2006

Assessing, Modifying, and Combining Data Fields from the Virginia Office of the Chief Medical Examiner (OCME) Dataset and the Virginia Department of Forensic Science (DFS) Datasets in Order to Compare Concentrations of Selected Drugs

Amy Elizabeth Herrin
Virginia Commonwealth University

ASSESSING, MODIFYING, AND COMBINING DATA FIELDS FROM THE VIRGINIA OFFICE OF THE CHIEF MEDICAL EXAMINER (OCME) DATASET AND THE VIRGINIA DEPARTMENT OF FORENSIC SCIENCE (DFS) DRIVING UNDER THE INFLUENCE OF DRUGS (DUI) DATASETS IN ORDER TO COMPARE CONCENTRATIONS OF SELECTED DRUGS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

AMY ELIZABETH HERRIN
B.S., Emory & Henry College, 2004

Director:
R.K. Elswick, Jr., Ph.D.
Associate Professor and Interim Chair
Department of Biostatistics

Virginia Commonwealth University
Richmond, Virginia
August 2006

Acknowledgements

I would like to thank Dr. Elswick for his support and guidance during the past two years in and out of school. I would like to thank Dr. Saady for giving me access to the data, for being willing to help with my questions, and showing an interest in my life.

I want to thank my parents, Susan and Scott Herrin, for believing in me and loving me for the past 24 years as I tried to decide what I wanted to do and where I wanted to go. Mom, thanks for always being there and offering words (or emails) of support and strength. Dad, thanks for picking up the phone in the middle of the night for me to cry when I felt like I would never finish. I love you both so much and could not have done it without you! You both have been there for me, and I know that will never change. Audrey, Lisa, and Rachel, you all have been a constant source of laughter, and I want to thank you for listening to me when I needed to vent, for coming to visit me in Richmond, and for giving me the best sisters in the world.

I could never have finished this degree without Ryan and Marcy. Without you guys, I never would have made it through the first year of classes. I owe you both so much. Thanks for going through this with me!

I also want to thank Buddy, whom I love dearly, for keeping me focused and on track and for taking time to sit with me when I was working on projects late into the night!

Last, but certainly not least, I want to thank God for giving me so many blessings in my life and for directing me where I needed to be. "I can do all things through Christ who strengthens me," Philippians 4:13.

Table of Contents

Page

List of Tables

Page

List of Figures

Page

Abstract

ASSESSING, MODIFYING, AND COMBINING DATA FIELDS FROM THE VIRGINIA OFFICE OF THE CHIEF MEDICAL EXAMINER (OCME) DATASET AND THE VIRGINIA DEPARTMENT OF FORENSIC SCIENCE (DFS) DATASETS IN ORDER TO COMPARE CONCENTRATIONS OF SELECTED DRUGS

By Amy Elizabeth Herrin

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2006

Major Director: R.K. Elswick, Jr., Ph.D.
Associate Professor and Interim Chair
Department of Biostatistics

The Medical Examiner of Virginia (ME) dataset and the Virginia Department of Forensic Science Driving Under the Influence of Drugs (DUI) datasets were used to determine whether people have the potential to develop tolerances to diphenhydramine, cocaine, oxycodone, hydrocodone, methadone, and morphine. These datasets included the years 2000-2004 and were used to compare the concentrations of these six drugs between people who died from a drug-related cause of death (of the drug of interest) and people who were pulled over for driving under the influence. Three drug pattern groups were

created to divide each of the six drug-specific datasets in order to compare concentrations between individuals with the drug alone, the drug and ethanol, or a poly pharmacy of drugs (multiple drugs). An ANOVA model was used to determine if there was an interaction effect between the source dataset (ME or DUI) and the drug pattern groups. For diphenhydramine and cocaine, an interaction was statistically significant, but for the other drugs, it was not significant. The other four drug-specific datasets showed that the DUI and ME were statistically significantly different from each other, and all of those datasets except for methadone showed that there was a statistically significant difference between at least two drug pattern groups. Showing that all of these datasets showed differences between the ME and DUI datasets did not provide sufficient evidence to suggest the development of tolerances to each of the six drugs. One exception was with methadone because there were 14 individuals that had what is defined as a "clinical 'lethal' blood concentration". These individuals provide some evidence for the possibility of developing tolerances.

The main outcomes of this study include suggesting changes to make to the ME datasets and the DUI datasets with regard to the way data is kept and collected. Several problems with the fields of these datasets arose before beginning the analysis and had to be corrected. Some of the changes suggested are currently being considered at the Virginia Office of the Chief Medical Examiner as they are beginning to restructure their database.

CHAPTER 1

Introduction

1.1 Introduction to Datasets

The datasets analyzed and presented in this thesis were provided by the Virginia Department of Forensic Science (DFS) and the Virginia Office of the Chief Medical Examiner (OCME). The Medical Examiner (ME) dataset came from the OCME and included data on all persons from the years 2000-2004 who received a toxicological screen as a result of an autopsy throughout the Commonwealth of Virginia. The Driving Under the Influence of drugs (DUI) datasets came from the DFS and contained the demographic and drug data on all persons suspected of, stopped for, and given a blood test for DUI during the years 2000-2004.

1.2 Initial Questions of Interest

One of the original goals for this project was to compare concentrations of tetrahydrocannabinol (THC) between the ME and DUI datasets and among different drug drug patterns groups in the DUI dataset in order to have information available for the Virginia General Assembly for crafting legislation dealing with THC and driving. THC is

the main active ingredient in marijuana, and currently, there are no laws in Virginia that address concentration limits for marijuana or THC. The General Assembly of Virginia uses information similar to the information in these datasets to create and amend current laws regarding drugs and legal limits.

Other goals for this project were to identify concentrations of THC that impair or intoxicate individuals, to see how ethanol combined with THC affects individuals, and to look at drug combinations of THC and other identifiable drugs to see if the effects are additive.

1.3 Problems with Initial Questions and the New Questions

Upon receiving the ME and DUI datasets, problems were encountered that did not allow the investigation of the questions regarding THC. Because THC is not considered lethal, there were no cases in the ME dataset that had THC data; therefore, the ME data was not helpful in answering questions regarding THC.

In the DUI dataset, there was no information about why the individual was stopped for DUI, so it was not possible to determine whether certain individuals were more or less intoxicated than others. This was problematic when trying to collect information about THC for the General Assembly that characterized the degree of intoxication; therefore, the current information contained in the DUI dataset was insufficient for assisting legislation or in the development of laws regarding concentrations of THC.

Since the datasets did not have enough information to answer the questions regarding THC, other questions were posed for this project. These questions included comparing concentrations between the ME and DUI datasets and comparing concentrations

between created drug pattern groups for diphenhydramine, cocaine, oxycodone, hydrocodone, methadone, and morphine. Answering these questions would be a first step in determining if individuals in the DUI dataset had developed tolerances to any of these drugs. These questions are answered in the following chapters.

1.4 Introduction to Subsequent Chapters

This chapter has described the motivation behind this thesis as well as a brief introduction to the datasets and the questions of interest. Chapter 2 gives a description of the datasets, the problems encountered with them, and the solutions to these problems. The analysis is detailed in Chapter 3 where there is a general introduction, a methods section, and results sections for each of the six drugs of interest. Chapter 4 is a summary chapter that includes recommendations to the DFS and OCME on how to better collect the data for their database and suggestions for future projects using these datasets. Throughout this thesis, there will be sections that seem to repeat what has already been stated, but the repeated material was deemed necessary for clarity.

Chapter 2

Sources of the Datasets and Preparation for Analysis

2.1 Sources

2.1.1 Medical Examiner Dataset

The Medical Examiner (ME) dataset is a subset of the database kept by the Office of the Chief Medical Examiner for the Commonwealth of Virginia. This dataset was provided, with permission, by Dr. Joseph J. Saady and the Office of the Chief Medical Examiner (OCME). The OCME collects data and maintains the ME database for the Commonwealth of Virginia. Located in Richmond, Virginia, the OCME is the central office for the medical examiner system throughout the Commonwealth. All autopsy information from any medical examiner in Virginia is entered into the ME database.

Currently, there are no laws that mandate the OCME keep and maintain the ME database, but annual reports concerning fatalities linked with family violence, children, and motor vehicles are required by the Virginia General Assembly. The Virginia General Assembly requires that the Chief Medical Examiner (and committee) "provide ongoing surveillance of fatal family violence occurrences and promulgate an annual report based on

accumulated data"[1]. Information on children's death is also required to be annually reported to the Virginia General Assembly by the Chief Medical Examiner [and committee][2]. These reports should include information on "(i) violent and unnatural child deaths, (ii) sudden child deaths occurring within the first 18 months of life, and (iii) those fatalities for which the cause or manner of death was not determined with reasonable medical certainty,"[2]. The reports concerning deaths of children will be public records, provided there is no identifying information[2]. The annual reports concerning fatalities involving family violence and deaths of children are to be supplied to the Governor and the General Assembly. All medical examiners must also make monthly reports to the Commissioner regarding deaths due to motor vehicle accidents[3]. The OCME also uses the ME database and the information it contains for other research projects in addition to the mandatory annual reports.

The ME dataset that was used for this project included demographic, drug, and death information on all persons receiving a toxicological screen as a result of an autopsy, in Virginia, during the years 2000-2004. The cases in this dataset were limited to those persons who received a toxicological screen during autopsy because the purpose of this study is to look at drug concentrations. Persons not receiving toxicological screens would not have any drug information in their records. The ME dataset received for this project consisted of 2642 individuals and 6026 records. In the original dataset, an individual had a record for each of the drugs found during the autopsy. This dataset was reformatted to include only one observation per case number, and this record combined all information from the multiple observations for each case. The process used to reformat the data is

described in Section 2.2. Once the data was reformatted, any person not receiving a toxicological screen was removed, and the final ME dataset had 2623 observations – one per case number.

Each record in the ME dataset contained the following information: case number, age, gender, fatality (general cause of death), cause of death (detailed description of cause of death), manner of death (e.g. suicide, homicide, natural, undetermined), place of injury and place of death (regional information), premise of death (location information such as home, water, hospital, etc.), drugs found during autopsy, concentrations of those drugs, and the tissue where the drug(s) was detected. Some of these variable fields required major restructuring for use in analysis, e.g. `fatality`, `premise of death`, and `cause of death`. The reorganization of these fields is described in Section 2.2.1. Other fields were used in the analysis in their present condition (with no corrections or restructuring), and some fields were corrected and then used only for descriptive purposes. The initial dataset was received as a Microsoft Excel ® file and was imported into JMP ® for initial inspection of distributions and frequencies of the different fields. Once a general view of the dataset was established, the data was imported into SAS ® for preparation and analysis.

2.1.2 Driving Under the Influence Dataset

The Driving Under the Influence of drugs (DUI) dataset used for this project originally was received in two separate datasets that are both collected and maintained by the Virginia Department of Forensic Science (DFS). The first of these datasets contained demographic information, and the second contained drug information. These datasets were supplied, with permission, by Dr. Joseph J. Saady and the DFS. The data contained in

these datasets are typically used to generate frequency reports to the Virginia General Assembly to help amend and create laws concerning the legal limits of drugs with respect to driving. Previous uses of these datasets have included descriptive reports of demographic information (age, race, gender, etc.) and locality reports of people suspected of DUI in Virginia, when this information was available. This information is used to draw conclusions about drivers who are using drugs and the areas where driving while intoxicated is a problem.

At this time, there are no laws concerning the collection or maintenance of these two datasets, but statistical reports from the DUI data are used by the Virginia General Assembly and sometimes results in laws concerning legal and illegal concentrations of drugs while driving. In the Code of Virginia, blood concentrations of the following substances have been deemed illegal, "(a) 0.02 milligrams of cocaine per liter of blood, (b) 0.1 milligrams of methamphetamine per liter of blood, (c) 0.01 milligrams of phencyclidine per liter of blood, or (d) 0.1 milligrams of 3,4-methylenedioxymethamphetamine per liter of blood,"[4]. It is also "unlawful to drive or operate any motor vehicle, engine or train if a blood alcohol concentration is 0.08 percent or more by weight by volume or 0.08 grams or more per 210 liters of breath as indicated by a chemical test administered,"[4]. By collecting the information contained in the DUI dataset, the DFS can ascertain more information about other drugs and the concentrations that impair drivers throughout the state of Virginia.

Both the demographic and drug datasets of the DUI data used for this project spanned the years 2000-2004. The demographic dataset contained the descriptive

information for all persons suspected of, stopped for, and given a blood test for DUI. This dataset included case numbers, age, race, gender, location of where the subject was detained, and the type of court. The drug dataset contained all drug information, including the case numbers, the drugs found in the blood test, the amount of the drug found, and whether the drug was detected at a concentration "less than", "greater than", "not detected", "quantitated", etc. for the amount that was reported. The drugs detected in the blood test were screened in a precise order specified by the tier system that is described in Section 2.2.5. The guidelines for the tier system are shown in Appendix A.

A driver can be stopped for suspicion of DUI more than one time during the four-year period encompassed in these datasets. In order to avoid repeating information in the datasets (including the same driver more than once), a random sample of case numbers were sent to the DFS to look for individuals who might appear in the dataset more than once. This potential problem is addressed in Section 2.2.4. By definition, being stopped by a police officer for suspicion of DUI does not mean that the driver was convicted of DUI.

The demographic and drug datasets were combined together, based on the case numbers, to form one DUI dataset. The DUI demographic dataset originally included 12,365 individuals, but only 11,819 were unique case numbers. This problem is addressed in Section 2.2.2. The DUI drug dataset contained 48,907 observations, but in this dataset, an identifiable individual could have as many observations as the number of drugs tested for during the blood test, whether or not these drugs were detected. The multiple observations problem was not addressed until the demographic and drug DUI datasets

were combined. When the two datasets were joined together, the pooled dataset contained 55,148 observations with all demographic and drug information in one dataset. This dataset was then reformatted to include only one observation per case number. The process used to reformat this dataset is described in detail in Section 2.2. There were 11,926 individuals in the new dataset, and each individual's record contained all information from the previous multiple observations. This DUI dataset with 11,926 unique individuals was then merged with the 2623 unique individuals in the ME dataset for use in the analysis of the specific drugs and their concentrations.

2.2 Problems

2.2.1 Problems with Free Text Fields

A field is defined as "a named subdivision of a record containing a specifically-defined piece of information within a system"[5]. A free text field is a field with data "containing no formal or predefined structure other than the normal use of grammar and punctuation"[5]. Generally, free text fields should be avoided in the development of databases because of the unlimited number of possible responses. "The data typed into the computer is often entered in a hurry. The language includes abbreviations, jargon, misspelled words, and incorrect grammar,"[6]. These multiple responses usually need to be combined to form general categories for use in analysis. For example, these non-uniform responses could include misspellings, different punctuation, dissimilar orderings of words, and the use of similar or detailed explanations. Although free text fields are valid and sometimes usable, the free text fields in this project required major reorganizing to make new fields that were functional for analysis.

Two of the free text fields in the ME dataset, `fatality` and `premise of death`, were restructured for use in analysis. This reconstruction consisted of the creation of a new field that combined several response levels of the original free text field. An example of this alteration was to the field `fatality`, which contained information regarding a general cause of death. Originally, the field contained 23 levels such as "asphyxia: aspiration/café coronary", "asphyxia: mechanical/positional", and "asphyxia: plastic bag". All three of these levels were joined to form "asphyxia" in the new variable field, `fatalnew`. The field `fatality` initially consisted of 23 levels that were reduced to 14 levels in `fatalnew`, including the new categories "asphyxia", "CO poisoning", "vehicle", and "drug poisoning" and some of the original categories. The new variable included the levels "alcohol/drug withdrawal", "asphyxia", "beating/blows/blunt instrument", "burns/fire", "carbon monoxide poisoning", "cutting instrument/stab", "drowning", "drug poisoning", "explosion", "fall/push", "gun: handgun", "lethal injection", "undetermined", and "vehicle".

A similar alteration was made to `premise of death`, which contained location information regarding the death of an individual. The initial field contained 45 responses such as "street: adjacent", "street: alley", street: bridge", "street: ditch", "street: driveway", "street: highway", "street: interstate", "street: nonspecified" "street: parking", and "street: sidewalk". These 10 levels were combined to form a new category "street" in a new field, `premisenew`. The 45 initial levels of `premise of death` were reduced to 28 levels in `premisenew` by creating four new categories: "street", "water", "home", and "hospital", and keeping some of the original categories from `premise of death`. The

four new categories combined 21 of the original levels from `premise of death`. An example of how SAS ® was used to decrease the number of levels is seen in the following code.

```
if fatality = 'ASPHYXIA: ASPIRATION/CAFE CORONARY' |
      fatality = 'ASPHYXIA: MECHANICAL/POSITIONAL' | fatality =
      'ASPHYXIA: PLASTIC BAG' then fatalnew = 'Asphyxia';
```

With regard to grammatical differences, the only field in the ME dataset that displayed this problem was `cause of death`. The original field `cause of death` had 1724 responses for 2642 individuals, and many of these levels were different only because of misspelled words, different punctuation, synonyms for the same words, or listing the same drugs in a different order. Some examples of misspelled words in the `cause of death` field were "intoxiciation", "intoxicatikon", "poinsoning", "poisoining", "herion", and "cintributing". Misspellings were found by visually searching through a frequency table of the `cause of death` field and looking for observations that were similar to other observations with the only difference being a misspelling.

Once these misspelled words were identified, SAS ® was used to scan the `cause of death` field for the misspelled word, and if it was found, it was replaced with the correct spelling. An example of how misspelled words were corrected is seen in the following code where the misspelled word "intoxiciation" is replaced with the correct spelling "intoxication". Not all misspelled words were corrected because there were too many misspellings to detect them all by scrolling through the frequency tables. Approximately 100 grammar problems were found and corrected, and 3003 changes were made to the `cause of death` field for the 2642 cases in the ME dataset.

```
index = find(cause_of_death, 'INTOXICIATION', 'i');
if index ne 0 then temp = substrn(cause_of_death,1,index-1) ||
      'INTOXICATION' ||
      substrn(cause_of_death,index+13,length(cause_of_death));
else temp = cause_of_death;
cause_of_death = temp;
```

Some observations for `cause of death` were only different in the punctuation. These punctuation differences included having double spaces between words, periods at the end of some cases, spaces between commas, and using "&", "and" "d/t", "d/t to" or "due to" inconsistently. In all of these cases, one example was chosen as the standard form, and the others were modified to match the standard form of the punctuation. An example of the grammatical fix is seen in the following SAS ® code.

```
index = find(cause_of_death, 'DUE TO', 'i');
if index ne 0 then temp = substrn(cause_of_death,1,index-1) ||
      'D/T' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'D/T TO', 'i');
if index ne 0 then temp = substrn(cause_of_death,1,index-1) ||
      'D/T' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
else temp = cause_of_death;
cause_of_death = temp;
```

Using synonyms for some words also posed a problem in the `cause of death` field. "A synonym is created when two different names are used for the same information (attribute). If an attribute resides in more than one entity [record], insure that all entities

[records] use the same attribute name,"[7]. For example, "overdose", "poisoning", "toxicity", and "intoxication" all have the same meaning in `cause of death`, but different medical examiners used these words interchangeably in their description for `cause of death`. "Using more than one name for the same attribute causes many problems,"[7]. In this example, all variations of "overdose", "poisoning", "toxicity", and "intoxication" were changed to "poisoning" in order to provide consistency within this field.

Yet another problem with this free text field was different orderings of words. The field `cause of death` contained a detailed description of the cause of death including the drugs that were attributable to death and the specific cause such as asphyxia, poisoning, lethal injection, drowning, etc. In several instances, a detailed cause of death was the same as another with the only difference being the ordering of the drugs. For example, one cause of death was "acute ethanol and opiate poisoning", and another cause of death was "acute opiate and ethanol poisoning". These two causes of death are the same for purposes of analysis, but they appear as two different responses.

Originally, one of the main variable fields of interest for analysis, `cause of death` contained a large number of non-uniform levels, and only parts of it were corrected. "Typographical errors are not trivial matters. A user can have little confidence in a database in which many typographical errors occur,"[8]. Because there were so many errors in `cause of death`, it was not the only field used to help choose the individuals from the ME dataset to include in analysis. Other fields were used to determine the cases where an individual died from an actual drug overdose. These fields included `fatalnew`

and `drug level`, which was renamed to `amount` in a process described in Section 2.2.6. Those cases where the person died of a drug overdose were the ones of interest from the ME dataset to include in analysis. The DUI datasets did not have problems with free text fields, but they did have problems with demographic information and missing data.

2.2.2 Problems with Demographic Fields

A demographic variable (or field) is "a varying characteristic that is a vital or social statistic of an individual, sample group, or population, for example, age, sex, socioeconomic status, racial origin, education,"[9]. The DUI demographic dataset contained demographic information on drivers who were stopped by police for suspicion of "driving under the influence" of drugs and received a blood test. When a suspect is stopped for driving under the influence, he or she "may be required to perform field sobriety tests, and the police officer usually asks them to blow into a Breathalyzer device to detect alcohol"[10]. If the police officer thinks the driver is impaired and the Breathalyzer device shows no ethanol or not enough to explain the impairment, then the driver is taken by the police officer to a location where a blood sample is collected[10].

Because demographic information does not change for a particular person, a case number in the demographic dataset should appear only one time; however, some case numbers were listed multiple times within this dataset. Upon investigation, it was discovered that any change in a record was accomplished by appending an updated record to the end of the database instead of correcting the existing entry. By appending these "new," updated records instead of replacing them, a case number could appear more than once in the dataset, one for every time the record was entered. To rectify this problem, the

last record added to the database for each case was assumed to have the most current information, and any previous records for the same case number were removed from the dataset.

Some of the demographic information in the DUI dataset was identified as being invalid due to laws concerning driver's license age. To possess a driver's license (including a driving permit) in the United States, a person must be at least fourteen years of age. The minimum driving age is usually 16, but this age varies by state[11]. Farm permits or school permits can be issued to drivers under 16 (but not younger than 14) in special circumstances in some states[12]. Using age 14 as the cutoff age for possessing a driver's license (or permit), there were some cases where the ages were too young to be legally driving, i.e. ages 0, 8, and 13. Eleven case numbers with ages less than 14 were sent to the DFS for verification. The ages were determined to be a clerical error, but a correct age could not be established. Because the correct age could not be found, the ages for these cases were set to missing. The problem of missing demographic information is discussed in the next section, but because of the magnitude of this problem, adding the 11 missing ages (due to clerical errors) did not pose any major problems for the dataset. These 11 cases were kept in the dataset (with missing ages) to keep the maximum number of observations possible for analysis.

Problems were also identified with the field `race`. Race can be determined by many methods, including self-report or observation by another party. In the ME dataset, race was not recorded, but a racial description was reported in the DUI demographic dataset. The `race` field in the DUI dataset was usually determined by the police officer.

"Some jurisdictions may require that the officer check the DMV [Department of Motor Vehicles] record, but mostly the officer decides,"[13]. In this dataset, race is not reported in a standard way because it was either self-reported by the individual (from the DMV record) or observed by the police officer; therefore, the information in the race field is prone to error.

2.2.3 Problem with Missing Data

A nearly universal problem with datasets or databases is missing observations, and this problem was especially apparent in the DUI datasets. The missing data problem appeared when the DUI demographic and drug datasets were combined to form one dataset with all DUI information. The DUI demographic dataset contained `case number`, `race`, `gender`, `age`, `court type description`, and `court name`. The DUI drug dataset contained `case number`, `drug`, `amount`, and `presence`. The datasets were joined by matching the case numbers in each dataset to create one dataset with both demographic and drug information.

Numerous case numbers appeared in the demographic dataset that did not appear in the drug dataset and vice versa. When the case numbers were matched, if a case number did not appear in the demographic dataset, then there was no demographic information to attach to the drug information, producing instances of missing demographic data. This missing information posed a problem because roughly half of the number of observations had missing values for `age`; therefore, `age` could not be used to accurately characterize the DUI dataset. About 100 observations were missing all demographic data. On the other hand, if a case number appeared in the demographic dataset and not in the drug dataset,

then there was no corresponding drug information for that case number when the two datasets were combined. Approximately 6000 of the 11,926 cases had missing drug data, which created a smaller pool of observations to select individuals for the analyses. Because of the high percentage of missing demographic information for `age`, the 11 ages that were set to missing (mentioned in the previous section) did not notably increase the amount of missing demographic data.

The ME dataset had some missing data, but a far lower percentage of the information was missing in this dataset. For example, `gender` had only one missing observation. There was not much missing data in the ME dataset, and consequently, it was not considered problematic for the analysis.

2.2.4 Problem with Independence Assumption

"Independence of observations refers to the notion that the value of one datum is unrelated to any other datum. In other words, knowing the value of one observation gives you no information about the value of any other,"[14]. Having an individual appear in the dataset more than once would violate the assumption of independence and cause problems in the analysis of the datasets. If any individual had been stopped for suspicion of driving under the influence more than once, then he or she appeared in the dataset under two different case numbers. Because an individual could be stopped for suspicion of DUI more than once over the course of four years and thus, appear in the dataset more than once, a random sample of 10% (1192) of the case numbers from the combined DUI dataset was sent to the DFS to cross-reference the case numbers with last names, looking for people who had been stopped more than once. The purpose for doing this cross-reference was to

estimate the number of individuals with multiple traffic stops and the impact (if any) on the independence assumption.

The results of this cross-reference test were inconclusive. The dataset "does not capture the proper data to ensure the query would produce good results"[15]. Because a unique identifier, such as social security number, was not collected in this dataset, the search could only be by name, gender, race, and date of birth[15]. Birthdates and names are not unique identifiers (neither are race are gender), so that combination is reliable to determine if people appear in the DUI dataset more than once. No combination of those four variables would guarantee unique individuals. From the list of case numbers provided to the DFS, "the names do not match between submission 1 and 2, submission 1 might have J. Smith with DOB of 6/12/1956 and Submission 2 might have John Smith without at DOB, so there is no reliable way to get this data [meaning the reoccurrence of an individual in the dataset]"[15].

2.2.5 Problem with Data Completeness (Underestimation)

Drugs listed in the DUI drug dataset were found using a tier system designed by Dr. Joseph J. Saady and the DFS (Appendix A). Because the DFS adopted this tier system to rapidly respond to DUI requests and maintain quality and validity within the program and to save time and money for the Commonwealth, some of the drug information in the DUI dataset is underestimated. Underestimation is "an estimation that is too low"[16] or "an estimate that is less than the true or actual value,"[16].

The tier system was designed so that once a drug, or group of drugs, was found at a concentration high enough for probable conviction of DUI, no other testing was done to

detect other drugs. The first drug tested was always ethanol, and if ethanol was found at the cutoff concentration (0.09 %), then the blood was not tested for the presence of other substances. Thus, there was a potential for all drugs other than ethanol to be underestimated. If the concentration of ethanol was not sent or high enough for probable conviction, then tests were done for a second group of drugs (Level II drugs). Level II drugs include, but are not limited to, Amphetamines, Barbiturates, Cannabinoids, Benzodiazepines, and Opiates. If no Level II drugs are found at a concentration high enough for probable conviction, then the blood is tested for Level III substances, such as Antihistimines, Antidepressants/Antipsychotics, hypnotics, muscle relaxants, and ketamine. If no Level III drugs are found at an illegal concentration, the blood is no longer tested and the drug information is likely not sufficient evidence for conviction.

This tier system presents an underestimation problem for the DUI dataset because the number of drugs found is clearly less than the "true or actual"[16] number of drugs that were present in an individual at the time of the traffic stop. Underestimation is not a problem for ethanol because all blood samples were tested for ethanol. Using the tier system's testing protocol causes Level II and Level III drugs to be underestimated, but due to the design of the tier system, Level III drugs will be more severely underestimated than the Level II drugs.

By using the tier system, the effect of underestimation on the DUI dataset results in several problems. First, the number of drugs that were found in the DUI dataset and the number of individuals using these drugs are both measures that are underestimated. The concentration levels are possibly underestimated as well in this dataset because while the

range of concentrations could span the entire range of drugs possibly found in blood, it is more likely that if more drugs were detected in more cases then the ranges of the drugs would expand.

A second problem with underestimation occurs with the reported concentration. When detecting a drug, the tier system's testing protocol (Appendix A) specifies lower and higher reporting limits for each drug (or group of drugs). The guidelines for the testing protocol instruct the laboratory to stop testing for a drug if the upper limit is reached. If the lower reporting limit for a drug is detected, then the lower reporting limit amount is recorded in the `amount` field. "Less than" is recorded in the `presence` field to report that the drug was detected below the lower reporting limit for that drug. On the other hand, if an upper limit for a drug is reached, then the upper limit amount is recorded in `amount`. "Greater than" is recorded in `presence`. When either of these situations takes places, an adjustment had to be made to the concentration amount to accurately reflect the amount of drug in a person's system. For concentrations that have a "less than" value for `presence`, the concentration is considered to be one half of the concentration reported, and for concentrations with a "greater than" value for `presence`, 15% of the amount reported is added to the concentration. This rule was determined by Dr. Saady in order to use the most accurate concentrations. The details of this adjustment are discussed in the last paragraph of Section 2.2.6.

2.2.6 Problem with Converting Character Fields to Numeric Fields

The ME dataset was joined with the DUI datasets by concatenating the two datasets. "*Concatenating* is combining two or more data sets one after the other into a

single data set. The number of observations in the new data set is the sum of the number of observations in the original data sets, and the order is all the observations from the first data set followed by all observations from the second data set and so on,"[17]. Problems with concatenating datasets occur when some fields appear in one dataset but not in the others and when some fields have different characteristics, attributes, or formats in the datasets. In order to concatenate the datasets, it must be determined what information should be assembled and collected from each of the source datasets to have in the final combined dataset. After determining what information to keep from each dataset, a format needs to be determined for the final dataset. Since all datasets being concatenated must fit into the chosen final format, any number of the fields from these datasets needs to be modified, including field names, types, and contents. Any extra information from either dataset is removed before concatenating the two sets.

In the ME dataset, the original variable field `drug level` contained a concentration amount and the units of measurement for each drug found. Thus, `drug level` was a character field. A character field can contain any series of letters, numbers, and special characters[18]. The corresponding variable field `amount` in the combined DUI dataset was a numeric field, "containing only numbers, including numbers in E-notation, and sometimes a decimal point or minus sign,"[18] because it contained only the concentration amount with no associated unit of measurement. Because `amount` (from the DUI dataset) was a numeric field and `drug level` (from the ME dataset) was a character field, `drug level` was converted to a numeric field in order to combine the two datasets. This conversion created the character to numeric problem discussed next.

To facilitate the conversion from character to numeric, new fields were created to contain the numeric part "concentration" of `drug level` and the character part "units of measurement" of `drug level` from the ME dataset. The concentration field of `drug level` (from the ME dataset) was named `amount` to match the name of the field in the DUI dataset, and `unit` was the field created for the unit of measurement.

Some of the case numbers in the ME dataset did not have a unit of measurement associated with the concentration, so a standard unit was fixed to each concentration. This standard unit of measurement was determined by the specific drug and the tissue where the drug was detected. Table 1 shows the tissues and the corresponding units of measurement. Certain cases in the ME dataset did not have a specific concentration and contained only "present" for `drug level`. These cases were not used in analysis for that drug because an exact concentration amount could not be determined.

Table 1: Tissues and the Corresponding Unit of Measurement

Drug	Tissue	Unit of Measurement
Ethanol	All	Grams % (g/100mL)
All (except ethanol)	Bile	mg/kg
All (except ethanol)	Blood	mg/L
All (except ethanol)	Brain	mg/kg
All (except ethanol)	Gastric	mg/L
All (except ethanol)	Liver	mg/kg
All (except ethanol)	Urine	mg/L
All (except ethanol)	Vitreous	mg/L

In the DUI dataset, amount was kept as numeric, and an additional field was created for the unit of measurement. In a blood test, the unit of measurement for all drugs is "mg/L" with the exception of ethanol, which is measured in "grams percent"[19]. This standard unit of measurement was added to the created field unit for all observations in the DUI dataset.

Another field in the DUI dataset was presence, and this field characterized amount as "less than" or "greater than" the concentration reported, as "quantitated" at that concentration, or "not detected". If presence was labeled as "less than", the amount was estimated to be one half of the value reported, which was usually the lower limit of detection. If presence was labeled as "greater than", then amount was calculated to be 15% higher than the reported concentration, which was usually the upper limit of detection. If presence was "not detected", then those cases were not used for the analysis of the particular drug. "Quantitated" values for presence were used as the values were reported. The field presence was only found in the DUI dataset. For the ME dataset, presence was always set to missing, and amount was assumed to be "quantitated."

2.3 Creation of New Fields (Compatibility for Concatenating)

Many new fields were created to use during analysis. These fields were created by grouping levels of the original variable fields, splitting existing fields into separate fields, or using information from one variable field in order to create multiple fields to divide the information into a more usable format.

The ME dataset was originally designed so that a case number could appear more than once in the dataset, with the maximum number of observations per case number equal to the number of drugs found during an autopsy. The original DUI drug dataset was designed in the same manner as the ME dataset, with the possibility of more than one observation per case number, and the maximum observations per case number equal to the number of drugs tested in the blood sample. When the two DUI datasets were merged together, the demographic information for a case number was replicated for the number of observations for that case in the DUI drug dataset. Both the ME and the combined DUI datasets had the possibility for multiple observations per case. For analysis, it was needed to combine the multiple observations for each case into a single observation that contained all of the pertinent information for each individual from the many observations.

Converting these datasets from multiple observations per case number into a dataset that had only one observation per case number involved the creation of 695 new variable fields to create a field for each of the 139 drugs, their concentrations, their presences, and their unit of measurement. Five arrays, each consisting of 139 fields (one for each of the 139 drugs found in either dataset), were created to sift through the multiple observations to combine all of the information into one observation for each case number. The first array, `names`, contained the names of all 139 drugs and was used to compare to the field `drug` to check whether or not a drug was listed (found) for a particular case number. The next array, `drugs`, contained 139 fields which either contained a "1" if the drug was listed for a case number or "." if the drug was not listed. The array `amounts` contained a numeric value for the concentration level, and `levels` contained a unit of

measurement for each drug found. Each of these two arrays, amounts and levels, contained 139 fields to match each of the 139 drugs. The last array, press, contained the information from the DUI field presence. Since presence was not in the ME dataset, when the two datasets were combined, press was set to missing for any observation that came from the ME dataset. The field names in all of the arrays (except drugs) were distinguished by the name of the drug followed by an underscore and then a letter that differentiated the variable fields between each array. The field names in drugs did not have an underscore or letter; they were merely the drugs themselves. For example, the fields in drugs include names like "acetaminophen," "ethanol," and "gamma-hydroxybutyrate." In the other arrays, the fields included names such as "acetaminophen_n" for names, "acetaminophen_c" for amounts, "acetaminophen_p" for press, and "acetaminophen_l" for levels. The SAS ® code that illustrates this process follows the next paragraph.

SAS ® was then used to create a single observation from the multiple observations by using a loop that retained the case number. If the subsequent observation and the previous observation had the same case number then the information for those records were combined. The arrays mentioned previously allowed for the new record (a single entry per case number) to contain information about whether a drug was tested (drugs), the concentration (amounts), the unit of measurement (levels), and the presence (press) for each drug. The following SAS ® code shows how the DUI dataset was transformed from a dataset with multiple observations per case number to a dataset with only one observation per case number. The SAS ® code for the ME dataset differs slightly

due to the different variable fields in the dataset. The full SAS ® code can be seen in Appendix D.

```
if first.case_number then do;
      do i = 1 to 139 by 1;
            drugs{i} = .;
            amounts{i} = .;
            levels{i} = '';
            press{i} = '';
            number_of_drugs = 0;
      end;
end;
do i =1 to 139 by 1;
      if drug=upcase(names{i}) then do;
            drugs{i}=1;
            amounts{i}=amount;
            levels{i} = unit;
            press{i} = presence;
            number_of_drugs = number_of_drugs+1;
      end;
end;
if last.case_number then output;
```

Other fields were also created for this project. As mentioned earlier, the variable fields `fatalnew` and `premisenew` were both created in the ME dataset to correct the original `fatality` and `premise of death` fields. The field `fatalnew` was created to assist in subsetting the data into separate drug datasets for analysis. `Premisenew` was created to have location information for the cause of death that could be used in the analysis. The two datasets (ME and DUI) contained some of the same information in fields with different names, so new fields were created to aid in combining the two datasets. For example, in the ME dataset, the information regarding a person's gender was in a field called `sex`, and in the DUI dataset the field was called `gender`. A new field

gender was created in the ME dataset and set equal to the values from sex. The field sex was then dropped from the ME dataset.

As mentioned in Section 2.2.6, drug level (ME dataset) was separated into two separate fields. The numeric piece of information (the part containing the concentration) became amount, which was the same in the DUI dataset. The field unit was created for both datasets, but the values for each dataset were different. The details creating the field unit for each of the datasets was discussed in Section 2.2.6 with the problem of converting character fields to numeric fields. Another variable field, database, was created in each of the datasets (ME and DUI). The database field contained character information regarding the original dataset affiliation and was used as an identifier of the source dataset after the ME and DUI datasets were combined.

2.4 Final Numbers

2.4.1 ME Dataset

The ME dataset initially contained 6026 records, 2642 unique case numbers, and 15 variable fields in a Microsoft Excel ® file. These fields contained information on age, race, gender, fatality, cause of death, premise of death, place of injury and place of death, drugs detected during an autopsy, drug concentration, and tissue where the drug was detected. After encountering numerous problems with the fields in the ME dataset, some of the fields were corrected for use in analysis.

The fields, fatality and premise of death, were restructured by combining multiple levels of the original free text fields. Creating new fields, such as fatalnew and premisenew combined many of the original levels (responses) into a

smaller number of levels in each of the two fields. Examples of combining the levels were discussed in Section 2.2.1 as well as other free text problems.

The new field `fatalnew` was used as a filter for the ME dataset to remove those persons who did not die from "drug poisoning". The purpose of `premisenew` was for descriptive information on the location for the cause of death.

Other free text problems encountered in the `cause of death` field consisted of many types of grammatical errors including misspelled words, different punctuation, different orderings of words, and synonyms for specific words. Once these grammatical problems were addressed, `cause of death` was used as another filter to create the specific drug datasets for use in analysis. Approximately 100 grammatical errors were corrected in the `cause of death` field. Section 2.2.1 described the details of the free text problems encountered and the respective fixes to these problems.

Some new fields were created in order to match variable field names between the two datasets because some fields were designed as character fields for one dataset and as numeric for another dataset. This problem was addressed in Section 2.2.6. An additional 695 fields were created to keep track of the 139 listed drugs and their information including names, whether a drug was detected during an autopsy, the concentrations of the drugs, units of measurement, and information regarding the "presence" of the drugs (Section 2.3). The 139 "presence" variable fields were created to use the information regarding lower and upper recording limits from the tier system from the DUI dataset. The array consisting of the 139 names of drugs was only used to compare the listed drug with

the drug field in each of the two datasets to see whether it was detected in an individual. Once the comparison was made, these 139 fields were deleted from the final dataset.

Some specific names of drugs in the ME dataset were renamed to match the names of the same drug in the DUI dataset. An example of this correction was "methylenedioxyamphetamine" which was changed to "mda". Not all of these drugs were used for analysis, but they were changed in the dataset for possible future use. Nineteen records in the ME dataset were removed from the dataset because the cases did not receive a toxicological screen as a result of an autopsy.

After all of these corrections and amendments were made, the final ME dataset consisted of 2623 records, 2623 unique individuals, and 577 variables.

2.4.2 DUI Dataset

The DUI dataset originally consisted of two datasets – one with demographic data and one with drug data. The DUI demographic dataset contained 12,365 records, but the records were not 12,365 unique individuals. In the demographic dataset, 11,819 case numbers were distinctive. Information in the DUI demographic dataset included case number, age, gender, race, court type, and court name (location of jurisdiction). The DUI drug dataset consisted of 48,907 records, and in this dataset, a case could have multiple observations up to the number of drugs tested for in the blood sample. The variable fields in the drug dataset included case number, drugs detected during blood test, concentration of drug, and presence of the drug.

The 11,819 unique case numbers from the DUI demographic dataset were matched with the case numbers in the DUI drug dataset to combine the two datasets into one dataset

with all possible information. Some of the case numbers appeared in one or the other of the two datasets (demographic and drug), but not in both. When this happened, sections of data in the combined dataset were missing. A section of case numbers had drug information with no demographic information and vice versa. This resulted in a problem with missing data (Section 2.2.3). Having missing drug data was more common than having missing demographic data, but both situations occurred when the two datasets were combined. Once the two DUI datasets were combined into one dataset, the resulting dataset had 55,148 records.

Other problems needed to be addressed in order to combine the collective DUI dataset with the ME dataset. These problems included creating new variable fields (both datasets) (Section 2.3 and 2.4.1), addressing problems with underestimation and independence (DUI only) (Sections 2.2.5 and 2.2.4), and converting character fields to numeric fields (ME only) (Section 2.2.6). After these problems were addressed and amendments were made, the DUI dataset needed to be reformatted to include only one observation per person. (When the demographic and drug datasets were combined, the drug dataset contained multiple observations per case, and the demographic information was duplicated to match the number of observations in the drug dataset for each record.) By reformatting the DUI dataset, each case number appeared only one time, and each of these records contained all the data for each drug tested and detected. This final dataset had 11,926 records and 577 variable fields. Close to half of the records (~6000) had missing drug information, and about other 100 records had completely missing demographic information.

2.4.3 Combined Dataset

With both the DUI and ME datasets in formats compatible with each other, the two datasets could be joined together for ease of analysis. The joint dataset that included all 2623 ME records and 11,926 DUI records, and had 14,549 records total. This dataset included all individuals from the source datasets, including some individuals from the ME dataset who did not die from "drug poisoning". Using the variable field `fatalnew` as a filter, 70 persons having a `cause of death` other than "drug poisoning" were eliminated from the dataset. The final dataset consisted of 14,479 total records and 577 variable fields from either the DUI or the ME dataset. These records compose the base dataset from which all of the smaller drug datasets were generated.

CHAPTER 3

Analysis of the Drug-Specific Datasets

3.0 The Analysis of the Combined Dataset

This chapter begins with an introduction of the combined dataset used for the analyses and the process that was used to select the smaller, drug-specific datasets. A short description of the combined dataset, including the number of observations from each of the contributing datasets (ME and DUI), and the variable fields in the final dataset are presented in Section 3.1. The remainder of this section (Section 3.1) describes how the six datasets (one for each drug of interest) were created. The methods section (Section 3.2) details the process used to break down each of the six datasets into categories to compare the concentrations of the drug between the two source datasets (ME and DUI) and between each of the three drug pattern groups. After specifying the process used to divide each dataset into the drug pattern groups, Section 3.2 describes the statistical methods that were used for each of the six datasets. Section 3.3 contains the results of the analyses of the six drug-specific datasets. This section is composed of six parts – one for each of the drug-

specific dataset. Section 3.4 summarizes the results from all six analyses and gives general conclusions of all analyses.

3.1 Introduction

The original ME and DUI datasets were combined into a final dataset that was subsequently divided into six smaller datasets (one for each drug of interest). The final dataset contained 14,479 individuals and 577 variable fields. (There were a couple of additional fields added to the drug datasets for transformation purposes and combining of the drug pattern groups.) The variable fields in the final dataset included `case number`, `age`, `race description`, `court name`, `court type description`, `current date`, `cause of death`, `manner of death`, `date of injury`, `date of death`, `place of injury`, `place of death`, `fatalnew`, `premisenew`, `gender`, `database`, `count`(used to determine how many corrections were made to the `cause of death` field), `bac` (listed the tissue where the drug was detected), 139 variable fields that indicate the name of the drug, 139 fields that give the concentrations of each of the drugs, 139 fields with the unit of measurement, 139 fields that give information on presence of the drug, `number of drugs` (a counter field that specifies the number of drugs that were detected in someone's system), `index fatal` (denoting whether or not the `cause of death` field was attributable to a drug poisoning), and `group` (indicates whether the individual had the only the drug of interest, the drug and ethanol, or the drug and any other combination of drugs). The variable field `group` is discussed in more detail in Section 3.2. (A table with all of these variable fields

and a description is seen in Appendix E.) Out of the 14,479 cases in the final dataset, 2553 came from the ME dataset and 11,926 came from the DUI datasets.

As mentioned in Chapter 1, there were six drugs of interest for this project. These drugs were diphenhydramine, cocaine, oxycodone, hydrocodone, methadone, and morphine. A separate dataset was created for each of the six drugs; each dataset was a subset of the final combined dataset where the subset contained all persons with the drug of interest. These six datasets were used to assess differences in concentrations of the drug of interest between the ME and DUI datasets and between the three drug pattern groups. Initially, the ME dataset contained cases that included other tissues such as liver, brain, urine, etc. These tissues were sampled when a blood sample was not available[20]. Because of differences in the unit of measurements between various tissue sources, only those samples taken in blood were used to compare concentrations since all individuals in the DUI dataset received blood tests. (Because blood was the oly tissue sampled for in the DUI dataset, any individual from the DUI dataset who tested positive for the drug of interest was included in the respective drug dataset.)

