A model of evaluative opinion to encourage greater transparency and justification of interpretation in post-mortem forensic toxicology.

Peter D Maskell[a,b*], Simon Elliott[c,d], Brigitte Desharnais[e], Martin Findell[f], Graham Jackson[f,g]

[a]Scottish Police Authority Forensic Services, Glasgow, UK

[b]Forensic Medicine and Science, University of Glasgow, Glasgow, UK

[c]Elliott Forensic Consulting, Birmingham, UK

[d]Department of Analytical, Environmental and Forensic Sciences, King's College London, London, UK

[e]Department of Toxicology, Laboratoire de sciences judiciaires et de médecine légale, 1701 Parthenais Street, Montréal, QC H2K 3S7, Canada

[f]Division of Psychology and Forensic Science, School of Applied Sciences, Abertay University, Dundee, UK

[g]Advance Forensic Science, St. Andrews, Scotland, UK

[*]Corresponding Author - peter.maskell.2@glasgow.ac.uk

FUNDING SOURCES

None

ACKNOWLEDGEMENTS

GJ wishes to gratefully acknowledge the support of the Alan Turing Institute for part of this project.

PDM wishes to thank Prof G Williams for introducing him to Bayes' Theorem and Alanna de Korompay for helpful discussions.

A model of evaluative opinion to encourage greater transparency and justification of interpretation in post-mortem forensic toxicology.

Abstract

Over the past decades, the calls to improve the robustness of interpretation in forensic science have increased in magnitude. Forensic toxicology has seen limited progress in this regard. In this work, we propose a transparent interpretive pathway for use in post-mortem forensic toxicology cases. This process allows the selection of the interpretive methodology based on the amount of previous information that is available for the substance(s) in question. One approach is an assessment of various pharmacological and circumstantial considerations resulting in a toxicological significance score (TSS), which is particularly useful in situations where limited information about a substance is available. When there is a robust amount of case data available, then a probabilistic approach, through the evaluation of likelihood ratios (LRs) by the toxicologist and of prior probabilities by the fact-finder, is utilized. This methodology provides a transparent means of making an interpretive decision on the role of a drug in the cause of death. This will allow the field of forensic toxicology to take a step forward in using best practice in evaluative reporting, a tool already used by many other forensic science disciplines.

Introduction

Scientists need to provide interpretive opinions that are as objective as possible, transparent and based on sound scientific reasoning (1). This has been highlighted as an area that is generally lacking in the field of forensic science (2–4). There have thus been increasing loud calls to improve the “science” in forensic science. In the field of forensic toxicology, modern analytical techniques have enabled the identification, with a very high degree of certainty, of the presence of a particular drug in a post-mortem sample. These techniques have also enabled the determination, with a high degree of accuracy and precision, of the amount of drug present in that sample.

Based on these analytical results, there has been a natural and understandable desire to define for each drug a lethal concentration (i.e., a concentration at which the drug would cause death). This would clarify the probative value of the drug’s presence and concentration, i.e., its importance as a piece of evidence in investigating the death as a whole. Defining a lethal concentration would imply that, for every individual, there is a definitive concentration of drug that will have by itself

caused, beyond doubt, the death of that individual. However, for logical, physiological and medical reasons, defining such a concentration for an individual is an unachievable aim (5). Controlled experiments, self-evidently, can never be performed, so no data will be available from that potential source of information. An alternative option, quoting general statistics derived from populations of past cases, provides some information but is of limited value. Merely stating “therapeutic”, “sub-therapeutic”, “toxic” and “fatal” concentrations observed in populations is of limited help to the fact finder in any one individual case. Thus, even with modern analytical techniques in post-mortem forensic toxicology, the issue of whether the detected drug (or substance) caused, or was a contributing factor in, the death of the individual being investigated remains problematic.

The decision on the cause of death is ultimately made by the trier of fact, also known as the fact finder: commonly, a forensic pathologist, coroner, or forensic medical examiner. This decision is based not only on the toxicological findings but also on the case history, scene investigation, gross and microscopic autopsy findings (6). The fact finder’s role is to resolve the uncertainty about the cause of death and, in so doing, they must consider the probative value of expert and non-expert evidence, and their uncertainty. How the fact finder does that, in a cognitive sense, is very personal but one would hope that fact finders are logical in their inferential process and in their use of evidence. When it comes to the contribution of the forensic toxicologist, it is vital that the interpretive process involved in forming an opinion on the significance of the analytical findings is based on whatever relevant data there may be and is as robust, transparent and as free from bias as possible, rather than being based on professional experience alone (4, 7, 8). This can be particularly difficult with the constant emergence of new psychoactive substances (NPS) for which there may be very limited toxicological and pharmacological information about the substance in question.

In this paper, we focus on just one part of the decision-making process in post-mortem forensic toxicology, namely, the way in which opinions about the probative force[1] of toxicological results may be formed logically and reliably. We will focus on

[1]**Probative force** is the power or tendency of evidence to prove or disprove a particular fact that is in issue in a case. In other words, it is the ability of the evidence to make a fact more or less likely to be true.

how we can give interpretation both for well-known drugs, in which there is a clear causative link to intoxication and death, and for the so-called new psychoactive substances, of which there is little known. We start with the previously proposed toxicological significance score, an interpretive model for qualitative evaluation of the role of a substance in deaths (9). This interpretive model is then built upon to achieve a more quantitative approach that, although having previously been proposed, to date has not been fully explored (10). The models used here are based on the central tenet of toxicology that "the dose makes the poison", or that the higher the concentration of a substance in blood, the higher the probability of adverse effects and death. The model also assumes that these are acute toxicological events (the adverse effects arise from exposure(s) to a substance in a short period of time).

This paper is structured into 5 sections. The first section explains how the analytical results can be assigned a Toxicological Significance Score to give information about the likely toxicological significance of the detected substance in the cause of death. The second section explains how an approach based on conditional probabilities for analytical results may be applied. It describes how these may be incorporated into the overall consideration of the primary issue of whether the detected substance caused the death of the individual. The third section provides a note on strengths and limitations of databases in the assignment of probabilities. The fourth section illustrates how one such database, *Toxicolist*, may be used to inform the assignment of probabilities for analytical results and gives an example case involving the drug ethanol (alcohol). Conclusions are presented in the final section.

Section 1: The Toxicological Significance Score

The Toxicological Significance Score was initially developed to enable a more systematic, reproducible, and transparent approach for assessing the toxicological significance of new psychoactive substances. While full details on the methodology can be found in (9), a high level overview is presented in this section. This approach is particularly important in cases where there is very limited information on the detected substance either in the literature or from other sources such as in-house databases. Whilst the intended application of the Toxicological Significance Score was to assist in reviewing the potential role a new psychoactive substance might have played in a death, the approach was based on concepts regularly used and

encountered in toxicology (9). These include the circumstances of the case, the nature of the specimen(s) analysed, the type of analysis performed, the availability and applicability of known concentration ranges, an assessment of tolerance to the substance, the impact of pharmacogenomics, the presence of other substances, stability of the substance in biological material and the possible *in vitro* or *in vivo* production of the substance (especially in post-mortem situations), in addition to other potential influencing factors such as post-mortem redistribution. Other drugs present that could also contribute to death to a lesser or higher degree than the specific drug being evaluated are also required to be investigated.

