Forensic Science International 311 (2020) 110292

Contents lists available at ScienceDirect

Forensic Science International

journal homepage: www.elsevier.com/locate/forsciint

The importance of sample size with regard to the robustness of postmortem reference values

C. Söderberg[a,b], A. Tillmar[c,d], A. Johansson[c], E. Wernvik[b], A.K. Jönsson[c], H. Druid[a,b,*]

[a] Department of Oncology-Pathology, Karolinska Institutet, Retzius v. 3, KI Campus Solna, 171 77 Stockholm, Sweden
[b] Department of Forensic Medicine, National Board of Forensic Medicine, Artillerigatan 12, 587 58 Linköping, Sweden
[c] Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Artillerigatan 12, 587 58 Linköping, Sweden
[d] Department of Biomedical and Clinical Sciences, Faculty of Health and Medical Sciences, Linköping University, Sandbäcksgatan 7, 582 35 Linköping, Sweden

ARTICLE INFO

Article history:
Received 16 January 2020
Received in revised form 1 April 2020
Accepted 3 April 2020
Available online 14 April 2020

Keywords:
Postmortem toxicology
Reference concentrations
Sample size
Power

ABSTRACT

Evaluating postmortem toxicological results is a challenging task due to multiple factors affecting blood concentrations after death. In order to improve the diagnostic accuracy in cases of suspected fatal intoxication different compilations of postmortem reference drug concentrations are often used. However, it is not clear what constitutes a reliable postmortem reference value.

The current study presents reference concentrations for 13 substances from seven substance groups according to a standardized protocol. The reference concentrations were gathered from 3767 autopsy cases and subdivided into intoxications by one substance only (Group A, n = 611), multi-substance intoxications (Group B, n = 1355) and postmortem controls, in which incapacitation by drugs were excluded (Group C, n = 1801). In particular, this study presents statistical information about the precision and conformity change with various sample sizes.

Based on the present data >10 detections are usually needed, for the substances examined, to differentiate between intoxication cases and controls. Repeated samplings show that the median of small samples (N = ≤5) has a high variation (normalized interquartile range 138–75%) and that a high number of detections (N = >20) in each group are needed to reduce the variation.

1. Introduction

Fatal intoxications, both intentional and unintentional, are a global issue. Of all deaths attributed to self-harm worldwide intoxications with pharmaceutical drugs are common in the western world, while pesticide intoxications predominate in the developing world [1,2].

Identifying the cause of death in intoxications is complicated by detections of multiple drugs, of which several may in isolation or in combination have caused the death, or sometimes not [3–5]. One challenge for the pathologist or toxicologist is to not only identify potentially fatal concentrations but also to be able to determine the levels of substances that have not contributed to a fatal intoxication. In the postmortem setting evaluating drug concentrations presents an additional challenge due to factors such as postmortem redistribution of drugs [6,7], bacterial metabolism [8] and the post mortem interval [9,10]. Forensic investigators have responded to these issues by compiling postmortem reference concentrations. Most often these compilations are descriptive presentations of detections [11–13] or specific case reports with an often low number of cases [14].

Small data sets are affected by uncertainty in how well the observed concentrations represent the "true" collection of concentrations for a given substance and its toxicological effect (e.g. intoxication and non-intoxication) on a population level. Furthermore, if the observed concentrations are scattered over a broad interval together with overlap between intoxications and non-intoxications the uncertainty is even higher. The sample size needed to provide a reliable postmortem reference value and the added benefit of large sample sizes have not been explored.

We have previously published both fatal and non-fatal postmortem reference concentrations based on the Swedish National Forensic Toxicology database using a standardized method including a rigorous case-by-case review [15–19].