3.2 Methods

Each dataset was created by excluding individuals from the final combined dataset that did not have a positive screen for the specific drug for the analysis. First, individuals who did not receive a toxicological screen as a result of an autopsy were excluded from the ME dataset prior to combining the ME and DUI datasets. For the ME dataset, a second inclusion criterion was that individuals must have died from a drug-related death due to the drug of interest (`fatalnew` must have "drug poisoning" as the factor). Third, individuals

who did not have the specific drug of interest in their system were excluded (i.e., the concentration was zero, missing, or "not detected"). Finally, some individuals were excluded based on clinical judgment. "It was a clinical decision. If a particular drug level was too low, it was determined to be an incidental finding and not a cause of death, and the case was excluded"[21]. This clinical decision was only for cases in the ME dataset. A detailed flow chart of the filtering system used to create the six smaller datasets is seen in Appendix B.

Following the creation of the six drug-specific datasets, the cases were divided into three drug pattern groups; the groupings were determined using clinical criteria. "The groups were chosen based on practicality. We definitely wanted to have the drug alone because it is informative and previous studies have been done with the drug alone. Ethanol is the most abused drug and there are many circumstances of ethanol and a drug. Because ethanol is so prevalent, that was the reason for the second group. The "poly pharmacy" drug pattern group was chosen because there was no other option left"[20]. Using this criteria, a new field, `group`, was created to distinguish between the different drug pattern groups. The first category in `group` consisted of those individuals who had only the drug of interest detected in their system. In the DUI dataset, all individuals were initially tested for ethanol, so for an individual from the DUI dataset to be in Group 1, the concentration for ethanol (and all other drugs) had to be "0 mg/L" or "missing" in the blood sample. (Section 2.2.5 details the tier system and the process for which drugs were screened in the DUI dataset). Group 2 contained individuals who had both the specific drug and ethanol in their system. Finally, Group 3 contained any other combination with the drug of interest

and thus, was considered a poly pharmacy group. All three of the created drug pattern groups were mutually exclusive, i.e., if an individual was in one of these drug pattern groups, then the same individual could not be in another drug pattern group (within the same drug dataset). It is important to note the possibility that individuals may appear in multiple drug-specific subgroups because of the nature of the poly pharmacy group (Group 3).

The analysis of the data began with a normal quantile plot constructed for each drug pattern group and each source dataset to assess if the blood concentration data followed an approximately normal distribution. Because the blood concentration data is bounded at zero and because of the likelihood of unusually high concentrations (especially in the ME dataset), it was anticipated that the blood concentration data would be positively skewed. If the blood concentrations were skewed positively then the drug concentration data were log-transformed in an attempt to transform the data to be approximately normal. The means of the log concentrations of all drugs for each of the drug pattern groups were calculated with the equivalent untransformed concentration amount. The log-transformed data was reported in "log (mg/L)" units, and the untransformed data was reported in "mg/L" units. For ease in interpretation of the concentrations, the untransformed (original metric) data in lieu of the log-transformed data is given whenever an amount is reported.

Using a statistical "rule of thumb," the minimum number needed to calculate a confidence interval, is 12. "The width of a confidence interval, involving estimation of variability and sample size, decreases rapidly until 12 observations are reached and then decreases less rapidly"[22]. In certain cases, samples sizes as small as 11 were accepted

based upon clinical judgment. Otherwise, if a drug pattern group had less than 12, an adjustment had to be made to the drug pattern groups. The options for adjustments were to combine the drug pattern group with a different drug pattern group or to omit that category from the analysis of that drug, so the number of drug pattern groups was either three (the number created) or two (if two drug pattern groups were combined or one was left out of the analysis).

A quantile box plot is a plot that summarizes the distribution of points for a specific factor level. The two ends of the box display the 25^{th} and 75^{th} quartiles. The space between these two ends is known as the *interquartile range*, and the line that crosses the middle of the box is the median (or 50^{th} percentile). Each box has lines that extend from each end of the box, called *whiskers,* that mark the furthest points in either direction[23]. Quantile box plots of the concentrations were produced for each of the drug pattern groups and each source dataset for both the log-transformed and untransformed data. The box plots were used to visually inspect for overlaps between the source datasets, within each drug pattern group. The box plots seen in the following sections show the concentrations before and after the log transformation. Each pair of box plots (transformed and untransformed) is displayed on the same scale, and some of the box plots show expanded versions for the DUI dataset. When the DUI datasets were put on the same scale as the ME datasets, the plots of the DUI data appeared condensed, and thus, it is frequently difficult to see the details of the plots.

A two-factor analysis of variance (ANOVA) model was fit to each of the six datasets in order to detect statistical differences between the mean blood concentration for

the two source datasets (ME and DUI) and for the three drug pattern groups. For this model, the dependent variable field (y) is always the log blood concentration of the drug of interest (due to the non normality of the untransformed concentrations) and the independent fields are `database` (β) and `group` (α). An additional term was also added to allow for an interaction (α β) between `database` and `group`. The two-factor ANOVA model is given,

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

where y_{ijk} is the log of the drug concentration from the i^{th} drug pattern group (i = 1, 2, 3), the j^{th} source dataset (j = 1, 2), and the k^{th} observation. The grand mean of the log concentrations is μ, α_i is the effect due to the i^{th} drug pattern group (1 = drug only, 2 = drug and ethanol, 3 = drug with any other drug combination), β_j is the effect due to the j^{th} source dataset (1 = ME, 2 = DUI), $(\alpha\beta)_{ij}$ is the interaction effect between the i^{th} level of α and the j^{th} level of β, and ε_{ijk} is the random error associated with y_{ijk}. The assumptions for the two-factor ANOVA model are that the errors (ε_{ijk}) are independent and identifcally distributed with a normal distribution ($\varepsilon_{ijk} \sim N(0, \sigma^2)$).

When fitting the two-factor ANOVA, an interaction term was considered significant at α = 0.10, and thus, it was retained in the model; otherwise, the interaction was removed from the model, and the model was rerun. The alpha was set at 0.10 for testing the interaction in order to mitigate the possibility of inappropriately removing the interaction term (a Type II error). In this case that the interaction term was removed from the model, the comparison of the two source datasets and the comparison of the three drug pattern groups is a straight-forward test of the main effects. A Tukey's multiple comparison

procedure was used to assess the nature of differences between the three levels of the drug pattern groups while controlling the overall alpha level.

If the p-value associated with the interaction was less than 0.10, then the comparison of the source datasets and the drug pattern groups were made within the interaction term. In other words, because of the presence of the interaction, the comparison of the source datasets must be made within each of the drug pattern groups. Similarly, the comparison of the drug pattern groups must be made within each of the source datasets. Again, a Tukey's multiple comparison procedure was used to determine individual differences while controlling the overall alpha level. Tukey's multiple comparison procedure makes all possible pairwise comparisons to detect any differences. The results of the ANOVA models and the Tukey tests are discussed in each of the sections respective to the drug of interest.

Residual by predicted plots and normal quantile plots of the residuals for all six drug datasets are seen in Appendix C. The residual by predicted plots are inspected to see if the residuals are randomly scattered above and below the horizontal line at zero and that there is not systematic pattern to the residuals; this is a check of the common variance assumption and of the model adequacy. The normal quantile plots of the residuals are used to verify the normality assumption of the residuals for the ANOVA model.

3.3 Results

The results section is divided into six subsections – one for each of the drug-specific datasets. Each section contains a short description of the dataset, a table with the number of observations per drug pattern group and per source dataset, and a table of

demographic information for each drug pattern group divided by dataset (ME and DUI). The normal quantile plots before and after the log transformations for each drug pattern group and source dataset are displayed as well as the pairs of box plots; these box plots show the effect of using the log-transformed concentrations and the untransformed concentrations for the ME and the DUI datasets and the levels in the `group` field. The results of the analyses will vary depending on the drug of interest.

3.3.1 Diphenhydramine (Level III)

"Diphenhydramine is an antihistamine, used to treat allergies, motion sickness, allergic reactions, insomnia, cough, nausea, and phenothiazine drug-induced abnormal muscle movement,"[24] Diphenhydramine is commonly known by its trade name, Benadryl ®, and is frequently used by cancer patients, people with allergies or nausea, or by people who have Parkinson's disease[24]. "Diphenhydramine (DPH)-related deaths in adults are extremely rare, and detailed autopsy studies are rarer still"[25]; however, there were some cases in the ME dataset with a cause of death attributable to diphenhydramine.

Using the filters mentioned in Section 3.2, the diphenhydramine dataset generated from the final dataset consisted of 213 cases. Of these 213 individuals, 102 came from the ME dataset and 111 came from the DUI dataset. Table 2 breaks down the 213 observations by drug pattern group and by dataset (ME or DUI). Because there were only two individuals in Group 2 for the ME dataset, Groups 2 and 3 were combined for the analysis of diphenhydramine. The combined group (Groups 2 and 3) consisted of 89 individuals from the ME dataset and 85 individuals from the DUI dataset for a total of 174 observations in Group 2. Table 3 shows the categorization of the demographics for the two

source datasets and the two new drug pattern groups. These two drug pattern groups are Group 1 (diphenhydramine only) and the new group (diphenhydramine with any combination of drug – including ethanol) that combined Group 2 and Group 3. In this table, the field `age` is displayed as a range of the youngest and the oldest, while the `gender` and `race` rows give frequencies of the individuals that fit into each of those categories. The percentages shown in this table are the percentage of the category (cell) in the source dataset.

Table 2: Number of Observations for Categories in the Diphenhydramine Dataset

Group	ME	DUI
1 (diphenhydramine only)	13	26
2 (diphenhydramine and ethanol)	2	31
3 (any other combination with diphenhydramine)	87	54
Total	102	111

Table 3: Demographics on the Diphenhydramine Dataset

	Source Dataset	ME		DUI	
	Drug Pattern Group	1	Combined (2&3)	1	Combined (2&3)
Race	White	--	--	13 (11.7%)	54 (48.6%)
	Black	--	--	3 (2.7%)	3 (2.7 %)
	Other	--	--	0	0
	Missing	13 (12.7%)	89 (87.3%)	10 (9%)	28 (25.2%)
Gender	Male	7 (6.9%)	32 (31.3%)	11 (9.9%)	35 (31.5%)
	Female	6 (5.9%)	57 (55.9%)	5 (4.5%)	23 (20.7%)
	Missing	0	0	10 (9%)	27 (24.3%)
Age (in years)	Range	18-66	14-87	18-60	20-69

The normal quantile plots for the drug pattern groups and the source datasets for diphenhydramine revealed a non-normal distribution for most of the categories. This type non-normality is typically corrected by a log transformation, so a log transformation was employed on the concentration of diphenhydramine. The normal quantile plots (before and after log transformation) are seen in Figures 1, 2, 3, and 4. From these figures, the log transformation appears to work very well; thus, the ANOVA was run on the log-transformed concentration for diphenhydramine. The first normal quantile plot for each of these figures shows the data for diphenhydramine concentration before the log transformation.

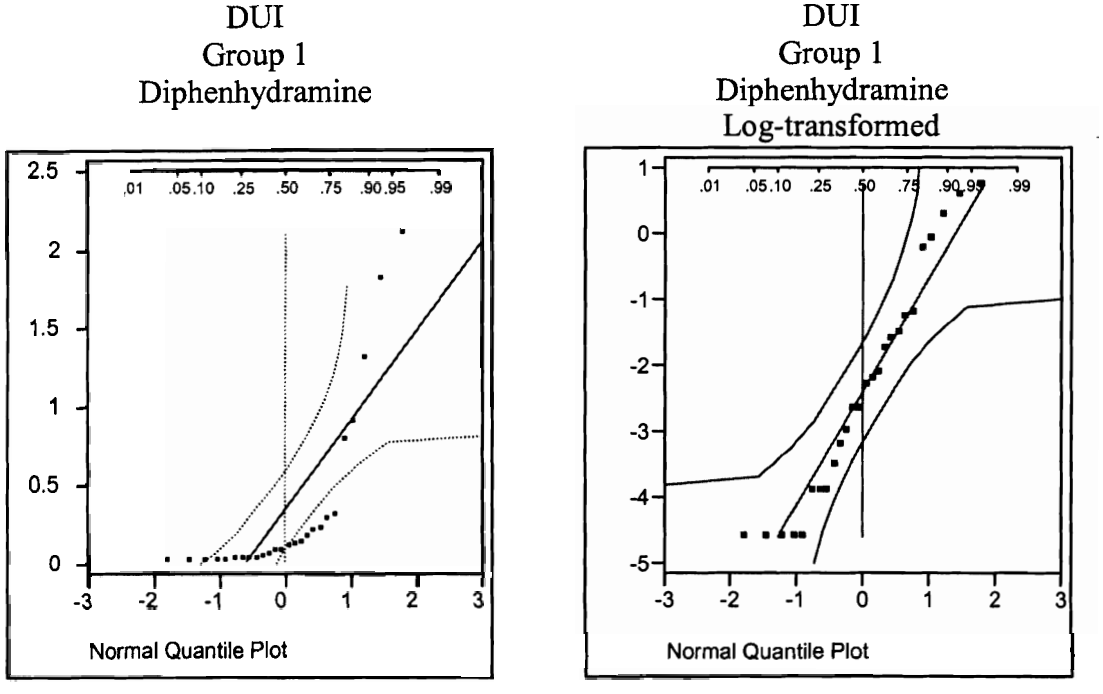


Figure 1: Normal quantile plots for DUI Group 1 (diphenhydramine only) before and after log transformation

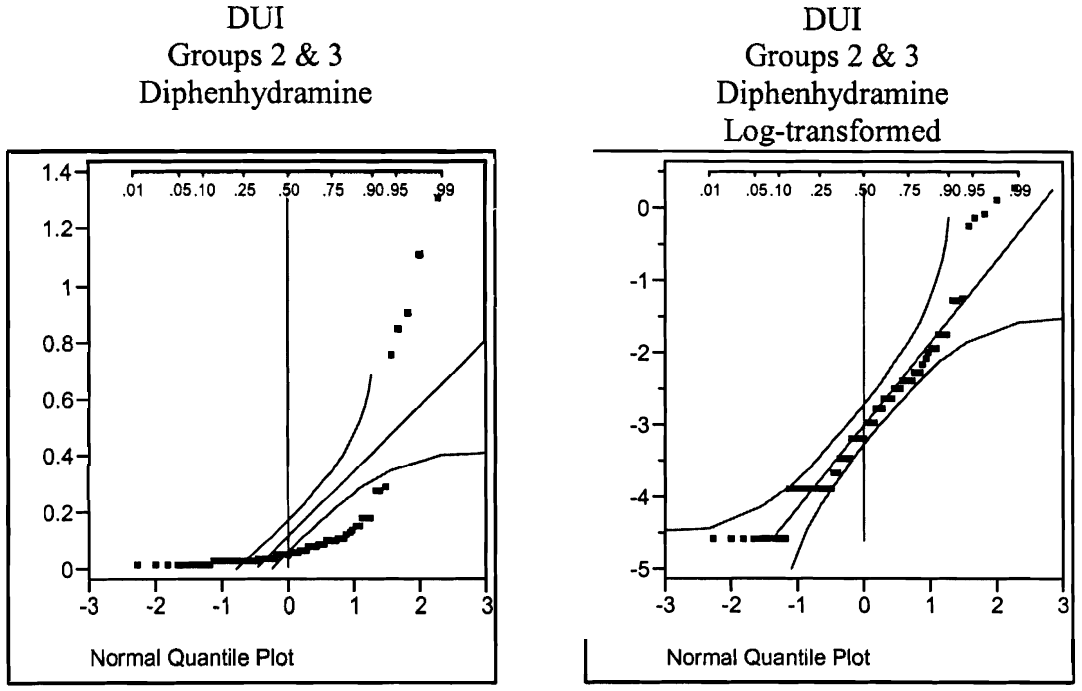


Figure 2: Normal quantile plots for DUI Groups 2 & 3 (diphenhydramine with any drug – including ethanol) before and after log transformation

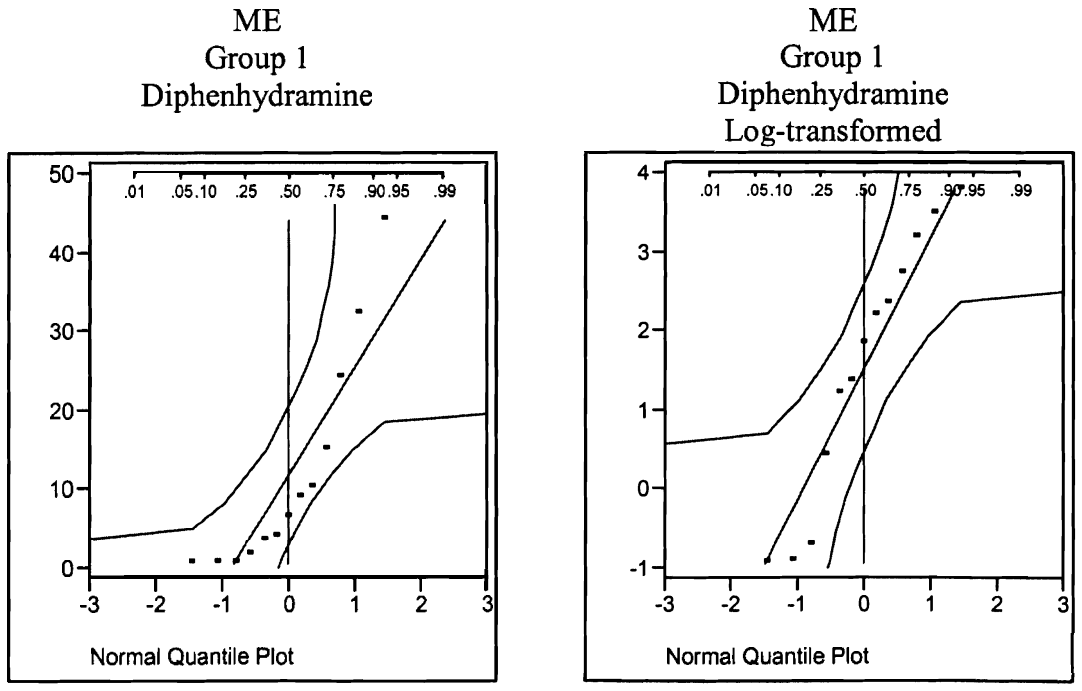


Figure 3: Normal quantile plots for ME Group 1 (diphenhydramine only) before and after log transformation

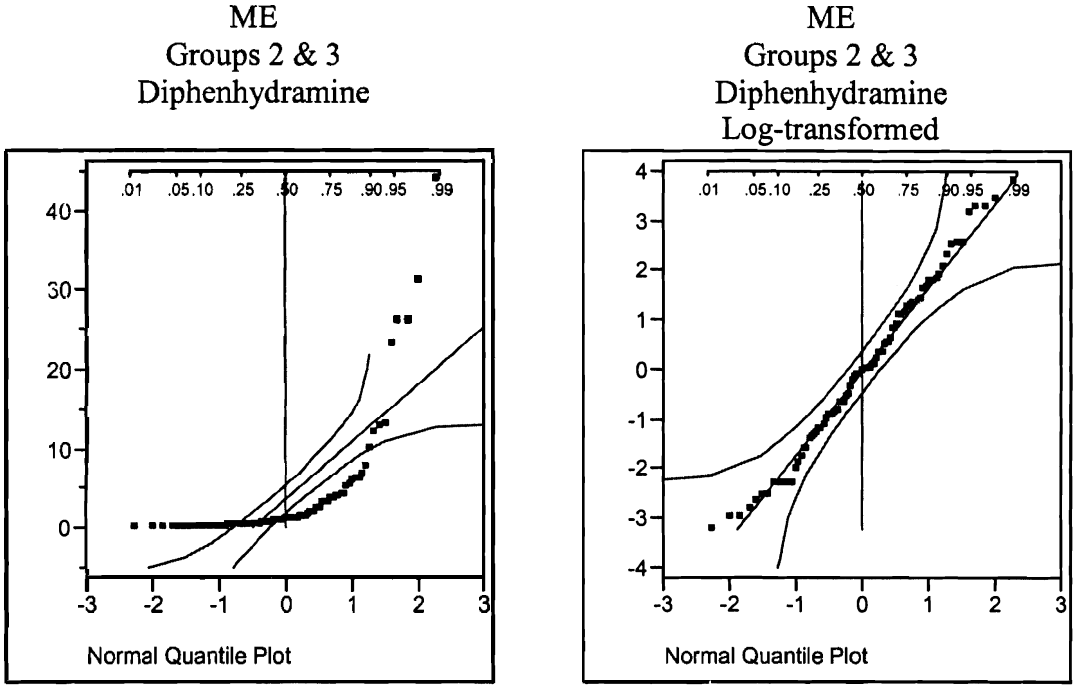


Figure 4: Normal quantile plots for ME Groups 2 & 3 (diphenhydramine with any drug – including ethanol) before and after log transformation

The box plots seen in Figure 5 are the set of box plots for the concentration of diphenhydramine for Group 1 for log-transformed and untransformed data. Each of the two sets of these box plots has been plotted on the same scale to simplify visual comparison. Figure 6 shows an expanded box plot for DUI Group 1 (untransformed data). Figure 7 shows the pair of box plots for the combined group – Groups 2 & 3 – for the ME and DUI datasets. The box plots in Figures 5 and 7 show the concentrations of the log of diphenhydramine concentration (left side) for each of the drug pattern groups in each of the source datasets, and the untransformed data (right side). Figure 8 shows an expanded box plot for the combined drug pattern group for DUI where the y-axis is modified to accommodate the inspection of the distribution.

For ME Group 1, the mean (with log transformation) is 1.53 log (mg/L) units (± 0.413 standard errors) which is equivalent to 4.62 mg/L, and the mean of DUI Group 1 is -2.406 log (mg/L) units (± 0.292 standard errors) [0.09 mg/L]. The range of the concentrations for the ME Group 1 was -0.942 log (mg/L) units to 3.784 log (mg/L) units [0.39 mg/L to 44 mg/L], and the range for the DUI dataset for Group 1 was -4.605 log (mg/L) units to 0.742 log (mg/L) units [0.01 mg/L to 2.1 mg/L]. (These means and standard errors are from the log-transformed concentrations and from untransforming those means; they will not necessarily match the raw means and standard errors seen in the table with the results of the Tukey tests.) There does seem to be some overlap between the two datasets in Group 1, but this is not enough evidence to tell whether or not there is a significant difference between the two datasets.

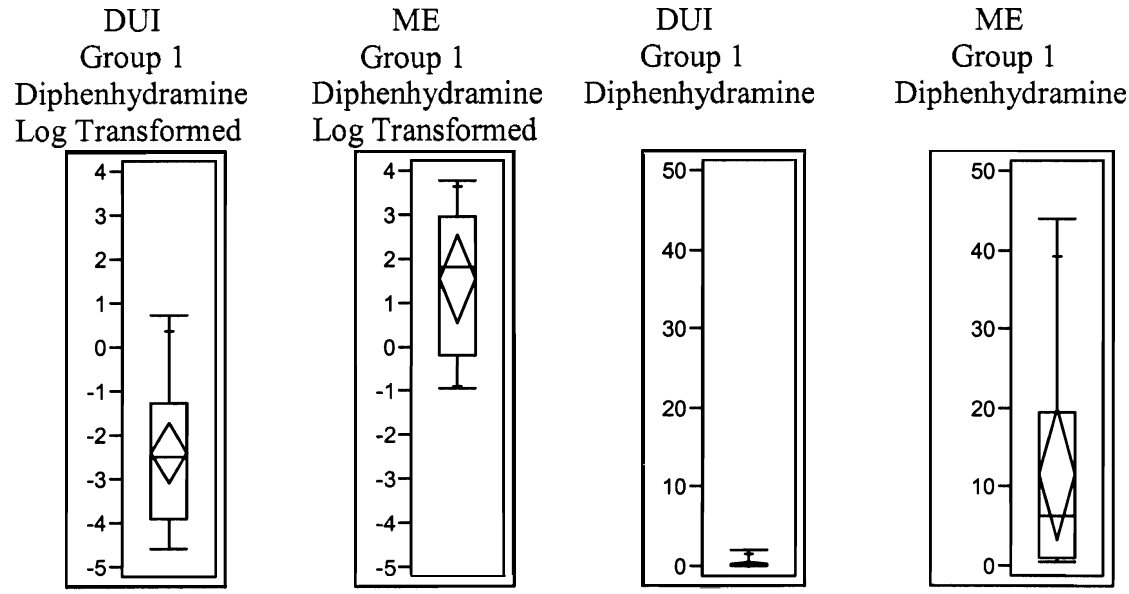


Figure 5: Box plots of Group 1 (diphenhydramine only) for DUI and ME – log transformed and untransformed data

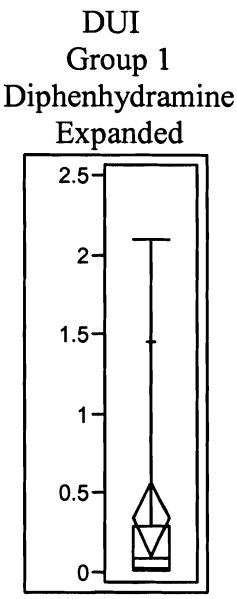


Figure 6: Expanded box plot for DUI Group 1 (diphenhydramine only) – untransformed data

Figure 7 shows the pair of box plots for log-transformed and untransformed data for Groups 2 & 3 combined. The mean of the concentration for the ME combined drug pattern group is -0.05 log (mg/L) units (± 0.158 standard errors) [0.95 mg/L] and the range is from -0.219 log (mg/L) units [0.04 mg/L] to 3.784 log mg/L [44.0 mg/L]. The mean of concentration for diphenhydramine for the DUI combined drug pattern group is -3.006 log (mg/L) units (± 0.162 standard errors) [0.05 mg/L] and the range is from -4.605 log (mg/L) units [0.01 mg/L] to 0.262 log (mg/L) units [1.3 mg/L]. The low ends of each of these ranges are close, so there is a definite overlap, but the high end of the ME range is much larger than that of the DUI range.

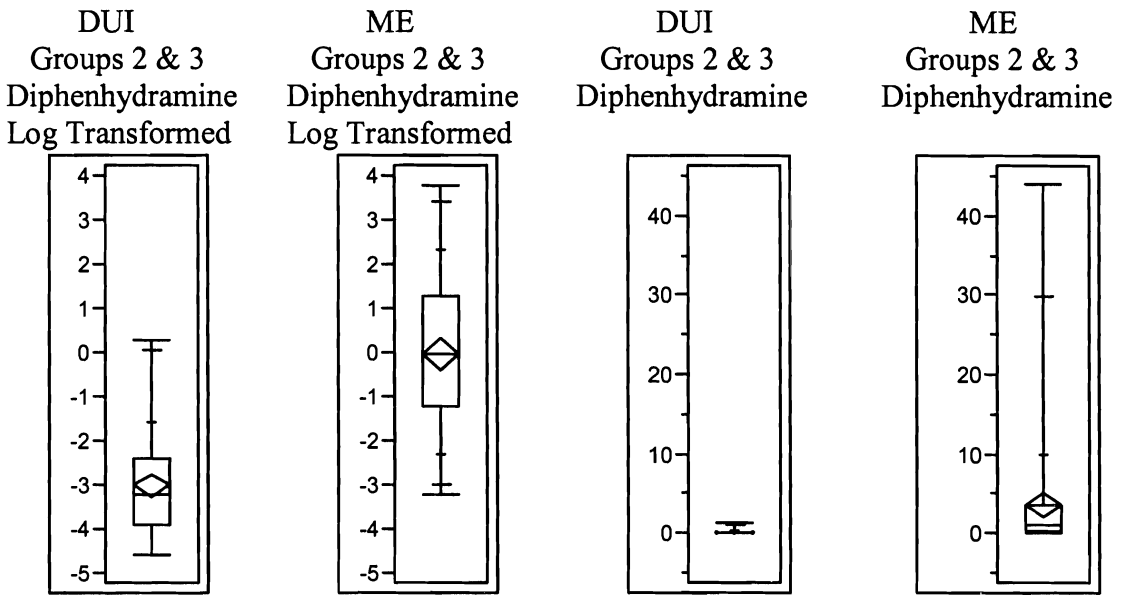


Figure 7: Box plots for Groups 2 & 3 (diphenhydramine with any drug - including ethanol) for DUI and ME – log transformed and untransformed data

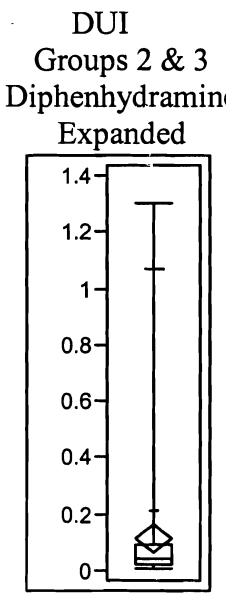


Figure 8: Expanded box plot for DUI Groups 2 & 3 (diphenhydramine with any drug - including ethanol) – untransformed data

When the ANOVA model was fit to the diphenhydramine data it was determined that the interaction between the `group` and `database` fields was significant at the pre-

determined $\alpha = 0.10$ level (p-value = 0.0793); therefore, it was kept in the model, and assessment of the differences in the blood concentrations between the drug pattern groups and the source datasets was within the interaction. The plot of the means in Figure 9 shows the interaction of the two fields `group` and `database`. From Figure 9, it can be seen that concentration of diphenhydramine between the ME and DUI datasets in Group 1 is a larger difference than the difference between the ME and DUI datasets in the combined drug pattern group.

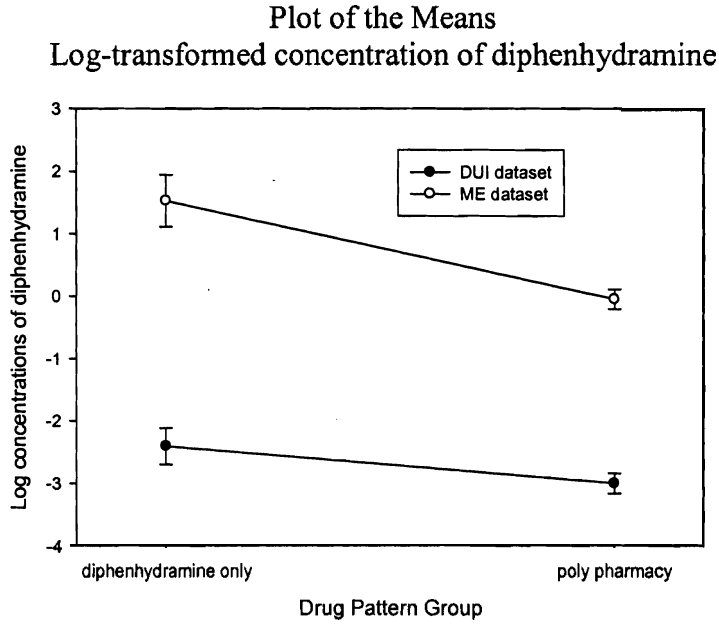


Figure 9: Plot of the means for log-transformed diphenhydramine - interaction

By interpreting the assessment of the differences in the blood concentrations between the drug pattern groups and the source datasets through the interaction, Tukey's test for multiple comparisons revealed that the ME and DUI dataset are statistically significantly different from each other both in Group 1 and the combined drug pattern group. Within the ME dataset, both drug pattern groups are different from each other, but in the DUI dataset,

Group 1 and the combined drug pattern group (Groups 2 & 3) are not different from each other. These results of the Tukey's test are summarized in Table 4 where levels not connected by the same letter are statistically significantly different from each other. Table 4 also shows the means and standard errors for the untransformed concentration of diphenhydramine.

Table 4: Levels of drug pattern groups and source datasets for diphenhydramine from Tukey's test

Group	Letter	Mean(Std. Error)
ME, Group 1	A	11.56(3.82)
ME, Combined Group	B	3.58(0.77)
DUI, Group 1	C	0.37(0.112)
DUI, Combined Group	C	0.115(0.025)

The residual by predicted plot for diphenhydramine is seen in Appendix C and does not show any pattern between the residuals and the predicted values from the model. It also seems that the assumption of constant variance is met because all four categories appear to have roughly the same span on this plot. A normal quantile plot of the residuals (Appendix C) was also inspected for deviations from normality, but for the diphenhydramine dataset, the residuals followed an approximately normal distribution.

A clinical range of 8-31 mg/L of diphenhydramine is considered a fatal blood concentration for an individual[26]. Using this limit, there were no individuals in the DUI dataset with a "lethal" amount of diphenhydramine in their system, and thus, there is no evidence to suggest a development of tolerance to diphenhydramine. The high end of the fatal blood concentration for diphenhydramine was 31 mg/L, but there were four cases in the ME dataset that had a concentration of 31 mg/L or higher, with the highest being 44

mg/L. All four of these cases listed "suicide" as a manner of death, but they might still be of clinical interest for a case study. Two of these four individuals had only diphenhydramine detected, one had diphenhydramine and ethanol, and the last had a poly pharmacy of drugs in their system.

3.3.2 Cocaine (Level II)

"Cocaine is a stimulant of the central nervous system and an appetite suppressant, creating what has been described as a euphoric sense of happiness and increased energy,"[27]. Cocaine is most often used as a "recreational drug", but it is sometimes used as a topical anesthetic for certain types of surgery[27]. One of the problems with determining information concerning cocaine in deaths is that "since cocaine in blood rapidly hydrolyzes to benzoylecgonine, cocaine concentrations determined in postmortem blood may not reflect the presence or true concentration of cocaine in the body at the time of death,"[28]. The main metabolite of cocaine is benzoylecgonine, and this metabolite has a longer half-life than cocaine, so the metabolite provides additional information when doing a drug test that searches for cocaine. The DUI drug tests were all administered via a blood sample, so using the ME concentrations from brain or tissues other than blood were not an appropriate comparator to the DUI blood concentrations; therefore, only the blood concentrations from the ME dataset were used in the comparison.

Using the same inclusion and exclusion criteria mentioned in Section 3.2, the final cocaine dataset contained 691 cases and 577 variable fields. Of the 691 observations, 359 came from the ME dataset and 332 from the DUI dataset. Detecting benzoylecgonine in a person's system could lead to the cause of death being "cocaine poisoning" (for an ME

case) or showing that the individual had taken cocaine prior to being stopped for DUI. If an individual had benzoylecgonine detected and not cocaine, then he or she was excluded from the cocaine dataset because a concentration amount of cocaine could not be established. This means that even if an individual had taken cocaine (as evidenced by the presence of benzoylecgonine), if no cocaine was determined from the blood sample, then this individual was excluded from the cocaine dataset.

Table 5 details the number of observations for the three drug pattern groups and the two datasets (ME and DUI) for the 691 observations. The next table, Table 6, shows the breakdown of `race`, `gender`, and `age` for the two source datasets and the three drug pattern groups. The information in Table 6 shows the frequencies for each of the categories except in the field `age`, which is displayed as a range from youngest to oldest. The demographic information shown in Table 6 is only displayed for descriptive purposes, and it is not used in the analysis. Because all six categories met the minimum number of observations requirement, no drug pattern groups were combined or removed for the analysis of the cocaine dataset.

Table 5: Number of Observations for Categories in the Cocaine Dataset

Group	ME	DUI
1 (cocaine only)	129	122
2 (cocaine and ethanol)	15	146
3 (any other combination with cocaine)	215	64
Total	359	332

Table 6: Demographics on the Cocaine Dataset

	Source Dataset	ME			DUI		
	Drug Pattern Group	1	2	3	1	2	3
Race	White	--	--	--	50 (15.1%)	50 (15.1%)	31 (9.3%)
	Black	--	--	--	25 (7.5%)	38 (11.4%)	13 (3.9%)
	Other	--	--	--	0	1 (.3%)	0
	Missing	129 (35.9%)	15 (4.2%)	215 (59.9%)	47 (14.2%)	57 (17.2%)	20 (6%)
Gender	Male	101 (28.1%)	14 (3.9%)	152 (42.3%)	54 (16.3%)	69 (20.8%)	35 (10.5%)
	Female	28 (7.8%)	1 (0.3%)	63 (17.5%)	22 (6.6%)	20 (6%)	9 (2.7%)
	Missing	0	0	0	46 (13.9%)	57 (17.2%)	20 (6%)
Age (in years)	Range	16-68	19-46	19-64	19-53	16-64	17-55

The majority of the normal quantile plots for the drug pattern groups showed non-normal distributions. Figures 10, 11, and 12 show the normal quantile plots for the three drug pattern groups for the DUI dataset before and after the log transformation. For the ME dataset drug pattern groups, Figures 13, 14, and 15 show the normal quantile plots for both the untransformed and log-transformed data. Using the log transformation for the concentration of cocaine, the data appears to be approximately normally distributed. This log transformation was used for the ANOVA model that was used for analysis of the cocaine dataset.

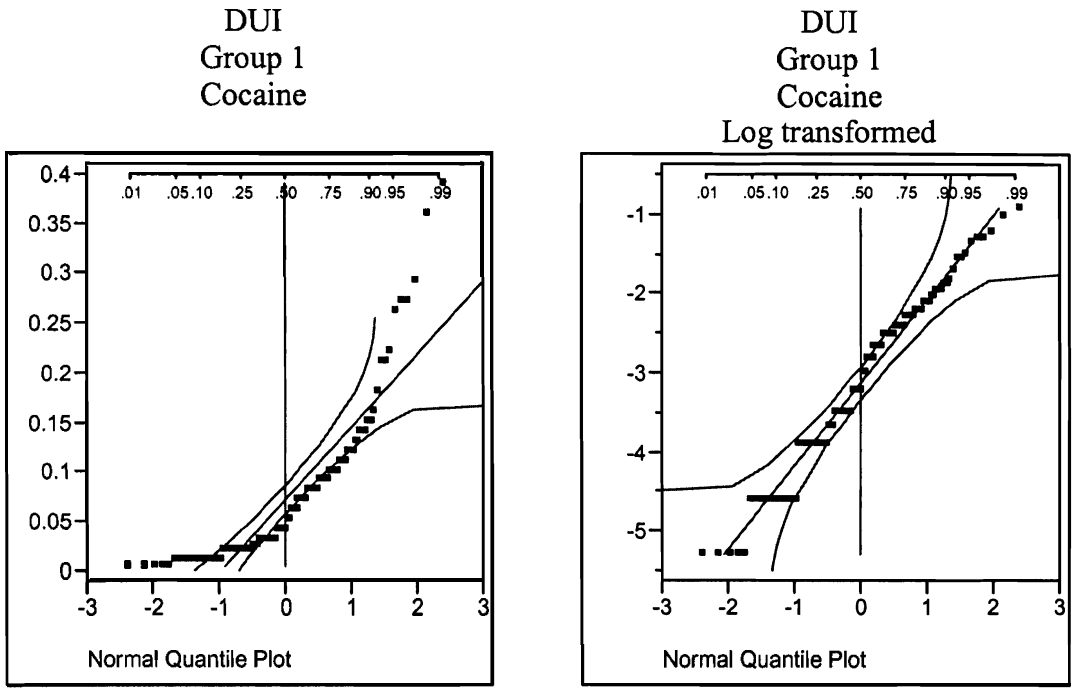


Figure 10: Normal quantile plots for DUI Group 1 (cocaine only) before and after log transformation

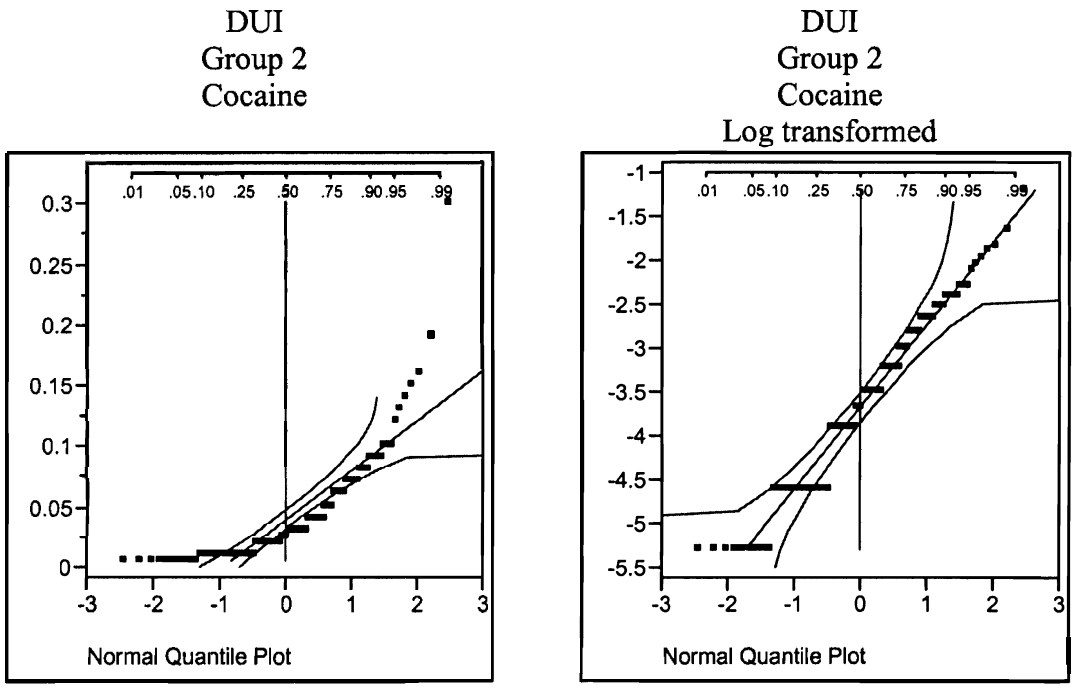


Figure 11: Normal quantile plots for DUI Group 2 (cocaine and ethanol) before and after log transformation

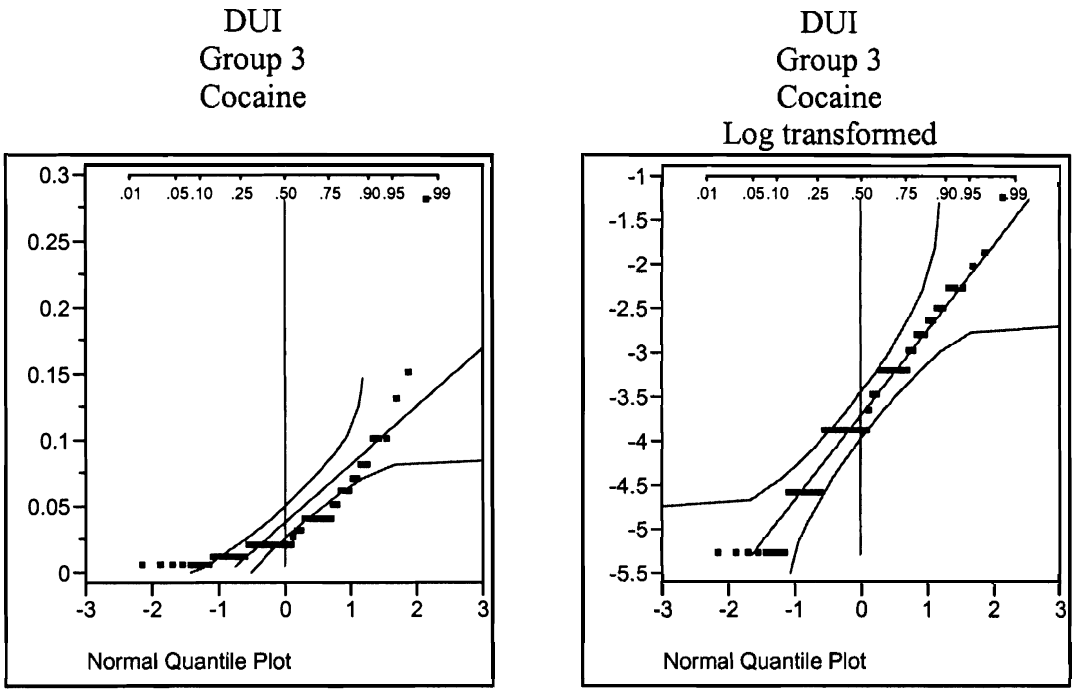


Figure 12: Normal quantile plots for DUI Group 3 (cocaine with any other combination of drug) before and after log transformation