The process for assessing the toxicological significance of a specific drug in fatalities requires the following factors to be considered:

1. Presence, concentration, and pharmacological nature of the specific drug under consideration in the analysed sample.
2. Presence, concentration, and pharmacological nature of any other drugs present (including alcohol) in the analysed sample.
3. Circumstances of death.
4. Potential cause of death, including pathology findings (especially important if it is an alternative cause to intoxication, such as drowning or mechanical suicide - e.g., hanging).
5. Depending on the number of cases, it may be possible to determine typical concentrations in varying circumstances (e.g., direct cause of death, or alternative/unrelated cause of death).
6. The deceased's drug/prescription history.
7. The deceased's age, sex, body mass index.
8. Any pre-existing medical conditions.

There are various inter-dependencies in the factors above, such that to assess whether a drug concentration is relevant or not, it is necessary to have comparative concentration datasets in varying circumstances. This includes post-mortem concentrations in fatalities with an alternative cause of death (e.g., hanging, gunshot, road traffic accident, trauma from a fall, etc.), post-mortem concentrations in drug-

related fatalities involving the drug(s) under review and in ante-mortem concentrations (e.g., non-fatal cases and suspected driving under the influence of drugs). These data may be available through published case reports, series, or compilations, but are arguably more appropriately obtained from in-house casework concentrations and experience, as this can minimise any inter-laboratory experimental differences. Aside from ensuring the absolute figure of the measured concentration is not taken at face value (i.e. some drugs are inherently associated with higher concentrations than others due to the dose), this also serves to incorporate the possibility of post-mortem redistribution and other changes that can occur after death which can result in artificially elevated drug concentrations depending on the post-mortem interval, nature of the drug and anatomical site of sampling (whereby blood obtained appropriately from the femoral vein is the preferred source as this is the post-mortem blood sampling site that is the least affected by post-mortem redistribution). If such datasets are not available, knowledge of the pharmacological nature of the drug can assist in predicting the possible actions and effects – as well as any interactions with any other pharmacodynamically similar drugs present. This enables an assessment even when only limited information is available.

Subsequently, it is possible, based on the considerations above, for an appropriately qualified and experienced toxicologist to assign a level of significance to the drug according to the following scores:

1 = Low - Drug unlikely to have contributed to death/alternative cause of death;

2 = Medium - Drug may have contributed to toxicity/death and/or other drugs present may be more toxicologically significant;

3 = High - Drug likely to have contributed to death, even in the presence of other drugs;

U = Unclassified –Insufficient data to allow assessment.

Although there may be some clear assignations during assessment, there will equally be some situations where scoring may be unclear. Aside from potential drug-drug interactions, much of this is due to the specific circumstances involved and drug/prescription history whereby tolerance is an essential consideration. This is

where a concentration for one individual would be “normal” and reflect therapeutic use, whereas the same concentration could be “fatal” for an individual not prescribed or used to the drug. Such instances result in significant overlap in concentrations in different circumstances but assessments and opinion can still be made if sufficient and supporting information is available.

Section 2: A conditional probability approach to interpretation

Assigning a likelihood ratio

A potential solution to the problem of providing transparent, useful opinions is the use of case-specific assignment of conditional probabilities for the analytical test results, when sufficient, relevant and reliable data are available (11). Conditional probabilities are used to evaluate a likelihood ratio (LR) – a measure of the probative force of the test results. As with the Toxicological Significance Score, the LR would be evaluated by an appropriately qualified and experienced toxicologist. There are numerous sources in the forensic science literature that explain the use of conditional probabilities, likelihood ratios and their use within the paradigm known as Bayes Theorem. We refer the reader to (12–17) for further information on this branch of statistics. In short, Bayes Theorem (see Fig. 1) shows how new evidence may be used logically to update a view about the uncertainty of a fact in issue. It is important to note that the approach we are suggesting in this paper is only one of several possible approaches when investigating issues that may be termed issues of “causality”. This type of issue seems to be particularly relevant in civil litigation (18, 19). It is worth noting that a probabilistic, LR approach has already been suggested for use in forensic pathology cases (20); the current paper extends that approach to evaluative opinion in forensic toxicology and provides illustrative examples.

The likelihood ratio may be defined broadly as:

> *The probability of obtaining the expert’s observation, given that a particular proposition were true, divided by the probability of obtaining the expert’s observation, given that a competing alternative proposition were true.*

It is important to stress that the proposition pair should be defined by the fact in issue in order for the LR to be directly relevant to that fact in issue. In cases of drug-

associated death, the fact in issue would appear generally to be whether the detected drug(s) caused the death. The LR would then be defined as:

> *The probability of observing the test result, given that the detected drug was the cause of death, divided by the probability of the test result, given that the detected drug was not the cause of death (death was through some other cause).*

In Bayesian terms, the LR is a measure of the amount by which a view on the truth of the proposition (that the drug caused the death) will be changed by the new test results. For example, an LR of 1 will not change the initial estimation of the proposition's truth, whereas an LR of 100 would strongly increase the initial estimation of the proposition's truth and an LR of 0.01 would strongly decrease that initial estimation.

Assigning a probability for the test result is not a wholly objective process – it is necessarily a mental process and personal to the "assigner". The assigner (in this case, the forensic toxicologist) will use their knowledge, experience and understanding of the evidence type to form a view of the probability but, crucially, they will draw upon whatever relevant, reliable data there may be. In our example here, data from two relevant populations would help the toxicologist assign probabilities for observing the test result given that the detected drug was or was not the cause of death. These two populations would be:

1) A population of deaths in which the drug was detected, quantified in the deceaseds' femoral blood samples and to which death was attributed.

2) A population of deaths in which the drug was detected, quantified in the deceaseds' femoral blood samples, but in which death was not attributed to the drug.

Because the propositions are predicated on the cause of death, this information should be known with certainty for all individuals in the dataset. Of course, knowing with absolute certainty the cause of a death is an ideal that is largely unachievable. The best we can hope for is an assessment of the cause of death which has been achieved through a robust, logical evaluation of the evidence that avoids any circular argumentation. From the datasets of drug concentrations within each of the two populations, one can build a probability density function (PDF) for each group (see

Fig. 2). For any given drug concentration, the ratio of the two densities gives us the LR for that concentration. The magnitude of the (log)LR is a measure of the degree of support provided by the observed concentration for one or the other proposition. It is important to note that for some drugs, such as morphine (21), there will be a broad overlap between the two distributions. Within this overlap, there will be a concentration range for which the two probabilities will be approximately equal, i.e., the observed concentration would be approximately equally probable if death was or was not caused by the drug. The resultant value of the LR for these concentrations will then be approximately 1 and the findings would therefore be of no probative force. However, as the two curves start to diverge from each other, the LR will move away from a value of 1, thereby providing probative value one way or the other. It is important, wherever possible, to use datasets that are as relevant as possible to the deceased's circumstances, e.g., age, sex, level of tolerance to drug, health of the deceased prior to death, etc. These data may not be available, so we must recognise and acknowledge the degree of uncertainty in probability assignments from such datasets. Greater uncertainty in the probability assignments should be reflected in LR values that are closer to 1. Further discussion about relevant populations and datasets is given in Section 3 of this paper.

