$ being essentially 0. If that is the case, the NONMEM® model would keep searching for an $f_2$ value but will never locate the global minimum because the low boundary of $f_2$ was set to be 0. However, the model still failed to converge after fixing $f_2$ as 0.

The oral arm seemed to have more data about the rising part of the plasma concentration-time curve, but the estimation of Pop PK parameters of the oral arm was not successful either. This may lead to doubts whether the two assumptions about the PK profiles of Pathways I and II were reasonable. The test of these two assumptions might be possible once the PK data for the parent compound are available.

While there has already been one method (two-step diagnostic approach) to identify subjects with substantial local exposure, the Pop PK analysis is still worth trying if more data become available. This is because from a pharmacokinetic-pharmacodynamic (PK-PD) point of view, the occurrence of lung cancer is the outcome (PD effect) of multiple years of cigarette carcinogen dosing including PAHs. The PK model of PAHs, together with the PK models of other major carcinogens like 4-(methylnitrosamino)-1-(3-pyridyl)-1-butane (NNK), might provide a comprehensive view of tobacco-related lung cancer and ultimately the prediction of lung cancer risk.

The results of the preliminary statistical analysis were interesting. The correlation between renal function and the measure of an individual’s capacity to activate Phe remains unexplained. Because urinary collections were carried out for 48 hr (> 5 half-lives of $[D_{10}]\text{PheT}$), the possibility that the correlation was due to incomplete collection
of [D_{10}]PheT has been ruled out. In order to further investigate the correlation, at least one blood sample should be collected in future large-scale clinical studies in order to quantitate the creatinine concentration in plasma and calculate the creatinine clearance.

III.E Conclusion
A NONMEM® dataset has been built for the 25 subjects in the pilot study and could be used as the platform for further model development. Various approaches have also been tried to develop the Pop PK model for the metabolic activation of [D_{10}]Phe. The inclusion of the PK profile of [D_{10}]Phe in the model may shed light on the diol epoxide activation of [D_{10}]Phe in humans. In addition, there seems to be a correlation between renal function and the measure of an individual’s capacity to active PAHs. Further studies are required to explain the correlation.
CHAPTER IV  The Optimization of the Study Design by Simulation

IV.A Introduction
The hypothesis of our research is that smokers with extensive activation of tobacco carcinogens are at higher lung cancer risk than others. In order to test the hypothesis, two projects have been proposed. The first project was the pilot study to characterize the diol epoxide metabolic activation of \([D_{10}]Phe\) in 25 subjects. This study has been completed and reported in Chapter II. The second project will be a large-scale study in which 350 smokers will be recruited, those with extensive activation identified, and bronchoscopy carried out to determine the existence of bronchial dysplasia, an early sign of the carcinogenesis process\(^1\). The correlation between extensive activation of Phe and high lung cancer risk will then be assessed.

Intensive sampling was performed in the pilot study to thoroughly characterize the diol epoxide metabolic activation of \([D_{10}]Phe\) in smokers. For example, both plasma and urine data were collected for 24 and 48 hr, respectively. In addition, \([D_{10}]Phe\) was administered by both oral dosing and smoking. Obviously, the study protocol of the first project cannot be simply applied to the second project without optimization, because of the large number of subjects involved in the second project. A simulation project was carried out which aimed to optimize the study protocol of the large-scale clinical trial so that the research hypothesis, if correct, could be successfully tested in a cost-efficient way.

The diol epoxide metabolic activation of \([D_{10}]Phe\) in 350 smokers after the administration of 10 \(\mu g\) \([D_{10}]Phe\) by oral dosing and smoking was simulated by a PK model with certain assumptions. The amount of parent compound converted to “active” intermediates after oral dosing and smoking \((A_{act,oral} \text{ and } A_{act,smk})\) as well as the values of PK parameters used in the model were based on the results of the pilot study reported in Chapter II. The
strategy of the simulation was to treat $A_{act, smk}$ or $A_{act, lung}$ (the difference between $A_{act, smk}$ and $A_{act, oral}$) as gold standards of extensive activation and presumably indicators of lung cancer susceptibility, and evaluate the robustness of alternative metrics in the determination of the correlation between extensive activation and high lung cancer risk.

IV.B Methods

The research hypothesis is that smokers with extensive activation are at higher lung cancer risk than others. Assuming that the hypothesis is correct, the simulation aimed to select a metric that was most likely to prove the correlation. The selected metric will provide the guidance for the design of the large-scale clinical trial.

In the simulation, the following assumptions were made: 1) all the “active” intermediate $[D_{10}]PheDE$ was converted to $[D_{10}]PheT$ so that the quantitation of $[D_{10}]PheT$ could provide a reasonable estimate of an individual’s capacity to activate Phe and other PAHs; 2) the formation and elimination of $[D_{10}]PheT$ followed linear PK; 3) systemic or local (lung) exposure to activated PAHs was one driving force for lung carcinogenesis; 4) subjects with the top 20% systemic or local exposure in the population were susceptible to lung cancer; 5) susceptible subjects have a probability of 30% of developing dysplasia while the others have a probability of 10% of developing dysplasia.

The rationale for assumption 4 was as follows. It has been reported that approximately 20% of life-long smokers develop lung cancer, indicating the existence of susceptible subpopulations. But the size of the susceptible subpopulation is unknown. For the convenience of discussion, in the simulation subjects with the top 20% systemic or local exposure were assumed to be susceptible to lung cancer.

The fifth assumption was based on the fact that smokers with large exposure to cigarette carcinogens appeared to have 3-fold higher lung cancer risk than those with low exposure. In addition, the prevalence of dysplasia in smokers appeared to be approximately 10% but the smoking status and smoking history of these subjects was not available (Roswell Park Cancer Institute, unpublished data). Another study indicated that
the prevalence of dysplasia in current and former heavy smokers (age: 45 - 74 years, > 30 pack-years) was about 19%\(^4\). It was speculated that the prevalence of dysplasia in current heavy smokers was higher (e.g. 30%) than former heavy smokers (e.g. 10%) but these two groups of subjects were not distinguished in the study. Collectively, for the convenience of discussion, it was assumed that the prevalence of dysplasia in susceptible smokers would be 30% and prevalence of dysplasia in other smokers would be 10%.

The simulation was performed using STELLA\textsuperscript{®} 9.1.3 (Isee Systems, Inc.; Lebanon, NH). Details of the simulation and equation derivations are described in the Appendix IV-A. Briefly, there were three variables in the PK model as shown in Figure IV-1: \(A_{act,smk}\), relative bioavailability (\(F\)) and clearance of [D\(_{10}\)]PheT (\(CL\)). These three variables followed log-normal distribution in the smoker population. The typical value and distribution of these variables used in the simulated population were based on the PK analysis of the pilot study reported in Chapter II. In other words, each subject in the simulated population had one set of \(A_{act,smk}\), \(F\) and \(CL\) values. The \(A_{act,oral}\) of each subject was determined by the product of \(A_{act,smk}\) and \(F\). It should be noticed that the dose used in the simulation was not the total dose (10 µg or 53.2 nmol). Instead, \(A_{act,smk}\) and \(A_{act,oral}\) were treated as the dose in the simulated smoking and oral arm because developing a measurement of diol epoxide metabolic activation was the focus of current study. The simulation was performed 350 times through the sensitivity test function of STELLA\textsuperscript{®} to simulate the PK profile of 350 subjects. In each run, a random value of \(A_{act,smk}\) was generated from the predetermined distribution of the variable. Similarly, random values of \(F\) and \(CL\) were also generated in the same run. The PK profile of one simulated subject was generated from one run, and 350 runs produced the PK profiles of 350 simulated subjects.