The present study provides postmortem reference concentrations for 13 substances, from different substance groups, subdivided

* Corresponding author at: Department of Oncology-Pathology, Karolinska Institutet, Retzius v. 3, KI Campus Solna, 171 77 Stockholm, Sweden.
E-mail addresses: carl.soderberg@rmv.se (C. Söderberg), andreas.tillmar@rmv.se (A. Tillmar), anna.johansson@rmv.se (A. Johansson), emma.wernvik@rmv.se (E. Wernvik), anna.jonsson@rmv.se (A.K. Jönsson), henrik.druid@ki.se (H. Druid).

http://dx.doi.org/10.1016/j.forsciint.2020.110292

into single intoxications, multi-drug intoxications and control cases using the Swedish National Forensic Toxicology database. This set of reference data was then used to study characteristics such as robustness and reliability for different sample sizes.

2. Material and methods

2.1. Study population

In Sweden all obvious or suspected unnatural death, deaths with unclear circumstances or unclear identity of the deceased should be reported to the police. The police may then request a forensic autopsy from the National Board of Forensic Medicine. At autopsy femoral blood is routinely collected (when available) and will be subjected to a broad toxicological screening at the Swedish national forensic toxicology laboratory, and positive findings will be verified. Case information, autopsy findings and toxicological results are entered into a case management system developed to serve as a real time database [20].

The present study considered all cases in which a selection of substances (alprazolam, amitriptyline, carbamazepine, citalopram, nitrazepam, olanzapine, oxazepam, oxycodone, phenytoin, quetiapine, tramadol, verapamil and zolpidem) had been detected in femoral blood between 1992 and 2010. Postmortem reference values for olanzapine and quetiapine were generated from the same dataset that has in part been presented in a previous publication [19]. The included substances were selected to represent different substance groups.

2.2. Sampling procedures

The sampling procedure has been described in our previous publications [15–19]. In brief, femoral blood was collected from the femoral vein at autopsy and potassium fluoride was added to each sample to inhibit bacterial metabolism. The samples were then transported to the national toxicology laboratory, under refrigerated conditions, where they were stored at 4 °C until analysis. During transport (1–3 days) the samples might have been exposed to higher temperatures.

2.3. Analytical methods

All analysis during the time period were conducted using gas chromatography with nitrogen phosphorus detection (GC-NPD). However, while the analytical method stayed the same the limit of quantification (LOQ) differed for certain drugs during the study period (1992–2010). To correct for this, the highest LOQ for each substance during the study period was used as a cut-off. For details regarding the used LOQ, see Table 1.

Table 1
The highest limit of quantification (LOQ) used for each substance during the study period (1992–2010).

Substance	Highest LOQ (μg/g)
Alprazolam	0.02
Amitriptyline	0.05
Carbamazepine	0.5
Citalopram	0.05
Nitrazepam	0.05
Olanzapine	0.05
Oxazepam	0.1
Oxycodone	0.05
Phenytoin	0.5
Quetiapine	0.05
Tramadol	0.05
Verapamil	0.1
Zolpidem	0.05

2.4. Selection of postmortem controls (group C)

The postmortem control cases are a subset of the study population and consist of cases in which one of the substances in this study was detected, but in which the death was unrelated to the substance. Hence, cause and manner of death in the control cases were selected to rule out that the deceased had been incapacitated by the drug detected. Cases in which the manner and/or cause of death with possible damage to a reservoir organ (heart, lungs, liver, intestines) were excluded to avoid contamination of the peripheral blood. This was accomplished by a set of exclusion criteria based on the cause of death code determined by the forensic pathologist who undertook the autopsy (see supplemental Tables S1 and S2).

All cases were then subject to a manual review by multiple independent reviewers. Cases that were hospitalized and subjected to care that might affect blood concentrations were excluded. Suicidal sharp force injuries to smaller vessels were excluded to rule out the effect of slow exsanguination on blood concentrations [21,22]. Cases of drowning were included only if the death was a witnessed suicidal jump into a body of water. Pulmonary embolism was only used as an exclusion criteria with regard to olanzapine and quetiapine as antipsychotics have shown to increase the risk of thromboembolism [23]. In unclear cases, or in cases in which unexpectedly high concentrations were found, the autopsy reports and police reports were reviewed. A case was only included in the control group if consensus between all reviewers had been achieved. Fig. 1 shows a general overview of the selection process.