Incorporating a likelihood ratio into the context of the fact in issue

At this point in the paper, we diverge from the central aim of assigning an LR value to that of placing the findings into the overall context of the issue facing the fact finder, i.e., deciding on the issue of whether the drug caused the death. It is important to clarify that whereas the LR is determined by the forensic expert, the placing of the LR into the overall context of the case is carried out by the trier of fact. When determining if the drug detected caused the death, Bayesian logic tell us that the posterior probability that a death was caused by the drug is a function of the LR and the prior probability of the drug being the cause of death.

For convenience, this relationship may be written in the Odds Form of Bayes Theorem, as:

$$\text{Odds}[D|C] = \text{Odds}[D] \times \text{LR}[C] \quad (1)$$

Where:

Odds[D|C] is the posterior odds of the death **[D]** having been caused by the drug, given the concentration **[C]** of drug present.

Odds[D] = prior odds (before the toxicology results are known) of the death having been caused by the drug.

$$\mathrm{LR[C]} = \frac{\text{Probability of observing measured drug concentration given drug was the cause of death}}{\text{Probability of observing measured drug concentration given drug was not the cause of death}} \quad (2)$$

Odds and probability are easily interchangeable, given that odds are simply the ratio of a probability to its complement.

It is important to note that the prior odds are case specific (22). While data such as Fatal Toxicity Index (FTI)[2] may be consulted to provide general values, the value assigned in any one case will be determined by what is known about the deceased, their medical history, whether any drugs or paraphernalia were found at the scene, and what witnesses may say (6). This can be illustrated, simplistically, with two cases of suspicious death with different case circumstances.

Case scenario 1

A body was found hanging by a noose from a high landing over a stairwell. A chair was found at the bottom of the stairs. The pathologist observed certain pathological features that, while not being specific for hanging, were commonly found in hangings. The deceased had a previous history of attempting to complete suicide through hanging. Based on the pathological evidence and the case circumstances (and not, as yet, the toxicology results), the pathologist was of the view that death was very probably caused by hanging. Let us assume that, if they were to be asked, they would assign a probability of 95% for this proposition. By implication, this would suggest a 5% probability for some other cause of death. We can readily convert probability into odds and say that the odds of death having been caused by hanging are 95:5 and the complementary odds of death having been caused though some other means are 5:95, or 1:19.

After analysis of a post-mortem femoral blood sample taken from the deceased, a certain concentration (C) of a drug was detected.

[2] The number of fatal poisonings caused by a drug divided the number of people consuming the drug [21].

To assign an LR, a pair of mutually exclusive propositions relating to the fact in issue needs to be defined. In our scenario here, it seems that there are only two possible causes of death – drug or hanging. If that is accepted, then the proposition and its alternative may be defined as:

D_1 – Death was caused by the drug

D_2 – Death was caused by hanging

and the LR will be given by:

$$\mathrm{LR[C]} = \frac{\text{Probability of observed drug concentration given the drug was the cause of death (D1)}}{\text{Probability of observed drug concentration given hanging was the cause of death (D2)}} \quad (3)$$

How may an LR for the observed drug concentration (C) be assigned? It would be useful to have data from two populations to inform the conditional probabilities of the LR – a population in which the cause of death was ascribed to the drug and a population of deaths in which the drug had been detected, but the cause of death was ascribed to something other than the drug. Assume that such data are available and that, on inspection, the observed concentration (C) was found to lie well within the range of deaths caused by the drug but was also found to lie in the tail of the range of concentrations observed in deaths through some other cause. From those data, the toxicologist decided that the observed concentration of the drug was of the order 10 times more probable if the cause of death was the drug (D_1) rather than if it was hanging (D_2) – the LR is therefore estimated at 10. An LR of value 10, by itself, would seem to provide good probative force for the view that death was caused by the drug. However, how does this LR interact with the prior view of the cause of death? We have, from the pathologist, the prior odds of death being due to some other means than hanging, such as the drug (D_1), of 1:19. To arrive at the posterior odds (Eqn. 1) of death being due to the drug, we follow the Odds Form of Bayes Theorem and multiply the prior odds of death having been caused by the drug (Odds [D_1] of 1:19) by the LR of 10 to reach posterior odds of the death having been caused by the drug, given that concentration C was observed (Odds [D_1|C] of 10:19).

10/19 (posterior odds) = 1/19 (prior odds) x 10 (LR)

Transforming these odds back to a probability, we obtain a posterior probability of approximately 0.34 of death having been caused by drugs (D_1) and a

complementary posterior probability of death having been cause by hanging (D_2) of approximately 0.66 (1 - 0.34 = 0.66).

Even though one would think the drug concentration by itself would be sufficient to ascertain a drug death, the LR provided by the concentration is insufficient to swing the prior probability of the cause of death from hanging over to drug toxicity. The posterior probability that the death was caused by hanging is still high at 0.66 but that probability has been reduced from 0.95 by the observation of the drug concentration.

Case scenario 2

A body was found on a sofa in the lounge. The deceased had no prior history of drug abuse, and no drug paraphernalia was found at the scene. The pathologist was uncertain about the cause of death since the autopsy gave no anatomical cause of death.

The prior odds of cause of death would be very different from scenario 1. The pathologist is unsure of the cause of death and therefore, while there may be a range of speculative causes, there are no specific propositions on the cause of death to consider prior to any results of drug analysis are known. It is therefore difficult to assign prior odds. Assume that the same drug and concentration as in scenario 1 was found in femoral blood. It seems we now have a potential cause of death, and we could define the proposition and its alternative as:

D_1 - Death was caused by the drug

D_2 - Death was not caused by the drug but by some other, unknown, cause

and the LR will be given by:

$$\mathrm{LR}[\mathrm{C}] = \frac{\text{Probability of observed drug concentration given drug was the cause of death (D1)}}{\text{Probability of observed drug concentration given some other cause of death (D2)}} \quad (4)$$

What would be reasonable, realistic, prior odds for the proposition that the drug was the cause of death? Assigning these prior odds would be in the remit of the fact finder, considering all the case circumstances. It may be appropriate to assign a probability reflecting the proportion of deaths of uncertain cause that were later shown to be drugs related. However, in this case, let us assume the fact finder

assigns odds of 50:50 (Odds[D_1] = 1) for the prior odds that death was caused by drugs.

The LR provided by the drug concentration was of the same probative force in favour of a drug death (LR of 10) as in scenario 1.

Multiplying the prior odds of 1 by the LR of 10 gives posterior odds of 10:1 on death having been caused by the drug. This converts to a posterior probability of 0.91 of death having been caused by the drug - a high probability that the drug was the cause of death.

For the same drug concentration, and hence same value of LR, the outcome in terms of the posterior probability of the cause of death is very different between the two case scenarios and is completely dependent on the prior probability (prior odds) of the cause of death. This in accordance with accepted practice of not interpreting toxicology results in a vacuum. The drug will not be the cause of death in all cases in which the drug concentration was above what would be considered a “lethal” level; it does not override all factors being the “non-drug” circumstances and evidence.