The term “extensive activation” may refer to large systemic exposure or substantial local exposure to activated PAHs. As such, two simulated scenarios were investigated. In scenario 1, subjects whose systemic exposure was within the top 20% of the group were considerate to be susceptible to lung cancer. In scenario 2, subjects whose local exposure
was within the top 20% of the group were considerate to be susceptible to lung cancer. \( A_{act,smk} \) was treated as the gold standard of systemic exposure. The difference between \( A_{act,smk} \) and \( A_{act,oral} \) was defined as \( A_{act,lung} \) (equation II-5) and treated as the gold standard of local exposure. \( A_{act,smk} \) and \( A_{act,oral} \) of each simulated subject were known from the simulation. Thus subjects with extensive activation (top 20% \( A_{act,smk} \) or \( A_{act,lung} \)) and presumably high lung cancer risk could be predetermined. This provided a unique opportunity to evaluate the relative robustness of metrics in the identification of subjects with extensive activation and eventually the determination of correlation between extensive activation and high lung cancer risk.

**Smoking arm**

\[
F = \frac{A_{act,smk}}{A_{act,oral}}
\]

**Oral arm**

In scenario 1, since systemic exposure to activated PAHs was assumed to be one driving force for lung carcinogenesis, the ideal metric to identify susceptible subjects
was \( A_{act,smk} \). But it was not possible to measure \( A_{act,smk} \) because the active form of [D\(_{10}\)Phe is [D\(_{10}\)PheDE, which readily reacts with water and is further converted to [D\(_{10}\)PheT. The quantitation of [D\(_{10}\)PheT in plasma and urine are possible approaches to estimate \( A_{act,smk} \). As such, three metrics were tested as the surrogates of \( A_{act,smk} \): 1) the cumulative amount of [D\(_{10}\)PheT excreted in the urine at time \( t \) after smoking (\( X_{usmk} \)); 2) the cumulative amount of [D\(_{10}\)PheT excreted in the urine at time \( t \) after oral dosing (\( X_{uoral} \)); and 3) the area under the curve of [D\(_{10}\)PheT from time 0 to infinity after the administration of [D\(_{10}\)Phe by smoking (\( AUC_{pht} \)). A total of 350 subjects were divided into two groups: groups 1 and 2. Subjects in the top 20% of the population in terms of \( X_{usmk}, X_{uoral} \) or \( AUC_{pht} \) were assigned to group 1. The remaining subjects were assigned to group 2. The number of subjects with dysplasia (\( D \)) in each group was calculated by equation IV-1 where \( n_1 \) was the number of susceptible subjects and \( n_2 \) was the number of the subjects not susceptible to lung cancer in each group. The percentage of subjects with dysplasia (\( P\% \)) in each group was calculated by equation IV-2 where \( n \) was the number of subjects in each group: \( n = 70 \) for group 1 and 280 for group 2. The relative risk ratio (\( RRR \)) was calculated by equation IV-3 where \( P_1\% \) and \( P_2\% \) were the percentage of subjects with dysplasia in group 1 and 2, respectively. The metric that produced a large \( RRR \) would be preferred in the clinical trial. Since susceptible subjects were assumed to have 3-fold higher lung cancer risk, the maximum value of \( RRR \) is 3. \( RRR \) based on 48-hr and 6-hr urine collection were compared to determine if the extension of urine collection time would substantially increase \( RRR \).

\[
D = (0.3 \times n_1) + (0.1 \times n_2) \quad \text{Equation IV-1}
\]

\[
P\% = \frac{D}{n} \times 100\% \quad \text{Equation IV-2}
\]

\[
RRR = \frac{P_1\%}{P_2\%} \quad \text{Equation IV-3}
\]

In scenario 2, subjects in the top 20% of the population in terms of local exposure (top 20% \( A_{act,lung} \)) in the population were considered as susceptible subjects. As in scenario 1, the
probability of developing dysplasia was 30% and 10% for susceptible subjects and the remaining subjects, respectively. While the ideal metric to identify susceptible subjects was $A_{act,lung}$, it was very difficult to estimate in the clinical study. Three metrics $X_{uoral}$, $X_{usmk}$ and $X_{ulung}$ were evaluated as the surrogates of $A_{act,lung}$. $X_{ulung}$ was defined as the difference between $X_{usmk}$ and $X_{uoral}$. A total of 350 subjects were divided into two groups: groups 1 and 2. Subjects with top 20% $X_{uoral}$, $X_{usmk}$ or $X_{ulung}$ were assigned to group 1. The remaining subjects were assigned to group 2. Equations IV-1, IV-2 and IV-3 were used to calculate the $RRR$ for each metric. The metric that produced a large $RRR$ would be preferred in the clinical trial. $RRR$ based on 48-hr and 6-hr urine collection were also compared to determine if the extension of urine collection time would substantially increase $RRR$.

Due to the large number of simulated subjects involved in the simulation, R 2.14.1 (open source software) was used to automatically calculate the $RRR$ for groups 1 and 2 in scenarios 1 and 2. Briefly, three groups of subjects were input into R: group EX (subjects with true extensive activation), group 1, and group 2. There were a total of 70 subjects in group EX because 20% of the simulated population were considered to be susceptible to lung cancer. Subjects in the top 20% of the population in terms of $A_{act,smk}$ (scenario 1) or $A_{act,lung}$ (scenario 2) were assigned to group EX. Group 1 and 2 had a total of 350 subjects, and the criteria for these two groups have been discussed previously. A cross-check was performed to determine the number of subjects that existed in both group EX and group 1 as well as those that existed in both group EX and group 2. In other words, R automatically counted the number of susceptible subjects in groups 1 and 2, and plugged these numbers into equations IV-1, IV-2 and IV-3 to calculate $RRR$ for each group.

**IV.C Results**

The plasma and urine data of a representative subject and a simulated subject after oral dosing and smoking were shown in Figures IV-2 and IV-3. The similarity of the PK properties of actual and simulated subjects supports the use of the current PK model for simulation.
Table IV-1 reports the results of scenario 1, in which subjects in the top 20% of the population in terms of $A_{act,smk}$ were considered to be susceptible to lung cancer. Table IV-1 indicated that the use of the top 20% $Xu_{smk}$ among the population as the metric to identify susceptible subjects produced the largest RRR compared to other metrics. The prolongation of urine collection time from 6 to 48 hr increased the RRR from 2.13 to 2.62.

Table IV-1. The evaluation of $Xu_{smk}, Xu_{oral}$ and $AUC_{Phet}$ in scenario 1

<table>
<thead>
<tr>
<th>Metrics</th>
<th>48-hr sample collection</th>
<th>6-hr sample collection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P%$</td>
<td>RRR</td>
</tr>
<tr>
<td>Top 20% $Xu_{smk}$</td>
<td>27.7</td>
<td>10.6</td>
</tr>
<tr>
<td>Top 20% $Xu_{oral}$</td>
<td>23.1</td>
<td>11.7</td>
</tr>
<tr>
<td>Top 20% $AUC_{Phet}$</td>
<td>24.6</td>
<td>11.4</td>
</tr>
</tbody>
</table>

NA: Not available

Table IV-2 reports the results of scenario 2, in which subjects in the top 20% of the population in terms of $A_{act,lung}$ were considered to be susceptible to lung cancer. Table IV-2 showed that the use of the top 20% $Xu_{lung}$ among the population as the metric to identify susceptible subjects produced the largest RRR compared to other metrics. The prolongation of urine collection time from 6 to 48 hr resulted in little increase in RRR.