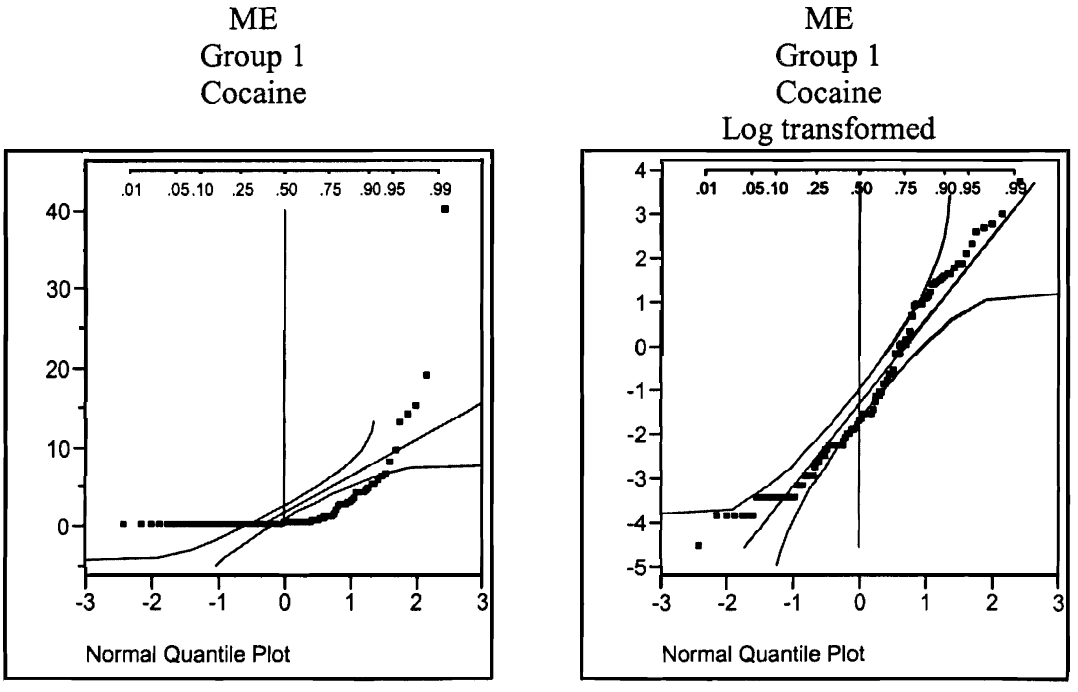


Figure 13: Normal quantile plots for ME Group 1 (cocaine only) before and after log transformation

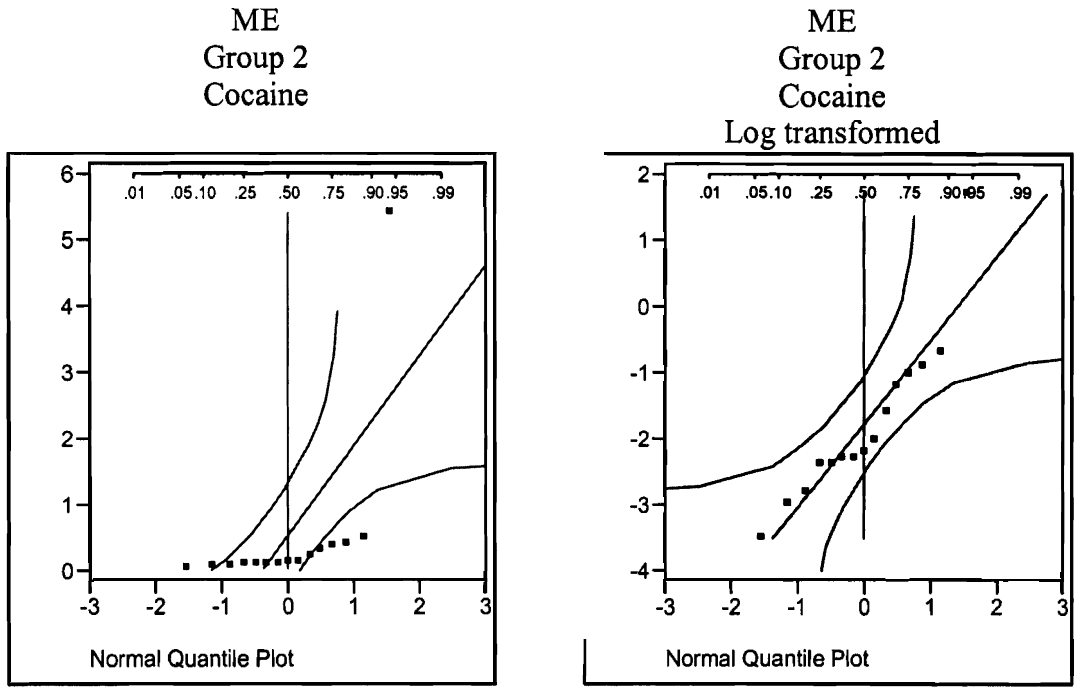


Figure 14: Normal quantile plots for ME Group 2 (cocaine and ethanol) before and after log transformation

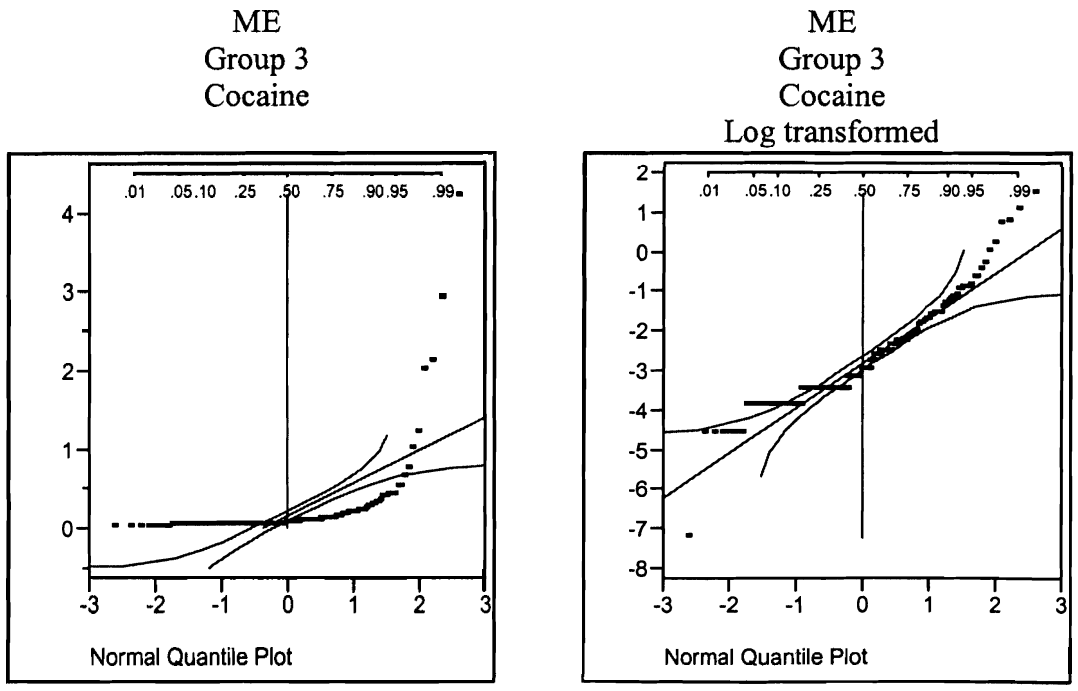


Figure 15: Normal quantile plots for ME Group 3 (cocaine with any other combination of drug) before and after log transformation

Box plots were created to compare the concentration of cocaine between the three drug pattern groups and between the two source datasets. The box plots shown in the following figures are of the log-transformed data and the untransformed data for each dataset (ME and DUI) and each drug pattern group. Figure 16 shows the comparisons between the two datasets for Group 1 (cocaine alone) for both the log-transformed and untransformed concentration amounts, and Figure 17 shows an expanded version of the untransformed data for DUI Group 1. When the log-transformed data was used, there was no need to expand any of the box plots, but when the untransformed data was used and the scales for the ME and DUI datasets were set to the same scale, the DUI box plot of the untransformed data needed to be expanded to view the details of the plot. The mean for the ME dataset for Group 1 is -1.333 log (mg/L) units (± 0.111 standard errors) [0.264 mg/L] with the range from -4.605 log (mg/L) units [0.01 mg/L] to 3.689 log (mg/L) units [40.0 mg/L]. The mean for DUI Group 1 is -3.135 log (mg/L) units (± 0.114 standard errors) [0.04 mg/L] with a range of -5.298 log (mg/L) units [0.005 mg/L] to -0.942 log (mg/L) units [0.39 mg/L]. For Group 1, the majority of the DUI cases fit into the ME range, showing some sort of overlap.

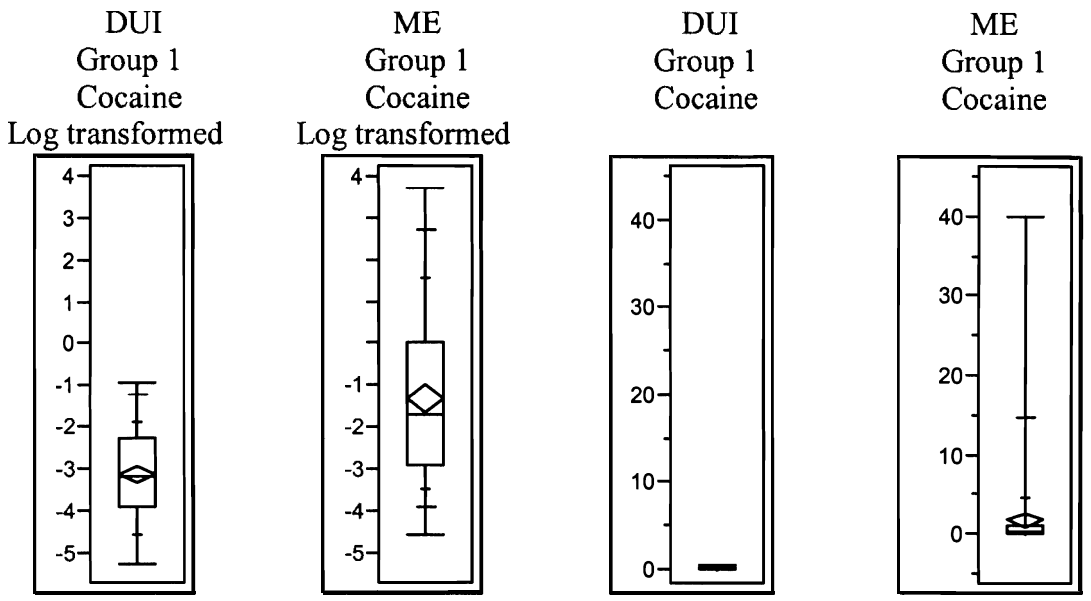


Figure 16: Box plots for Group 1 (cocaine only) for DUI and ME – log-transformed and untransformed data

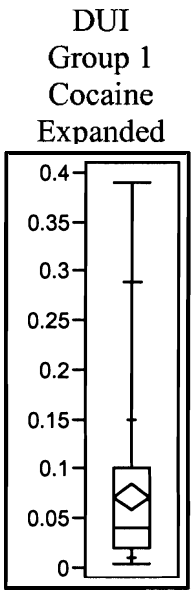


Figure 17: Expanded box plot for DUI Group 1 (cocaine only) – untransformed data

In the next figure, Figure 18, the two pairs of box plots for Group 2 of the cocaine dataset are shown; Figure 19 shows the expanded box plot for the DUI untransformed data. The range for the DUI dataset is larger than it was in Group 1, but the overlap between the ME

and DUI datasets is still there. For the ME dataset Group 2, the mean is -1.783 log (mg/L) units (± 0.324 standard errors) [0.17 mg/L], and the DUI mean for Group 2 is -3.68 log (mg/L) units (± 0.104 standard errors) [0.025 mg/L]. The range for the ME dataset for Group 2 is from -3.507 log (mg/L) units [0.03 mg/L] to 1.686 log (mg/L) units [5.40 mg/L], and the range for DUI Group 2 is from -5.298 log (mg/L) units [0.005 mg/L] to -1.204 log (mg/L) units [0.3 mg/L]. These ranges have some overlap, but the range for this drug pattern group in the DUI dataset was smaller than that for the ME dataset.

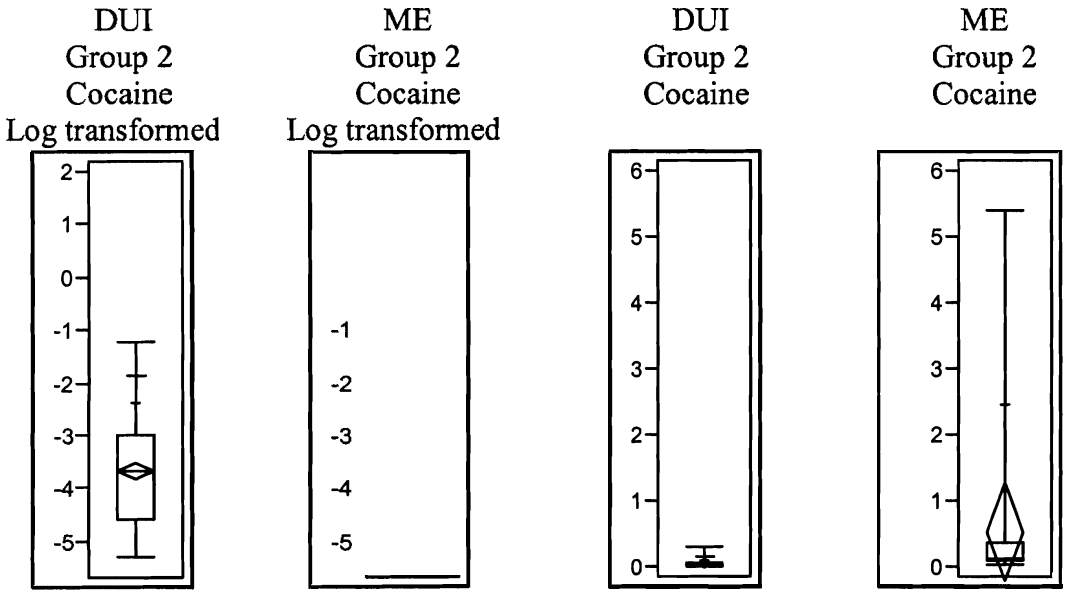


Figure 18: Box plots for Group 2 (cocaine and ethanol) for DUI and ME – log-transformed and untransformed data

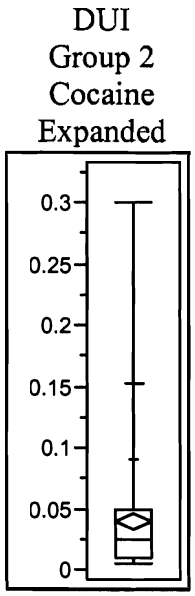


Figure 19: Expanded box plot for DUI Group 2 (cocaine and ethanol) – untransformed data

The last figures for the cocaine dataset show the box plots for the log-transformed and untransformed concentrations for Group 3. Figure 20 shows that the range for DUI Group 3 can be completely encompassed within the range for ME Group 3. Figure 21 shows an expanded box plot for DUI Group 3 of the untransformed data. The means for the ME and DUI dataset are -2.812 log (mg/L) units (± 0.086 standard errors) [0.06 mg/L] and -3.723 log (mg/L) units (± 0.157 standard errors) [0.024 mg/L] respectively. The range for Group 3 in the DUI dataset is from -5.298 log (mg/L) units [0.005 mg/L] to -1.273 log (mg/L) units [0.28 mg/L], while the range for Group 3 in the ME dataset is from -7.264 log (mg/L) units [0.0007 mg/L] to 1.435 log (mg/L) units [4.2 mg/L]. (These means and ranges are determined from the log-transformed concentrations of cocaine and by untransforming these means; they will not necessarily match the means for the untransformed concentrations that are seen in the table with the Tukey results.)

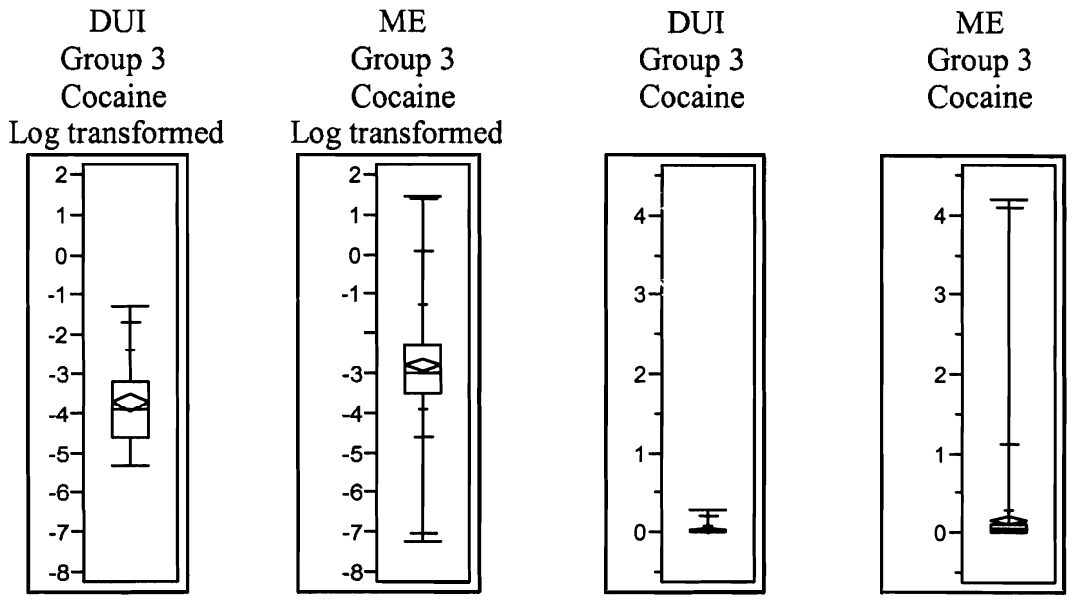


Figure 20: Box plots for Group 3 (cocaine with any other combination of drug) for DUI and ME – log-transformed and untransformed data

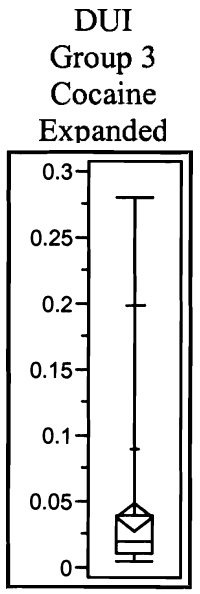


Figure 21: Expanded box plot for DUI Group 3 (cocaine with any other combination of drug) – untransformed data

Based only on the box plots, it seems that there might be a difference between the two source datasets for some drug pattern groups, but perhaps, not all, and that perhaps the

drug pattern groups are different from each other, within the same dataset. An ANOVA model was fit to the log-transformed data to see if these conjectures were correct.

When the ANOVA model was fit to the cocaine dataset, it was determined that the interaction between `group` and `database` was significant (p-value = 0.0004). Since this interaction was significant, assessments of the differences in the blood concentrations between drug pattern groups and source datasets were tested through the interaction. The plot of the means for the cocaine dataset is seen in Figure 22, and it shows that the difference in concentrations between the source datasets (ME and DUI) for Group 3 is different than the difference in concentrations between the source datasets for Groups 1 and 2. Table 7 shows the results of the Tukey's test with different letters signifying statistical significant differences between the different levels; categories that do not share the same letter are statistically significantly different from each other. This table shows that ME Group 1 and ME Group 2 are different from all of the drug pattern groups in the DUI dataset, but ME Group 3 is not different from DUI Group 1. Within the ME dataset, Groups 1 and 2 are not different from each other, but both are different from ME Group 3. For the DUI dataset, Groups 2 and 3 are not different from each other, but both are different from Group 1. Table 7 also shows the raw means and standard errors (untransformed concentrations) for cocaine; the units for these means and standard errors are mg/L.

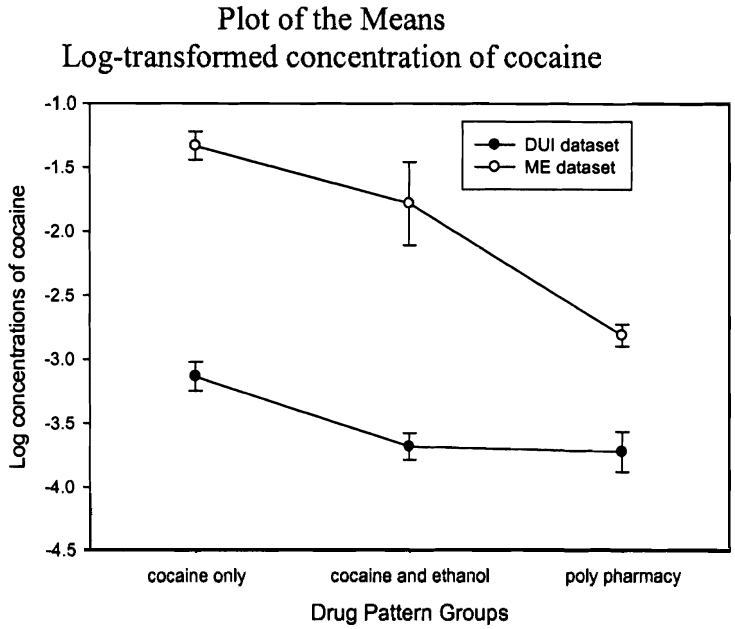


Figure 22: Plot of the means for log-transformed cocaine - interaction

Table 7: Levels of drug pattern groups and source datasets for cocaine from Tukey's test

Group	Letter	Mean(Std. Error)
ME, Group 1	A	1.71(0.403)
ME, Group 2	A	0.528(0.35)
ME, Group 3	B	0.149(0.028)
DUI, Group 1	B	0.071(0.007)
DUI, Group 2	C	0.039(0.003)
DUI, Group 3	C	0.038(0.005)

The residual by predicted plot for the cocaine dataset is seen in Appendix C, and this plot shows no obvious pattern of the residuals with the predicted values from the model. Because there is no pattern of the residuals and there does not seem to be a problem with constant variance, then the modeling assumptions are verified. The constant variance assumption is met because for each of the combinations, the variance appears to be the same. The adequacy of the model is verifed through the fact that there not a systematic pattern of residuals and the normal quantile plot. The normal quantile plot of the residuals

for the ANOVA model, also seen in Appendix C, shows an approximately normal distribution because the data points fall along the diagonal, so the normality assumption in not violatedin the fitted model.

Using 0.9 - 21 mg/L as the clinically defined, fatal blood concentration for cocaine[26], no cases in the DUI dataset fell within this range, but there was one case in the ME dataset that exceeded the upper end of the range (21 mg/L) by having 40 mg/L. This case was a single drug overdose and might be of interest in a case study.

3.3.3 Oxycodone (Level II)

Oxycodone, commonly known as OxyContin ®, Percocet ®, or Percodan ® is a "central nervous system depressant that appears to work through stimulating the opioid receptors found in the central nervous system that activate responses ranging from analgesia to respiratory depression to euphoria,"[29]. According to an article from www.streetdrugs.org, "people who take the drug repeatedly can develop a tolerance or resistance to the drug's effects,"[29].

The dataset created for oxycodone consisted of 487 individuals and 577 fields. This dataset was created using the same criteria as the previous datasets and excluded those individuals who did not die from an oxycodone-related drug death, those individuals not receiving a toxicological screen as a result of an autopsy, and those individuals in the DUI dataset that did not register any level of concentration for oxycodone. From the ME dataset, 263 individuals were included, and from the DUI dataset, 224 individuals were included. Table 8 shows the number of observations that appear in each drug pattern group (oxycodone only, oxycodone and ethanol, and oxycodone with any other combination of

drug) for each of the source datasets, and Table 9 shows the demographic descriptions for each of these drug pattern groups. In Table 8, the category for oxycodone and ethanol (Group 2) from the ME dataset had only 11 observations. The general rule of thumb for a minimum number of observations was 12[22], but the clinical decision was to leave the drug pattern groups as is, without combining Group 2 with any of the other drug pattern groups. One interesting fact to notice in Table 9 is that in ME Group 1 (oxycodone only), there was an individual with an age of zero; this might be an interesting individual to use for a case study, once the age is verified.

Table 8: Number of Observations for Categories in the Oxycodone Dataset

Group	ME	DUI
1 (oxycodone only)	53	31
2 (oxycodone and ethanol)	11	68
3 (any other combination with oxycodone)	199	125
Total	263	224

Table 9: Demographics on theOxycodone Dataset

	Source Dataset	ME			DUI		
	Drug Pattern Group	1	2	3	1	2	3
Race	White	--	--	--	18 (8%)	46 (20.6%)	77 (34.4%)
	Black	--	--	--	1 (0.4%)	3 (1.3%)	3 (1.3%)
	Other	--	--	--	0	0	0
	Missing	53 (20.1%)	11 (4.2%)	199 (75.7%)	12 (5.4%)	19 (8.5%)	45 (20.1%)
Gender	Male	40 (15.2%)	11 (4.2%)	120 (45.6%)	11 (4.9%)	36 (16.1%)	58 (25.9%)
	Female	13 (4.9%)	0	79 (30.1%)	8 (3.5%)	13 (5.8%)	22 (9.8%)
	Missing	0	0	0	12 (5.4%)	19 (8.5%)	45 (20.1%)
Age (in years)	Range	0-57	20-46	19-73	20-57	17-59	18-68

The normal quantile plots for the concentration of oxycodone showed data that did not appear normally distributed, so a log transformation was done on the concentration of oxycodone. The normal quantile plots, before and after the log transformation, are seen in Figures 23, 24, 25, 26, 27, and 28. The first three figures show the normal quantile plots for the DUI dataset for Groups 1, 2, and 3, with the untransformed data on the left, and the log-transformed data on the right. Figures 26, 27, and 28 show the normal quantile plots for the ME dataset. As these figures show, the normal quantile plots for the log-transformed concentration appear approximately normally distributed, while the normal quantile plots for the untransformed data show data that is not normally distributed.

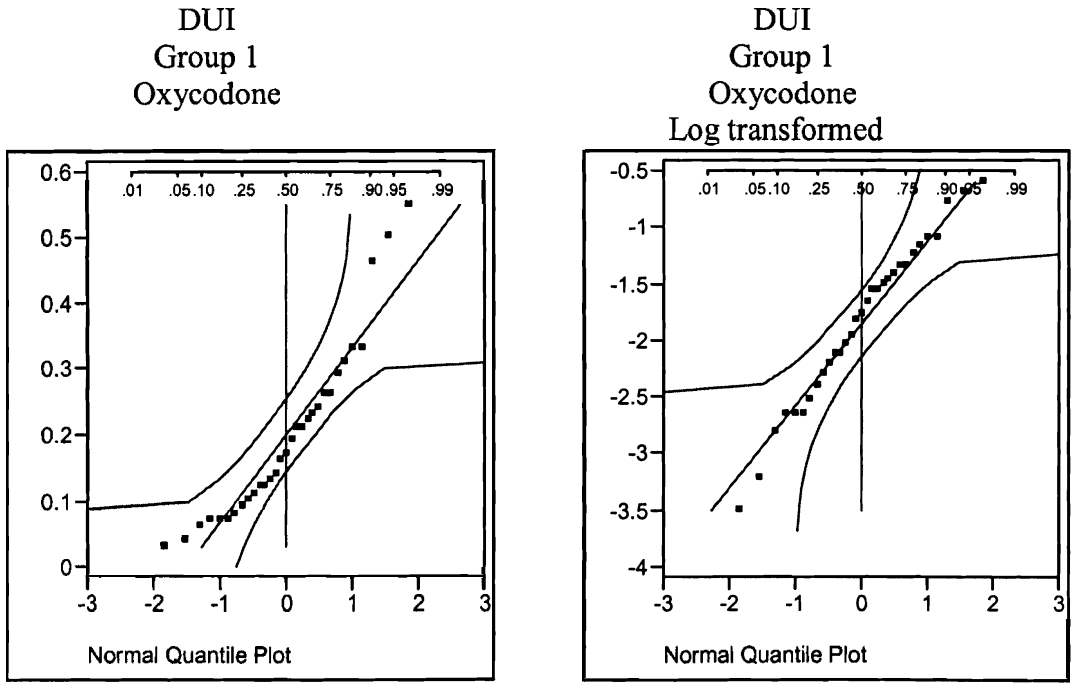


Figure 23: Normal quantile plots for DUI Group 1 (oxycodone only) before and after log transformation

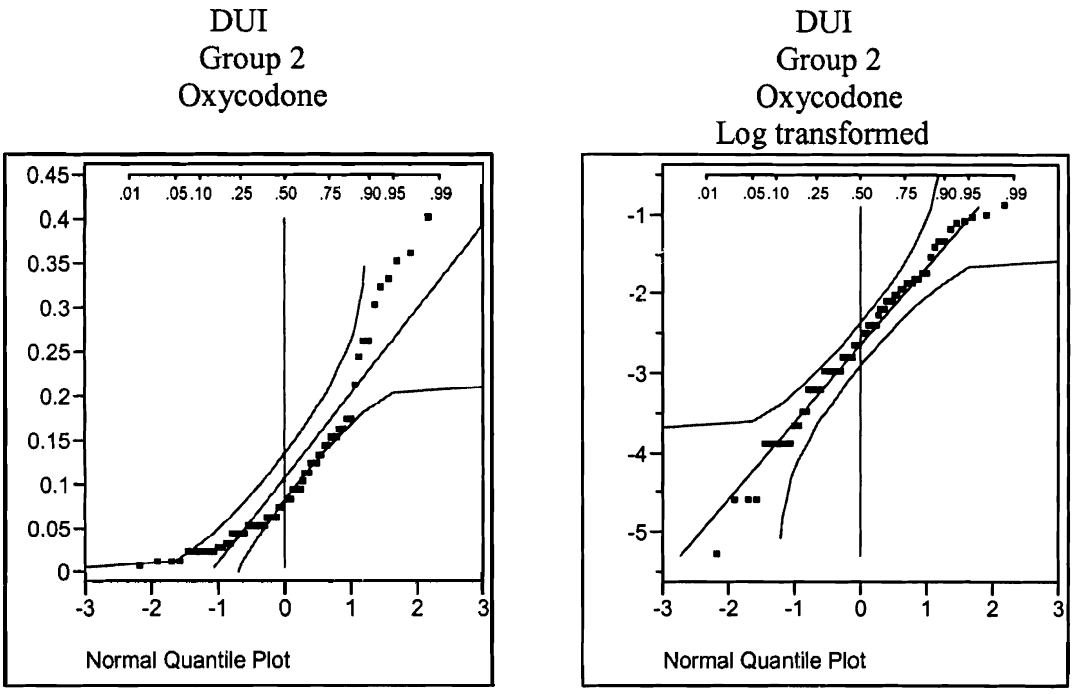


Figure 24: Normal quantile plots for DUI Group 2 (oxycodone and ethanol) before and after log transformation

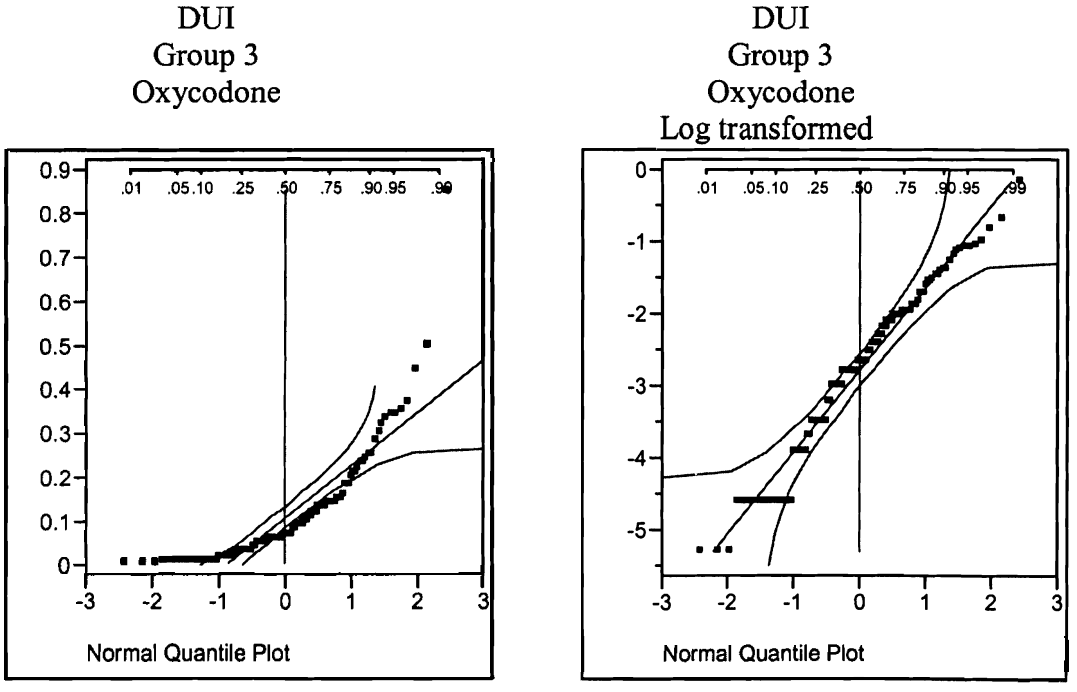


Figure 25: Normal quantile plots for DUI Group 3 (oxycodone with any other combination of drug) before and after log transformation

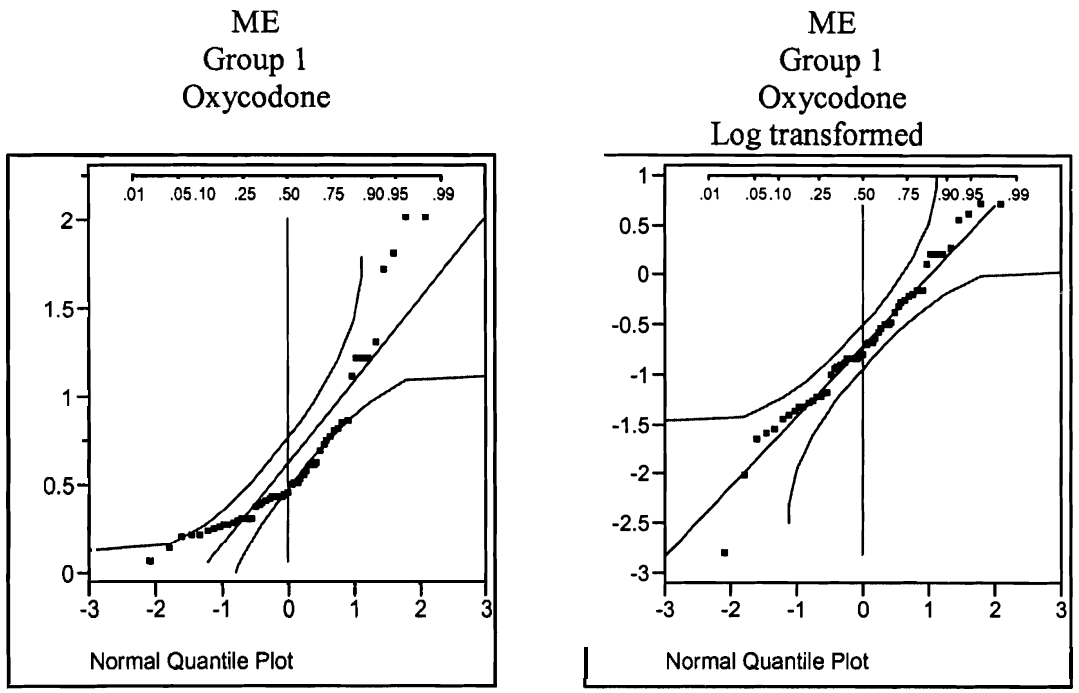


Figure 26: Normal quantile plots for ME Group 1 (oxycodone only) before and after log transformation

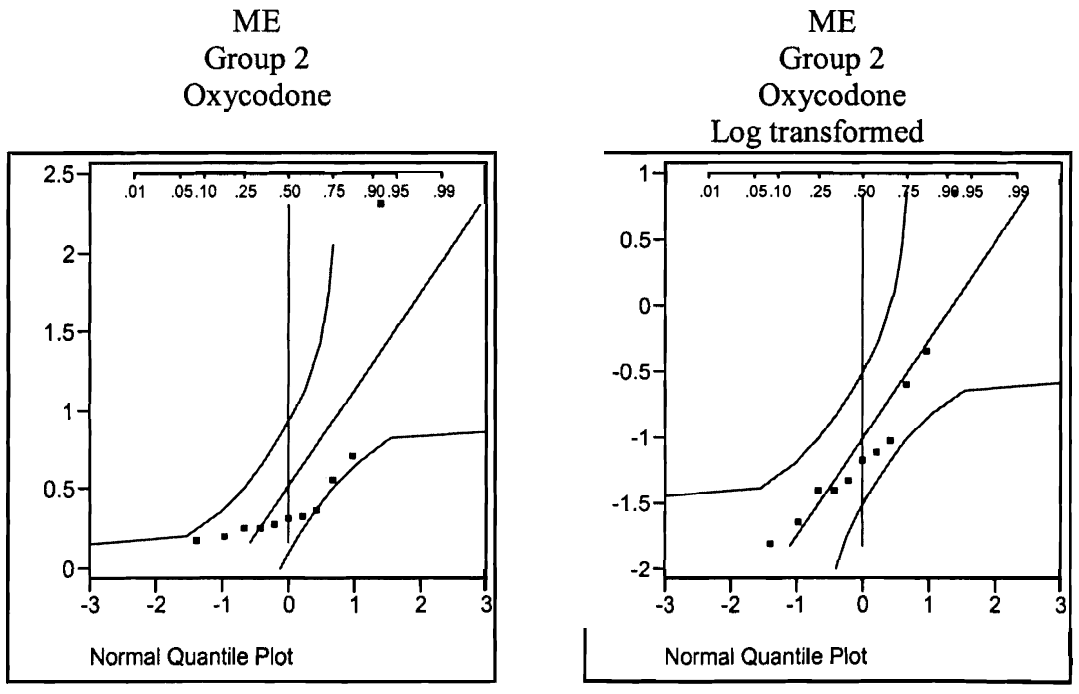


Figure 27: Normal quantile plots for ME Group 2 (oxycodone and ethanol) before and after log transformation

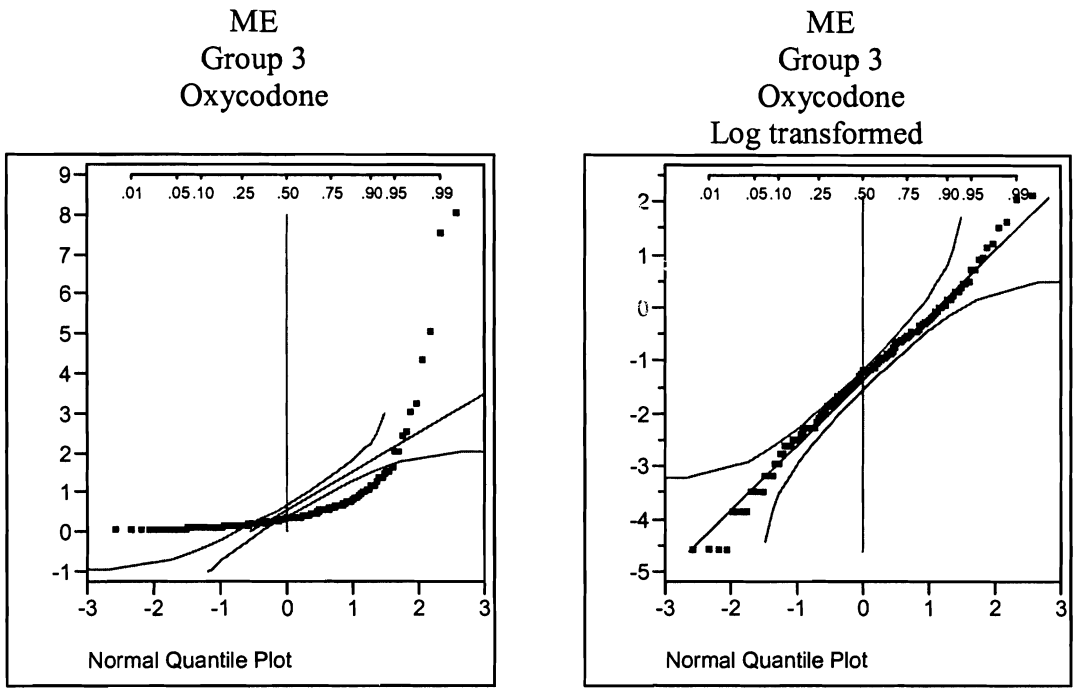


Figure 28: Normal quantile plots for ME Group 3 (oxycodone with any other combination of drug) before and after log transformation

The following figures show two pairs of box plots for the three drug pattern groups before and after the log transformation of cocaine. Figure 29 shows the pair of box plots for Group 1, and Figure 30 shows the expanded box plot for the DUI dataset for Group 1.

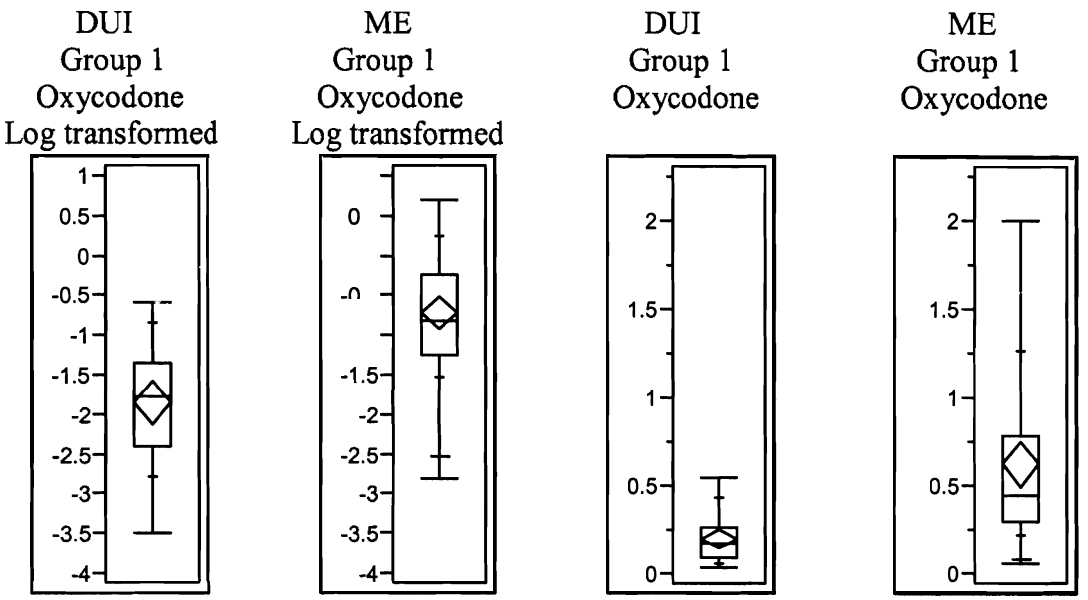


Figure 29: Box plots for Group 1 (oxycodone only) for DUI and ME – log-transformed and untransformed data

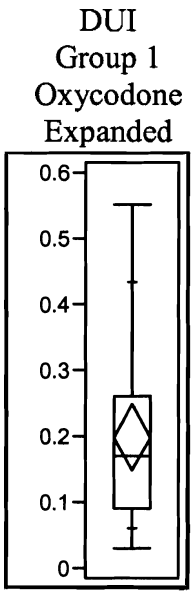


Figure 30: Expanded box plot for DUI Group 1 (oxycodone only) – untransformed data

The mean of DUI Group 1 is -1.85 log (mg/L) units (± 0.131 standard errors) [0.16 mg/L], ranging from -3.507 log (mg/L) units [0.03 mg/L] to -0.598 log (mg/L) units [0.55 mg/L]. The mean of ME Group 1 is -0.7185 log (mg/L) units (± 0.0974 standard errors) [0.49

mg/L] with a range from -2.813 log (mg/L) units [0.06 mg/L] to 0.693 log (mg/L) units [2.0 mg/L]. The majority of the cases in the DUI dataset for Group 1 fit inside the range determined by ME Group 1. The box plots seen in Figure 31 show Group 2 for the ME and DUI datasets with and without the log transformation; Figure 32 shows the box plot for untransformed DUI Group 2 on an expanded scale.