2.5. Selection of intoxication cases (group A and B)

Intoxication cases were those cases in which one of the included substances was found and in which the cause of death had been attributed to intoxication by pharmaceutical drugs (either alone or in combination with alcohol) by the forensic pathologist responsible for the autopsy. All cases were then subject to a manual review by multiple independent reviewers. This review served to identify the responsible substance(s) in each case. The reference concentrations from the control group (group C) were used as a guideline for the evaluation. As with the control group, cases that were hospitalized and subjected to hospital care that might affect blood concentrations were excluded. Additionally, cases with concomitant gas poisoning and cases where opioid intake could have contributed to the cause of death were excluded. Autopsy reports and police reports were reviewed in unclear cases.

Cases of single substance intoxication in which the contribution of ethanol and/or other substances were excluded were included in group A. Cases of multi-substance intoxication or cases of substance intoxication in combination with alcohol (BAC >1‰) were included in group B. Fig. 1 shows a general overview of the selection process.

The A, B, and C groups are mutually exclusive. A single individual can only contribute to a single group (for example a single case cannot both be classified as A and B). However, a single case can be classified as group B or group C for multiple substances, as the same intoxication case might contain multiple key substances and the same control case can involve more than one of the included substances. Therefore, the number of individuals included in the study does not necessarily match the number of cases in group B and group C.

2.6. Statistical analysis

The concentrations in group A, B and C did not show a normal distribution. Thus, only non-parametric metrics were applied for all statistical analyses.

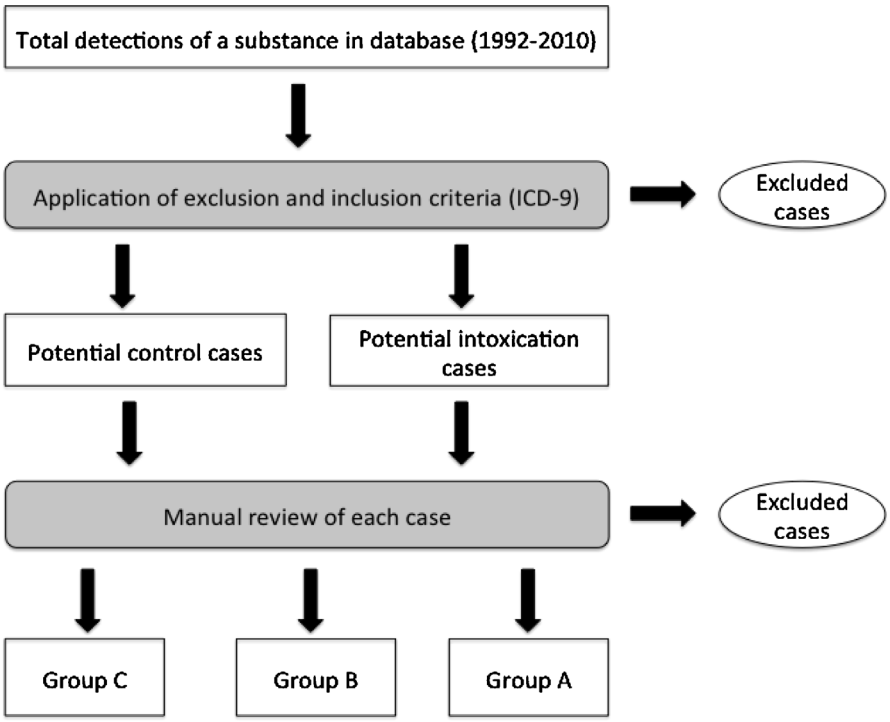

Fig. 1. Schematic workflow from initial database search to finished reference values.
Group A = single-substance intoxication.
Group B = multi-substance (including alcohol) intoxication.
Group C = control cases.