Following these two examples that illustrate the conditional probability approach to interpretation, a few points are worth discussing. The first is how the notion of bias interplays with the suggested procedure. To avoid any subconscious bias, the prior probability should be assigned based on circumstantial and other information before the concentration of the drug is known. Interestingly, the LR calculated by the toxicologist is based strictly on concentration distributions and does not consider case circumstances. In this way, a so called “snowball effect” is avoided, whereby the pathologist, toxicologist and fact-finder all include the same information into their decision making process, leading to an undesired multiple counting (8, 23).

Furthermore, we must stress that the numbers quoted for LRs should not be construed as having a precision that is simply not justified. Indeed, the UK Forensic Science Regulator’s Standard on the development of evaluative opinions (4) says at 7.2.9:

> *Communication of the outcomes of evaluation needs to take account of the fact that lay jurors and other court participants may equate a numerical value of a LR with a level of precision that was not intended by the scientist. For this*

reason, where the LR is based on a data set that is not adequately representative of the case circumstances or not of sufficient quality or size to enable a precise numerical assessment, it cannot be reported as a single figure; instead, results will be reported at an order of magnitude level, alongside a verbal equivalent (see 8.5.11 - 8.5.12).

and at 8.3.8:

There will be examples, such as probabilistic evaluation software for DNA or fingerprint comparison, which may assign probabilities on the basis of probability density functions, resulting in many significant figures. In such circumstances the LR, when determined, should in general be reduced to one significant figure.

In accordance with these recommendations, we suggest the use of the TSS for cases in which there is not "*a data set (…) adequately representative of the case circumstances or (…) of sufficient quality or size"*. One obvious such case would be NPS which either have never been licenced for clinical use (such as etonitazene (24)) or that have only been licenced in a small number of countries (such as phenazepam and etizolam (25, 26)).

Section 3: Strengths and limitations of available data

There are various sources of drug concentrations quantified in a variety of case types in living and deceased subjects that may be consulted to help inform assignments of probabilities. These may include case studies from journal articles, in-house databases, and published compilations of case studies. As mentioned previously, these data cannot be complemented by experimental studies on lethal concentrations, due to obvious ethical concerns. While clinical reference data about therapeutic concentrations are often available, they are not appropriate to use in drug-death investigations due to a host of potential post-mortem effects well known to forensic toxicologists (5). Data based on post-mortem analytical results will be the most valuable, especially when all the blood samples are from the femoral vein (generally agreed to be the least affected by post-mortem redistribution and for which there are relatively large data sets) and have been analysed with similar

validated analytical procedures. Ideally, we would want a data set from a population of people who were similar to the deceased in terms of factors such as age, gender, ethnicity, state of health, lifestyle. However, there will never be a perfect database for every individual case because each case has unique circumstances that cannot be replicated. Therefore, imperfect existing data sources will need to be appraised by an expert for their relevance to the case at hand and in terms of their size, robustness, and ease of access. The issue of the required sample size and how to improve the reliability of post-mortem drug concentration databases have been discussed by Söderburg *et al.* (11) and Desharnais *et al* (27).

However, a significant limitation of cause-of-death data is their dependence on a judgement by the certifying fact finder, whether that be a pathologist, a coroner or other agent. While definitions of causes of death are standardised by the World Health Organisations (WHO) ICD-11 (28), the reliability and relevance of the final ruling depends on the fact finder's judgement being correct. In many instances, whether drug(s) directly caused a death may be unknowable or surrounded by a significant degree of uncertainty and complexity. Furthermore, there may be an element of circularity to the argument when a death is ascribed to drugs. As far as we are aware, there are no reliable data from populations in which the death, without consideration of the surrounding circumstances, was directly caused by a drug. Statistics on cause of death are based on the certified "immediate" cause of death, as recommended by WHO, but drug intoxication may be listed as an intermediary cause in the sequence or as a contributory factor (29). WHO states that the certified cause of death is "(…) the best medical opinion". In this situation, we must retreat to a broader specification of the fact in issue in toxicological cases - the issue necessarily becomes whether the detected drug(s) was the immediate/intermediate/contributory cause of death. This may seem to be largely a case of semantics, but definition of the issue, and the resultant propositions, is intimately connected with the form of data that are available to inform assignments of probabilities. For the remainder of this paper, we shall use the phrase "cause of death" to accommodate both "immediate" and "intermediate" causes of death.

In summary, while there will always be limitations to available datasets, the expert's role is to assess possible sources and select the most adequate one. One such source of data is presented in the next section.

Section 4: Potential application of "*Toxicolist"* and other sources to inform probability assignments

Druid and Holmgren developed a classification methodology (30) that provides a useful model for the creation of "reference ranges" for the interpretation of post-mortem femoral blood drug concentrations. These data are now in a database called "Toxicolist" (31), which, containing a large number of cases, could be used to assign conditional probabilities in the manner described below.

This dataset comprises multiple groups, all measured in femoral blood:

Group A – Certified deaths by intoxication involving only one detected substance, and other contributory factors could be ruled out. This group could be designated as "Single drug intoxication".

Group B – Certified deaths by intoxication in which more than one substance was detected. This group could be designated as "Multiple drug intoxication".

Group C – Certified other causes of death in which the relevant substance was detected but in which the circumstances of death, in the view of the fact finder, exclude the possibility of incapacitation by the substance, or other substances. This group could be viewed and designated as a "Control group".

It is important to note that, for all cases in groups A, B and C, the certification of cause of death was carried out by the authorised fact finder and was based on all the evidence in the case. These decisions were reviewed by the database compilers. Most substances and substance groupings will show an increasing median drug concentration when going from group C to B to A.

There are several important considerations to remember when assessing whether to use the data from these groups:

- The frequency distributions of "non-fatal" (Groups C) and "fatal" (Groups A and B) concentrations of drugs will inevitably overlap and there is no sharp dividing line between the two sets. However, that is not necessarily a limitation – it is a fact of life (or death) that is accommodated through a probabilistic approach.

- Groups C may include “therapeutic”, “sub-therapeutic” or “toxic” concentrations.
- As causes of death are the opinion of the forensic pathologist (or other fact finder), every approach to categorisation and compilation of fatal and toxic drug concentrations is susceptible to some degree of circular reasoning. For example, a fact finder may decide, given the drug concentration and possibly other evidence, that a particular death was caused by the drug, even though no-one can say for certain that the drug was the cause. This concentration would then be added to the data set of the “fatal due to the drug” group, thereby influencing the data distribution of this grouping for future users. Unless control trials are performed (and they will not be performed for obvious ethical reasons), it is impossible to have datasets of deaths where the drug was and was not, without doubt, the cause of death. However, due to the phenomena of “regression to the mean”, if enough observations are made, an overall reliable mean should be obtainable.
- Several variables are ignored or grouped in such a dataset, e.g., post-mortem sampling intervals and drug metabolism profiles. But, as discussed in Section 3, the perfect database for every individual case is an impossible goal. Assuming variables take similar ranges of values in each group, and if the groups are large enough, then the impact of these variables may be cancelled out. For example, in clinical studies it is recommended that there are 120 samples for each variable (such as sex or age (32, 33)). For forensic studies, this sample size may not be possible. For post-mortem toxicology cases, at least 20 cases are needed for each of the drug groupings (A, B & C) for the creation of usable data sets for probability density functions (11, 34).