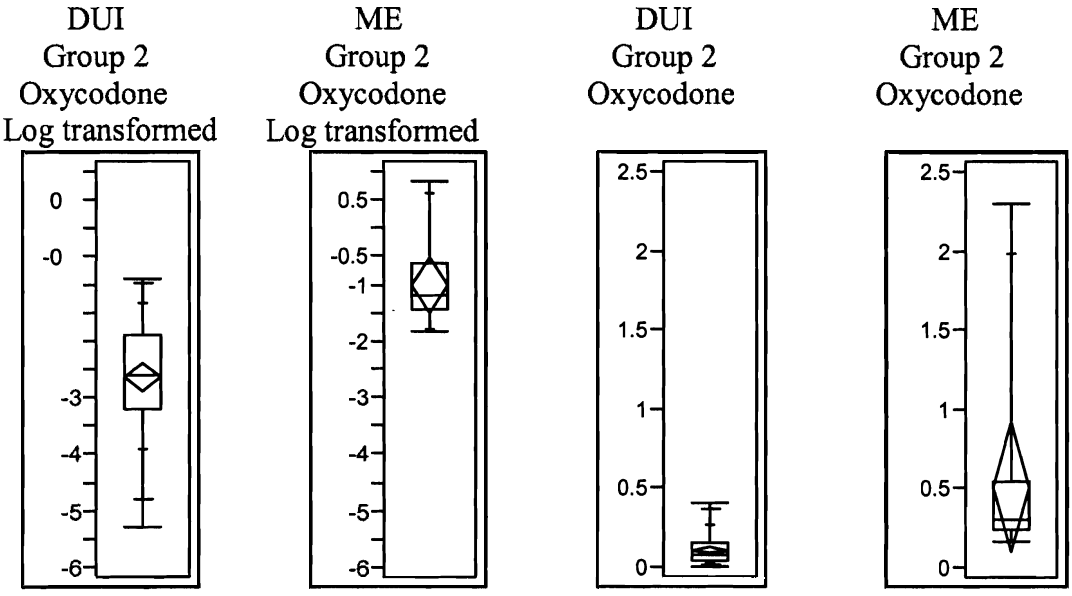


Figure 31: Box plots for Group 2 (oxycodone and ethanol) for DUI and ME – log-transformed and untransformed data

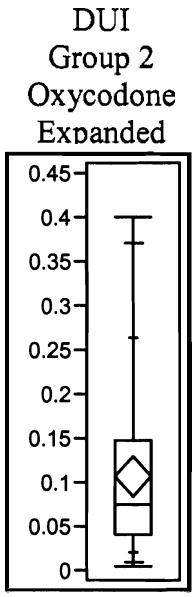


Figure 32: Expanded box plot for DUI Group 2 (oxycodone and ethanol) – untransformed data

For Group 2, the means for the DUI and ME datasets are -2.638 log (mg/L) units (± 0.118 standard errors) [0.07 mg/L] and -1.02 log (mg/L) units (± 0.226 standard errors) [0.36 mg/L], respectively. The range for the DUI dataset for this drug pattern group is from -5.298 log (mg/L) units [0.005 mg/L] to -0.916 log (mg/L) units [0.4 mg/L], while the ME dataset for Group 2 ranges from -1.833 log (mg/L) units [0.16 mg/L] to 0.833 log (mg/L) units [2.3 mg/L]. There seems to be some overlapping between the two datasets (ME and DUI) for Group 2.

The box plots for Group 3 for the oxycodone dataset are seen in Figures 33 and 34, with Figure 34 showing an expanded box plot for the untransformed data for DUI Group 3.

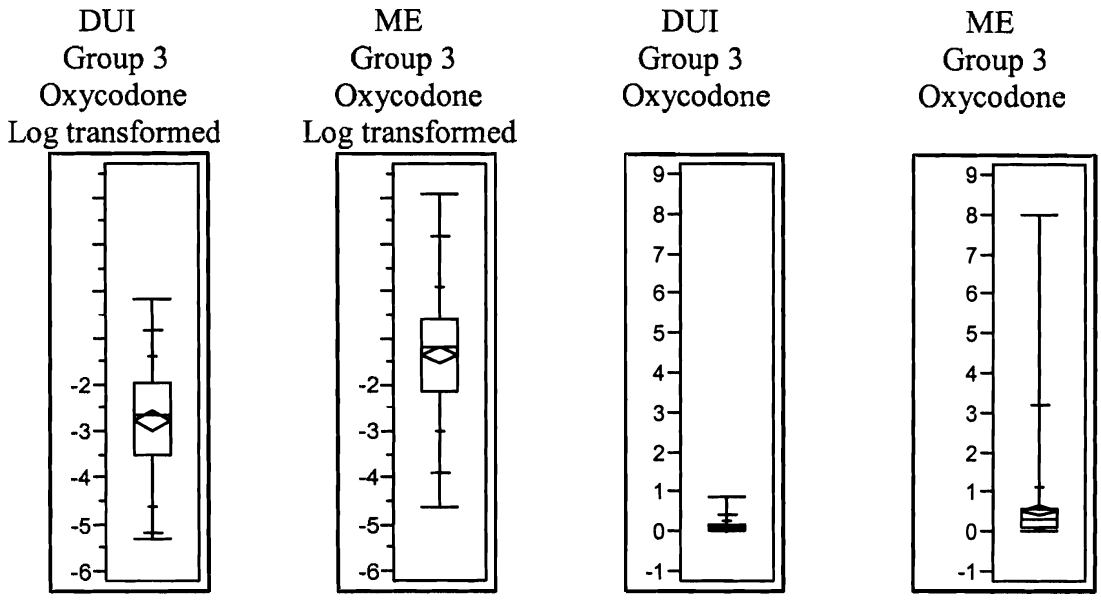


Figure 33: Box plots for Group 3 (oxycodone with any other combination of drug) for DUI and ME – log-transformed and untransformed data

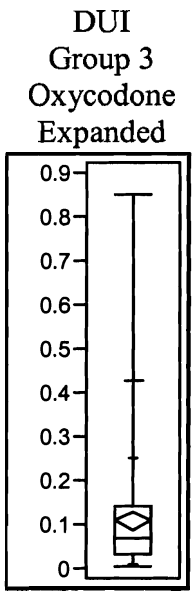


Figure 34: Expanded box plot for DUI Group 3 (oxycodone with any other combination of drug) – untransformed data

For DUI Group 3, the mean is -2.782 log (mg/L) units (± 0.102 standard errors) [0.06 mg/L] and the mean for ME Group 3 is -1.366 log (mg/L) units (± 0.087 standard errors)

[0.255 mg/L]. The range for DUI Group 3 is -5.298 log (mg/L) units [0.005 mg/L] to -0.163 log (mg/L) units [0.85 mg/L]. The range for ME Group 3 is from -4.605 log (mg/L) units [0.01 mg/L] to 2.079 log (mg/L) units [8.0 mg/L]. As the ranges show, there is a much larger spread of the oxycodone concentration for the ME dataset than for the DUI dataset. Note that the box plots of the log-transformed data do not always show the differences in the concentrations as well as the box plots of the untransformed concentration amounts. (All of these means and ranges are determined by the log transformed concentrations of oxycodone and by untransforming these values; they are not necessarily the same values as the untransformed means and standard errors as the ones seen in the table that displays the results of Tukey's test.)

After the box plots were created to visually inspect for obvious differences between the ME and DUI datasets, an ANOVA model was fit to the log-transformed concentration. The interaction term between `group` and `database` was not significant (p-value = 0.4625), so it was removed from the model. The new model contained only the two independent variables, and both of these were significant (p-values were both less than 0.0001); therefore, there is a statistically significant difference between the two datasets (ME and DUI), and there is also a statistically significant difference between at least two of the three drug pattern groups. The plots of the means for each of the independent fields are seen in Figures 35 and 36. Because there are only two source datasets, the statistically significant difference is between the ME and DUI datasets. The drug concentrations from the ME dataset are significantly higher than the drug concentrations from the DUI dataset (see Figure 36). With respect to the three levels in the drug pattern groups, the ANOVA

indicates that there is a significant difference between the levels, but it does not specify which levels are different from each other. A Tukey's test was done to determine which differences between the means of the drug pattern groups are statistically significantly different. Table 10 shows the results of the Tukey's test where levels with different letter labels are statistically significantly different, and thus, levels that are not connected by the same letter are significantly different from each other (regardless of the source dataset). Referring to Table 10, Groups 2 and 3 are not different from each other, but both are different from Group 1, and from the plot of the means in Figure 35, it seems that Group 1 has a higher concentration of oxycodone than Groups 2 and 3. Table 10 also displays the means and standard errors of the untransformed, or raw, concentrations of oxycodone. The means and standard errors reported in Table 10 are in mg/L units.

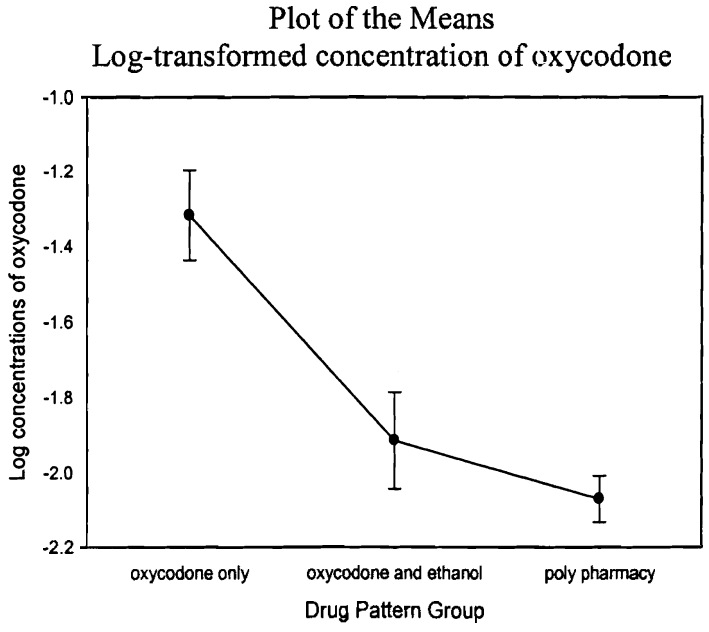


Figure 35: Plot of the means for "group" field for log-transformed oxycodone

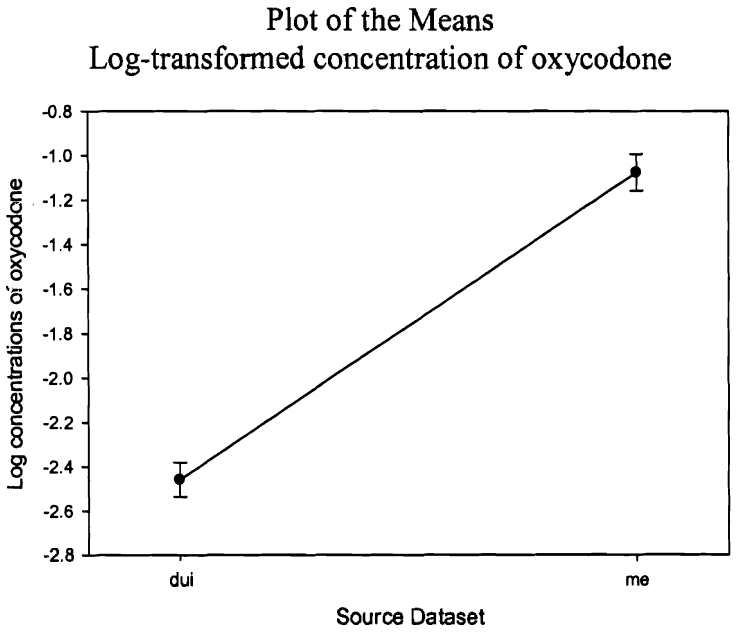


Figure 36: Plot of the means for "database" field for log-transformed oxycodone

Table 10: Levels of drug pattern groups for oxycodone from Tukey's test

Group	Letter	Mean(Std. Error)
Group 1	A	0.465(0.045)
Group 2	B	0.163(0.031)
Group 3	B	0.373(0.044)

The residual by predicted plot for oxycodone is seen in Appendix C and shows that there does not seem to be any systematic pattern between the predicted values and the residuals. The residuals also appear to follow the constant variance assumption since there seems to be roughly the same span of residuals in each of the categories. A normal quantile plot of the residuals was also constructed for the ANOVA model (see Appendix C), and it did not reveal any non-normality of the residuals.

There is not a specified clinical limit that indicates a lethal amount of oxycodone, so there was not an amount or range to compare the concentrations of oxycodone to for

either dataset. Since the datasets were determined to be different when it came to oxycodone concentration, the DUI dataset and the ME dataset did not have enough evidence to suggest development of tolerance to oxycodone.

3.3.4 Hydrocodone (Level II)

Hydrocodone is marketed as Vicodin ®, Lorcet ®, Lortab ®, and many other trade names. This drug is used as "an orally active narcotic analgesic and antitussive,"[30]. Hydrocodone can be "habit forming and can lead to physical and psychological addiction"[30]. Mixing alcohol with hydrocodone can also cause other health problems[30].

The hydrocodone dataset consisted of 542 individuals and 577 variable fields, including 200 cases from the ME dataset and 342 cases from the DUI dataset. The dataset for hydrocodone was created using the same criteria as the other datasets, and includes all individuals from the ME dataset who died from a hydrocodone-related drug death and all DUI individuals that had hydrocodone detected in their system. Table 11 shows the number of individuals in each of the drug pattern groups and each source dataset. As Table 11 shows, Group 2 for the ME dataset contained only six individuals, so Groups 2 and 3 were combined for analysis. The combined drug pattern group consisted of 166 individuals from the ME dataset and 315 from the DUI dataset for a total of 481 observations. Table 12 shows the demographic breakdown for Group 1 and the combined drug pattern group (Groups 2 & 3) for the ME and DUI datasets for the hydrocodone.

Table 11: Number of Observations for Categories in the Hydrocodone Dataset

Group	ME	DUI
1 (hydrocodone only)	34	27
2 (hydrocodone and ethanol)	6	121
3 (any other combination with hydrocodone)	160	194
Total	200	342

Table 12: Demographics on the Hydrocodone Dataset

	Source Dataset	ME		DUI	
	Drug Pattern Group	1	Combined (2&3)	1	Combined (2&3)
Race	White	--	--	19 (5.6%)	205 (59.9%)
	Black	--	--	1 (0.3%)	6 (1.8%)
	Other	--	--	0	0
	Missing	34 (17%)	166 (83%)	7 (2%)	104 (30.4%)
Gender	Male	18 (9%)	104 (52%)	16 (4.7%)	136 (39.8%)
	Female	16 (8%)	62 (31%)	4 (1.2%)	75 (21.9%)
	Missing	0	0	7 (2%)	104 (30.4%)
Age (in years)	Range	19-77	17-73	18-46	17-68

The normal quantile plots for the four categories shown in the tables, showed non-normal distributions for the drug pattern groups and datasets, so a log transformation was used for the hydrocodone concentration. The normal quantile plots for all categories are seen in the following figures with the untransformed data and the log-transformed data for the concentration of hydrocodone. As Figures 37, 38, 39, and 40 show, the log-transformed data followed an approximately normal distribution for all drug pattern groups and datasets, while the untransformed data did not follow an approximately normal distribution.

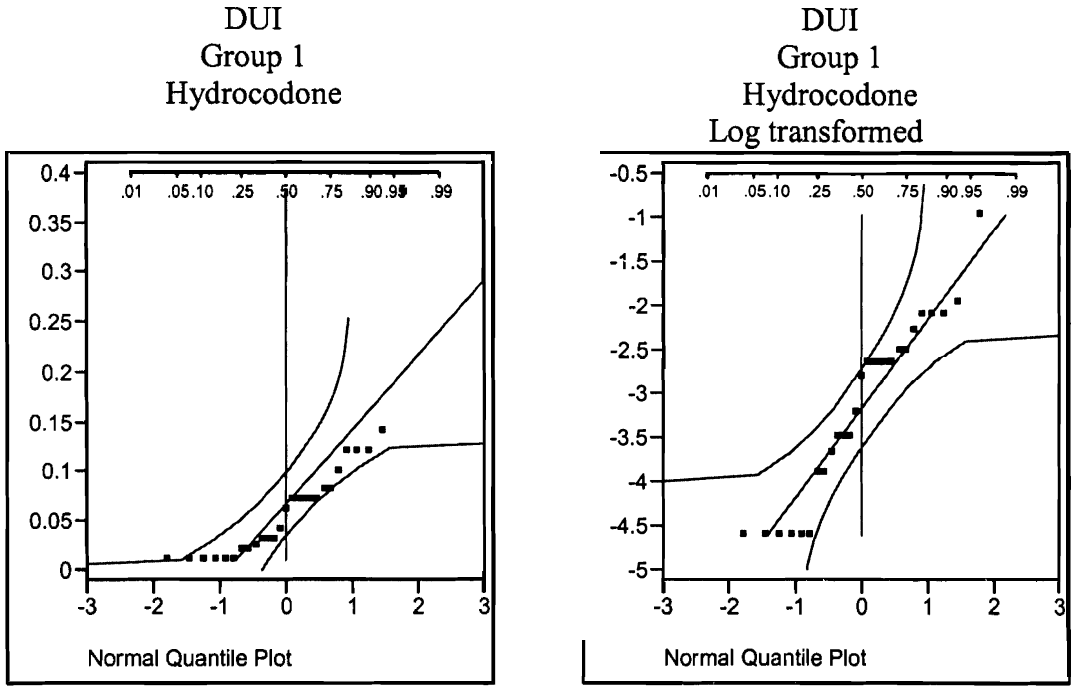


Figure 37: Normal quantile plots for DUI Group 1 (hydrocodone only) before and after log transformation

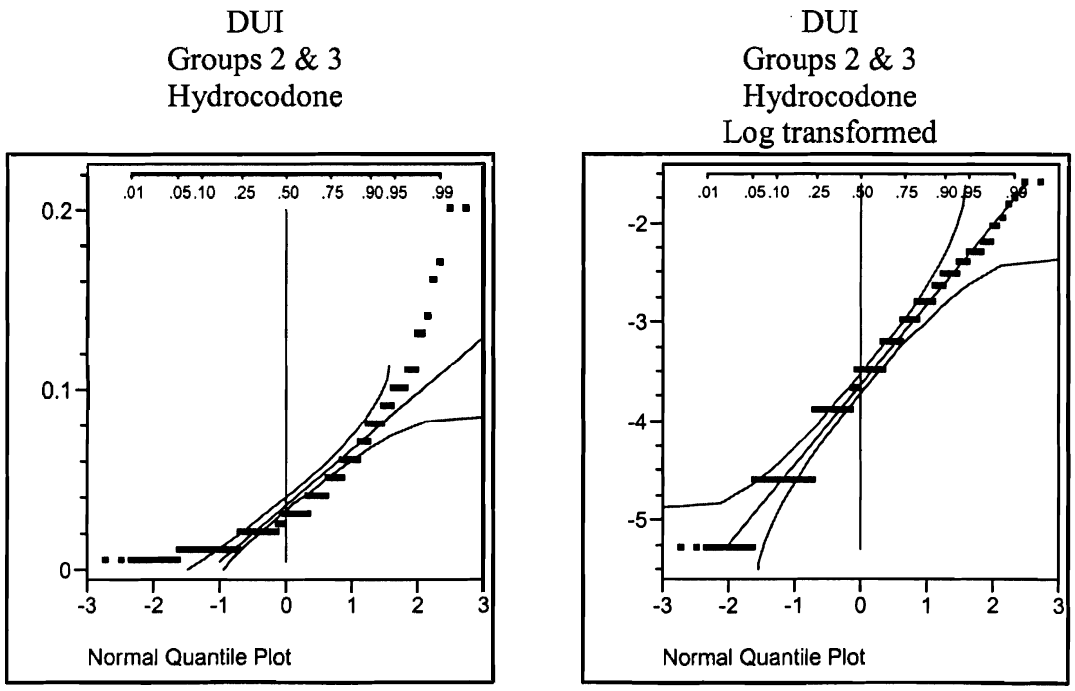


Figure 38: Normal quantile plots for DUI Groups 2 & 3 (hydrocodone with any drug - including ethanol) before and after log transformation

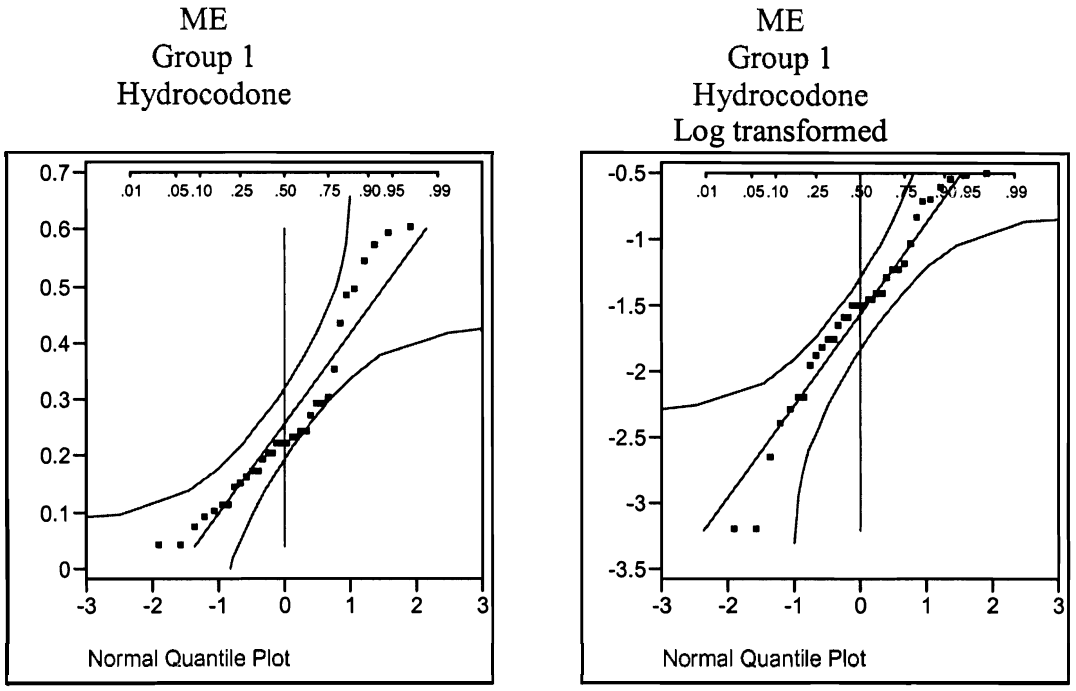


Figure 39: Normal quantile plots for ME Group 1 (hydrocodone only) before and after log transformation

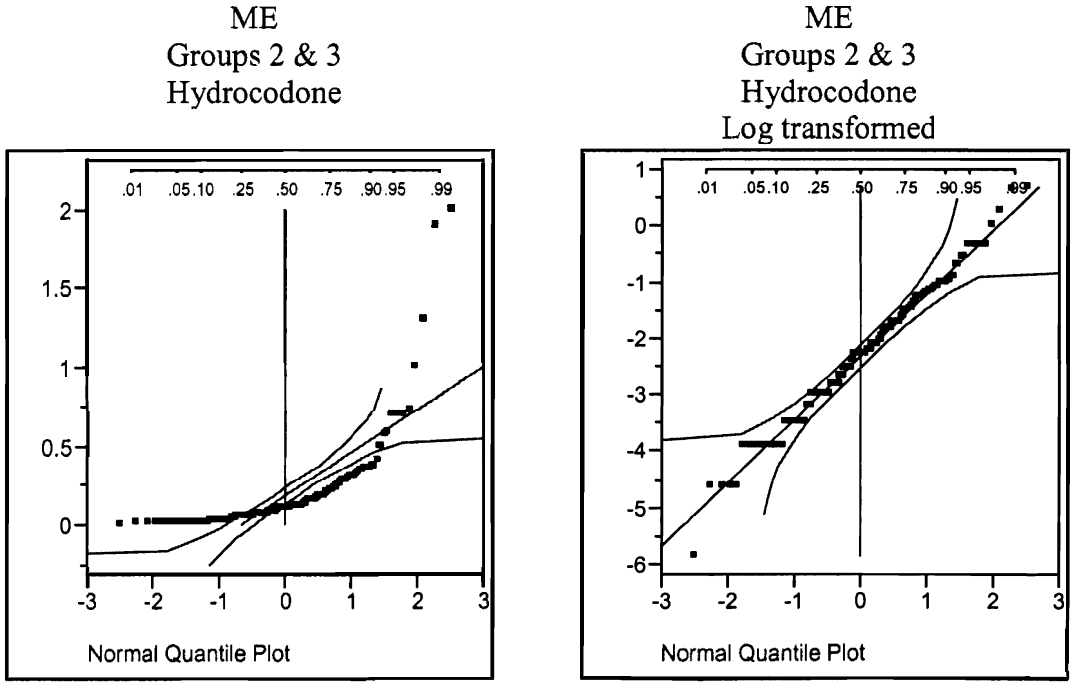


Figure 40: Normal quantile plots for ME Groups 2 & 3 (hydrocodone with any drug - including ethanol) before and after log transformation

The box plots comparing the log-transformed concentrations and the untransformed concentrations for the source datasets for each of the drug pattern groups are seen in the following figures. Figure 41 shows the box plots for Group 1, and Figure 42 shows the box plots for the combined drug pattern group (Groups 2 and 3). Figure 43 shows an expanded version of the untransformed DUI box plot for the combined drug pattern group, but there is not an expanded DUI box plot for Group 1 because Figure 41 shows the details with no need to expand the scale.

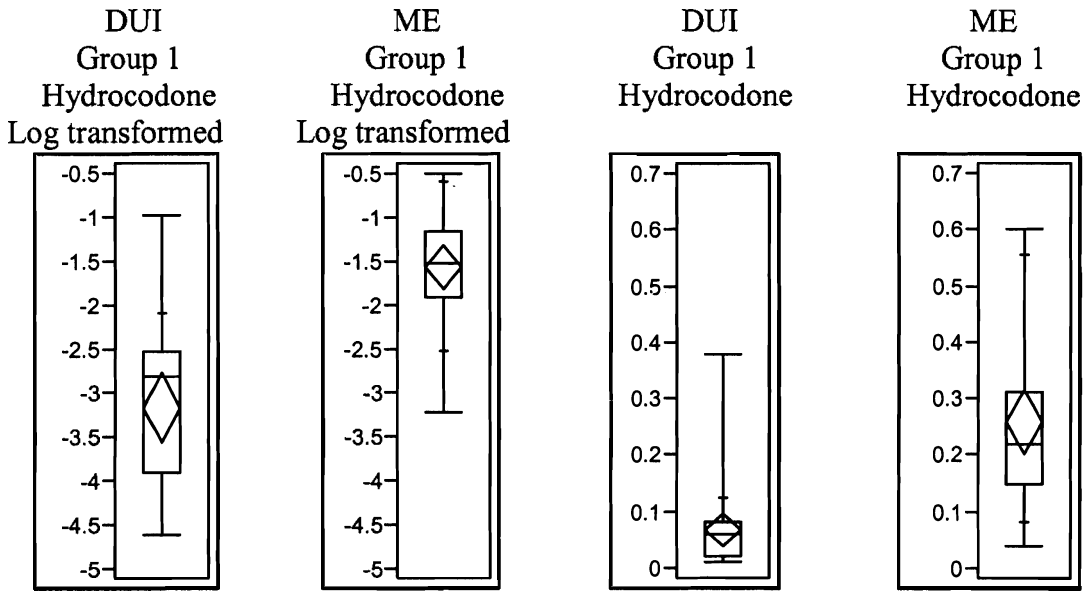


Figure 41: Box plots for Group 1 (hydrocodone only) for DUI and ME – log-transformed and untransformed data

The mean for Group 1 for the ME dataset is -1.565 log (mg/L) units (± 0.119 standard errors) [0.21 mg/L], and the mean for Group 1 for the DUI dataset is -3.172 log (mg/L) units (± 0.195 standard errors) [0.04 mg/L]. The range for the ME dataset for Group 1 spans from -3.219 log (mg/L) units [0.04 mg/L] to -0.511 log (mg/L) units [0.6 mg/L], while the range for the DUI dataset for this drug pattern group is from -4.605 log (mg/L)

units [0.01 mg/L] to -0.968 log (mg/L) units [0.38 mg/L]. Group 1 for the ME dataset spans a larger region than the DUI dataset, but the two datasets seem to follow a similar pattern for Group 1.

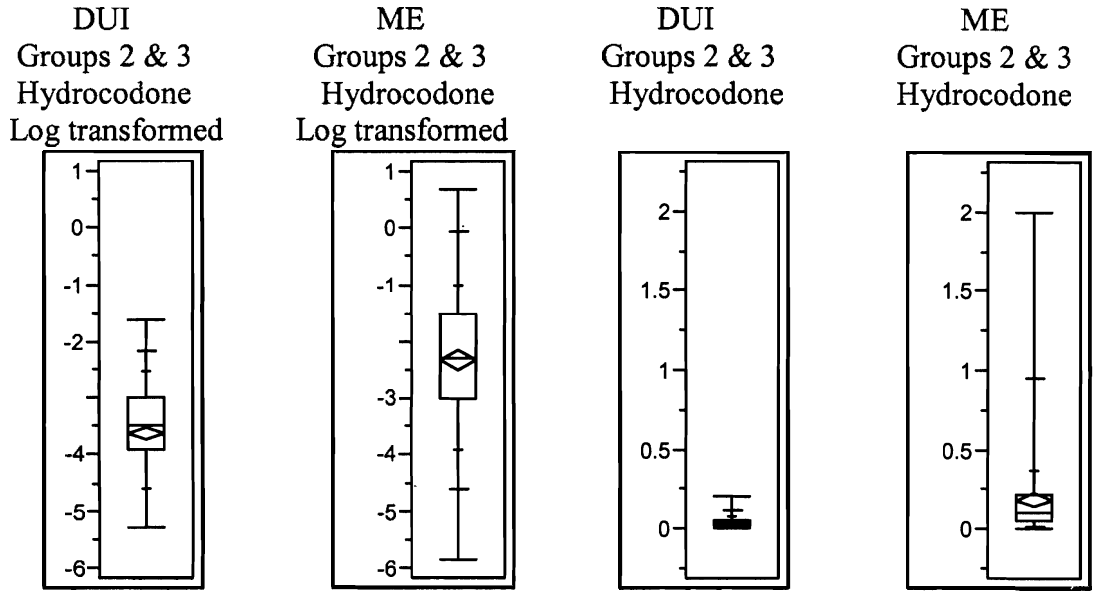


Figure 42: Box plots for Groups 2 & 3 (hydrocodone with any drug - including ethanol) for DUI and ME – log-transformed and untransformed data

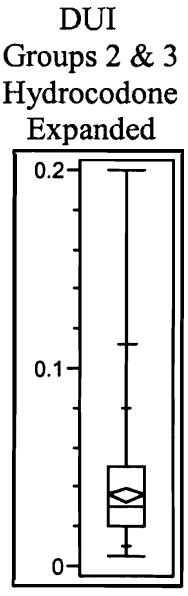


Figure 43: Expanded box plot for DUI Groups 2 & 3 (hydrocodone with any drug – including ethanol) – untransformed data

For the drug pattern group that combined Groups 2 & 3, the mean for the ME dataset was -2.317 log (mg/L) units (± 0.087 standard errors) [0.10 mg/L], and the mean for the DUI dataset was -3.63 log (mg/L) units (± 0.045 standard errors) [0.027 mg/L]. The range for the ME dataset was from -5.843 log (mg/L) units [0.0029 mg/L] to -0.693 log (mg/L) units [2.0 mg/L]; the range for the DUI dataset is from -5.298 log (mg/L) units [0.005 mg/L] to -1.609 log (mg/L) units [0.2 mg/L]. It is interesting to note that the range for DUI combined drug pattern group is completely contained in the range for the ME dataset. (All of these means and ranges are determined by using the log-transformed concentrations of hydrocodone and by untransforming these values; they will not always be the same as the raw means reported with the results of the ANOVA model.

An ANOVA model was fit to the hydrocodone dataset. The interaction between `group` and `database` was not statistically significant (p-value = 0.2475), so it was removed from the model. With only the independent fields left in the model, both were determined to be statistically significant (p-values each less than 0.0001). Since both fields were statistically significant, then the ME dataset and the DUI datasets were determined to be statistically significantly different from each other within the hydrocodone dataset. Group 1 was also determined to be statistically different from the combined drug pattern group (Groups 2 & 3). The plots of the means for each of the variable fields are shown in Figures 44 and 45. The concentration of hydrocodone is statistically significantly higher for the ME dataset (mean = 0.196 ± 0.018 standard errors) than for the DUI dataset (mean = 0.039 ± 0.002 standard errors), and the concentration is statistically significantly higher

in Group 1 (mean = 0.173 ± 0.02 standard errors) than it is in the combined drug pattern group (Groups 2 & 3) (mean = 0.087 ± 0.008 standard errors).

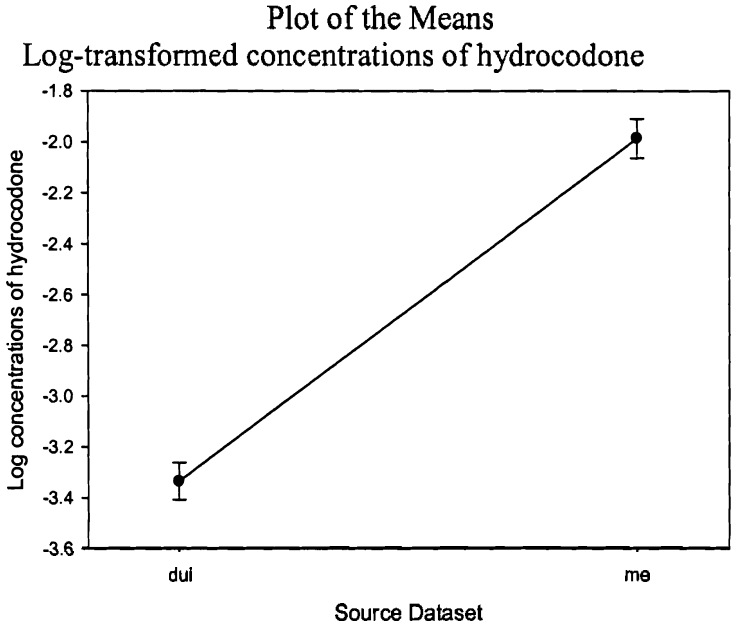


Figure 44: Plot of the means for "database" field for log-transformed hydrocodone

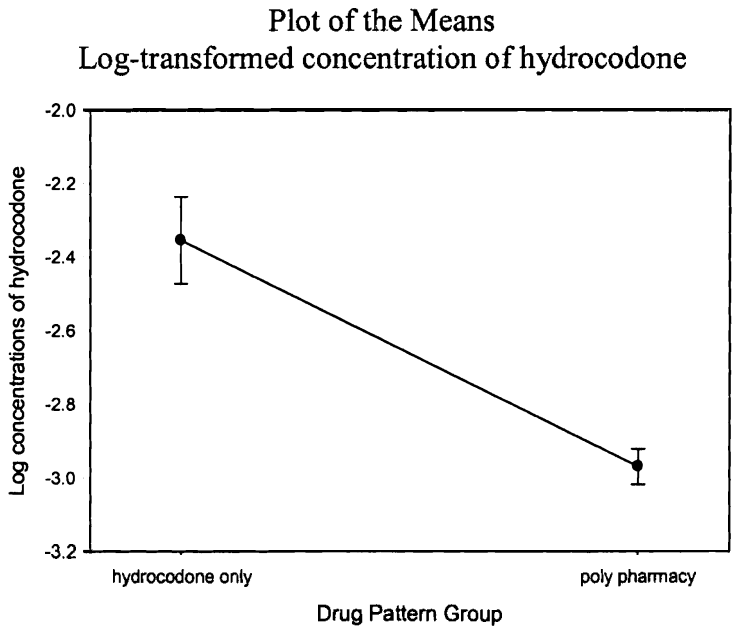


Figure 45: Plot of the means for "group" field for log-transformed hydrocodone

Once the model was fitted to the data, a residual by predicted plot was constructed (Appendix C). This plot showed that the constant variance assumption was met, and since there did not seem to be overall systematic pattern, the model was adequate for the data. The residuals for this model were also plotted in a normal quantile plot to check for normality (Appendix C), and the residuals appeared to be approximately normal.

Because the ME and DUI datasets are statistically different with respect to hydrocodone, there is not sufficient evidence to suggest tolerances to hydrocodone. There is not a defined clinical concentration of hydrocodone that is considered fatal.

3.3.5 Methadone (Level III)

"German scientists synthesized methadone during World War II because of a shortage of morphine,"[31]. The trade name for methadone is Dolophine ®, and it is "used as an analgesic and in the treatment of narcotic addiction"[32]. "Ironically, methadone used to control narcotic addiction is frequently encountered on the illicit market and has been associated with a number of overdose deaths"[31].

The methadone dataset contained 522 observations and 577 variable fields. Consisting of 407 individuals from the ME dataset and 115 from the DUI dataset, the methadone dataset followed the same inclusion and exclusion rules that the previous datasets have mentioned. Three drug pattern groups were created: methadone only, methadone and alcohol, and methadone with any other combination of drugs. The following two tables show the frequency of observations for each category and the demographics of each of the categories. The number of observations in each drug pattern group and each source dataset is shown in Table 13, and Table 14 shows the itemization of

the three drug pattern groups and the two source datasets with respect to the demographic variable fields `race`, `gender`, and `age`.

Table 13: Number of Observations for Categories in the Methadone Dataset

Group	ME	DUI
1 (methadone only)	142	14
2 (methadone and ethanol)	20	27
3 (any other combination with methadone)	245	74
Total	407	115

Table 14: Demographics on the Methadone Dataset

	Source Dataset	ME			DUI		
	Drug Pattern Group	1	2	3	1	2	3
Race	White	--	--	--	7 (6.1%)	13 (11.3%)	45 (39.1%)
	Black	--	--	--	0	1 (0.9%)	2 (1.7%)
	Other	--	--	--	0	0	1 (0.9%)
	Missing	142 (34.9%)	20 (4.9%)	245 (60.2%)	7 (6.1%)	13 (11.3%)	26 (22.6%)
Gender	Male	97 (23.8%)	17 (4.2%)	153 (37.6%)	4 (3.5%)	9 (7.8%)	31 (27%)
	Female	45 (11.1%)	3 (0.7%)	91 (22.4%)	3 (2.6%)	5 (4.3%)	18 (15.7%)
	Missing	0	0	1 (0.2%)	7 (6.1%)	13 (11.3%)	25 (21.7%)
Age (in years)	Range	18-59	18-45	17-71	18-44	22-48	18-64

Because all of the drug pattern groups contained at least 12 cases, none of them were combined, and the analysis was done using all three drug pattern groups. The normal quantile plots of the drug pattern groups and the source datasets showed non-normal distributions, so a log transformation was done on the concentration of methadone for the

analysis. The normal quantile plots for each of the datasets and drug pattern groups are displayed in the following figures. Figures 46 through 48 show the normal quantile plots for all of the DUI drug pattern groups with both the untransformed and the log-transformed data. Figures 49 through 51 show the normal quantile plots for the ME drug pattern groups. The normal quantile plots of the log-transformed data show an approximately normal distribution, and this transformed data was used for the ANOVA model.

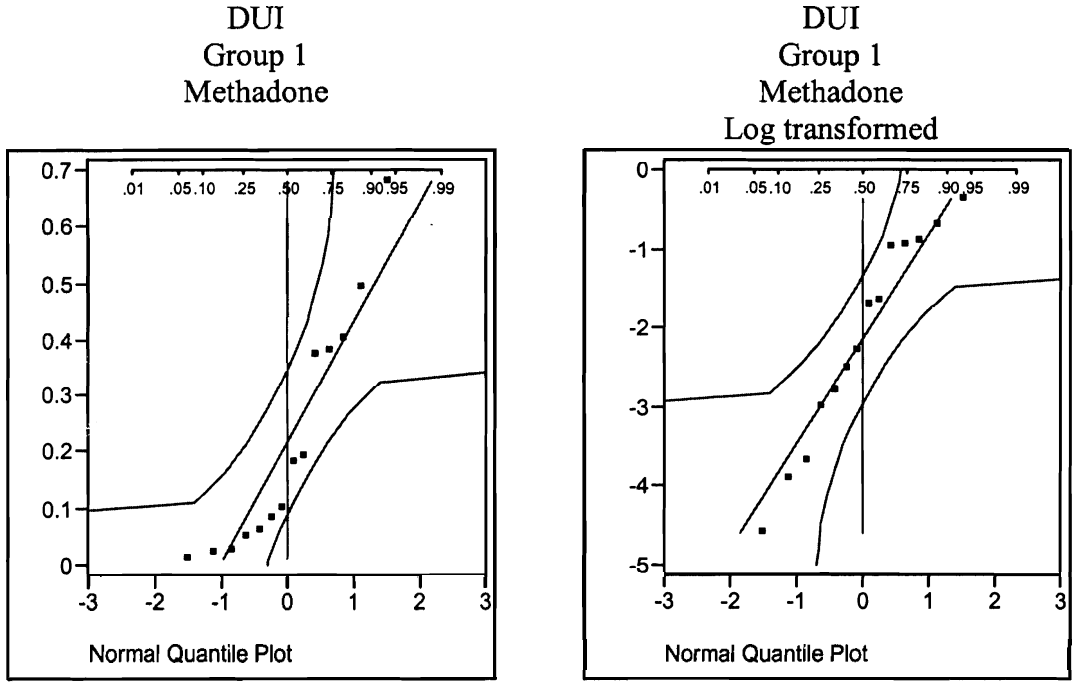


Figure 46: Normal quantile plots for DUI Group 1 (methadone only) before and after log transformation

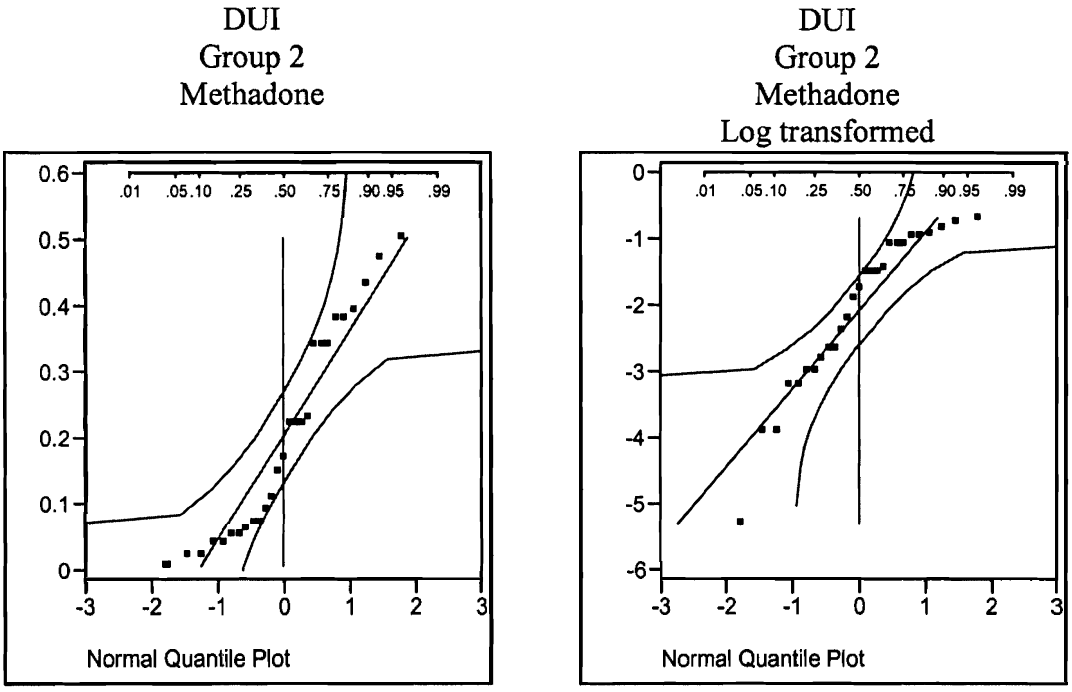


Figure 47: Normal quantile plots for DUI Group 2 (methadone and ethanol) before and after log transformation

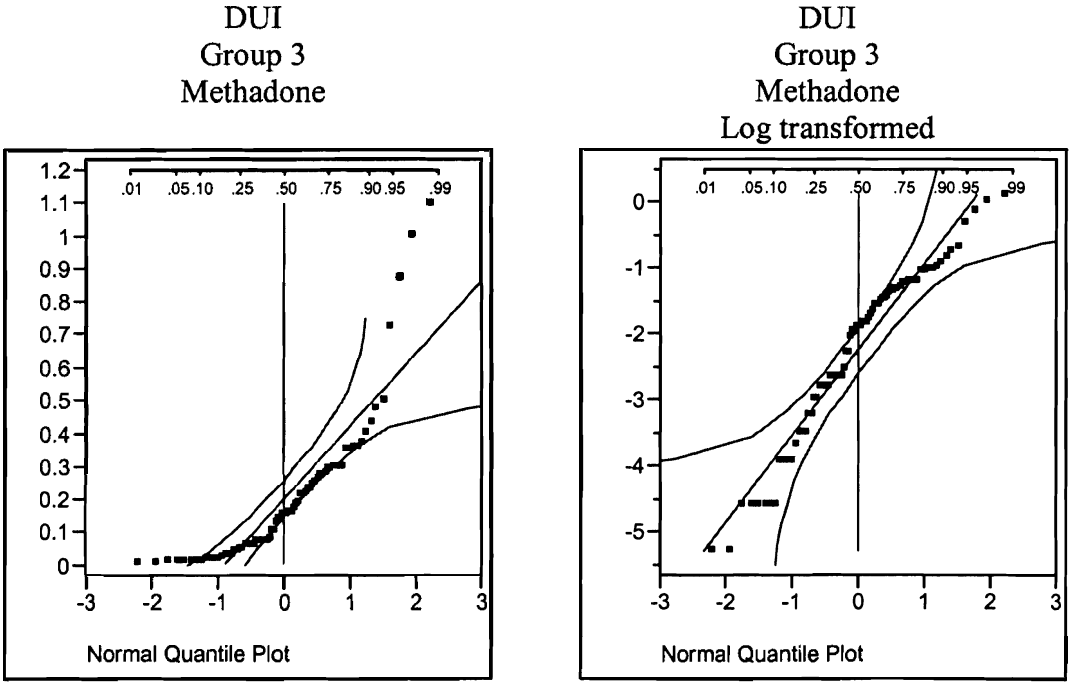


Figure 48: Normal quantile plots for DUI Group 3 (methadone with any other combination of drug) before and after log transformation