Table IV-2. The evaluation of $Xu_{smk}, Xu_{oral}$ and $Xu_{lung}$ in scenario 2

<table>
<thead>
<tr>
<th>Metrics</th>
<th>48-hr sample collection</th>
<th>6-hr sample collection</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P%$</td>
<td>RRR</td>
</tr>
<tr>
<td>Top 20% $Xu_{smk}$</td>
<td>18.3</td>
<td>12.9</td>
</tr>
<tr>
<td>Top 20% $Xu_{oral}$</td>
<td>12.9</td>
<td>14.3</td>
</tr>
<tr>
<td>Top 20% $Xu_{lung}$</td>
<td>26.9</td>
<td>10.8</td>
</tr>
</tbody>
</table>
Figure IV-2. The plasma concentration-time profile of [D$_{10}$]PheT of a representative subject from the pilot study (top) and a simulated subject (bottom)
Figure IV-3. Cumulative amount of [D₁₀]PheT in urine-time profile of a representative subject from the pilot study (top) and a simulated subject (bottom)
### IV.D Discussion

The hypothesis of our research was that subjects with extensive activation have higher lung cancer risk than others. This hypothesis will be further tested in a large-scale clinical trial with 350 subjects receiving 10 µg \([\text{D}_{10}]\text{Phe}\) and bronchoscopy screening. According to the research proposal by Dr. Stephen Hecht (personal communication), one method to prove the hypothesis is the comparison of the amount of \([\text{D}_{10}]\text{PheT}\) formed between two groups: subjects with dysplasia (group A) and subjects with normal bronchoscopies (group B). If the hypothesis is correct, then group A should produce more \([\text{D}_{10}]\text{PheT}\) than group B. A second approach to prove the hypothesis is to compare the lung cancer risk measured by bronchoscopy between two groups: subjects with extensive activation (group 1) and the remaining subjects (group 2). If the hypothesis is correct, then the \(P\%\) in group 1 should be higher than that of group 2. Current simulations focused on the second method.

In the simulations, two scenarios were investigated. In scenario 1, \(X_u_{\text{smk}}\), \(AUC_{\text{smk}}\), \(X_u_{\text{oral}}\) were tested as surrogates of \(A_{\text{act,smk}}\) to identify susceptible subjects. According to equation IV-4 (derivation listed in Appendix IV-A), a subject whose \(A_{\text{act,smk}}\) metric is in the top 20% of the population will not necessarily be in the top 20% of the population with regards to \(X_u_{\text{smk}}\). In other words, whether \(X_u_{\text{smk}}\) can accurately identify subjects with large systemic exposure depends on two factors: 1) the urine collection time \((t)\); 2) the between-subject variability (BSV) of \(k_2\). One additional factor that might also influence the accuracy of the identification of susceptible subjects is the BSV of \(k_{1,\text{smk}}\). But due to the lack of sufficient early data in the pilot study, it cannot be estimated. As such, \(k_{1,\text{smk}}\) was fixed to be 10 hr\(^{-1}\) in the simulation. Only the influence of urine collection time and BSV of \(k_2\) were investigated.

\[
X_u_{\text{smk}} = A_{\text{act,smk}} \times \left[1 - \frac{k_{1,\text{smk}} e^{-k_2 t}}{k_{1,\text{smk}} - k_2}\right] \quad \text{Equation IV-4}
\]
Ideally, if the urine collection time is long enough \( t > 5 \) half-lives), then \( X_{u,smk} \) would equal to \( A_{act,smk} \). In that case, all the subjects assigned to group 1 by \( X_{u,smk} \) would be susceptible subjects and the \( P\% \) of groups 1 and 2 would be 30\% and 10\%, respectively. This would lead to the maximum \( RRR \) of 3. This explains the large value of \( RRR \) by \( X_{u,smk} \) (2.62) based on a 48-hr urine collection (Table IV-1).

Figure IV-4. An example of the effect of BSV of \( k_2 \) on \( X_{u,smk} \)

But it is not practical to collect urine samples for 5 half-lives of \([D_{10}]PheT\) in the large-scale clinical trial. The BSV of \( k_2 \) begins to influence the ability of \( X_{u,smk} \) to identify susceptible subjects as the urine collection time decreases. For example, Figure IV-4 shows the simulation of two subjects with \( A_{act,smk} \) of 3 and 1 nmol, and \( k_2 \) of 0.0578 and 0.3465 hr\(^{-1}\), respectively. The elimination half-life in the two subjects is 12 and 2 hr, respectively. A urine collection longer than 6 hr would correctly identify subject B as the one with larger systemic exposure (larger \( A_{act,smk} \)). If the urine collection time was only 6 hr, then the cumulative amount of \([D_{10}]PheT\) in the urine \( (X_{u,smk}) \) of subject A would be the same as that of subject B. If the urine collection time was less than 6 hr, then
would falsely identify subject A as the one with larger systemic exposure because subject A excreted [D$_{10}$]PheT more quickly into the urine in the first few hours. This explained why the RRR of $X_{u, smk}$ based on 6-hr urine collection was somewhat smaller than that based on 48-hr urine collection as shown in Table IV-1.

Unlike urine data, equation IV-5 (derivation listed in the Appendix IV-A) indicates that even if plasma data is collected from time 0 to infinity ($t > 5$ half-lives of [D$_{10}$]PheT) $AUC_{PheT}$ may not accurately identify susceptible subjects, whose criterion in scenario 1 was only based on $A_{act, smk}$. In other words, subjects with values of $AUC_{PheT}$ in the top 20% of the population will not necessarily have $A_{act, smk}$ in the top 20% due to the BSV of CL. A subject with small $A_{act, smk}$ can still have very large $AUC_{PheT}$ because of the low clearance and therefore be falsely identified as a susceptible subject. This explains the relatively low RRR of $AUC_{PheT}$ despite a 48-hr plasma collection. Since the use of 48-hr plasma data produced a similar RRR as 6-hr urine data (Table IV-1), the use of urine data was a better choice to identify susceptible subjects.

$$A_{act, smk} = CL \times AUC_{PheT} \quad \text{Equation IV-5}$$

The use of $X_{u, oral}$ to identify susceptible subjects was also investigated in scenario 1. Equation IV-6 (derivation listed in the Appendix IV-A) indicated that the accuracy of having $X_{u, oral}$ measurement in the top 20% as the indicator to identify subjects in the top 20% of the population in terms of $A_{act, smk}$ could be influenced by two factors: 1) the urine collection time ($t$); 2) the BSV of $k_2$. This is the same as the $X_{u, smk}$. However, the BSV of relative bioavailability ($F$) adds extra noise to $X_{u, oral}$, leading to a relatively low RRR of $X_{u, oral}$ compared to $X_{u, smk}$ as shown in Table IV-1.

$$X_{u, oral} = A_{act, smk} \times F \times \left[1 - \frac{k_{1, oral} e^{-k_2 t}}{k_{1, oral} - k_2}\right] \quad \text{Equation IV-6}$$

While systemic exposure ($A_{act, smk}$) seemed to be a reasonable indicator of an individual’s capacity to activate PAHs, local exposure might be physiologically more
relevant to lung carcinogenesis. Scenario 2 aimed to answer the question: what would be the outcome of the clinical trial if local exposure rather than systemic exposure is the driving force for lung carcinogenesis? Therefore, in the following simulation the lung cancer risk of subjects with substantial local exposure \( A_{act,lung} \) was assumed to be 3-fold of the remaining subjects.