Example calculation: Ethanol

Following our earlier definition of an LR evaluation in cause-of-death, a case involving ethanol may be written as:

$$\mathrm{LR}=\frac{\text{Probability of the measured ethanol concentration given ethanol was the cause of death}}{\text{Probability of the measured ethanol concentration given ethanol was not the cause of death}}$$

To help assign a probability for the numerator of the LR, Group A-type data was obtained from the study by Jones and Holmgren with 593 cases that gives the concentration of ethanol in femoral blood where the death was attributed to ethanol

intoxication (35). For the denominator of the LR, Group C-type data was obtained from a study by Jones *et al.* with 1343 cases in which ethanol was quantitated in femoral blood but the cause of death was hanging (36). The Group A data set was normally distributed whereas the Group C data set had a non-standard distribution (Fig. 3A). From these probability density functions, it is possible to evaluate the LR for any given concentration (Fig. 3B). As expected, the LR increases with increasing blood alcohol concentration (BAC); but note that it is only at BACs more than ~245 mg/100mL that the LR starts to exceed 1. It is only at this concentration that the probability of that BAC, given death was caused by ethanol, begins to exceed the probability of that BAC, given death was not caused by ethanol. At this BAC, the LR begins to provide evidence supporting the proposition that ethanol caused the death rather than it did not. As BAC increases above 245 mg/100mL, the magnitude of the LR increases, the probative force of the BAC increases, and the LR has an increasing effect on the posterior odds of death having been caused by ethanol. There is a large increase in the LR at a BAC >360 mg/100ml, due to the small probability density in the group A data set. As the probability of adverse effects from a drug increases with blood concentration, if the BAC is >380 mg/100ml, then the maximum likelihood ratio observed would be utilised. In the ethanol data set, the values in Groups A and C are large enough for robust LRs to be produced in case work.

Cases with polydrug use

The examples we have given above are solo drug intoxications. However, it is more common in post-mortem toxicology to see polypharmacy where multiple drugs have been consumed and detected in the toxicological sample. In these cases, there is added complexity in interpretation as some drugs may have synergistic rather than additive actions (i.e., morphine and alcohol, where both drugs have potential central nervous depressant effects) and other pharmacological actions such as antagonism and potentiation. To accommodate this complexity, the group B (multiple drug intoxication) data set could be used to assign probabilities and be compared with a multi-drug subset of the group C (control) data set. This would potentially allow assignment of LRs individually for each of the drugs in the multi-drug intoxication but would not allow assignment of an LR for the specific combination of drugs observed in a given case. Combining individual drug LRs can only be performed if there is

independence between the probabilities for the individual drugs, an independence which almost certainly is not present. The role of the combination of multiple drugs in a death could be assessed if conditional probability distributions could be generated for common drug combinations. In the meantime, a simpler approach has been proposed by some researchers that may be used when multiple drugs of the same class (such as benzodiazepines) have been consumed (such as etizolam and diazepam). In these cases, “pooled data” may be used. In this methodology, for which background and assumptions are given in (37, 38), there is an assumption that each drug in a class has an equipotent concentration and that a certain concentration of a drug in a class will give the same effect. An example is that etizolam is considered to be 3 times as potent as diazepam (38) and so a blood concentration of etizolam of 1 mg/L would be equivalent to 3 mg/L of diazepam. Therefore, if a concentration of 0.5 mg/L (etizolam) and 1 mg/L (diazepam) were detected in a femoral blood sample, this would give a pooled diazepam equivalent benzodiazepine concentration of 2.5 mg/L ((0.5 x 3) + 1). This “pooled data” approach allows equivalent group A, B and C data to be compiled and used in a probabilistic approach to post-mortem drug interpretation for drug classes.

The reliability of a data set to determine LRs

As mentioned above, it is important to have a robust, reliable and sufficiently large data set in order to assign robust, reliable LRs for specific drugs. However, much more research needs to be carried out to characterize what would constitute a suitable data set for the assignment of LRs. Considerations would include the data set’s size and characteristics (e.g., validated data, measurement uncertainty, the suitability of this model for use with other blood/tissue sampling sites etc.). While the information presented in this paper suggests that a single external database could be used, other avenues are certainly possible, e.g., the use/compilation of an internal laboratory database, or the compilation of relevant case reports from the literature (although this is necessarily “uncontrolled” information). What is presented in this paper is merely an example to help support and illustrate a probabilistic interpretive model. With the compilation of a large, robust and reliable data sets, with suitable variables, in the future it may be possible to utilise artificial intelligence (AI) to help inform a decision on the role of a drug in a death but to date no such relevant, large data sets exist. Commonly in forensic toxicology, especially with the emergence of

new NPS, there is a lack of data, hence the need for the TSS model. Furthermore, AI suffers from being a “black box” system, possibly lacking the required transparency for the legal system. Therefore, with the data and resources available at the present time, the approach suggested in this paper provides a realistic, logical and transparent model for forensic toxicology interpretation. The advantage of such a model is the possibility of continuous refinement as increasing amounts of data become available (such as markers for tolerance, metabolomics data etc.) allowing improvements in forensic toxicology interpretation in a transparent manner.

Section 5. Conclusions

The interpretive model presented here provides a transparent evaluation of the femoral blood concentrations of drugs detected in post-mortem cases in situations where there is limited, pre-existing case data (such as in the case of NPS) and for substances with greater amounts of data. Where robust data sets are available, evaluation of the probative force of observations, through the assignment of likelihood ratios, is suggested, conforming to a standard set out by the UK Forensic Science Regulator. This paper also illustrates how the LR may be combined with the relevant circumstances, avoids the definition of a “lethal concentration”, and provides an alternative means of assessing the potential role of a drug in the cause of death. The methodology outlined here would also be easily adaptable to current databases such as “Toxicolist” to help assign LRs for the blood drug concentrations detected in individual cases. Further work does need to be carried out to allow implementation of the approach presented in this paper in routine forensic casework. A standardized approach to the generation of PDFs, the generation of LRs from the PDFs and the validation of LRs for specific drugs, is required for consistent intra and inter-laboratory application and for any potential AI modelling. Also, much work remains to be done with regards to 1) characterizing the underlying data necessary for a reliable interpretive model, 2) refining the Bayesian model with known forensic toxicology considerations (tolerance, PMR, etc.), and finally 3) exploring the best approaches to use to communicate LR to other forensic practitioners and the wider judicial system in a clear and understandable manner. It is important to note that any conclusions reached by the approach proposed in this paper may not necessarily be expected to be better than the traditional interpretive process that require an experienced forensic toxicologist. It will, however, support experienced forensic toxicologist and

allow them to be more transparent about the conclusions they have reached. The databases/data and other factors used to reach conclusions on, for example, probative force (LR) of the results, would be explicitly laid out and available for consultation, as is common with the analytical results of post-mortem samples and the associated methodological validation. This work represents a first, very small, but important step towards full evaluative reporting in forensic toxicology; but there remains, however, a lot of work to be done.

Data Availability

The data underlying this article are available in the article.