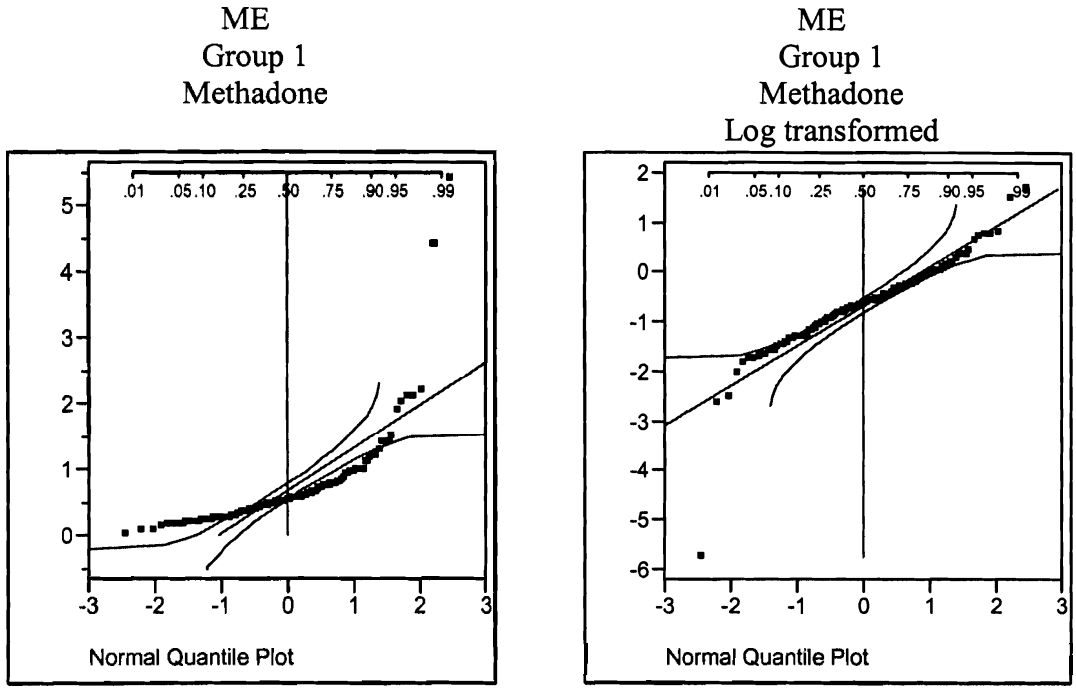


Figure 49: Normal quantile plots for ME Group 1 (methadone only) before and after log transformation

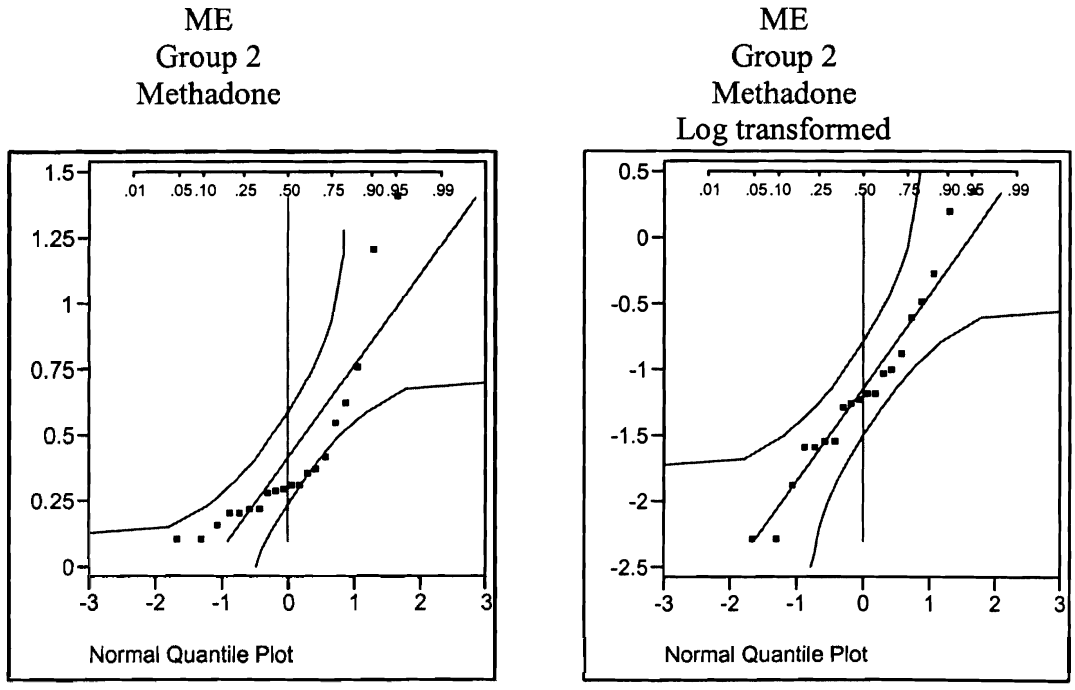


Figure 50: Normal quantile plots for ME Group 2 (methadone and ethanol) before and after log transformation

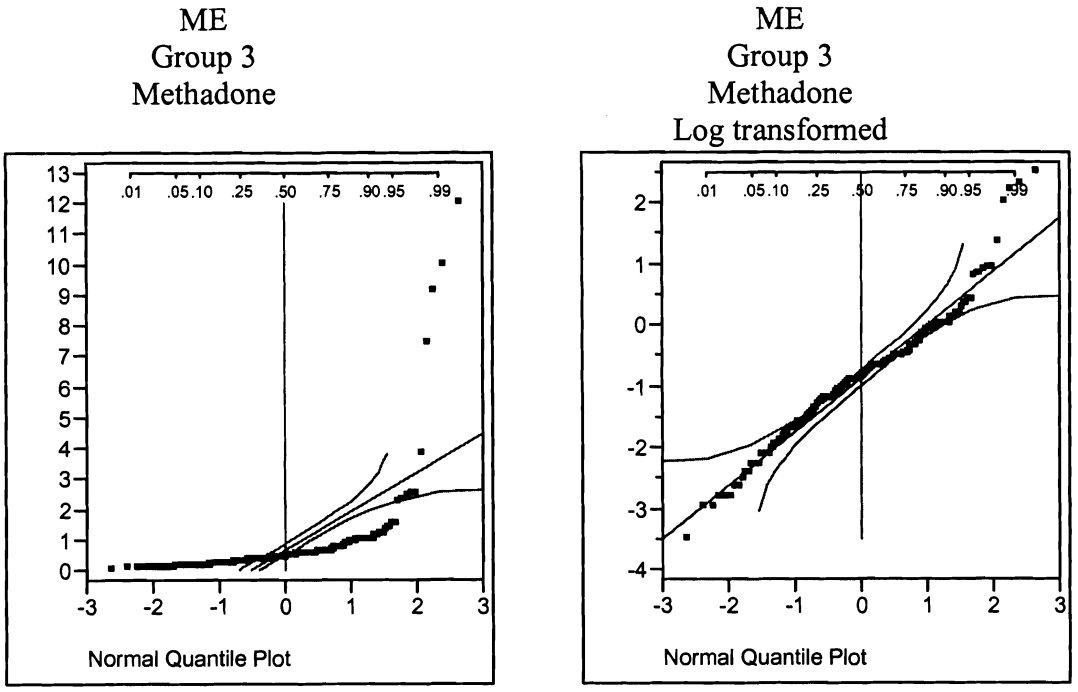


Figure 51: Normal quantile plots for ME Group 3 (methadone with any other combination of drug) before and after log transformation

The next sets of figures show the box plots (log-transformed and untransformed) comparing the ME and DUI datasets for each of the three drug pattern groups. Figure 52 displays two pairs of box plots for Group 1 (methadone only) for the log-transformed data (left side) and the untransformed data (right side), and Figure 53 shows an expanded box plot of the untransformed DUI Group 1 data. The expanded box plot is displayed to show the details of the untransformed DUI data. Figures 54 and 55 show the box plots for Group 2 (methadone and ethanol).

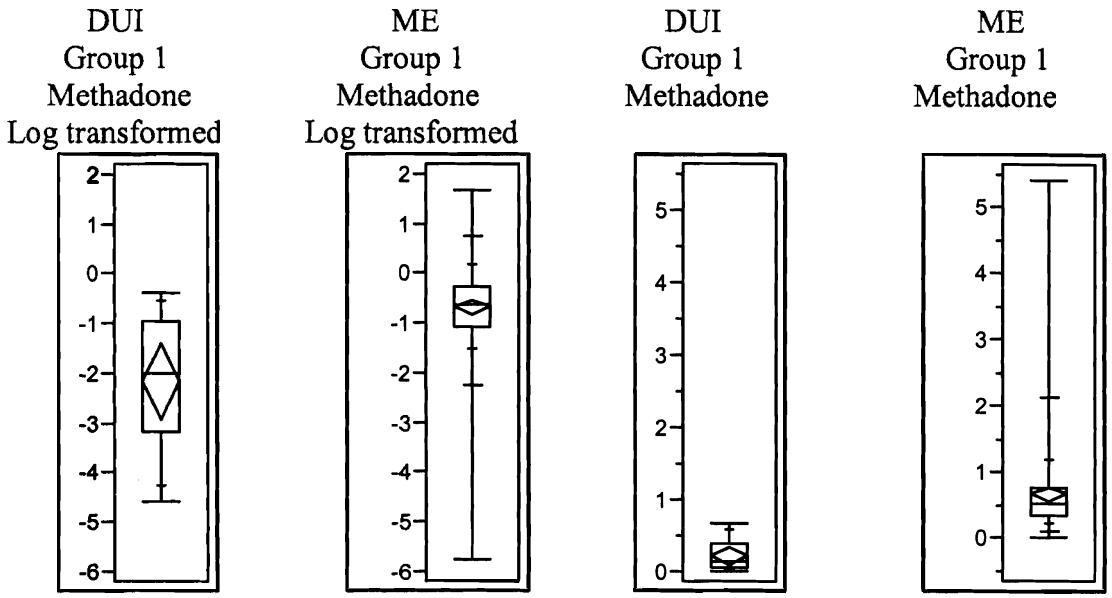


Figure 52: Box plots for Group 1 (methadone only) for DUI and ME – log-transformed and untransformed data

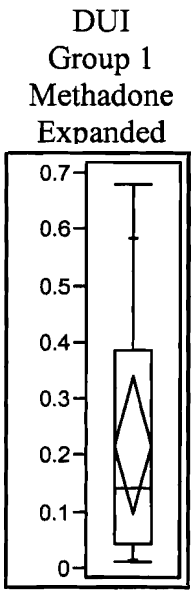


Figure 53: Expanded box plot for DUI Group 1 (methadone only) – untransformed data

The means for Group 1 of the methadone dataset are -0.678 log (mg/L) units (± 0.068 standard errors) [0.51 mg/L] for the ME dataset and -2.157 log (mg/L) units (± 0.353 standard errors) [0.116 mg/L] for the DUI dataset. The range for Group 1 for the ME

dataset is from -5.776 log (mg/L) units [0.003 mg/L] to 1.686 log (mg/L) units [5.4 mg/L], and the range for Group 1 for the DUI dataset is from -4.605 log (mg/L) units [0.01 mg/L] to -0.386 log (mg/L) units [0.68 mg/L]. As these ranges show, the entire range of the DUI dataset fits within the boundaries defined by the ME range; however, the lower boundary for the DUI dataset is higher than the lower boundary for the ME dataset. It might be of clinical interest to inspect the ME case that has a methadone-related death with a concentration of methadone equal to 0.0003 mg/L because it seems that the concentration is not a "lethal" dose.

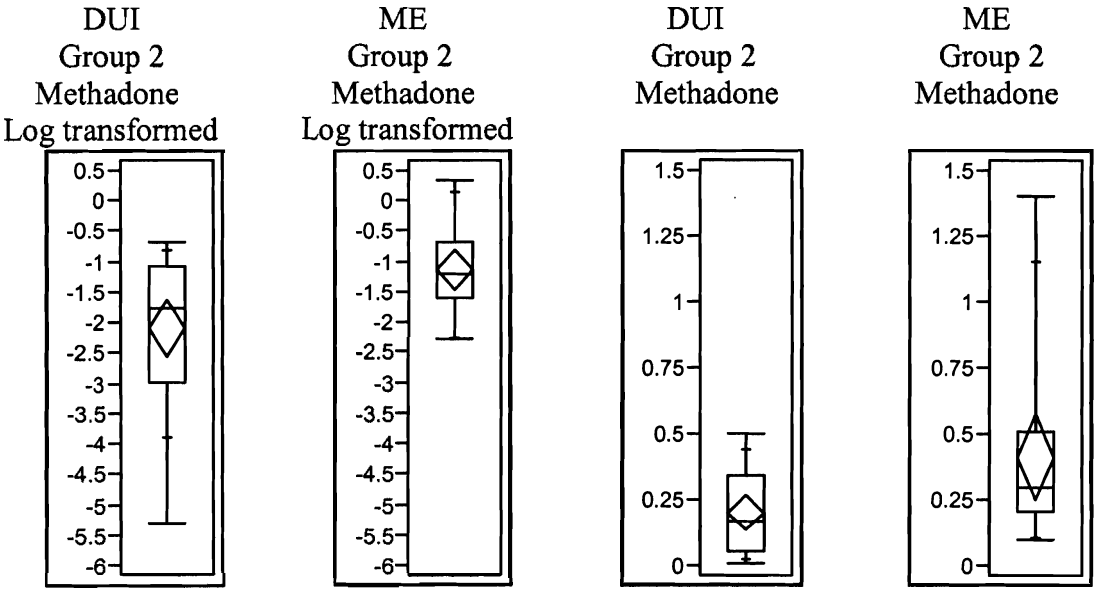


Figure 54: Box plots for Group 2 (methadone and ethanol) for DUI and ME – log-transformed and untransformed data

For Group 2, the mean for the ME dataset is -1.146 log (mg/L) units (± 0.159 standard errors) [0.32 mg/L], while the mean for the DUI dataset is -2.088 log (mg/L) units (± 0.227 standard errors) [0.124 mg/L]. The span of the concentrations for methadone in the ME dataset ranges from -2.303 log (mg/L) units [0.1 mg/L] to 0.336 log (mg/L) units

[1.4 mg/L]. For the DUI dataset for Group 2, the range is from -5.298 log (mg/L) units [0.005 mg/L] to -0.693 log (mg/L) units [0.5 mg/L]. In this drug pattern group, unlike Group 1, the range for the DUI dataset starts at a lower concentration than that of the ME dataset.

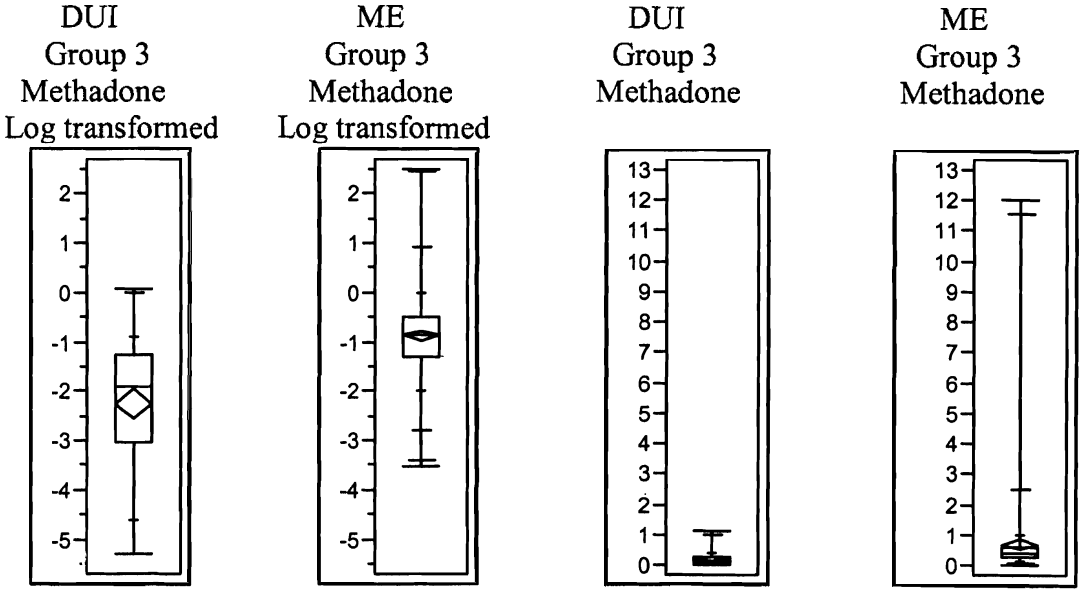


Figure 55: Box plots for Group 3 (methadone with any other combination of drug) for DUI and ME – log-transformed and untransformed data

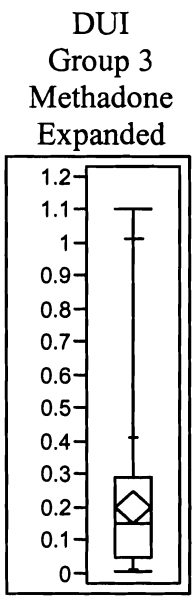


Figure 56: Expanded box plot for DUI Group 3 (methadone with any other combination of drug) – untransformed data

In Group 3, the means for the two datasets are -0.87 log (mg/L) units (± 0.056 standard errors) [0.419 mg/L] for the ME dataset and -2.255 log (mg/L) units (± 0.152 standard errors) [0.105 mg/L] for the DUI dataset. The concentrations of methadone for Group 3 for the ME dataset ranges from -3.507 log (mg/L) units [0.03 mg/L] to 2.485 log (mg/L) units [12 mg/L], and the range for the DUI dataset is from -5.298 log (mg/L) units [0.005 mg/L] to 0.095 log (mg/L) units [1.1 mg/L]. The span of the range for the DUI dataset for this particular drug pattern group is smaller and begins at a lower dose than the range for the ME dataset. (All of these means and ranges were determined from the log-transformed concentrations of methadone and by untransforming these means.)

An ANOVA model was fit to the data to determine whether or not there was a statistically significant difference in concentrations of methadone between the two source datasets and between the three drug pattern groups. Once this model was fit, the

interaction term was determined to be not statistically significant (p-value = 0.3064), and thus, it was removed from the model. After removing the interaction term from the ANOVA model, the `group` variable field was also noted as not significant (p-value = 0.1117), and it was removed from the model as well. The only significant variable field left in the model was the source dataset, implying that the ME and DUI datasets are different from each other with regard to concentration of methadone. Figure 57 shows the plot of the means for the `database` field. From this plot, it can be seen that the ME dataset has a statistically significant higher mean for concentration of methadone (mean = 0.664 mg/L ± 0.052 standard errors) than the DUI dataset (mean = 0.202 mg/L ± 0.019 standard errors). (These means and standard errors are the untransformed concentrations for methadone and are given in mg/L units.) Because the ME and the DUI datasets are statistically significantly different from each other, there is not sufficient evidence to suggest tolerances to methadone.

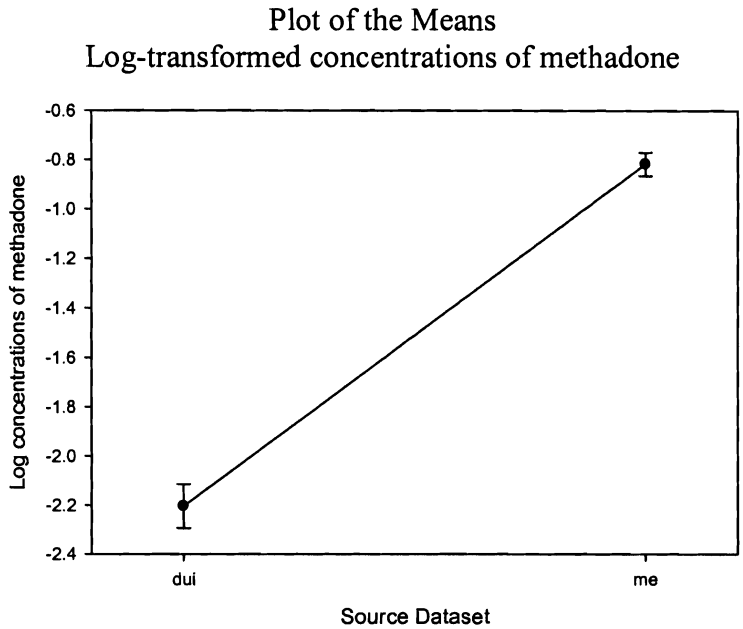


Figure 57: Plot of the means for "database" field for log-transformed methadone

A residual by predicted plot was constructed (Appendix C) from the model with only the field for the source datasets. This plot verified the constant variance assumption and did not show any systematic pattern of the residuals with the predicted values that would imply any inadequacy with the model did not fit. A normal quantile plot of the residuals from this model was also constructed and can be seen in Appendix C. This plot shows that the assumption of normality for the residuals is met because the residuals follow the diagonal.

The clinically determined fatal blood concentration for methadone ranges from 0.4 to 1.8 mg/L[26]. Using this as the guideline to determine whether there are any interesting DUI cases with high levels of methadone, 14 individuals (in the DUI dataset) were found to have a concentration of methadone that fell within the "lethal" concentration; this is visible on the box plots in Figures 53, 54, and 56 where data points for the DUI dataset fall

within this clinically defined range. These cases offer some evidence, regardless of the results of the analysis, that it is possible to develop a tolerance to methadone; these 14 individuals have "lethal" doses of methadone in their systems and are still attempting to operate a motor vehicle.

3.3.6 Morphine (Level II)

"Morphine is the naturally occurring substance in the opium poppy, Papaver Somniferous. It is a potent narcotic analgesic, and its primary clinical use is in the management of moderately severe and severe pain. After heroin, morphine has the greatest dependence liability of the narcotic analgesics in common use"[33]. It is thought that regular use of morphine can result in developing tolerances to the drug[33].

The morphine dataset consisted of the individuals testing positive for morphine and contained 742 cases and 577 variable fields. The morphine dataset was divided into the same three types of drug pattern groups as the other drugs: morphine only, morphine and ethanol, and morphine with any other combination of drugs. These drug pattern groups were used to compare the concentration of morphine between the ME and DUI datasets and among the three drug pattern groups. The number of observations in each of the three drug pattern groups is seen in Table 15. The ME dataset contributed 611 cases to the morphine dataset, while the DUI dataset contributed 131 cases. From the numbers in Table 15, it is obvious that none of the categories had less than 12 cases in them, so no drug pattern groups were combined or removed for analysis of the morphine dataset. The demographics of `age`, `gender`, and `race`, are broken down by drug pattern group and dataset in Table 16.

Table 15: Number of Observations for Categories in the Morphine Dataset

Group	ME	DUI
1 (morphine only)	143	25
2 (morphine and ethanol)	50	35
3 (any other combination with morphine)	418	71
Total	611	131

Table 16: Demographics on the Morphine Dataset

	Source Dataset	ME			DUI		
	Drug Pattern Group	1	2	3	1	2	3
Race	White	--	--	--	15 (11.5%)	16 (12.2%)	37 (28.2%)
	Black	--	--	--	2 (1.5%)	3 (2.3%)	9 (6.9%)
	Other	--	--	--	0	0	0
	Missing	143 (23.4%)	50 (8.2%)	418 (68.4%)	8 (6.1%)	16 (12.2%)	25 (19.1%)
Gender	Male	119 (19.5%)	47 (7.7%)	300 (49.1%)	10 (7.6%)	15 (11.5%)	38 (29%)
	Female	24 (3.9%)	3 (0.5%)	118 (19.3%)	7 (5.3%)	4 (3.1%)	8 (6.1%)
	Missing	0	0	0	8 (6.1%)	16 (12.2%)	25 (19.1%)
Age (in years)	Range	20-76	19-59	15-69	18-54	18-49	16-59

The normal quantile plots for the concentrations of morphine by drug pattern group and by source dataset showed non-normal distributions, so a log transformation was used for the concentration of morphine. The normal quantile plots seen in the following figures show both the untransformed data and the log-transformed data. Figures 58, 59, and 60 show the untransformed and log-transformed box plots for the DUI dataset for Groups 1, 2, and 3. The box plots for the ME dataset are seen in Figures 61, 62, and 63.

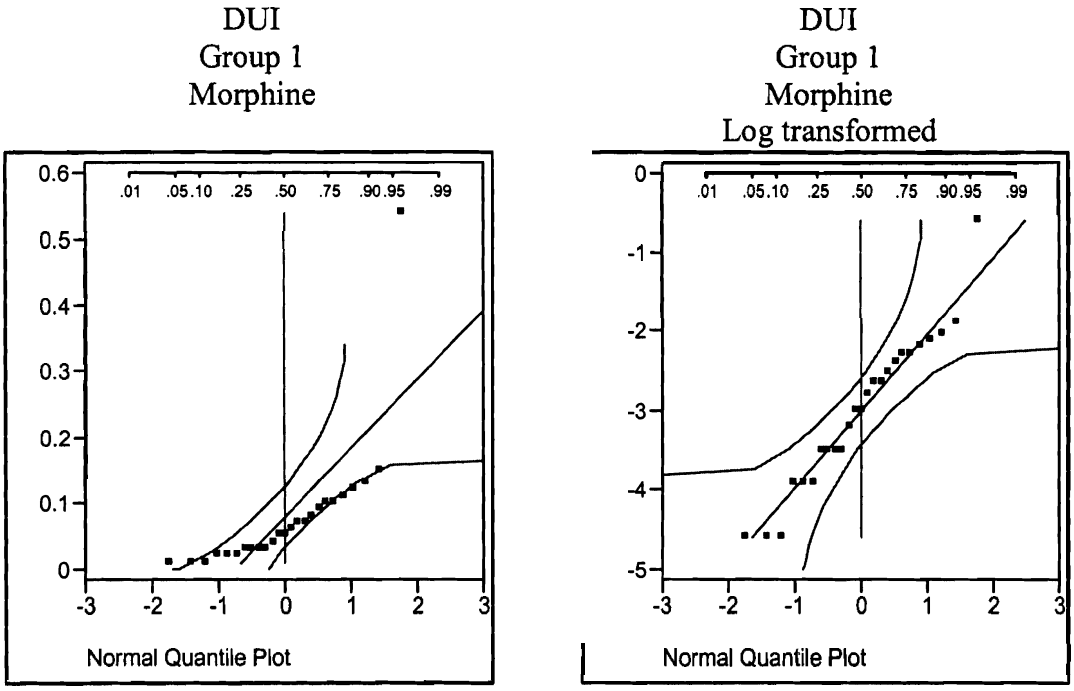


Figure 58: Normal quantile plots for DUI Group 1 (morphine only) before and after log transformation

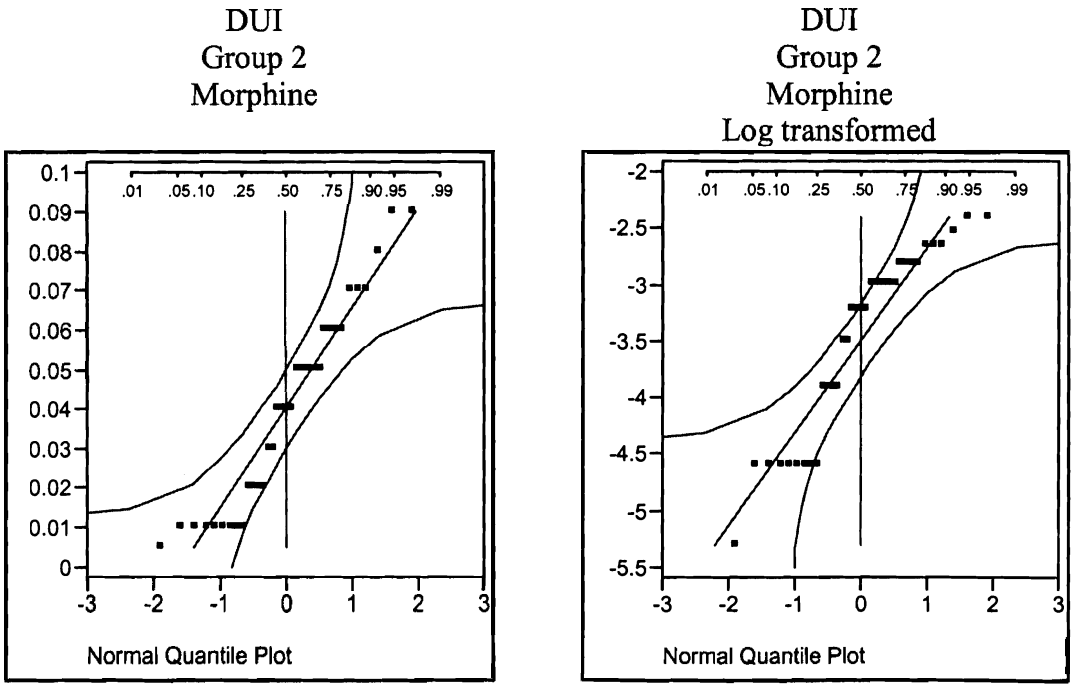


Figure 59: Normal quantile plots for DUI Group 2 (morphine and ethanol) before and after log transformation

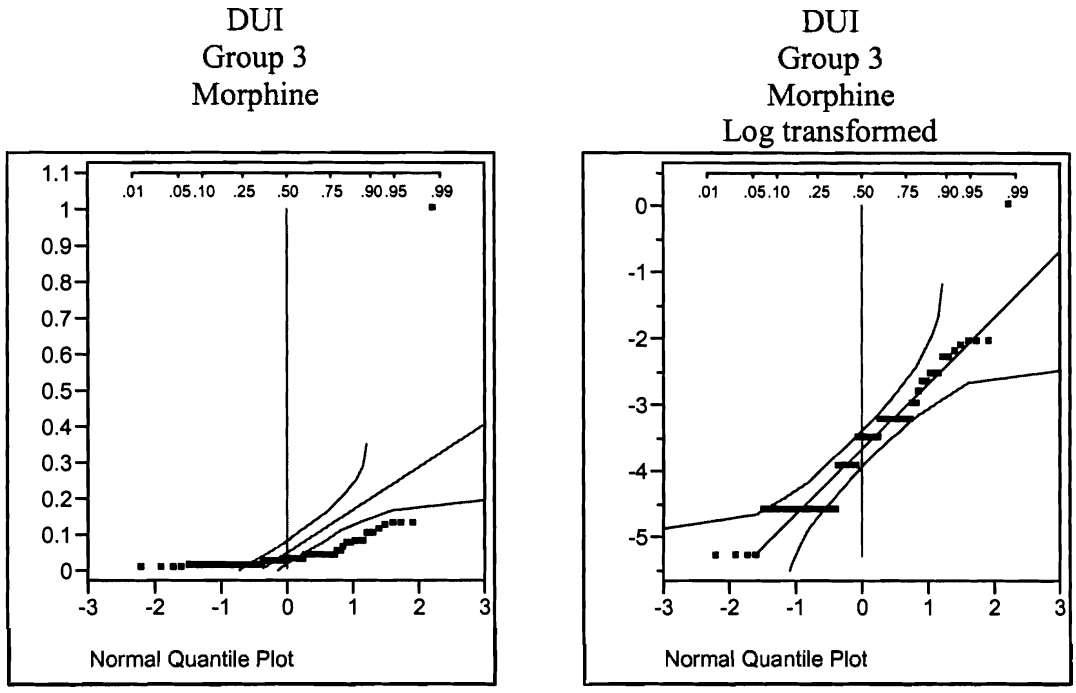


Figure 60: Normal quantile plots for DUI Group 3 (morphine with any other combination of drug) before and after log transformation

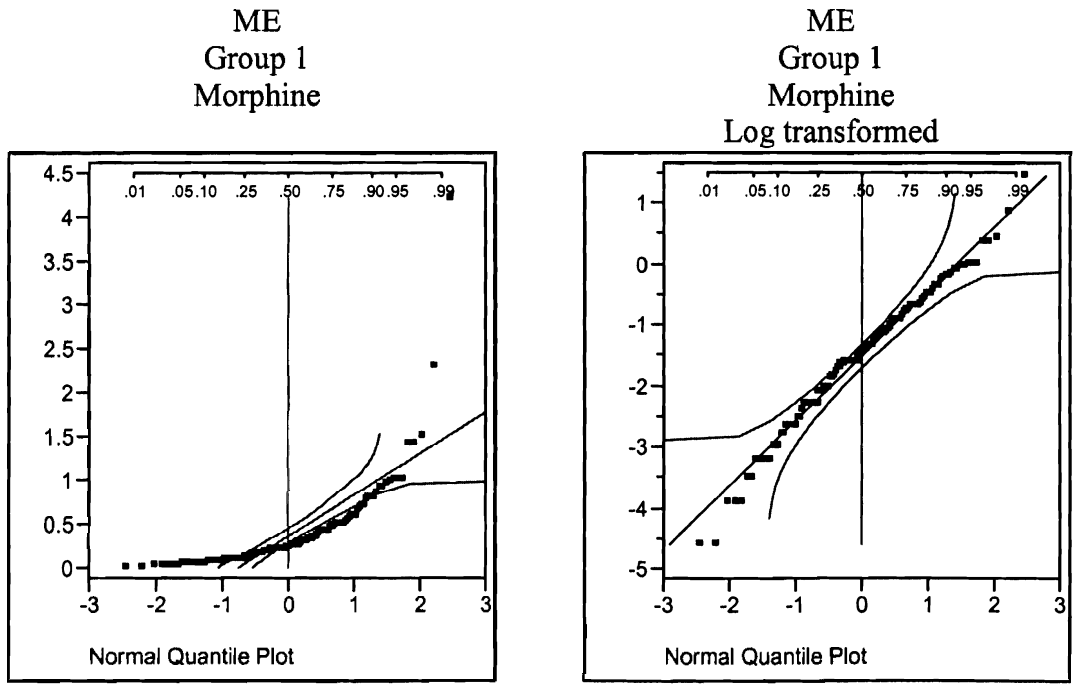


Figure 61: Normal quantile plots for ME Group 1 (morphine only) before and after log transformation

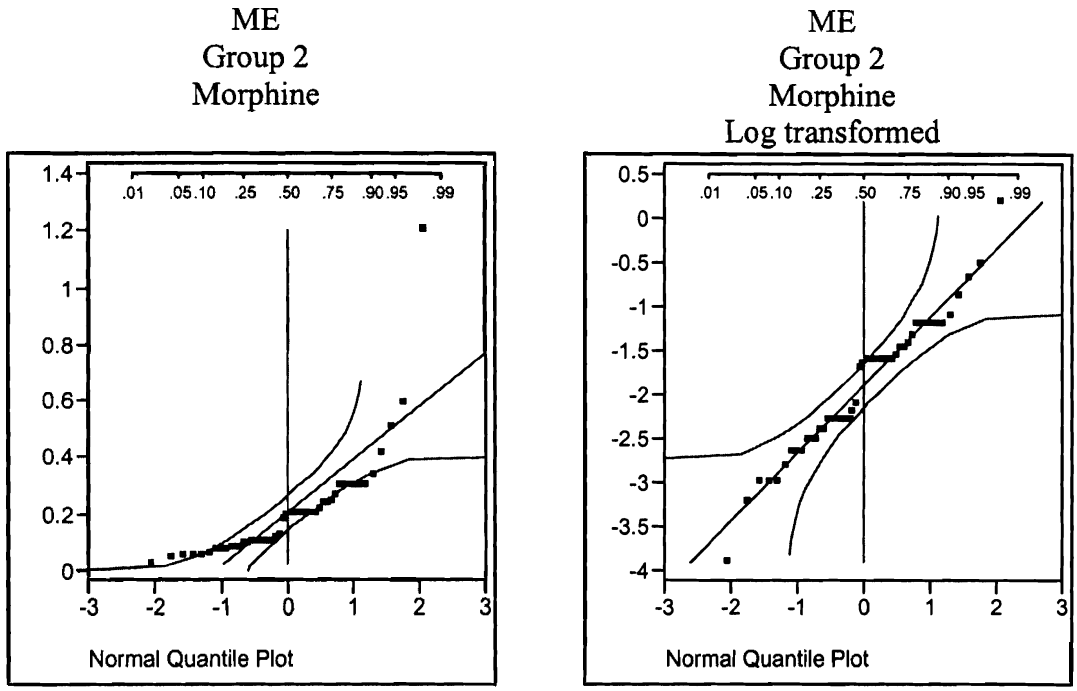


Figure 62: Normal quantile plots for ME Group 2 (morphine and ethanol) before and after log transformation

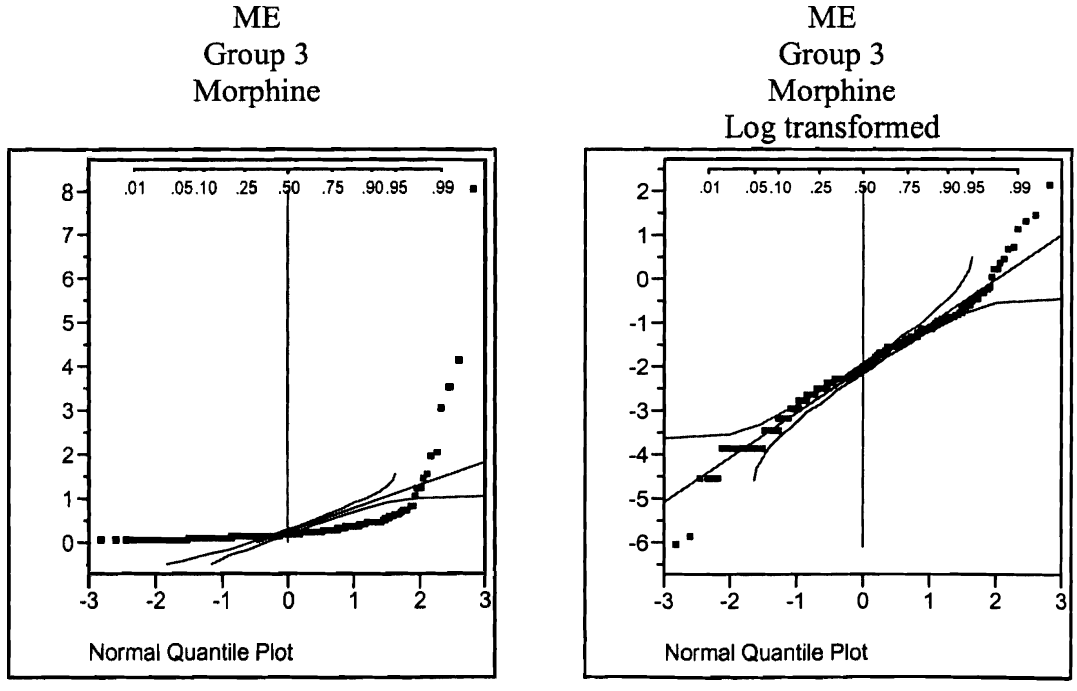


Figure 63: Normal quantile plots for ME Group 3 (morphine with any other combination of drug) before and after log transformation

The next section of figures shows the pairs of box plots for each of three drug pattern groups and compares the ME and DUI datasets. The set of box plots for Group 1 (morphine only) is seen in Figure 64. Figures 66 and 68 show the set of box plots for Group 2 and Group 3, for the DUI and ME datasets before and after the log transformation. For each of the three drug pattern groups, an expanded version of the untransformed DUI data is displayed to better view the details of the box plot.

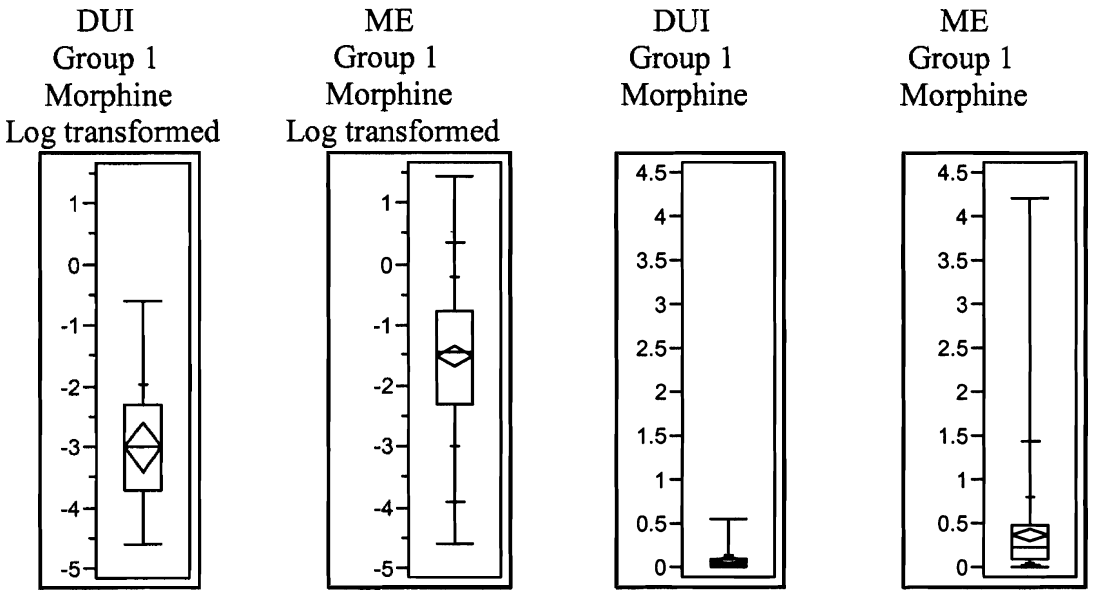


Figure 64: Box plots for Group 1 (morphine only) for DUI and ME – log-transformed and untransformed data

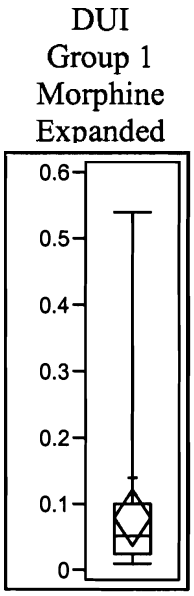


Figure 65: Expanded box plot for DUI Group 1 (morphine only) – untransformed data

The mean for Group 1 for the ME dataset is -1.52 log (mg/L) units (± 0.089 standard errors) [0.22 mg/L], and the mean for Group 1 for the DUI dataset is -3.01 log (mg/L) units (± 0.193 standard errors) [0.05 mg/L]. The range for the ME dataset for this group is from -4.605 log (mg/L) units [0.01 mg/L] to 1.435 log (mg/L) units [4.2 mg/L], while the range for the DUI dataset is from -4.605 log (mg/L) units [0.01 mg/L] to -0.616 log (mg/L) units [0.54 mg/L]. Both datasets have the same lower boundary for the range, but the ME dataset spans a much larger region than the DUI dataset.

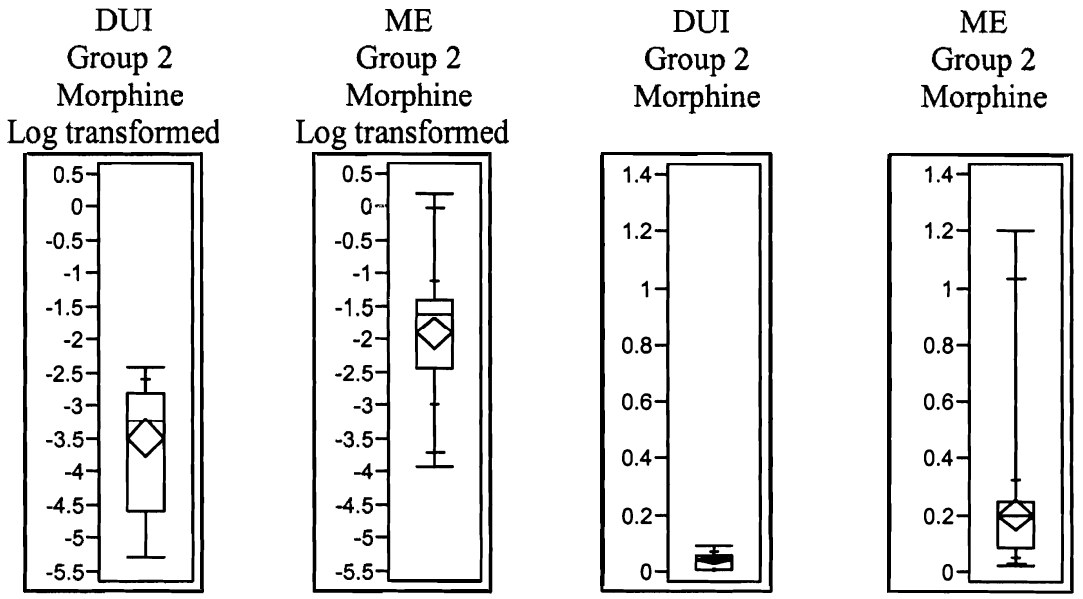


Figure 66: Box plots for Group 2 (morphine and ethanol) for DUI and ME – log-transformed and untransformed data

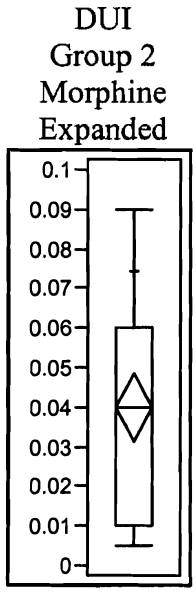


Figure 657: Expanded box plot for DUI Group 2 (morphine and ethanol) – untransformed data

Group 2 has a mean of -1.903 log (mg/L) units (± 0.109 standard errors) [0.15 mg/L] for the ME dataset and a mean of -3.492 log (mg/L) units (± 0.139 standard errors) [0.03 mg/L] for the DUI dataset. The range of the ME dataset is from -3.912 log (mg/L)

units [0.02 mg/L] to 0.182 log (mg/L) units [1.12 mg/L], and the range for the DUI dataset is from -5.298 log (mg/L) units [0.005 mg/L] to -2.408 log (mg/L) units [0.09 mg/L]. The range for the DUI dataset starts at a lower concentration, but it also has a lower maximum log concentration than the ME dataset.