For each substance, the impact of the size of the data set was studied by randomly drawing (with replacement) N samples from the original set of concentrations in each group. N varied from 1 up to 50. For each sample size (N), the sampling process was repeated 1000 times. The sample size (N) was limited to less than the total number of cases in each group for each substance (e.g. a substance with 10 detections in a group would only contribute data to sample sizes <10), in order to reduce the amount of resampling in the random draw. In each such sampling round, the median concentration was registered. The median of all medians and the interquartile range (IQR) of the medians were studied. Further, a normalized IQR was studied by dividing IQR with the median concentration for each value of N.

If a significant difference was observed between the different groups in the original data set of concentrations, the power to detect that difference with smaller sample sizes was studied. Since the data were not normally distributed the Mann-Whitney *U* test was applied to statistically compare the concentrations for the A, B and C groups for each drug. A p-value below 0.05 was considered to be a statistically significant difference and the proportion of such differences was estimated for each sample size, N, based on the 1000 replicates, which represented the power.

All statistical analyses were performed using the R software v. 3.5.0 [24].

2.7. Ethical considerations

The study was approved by the Regional Ethics Review Board in Linköping, Sweden, No 2012/343-31.

Table 2
Number of detections and number of included and excluded cases.

Substance	Total detections in database (1992–2010)	Included cases	Excluded cases	Detections per included case	Excluded cases (% of total detections)
Alprazolam	922	194	728	4.8	79.0
Amitriptyline	1109	459	650	2.4	58.6
Carbamazepine	2526	206	2320	12.3	91.8
Citalopram	4284	1363	2921	3.1	68.2
Nitrazepam	1910	568	1342	3.4	70.3
Olanzapine	473	138	335	3.4	70.8
Oxazepam	823	237	586	3.5	71.2
Oxycodone	188	49	139	3.8	73.9
Phenytoin	622	38	584	16.4	93.9
Quetiapine	133	41	92	3.2	69.2
Tramadol	1737	519	1218	3.3	70.1
Verapamil	398	68	330	5.9	82.9
Zolpidem	1177	337	840	3.5	71.4
Summary					
Total	16,302	4217	12,085		
Mean				5.3	74.7

3. Results

3.1. General

During the study period toxicological analysis in femoral blood was performed in 93 623 autopsy cases, 13 925 (14.9%) of which were positive for one or more of the included substances.

After application of exclusion/inclusion criteria and manual review a total of 3767 individuals were included in the study. Of these 611 (16%) were included in group A, 1355 (36%) in group B and 1801 (48%) in group C. The median age was 54 years across all groups, and was 54 years in group A, 51 years in group B and 57 years in group C. The sex of the deceased was male in 2603 cases (69.1%) across all groups, and the proportion of males was 62.7% in group A, 56.8% in group B, and 80.6% in group C respectively.

Table 2 presents the total number of detections for each substance and the number of detections included in the study for each substance. Note that a single individual may contribute to more than one detection of the included substances. A majority of cases (74.7%) in which the included substances had been detected were removed after application of exclusion criteria and manual review. Across all substances an average of 5.3 detections in the database were needed for each case included in group A, B or C. Carbamazepine and phenytoin presented a high degree of excluded detections (91.8% and 93.9% respectively) and therefore had a higher amount of detections per included case (12.3 and 16.4%,

Table 3
Femoral blood concentrations (μg/g) of substances.