In scenario 2, subjects in the top 20% of the population in terms of \( A_{act,lung} \) were considered to be susceptible to lung cancer. According to equation IV-7 (derivation listed in the Appendix IV-A), if local exposure is the driving force for lung carcinogenesis and oral dose alone is used in the large-scale trial, then the study is most likely to fail. The use of \( Xu_{oral} \) as the metric to identify subjects with substantial local exposure is based on the assumption that subjects with large \( Xu_{oral} \) also had large \( A_{act,lung} \). Unfortunately, this is not a reasonable assumption. The simulation showed that very few subjects with top 20% of \( Xu_{oral} \) values had \( Xu_{smk} \) values that would guarantee substantial local exposure (\( A_{act,lung} \)). Actually, subjects with large \( Xu_{oral} \) are more likely to have small \( A_{act,lung} \) and low local exposure. This explains why the \( P\% \) of group 1 was lower than that of group 2 (Table IV-2) when the top 20% of \( Xu_{oral} \) values based on 48-hr urine collection was used as the metric to identify susceptible subjects.

\[
A_{act,lung} = Xu_{smk} - Xu_{oral} \quad \text{Equation IV-7}
\]

\( Xu_{smk} \) might be a useful indicator of local exposure. According to equation IV-7 subjects in the top 20% of the population in terms of \( Xu_{smk} \) do not always have \( A_{act,lung} \) values in the top 20%. Nevertheless, as long as these subjects do not demonstrate extremely high \( Xu_{oral} \) values, they are likely to have fairly large \( A_{act,lung} \) values which could still be within the top 20% of the population. This explained the moderate difference in lung cancer risk between group 1 and 2 reflected by a \( RRR \) value of 1.4 when the top 20% of the \( Xu_{smk} \) measurement was used as a metric to identify susceptible subjects (Table IV-2).
$X_{u\text{lung}}$ is a better indicator of local exposure than $X_{u\text{oral}}$ and $X_{u\text{smk}}$. According to equation IV-8 (derivation listed in the Appendix IV-A), when urine collection time is long enough, $X_{u\text{lung}}$ equals to $A_{act,lung}$ and the use of $X_{u\text{lung}}$ can accurately identify subjects with substantial local exposure and presumably high lung cancer risk. This explains why the $P\%$ of group 1 was very close to the maximum value of 30$\%$ when the top 20$\%$ of $X_{u\text{lung}}$ values based on 48-hr urine collection was selected as the metric (Table IV-2). In addition, a 6-hr urine collection produced a similar $RRR$, which justifies the reduction of urine collection time from 48 hr to 6 hr.

\[ A_{act,lung} = X_{u\text{lung}} \] \hspace{1cm} Equation IV-8

**IV.E Conclusion**

The simulations indicated that if systemic exposure is the driving force for lung carcinogenesis, then $X_{u\text{smk}}$ is the best metric followed by $X_{u\text{oral}}$. Blood sampling is not recommended because $AUC_{\text{PheT}}$ does not provide more information than $X_{u\text{smk}}$ but requires invasive sampling. This recommendation might be modified if $[D_{10}]\text{Phe}$ itself was measurable in plasma.

If the local exposure is the driving force, then the use of both oral and smoking arms would provide most information about susceptible subjects. Oral dosing alone is not recommended because $X_{u\text{oral}}$ provided little information on local exposure. On the other hand, $X_{u\text{smk}}$ seems to be a useful metric for local exposure.

Collectively, since it is unknown whether systemic or local exposure is the driving force, the smoking arm is necessary in the clinical trial. The addition of one oral arm does not lead to a substantial increase of resources required, and therefore a 6-hr urine collection of both arms is highly recommended. The use of oral arm alone is not suggested because if the local exposure is the driving force for lung carcinogenesis, then $X_{u\text{act,oral}}$ might not be an appropriate metric to identify subjects with substantial local exposure and presumably high lung cancer risk.
CHAPTER V  Conclusion

In the present study the diol epoxide metabolic activation of [D_{10}]Phe in 25 smokers has been investigated. The pharmacokinetic (PK) study demonstrated that in both the oral and smoking arms approximately 6% of the dose was metabolically converted to diol epoxides, with a large inter-subject variability in the formation of [D_{10}]PheT observed. A two-step diagnostic plot approach has been developed to identify subjects with substantial lung exposure. This approach may have significant potential in the prediction of lung cancer risk. The integration of phenotyping and genotyping results indicated that GSTM1 played an important role in the detoxification of Phe.

In addition to the two-step diagnostic plot approach, a different method, the population (Pop) PK approach, was also tried to identify subjects with extensive activation but was not completely successful. The failure to obtain final PK estimates might be due to the lack of appropriate data (e.g. early blood samples taken during the smoking administration), unreasonable assumptions of the Pop PK model or both. The integration of the plasma and urine data of the parent compound, if available, might provide an opportunity to further develop the Pop PK model and identify subjects with substantial local exposure. The preliminary statistical analysis of the population data indicated an association between renal function and the measurement of an individual’s capacity to activate PAHs. This might need to be considered in the design of the future large-scale clinical trial. At least creatinine concentration in the plasma should be measured during the screening visit as was conducted in the pilot PK study. This would allow an estimate of an individual’s creatinine clearance.

The simulation of the diol epoxide metabolic activation of Phe in 350 subjects recommended that both oral and smoking arms should be used in the future large-scale trial in order to prove the hypothesized correlation between extensive activation and metabolic activation, although the use of the smoking arm alone may also provide some information on the correlation. The use of the oral arm alone is not suggested. The simulation also indicated that a 6-hr urine collection would be sufficient to identify
subjects with extensive activation and presumably high lung cancer risk. The collection of plasma samples during clinical study is not necessary unless an assay for the parent compound [D_{10}]Phe in plasma is developed. But the plasma samples collected at the screening visit should be carefully stored in order to measure genotypes and the creatinine clearance, which could further confirm the results discovered in the pilot study.
Bibliography

References Chapter I


12. Tammemagi CM, Pinsky PF, Caporaso NE, Kvale PA, Hocking WG, Church TR, Riley TL, Commins J, Oken MM, Berg CD, Prorok PC. *Lung cancer risk prediction: Prostate, Lung, Colorectal And Ovarian Cancer Screening Trial models and validation.* J Natl Cancer Inst 2011; 103:1058-68


38. Yang SK, McCourt DW, Roller PP, Gelboin HV. Enzymatic conversion of benzo(a)pyrene leading predominantly to the diol-epoxide r-7,t-8-dihydroxy-t-9,10-oxy-7,8,9,10-tetrahydrobenzo(a)pyrene through a single enantiomer of r-7, t-8-dihydroxy-7,8-dihydrobenzo(a)pyrene. Proc Natl Acad Sci USA 1976: 2594-8


44. Choudhary D, Jansson I, Stoilov I, Sarfarazi M, Schenkman JB. Expression patterns of mouse and human CYP orthologs (families 1-4) during development and in different adult tissues. Arch Biochem Biophys 2005; 436:50-61


50. Bouchard M, Viau C. *Urinary excretion kinetics of pyrene and benzo[a]pyrene metabolites following intravenous administration of the parent compounds or the metabolites.* Toxicol Appl Pharmacol 1996; 139:301-9


54. Upadhyaya P, Hochalter JB, Balbo S, McIntee EJ, Hecht SS. *Preferential glutathione conjugation of a reverse diol epoxide compared to a bay region diol epoxide of benzo[a]pyrene in human hepatocytes.* Drug Metab Dispos 2010; 38:1397-402


70. Schober W, Pusch G, Oeder S, Reindl H, Behrendt H, Buters JT. Metabolic activation of phenanthrene by human and mouse cytochromes P450 and pharmacokinetics in CYP1A2 knockout mice. Chem Biol Interact 2010; 183:57-66


97. Smith GB, Harper PA, Wong JM, Lam MS, Reid KR, Petsikas D, Massey TE.


99. Sherson D, Sigsgaard T, Overgaard E, Loft S, Poulsen HE, Jongeneelen FJ.


102. Hecht SS, Chen M, Yoder A, Jensen J, Hatsu kami D, Le C, Carmella SG.
*Longitudinal study of urinary phenanthrene metabolite ratios: effect of smoking on the diol epoxide pathway.* Cancer Epidemiol Biomarkers Prev 2005; 14:2969-74