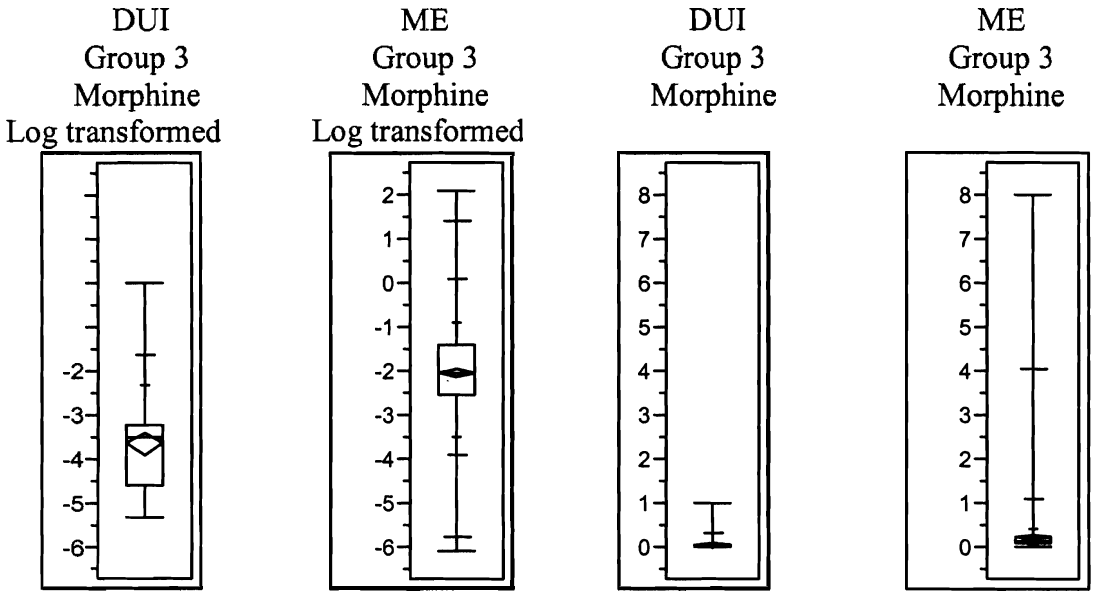


Figure 68: Box plots for Group 3 (morphine with any other combination of drug) for DUI and ME – log-transformed and untransformed data

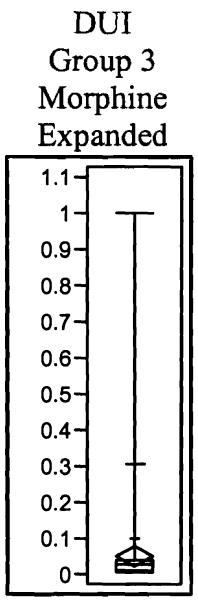


Figure 69: Expanded box plot for DUI Group 3 (morphine with any other combination of drug) – untransformed data

The means for Group 3 for the ME and DUI datasets are -2.047 log (mg/L) units (± 0.05 standard errors) [0.129 mg/L] and -3.657 log (mg/L) units (± 0.118 standard errors) [0.026 mg/L], respectively. The range for the ME dataset for this group is from -6.075 log (mg/L) units [0.0023 mg/L] to 2.079 log (mg/L) units [8 mg/L]. The DUI dataset had a smaller range than the range for the ME dataset, and it was from -5.298 log (mg/L) units [0.005 mg/L] to 0 log (mg/L) units [1.0 mg/L]. (The means and ranges reported in the previous sections came from the log-transformed concentrations for morphine and by untransforming these values; they are not necessarily the same as the means of the untransformed concentrations of morphine seen in Table 17.)

An ANOVA model was run on the log-transformed data for the concentration of morphine. Upon running the model, the interaction term was not determined to be significant (p-value = 0.897), so it was removed from the model. After the interaction term

was removed, the two independent fields, `group` and `database`, were both determined to be significant (p-values both less than 0.0001); therefore, there is a statistically significant difference between the ME and DUI datasets and a statistically significant difference between at least two of the three drug pattern groups. The plots of the means for each of these variable fields are seen in Figures 70 and 71. In Figure 70, the plot shows that the concentration of morphine is statistically higher for the ME dataset than for the DUI dataset. Because the ME and DUI datasets have significantly different means, there is not sufficient evidence to suggest of development of tolerance to morphine. The ANOVA indicated that there was a significant difference between the levels of the drug pattern groups, but it does not specify which differences between the means of the drug pattern groups are statistically significantly different. A Tukey's test was done to test for all pairwise differences, and the results of this test are seen in Table 17 where levels not connected by the same letter are significantly different. Group 1 is different from both Groups 2 and 3, but Groups 2 and 3 are not different from each other. (The values for the means seen in Table 17 are of the untransformed, or raw, concentrations of morphine, and they are reported in mg/L units.)

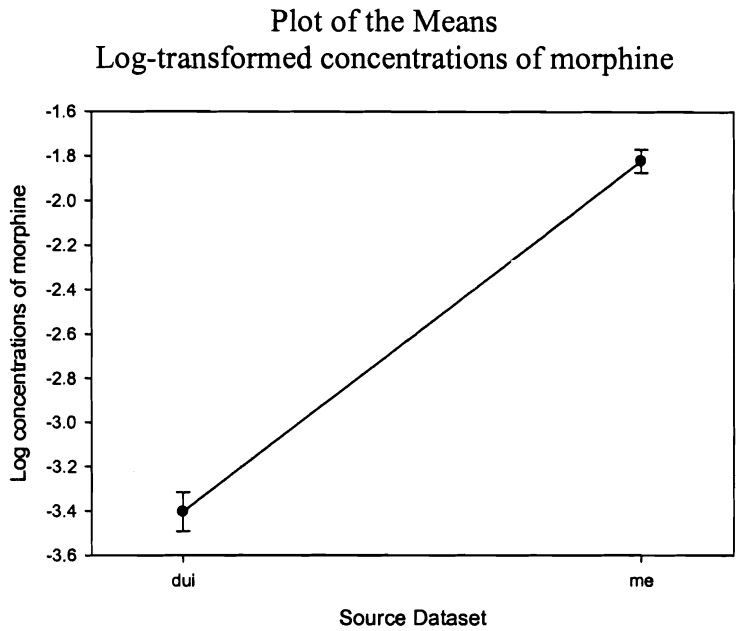


Figure 70: Plot of the means for "database" field for log-transformed morphine

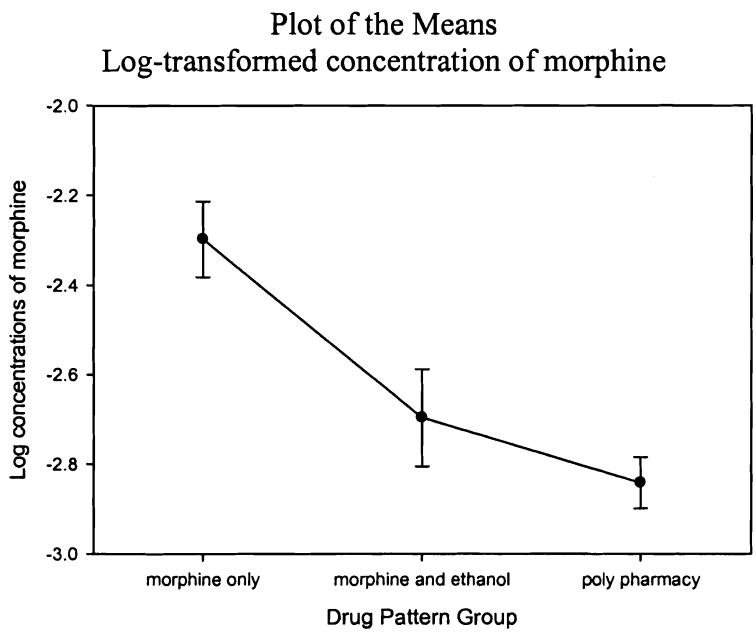


Figure 71: Plot of the means for "group" field for log-transformed morphine

Table 17: Levels of drug pattern groups for morphine from Tukey's test

Group	Letter	Mean(Std. Error)
Group 1	A	0.322(0.034)
Group 2	B	0.134(0.018)
Group 3	B	0.207(0.022)

The residual by predicted plot shows that the constant variance assumption is met and that there does not show a systematic pattern. There also does not seem to be any inadequacy with the ANOVA model. A normal quantile plot of the residuals for the fitted model was also constructed (Appendix C). This plot showed no obvious problems with the normality assumption as the residuals followed the diagonal in the plot.

There is not an exact range of concentration for morphine that clinicians have determined "fatal", so there is no definite way to know if there are individuals in the DUI dataset that had a concentration of morphine that would be termed be "lethal". Since there was a difference in the drug pattern groups in the ANOVA model, it implies that the concentration of morphine varies depending on whether the individual had morphine only or morphine with any drug. (Groups 2 and 3 were not statistically significantly different from each other). Because the `database` field was significant, then there is a significant statistical difference between the ME and DUI datasets, with the ME dataset having a higher concentration of morphine.

3.4 Conclusions

Based on the evidence provided by the ANOVA models discussed in the previous sections, there is not sufficient evidence to suggest the development of tolerances to diphenhydramine, cocaine, oxycodone, hydrocodone, methadone, or morphine. The only

exception to this is in the methadone dataset, where 14 individuals from the DUI dataset had "lethal" concentrations of methadone (determined by clinical evidence). Because there is not sufficient evidence for the development of tolerances with the ANOVA models does not imply that tolerances cannot be developed for any of these drugs. The collection of and data contained in the ME and DUI datasets do not provide adequate information to fully answer the question of tolerance with respect to these drugs. One of the reasons that the DUI dataset does not provide adequate information is because of the underestimation bias due to the tier system. This is a downward bias that underestimates the number of drugs found in persons suspected of DUI, which may have been a cause as to why the source datasets were determined to be statistically significantly different. The analyses of these six drug-specific datasets did provide selected cases that might be of interest to analyze further for case studies. These cases included the fourteen individuals in the DUI dataset with "lethal" concentrations of methadone in their system, the infant with an oxycodone death, the case in the ME dataset with a concentration of hydrocodone of 0.0003 mg/L, and the individuals in the ME dataset with diphenhydramine and cocaine concentrations higher than the upper boundary of the clinical "fatal" blood concentration. All of these cases would be of further interest to study for specific case analyses.

CHAPTER 4

Recommendations and Future Work

4.1 Summary of Project

The datasets used for this project came from the Office of the Chief Medical Examiner (ME dataset) and the Department of Forensic Science (DUI datasets). The original goals for this project were to investigate concentrations of THC in order to provide information to the Virginia General Assembly for the purpose of enacting legislation similar to that for ethanol intoxication. This information was anticipated to pertain to THC concentration in drivers, to concentrations of THC that "intoxicate" individuals, how ethanol affects THC, and to see if combinations of multiple drugs have additive affects. Unfortunately, due to the form of the database, the data collection procedures, and the overall quality of the data, this goal was unable to be achieved. Because the original goals could not be met, secondary objectives were set for this project. The new goals included comparing the concentrations of diphenhydramine, cocaine, oxycodone, hydrocodone, methadone, and morphine between the ME and DUI datasets and between three groups: drug only, drug and ethanol, or drug with any other combination of drug. The motivation

for this analysis was to investigate the possibility that some individuals in the DUI dataset have developed tolerances to some of these drugs to the point that the amount in their system approximates the amount that would be considered a lethal concentration for most individuals.

The first part of this thesis includes a description of each of the datasets and the modifications made to use them in analysis. Attempts were made to identify and correct all problems with these datasets in order to address the research questions. These problems in the datasets included changing character fields to numeric fields, correcting grammatical errors, addressing the missing data and underestimation issues, and creating new variable fields needed for analysis. Multiple changes had to be made to existing data fields in order to combine the data across the various datasets to allow the data analysis.

Once these changes were made to the datasets, a two-factor ANOVA was done to look for differences between the ME and DUI datasets and between the drug pattern groups created to divide each drug dataset. It was stated a priori that if the source datasets were not significantly different or if the concentrations in DUI dataset tended to be higher than the cases in the ME dataset, then this would constitute evidence to suggest that individuals in the DUI dataset that have built up tolerances to the drugs.

In the analyses of the six drug datasets, all of the datasets showed that the `database` field was significant, implying that the two datasets had statistically significant differences with respect to the drug concentration levels. Thus, there was no evidence to suggest that individuals in the DUI datasets had developed tolerances to any of the six drugs. In all of the datasets, except for morphine, the drug pattern group variable

field was significant, meaning that depending on whether an individual had the drug only, the drug and ethanol, or the drug with any other combination of drug, the concentration for that drug would vary. The analyses did not provide any evidence to support the original hypothesis, but they did pinpoint some individuals in the DUI and ME datasets that might be of further interest in a case study.

The main outcome of the project was that the datasets were unable to establish evidence of tolerance. This does not mean that those stopped for DUI do not have a tolerance to the six drugs studied, just that the analysis did not show any evidence for it. Several reasons are possible for this result, including that the DUI dataset was collected based on a tier system, and thus, not all concentrations for all drug were recorded in the dataset. The tier system creates mean concentrations that are biased downward (number of drugs and concentrations of drugs were all underestimated). Suggestions and recommendations will be made in the next section (Section 4.2) on how to change the datasets and data collection procedures in order to possibly allow the question of tolerance to be answered with a future dataset. Some of these changes are already being implemented into the ME database, and others will be given the OCME to suggest other ways to improve their data quality.

4.2 Recommendations for Change

On average, one out of every eight death certificates in Virginia is sent back to the medical examiner because it is "uninterpretable"[34]. This is evidence of the problems in the ME dataset. Another reason that the ME dataset had problems is because of the number of people who enter the information into the database. At this time, there are no rules

concerning how the data is entered or who enters it. Most of the fields in this database are free text fields, which allows for any character, number, or symbol to be entered into this field. If the database was reorganized so that it had drop-down menus instead of free text fields, many of the free-text problems would be eliminated. For example, in the `cause of death` field, instead of having a free text field where an unlimited number of characters or words can be entered, a menu that allows the data entry person to choose from "drug overdose", "lethal injection", "drowning", "fall", "gunshot wound", etc., would be more beneficial for analysis than the current design of the database. Eliminating free text fields would ensure that all entries into the database are standardized – a more useful form for analysis. The DUI datasets did not have many free text problems, but a drop-down menu for drugs or concentration amounts would be beneficial and a time-saving technique for future use of the dataset.

The DUI dataset is currently designed to test blood samples using the tier system described in Chapter 2. Because of the tier system, all drugs except ethanol are underestimated in the DUI dataset. The tier system was designed to save time and money for the state laboratories that do the analysis of the blood samples, and once a drug is found at a concentration high enough for conviction, then no other drugs are tested. To avoid the underestimation problem in the future, a random sample of blood samples from DUI suspects could be tested for all possible drugs. Thus, the drug concentrations from those in the random sample would not be underestimated. This data could then be used to give a better idea of the number of people in Virginia that drive under the influence of drugs, and to have a fuller range of the drugs that are being used across the state.

The Commonwealth of Virginia sends all death certificate information to the National Center for Health Statistics where the death certificates are used to generate National Death Index (NDI) codes. Currently, the OCME does not automatically receive these death codes, unless they specifically request them for research purposes. One suggestion for the OCME is to request these codes regularly and incorporate them into the ME database so that when research projects are being done, no further categorization on `cause of death` fields has to be done. Having the NDI codes would allow researchers to compare Virginia death statistics nationally or with specific states.

Another recommendation for the ME database is to attempt to detect the drugs in a consistent manner, i.e. in the same tissues. Usually, if the drug was detected in a tissue other than blood, it was because a blood sample was not available or more information was needed regarding the nature of the death (i.e. suicide vs. accident)[20]. If a standard system existed for routinely testing drugs in specific tissues, then fewer cases would be lost when trying to compare datasets. If a drug was tested and detected in the brain tissue, but not in blood, then that individual was not included in the analysis for this project because it is extremely problematic to convert brain concentrations to blood concentrations for analysis. Another solution to this problem is to have a reference tool designed that allows drug concentrations from one tissue to be converted into an equivalent amount for the tissue desired. This solution would be difficult because the clinical evidence for finding equivalent concentrations is inconclusive.

In the DUI dataset, the data collected does not provide the information needed to check for individuals appearing in the dataset more than once. To fix this problem, the

DUI dataset should collect and keep cross-referencing information to look for these individuals. This information could be a combination of driver's license number, social security number, date of birth, or full name. Keeping this information would help in analysis to verify the independence assumption.

Some of the changes that have been suggested are currently being made to the ME dataset. The OCME is reorganizing their database to be more "user-friendly" (including using drop-down menus), and the Chief Medical Examiner, Dr. Marcella Fierro, sent a statewide message to medical examiners concerning the consistency of their reporting[35]. "Furthermore, they are writing an SOP [Standard Operating Procedure] for the MEs to follow"[35] because of what this project uncovered.

4.3 Future Work

There are many questions that were left unanswered at the end of this project, and they are all ideas for future work with the ME and DUI datasets. One proposal is to use the location information (`court name` or `place of death`) to look at prevalence of drugs in certain regions of Virginia. This information could be displayed by means of a map with certain areas highlighted that show problem areas for specific drugs, displaying the number of drugs found in a certain area, or the ratio of the number of drugs found to the number of people (or children) in an area.

Another suggestion for future work is to redo the analysis done in this thesis, but without using the tier system mentioned in Section 2.2.5. To have more accurate results, a random sample of DUI suspects is needed that tests the blood samples for all possible drugs. By doing this, the blood concentration data for the DUI dataset would not be

underestimated. An alternative way to do this analysis would be to answer the question "Are the two source datasets equivalent when it comes to the selected drugs?" This was perhaps a better way to do this study.

References

References

1. Virginia General Assembly. Family violence fatality review teams established; model protocol and data management; membership; authority; confidentiality; etc. 1999;32.1-283.3:A-B. Available from: http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+32.1-283.3. Accessed May 25, 2006.

2. Virginia General Assembly. State child fatality review team established; membership; access to and maintenance of records; confidentiality; etc. 2004;32.1-283.1:A-E. Available from: http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+32.1-283.1. Accessed May 25, 2006.

3. Virginia General Assembly. Reports by medical examiners of deaths resulting from accidents. 1989;46.2-375. Available from: http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+46.2-375. Accessed May 25, 2006.

4. Virginia General Assembly. Driving motor vehicle, engine, etc., while intoxicated, etc. 2005;18.2-266. Available from: http://leg1.state.va.us/cgi-bin/legp504.exe?000+cod+18.2-266. Accessed June 20, 2006.

5. CIDOC guidelines for museum object information: Glossary. Available at: http://www.willpowerinfo.myby.co.uk/cidoc/guide/guideglo/htm. Accessed May 26, 2006.

6. Teorey T, Lightstone S, Nadeau T. *Database Modeling and Design.* Fourth Edition. San Francisco, CA 94111: Morgan Kaufmann Publishers; 2006.

7. Hogan R. *A Practical Guide to Data Base Design.* Englewood Cliffs, New Jersey 07632: Prentice-Hall, Inc.; 1990.

8. Jasc`o P, Lancaster FW. *Build Your Own Database.* Chicago and London: American Library Association; 1999.

9. Demographic variable. Available at: http://www.webref.org/psychology/d/demographic_variable.htm. Accessed June 2, 2006.

10. Saady JJ. 2006; email.

11. Minimum driving age USA. Available at: http://www.2pass.co.uk/ages2.htm. Accessed May 29, 2006.

12. Driver's license. Available at: http://en.wikipedia.org/wiki/Driver's_license. Accessed May 29, 2006.

13. Saady JJ. 2006; email.

14. Hopkins WG. A new view of statistics. Available at:

http://sportsci.org/resource/stats/modelsdetail.html. Accessed June 15, 2006.

15. Chandler D. 2006; email from Doug Chandler to Joe Saady. June 20, 2006.

16. Underestimation. Available at:

http://wordenet.princeton.edu/perl/webwn?s=underestimation. Accessed June 2, 2006.

17. Concatenating. In: *SAS Language: Reference, Version 6, First Edition.* First Edition. Cary, NC: SAS Institute Inc.; 1990:139-140.

18. Kinds of data. In: *SAS Language: Reference, Version 6, First Edition.* First Edition. Cary, NC: SAS Institute Inc.; 1990:29.

19. Saady JJ. 2006; meeting.

20. Saady JJ. July 14, 2006; meeting.

21. Saady JJ. July 12, 2006; phone conversation.

22. van Belle G. Use at least twelve observations in constructing a confidence interval. In: *Statistical Rules of Thumb.* United States of America: John Wiley & Sons, Inc; 2002:18-19.

23. JMP. Quantiles and box plots. Accessed July 11, 2006.

24. Wood D. Diphenhydramine health article. Available at: http://www.healthline.com/gale.jsp?utm_source=healthline&utm_medium=google&utm_campaign=nidus&utm_term=diphenhydramine&pagenum=1&contentId=gec_01_00153. Accessed July 22, 2006.

25. Karch SB. Diphenhydramine toxicity: Comparisons of postmortem findins in diphenhydramine-, cocaine-, and heroin-related deaths. *Am J Forensic Med Pathol.* 1998;19:143 Accessed May 24, 2006.

26. Baselt RC. *Disposition of Toxic Drugs & Chemicals in Man.* 7th Ed. ed. Foster City, CA: Biomedical Publications; 2004.

27. Cocaine. Available at: http://en.wikipedia.org/w/index.php?title=Cocaine&oldid=65441555. Accessed July 23, 2006.

28. Spiehler VR, Reed D. Brain concentrations of cocaine and benzoylecognine in fatal cases. *J Forensic Sci.* 1985;30:1003-1011 Accessed July 23, 2006.

29. Oxycodone. Available at: http://www.streetdrugs.org/oxycodone.htm. Accessed July 23, 2006.

30. Hydrocodone. Available at: http://en.wikipedia.org/wiki/Hydrocodone. Accessed July 23, 2006.

31. Methadone. Available at: http://www.streetdrugs.org/methadone.htm. Accessed July 23, 2006.

32. Methadone. Available at: http://en.wikipedia.org/wiki/Methadone. Accessed July 23, 2006.

33. Morphine. Available at: http://www.chamisamesa.net/morphine.html. Accessed July 23, 2006.

34. Powell V. April 12, 2006; meeting.

35. Saady JJ. May 9, 2006;email.

APPENDIX A

Testing Protocol for Drugs in the DUI Dataset

FEE SCHEDULE FOR APPROVED LABORATORY ANALYSES OF BLOOD FOR ALCOHOL AND DRUGS

ALLOWED FEES

Level I	Alcohol Testing	$50
Level II	Drug Screening	$50
	Drug Identification/Quantitation	$75 / Drug Class *
Level III	Drug Screening	$50
	Drug Identification/Quantitation	$75 / Drug Class *

* Includes results for all drugs reported in a class

ANALYTICAL SCHEME

All analyses shall be for the unconjugated (free) form of drugs.

1. Analyze all samples for ethanol (Alcohol screening may be performed, if desired, using immunoassay.)
   a. If ethanol is less than 0.09%, include it in the report, and go to 2.
   b. If ethanol is at or above 0.09%, stop; report results.

2. Perform Level II Drug Screening:
   a. If no drug classes are detected, go to 4.
   b. If any drug is tentatively present at or above the reporting limit, go to 3.

3. Perform Level II Identification/Quantitation:
   a. If no drugs are present, or are identified as present but at a concentration below the reporting limit, go to 4.
   b. If drugs are identified as present at a concentration at or above the reporting limit, but below the stop analysis limit, include them in the report, and go to 4.
   c. If drugs are identified as present at a concentration at or above the stop analysis limit, stop; report results.

4. Perform Level III Screening:
   a. If any drug is tentatively identified as present, go to 5.
   b. If no drugs are present, stop; report results.

5. Perform Level III Identification/Quantitation:
   a. Report results.

ANALYTICAL LIMITS

Level I Alcohol Testing (Limits are in units of % by weight by volume)

	REPORTING LIMIT	STOP ANALYSIS LIMIT**
Ethanol	0.01	0.09

***** Do not proceed further in the analytical scheme when results at or above this concentration are obtained.***

Level II Drug Screening and Identification/Quantitation (Limits are in units of mg/L)

SCREENING	**IDENTIFICATION/QUANTITATION**		
DRUG CLASS	DRUG	REPORTING LIMIT	STOP ANALYSIS LIMIT **
Amphetamine	Amphetamine	0.01	0.1
	Methamphetamine	0.01	0.2
	3,4-Methylenedioxymethamphetamine	0.01	0.02
	Methylenedioxyamphetamine	0.01	0.02
Barbiturate	Amobarbital	0.5	2
	Butabarbital	0.5	2
	Butalbital	0.5	10
	Pentobarbital	0.5	2
	Phenobarbital	0.5	30
	Secobarbital	0.5	2
Cocaine/ Benzoylecgonine	Cocaine	0.01	0.1
	Benzoylecgonine	0.1	-
	Cocaethylene	0.01	0.1
Benzodiazepine	Alprazolam	0.02	0.06
	Chlordiazepoxide	0.2	4
	Clonazepam	0.01	0.06
	Diazepam	0.1	1
	Lorazepam	0.01	0.3
	n-Desalkylflurazepam	0.01	-
	Nordiazepam	0.1	1
	Oxazepam	0.1	2
	Temazepam	0.02	1
Cannabinoid	Tetrahydrocannabinol	0.001	0.002
	THC carboxylic acid	0.005	-
Opiate	Codeine	0.02	0.5
	Hydrocodone	0.02	0.06
	Morphine	0.02	0.1
	6-Acetylmorphine	0.01	0.01
	Oxycodone	0.02	0.06

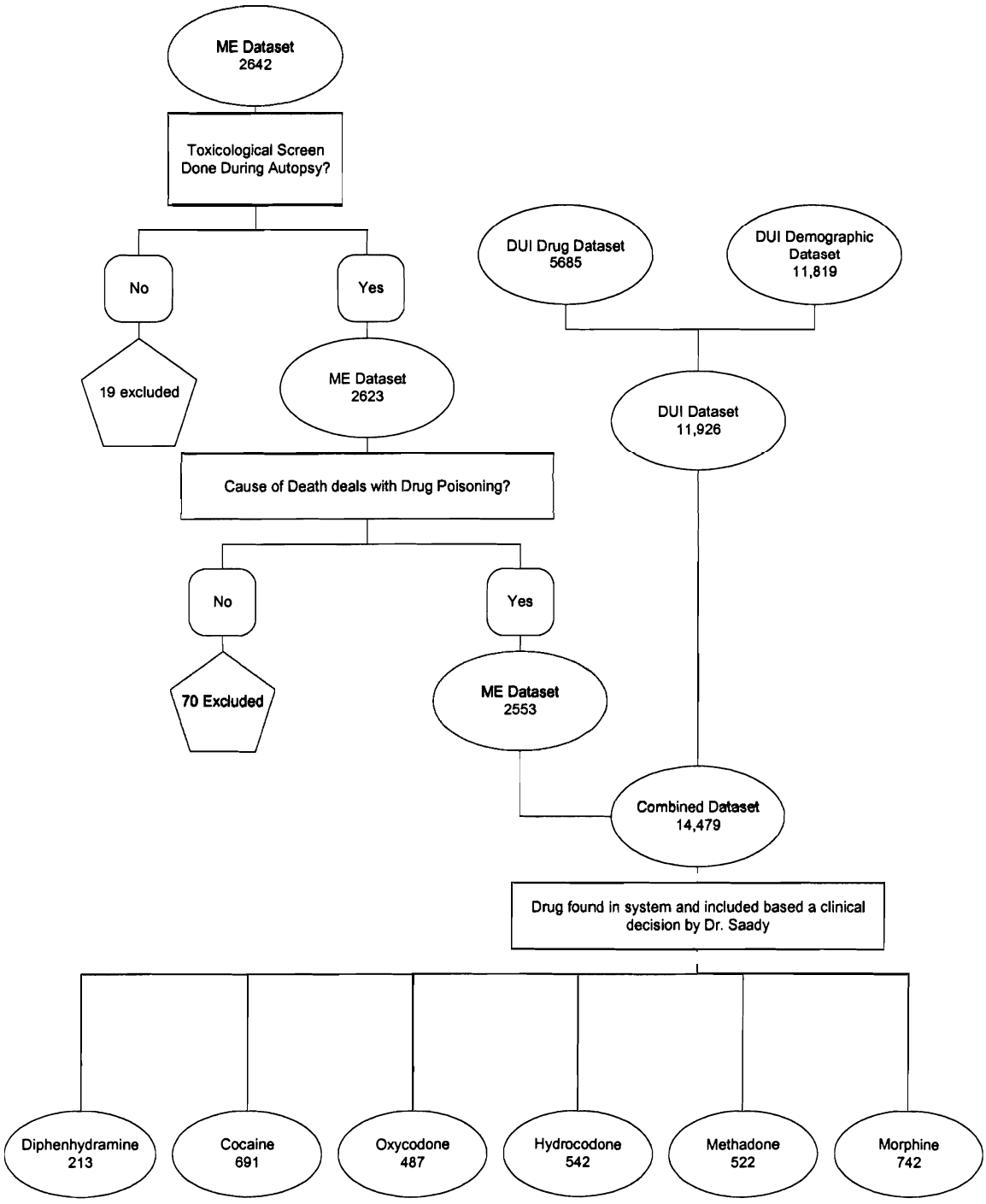
***** Do not proceed further in the analytical scheme when results at or above this concentration are obtained.***

Level III Drug Identification/Quantitation (Limits are in units of mg/L)

DRUG CLASS	DRUG	REPORTING LIMIT
Antihistamine	Chlorpheniramine	0.02
	Brompheniramine	0.02
	Dextromethorphan	0.1
	Diphenhydramine	0.02
	Promethazine	0.02
Antidepressant/ Antipsychotic	Amitriptyline	0.05
	Buproprion	0.05
	Clomipramine	0.05
	Clozapine	0.05
	Desipramine	0.05
	Doxepin	0.05
	Haloperidol	0.02
	Imipramine	0.05
	Mirtazapine	0.02
	Nortriptyline	0.05
	Nordoxepin	0.05
	Trazodone	0.05
Hypnotic	Zolpidem	0.02
Muscle Relaxant	Carisoprodol	1
	Meprobamate	1
	Methocarbamol	1
	Cyclobenzaprine	0.02
Opiate/ Opiate-like	Dihydrocodeine	0.02
	Hydromorphone	0.02
	Oxymorphone	0.02
	Meperidine	0.02
	Methadone	0.02
	Pentazocine	0.02
	Propoxyphene	0.02
	Tramadol	0.02
Ketamine	Ketamine	0.05
Phencyclidine	Phencyclidine	0.01