Substance	Group	N	10th percentile[a]	Median	90th percentile[a]	p-value
Alprazolam	A	6	0.17	0.30	0.40	vs B 0.995
	B	113	0.11	0.20	1.60	Vs C <0.001
	C	75	0.02	0.05	0.16	Vs A <0.001
Amitriptyline	A	159	0.90	2.40	6.86	vs B <0.001
	B	175	0.60	1.60	5.46	Vs C <0.001
	C	125	0.10	0.20	0.40	Vs A <0.001
Carbamazapine	A	24	16.6	36.5	80.0	Vs B 0.004
	B	46	12.0	19.5	79.5	Vs C <0.001
	C	136	1.05	5.00	11.0	Vs A <0.001
Citalopram	A	70	1.09	4.45	21.2	Vs B <0.001
	B	377	0.70	1.40	9.04	Vs C <0.001
	C	916	0.10	0.30	0.70	Vs A <0.001
Nitrazepam[b]	A	69	0.40	1.13	3.20	Vs B <0.001
	B	289	0.30	0.60	1.70	Vs C <0.001
	C	210	0.06	0.10	0.28	Vs A <0.001
Olanzapine	A	10	0.29	0.55	2.71	Vs B 0.138
	B	37	0.20	0.40	0.90	Vs C <0.001
	C	91	0.06	0.10	0.20	Vs A <0.001
Oxazepam	A	5	2.00	3.50	6.10	Vs B 0.07
	B	109	1.08	1.70	3.64	Vs C <0.001
	C	123	0.10	0.30	0.90	Vs A <0.001
Oxycodone	A	19	0.50	0.70	5.22	Vs B 0.508
	B	26	0.40	0.65	2.60	Vs C 0.007
	C	4	0.10	0.10	0.40	Vs A 0.007
Phenytoin	A	2	35.0	38.5	43.0	Vs B 0.8
	B	3	21.0	26.0	80.0	Vs C 0.018
	C	33	1.50	6.00	18.2	Vs A 0.041
Quetiapine	A	9	0.70	2.30	22.0	Vs B 0.59
	B	17	0.96	5.60	42.8	Vs C <0.001
	C	15	0.09	0.20	0.66	Vs A <0.001
Tramadol	A	161	2.00	5.10	25.0	Vs B 0.575
	B	184	1.80	4.75	35.7	Vs C <0.001
	C	174	0.10	0.50	1.80	Vs <0.001
Verapamil	A	16	2.10	4.10	12.1	Vs B 0.066
	B	9	0.60	2.80	15.0	Vs C <0.001
	C	43	0.10	0.20	0.50	Vs A <0.001
Zolpidem	A	51	0.60	1.50	5.10	Vs B 0.08
	B	205	0.40	1.00	3.36	Vs C <0.001
	C	81	0.06	0.16	0.50	Vs A <0.001

Group A = single-substance intoxication.
Group B = multi-substance (including alcohol) intoxication.
Group C = control cases.
[a] If n < 10 then minimum and maximum values are presented. Otherwise the 10th and 90th percentiles are presented.
[b] For nitrazepam the concentrations of nitrazepam and its metabolite (7-amino-nitrazepam) were added together for all groups.

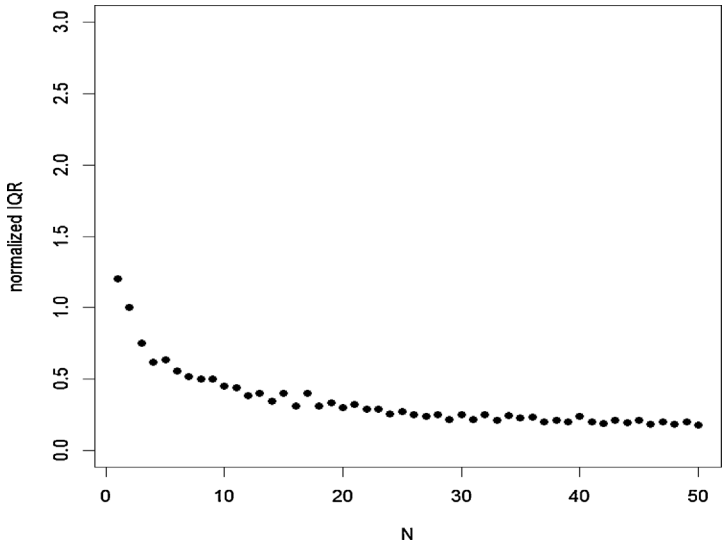

Fig. 2. Typical normalized interquartile range (IQR) of the median values, taking all substances and groups into account, as a function of the sample size, N.

respectively) than average. In contrast, amitriptyline only excluded 58.6% of the total detections and only had 2.4 detections for each included case.

3.2. Postmortem reference values

Table 3 lists the femoral blood concentrations of the included substances subdivided into groups A, B and C.