References

1. International Laboratory Accreditation Cooperation (ILAC) (2022) ILAC-G19:06/2022 Modules in a Forensic Science Process. https://ilac.org/?ddownload=124605 (accessed July 17, 2023).

2. National Academy of Sciences (2009) Strengthening Forensic Science in the United States: A Path Forward. https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf (accessed July 17, 2023).

3. President's Council of Advisors on Science and Technology (PCAST) (2016) Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (accessed July 17, 2023)

4. Forensic Science Regulator (2021) Development of Evaluative Opinions (FSR-C-118).

5. Ferner, R.E. (2008) Post-mortem clinical pharmacology. *British Journal of Clinical Pharmacology*, **66**, 430–443.

6. Davis, G.G., Cadwallader, A.B., Fligner, C.L., Gilson, T.P., Hall, E.R., Harshbarger, K.E., et al. (2020) Position Paper: Recommendations for the Investigation, Diagnosis, and Certification of Deaths Related to Opioid and Other Drugs. *The American Journal of Forensic Medicine and Pathology*, **41**, 152–159.

7. Hamnett, H.J. and Dror, I.E. (2020) The effect of contextual information on decision-making in forensic toxicology. *Forensic Science International. Synergy*, **2**, 339–348.

8. Dror, I., Melinek, J., Arden, J.L., Kukucka, J., Hawkins, S., Carter, J., et al. (2021) Cognitive bias in forensic pathology decisions. *Journal of Forensic Sciences*, **66**, 1751–1757.

9. Elliott, S., Sedefov, R. and Evans-Brown, M. (2018) Assessing the toxicological significance of new psychoactive substances in fatalities. *Drug Testing and Analysis*, **10**, 120–126.

10. Fuller, D.C., Bishop-Freeman, S.C. and Winecker, R.E. (2016) A Potential Application of Inferential Statistics to Assist in the Interpretation of Postmortem Drug Concentrations. In Society of Forensic Toxicologists Annual Meeting, Dallas. p. S54. http://www.soft-tox.org/files/meeting_abstracts/SOFT_2016_meeting_abstracts.pdf. (accessed July 17, 2023)

11. Söderberg, C., Tillmar, A., Johansson, A., Wernvik, E., Jönsson, A.K. and Druid, H. (2020) The importance of sample size with regard to the robustness of postmortem reference values. *Forensic Science International*, **311**, 110292.

12. The Council of the Inns of Court and The Royal Statistical Society (2017) Statistics and probability for advocates: Understanding the use of statistical evidence in courts and tribunals. https://www.icca.ac.uk/wp-content/uploads/2019/11/RSS-Guide-to-Statistics-and-Probability-for-Advocates.pdf (accessed July 17, 2023).

13. Aitken, C.G.G., Taroni, F. and Bozza, S. (2020) Statistics and the Evaluation of Evidence for Forensic Scientists: Second Edition. 3rd ed., Wiley.

14. Robertson, B., Vignaux, G.A. and Berger, C.E.H. (2016) Interpreting Evidence: Evaluating Forensic Science in the Courtroom: Second Edition. Wiley-Blackwell. https://www.wiley.com/en-gb/Interpreting+Evidence%3A+Evaluating+Forensic+Science+in+the+Courtroom%2C+2nd+Edition-p-9781118492437 (Accessed September 17, 2023).

15. ENFSI (2015) ENFSI guideline for evaluative reporting in forensic science - Strengthening the evaluation of forensic results across Europe (STEOFRAE). http://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf (accessed July 17, 2023).

16. Jackson, G., Aitken, C. and Roberts, P. (2015) Communicating and Interpreting Statistical Evidence in the Administration of Criminal Justice 4. Case Assessment and Interpretation of Expert Evidence Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses. https://rss.org.uk/RSS/media/File-library/Publications/rss-case-assessment-interpretation-expert-evidence.pdf (accessed July 17, 2023).

17. Johnson, A.A., Ott, M.Q. and Mine Dogucu (2022) Bayes Rules! An Introduction to Applied Bayesian Modeling. 1st ed., Chapman and Hall/CRC, New York.

18. Dawid, A.P., Faigman, D.L. and Fienberg, S.E. (2013) Fitting Science Into Legal Contexts: Assessing Effects of Causes or Causes of Effects? *Sociological Methods & Research*, **43**, 359–390. https://doi.org/10.1177/0049124113515188.

19. Dawid, A.P. and Musio, M. (2022) Effects of Causes and Causes of Effects. *Annual Review of Statistics and Its Application*, **9**, 261–287. https://doi.org/10.1146/annurev-statistics-070121-061120.

20. de Boer, H.H., Fronczek, J., Berger, C.E.H. and Sjerps, M. (2022) The logic of forensic pathology opinion. *International Journal of Legal Medicine*, **136**, 1027–1036.

21. Jones, A.W., Holmgren, A. and Ahlner, J. (2012) Concentrations of free-morphine in peripheral blood after recent use of heroin in overdose deaths and in apprehended drivers. *Forensic Science International*, **215**, 18–24.

22. Maskell, P.D. and Jackson, G. (2020) Presumptive drug testing - The importance of considering prior probabilities. *WIREs Forensic Science*, **2**, e1371.

23. Dror, I.E., Morgan, R.M., Rando, C. and Nakhaeizadeh, S. (2017) Letter to the Editor - The Bias Snowball and the Bias Cascade Effects: Two Distinct Biases that May Impact Forensic Decision Making. *Journal of Forensic Sciences*, **62**, 832–833.

24. Vandeputte, M.M., Verougstraete, N., Walther, D., Glatfelter, G.C., Malfliet, J., Baumann, M.H., et al. (2022) First identification, chemical analysis and pharmacological characterization of N-piperidinyl etonitazene (etonitazepipne), a recent addition to the 2-benzylbenzimidazole opioid subclass. *Archives of Toxicology*, **96**, 1865–1880.

25. Crichton, M.L., Shenton, C.F., Drummond, G., Beer, L.J., Seetohul, L.N. and Maskell, P.D. (2015) Analysis of phenazepam and 3-hydroxyphenazepam in post-mortem fluids and tissues. *Drug Testing and Analysis*, **7**, 926–936.

26. Hikin, L.J., Smith, P.R., Maskell, P.D., Kurimbokus, H., Ashong, E., Couchman, L., et al. (2021) Femoral blood concentrations of the designer benzodiazepine

etizolam in post-mortem cases. *Medicine, Science, and the Law*, **61**, 122–129.

27. Desharnais, B., Camirand-Lemyre, F., Mireault, P. and Skinner, C.D. (2015) Determination of Confidence Intervals in Non-normal Data: Application of the Bootstrap to Cocaine Concentration in Femoral Blood. *Journal of Analytical Toxicology*, **39**, 113–117.

28. World Health Organisation (2022) ICD-11 for Mortality and Morbidity Statistics (ICD-11 MMS). https://icd.who.int/browse11/l-m/en (accessed July 17, 2023).

29. Organization, W.H. (1979) Medical certification of cause of death : instructions for physicians on use of international form of medical certificate of cause of death. 1979. https://apps.who.int/iris/handle/10665/40557 (accessed July 17, 2023).

30. Druid, H. and Holmgren, P. (1997) A Compilation of Fatal and Control Concentrations of Drugs in Postmortem Femoral Blood. *Journal of Forensic Sciences*, **42**, 79–87.