103. Kim JH, Sherman ME, Curriero FC, Guengerich FP, Strickland PT, Sutter TR.
*Expression of cytochromes P450 1A1 and 1B1 in human lung from smokers, non-smokers, and ex-smokers.* Toxicol Appl Pharmacol 2004; 199:210-9
References Chapter II


**References Chapter III**


**References Chapter IV**

Appendix II-A

In the PK approach to identify susceptible subjects, one primary objective was to quantitate the \([\text{D}_{10}]\text{PheT}\) formed and use it as a surrogate for the amount of “active” metabolite \([\text{D}_{10}]\text{PheDE}\) formed during activation \((A_{act})\). This is an indicator of systemic exposure. As shown in Figure I-5, \([\text{D}_{10}]\text{PheDE}\) is the surrogate for the active intermediate BPDE, which binds to DNA and initiates carcinogenesis. However, direct measurement of \([\text{D}_{10}]\text{PheDE}\) was very difficult because of the poor stability of this intermediate. Our investigation turned to \([\text{D}_{10}]\text{PheT}\), the end product of activation pathway. Based on our understanding of the metabolic activation of Phe as shown in Figure II-1, \(A_{act}\) could be estimated by quantitating \([\text{D}_{10}]\text{PheT}\) in plasma or urine.

The absorption and disposition of \([\text{D}_{10}]\text{Phe}\) after extravascular administration (oral or smoking) could be summarized in 4 steps (Figure II-1). Step 1: after delivery to the site of absorption, \([\text{D}_{10}]\text{Phe}\) was transferred to systemic circulation intact (absorption), in the form of \([\text{D}_{10}]\text{PheT}\) (first-pass activation), and in the form of other metabolites; the clearances of these three pathways of step 1 were \(CL_1\), \(CL_2\) and \(CL_3\), respectively. Step 2: after entering the systemic circulation, a small fraction of \([\text{D}_{10}]\text{Phe}\) was converted to “active” metabolite \([\text{D}_{10}]\text{PheDE}\) (activation) and the majority of \([\text{D}_{10}]\text{Phe}\) was transformed to other metabolites; the clearances of these two pathways of step 2 were \(CL_5\) and \(CL_4\), respectively. Step 3: \([\text{D}_{10}]\text{PheDE}\) was converted to \([\text{D}_{10}]\text{PheT}\) (hydrolysis) and the clearance of this process was \(CL_6\); Step 4: \([\text{D}_{10}]\text{PheT}\) was further converted to sulfate or glucuronide conjugates followed by renal excretion. The clearance of the process that eliminates \([\text{D}_{10}]\text{PheT}\) from the body was \(CL_7\).

Several assumptions were made in step 3 and 4, respectively. The first assumption is that in step 3 all \([\text{D}_{10}]\text{PheDE}\) was converted to \([\text{D}_{10}]\text{PheT}\). The second assumption is that in step 4 all the \([\text{D}_{10}]\text{PheT}\) formed was excreted in urine in 48 hr.

As such, the following differential equations were derived
\[
\frac{dA_{(\text{Phe}_{\text{abs}})}}{dt} = -(CL_1 + CL_2 + CL_3) \times C_{(\text{Phe}_{\text{abs}})} (\text{II-A1})
\]

\[
\frac{dA_{(\text{Phe})}}{dt} = CL_3 \times C_{(\text{Phe}_{\text{abs}})} - (CL_4 + CL_5) \times C_{(\text{Phe})} (\text{II-A2})
\]

\[
\frac{dA_{(\text{PheDE})}}{dt} = CL_2 \times C_{(\text{Phe}_{\text{abs}})} + CL_5 \times C_{(\text{Phe})} - CL_6 \times C_{(\text{PheDE})} (\text{II-A3})
\]

\[
\frac{dA_{(\text{PheT})}}{dt} = CL_6 \times C_{(\text{PheDE})} - CL_7 \times C_{(\text{PheT})} (\text{II-A4})
\]

\[
\frac{dXu_{(\text{PheT})}}{dt} = CL_7 \times C_{(\text{PheT})} (\text{II-A5})
\]

Equation II-A1 could also be written as

\[
dA_{(\text{Phe}_{\text{abs}})} = -(CL_1 + CL_2 + CL_3) \times C_{(\text{Phe}_{\text{abs}})}dt (\text{II-A6})
\]

The integration of equation II-A6 with regard to time from \( t = 0 \) to \( t = \infty \) provides the following expression:

\[
\int_{A_{(\text{Phe}_{\text{abs}}), t = 0}}^{A_{(\text{Phe}_{\text{abs}}), t = \infty}} dA_{(\text{Phe}_{\text{abs}})} = -(CL_1 + CL_2 + CL_3) \int_0^\infty C_{(\text{Phe}_{\text{abs}})}dt
\]

\( A_{(\text{Phe}_{\text{abs}})} \) integrated between \( t = 0 \) and \( t = \infty \) is equal to the dose of \([D_{10}]\text{Phe} \) (D).

Therefore, \( D = (CL_1 + CL_2 + CL_3) \times AUC_{(\text{Phe}_{\text{abs}})} \)

\[
AUC_{(\text{Phe}_{\text{abs}})} = \frac{D}{CL_1 + CL_2 + CL_3} (\text{II-A7})
\]

Similarly, the integration of equations II-A2, II-A3, II-A4 and II-A5 from \( t = 0 \) to \( t = \infty \) provides the following expressions:
\[ CL_A \times AUC_{\text{Phe,abs}} = (CL_A + CL_S) \times AUC_{\text{Phe}} \]  \hspace{1cm} (II-A8)

\[ CL_S \times AUC_{\text{Phe,abs}} + CL_S \times AUC_{\text{Phe}} = CL_S \times AUC_{\text{PheDE}} \]  \hspace{1cm} (II-A9)

\[ CL_S \times AUC_{\text{PheDE}} = CL_T \times AUC_{\text{PheT}} \]  \hspace{1cm} (II-A10)

\[ Xu_{\text{(PheT), } t=\infty} = CL_T \times AUC_{\text{PheT}} \]  \hspace{1cm} (II-A11)

It is assumed that \( Xu_{\text{(PheT), } t=\infty} \approx Xu_{\text{(PheT), } t=48} \); that is, almost all of the \([D_{10}]\text{PheT}\) formed is excreted into the urine in 48 hours.

The integration of equation II-A3 from \( t = 0 \) to \( t = T \) (\( 0 < T < \infty \)) provides the following expressions:

\[ A_{\text{(PheDE), } t=T} - A_{\text{(PheDE), } t=0} = CL_2 \times AUC_{0}^{T \text{(Phe, abs)}} + CL_S \times AUC_{0}^{T \text{(Phe)}} - CL_S \times AUC_{0}^{T \text{(PheDE)}} \]

Since \( A_{\text{(PheDE), } t=0} \) equals 0, then

\[ A_{\text{(PheDE), } t=T} = CL_2 \times AUC_{0}^{T \text{(Phe, abs)}} + CL_S \times AUC_{0}^{T \text{(Phe)}} - CL_S \times AUC_{0}^{T \text{(PheDE)}} \]  \hspace{1cm} (II-A12)

Similarly, the integration of equations II-A4 and II-A5 from \( t = 0 \) to \( t = T \) could result in equations II-A13 and II-A14:

\[ A_{\text{(PheT), } t=T} = CL_S \times AUC_{0}^{T \text{(PheDE)}} - CL_T \times AUC_{0}^{T \text{(PheT)}} \]  \hspace{1cm} (II-A13)

\[ Xu_{\text{(PheT), } t=T} = CL_T \times AUC_{0}^{T \text{(PheT)}} \]  \hspace{1cm} (II-A14)

At time \( T \), the \([D_{10}]\text{PheDE}\) that has been produced from activation can exist in three different forms: \([D_{10}]\text{PheDE}\) in the body, \([D_{10}]\text{PheT}\) in the body and \([D_{10}]\text{PheT}\) in urine. Therefore, the amount of \([D_{10}]\text{PheDE}\) formed at time \( T \) is the sum of \( A_{\text{(PheDE), } t=T} \), \( A_{\text{(PheT), } t=T} \) and \( Xu_{\text{(PheT), } t=T} \).