There were significant concentration differences ($p < 0.05$) between intoxications (group A and group B) and controls (group C) for all substances. However, there was no significant difference between single intoxications (group A) and multi-substance intoxications (group B) for alprazolam, olanzapine, oxycodone, phenytoin, quetiapine, tramadol and verapamil.

Supplemental Table S3 provides examples of the concentration distributions across groups A, B and C for citalopram, carbamazepine and oxazepam.

3.3. Statistical evaluation of number of cases needed for stable reference values in each group

Fig. 2 displays the variation of medians, when the medians have been sampled 1000 times from the cases in each substance and group. Supplemental Tables S4 and S5 present substance specific variations of the corresponding medians. Table S4 presents this data numerically in a 95% interval, and Table S5 presents the normalized IQR from the same data across all substances and groups. The normalized IQR with a sample size of 5 or 10 cases, averaged among all groups, was high (75.3% and 47.0% respectively) compared to 20 or 30 cases in each group (32.3% and 27.7% respectively). It is worth noting that the rate of improvement was reduced above 20–30 cases in each group. In addition, even with 50 cases in each group, there was still a moderate amount of variation (17.5%). As an example, Fig. 3 shows the impact of the sample size on the distribution of the median values for a typical substance (citalopram).

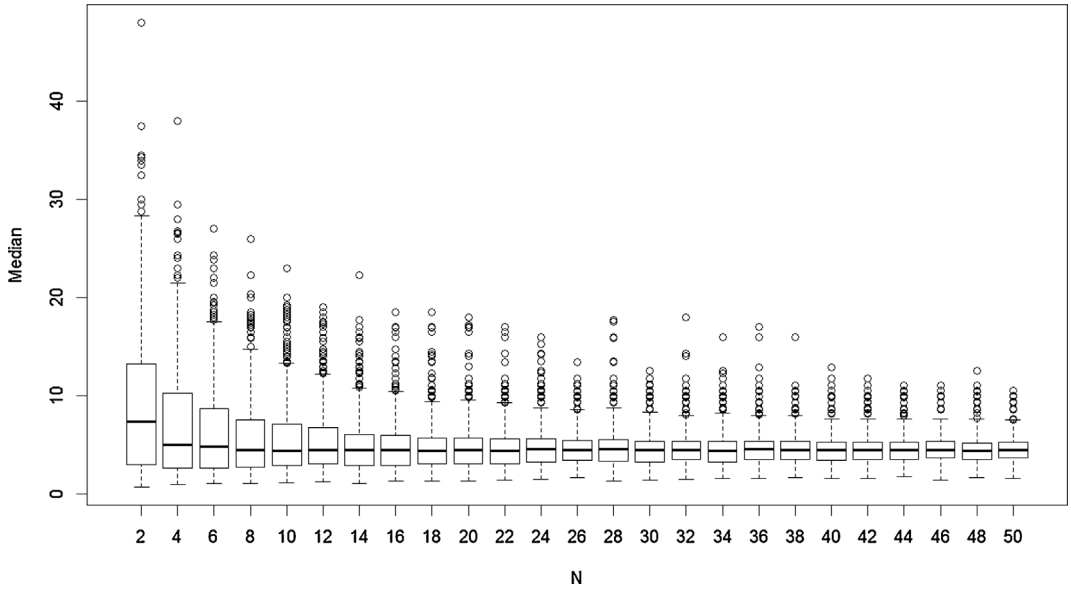

Fig. 3. Boxplot of the distribution of median values when performing replicative sampling (n = 1000) of sample sizes N = 2, 4, 6 . . . 50 for citalopram (Group A cases). The reference median value in the complete data set is 4.5. The deviation from the expected median decreases as the sample size increases. In the boxplot, the whiskers are 1.5 times the interquartile range (IQR), and outliers are shown as circles.

Supplemental Table S6 presents power as a function of the number of cases in each group. As an example, Fig. 4 presents the power to detect differences between groups (A, B and C cases) for the substance citalopram, as a function of the sample size, N. In general, separating group A from group B is often difficult requiring either a large amount of cases, or not being possible at all. However, separating the intoxications (groups A and B) from the postmortem control (group C) requires less cases, with 10 cases providing a high degree of power (>0.95) in all but one case. The exception being separating group B and group C for olanzapine, which requires 17 cases to reach a high degree of power.