31. Druid, H. (2018) Postmortem Toxicology: The Toxicolist Story. *TIAFT Bulletin*, **48**, 16–20.

32. Clinical and Laboratory Standards Institute (2010) EP28-A3C Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory, 3rd Edition.

33. Poulson, O.M., Holst, E. and Christensen, J.M. (1997) Calculation and application of coverage intervals for biological reference values. *Pure and Applied Chemistry*, **69**, 1601–1611.

34. Horn, P.S., Pesce, A.J. and Copeland, B.E. (1998) A robust approach to reference interval estimation and evaluation. *Clinical Chemistry*, **44**, 622–631.

35. Jones, A.W. and Holmgren, P. (2003) Comparison of Blood-Ethanol Concentration in Deaths Attributed to Acute Alcohol Poisoning and Chronic Alcoholism. *Journal of Forensic Sciences*, **48**, 2002420.

36. Jones, A.W., Holmgren, A. and Ahlner, J. (2013) Toxicology findings in suicides: concentrations of ethanol and other drugs in femoral blood in victims of hanging and poisoning in relation to age and gender of the deceased. *Journal of Forensic and Legal Medicine*, **20**, 842–847.

37. Bech, A.B., Clausen, T., Waal, H., Vindenes, V., Edvardsen, H.E., Frost, J., et al. (2021) Postmortem toxicological analyses of blood samples from 107 patients receiving opioid agonist treatment: substances detected and pooled opioid and benzodiazepine concentrations. *Addiction*, **116**, 845–855.

38. Strand, M.C., Mørland, J., Slørdal, L., Riedel, B., Innerdal, C., Aamo, T., et al. (2017) Conversion factors for assessment of driving impairment after exposure to multiple benzodiazepines/z-hypnotics or opioids. *Forensic Science International*, **281**, 29–36.

Figure 1: Bayes Theorem, with indications on roles in death investigation process.

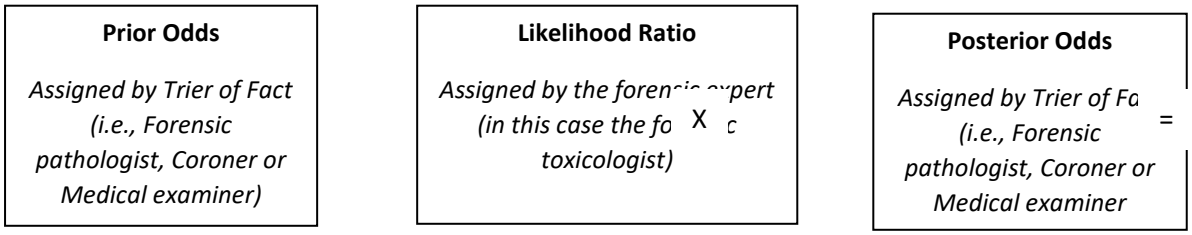

Figure 2A: A fictitious example of probability density functions (PDF) for a drug. The two populations are: 1) A population of deaths in which the drug was detected, quantified in the deceased's femoral blood sample and to which death was attributed ("Death group"). 2) A population of deaths in which the drug was detected, quantified in the deceased's femoral blood sample, but in which death was not attributed to the drug ("Control group"). Both drug populations are plotted against the concentration of the drug (arbitrary units). The lines show how a relevant drug concentration of interest can be used to determine the probabilities for use in likelihood ratio (LR) calculations.

$$LR = \frac{Probability\ of\ observing\ measured\ drug\ concentration\ given\ drug\ was\ the\ \ cause\ of\ death\ (Y)}{Probability\ of\ observing\ measured\ drug\ concentration\ \ given\ drug\ was\ not\ the\ \ cause\ of\ death\ (X)}$$

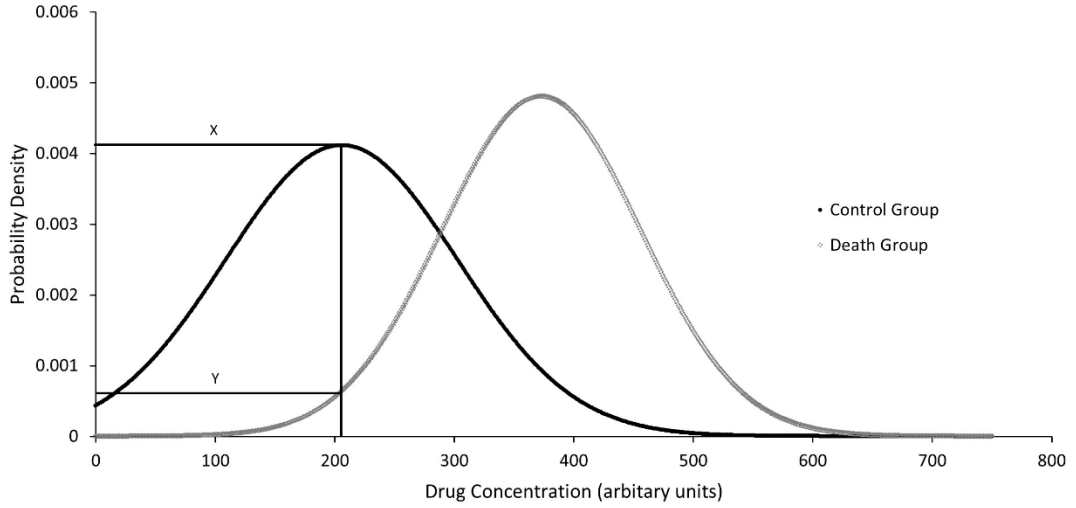

Figure 2B: The likelihood ratio (LR) calculated using probabilities obtained from graph A.

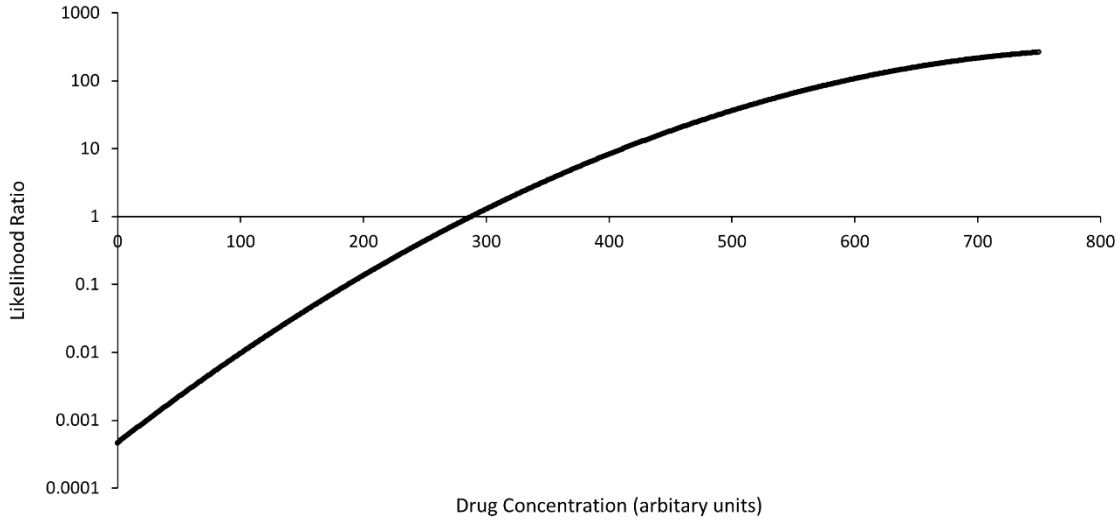

Figure 3A: The probability density distributions based on the relative frequency of cases where alcohol was detected, but in which the circumstances of death were not due to alcohol (Group C) and in which cause of death was given as acute alcohol consumption (Group A).