Appendix B

Flow Chart of Exclusion Criteria



Appendix C

Analyses of the Residuals

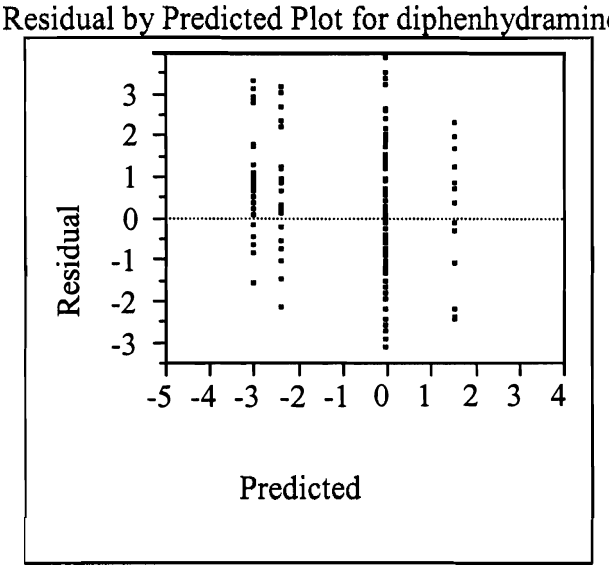


Figure 66c: Residual by predicted plot for diphenhydramine

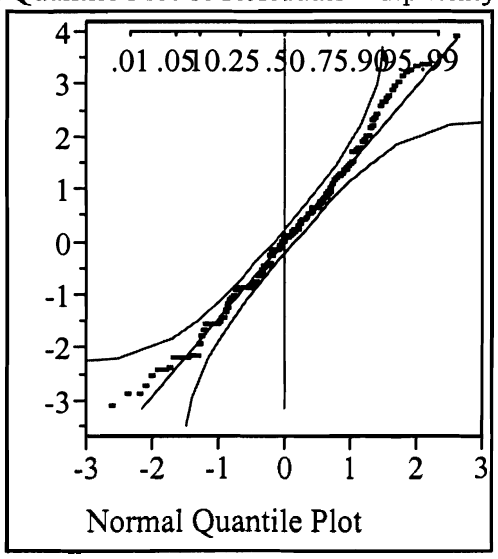


Figure 67c: Normal quantile plot of residuals for diphenhydramine

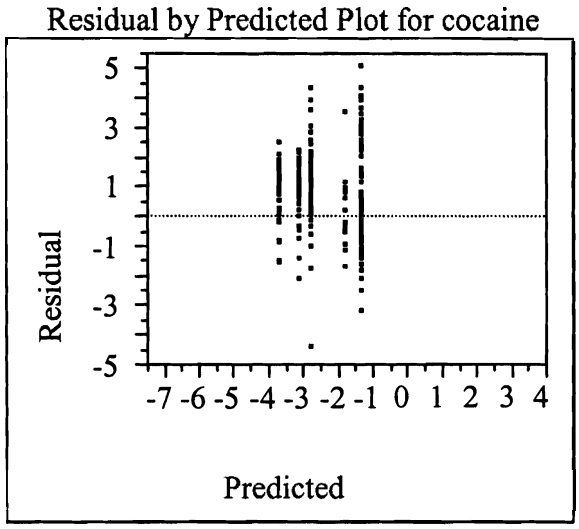


Figure 68c: Residual by predicted plot for cocaine

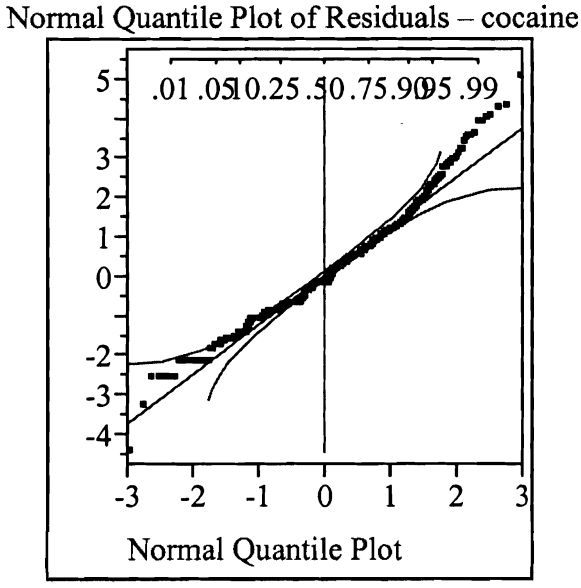


Figure 69c: Normal quantile plot of residuals for cocaine

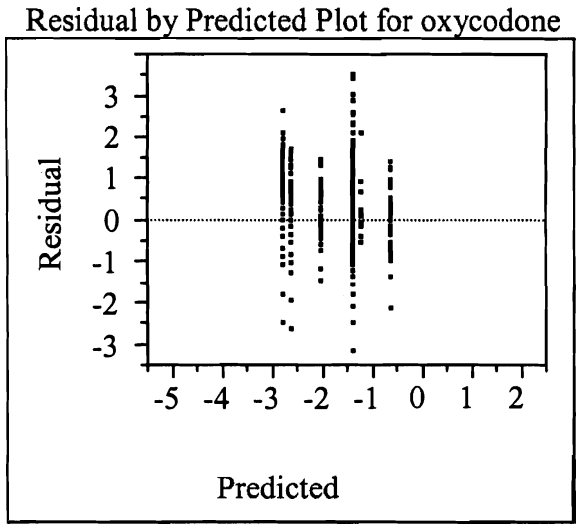


Figure 70c: Residual by predicted plot for oxycodone

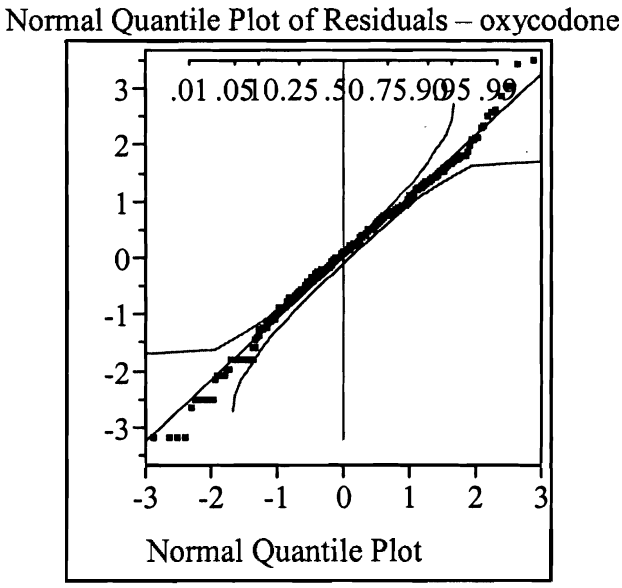


Figure 71c: Normal quantile plot of residuals for oxycodone

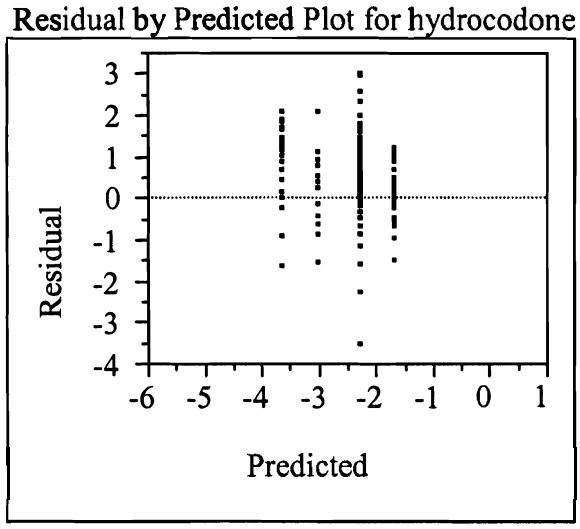


Figure 72c: Residual by predicted plot for hydrocodone

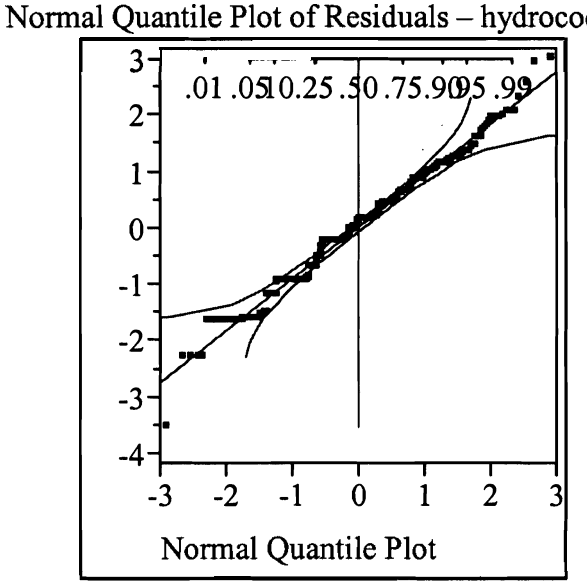


Figure 73c: Normal quantile plot of residuals for hydrocodone

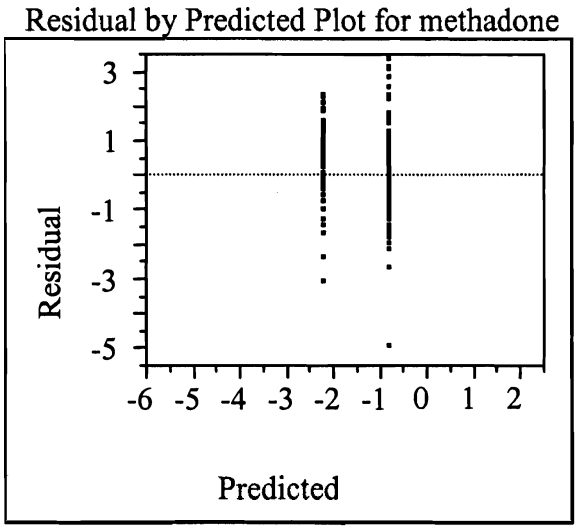


Figure 74c: Residual by predicted plot for methadone

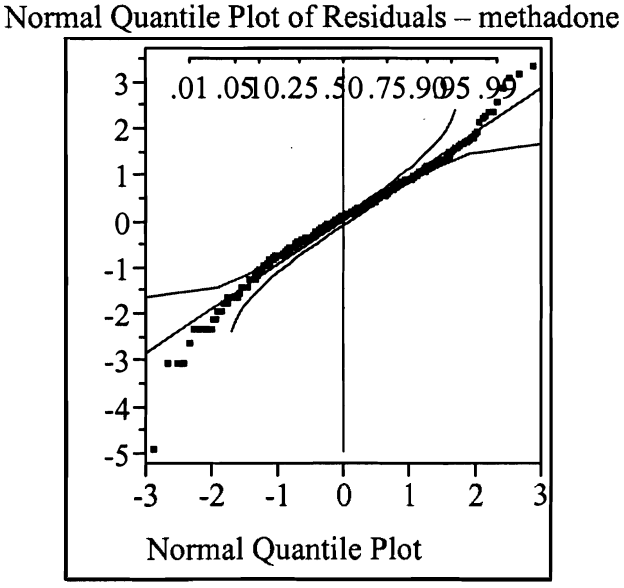


Figure 75c: Normal quantile plot of residuals for methadone

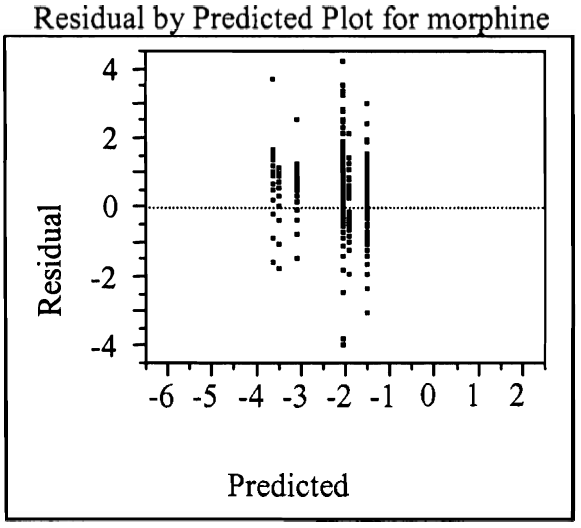


Figure 76c: Residual by predicted plot for morphine

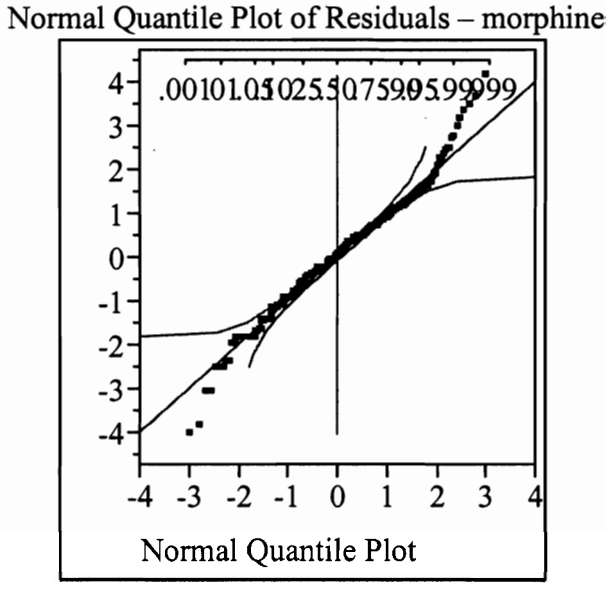


Figure 77c: Normal quantile plot of residuals for morphine

Appendix D

SAS Code

```
libname thesis 'C:\Documents and Settings\Owner\Desktop';
*ME DATABASE;
*Import the JMP file - from the Excel file;
PROC IMPORT OUT= thesis.ME1
            DATAFILE= "E:\Thesis Files\MEoriginal.jmp"
            DBMS=JMP REPLACE;
RUN;

proc freq data = thesis.me1;
      tables fatality premise_of_death;
run;

/*Combine asphyxia, poisoning, vehicular fatality groups and combine
premise of death street, house, water, and hospital groups*/

data thesis.meclean1;
      set thesis.me1;

      attrib fatalnew format=$58.;
      attrib premisenew format=$19.;

      if fatality = 'ASPHYXIA: ASPIRATION/CAFE CORONARY' | fatality =
            'ASPHYXIA: MECHANICAL/POSITIONAL' | fatality =
            'ASPHYXIA: PLASTIC BAG' then fatalnew = 'Asphyxia';
      else if fatality = 'POISONING: CARBON MONOXIDE (FAULTY HEATER)' |
            fatality = 'POISONING: CARBON MONOXIDE (FIRE SMOKE
            INHALATION)' | fatality = 'POISONING: CARBON MONOXIDE
            (GENERATOR)' | fatality = 'POISONING: CARBON MONOXIDE (MOTOR
            VEHICLE EXHAUST)' then fatalnew = 'CO poisoning';
      else if fatality = 'POISONING: ALCOHOL-ETHANOL' | fatality =
            'POISONING: DRUGS, OTHER POISONS' | fatality =
            'POISONING: OTHER ALCOHOLS' then fatalnew =
            'Drug Poisoning';
      else if fatality = 'VEHICULAR: AUTO/TRUCK (DRIVER)' | fatality =
            'VEHICULAR: AUTO/TRUCK (PASSENGER)' |
            fatality = 'VEHICULAR: AUTO/TRUCK (PEDESTRIAN)'
            then fatalnew = 'Vechicle';
      else fatalnew = fatality;
      if premise_of_death = 'STREET: ADJACEN' | premise_of_death =
            'STREET: ALLEY' | premise_of_death = 'STREET: BRIDGE' |
            premise_of_death = 'STREET: DITCH' | premise_of_death =
            'STREET: DRIVEWA' | premise_of_death = 'STREET: HIGHWAY' |
            premise_of_death = 'STREET: INTERST' | premise_of_death =
            'STREET: NONSPEC' | premise_of_death = 'STREET: PARKING' |
            premise_of_death = 'STREET: SIDEWAL' then premisenew =
            'Street';
```

```
      else if premise_of_death = 'WATER: LAKE' | premise_of_death =
            'WATER: OTHER' then premisenew = 'Water';
      else if premise_of_death = 'HOME: APARTMENT' | premise_of_death =
            'HOME: GARAGE/SH' | premise_of_death = 'HOME: HOUSE/RES' |
            premise_of_death = 'HOME: NONSPECIF' | premise_of_death =
            'HOME: TRAILER' | premise_of_death = 'HOME: YARD/PORC'
            then premisenew = 'Home';
      else if premise_of_death = 'HOSPITAL/INPATI' |
            premise_of_death = 'HOSPITAL/MENTAL' | premise_of_death =
            'HOSPITAL/OUTPAT' then premisenew = 'Hospital';
      else premisenew = premise_of_death;
      drop premise_of_death fatality;

*Remove all double spaces and any periods at the end;
      index = find(cause_of_death, '  ', 'i');
      new = substrn(cause_of_death, 1, index) ||
            substrn(cause_of_death, index+2, length(cause_of_death));
      cause_of_death = new;
      new2 = cause_of_death;
      end = length(cause_of_death);
      str = substrn(new2,end);
      if str eq '.' then per = 1;
            else per = 0;
      if per = 1 then substr(new2, end) = '';
      cause_of_death = new2;
      drop index new new2 per str end;


*Fixing drug names to match in both datasets;
      if drug_s_ = "METHYLENEDIOXYAMPHETAMINE" then drug_s_ = "MDA";
      if drug_s_ = "METHYLENEDIOXYMETHAMPHETAMINE" then drug_s_ =
            "MDMA";
      if drug_s_ = "BENZOYL ECGONINE" then drug_s_ = "BENZOYLECGONINE";
      if drug_s_ = "DESOXYCHLORDIAZEPOXIDE" then drug_s_ =
            "CHLORDIAZEPOXIDE";
      if drug_s_ = "MONOACETYLMORPHINE" then drug_s_ = "ACETYLMORPHINE";
      if drug_s_ = "BUPROPION (WELLBUTRIN)" then drug_s_ = "BUPROPION";

      attrib gender format = $4.;
      if sex = 'FEMALE' then gender = 'F';
      else if sex = 'MALE' then gender = 'M';
      else if sex = 'NULL' then gender = 'N';
      else gender = '';


*so that blood will come first alphabetically (or last -
      descending);
      if bac = 'BLOOD(POST)' then bac = 'aBLOOD(POST)';
      if bac = 'BLOOD(PRE)' then bac = 'bBLOOD(PRE)';
      if bac = 'BILE' then bac = 'cBILE';
      if bac = 'LIVER' then bac = 'dLIVER';
      if bac = 'URINE' then bac = 'eURINE';
      if bac = 'GASTRIC' then bac = 'fGASTRIC';
```

```
      if bac = 'VITREOUS' then bac = 'gVITREOUS';
      if bac = 'BRAIN' then bac = 'hBRAIN';


      drug = drug_s_;

      attrib database format = $3.;
      database = 'me';

      attrib presence format = $31.;
      presence = '';

      *Count will keep track of the number of grammatical fixes
            for each observation;
      count = 0;

      drop sex drug_s_;

run;

proc freq data=thesis.meclean1;
      tables fatalnew premisenew;
run;

/*match tissues to correct drug & fix cause of death misspellings
      and create units of measurement*/
data thesis.metissues;
      set thesis.meclean1;
      if drug = 'ETHANOL' then unit = 'Grams%';
      else if bac = 'aBLOOD(POST)' | bac = 'fGASTRIC' | bac = 'gVITREOUS'
            | bac = 'bBLOOD(PRE)' | bac = 'eURINE' then unit = 'mg/L';
      else if bac = 'dLIVER' | bac = 'cBILE' | bac = 'hBRAIN' then unit =
            'mg/kg';
      else unit = '';

      index = find(cause_of_death, 'AMTRIPTYLINE', 'i');
      if index ne 0 then do;
            temp = substrn(cause_of_death,1,index-1) ||
            'AMITRIPTYLINE' ||
            substrn(cause_of_death,index+12,length(cause_of_death));
            count = count + 1;
      end;
      else temp = cause_of_death;
      cause_of_death = temp;

      index = find(cause_of_death, 'INTOXICIATION', 'i');
      if index ne 0 then do;
            temp = substrn(cause_of_death,1,index-1) ||
            'INTOXICATION' ||
            substrn(cause_of_death,index+13,length(cause_of_death));
            count = count + 1;
      end;
      else temp = cause_of_death;
```

```
cause_of_death = temp;

index = find(cause_of_death, 'INTOXICATIKON', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'INTOXICATION' ||
      substrn(cause_of_death,index+13,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'POINSONING', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'POISONING' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'POISIONING', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'POISONING' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HOUSE FIRE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HOUSEFIRE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HERION', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HEROIN' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HERION', 'i');
```

```
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HEROIN' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HREOIN', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HEROIN' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HEROIN, POISONING ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HEROIN POISONING' ||
      substrn(cause_of_death,index+18,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'LIKELY ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'LIKELY,' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'AND', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      '&' ||
      substrn(cause_of_death,index+3,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

*Occurs more than once - see comment below;
index = find(cause_of_death, 'AND', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
```

```
      '&' ||
      substrn(cause_of_death,index+3,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DUE TO', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'D/T' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

*if 'DUE TO' occurred more than once, find only picks up the
      first time;

index = find(cause_of_death, 'DUE TO', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'D/T' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'D/T TO', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'D/T' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, ', & ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      ' & ' ||
      substrn(cause_of_death,index+4,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

*Delete all spaces after commas;
index = find(cause_of_death, ', ', 'i');
if index ne 0 then do;
```

```
      temp = substrn(cause_of_death,1,index-1) ||
      ',' ||
      substrn(cause_of_death,index+2,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, ', ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      ',' ||
      substrn(cause_of_death,index+2,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

*Remove all periods and replace with commas;
index = find(cause_of_death, '.', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      ',' ||
      substrn(cause_of_death,index+1,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, '))', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      ')' ||
      substrn(cause_of_death,index+2,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OLAANZAPINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OLANZAPINE' ||
      substrn(cause_of_death,index+11,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'CINTRIBUTING', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CONTRIBUTING' ||
```

```
      substrn(cause_of_death,index+12,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'AUTOMOBILE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CAR' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'AUTO ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CAR' ||
      substrn(cause_of_death,index+5,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DEPHENHYDRAMINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DIPHENHYDRAMINE' ||
      substrn(cause_of_death,index+15,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DIPENHYDRAMINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DIPHENHYDRAMINE' ||
      substrn(cause_of_death,index+14,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;


index = find(cause_of_death, 'BENADRYL', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DIPHENHYDRAMINE' ||
      substrn(cause_of_death,index+8,length(cause_of_death));
      count = count + 1;
```

```
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DIPHENYDRAMINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DIPHENHYDRAMINE' ||
      substrn(cause_of_death,index+14,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DIPHENYHDRAMINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DIPHENHYDRAMINE' ||
      substrn(cause_of_death,index+15,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'COCANE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'COCAINE' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'CICAINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'COCAINE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'COCAIN ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'COCAINE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;
```

```
index = find(cause_of_death, 'COCIANE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'COCAINE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HEARD', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HEART' ||
      substrn(cause_of_death,index+5,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'SYSTEMMIC', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'SYSTEMIC' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HEROINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HEROIN' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OXCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+8,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'EXYCODONE', 'i');
if index ne 0 then do;
```

```
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'PXYCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OXYCONTIN', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OCYCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPROZOLAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OXYCOCONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
```

```
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OXYCONDONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'OXYODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'OXYCODONE' ||
      substrn(cause_of_death,index+8,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HYDROCONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HYRDOCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+11,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HYDROCDONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
```

```
cause_of_death = temp;

index = find(cause_of_death, 'HYDOCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HYDROCOONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'HYROCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'MYODROCODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'HYDROCODONE' ||
      substrn(cause_of_death,index+12,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'METHADON ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'METHADONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'COCAINEE', 'i');
```

```
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'COCAINE' ||
      substrn(cause_of_death,index+8,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'MENTADONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'METHADONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'METHADNE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'METHADONE' ||
      substrn(cause_of_death,index+8,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'METHODONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'METHADONE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'METHADDONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'METHADONE' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'MEHTADONE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'METHADONE' ||
```

```
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'COODEINE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CODEINE' ||
      substrn(cause_of_death,index+8,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALCOHOL', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ETHANOL' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'POISONINGN', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'POISONING' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'EHTANOL', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ETHANOL' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'POIS0NING', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'POISONING' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
```

```
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'SBBSTANCE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'SUBSTANCE' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'MORPH ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'MORPHINE' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DEXTROMETH ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DEXTROMETHORPHAN' ||
      substrn(cause_of_death,index+11,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'COC ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'COCAINE' ||
      substrn(cause_of_death,index+4,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPRAZOLM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;
```

```
index = find(cause_of_death, 'ALPAZOLAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPRAXOLAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'APRAZOLAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPROAZOLAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+11,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPRAZOLEM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPRA ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
```

```
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+6,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'ALPAZOLAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'ALPRAZOLAM' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'CODIENE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CODEINE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'CODEIN ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CODEINE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'PROP ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'PROPOXYPHENE' ||
      substrn(cause_of_death,index+5,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'PROPOXYHENE', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'PROPOXYPHENE' ||
      substrn(cause_of_death,index+11,length(cause_of_death));
      count = count + 1;
```

```
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'PROPOX ', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'PROPOXYPHENE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'PROPOX,', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'PROPOXYPHENE' ||
      substrn(cause_of_death,index+7,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'CITALOPROM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CITALOPRAM' ||
      substrn(cause_of_death,index+10,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'CIALOPRAM', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'CITALOPRAM' ||
      substrn(cause_of_death,index+9,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;

index = find(cause_of_death, 'DRUB', 'i');
if index ne 0 then do;
      temp = substrn(cause_of_death,1,index-1) ||
      'DRUG' ||
      substrn(cause_of_death,index+4,length(cause_of_death));
      count = count + 1;
end;
else temp = cause_of_death;
cause_of_death = temp;
```

```
      *Changinng all variations to poisoning;
      index = find(cause_of_death, 'TOXICITY', 'i');
      if index ne 0 then temp = substrn(cause_of_death,1,index-1) ||
            'POISONING' ||
            substrn(cause_of_death,index+8,length(cause_of_death));
      else temp = cause_of_death;
      cause_of_death = temp;

      index = find(cause_of_death, 'INTOXICATION', 'i');
      if index ne 0 then temp = substrn(cause_of_death,1,index-1) ||
            'POISONING' ||
            substrn(cause_of_death,index+12,length(cause_of_death));
      else temp = cause_of_death;
      cause_of_death = temp;

      index = find(cause_of_death, 'OVERDOSE', 'i');
      if index ne 0 then temp = substrn(cause_of_death,1,index-1) ||
            'POISONING' ||
            substrn(cause_of_death,index+11,length(cause_of_death));
      else temp = cause_of_death;
      cause_of_death = temp;


      if drug_s_ = 'DIPHENYLAMINE' then drug_s_ = 'DIPHENHYDRAMINE';

      retain count;

      drop temp;
      drop index;
run;

proc print data=thesis.metissues;
      var drug unit bac presence;
run;

proc freq data=thesis.metissues;
      tables bac;
run;

*sort to be in correct order to next data statement;
proc sort data=thesis.metissues;
      by case_number drug descending bac;
run;

/*Creation of new variable arrays*/
data thesis.mefinal;
      set thesis.metissues;
      by case_number drug descending bac;
      array names {139} $100. acetaminophen_n acetone_n acetylmorphine_n
            alprazolam_n amitriptyline_n amoxapine_n amphetamine_n
            benzoylecgonine_n benztropine_n brompheniramine_n bupropion_n
            busiprone_n butabarbital_n butalbital_n caffeine_n
```

```
carbamazepine_n carbondioxide_n carbonmonoxide_n
carisoprodol_n chlordiazepoxide_n chlorodifluoromethane_n
chlorpheniramine_n chlorpromazine_n citalopram_n
clomipramine_n clonzepam_n clozapine_n cocaethlyene_n
cocaine_n codeine_n cyanide_n cyclobenzaprine_n
desalkyflurazepam_n desipramine_n dextromethorphan_n
diazepam_n difluoroethane_n dihydrocodeine_n diltiazem_n
diphenhydramine_n doxepin_n doxylamine_n ephedrine_n
ethanol_n ethyleneglycol_n fenfluramine_n fentanyl_n
fluoxetine_n fluvoxamine_n gabapentin_n
gammahydroxybutyrate_n guaifenesin_n hydrocodone_n
hydromorphone_n hydroxyzine_n ibuprofen_n imipramine_n
insulin_n isopropanol_n ketamine_n lithium_n lorazepam_n
lysergicaciddiethylamide_n mda_n mdma_n meclizine_n
meperidine_n meprobamate_n metaxalone_n methadone_n
methamphetamine_n methanol_n methocarbamol_n methotrexate_n
methylphenidate_n midazolam_n mirtazapine_n molidone_n
morphine_n naproxen_n nefazodone_n nicotine_n nitricacid_n
norchlorcyclizine_n nordiazepam_n nordoxepin_n
norfluoxetine_n normeperidine_n norpropoxyphene_n
nortriptylene_n olanzapine_n orphenadrine_n oxazepam_n
oxycodone_n oxymorphone_n pancuroniumbromide_n paraquat_n
paroxetine_n pentazocine_n pentobarbital_n pentoxifylline_n
phencyclidine_n phenobarbital_n phentermine_n
phenyltoloxamine_n phenytoin_n primidone_n promethazine_n
propane_n propoxyphene_n propranolol_n propyleneglycol_n
quetiapine_n quinine_n ritalinicacid_n salicylate_n
salicylicacid_n secobarbital_n sertraline_n temazepam_n
tetrafluroethane_n tetrahydrocannabinol_n thccarboxylicacid_n
theophylline_n thiopental_n thioridazine_n toluene_n
tramadol_n tranylcypromine_n trazodone_n triazolam_n
tricyclic_n trifluoperazine_n trihexyphenidyl_n
trimipramine_n valproic_n venlafaxine_n verapamil_n
zolpidem_n ('acetaminophen', 'acetone', 'acetylmorphine',
'alprazolam', 'amitriptyline', 'amoxapine', 'amphetamine',
'benzoylecgonine', 'benztropine', 'brompheniramine',
'bupropion', 'busiprone', 'butabarbital', 'butalbital',
'caffeine', 'carbamazepine', 'carbon dioxide', 'carbon
monoxide', 'carisoprodol', 'chlordiazepoxide',
'chlorodifluoromethane', 'chlorpheniramine',
'chlorpromazine', 'citalopram', 'clomipramine', 'clonzepam',
'clozapine', 'cocaethlyene', 'cocaine', 'codeine', 'cyanide',
'cyclobenzaprine', 'desalkyflurazepam', 'desipramine',
'dextromethorphan', 'diazepam', 'difluoroethane',
'dihydrocodeine', 'diltiazem', 'diphenhydramine', 'doxepin',
'doxylamine', 'ephedrine', 'ethanol', 'ethyleneglycol',
'fenfluramine', 'fentanyl', 'fluoxetine', 'fluvoxamine',
'gabapentin', 'gammahydroxybutyrate', 'guaifenesin',
'hydrocodone', 'hydromorphone', 'hydroxyzine', 'ibuprofen',
'imipramine', 'insulin', 'isopropanol', 'ketamine',
'lithium', 'lorazepam', 'lysergicaciddiethylamide', 'mda',
'mdma', 'meclizine', 'meperidine', 'meprobamate',
'metaxalone', 'methadone', 'methamphetamine', 'methanol',
```

```
        'methocarbamol', 'methotrexate', 'methylphenidate',
        'midazolam', 'mirtazapine', 'molidone', 'morphine',
        'naproxen', 'nefazodone', 'nicotine', 'nitricacid',
        'norchlorcyclizine', 'nordiazepam', 'nordoxepin',
        'norfluoxetine', 'normeperidine', 'norpropoxyphene',
        'nortriptylene', 'olanzapine', 'orphenadrine', 'oxazepam',
        'oxycodone', 'oxymorphone', 'pancuroniumbromide', 'paraquat',
        'paroxetine', 'pentazocine', 'pentobarbital',
        'pentoxifylline', 'phencyclidine', 'phenobarbital',
        'phentermine', 'phenyltoloxamine', 'phenytoin', 'primidone',
        'promethazine', 'propane', 'propoxyphene', 'propranolol',
        'propyleneglycol', 'quetiapine', 'quinine', 'ritalinicacid',
        'salicylate', 'salicylicacid', 'secobarbital', 'sertraline',
        'temazepam', 'tetrafluroethane', 'tetrahydrocannabinol',
        'thccarboxylicacid', 'theophylline', 'thiopental',
        'thioridazine', 'toluene', 'tramadol', 'tranylcypromine',
        'trazodone', 'triazolam', 'tricyclic', 'trifluoperazine',
        'trihexyphenidyl', 'trimipramine', 'valproic', 'venlafaxine',
        'verapamil', 'zolpidem');

array drugs {139} acetaminophen acetone acetylmorphine alprazolam
      amitriptyline amoxapine amphetamine benzoylecgonine
      benztropine brompheniramine bupropion busiprone butabarbital
      butalbital caffeine carbamazepine carbondioxide
      carbonmonoxide carisoprodol chlordiazepoxide
      chlorodifluoromethane chlorpheniramine chlorpromazine
      citalopram clomipramine clonzepam clozapine cocaethlyene
      cocaine codeine cyanide cyclobenzaprine desalkyflurazepam
      desipramine dextromethorphan diazepam difluoroethane
      dihydrocodeine diltiazem diphenhydramine doxepin doxylamine
      ephedrine ethanol ethyleneglycol fenfluramine fentanyl
      fluoxetine fluvoxamine gabapentin gammahydroxybutyrate
      guaifenesin hydrocodone hydromorphone hydroxyzine ibuprofen
      imipramine insulin isopropanol ketamine lithium lorazepam
      lysergicaciddiethylamide mda mdma meclizine meperidine
      meprobamate metaxalone methadone methamphetamine methanol
      methocarbamol methotrexate methylphenidate midazolam
      mirtazapine molidone morphine naproxen nefazodone nicotine
      nitricacid norchlorcyclizine nordiazepam nordoxepin
      norfluoxetine normeperidine norpropoxyphene nortriptylene
      olanzapine orphenadrine oxazepam oxycodone oxymorphone
      pancuroniumbromide paraquat paroxetine pentazocine
      pentobarbital pentoxifylline phencyclidine phenobarbital
      phentermine phenyltoloxamine phenytoin primidone promethazine
      propane propoxyphene propranolol propyleneglycol quetiapine
      quinine ritalinicacid salicylate salicylicacid secobarbital
      sertraline temazepam tetrafluroethane tetrahydrocannabinol
      thccarboxylicacid theophylline thiopental thioridazine
      toluene tramadol tranylcypromine trazodone triazolam
      tricyclic trifluoperazine trihexyphenidyl trimipramine
      valproic venlafaxine verapamil zolpidem;

array amounts {139} acetaminophen_c acetone_c acetylmorphine_c
```

```
        alprazolam_c amitriptyline_c amoxapine_c amphetamine_c
        benzoylecgonine_c benztropine_c brompheniramine_c bupropion_c
        busiprone_c butabarbital_c butalbital_c caffeine_c
        carbamazepine_c carbondioxide_c carbonmonoxide_c
        carisoprodol_c chlordiazepoxide_c chlorodifluoromethane_c
        chlorpheniramine_c chlorpromazine_c citalopram_c
        clomipramine_c clonzepam_c clozapine_c cocaethlyene_c
        cocaine_c codeine_c cyanide_c cyclobenzaprine_c
        desalkyflurazepam_c desipramine_c dextromethorphan_c
        diazepam_c difluoroethane_c dihydrocodeine_c diltiazem_c
        diphenhydramine_c doxepin_c doxylamine_c ephedrine_c
        ethanol_c ethyleneglycol_c fenfluramine_c fentanyl_c
        fluoxetine_c fluvoxamine_c gabapentin_c
        gammahydroxybutyrate_c guaifenesin_c hydrocodone_c
        hydromorphone_c hydroxyzine_c ibuprofen_c imipramine_c
        insulin_c isopropanol_c ketamine_c lithium_c lorazepam_c
        lysergicaciddiethylamide_c mda_c mdma_c meclizine_c
        meperidine_c meprobamate_c metaxalone_c methadone_c
        methamphetamine_c methanol_c methocarbamol_c methotrexate_c
        methylphenidate_c midazolam_c mirtazapine_c molidone_c
        morphine_c naproxen_c nefazodone_c nicotine_c nitricacid_c
        norchlorcyclizine_c nordiazepam_c nordoxepin_c
        norfluoxetine_c normeperidine_c norpropoxyphene_c
        nortriptylene_c olanzapine_c orphenadrine_c oxazepam_c
        oxycodone_c oxymorphone_c pancuroniumbromide_c paraquat_c
        paroxetine_c pentazocine_c pentobarbital_c pentoxifylline_c
        phencyclidine_c phenobarbital_c phentermine_c
        phenyltoloxamine_c phenytoin_c primidone_c promethazine_c
        propane_c propoxyphene_c propranolol_c propyleneglycol_c
        quetiapine_c quinine_c ritalinicacid_c salicylate_c
        salicylicacid_c secobarbital_c sertraline_c temazepam_c
        tetrafluroethane_c tetrahydrocannabinol_c thccarboxylicacid_c
        theophylline_c thiopental_c thioridazine_c toluene_c
        tramadol_c tranylcypromine_c trazodone_c triazolam_c
        tricyclic_c trifluoperazine_c trihexyphenidyl_c
        trimipramine_c valproic_c venlafaxine_c verapamil_c
        zolpidem_c;

array levels {139} $9. acetaminophen_1 acetone_1 acetylmorphine_1
        alprazolam_1 amitriptyline_1 amoxapine_1 amphetamine_1
        benzoylecgonine_1 benztropine_1 brompheniramine_1 bupropion_1
        busiprone_1 butabarbital_1 butalbital_1 caffeine_1
        carbamazepine_1 carbondioxide_1 carbonmonoxide_1
        carisoprodol_1 chlordiazepoxide_1 chlorodifluoromethane_1
        chlorpheniramine_1 chlorpromazine_1 citalopram_1
        clomipramine_1 clonzepam_1 clozapine_1 cocaethlyene_1
        cocaine_1 codeine_1 cyanide_1 cyclobenzaprine_1
        desalkyflurazepam_1 desipramine_1 dextromethorphan_1
        diazepam_1 difluoroethane_1 dihydrocodeine_1 diltiazem_1
        diphenhydramine_1 doxepin_1 doxylamine_1 ephedrine_1
        ethanol_1 ethyleneglycol_1 fenfluramine_1 fentanyl_1
        fluoxetine_1 fluvoxamine_1 gabapentin_1
```

```
        gammahydroxybutyrate_1 guaifenesin_1 hydrocodone_1
        hydromorphone_1 hydroxyzine_1 ibuprofen_1 imipramine_1
        insulin_1 isopropanol_1 ketamine_1 lithium_1 lorazepam_1
        lysergicaciddiethylamide_1 mda_1 mdma_1 meclizine_1
        meperidine_1 meprobamate_1 metaxalone_1 methadone_1
        methamphetamine_1 methanol_1 methocarbamol_1 methotrexate_1
        methylphenidate_1 midazolam_1 mirtazapine_1 molidone_1
        morphine_1 naproxen_1 nefazodone_1 nicotine_1 nitricacid_1
        norchlorcyclizine_1 nordiazepam_1 nordoxepin_1
        norfluoxetine_1 normeperidine_1 norpropoxyphene_1
        nortriptylene_1 olanzapine_1 orphenadrine_1 oxazepam_1
        oxycodone_1 oxymorphone_1 pancuroniumbromide_1 paraquat_1
        paroxetine_1 pentazocine_1 pentobarbital_1 pentoxifylline_1
        phencyclidine_1 phenobarbital_1 phentermine_1
        phenyltoloxamine_1 phenytoin_1 primidone_1 promethazine_1
        propane_1 propoxyphene_1 propranolol_1 propyleneglycol_1
        quetiapine_1 quinine_1 ritalinicacid_1 salicylate_1
        salicylicacid_1 secobarbital_1 sertraline_1 temazepam_1
        tetrafluroethane_1 tetrahydrocannabinol_1 thccarboxylicacid_1
        theophylline_1 thiopental_1 thioridazine_1 toluene_1
        tramadol_1 tranylcypromine_1 trazodone_1 triazolam_1
        tricyclic_1 trifluoperazine_1 trihexyphenidyl_1
        trimipramine_1 valproic_1 venlafaxine_1 verapamil_1
        zolpidem_1;

array press {139} $31. acetaminophen_p acetone_p acetylmorphine_p
        alprazolam_p amitriptyline_p amoxapine_p amphetamine_p
        benzoylecgonine_p benztropine_p brompheniramine_p bupropion_p
        busiprone_p butabarbital_p butalbital_p caffeine_p
        carbamazepine_p carbondioxide_p carbonmonoxide_p
        carisoprodol_p chlordiazepoxide_p chlorodifluoromethane_p
        chlorpheniramine_p chlorpromazine_p citalopram_p
        clomipramine_p clonzepam_p clozapine_p cocaethlyene_p
        cocaine_p codeine_p cyanide_p cyclobenzaprine_p
        desalkyflurazepam_p desipramine_p dextromethorphan_p
        diazepam_p difluoroethane_p dihydrocodeine_p
        diltiazem_p diphenhydramine_p doxepin_p doxylamine_p
        ephedrine_p ethanol_p ethyleneglycol_p fenfluramine_p
        fentanyl_p fluoxetine_p fluvoxamine_p gabapentin_p
        gammahydroxybutyrate_p guaifenesin_p hydrocodone_p
        hydromorphone_p hydroxyzine_p ibuprofen_p imipramine_p
        insulin_p isopropanol_p ketamine_p lithium_p lorazepam_p
        lysergicaciddiethylamide_p mda_p mdma_p meclizine_p
        meperidine_p meprobamate_p metaxalone_p methadone_p
        methamphetamine_p methanol_p methocarbamol_p methotrexate_p
        methylphenidate_p midazolam_p mirtazapine_p molidone_p
        morphine_p naproxen_p nefazodone_p nicotine_p nitricacid_p
        norchlorcyclizine_p nordiazepam_p nordoxepin_p
        norfluoxetine_p normeperidine_p norpropoxyphene_p
        nortriptylene_p olanzapine_p orphenadrine_p oxazepam_p
        oxycodone_p oxymorphone_p pancuroniumbromide_p paraquat_p
        paroxetine_p pentazocine_p pentobarbital_p pentoxifylline_p
        phencyclidine_p phenobarbital_p phentermine_p
```

```
            phenyltoloxamine_p phenytoin_p primidone_p promethazine_p
            propane_p propoxyphene_p propranolol_p propyleneglycol_p
            quetiapine_p quinine_p ritalinicacid_p salicylate_p
            salicylicacid_p secobarbital_p sertraline_p temazepam_p
            tetrafluroethane_p tetrahydrocannabinol_p thccarboxylicacid_p
            theophylline_p thiopental_p thioridazine_p toluene_p
            tramadol_p tranylcypromine_p trazodone_p triazolam_p
            tricyclic_p trifluoperazine_p trihexyphenidyl_p
            trimipramine_p valproic_p venlafaxine_p verapamil_p
            zolpidem_p;

            if first.case_number then do;
                  do i = 1 to 139 by 1;
                        drugs{i} = .;
                        amounts{i} = .;
                        levels{i} = '';
                        press{i} = '';
*number_of_drugs keeps track of the number of drugs detected (not just
tested for);
                        number_of_drugs = 0;
                  end;
            end;

            do i =1 to 139 by 1;
                  if drug=upcase(names{i}) then do;
                        drugs{i} = 1;
                        press{i} = presence;
                        levels{i} = unit;

                        index1 = find(drug_level, 'M', 'i');
                        index2 = find(drug_level, 'P', 'i');
                        index3 = find(drug_level, '/', 'i');

                        if index1 ne 0 then do;
                           amounts{i}=input(substrn(
                              drug_level,1,index1-1),4.);
                        end;

                        else if index2 ne 0 then do;
                           amounts{i} = .;
                        end;

                        else if index3 ne 0 then do;
                           amounts{i}=input(substrn(
                              drug_level,1,index3-1),4.);
                        end;

                        else do;
                           amounts{i} = input(drug_level, 8.);
                        end;

                        if amounts{i} ne 0 & press{i} ne 'not detected'
                              then number_of_drugs = number_of_drugs+1;
```

```
      end;

end;

if last.case_number then output;

retain case_number acetaminophen acetone acetylmorphine
alprazolam amitriptyline amoxapine amphetamine
benzoylecgonine benztropine brompheniramine bupropion
busiprone butabarbital butalbital caffeine carbamazepine
carbondioxide carbonmonoxide carisoprodol chlordiazepoxide
chlorodifluoromethane chlorpheniramine chlorpromazine
citalopram clomipramine clonzepam clozapine cocaethlyene
cocaine codeine cyanide cyclobenzaprine desalkyflurazepam
desipramine dextromethorphan diazepam difluoroethane
dihydrocodeine diltiazem diphenhydramine doxepin doxylamine
ephedrine ethanol ethyleneglycol fenfluramine fentanyl
fluoxetine fluvoxamine gabapentin gammahydroxybutyrate
guaifenesin hydrocodone hydromorphone hydroxyzine ibuprofen
imipramine insulin isopropanol ketamine lithium lorazepam
lysergicaciddiethylamide mda mdma meclizine meperidine
meprobamate metaxalone methadone methamphetamine methanol
methocarbamol methotrexate methylphenidate midazolam
mirtazapine molidone morphine naproxen nefazodone nicotine
nitricacid norchlorcyclizine nordiazepam nordoxepin
norfluoxetine normeperidine norpropoxyphene nortriptylene
olanzapine orphenadrine oxazepam oxycodone oxymorphone
pancuroniumbromide paraquat paroxetine pentazocine
pentobarbital pentoxifylline phencyclidine phenobarbital
phentermine phenyltoloxamine phenytoin primidone promethazine
propane propoxyphene propranolol propyleneglycol quetiapine
quinine ritalinicacid salicylate salicylicacid secobarbital
sertraline temazepam tetrafluroethane tetrahydrocannabinol
thccarboxylicacid theophylline thiopental thioridazine
toluene tramadol tranylcypromine trazodone triazolam
tricyclic trifluoperazine trihexyphenidyl trimipramine
valproic venlafaxine verapamil zolpidem acetaminophen_c
acetone_c acetylmorphine_c alprazolam_c amitriptyline_c
amoxapine_c amphetamine_c benzoylecgonine_c benztropine_c
brompheniramine_c bupropion_c busiprone_c butabarbital_c
butalbital_c caffeine_c carbamazepine_c carbondioxide_c
carbonmonoxide_c carisoprodol_c chlordiazepoxide_c
chlorodifluoromethane_c chlorpheniramine_c chlorpromazine_c
citalopram_c clomipramine_c clonzepam_c clozapine_c
cocaethlyene_c cocaine_c codeine_c cyanide_c
cyclobenzaprine_c desalkyflurazepam_c desipramine_c
dextromethorphan_c diazepam_c difluoroethane_c
dihydrocodeine_c diltiazem_c diphenhydramine_c doxepin_c
doxylamine_c ephedrine_c ethanol_c ethyleneglycol_c
fenfluramine_c fentanyl_c fluoxetine_c fluvoxamine_c
gabapentin_c gammahydroxybutyrate_c guaifenesin_c
hydrocodone_c hydromorphone_c hydroxyzine_c ibuprofen_c
imipramine_c insulin_c isopropanol_c ketamine_c lithium_c
```