Combining equations II-A12, II-A13 and II-A14
\[ A_{(PheDE),t=T} + A_{(PheT),t=T} + Xu_{(PheT),t=T} = CL_2 \times AUC_T^{(Phe, abs)} + CL_5 \times AUC_T^{(Phe)} \]  
(II-A15)

When time T approaches infinity, equation II-A15 can be written as

\[ A_{(PheDE),t=\infty} + A_{(PheT),t=\infty} + Xu_{(PheT),t=\infty} = CL_2 \times AUC_{(Phe, abs)} + CL_5 \times AUC_{(Phe)} \]  
(II-A16)

By definition, the left side of equation II-A16 is equal to \( A_{act} \):

\[ A_{act} = A_{(PheDE),t=\infty} + A_{(PheT),t=\infty} + Xu_{(PheT),t=\infty} \]  
(II-A17)

Combining equations II-A16 and II-A17 to obtain equation II-A18:

\[ A_{act} = CL_2 \times AUC_{(Phe, abs)} + CL_5 \times AUC_{(Phe)} \]  
(II-A18)

Substituting equations II-A7 and II-A8 into II-A18 to obtain equation II-A19

\[ A_{act} = \frac{D}{CL_1 + CL_2 + CL_3} + CL_5 \times \frac{CL_1}{CL_4 + CL_5} \times \frac{D}{CL_1 + CL_2 + CL_3} \]  
(II-A19)

If \( f_a \), \( f_{act} \) and \( f_m \) are defined as the following

\[ f_a = \frac{CL_1}{CL_1 + CL_2 + CL_3} \]  
(II-A20)

\[ f_{act} = \frac{CL_2}{CL_1 + CL_2 + CL_3} \]  
(II-A21)

\[ f_m = \frac{CL_5}{CL_4 + CL_5} \]  
(II-A22)

then

\[ A_{act} = [(f_a \times f_m) + f_{act}] \times D \]  
(II-A23)
where $f_a$, $f_{act}$ and $f_m$ are the fraction of the dose absorbed, the fraction of the dose converted to $[\text{D}_{10}]\text{PheDE}$ during first-pass activation and the fraction of $[\text{D}_{10}]\text{Phe}$ converted to $[\text{D}_{10}]\text{PheDE}$ in the systemic circulation, respectively.

Since $A_{(\text{PheDE})_{t=\infty}}$ and $A_{(\text{PheT})_{t=\infty}}$ both equal 0, equation II-A17 could be reduced to

$$A_{act} = X\mu_{(\text{PheT})_{t=\infty}} \quad (\text{II-A24})$$

In addition, it is assumed that all the $[\text{D}_{10}]\text{PheT}$ formed was excreted in urine in 48 hr. Therefore,

$$A_{act} = X\mu_{(\text{PheT})_{t=48}} \quad (\text{II-A25, equation II-3 in text})$$

The combination of equations II-A25 and II-A11 leads to equation A26

$$A_{act} = AUC_{(\text{PheT})} \times CL_7 \quad (\text{II-A26, equation II-2 in text})$$

Abbreviations (all mass and concentration units are on a molar basis)

- $A_{act}$: total amount of $[\text{D}_{10}]\text{PheDE}$ formed during metabolic activation
- $A_{(\text{Phe,abs})}$: amount of $[\text{D}_{10}]\text{Phe}$ at the site of absorption
- $A_{(\text{Phe,abs})_{t=0}}$: amount of $[\text{D}_{10}]\text{Phe}$ at the site of absorption at time 0
- $A_{(\text{Phe,abs})_{t=T}}$: amount of $[\text{D}_{10}]\text{Phe}$ at the site of absorption at time T
- $A_{(\text{PheDE})}$: amount of $[\text{D}_{10}]\text{Phe}$ in the body
- $A_{(\text{PheDE})_{t=0}}$: amount of $[\text{D}_{10}]\text{PheDE}$ in the body at time 0
- $A_{(\text{PheDE})_{t=T}}$: amount of $[\text{D}_{10}]\text{PheDE}$ in the body at time T
- $A_{(\text{PheDE})_{t=\infty}}$: amount of $[\text{D}_{10}]\text{PheDE}$ in the body at time infinity
$A_{(PheT)}$: amount of $[D_{10}]PheT$ in the body

$A_{(PheT),t=0}$: amount of $[D_{10}]PheT$ in the body at time 0

$A_{(PheT),t=T}$: amount of $[D_{10}]PheT$ in the body at time T

$A_{(PheT),t=\infty}$: amount of $[D_{10}]PheT$ in the body at time infinity

$AUC_{(Phe,abs)}$: area under $C_{(Phe,abs)}$-time curve from time 0 to infinity

$AUC_{(Phe,abs)}^T$: area under $C_{(Phe,abs)}$-time curve from time 0 to T

$AUC_{(Phe)}$: area under $C_{(Phe)}$-time curve from time 0 to infinity

$AUC_{(Phe)}^T$: area under $C_{(Phe)}$-time curve from time 0 to T

$AUC_{(PheDE)}$: area under $C_{(PheDE)}$-time curve from time 0 to infinity

$AUC_{(PheDE)}^T$: area under $C_{(PheDE)}$-time curve from time 0 to T

$AUC_{(PheT)}$: area under $C_{(PheT)}$-time curve from time 0 to infinity

$AUC_{(PheT)}^T$: area under $C_{(PheT)}$-time curve from time 0 to T

$C_{(Phe,abs)}$: concentration of $[D_{10}]Phe$ at the site of absorption

$C_{(Phe)}$: plasma concentration of $[D_{10}]Phe$

$C_{(PheDE)}$: plasma concentration of $[D_{10}]PheDE$

$C_{(PheT)}$: plasma concentration of $[D_{10}]PheT$

$f_d$: the fraction of dose absorbed

$f_{act}$: the fraction of dose converted to $[D_{10}]PheDE$ during first-pass activation
$f_{m}$: the fraction of $[D_{10}]$Phe converted to $[D_{10}]$PheDE in the systemic circulation

$X_{U(Phe)}$: cumulative amount of $[D_{10}]$PheT collected in the urine

$X_{U(PheT), t=T}$: amount of $[D_{10}]$PheT collected in the urine at time $T$

$X_{U(PheT), t=48}$: amount of $[D_{10}]$PheT collected in the urine at 48 hr

$X_{U(PheT), t=\infty}$: amount of $[D_{10}]$PheT collected in the urine at time infinity
Appendix III-A

A subset of the NONMEM® dataset was listed as the following:

**Table III-A-1.** An example of the NONMEM® dataset

<table>
<thead>
<tr>
<th>C</th>
<th>ID</th>
<th>Route</th>
<th>Time</th>
<th>DV</th>
<th>Data.type</th>
<th>EVID</th>
<th>AMT</th>
<th>CMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>53200</td>
<td>1</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>.</td>
<td>1</td>
<td>1</td>
<td>53200</td>
<td>2</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0.25</td>
<td>142</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>121</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>1215.5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0.5</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>0.75</td>
<td>127.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>120.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>620.975</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>110.9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>99</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>265.125</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>2.5</td>
<td>91.2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>73.8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>59.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>2025.83</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>36.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>1488.24</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>18.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>1387.8</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>24</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>48</td>
<td>1001</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>48</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>-4</td>
</tr>
<tr>
<td>.</td>
<td>1</td>
<td>1</td>
<td>48</td>
<td>.</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>
DV: dependable variable
The units of DV: pmol/L for plasma data and pmol for urine data
Data.type: 1 for plasma data and 2 for urine data
EVID: 0 for observation, 1 for dosing event and 2 for urine collection
CMT: the code of the compartment; a negative value indicates that the compartment is
turned off
AMT: the dose (pmol)