4. Discussion

4.1. Cases needed for reliable reference values

In this study we provide reference values for 13 substances using a standardized method of selection and evaluation. A key feature of the current study and the previous studies using the same method [15–19] is not only the presentation of reference concentrations of "postmortem controls" and intoxication cases, but also the number of detections of each substance on which we base our research.

It has previously not been known how many detections of a substance that are needed to produce reliable reference values using our method. Indeed, the impact of population size with regard to postmortem reference values in general has only been discussed briefly [14]. However, it should be noted that the current international recommendations for clinical laboratories is 120 samples for each partition (such as age and sex) in order to both determine the central 95% of the distribution as well as to provide 90% confidence limits on the endpoints [25,26]. In the forensic setting it is very difficult to generate the amount of cases needed that are both reliable, takes into account not only clinical aspects but also forensic aspects such as postmortem redistribution and the postmortem interval. The present study provides a first step in discussing sample size and reliable reference concentrations in postmortem toxicology.

From Table 2 it can be seen that approximately 74.7% of the detections for a given substance is lost when inclusion/exclusion criteria are applied or upon manual review. Once detections have been ordered into groups A, B and C there is a need to ensure that the number of included cases are sufficient for a reliable reference value. As can be seen in Table S6 ten cases (n = 10) are in almost all cases sufficient for high power (>0.95) when separating intoxications (group A and B) from non-intoxications (group C). Separating single intoxications (group A) from multi-intoxications (group B) is much harder, and is not always possible even with 50 cases in each group. This is to be expected since the underlying difference in antemortem action between the intoxication groups is slight, with the same modus of taking a large amount of available medication prevalent in both groups.

Over multiple samplings from the larger population it can be seen that the variation of the normalized IQR decreases with increased sample size, and that the rate of improvement stabilizes around a sample size of 20–30. Of special note is the very high variation in low sample sizes (<5), indicating that for substances in this group much can be gained by increasing the number of cases.