lorazepam_c lysergicaciddiethylamide_c mda_c mdma_c
meclizine_c meperidine_c meprobamate_c metaxalone_c
methadone_c methamphetamine_c methanol_c methocarbamol_c
methotrexate_c methylphenidate_c midazolam_c mirtazapine_c
molidone_c morphine_c naproxen_c nefazodone_c nicotine_c
nitricacid_c norchlorcyclizine_c nordiazepam_c nordoxepin_c
norfluoxetine_c normeperidine_c norpropoxyphene_c
nortriptylene_c olanzapine_c orphenadrine_c oxazepam_c
oxycodone_c oxymorphone_c pancuroniumbromide_c paraquat_c
paroxetine_c pentazocine_c pentobarbital_c pentoxifylline_c
phencyclidine_c phenobarbital_c phentermine_c
phenyltoloxamine_c phenytoin_c primidone_c promethazine_c
propane_c propoxyphene_c propranolol_c propyleneglycol_c
quetiapine_c quinine_c ritalinicacid_c salicylate_c
salicylicacid_c secobarbital_c sertraline_c temazepam_c
tetrafluroethane_c tetrahydrocannabinol_c thccarboxylicacid_c
theophylline_c thiopental_c thioridazine_c toluene_c
tramadol_c tranylcypromine_c trazodone_c triazolam_c
tricyclic_c trifluoperazine_c trihexyphenidyl_c
trimipramine_c valproic_c venlafaxine_c verapamil_c
zolpidem_c acetaminophen_p acetone_p acetylmorphine_p
alprazolam_p amitriptyline_p amoxapine_p amphetamine_p
benzoylecgonine_p benztropine_p brompheniramine_p bupropion_p
busiprone_p butabarbital_p butalbital_p caffeine_p
carbamazepine_p carbondioxide_p carbonmonoxide_p
carisoprodol_p chlordiazepoxide_p chlorodifluoromethane_p
chlorpheniramine_p chlorpromazine_p citalopram_p
clomipramine_p clonzepam_p clozapine_p cocaethlyene_p
cocaine_p codeine_p cyanide_p cyclobenzaprine_p
desalkyflurazepam_p desipramine_p dextromethorphan_p
diazepam_p difluoroethane_p dihydrocodeine_p diltiazem_p
diphenhydramine_p doxepin_p doxylamine_p ephedrine_p
ethanol_p ethyleneglycol_p fenfluramine_p fentanyl_p
fluoxetine_p fluvoxamine_p gabapentin_p
gammahydroxybutyrate_p guaifenesin_p hydrocodone_p
hydromorphone_p hydroxyzine_p ibuprofen_p imipramine_p
insulin_p isopropanol_p ketamine_p lithium_p lorazepam_p
lysergicaciddiethylamide_p mda_p mdma_p meclizine_p
meperidine_p meprobamate_p metaxalone_p methadone_p
methamphetamine_p methanol_p methocarbamol_p methotrexate_p
methylphenidate_p midazolam_p mirtazapine_p molidone_p
morphine_p naproxen_p nefazodone_p nicotine_p nitricacid_p
norchlorcyclizine_p nordiazepam_p nordoxepin_p
norfluoxetine_p normeperidine_p norpropoxyphene_p
nortriptylene_p olanzapine_p orphenadrine_p oxazepam_p
oxycodone_p oxymorphone_p pancuroniumbromide_p paraquat_p
paroxetine_p pentazocine_p pentobarbital_p pentoxifylline_p
phencyclidine_p phenobarbital_p phentermine_p
phenyltoloxamine_p phenytoin_p primidone_p promethazine_p
propane_p propoxyphene_p propranolol_p propyleneglycol_p
quetiapine_p quinine_p ritalinicacid_p salicylate_p
salicylicacid_p secobarbital_p sertraline_p temazepam_p
tetrafluroethane_p tetrahydrocannabinol_p thccarboxylicacid_p

```
theophylline_p thiopental_p thioridazine_p toluene_p
tramadol_p tranylcypromine_p trazodone_p triazolam_p
tricyclic_p trifluoperazine_p trihexyphenidyl_p
trimipramine_p valproic_p venlafaxine_p verapamil_p
zolpidem_p acetaminophen_l acetone_l acetylmorphine_l
alprazolam_l amitriptyline_l amoxapine_l amphetamine_l
benzoylecgonine_l benztropine_l brompheniramine_l
bupropion_l busiprone_l butabarbital_l butalbital_l
caffeine_l carbamazepine_l carbondioxide_l carbonmonoxide_l
carisoprodol_l chlordiazepoxide_l chlorodifluoromethane_l
chlorpheniramine_l chlorpromazine_l citalopram_l
clomipramine_l clonzepam_l clozapine_l cocaethlyene_l
cocaine_l codeine_l cyanide_l cyclobenzaprine_l
desalkyflurazepam_l desipramine_l dextromethorphan_l
diazepam_l difluoroethane_l dihydrocodeine_l
diltiazem_l diphenhydramine_l doxepin_l doxylamine_l
ephedrine_l ethanol_l ethyleneglycol_l fenfluramine_l
fentanyl_l fluoxetine_l fluvoxamine_l gabapentin_l
gammahydroxybutyrate_l guaifenesin_l hydrocodone_l
hydromorphone_l hydroxyzine_l ibuprofen_l imipramine_l
insulin_l isopropanol_l ketamine_l lithium_l lorazepam_l
lysergicaciddiethylamide_l mda_l mdma_l meclizine_l
meperidine_l meprobamate_l metaxalone_l methadone_l
methamphetamine_l methanol_l methocarbamol_l methotrexate_l
methylphenidate_l midazolam_l mirtazapine_l molidone_l
morphine_l naproxen_l nefazodone_l nicotine_l nitricacid_l
norchlorcyclizine_l nordiazepam_l nordoxepin_l
norfluoxetine_l normeperidine_l norpropoxyphene_l
nortriptylene_l olanzapine_l orphenadrine_l oxazepam_l
oxycodone_l oxymorphone_l pancuroniumbromide_l paraquat_l
paroxetine_l pentazocine_l pentobarbital_l pentoxifylline_l
phencyclidine_l phenobarbital_l phentermine_l
phenyltoloxamine_l phenytoin_l primidone_l promethazine_l
propane_l propoxyphene_l propranolol_l propyleneglycol_l
quetiapine_l quinine_l ritalinicacid_l salicylate_l
salicylicacid_l secobarbital_l sertraline_l temazepam_l
tetrafluroethane_l tetrahydrocannabinol_l thccarboxylicacid_l
theophylline_l thiopental_l thioridazine_l toluene_l
tramadol_l tranylcypromine_l trazodone_l triazolam_l
tricyclic_l trifluoperazine_l trihexyphenidyl_l
trimipramine_l valproic_l venlafaxine_l verapamil_l
zolpidem_l number_of_drugs;

drop acetaminophen_n acetone_n acetylmorphine_n alprazolam_n
amitriptyline_n amoxapine_n amphetamine_n benzoylecgonine_n
benztropine_n brompheniramine_n bupropion_n busiprone_n
butabarbital_n butalbital_n caffeine_n carbamazepine_n
carbondioxide_n carbonmonoxide_n carisoprodol_n
chlordiazepoxide_n chlorodifluoromethane_n chlorpheniramine_n
chlorpromazine_n citalopram_n clomipramine_n clonzepam_n
clozapine_n cocaethlyene_n cocaine_n codeine_n cyanide_n
cyclobenzaprine_n desalkyflurazepam_n desipramine_n
dextromethorphan_n diazepam_n difluoroethane_n
```

```
            dihydrocodeine_n diltiazem_n diphenhydramine_n doxepin_n
            doxylamine_n ephedrine_n ethanol_n ethyleneglycol_n
            fenfluramine_n fentanyl_n fluoxetine_n fluvoxamine_n
            gabapentin_n gammahydroxybutyrate_n guaifenesin_n
            hydrocodone_n hydromorphone_n hydroxyzine_n ibuprofen_n
            imipramine_n insulin_n isopropanol_n ketamine_n lithium_n
            lorazepam_n lysergicaciddiethylamide_n mda_n mdma_n
            meclizine_n meperidine_n meprobamate_n metaxalone_n
            methadone_n methamphetamine_n methanol_n methocarbamol_n
            methotrexate_n methylphenidate_n midazolam_n mirtazapine_n
            molidone_n morphine_n naproxen_n nefazodone_n nicotine_n
            nitricacid_n norchlorcyclizine_n nordiazepam_n nordoxepin_n
            norfluoxetine_n normeperidine_n norpropoxyphene_n
            nortriptylene_n olanzapine_n orphenadrine_n oxazepam_n
            oxycodone_n oxymorphone_n pancuroniumbromide_n paraquat_n
            paroxetine_n pentazocine_n pentobarbital_n pentoxifylline_n
            phencyclidine_n phenobarbital_n phentermine_n
            phenyltoloxamine_n phenytoin_n primidone_n promethazine_n
            propane_n propoxyphene_n propranolol_n propyleneglycol_n
            quetiapine_n quinine_n ritalinicacid_n salicylate_n
            salicylicacid_n secobarbital_n sertraline_n temazepam_n
            tetrafluroethane_n tetrahydrocannabinol_n thccarboxylicacid_n
            theophylline_n thiopental_n thioridazine_n toluene_n
            tramadol_n tranylcypromine_n trazodone_n triazolam_n
            tricyclic_n trifluoperazine_n trihexyphenidyl_n
            trimipramine_n valproic_n venlafaxine_n verapamil_n
            zolpidem_n i;
run;

data thesis.mefinal1;
      set thesis.mefinal;
      if upcase(tox_) ne 'N';
run;

proc freq data=thesis.mefinal1;
      tables fatalnew;
run;

/*Checking a drug to see if the above code worked*/
proc print data=thesis.mefinal1;
      var case_number doxepin doxepin_c doxepin_l doxepin_p;
run;

*DUI DATASETS;
PROC IMPORT OUT= thesis.DUIDrug1
            DATAFILE= "E:\Thesis Files\DUIoriginalDrug.jmp"
            DBMS=JMP REPLACE;
RUN;

PROC IMPORT OUT= thesis.DUIDescrip1
            DATAFILE= "E:\Thesis Files\DUIoriginalDescrip.jmp"
            DBMS=JMP REPLACE;
RUN;
```

```
/*Looking for duplicates in descriptive information -
keeping the most recent observation*/
data thesis.DUIDescrip2;
      set thesis.DUIDescrip1;
      by fslabnum;
      *last = last.fslabnum;
      if last.fslabnum then output;
run;

/*Sort in order to merge two datasets*/
proc sort data=thesis.DUIDrug1;
      by FSLabNum;
run;
proc sort data=thesis.DUIDescrip2;
      by FSLabNum datesubmitted;
run;

/*Merge the descriptive and the drug sets into one*/
data thesis.DUI1;
      merge thesis.DUIDescrip2 thesis.DUIDrug1;
      by FSLabNum;
run;

/*Calculating age, using current date as the last possible
date if no other date is available to calculate age*/
data thesis.DUIclean1;
      set thesis.DUI1;
      attrib currentdate format=MMDDYY10.;
      currentdate = 01/01/2005;
      if dateofbirth ne . & datesubmitted ne . then
            age_all = yrdif(dateofbirth, datesubmitted, 'ACT/ACT');
      else if dateofbirth ne . & date_case_completed ne . then
            age_all = yrdif(dateofbirth, date_case_completed,
                  'ACT/ACT');
      else if dateofbirth ne . then age_all = yrdif(dateofbirth,
            currentdate, 'ACT/ACT');
      else age_all = .;
      age = floor(age_all);

      drop age_all dateofbirth datesubmitted submissionnum
            date_case_completed;

run;

data thesis.DUIclean2;
      set thesis.DUIclean1;
      *Fixing drug names to simplify;
      if drug = '1,1-difluoroethane' then drug = 'Difluoroethane';
      if drug = '6-Acetyl Morphine' then drug = 'AcetylMorphine';
      if drug = 'Blood Alcohol' then drug = 'Ethanol';
      if drug = 'Gamma-Hydroxybutyrate/Lactone' then drug =
            'Gamma-Hydroxybutyrate';
```

```
      if drug = 'N-Desalkyl Flurazepam' then drug =
            'Desalkyflurazepam';

run;

data thesis.DUIclean4;
      set thesis.DUIclean2;

      *Create new database variable;
      database = 'dui';
      if database = 'dui' then bac = 'BLOOD';

      *Format variables to match other dataset;
      drug = upcase(drug);
      attrib case_number format=$10.;
      case_number = fslabnum;

      *Create units of measurement - all in blood;
      if drug = 'ETHANOL' then unit = 'Grams%';
      else unit = 'mg/L';

      *Set illegal ages to missing;
      if age < 14 then age = .;

      drop fslabnum;
run;

proc freq data=thesis.DUIclean4;
      tables presence;
run;

data thesis.DUIclean5;
      set thesis.DUIclean4;

      *Edit the concentrations to reflect a more accurate amount
      based on the presence variable;
      if presence = 'less than' | presence = 'present less than' then
            amount = amount*.5;
      else if presence = 'greater than' then amount = amount*1.15;
      else if presence = 'present' | presence = 'quantitated' |
            presence = 'not detected' | presence = ' ' then
amount=amount;
      else amount = .;

run;

proc sort data=thesis.DUIclean5;
      by case_number drug descending bac;
run;


data thesis.DUIfinal;
      set thesis.DUIclean5;
```

```
by case_number drug descending bac;
array names {139} $100. acetaminophen_n acetone_n acetylmorphine_n
      alprazolam_n amitriptyline_n amoxapine_n amphetamine_n
      benzoylecgonine_n benztropine_n brompheniramine_n bupropion_n
      busiprone_n butabarbital_n butalbital_n caffeine_n
      carbamazepine_n carbondioxide_n carbonmonoxide_n
      carisoprodol_n chlordiazepoxide_n chlorodifluoromethane_n
      chlorpheniramine_n chlorpromazine_n citalopram_n
      clomipramine_n clonzepam_n clozapine_n cocaethlyene_n
      cocaine_n codeine_n cyanide_n cyclobenzaprine_n
      desalkyflurazepam_n desipramine_n dextromethorphan_n
      diazepam_n difluoroethane_n dihydrocodeine_n diltiazem_n
      diphenhydramine_n doxepin_n doxylamine_n ephedrine_n
      ethanol_n ethyleneglycol_n fenfluramine_n fentanyl_n
      fluoxetine_n fluvoxamine_n gabapentin_n
      gammahydroxybutyrate_n guaifenesin_n hydrocodone_n
      hydromorphone_n hydroxyzine_n ibuprofen_n imipramine_n
      insulin_n isopropanol_n ketamine_n lithium_n lorazepam_n
      lysergicaciddiethylamide_n mda_n mdma_n meclizine_n
      meperidine_n meprobamate_n metaxalone_n methadone_n
      methamphetamine_n methanol_n methocarbamol_n methotrexate_n
      methylphenidate_n midazolam_n mirtazapine_n molidone_n
      morphine_n naproxen_n nefazodone_n nicotine_n nitricacid_n
      norchlorcyclizine_n nordiazepam_n nordoxepin_n
      norfluoxetine_n normeperidine_n norpropoxyphene_n
      nortriptylene_n olanzapine_n orphenadrine_n oxazepam_n
      oxycodone_n oxymorphone_n pancuroniumbromide_n paraquat_n
      paroxetine_n pentazocine_n pentobarbital_n pentoxifylline_n
      phencyclidine_n phenobarbital_n phentermine_n
      phenyltoloxamine_n phenytoin_n primidone_n promethazine_n
      propane_n propoxyphene_n propranolol_n propyleneglycol_n
      quetiapine_n quinine_n ritalinicacid_n salicylate_n
      salicylicacid_n secobarbital_n sertraline_n temazepam_n
      tetrafluroethane_n tetrahydrocannabinol_n thccarboxylicacid_n
      theophylline_n thiopental_n thioridazine_n toluene_n
      tramadol_n tranylcypromine_n trazodone_n triazolam_n
      tricyclic_n trifluoperazine_n trihexyphenidyl_n
      trimipramine_n valproic_n venlafaxine_n verapamil_n
      zolpidem_n ('acetaminophen', 'acetone', 'acetylmorphine',
      'alprazolam', 'amitriptyline', 'amoxapine', 'amphetamine',
      'benzoylecgonine', 'benztropine', 'brompheniramine',
      'bupropion', 'busiprone', 'butabarbital', 'butalbital',
      'caffeine', 'carbamazepine', 'carbon dioxide', 'carbon
      monoxide', 'carisoprodol', 'chlordiazepoxide',
      'chlorodifluoromethane', 'chlorpheniramine',
      'chlorpromazine', 'citalopram', 'clomipramine', 'clonzepam',
      'clozapine', 'cocaethlyene', 'cocaine', 'codeine', 'cyanide',
      'cyclobenzaprine', 'desalkyflurazepam', 'desipramine',
      'dextromethorphan', 'diazepam', 'difluoroethane',
      'dihydrocodeine', 'diltiazem', 'diphenhydramine', 'doxepin',
      'doxylamine', 'ephedrine', 'ethanol', 'ethyleneglycol',
      'fenfluramine', 'fentanyl', 'fluoxetine', 'fluvoxamine',
      'gabapentin', 'gammahydroxybutyrate', 'guaifenesin',
```

```
        'hydrocodone', 'hydromorphone', 'hydroxyzine', 'ibuprofen',
        'imipramine', 'insulin', 'isopropanol', 'ketamine',
        'lithium', 'lorazepam', 'lysergicaciddiethylamide', 'mda',
        'mdma', 'meclizine', 'meperidine', 'meprobamate',
        'metaxalone', 'methadone', 'methamphetamine', 'methanol',
        'methocarbamol', 'methotrexate', 'methylphenidate',
        'midazolam', 'mirtazapine', 'molidone', 'morphine',
        'naproxen', 'nefazodone', 'nicotine', 'nitricacid',
        'norchlorcyclizine', 'nordiazepam', 'nordoxepin',
        'norfluoxetine', 'normeperidine', 'norpropoxyphene',
        'nortriptylene', 'olanzapine', 'orphenadrine', 'oxazepam',
        'oxycodone', 'oxymorphone', 'pancuroniumbromide',
        'paraquat', 'paroxetine', 'pentazocine', 'pentobarbital',
        'pentoxifylline', 'phencyclidine', 'phenobarbital',
        'phentermine', 'phenyltoloxamine', 'phenytoin',
        'primidone', 'promethazine', 'propane', 'propoxyphene',
        'propranolol', 'propyleneglycol', 'quetiapine', 'quinine',
        'ritalinicacid', 'salicylate', 'salicylicacid',
        'secobarbital', 'sertraline', 'temazepam',
        'tetrafluroethane', 'tetrahydrocannabinol',
        'thccarboxylicacid', 'theophylline', 'thiopental',
        'thioridazine', 'toluene', 'tramadol', 'tranylcypromine',
        'trazodone', 'triazolam', 'tricyclic', 'trifluoperazine',
        'trihexyphenidyl', 'trimipramine', 'valproic', 'venlafaxine',
        'verapamil', 'zolpidem');

array drugs {139} acetaminophen acetone acetylmorphine alprazolam
      amitriptyline amoxapine amphetamine benzoylecgonine
      benztropine brompheniramine bupropion busiprone butabarbital
      butalbital caffeine carbamazepine carbondioxide
      carbonmonoxide carisoprodol chlordiazepoxide
      chlorodifluoromethane chlorpheniramine chlorpromazine
      citalopram clomipramine clonzepam clozapine cocaethlyene
      cocaine codeine cyanide cyclobenzaprine desalkyflurazepam
      desipramine dextromethorphan diazepam difluoroethane
      dihydrocodeine diltiazem diphenhydramine doxepin doxylamine
      ephedrine ethanol ethyleneglycol fenfluramine fentanyl
      fluoxetine fluvoxamine gabapentin gammahydroxybutyrate
      guaifenesin hydrocodone hydromorphone hydroxyzine ibuprofen
      imipramine insulin isopropanol ketamine lithium lorazepam
      lysergicaciddiethylamide mda mdma meclizine meperidine
      meprobamate metaxalone methadone methamphetamine methanol
      methocarbamol methotrexate methylphenidate midazolam
      mirtazapine molidone morphine naproxen nefazodone nicotine
      nitricacid norchlorcyclizine nordiazepam nordoxepin
      norfluoxetine normeperidine norpropoxyphene nortriptylene
      olanzapine orphenadrine oxazepam oxycodone oxymorphone
      pancuroniumbromide paraquat paroxetine pentazocine
      pentobarbital pentoxifylline phencyclidine phenobarbital
      phentermine phenyltoloxamine phenytoin primidone promethazine
      propane propoxyphene propranolol propyleneglycol quetiapine
      quinine ritalinicacid salicylate salicylicacid secobarbital
      sertraline temazepam tetrafluroethane tetrahydrocannabinol
```

```
      thccarboxylicacid theophylline thiopental thioridazine
      toluene tramadol tranylcypromine trazodone triazolam
      tricyclic trifluoperazine trihexyphenidyl trimipramine
      valproic venlafaxine verapamil zolpidem;

array amounts {139} acetaminophen_c acetone_c acetylmorphine_c
      alprazolam_c amitriptyline_c amoxapine_c amphetamine_c
      benzoylecgonine_c benztropine_c brompheniramine_c bupropion_c
      busiprone_c butabarbital_c butalbital_c caffeine_c
      carbamazepine_c carbondioxide_c carbonmonoxide_c
      carisoprodol_c chlordiazepoxide_c chlorodifluoromethane_c
      chlorpheniramine_c chlorpromazine_c citalopram_c
      clomipramine_c clonzepam_c clozapine_c cocaethlyene_c
      cocaine_c codeine_c cyanide_c cyclobenzaprine_c
      desalkyflurazepam_c desipramine_c dextromethorphan_c
      diazepam_c difluoroethane_c dihydrocodeine_c diltiazem_c
      diphenhydramine_c doxepin_c doxylamine_c ephedrine_c
      ethanol_c ethyleneglycol_c fenfluramine_c fentanyl_c
      fluoxetine_c fluvoxamine_c gabapentin_c
      gammahydroxybutyrate_c guaifenesin_c hydrocodone_c
      hydromorphone_c hydroxyzine_c ibuprofen_c imipramine_c
      insulin_c isopropanol_c ketamine_c lithium_c lorazepam_c
      lysergicaciddiethylamide_c mda_c mdma_c meclizine_c
      meperidine_c meprobamate_c metaxalone_c methadone_c
      methamphetamine_c methanol_c methocarbamol_c methotrexate_c
      methylphenidate_c midazolam_c mirtazapine_c molidone_c
      morphine_c naproxen_c nefazodone_c nicotine_c nitricacid_c
      norchlorcyclizine_c nordiazepam_c nordoxepin_c
      norfluoxetine_c normeperidine_c norpropoxyphene_c
      nortriptylene_c olanzapine_c orphenadrine_c oxazepam_c
      oxycodone_c oxymorphone_c pancuroniumbromide_c paraquat_c
      paroxetine_c pentazocine_c pentobarbital_c pentoxifylline_c
      phencyclidine_c phenobarbital_c phentermine_c
      phenyltoloxamine_c phenytoin_c primidone_c promethazine_c
      propane_c propoxyphene_c propranolol_c propyleneglycol_c
      quetiapine_c quinine_c ritalinicacid_c salicylate_c
      salicylicacid_c secobarbital_c sertraline_c temazepam_c
      tetrafluroethane_c tetrahydrocannabinol_c thccarboxylicacid_c
      theophylline_c thiopental_c thioridazine_c toluene_c
      tramadol_c tranylcypromine_c trazodone_c triazolam_c
      tricyclic_c trifluoperazine_c trihexyphenidyl_c
      trimipramine_c valproic_c venlafaxine_c verapamil_c
      zolpidem_c;

array levels {139} $14. acetaminophen_1 acetone_1 acetylmorphine_1
      alprazolam_1 amitriptyline_1 amoxapine_1 amphetamine_1
      benzoylecgonine_1 benztropine_1 brompheniramine_1 bupropion_1
      busiprone_1 butabarbital_1 butalbital_1 caffeine_1
      carbamazepine_1 carbondioxide_1 carbonmonoxide_1
      carisoprodol_1 chlordiazepoxide_1 chlorodifluoromethane_1
      chlorpheniramine_1 chlorpromazine_1 citalopram_1
      clomipramine_1 clonzepam_1 clozapine_1 cocaethlyene_1
```

```
        cocaine_1 codeine_1 cyanide_1 cyclobenzaprine_1
        desalkyflurazepam_1 desipramine_1 dextromethorphan_1
        diazepam_1 difluoroethane_1 dihydrocodeine_1 diltiazem_1
        diphenhydramine_1 doxepin_1 doxylamine_1 ephedrine_1
        ethanol_1 ethyleneglycol_1 fenfluramine_1 fentanyl_1
        fluoxetine_1 fluvoxamine_1 gabapentin_1
        gammahydroxybutyrate_1 guaifenesin_1 hydrocodone_1
        hydromorphone_1 hydroxyzine_1 ibuprofen_1 imipramine_1
        insulin_1 isopropanol_1 ketamine_1 lithium_1 lorazepam_1
        lysergicaciddiethylamide_1 mda_1 mdma_1 meclizine_1
        meperidine_1 meprobamate_1 metaxalone_1 methadone_1
        methamphetamine_1 methanol_1 methocarbamol_1 methotrexate_1
        methylphenidate_1 midazolam_1 mirtazapine_1 molidone_1
        morphine_1 naproxen_1 nefazodone_1 nicotine_1 nitricacid_1
        norchlorcyclizine_1 nordiazepam_1 nordoxepin_1
        norfluoxetine_1 normeperidine_1 norpropoxyphene_1
        nortriptylene_1 olanzapine_1 orphenadrine_1 oxazepam_1
        oxycodone_1 oxymorphone_1 pancuroniumbromide_1 paraquat_1
        paroxetine_1 pentazocine_1 pentobarbital_1 pentoxifylline_1
        phencyclidine_1 phenobarbital_1 phentermine_1
        phenyltoloxamine_1 phenytoin_1 primidone_1 promethazine_1
        propane_1 propoxyphene_1 propranolol_1 propyleneglycol_1
        quetiapine_1 quinine_1 ritalinicacid_1 salicylate_1
        salicylicacid_1 secobarbital_1 sertraline_1 temazepam_1
        tetrafluroethane_1 tetrahydrocannabinol_1 thccarboxylicacid_1
        theophylline_1 thiopental_1 thioridazine_1 toluene_1
        tramadol_1 tranylcypromine_1 trazodone_1 triazolam_1
        tricyclic_1 trifluoperazine_1 trihexyphenidyl_1
        trimipramine_1 valproic_1 venlafaxine_1 verapamil_1
        zolpidem_1;

array press {139} $31. acetaminophen_p acetone_p acetylmorphine_p
        alprazolam_p amitriptyline_p amoxapine_p amphetamine_p
        benzoylecgonine_p benztropine_p brompheniramine_p bupropion_p
        busiprone_p butabarbital_p butalbital_p caffeine_p
        carbamazepine_p carbondioxide_p carbonmonoxide_p
        carisoprodol_p chlordiazepoxide_p chlorodifluoromethane_p
        chlorpheniramine_p chlorpromazine_p citalopram_p
        clomipramine_p clonzepam_p clozapine_p cocaethlyene_p
        cocaine_p codeine_p cyanide_p cyclobenzaprine_p
        desalkyflurazepam_p desipramine_p dextromethorphan_p
        diazepam_p difluoroethane_p dihydrocodeine_p diltiazem_p
        diphenhydramine_p doxepin_p doxylamine_p ephedrine_p
        ethanol_p ethyleneglycol_p fenfluramine_p fentanyl_p
        fluoxetine_p fluvoxamine_p gabapentin_p
        gammahydroxybutyrate_p guaifenesin_p hydrocodone_p
        hydromorphone_p hydroxyzine_p ibuprofen_p imipramine_p
        insulin_p isopropanol_p ketamine_p lithium_p lorazepam_p
        lysergicaciddiethylamide_p mda_p mdma_p meclizine_p
        meperidine_p meprobamate_p metaxalone_p methadone_p
        methamphetamine_p methanol_p methocarbamol_p methotrexate_p
        methylphenidate_p midazolam_p mirtazapine_p molidone_p
        morphine_p naproxen_p nefazodone_p nicotine_p nitricacid_p
```

```
norchlorcyclizine_p nordiazepam_p nordoxepin_p
norfluoxetine_p normeperidine_p norpropoxyphene_p
nortriptylene_p olanzapine_p orphenadrine_p oxazepam_p
oxycodone_p oxymorphone_p pancuroniumbromide_p paraquat_p
paroxetine_p pentazocine_p pentobarbital_p pentoxifylline_p
phencyclidine_p phenobarbital_p phentermine_p
phenyltoloxamine_p phenytoin_p primidone_p promethazine_p
propane_p propoxyphene_p propranolol_p propyleneglycol_p
quetiapine_p quinine_p ritalinicacid_p salicylate_p
salicylicacid_p secobarbital_p sertraline_p temazepam_p
tetrafluroethane_p tetrahydrocannabinol_p thccarboxylicacid_p
theophylline_p thiopental_p thioridazine_p toluene_p
tramadol_p tranylcypromine_p trazodone_p triazolam_p
tricyclic_p trifluoperazine_p trihexyphenidyl_p
trimipramine_p valproic_p venlafaxine_p verapamil_p
zolpidem_p;

if first.case_number then do;
      do i = 1 to 139 by 1;
            drugs{i} = .;
            amounts{i} = .;
            levels{i} = '';
            press{i} = '';
            number_of_drugs = 0;
      end;
end;

do i =1 to 139 by 1;
      if drug=upcase(names{i}) then do;
            drugs{i}=1;
            amounts{i}=amount;
            levels{i} = unit;
            press{i} = presence;

            if amounts{i} ne 0 & press{i} ne 'not detected'
                  then number_of_drugs = number_of_drugs+1;
      end;


end;

if last.case_number then output;

retain case_number acetaminophen acetone acetylmorphine
alprazolam amitriptyline amoxapine amphetamine
benzoylecgonine benztropine brompheniramine bupropion
busiprone butabarbital butalbital caffeine carbamazepine
carbondioxide carbonmonoxide carisoprodol chlordiazepoxide
chlorodifluoromethane chlorpheniramine chlorpromazine
citalopram clomipramine clonzepam clozapine cocaethlyene
cocaine codeine cyanide cyclobenzaprine desalkyflurazepam
desipramine dextromethorphan diazepam difluoroethane
dihydrocodeine diltiazem diphenhydramine doxepin doxylamine
```

ephedrine ethanol ethyleneglycol fenfluramine fentanyl
fluoxetine fluvoxamine gabapentin gammahydroxybutyrate
guaifenesin hydrocodone hydromorphone hydroxyzine ibuprofen
imipramine insulin isopropanol ketamine lithium lorazepam
lysergicaciddiethylamide mda mdma meclizine meperidine
meprobamate metaxalone methadone methamphetamine methanol
methocarbamol methotrexate methylphenidate midazolam
mirtazapine molidone morphine naproxen nefazodone nicotine
nitricacid norchlorcyclizine nordiazepam nordoxepin
norfluoxetine normeperidine norpropoxyphene nortriptylene
olanzapine orphenadrine oxazepam oxycodone oxymorphone
pancuroniumbromide paraquat paroxetine pentazocine
pentobarbital pentoxifylline phencyclidine phenobarbital
phentermine phenyltoloxamine phenytoin primidone promethazine
propane propoxyphene propranolol propyleneglycol quetiapine
quinine ritalinicacid salicylate salicylicacid secobarbital
sertraline temazepam tetrafluroethane tetrahydrocannabinol
thccarboxylicacid theophylline thiopental thioridazine
toluene tramadol tranylcypromine trazodone triazolam
tricyclic trifluoperazine trihexyphenidyl trimipramine
valproic venlafaxine verapamil zolpidem acetaminophen_c
acetone_c acetylmorphine_c alprazolam_c amitriptyline_c
amoxapine_c amphetamine_c benzoylecgonine_c benztropine_c
brompheniramine_c bupropion_c busiprone_c butabarbital_c
butalbital_c caffeine_c carbamazepine_c carbondioxide_c
carbonmonoxide_c carisoprodol_c chlordiazepoxide_c
chlorodifluoromethane_c chlorpheniramine_c chlorpromazine_c
citalopram_c clomipramine_c clonzepam_c clozapine_c
cocaethlyene_c cocaine_c codeine_c cyanide_c
cyclobenzaprine_c desalkylflurazepam_c desipramine_c
dextromethorphan_c diazepam_c difluoroethane_c
dihydrocodeine_c diltiazem_c diphenhydramine_c doxepin_c
doxylamine_c ephedrine_c ethanol_c ethyleneglycol_c
fenfluramine_c fentanyl_c fluoxetine_c fluvoxamine_c
gabapentin_c gammahydroxybutyrate_c guaifenesin_c
hydrocodone_c hydromorphone_c hydroxyzine_c ibuprofen_c
imipramine_c insulin_c isopropanol_c ketamine_c lithium_c
lorazepam_c lysergicaciddiethylamide_c mda_c mdma_c
meclizine_c meperidine_c meprobamate_c metaxalone_c
methadone_c methamphetamine_c methanol_c methocarbamol_c
methotrexate_c methylphenidate_c midazolam_c mirtazapine_c
molidone_c morphine_c naproxen_c nefazodone_c nicotine_c
nitricacid_c norchlorcyclizine_c nordiazepam_c nordoxepin_c
norfluoxetine_c normeperidine_c norpropoxyphene_c
nortriptylene_c olanzapine_c orphenadrine_c oxazepam_c
oxycodone_c oxymorphone_c pancuroniumbromide_c paraquat_c
paroxetine_c pentazocine_c pentobarbital_c pentoxifylline_c
phencyclidine_c phenobarbital_c phentermine_c
phenyltoloxamine_c phenytoin_c primidone_c promethazine_c
propane_c propoxyphene_c propranolol_c propyleneglycol_c
quetiapine_c quinine_c ritalinicacid_c salicylate_c
salicylicacid_c secobarbital_c sertraline_c temazepam_c
tetrafluroethane_c tetrahydrocannabinol_c thccarboxylicacid_c

theophylline_c thiopental_c thioridazine_c toluene_c
tramadol_c tranylcypromine_c trazodone_c triazolam_c
tricyclic_c trifluoperazine_c trihexyphenidyl_c
trimipramine_c valproic_c venlafaxine_c verapamil_c
zolpidem_c acetaminophen_p acetone_p acetylmorphine_p
alprazolam_p amitriptyline_p amoxapine_p amphetamine_p
benzoylecgonine_p benztropine_p brompheniramine_p bupropion_p
busiprone_p butabarbital_p butalbital_p caffeine_p
carbamazepine_p carbondioxide_p carbonmonoxide_p
carisoprodol_p chlordiazepoxide_p chlorodifluoromethane_p
chlorpheniramine_p chlorpromazine_p citalopram_p
clomipramine_p clonzepam_p clozapine_p cocaethlyene_p
cocaine_p codeine_p cyanide_p cyclobenzaprine_p
desalkyflurazepam_p desipramine_p dextromethorphan_p
diazepam_p difluoroethane_p dihydrocodeine_p diltiazem_p
diphenhydramine_p doxepin_p doxylamine_p ephedrine_p
ethanol_p ethyleneglycol_p fenfluramine_p fentanyl_p
fluoxetine_p fluvoxamine_p gabapentin_p
gammahydroxybutyrate_p guaifenesin_p hydrocodone_p
hydromorphone_p hydroxyzine_p ibuprofen_p imipramine_p
insulin_p isopropanol_p ketamine_p lithium_p lorazepam_p
lysergicaciddiethylamide_p mda_p mdma_p meclizine_p
meperidine_p meprobamate_p metaxalone_p methadone_p
methamphetamine_p methanol_p methocarbamol_p methotrexate_p
methylphenidate_p midazolam_p mirtazapine_p molidone_p
morphine_p naproxen_p nefazodone_p nicotine_p nitricacid_p
norchlorcyclizine_p nordiazepam_p nordoxepin_p
norfluoxetine_p normeperidine_p norpropoxyphene_p
nortriptylene_p olanzapine_p orphenadrine_p oxazepam_p
oxycodone_p oxymorphone_p pancuroniumbromide_p paraquat_p
paroxetine_p pentazocine_p pentobarbital_p pentoxifylline_p
phencyclidine_p phenobarbital_p phentermine_p
phenyltoloxamine_p phenytoin_p primidone_p promethazine_p
propane_p propoxyphene_p propranolol_p propyleneglycol_p
quetiapine_p quinine_p ritalinicacid_p salicylate_p
salicylicacid_p secobarbital_p sertraline_p temazepam_p
tetrafluroethane_p tetrahydrocannabinol_p thccarboxylicacid_p
theophylline_p thiopental_p thioridazine_p toluene_p
tramadol_p tranylcypromine_p trazodone_p triazolam_p
tricyclic_p trifluoperazine_p trihexyphenidyl_p
trimipramine_p valproic_p venlafaxine_p verapamil_p
zolpidem_p acetaminophen_l acetone_l acetylmorphine_l
alprazolam_l amitriptyline_l amoxapine_l amphetamine_l
benzoylecgonine_l benztropine_l brompheniramine_l bupropion_l
busiprone_l butabarbital_l butalbital_l caffeine_l
carbamazepine_l carbondioxide_l carbonmonoxide_l
carisoprodol_l chlordiazepoxide_l chlorodifluoromethane_l
chlorpheniramine_l chlorpromazine_l citalopram_l
clomipramine_l clonzepam_l clozapine_l cocaethlyene_l
cocaine_l codeine_l cyanide_l cyclobenzaprine_l
desalkyflurazepam_l desipramine_l dextromethorphan_l
diazepam_l difluoroethane_l dihydrocodeine_l diltiazem_l
diphenhydramine_l doxepin_l doxylamine_l ephedrine_l

```
ethanol_1 ethyleneglycol_1 fenfluramine_1 fentanyl_1
fluoxetine_1 fluvoxamine_1 gabapentin_1
gammahydroxybutyrate_1 guaifenesin_1 hydrocodone_1
hydromorphone_1 hydroxyzine_1 ibuprofen_1 imipramine_1
insulin_1 isopropanol_1 ketamine_1 lithium_1 lorazepam_1
lysergicaciddiethylamide_1 mda_1 mdma_1 meclizine_1
meperidine_1 meprobamate_1 metaxalone_1 methadone_1
methamphetamine_1 methanol_1 methocarbamol_1 methotrexate_1
methylphenidate_1 midazolam_1 mirtazapine_1 molidone_1
morphine_1 naproxen_1 nefazodone_1 nicotine_1 nitricacid_1
norchlorcyclizine_1 nordiazepam_1 nordoxepin_1
norfluoxetine_1 normeperidine_1 norpropoxyphene_1
nortriptylene_1 olanzapine_1 orphenadrine_1 oxazepam_1
oxycodone_1 oxymorphone_1 pancuroniumbromide_1 paraquat_1
paroxetine_1 pentazocine_1 pentobarbital_1 pentoxifylline_1
phencyclidine_1 phenobarbital_1 phentermine_1
phenyltoloxamine_1 phenytoin_1 primidone_1 promethazine_1
propane_1 propoxyphene_1 propranolol_1 propyleneglycol_1
quetiapine_1 quinine_1 ritalinicacid_1 salicylate_1
salicylicacid_1 secobarbital_1 sertraline_1 temazepam_1
tetrafluroethane_1 tetrahydrocannabinol_1 thccarboxylicacid_1
theophylline_1 thiopental_1 thioridazine_1 toluene_1
tramadol_1 tranylcypromine_1 trazodone_1 triazolam_1
tricyclic_1 trifluoperazine_1 trihexyphenidyl_1
trimipramine_1 valproic_1 venlafaxine_1 verapamil_1
zolpidem_1 number_of_drugs;

drop acetaminophen_n acetone_n acetylmorphine_n alprazolam_n
amitriptyline_n amoxapine_n amphetamine_n benzoylecgonine_n
benztropine_n brompheniramine_n bupropion_n busiprone_n
butabarbital_n butalbital_n caffeine_n carbamazepine_n
carbondioxide_n carbonmonoxide_n carisoprodol_n
chlordiazepoxide_n chlorodifluoromethane_n chlorpheniramine_n
chlorpromazine_n citalopram_n clomipramine_n clonzepam_n
clozapine_n cocaethlyene_n cocaine_n codeine_n cyanide_n
cyclobenzaprine_n desalkyflurazepam_n desipramine_n
dextromethorphan_n diazepam_n difluoroethane_n
dihydrocodeine_n diltiazem_n diphenhydramine_n doxepin_n
doxylamine_n ephedrine_n ethanol_n ethyleneglycol_n
fenfluramine_n fentanyl_n fluoxetine_n fluvoxamine_n
gabapentin_n gammahydroxybutyrate_n guaifenesin_n
hydrocodone_n hydromorphone_n hydroxyzine_n ibuprofen_n
imipramine_n insulin_n isopropanol_n ketamine_n lithium_n
lorazepam_n lysergicaciddiethylamide_n mda_n mdma_n
meclizine_n meperidine_n meprobamate_n metaxalone_n
methadone_n methamphetamine_n methanol_n methocarbamol_n
methotrexate_n methylphenidate_n midazolam_n mirtazapine_n
molidone_n morphine_n naproxen_n nefazodone_n nicotine_n
nitricacid_n norchlorcyclizine_n nordiazepam_n nordoxepin_n
norfluoxetine_n normeperidine_n norpropoxyphene_n
nortriptylene_n olanzapine_n orphenadrine_n oxazepam_n
oxycodone_n oxymorphone_n pancuroniumbromide_n paraquat_n
paroxetine_n pentazocine_n pentobarbital_n pentoxifylline_n
```

```
            phencyclidine_n phenobarbital_n phentermine_n
            phenyltoloxamine_n phenytoin_n primidone_n promethazine_n
            propane_n propoxyphene_n propranolol_n propyleneglycol_n
            quetiapine_n quinine_n ritalinicacid_n salicylate_n
            salicylicacid_n secobarbital_n sertraline_n temazepam_n
            tetrafluroethane_n tetrahydrocannabinol_n thccarboxylicacid_n
            theophylline_n thiopental_n thioridazine_n toluene_n
            tramadol_n tranylcypromine_n trazodone_n triazolam_n
            tricyclic_n trifluoperazine_n trihexyphenidyl_n
            trimipramine_n valproic_n venlafaxine_n verapamil_n
            zolpidem_n i;
run;

proc freq data=thesis.DUIfinal;
      tables age;
run;

/*Merge two datasets together*/
data thesis.combined;
      set thesis.Mefinal1 thesis.DUIfinal;
      indexfatal = find(fatalnew, 'POISONING', 'i');
      drop tox_  presence amount drug_level drug unit;

run;

*Can either be a drug death or in the dui dataset;
data thesis.final1;
      set thesis.combined;
      if indexfatal ne 0 | database = 'dui';

      drop drug_s_ index1 index2 index3;
run;

data thesis.diphenhydramine2;
      set thesis.final1;
      if diphenhydramine = 1 & diphenhydramine_c ne 0 &
            diphenhydramine_p ne 'not detected';

      if (number_of_drugs = 1 & diphenhydramine = 1) then group =
            'single';
      else if number_of_drugs = 2 & (diphenhydramine = 1 & ethanol = 1)
            then group = 'with alcohol';
      else if number_of_drugs >=2 then group = 'et al';

run;

data thesis.diphenhydraminedata;
      set thesis.diphenhydramine2; *or diphenhydramine1;
      if bac = 'BLOOD' | bac = 'aBLOOD(POST)' |
            bac = 'bBLOOD(PRE)';
run;

data thesis.diphenhydraminefinal;
```

```
      set thesis.diphenhydraminedata;
      if case_number ne 'C0133260' & case_number ne 'C0133349' &
            case_number ne 'C0136742' & case_number ne 'NV045033' &
            case_number ne 'NV046032' & case_number ne 'NV046331' &
            case_number ne 'NV046619' & case_number ne 'NV047201' &
            case_number ne 'NV047229' & case_number ne 'NV050545' &
            case_number ne 'NV050657' & case_number ne 'NV050687' &
            case_number ne 'NV051115' & case_number ne 'T0058619' &
            case_number ne 'T0059266' & case_number ne 'T0063020' &
            case_number ne 'W0067678' & case_number ne 'W0067902' &
            case_number ne 'W0068316' & case_number ne 'W0068326';
run;

data thesis.diphenfinal;
      set thesis.diphenhydraminefinal;
      if group = 'single' then newgroup = 1;
      if group = 'with a' | group = 'et al' then newgroup = 2;
      logdiphen = log(diphenhydramine_c);

run;

proc glm data=thesis.diphenfinal;
      class newgroup database;
      model logdiphen = newgroup database newgroup*database;
run;

data thesis.cocaine2;
      set thesis.final1;
      if (cocaine = 1 | benzoylecgonine = 1) & cocaine_c ne 0 &
            cocaine_c ne . & cocaine_p ne 'not detected';

      if (number_of_drugs = 1 & cocaine = 1) | (number_of_drugs = 2
            & (cocaine = 1 & benzoylecgonine = 1)) then group =
            'single';
      else if (number_of_drugs = 2 & (cocaine = 1 & ethanol = 1)) |
            (number_of_drugs = 3 & (cocaine = 1 & benzoylecgonine = 1
            & ethanol = 1)) then group = 'with alcohol';
      else if number_of_drugs >=2 then group = 'et al';

run;

data thesis.cocainedata;
      set thesis.cocaine2;
      if bac = 'BLOOD' | bac = 'aBLOOD(POST)' |
            bac = 'bBLOOD(PRE)';
run;

data thesis.cocainefinal;
      set thesis.cocainedata;
      if case_number ne 'C0131292' & case_number ne 'C0131309' &
            case_number ne 'C0132494' & case_number ne 'C0133717' &
            case_number ne 'C0133871' & case_number ne 'C0135112' &
            case_number ne 'NV045054' & case_number ne 'NV045197' &
```

```
            case_number ne 'NV045270' & case_number ne 'NV045356' &
            case_number ne 'NV045619' & case_number ne 'NV045869' &
            case_number ne 'NV046009' & case_number ne 'NV046291' &
            case_number ne 'NV046329' & case_number ne 'NV046405' &
            case_number ne 'NV046632' & case_number ne 'NV047112' &
            case_number ne 'NV047201' & case_number ne 'NV047579' &
            case_number ne 'NV047622' & case_number ne 'NV047821' &
            case_number ne 'NV047878' & case_number ne 'NV047931' &
            case_number ne 'NV047935' & case_number ne 'NV048054' &
            case_number ne 'NV048238' & case_number ne 'NV048252' &
            case_number ne 'NV048292' & case_number ne 'NV048526' &
            case_number ne 'NV048603' & case_number ne 'NV048773' &
            case_number ne 'NV048882' & case_number ne 'NV048906' &
            case_number ne 'NV049148' & case_number ne 'NV050053' &
            case_number ne 'NV050361' & case_number ne 'NV050376' &
            case_number ne 'NV050390' & case_number ne 'NV050414' &
            case_number ne 'NV050687' & case_number ne 'NV050963' &
            case_number ne 'NV051034' & case_number ne 'NV051160' &
            case_number ne 'NV051257' & case_number ne 'T0061889' &
            case_number ne 'T0063321' & case_number ne 'W0064705' &
            case_number ne 'W0065652' & case_number ne 'W0066365' &
            case_number ne 'W0066710' & case_number ne 'W0066766' &
            case_number ne 'W0067005' & case_number ne 'W0067031' &
            case_number ne 'W0067044' & case_number ne 'W0067455' &
            case_number ne 'W0067497' & case_number ne 'W0067551' &
            case_number ne 'W0067734' & case_number ne 'W0067778' &
            case_number ne 'W0067820' & case_number ne 'W0067912' &
            case_number ne 'W0068023' & case_number ne 'W0068095' &
            case_number ne 'W0068343' & case_number ne 'W0068394' &
            case_number ne 'W0068443' & case_number ne 'W0068581' &
            case_number ne 'W0068708' & case_number ne 'W0068792' &
            case_number ne 'W0068837' & case_number ne 'W0068861' &
            case_number ne 'W0068937' & case_number ne 'W0068952' &
            case_number ne 'W0069006' & case_number ne 'W0069040' &
            case_number ne 'W0069408' ;
run;

data thesis.cocfinal;
      set thesis.cocainefinal;
      logcocaine = log(cocaine_c);
run;

proc glm data=thesis.cocfinal;
      class group database;
      model logcocaine = group database group*database;
run;

data thesis.oxycodone2;
      set thesis.final1;
      if oxycodone = 1 & oxycodone_p ne 'not detected' &
            oxycodone_c ne 0 & oxycodone_c ne .;

      if (number_of_drugs = 1 & oxycodone = 1) then group =
```

```
            'single';
      else if number_of_drugs = 2 & (oxycodone = 1 & ethanol = 1)
            then group = 'with alcohol';
      else if number_of_drugs >=2 then group = 'et al';
run;

data thesis.oxycodonedata;
      set thesis.oxycodone2;
      if bac = 'BLOOD' | bac = 'aBLOOD(POST)' |
            bac = 'bBLOOD(PRE)';
run;

data thesis.oxycodonefinal;
      set thesis.oxycodonedata;
      if case_number ne 'T0060024' & case_number ne 'T0060223' &
            case_number ne 'T0063930' & case_number ne 'T0063935' &
            case_number ne 'T0063980' & case_number ne 'T0064036' &
            case_number ne 'T0064217' & case_number ne 'W0063194' &
            case_number ne 'W0063955' & case_number ne 'W0064128' &
            case_number ne 'W0064135' & case_number ne 'W0064844' &
            case_number ne 'W0064884' & case_number ne 'W0064912' &
            case_number ne 'W0064992' & case_number ne 'W0065793' &
            case_number ne 'W0067575' & case_number ne 'W0067766' &
            case_number ne 'W0069583';
run;

data thesis.oxyfinal;
      set thesis.oxycodonefinal;
      logoxy = log(oxycodone_c);
run;

proc glm data=thesis.oxyfinal;
      class group database;
      model logoxy = group database group*database;
run;

data thesis.hydro2;
      set thesis.final1;
      if hydrocodone = 1 & hydrocodone_p ne 'not detected' &
            hydrocodone_c ne 0 & hydrocodone_c ne .;

      if (number_of_drugs = 1 & hydrocodone = 1) then group =
            'single';
      else if number_of_drugs = 2 & (hydrocodone = 1 &
            ethanol = 1) then group = 'with alcohol';
      else if number_of_drugs >=2 then group = 'et al';
run;

data thesis.hydrodata;
      set thesis.hydro2;
      if bac = 'BLOOD' | bac = 'aBLOOD(POST)' |
            bac = 'bBLOOD(PRE)';
```

```
data thesis.hydrocodonefinal;
      set thesis.hydrodata;
      if case_number ne 'C0134397' & case_number ne 'C0134883' &
            case_number ne 'C0136093' & case_number ne 'T0059266' &
            case_number ne 'T0061306' & case_number ne 'T0063871' &
            case_number ne 'T0063980' & case_number ne 'T0064217' &
            case_number ne 'W0064613' & case_number ne 'W0064705' &
            case_number ne 'W0065937' & case_number ne 'W0066627' &
            case_number ne 'W0066632' & case_number ne 'W0067364' &
            case_number ne 'W0067426' & case_number ne 'W0067766' &
            case_number ne 'W0067977' & case_number ne 'W0068235' &
            case_number ne 'W0068326' & case_number ne 'W0068348' &
            case_number ne 'W0068382' & case_number ne 'W0068421' &
            case_number ne 'W0068708' & case_number ne 'W0068797' &
            case_number ne 'W0068849' & case_number ne 'W0068999' &
            case_number ne 'W0069082' & case_number ne 'W0069583';
run;

data thesis.hydrofinal;
      set thesis.hydrocodonefinal;
      if group = 'single' then newgroup = 1;
      if group = 'with a' | group = 'et al' then newgroup = 2;
      else group = .;
      loghydro = log(hydrocodone_c);
run;

proc glm data=thesis.hydrofinal;
      class newgroup database;
      model loghydro = newgroup database newgroup*database;
run;

data thesis.meth2;
      set thesis.final1;
      if methadone = 1 & methadone_p ne 'not detected' &
            methadone_c ne 0 & methadone_c ne .;

      if (number_of_drugs = 1 & methadone = 1) then group = 'single';
      else if number_of_drugs = 2 & (methadone = 1 & ethanol = 1)
            then group = 'with alcohol';
      else if number_of_drugs >=2 then group = 'et al';
run;

data thesis.methdata;
      set thesis.meth2;
      if bac = 'BLOOD' | bac = 'aBLOOD(POST)' |
            bac = 'bBLOOD(PRE)';
run;

data thesis.methadonefinal;
      set thesis.methdata;
      if case_number ne 'NV046574' & case_number ne 'T0059478' &
            case_number ne 'T0062596' & case_number ne 'T0063441' &
```

```
            case_number ne 'T0063548' & case_number ne 'T0063555' &
            case_number ne 'W0064013' & case_number ne 'W0064530' &
            case_number ne 'W0065458' & case_number ne 'W0065525' &
            case_number ne 'W0065559' & case_number ne 'W0067039' &
            case_number ne 'W0067338' & case_number ne 'W0068276';
run;

data thesis.methfinal;
      set thesis.methadonefinal;
      logmeth = log(methadone_c);
run;

proc glm data=thesis.methfinal;
      class group database;
      model logmeth = group database group*database;
run;

proc glm data=thesis.methfinal;
      class group database;
      model logmeth = group database;
run;

data thesis.morphine2;
      set thesis.final1;
      if morphine = 1 & morphine_p ne 'not detected' &
            morphine_c ne 0 & morphine_c ne .;

      if number_of_drugs = 1 & morphine = 1 then group = 'single';
      else if number_of_drugs = 2 & (morphine = 1 & ethanol = 1)
            then group = 'with alcohol';
      else if number_of_drugs >=2 then group = 'et al';
run;

data thesis.morphinedata;
      set thesis.morphine2;
      if bac = 'BLOOD' | bac = 'aBLOOD(POST)' |
            bac = 'bBLOOD(PRE)';
run;

data thesis.morphinefinal;
      set thesis.morphinedata;
run;

data thesis.morphfinal;
      set thesis.morphinefinal;
      logmorph = log(morphine_c);
run;

proc glm data=thesis.morphfinal;
      class group database;
      model logmorph = group database group*database;
run;
```

```
proc glm data=thesis.morphfinal;
      class group database;
      model logmorph = group database;
run;
```

APPENDIX E

Table of Variable Fields

Field	Description
case number	identifier - character
age	numeric
race description	character
gender	character
court name	character
court type description	character
current date	numeric – date format
cause of death	detailed description of cause of death
manner of death	Suicide, homicide, undetermined, accidental, undetermined
date of injury	numeric – date format
date of death	numeric – date format
place of injury	character
place of death	character
fatalnew	generic cause of death
premisenew	location information
database	source dataset (ME or DUI)
count	used to determine how many corrections were made to cause of death field
bac	listed the tissue where the drug was detected
"drug" array	139 names of drugs (1 for present, 0 for missing)
"concentrations" array	139 concentrations (for each drug)
"units" array	139 units of measurements (for each drug)
"presence" array	139 presences (for each drug)
number of drugs	counter field that specifies the number of drugs detected in an individual
indexfatal	denoted whether of not cause of death was attributable to a drug poisoning
group	indicated drug pattern group

VITA

Amy Elizabeth Herrin grew up in Knoxville, Tennessee, where she attended Bearden High School. After high school, she went to Emory & Henry College in southwestern Virginia. She graduated magna cum laude with a Bachelor's of Science degree in both Mathematics and Computer Science and a minor in Chemistry in May 2004. Amy then moved to Richmond, where she started in the Biostatistics Department at VCU. She received the GlaxoSmithKline first-year award in 2004. She worked for Technology Services as a statistics consultant for faculty, staff, and students who need help with their data from 2005-2006. She is moving to Charleston, South Carolina, where she will be a faculty-level Research Associate at the Medical University of South Carolina in the Department of Biostatistics, Bioinformatics, and Epidemiology.