The control file of the NONMEM® model for the smoking arm was listed as the
following:

;Model Desc: PheT, two dosing CMT
$PROB RUN# J135
$INPUT C ID Route TIME DV Datatype EVID AMT CMT
$DATA d20.csv IGNORE=C
$SUBROUTINE ADVAN6
$MODEL COMP=(pathwayI) COMP=(pathwayII) COMP=(PheT) COMP=(Xu)

$PK
;FR1- fraction of PheT from pathway I
;FR2- fraction of PheT from pathway II
;K1- input rate constant of pathway I
;K2- input rate constant of pathway II
;V- volume of distribution of PheT
;CL- clearance of PheT

;-----------------------------------------------------------------------
TVFR1=THETA(1)
FR1=TVFR1*EXP(ETA(1))
TVFR2=THETA(2)
FR2=TVFR2*EXP(ETA(2))
TVK1=THETA(3)
K1=TVK1*EXP(ETA(3))
TVK2=THETA(4)
K2=TVK2*EXP(ETA(4))
TVV=THETA(5)
V=THETA(5)*EXP(ETA(5))
S3=V
S4=1
TVCL=THETA(6)
CL=THETA(6)*EXP(ETA(6))
F1=FR1
F2=FR2

$DES$
DADT(1)=-K1*A(1)
DADT(2)=-K2*A(2)
DADT(3)=K1*A(1)+K2*A(2)-(CL*A(3)/V)
DADT(4)=CL*A(3)/V

$ERROR$
IF (Datatype.EQ.1) THEN
Y=(A(3)/V)+(A(3)/V)*EPS(1)+EPS(3)
ELSE
Y=A(4)+A(4)*EPS(2)+EPS(4)
ENDIF
$\text{STHETA}$

$(0,0.03,0.2) \quad \text{[TVFR1]}$

$(0,0.03,0.2) \quad \text{[TVFR2]}$

$1000 \text{ FIX} \quad \text{[TVK1 (/hr)]}$

$(0.1,2) \quad \text{[TVK2 (/hr)]}$

$(3,50) \quad \text{[TVV (L)]}$

$(0,4) \quad \text{[TVCL (L/hr)]}$

$\text{SOMEGA}$

0.01;

0.01;

0 FIX;

0 FIX;

0 FIX;

0 FIX;

$\text{SIGMA}$

0.1 ;

0.1 ;

0.1 ;

0.1 ;

$\text{SESTIMATION METHOD=1 MAXEVAL=9999 PRINT=5 SIG=2 MSFO=J135.MSF NOABORT}$

$\text{SCOV}$
<table>
<thead>
<tr>
<th>ID</th>
<th>Route</th>
<th>TIME</th>
<th>Datatype</th>
<th>DV PRED</th>
<th>FR1</th>
<th>FR2</th>
<th>V CL</th>
<th>CWRES</th>
<th>Datatype</th>
<th>EVID</th>
<th>FILE=J135.TAB</th>
</tr>
</thead>
</table>
Appendix IV-A

STELLA® 9.1.3 (isee systems, inc.; Lebanon, NH) and R 2.14.1 were used for simulation and data analysis, respectively. As shown in Figure IV-5, two models were built in STELLA® to simulate the ability to quantitate the diol epoxide metabolic activation of [D\textsubscript{10}]Phe after smoking (top) and oral dosing (bottom), respectively.

**Figure IV-5.** Stella model of metabolic activation of [D\textsubscript{10}]Phe after smoking (top) and oral dosing (bottom)
The abbreviations of PK parameters and symbols used in the simulation are as follows:

\( A_{act, smk} \): the total amount of \([D_{10}]Phe\) activated after the administration of \([D_{10}]Phe\) by smoking (nmol)

\( A_{act, oral} \): the total amount of \([D_{10}]Phe\) activated after the administration of \([D_{10}]Phe\) by oral dosing (nmol)

\( A_{act, lung} \): the difference between \(A_{act, smk}\) and \(A_{act, oral}\) (nmol)

\( Depo_{smk} \): the amount of \([D_{10}]Phe\) at the site of administration after smoking (nmol)

\( Depo_{oral} \): the amount of \([D_{10}]Phe\) at the site of administration after oral dosing (nmol)

\( PheT_{smk} \): the amount of \([D_{10}]PheT\) in the body after smoking (nmol)

\( PheT_{oral} \): the amount of \([D_{10}]PheT\) in the body after oral dosing (nmol)

\( Ln A_{act, smk} \): the natural logarithmic (log) transformation of \(A_{act, smk}\)

\( Ln A_{act, oral} \): the natural log transformation of \(A_{act, oral}\)

\( F \): the relative bioavailability calculated by the ratio of \(A_{act, oral}\) to \(A_{act, smk}\)

\( Ln F \): the natural log transformation of \(F\)

\( k_{1, smk} \): the input rate constant that describes the transformation of parent compound at the site of absorption to \([D_{10}]PheT\) in the blood after smoking (hr\(^{-1}\))

\( k_{1, oral} \): the input rate constant that describes the transformation of parent compound at the site of absorption to \([D_{10}]PheT\) in the blood after oral dosing (hr\(^{-1}\))

\( k_2 \): the elimination rate constant of \([D_{10}]PheT\) (hr\(^{-1}\))

\( V \): the volume of distribution of \([D_{10}]PheT\) (L)

\( C_{PheT, smk} \): the plasma concentration of \([D_{10}]PheT\) after smoking (pmol/L)
\( C_{\text{PheT, oral}} \): the plasma concentration of [D\textsubscript{10}]PheT after oral dosing (pmol/L)

\( AUC_{\text{PheT}} \): the area under the curve of [D\textsubscript{10}]PheT from time 0 to infinity after the administration of [D\textsubscript{10}]Phe by smoking (pmol*hr/L)

\( CL \): the clearance of [D\textsubscript{10}]PheT (L/hr)

\( \ln CL \): the natural log transformation of CL

\( Xu_{\text{smk}} \): the cumulative amount of [D\textsubscript{10}]PheT excreted in the urine at time t after smoking (nmol)

\( Xu_{\text{oral}} \): the cumulative amount of [D\textsubscript{10}]PheT excreted in the urine at time t after oral dosing (nmol)

\( Xu_{\text{lung}} \): the difference between \( Xu_{\text{smk}} \) and \( Xu_{\text{oral}} \) (nmol)

Since diol epoxide metabolic activation was the focus of current study, the dose used in the simulation was not the total dose (10 µg or 53.2 nmol) but the fraction of [D\textsubscript{10}]Phe dose that was converted to [D\textsubscript{10}]PheT after smoking and oral dosing. In the pilot study as discussed in chapter II, 3.06 ± 1.91 nmol [D\textsubscript{10}]Phe was activated after smoking, which was 0.875 ± 0.81 after natural log transformation. Therefore, in the smoking model (Figure IV-5, top) \( \ln A_{\text{act, smk}} \) was set to be 0.875 with a standard deviation of 0.81. In addition, the pilot study indicated a relative bioavailability of 1.35 ± 0.96, which was 0.136 ± 0.536 after natural log transformation. Hence in the oral model (Figure IV-5, bottom) \( \ln F \) was set to be 0.136 with the standard deviation of 0.536. \( A_{\text{act, oral}} \) was determined by the product of \( F \) and \( A_{\text{act, smk}} \cdot k_{1, \text{smk}} \) and \( k_{1, \text{oral}} \) were set to be 10 and 2 hr\textsuperscript{-1} in the simulation respectively.