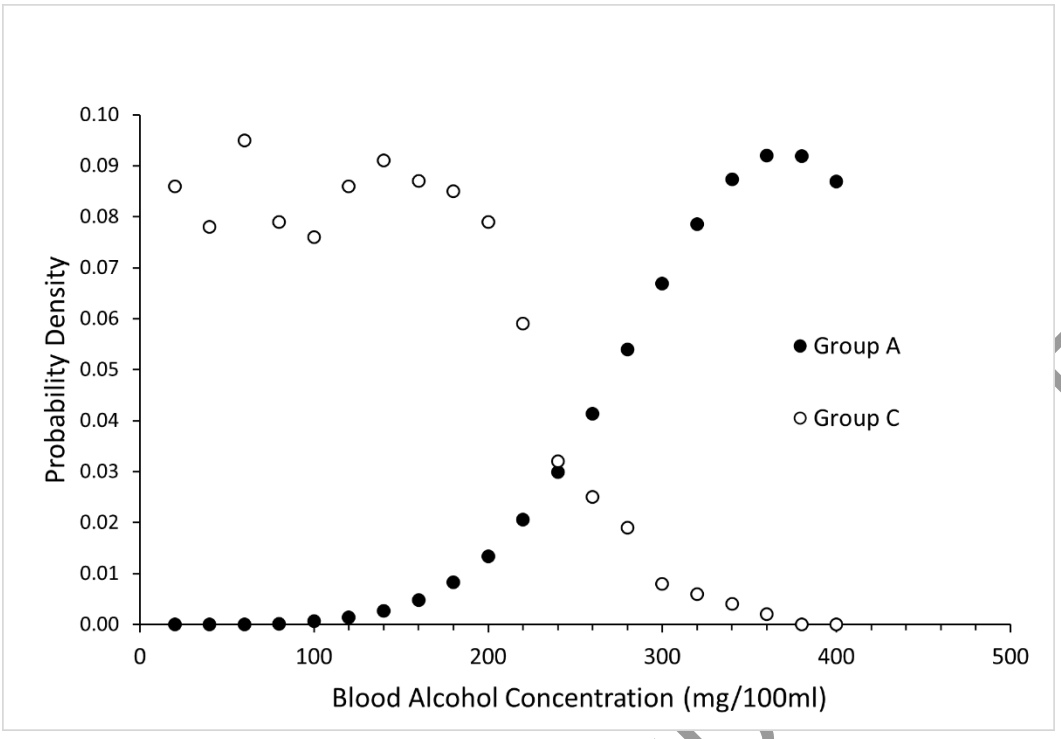

Figure 3B: Calculated likelihood ratio against measured blood alcohol concentrations.

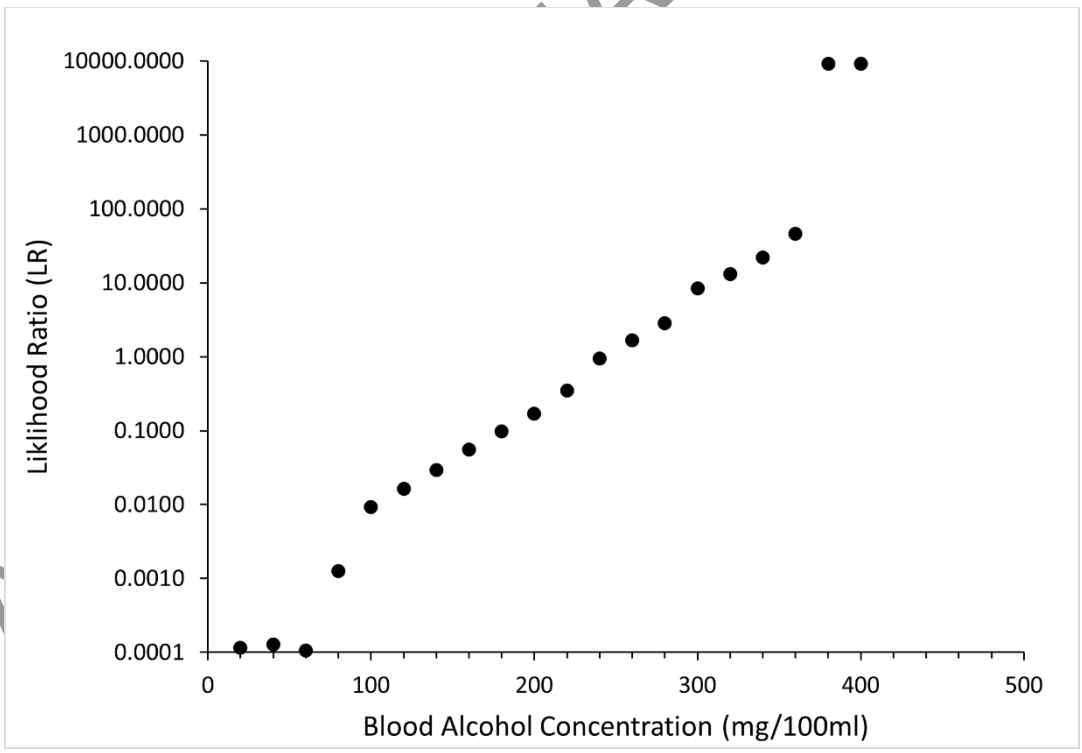

Table 1: Table of the calculated likelihood ratios for given blood alcohol concentrations based on probability density distributions derived from relative frequency of A (solo ethanol intoxication cases) and group C (control cases) based on data from (35, 36)

BAC (mg/100ml)	Probability Density		LR
	Group A	Group C	
<20	0.00001	0.08600	0.0001
21 - 40	0.00001	0.07800	0.0001
41 -60	0.00001	0.09500	0.0001
61- 80	0.00010	0.07900	0.0013
81 -100	0.00070	0.07600	0.0092
101 - 120	0.00140	0.08600	0.0163
121 - 140	0.00267	0.09100	0.0293
141 - 160	0.00483	0.08700	0.0555
161 - 180	0.00826	0.08500	0.0972
180 - 200	0.01339	0.07900	0.1695
201 - 220	0.02057	0.05900	0.3486
221 - 240	0.02994	0.03200	0.9356
241 - 260	0.04131	0.02500	1.6524
261 - 280	0.05399	0.01900	2.8416
281 - 300	0.06688	0.00800	8.3600
301 - 320	0.07850	0.00600	13.0833
321 - 340	0.08730	0.00400	21.8250
341 - 360	0.09201	0.00200	46.0050
361 - 380	0.09188	0.00001	9188.0000
>380	0.08695	0.00001	9188.0000