The pilot study indicated that the clearance of [D\textsubscript{10}]PheT was estimated to be 5.2 ± 3.0 L/hr after the administration of [D\textsubscript{10}]Phe by smoking, which was 1.51 ± 0.55 after natural log transformation. Accordingly, in the smoking model \( \ln CL \) was set to be 1.51 with a
standard deviation of 0.55. \( CL \) was set to be the same in both oral and smoking arms. Volume of distribution of \([D_{10}]\text{PheT}\) was set to be 40 L in both smoking and oral arms.

The simulation of 350 subjects was achieved through the sensitivity function of STELLA\textsuperscript{\textregistered}. \( A_{act,smk} \), \( F \) and \( CL \) were three variables that followed log-normal distribution during the 350 sensitivity runs. The distributions of these three variables (means and standard deviations) were based on the PK analysis of the pilot study reported in Chapter II. The remaining PK parameters including \( k_{1,smk} \), \( k_{1,oral} \) and \( V \) were set to be constant. In addition, assay error was also introduced with a CV\% of 15\% for plasma data \((C_{\text{PheT,smk}}\) and \( C_{\text{PheT,oral}}\)) and 10\% for urinary data \((X_{\text{oral}}\) and \( X_{\text{smk}}\)). A total of 350 runs were carried out to simulate the PK profiles of 350 subjects. For each run, a random value of \( A_{act,smk} \) was selected from the predetermined log-normal distribution by STELLA\textsuperscript{\textregistered}. At the same time, random values of \( F \) and \( CL \) were also generated in the similar way. As such, each subject was assigned a unique set of \( A_{act,smk} \), \( F \) and \( CL \) values. Then the simulation was initiated to simulate the plasma and urine data of one subject. After the completion of one single run, the PK profile of one simulated subject was generated. The completion of 350 runs produced the PK profiles of 350 simulated subjects.

According to the smoking model of Figure IV-5 (top), the following equations could be written:

\[
\frac{d Depo_{smk}}{dt} = -k_{1,smk} \times Depo_{smk} \quad \text{(IV-A1)}
\]

\[
\frac{d PheT_{smk}}{dt} = k_{1,smk} \times Depo_{smk} - CL \times C_{\text{PheT, smk}} \quad \text{(IV-A2)}
\]

\[
\frac{d Xu_{smk}}{dt} = CL \times C_{\text{PheT, smk}} \quad \text{(IV-A3)}
\]
In addition, elimination constant \( k_2 \) is defined as:

\[
k_2 = \frac{CL}{V}
\]  

(IV-A4)

Laplace transforms were used to solve these differential equations and to obtain the following expression:

\[
X_{u_{smk}} = A_{act,smk} \times \left[ 1 + \frac{k_2 e^{-k_{1,smk} t}}{k_{1,smk} - k_2} - \frac{k_{1,smk} e^{-k_2 t}}{k_{1,smk} - k_2} \right]
\]  

(IV-A5)

Similarly, the following expression could also be derived:

\[
X_{u_{oral}} = A_{act,oral} \times \left[ 1 + \frac{k_2 e^{-k_{1,oral} t}}{k_{1,oral} - k_2} - \frac{k_{1,oral} e^{-k_2 t}}{k_{1,oral} - k_2} \right]
\]  

(IV-A6)

In equation IV-A5, approximately 0.35 hr after dosing, the term \( \frac{k_2 e^{-k_{1,smk} t}}{k_{1,smk} - k_2} \) approaches 0 as \( k_{1,smk} \) was set to be 10 hr\(^{-1}\). The high value for \( k_{1,smk} \) was supported by the pilot study which showed the absence of a rising part (“absorption” phase) in the plasma concentration - time profile of [D10]PheT of the smoking arm. Equation IV-A5 can then be further simplified to:

\[
X_{u_{smk}} = A_{act,smk} \times \left[ 1 - \frac{k_{1,smk} e^{-k_2 t}}{k_{1,smk} - k_2} \right]
\]  

(IV-A7, equation III-4 in text)

Equation IV-A3 could be rearranged as the following:

\[
d X_{u_{smk}} = CL \times C_{PheT, smk} \ dt
\]  

(IV-A8)

The integration of equation IV-A8 leads to equation IV-A9:
\[
\int_{Xu_{smk} \text{ at } t=0}^{Xu_{smk} \text{ at } t=\infty} dXu_{smk} = CL \times \int_{t=0}^{t=\infty} C_{PheT, smk} \, dt
\]  

(IV-A9)

Since \( Xu_{smk} \) equals 0 and \( A_{act,smk} \) at time 0 and infinity, respectively, equation IV-A9 could be written as:

\[
A_{act,smk} = CL \times AUC_{PheT}
\]  

(IV-A10, equation III-5 in text)

The relative bioavailability was defined in equation IV-A11:

\[
F = \frac{A_{act,oral}}{A_{act,smk}}
\]  

(IV-A11)

The combination of equations IV-A6 and IV-A11 led to equation IV-A12.

\[
Xu_{oral} = A_{act,smk} \times F \times \left[1 + \frac{k_2 e^{-k_{1,oral} t}}{k_{1,oral} - k_2} - \frac{k_{1,oral} e^{-k_2 t}}{k_{1,oral} - k_2}\right]
\]  

(IV-A12)

In equation IV-A12, approximately 1.7 hr after dosing the term \( \frac{k_2 e^{-k_{1,oral} t}}{k_{1,oral} - k_2} \) approaches 0 as \( k_{1,oral} \) was set to 2 hr\(^{-1}\). The value of 2 hr\(^{-1}\) of \( k_{1,oral} \) used in the simulation was supported by the pilot study which showed the maximum level of [D\(_{10}\)]PheT in plasma at 45 - 60 min after oral dosing of [D\(_{10}\)]Phe. Equation IV-A12 can then be further simplified to equation IV-A13.

\[
Xu_{oral} = A_{act,smk} \times F \times \left[1 - \frac{k_{1,oral} e^{-k_2 t}}{k_{1,oral} - k_2}\right]
\]  

(IV-A13, equation III-6 in text)

As discussed in Chapter II, \( A_{act,lung} \) is the indicator of local exposure defined by equation IV-A14:
\[ A_{act,\text{lung}} = A_{act,\text{smk}} - A_{act,\text{oral}} \]  \hspace{1cm} (IV-A14)

The combination of equations IV-A7, IV-A11, IV-A13 and IV-A14 led to the following expression:

\[
A_{act,\text{lung}} = \frac{Xu_{act,\text{smk}}}{k_{1,\text{smk}} - k_2} \frac{e^{-k_2 t}}{1 - k_1,\text{smk} e^{-k_2 t}} - \frac{Xu_{act,\text{oral}}}{k_{1,\text{oral}} - k_2} \frac{e^{-k_2 t}}{1 - k_1,\text{oral} e^{-k_2 t}} \]  \hspace{1cm} (IV-A15)

When urine collection time \( t \) reaches infinity, then equation IV-A15 can be further simplified to equation IV-A16:

\[ A_{act,\text{lung}} = Xu_{\text{smk}} - Xu_{\text{oral}} \]  \hspace{1cm} (IV-A16, equation III-7 in text)

By definition,

\[ Xu_{\text{lung}} = Xu_{\text{smk}} - Xu_{\text{oral}} \]  \hspace{1cm} (IV-A17)

The combination of equations IV-A16 and IV-A17 led to:

\[ A_{act,\text{lung}} = Xu_{\text{lung}} \]  \hspace{1cm} (IV-A18, equation III-8 in text)