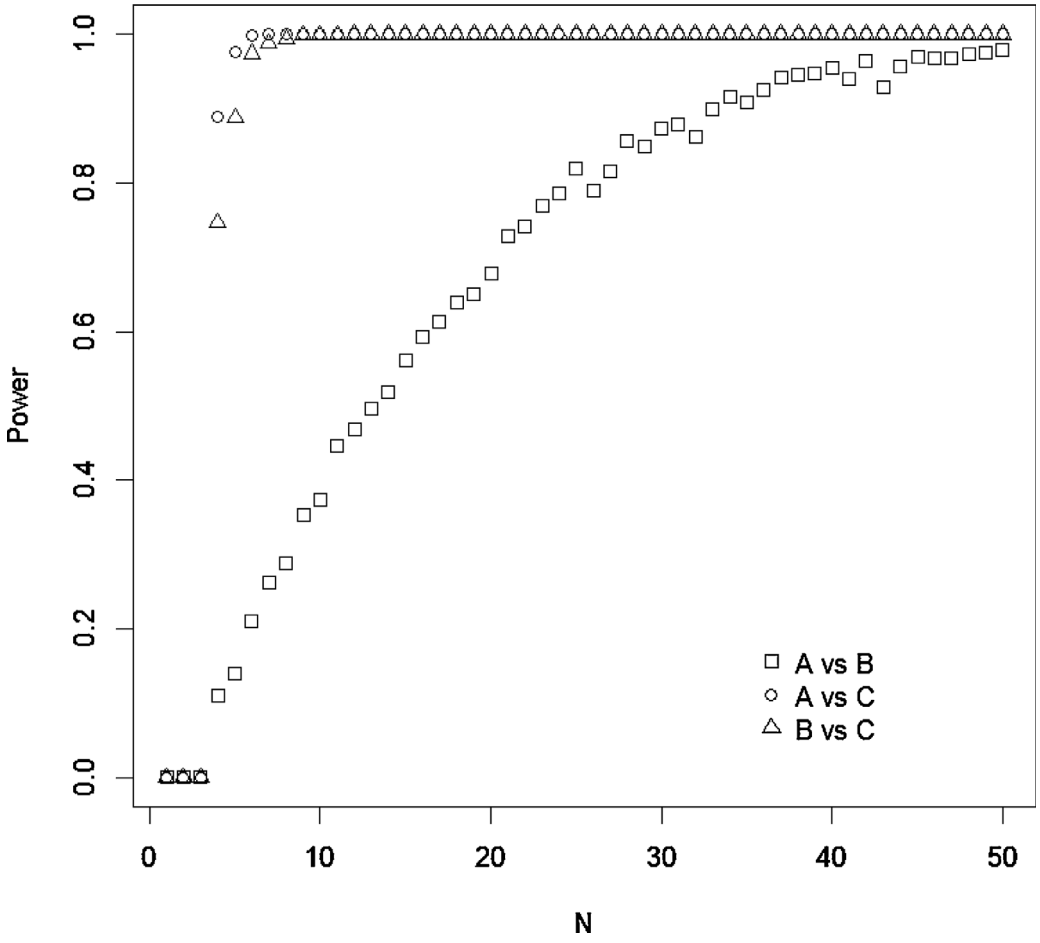

Fig. 4. The power to detect a significant difference, between groups (A, B and C cases) for the substance citalopram, as a function of the sample size, N.

How much variance that can be accepted in a sample with regard to a larger population is difficult to address, and is likely dependent on the extent of overlap between normal and toxic concentrations.

4.2. Comparison with results from our previous studies

All substances except oxycodone and tramadol have previously been evaluated using our method in one of our previous publications [15–17,19]. However, except for olanzapine and quetiapine, which in our previous study [19] already contained cases from 1992–2010, the longer study period in the present study has resulted in an increased number of included cases. Comparing the present results with the earlier iterations of reference values the corresponding median values are the same, or with only slight variation (≤10%), for alprazolam (group A and group C), amitriptyline (all groups), citalopram (group B and group C), nitrazepam (all groups), olanzapine (all groups), oxazepam (all groups), verapamil (group A and group C) and zolpidem (group A and group C). All these variations are well within the expected range based on the sample size, as described above. However, in the case of some substances and groups there are more pronounced differences.

In the case of alprazolam (group B) the median concentration has increased by 25.0% (0.16 μg/g to 0.20 μg/g) between the previous and the present study. The number of cases increased from 67 to 113. The median of group B alprazolam concentrations is slightly lower than expected based on the sample size of our previous study [16], but the increase to 0.20 μg/g in the present study is within the expected range of variation. With regard to citalopram, the median of group A has decreased 31.5% (from 6.50 μg/g to 4.45 μg/g) and the number of included cases increased from 50 to 70. However, based on the sample size of our original study [17], this variation can be expected. Of note is that the standardized IQR for alprazolam group B and citalopram group A at a sample size of 50 was high (40.0% and 36.0% respectively) compared to the average (17.5%) suggesting a large spread of the data even at large sample sizes.

For carbamazepine the median of group A decreased 18.9% (from 45.0 μg/g to 36.5 μg/g), the median of group B increased 39.3% (from 14.0 μg/g to 19.5 μg/g) when compared to the original study [15]. The number of cases has changed significantly in group A (from 7 to 24) and group B (from 9 to 46).

In the case of phenytoin there are shifts in all groups compared to the original study [15]. The median of group A decreased 10.5% (from 43.0 μg/g to 38.5 μg/g), the median of group B decreased 67.5% (from 80 μg/g to 26 μg/g) and the median of group C increased by 20% (from 5 μg/g to 6 μg/g). However, the number of cases included in group A and group B are still small (2 and 3 cases, respectively), which can explain the variability of the median. Group C showed a more significant increase in the number of cases compared to the original study (from 14 to 33). For verapamil the group B median concentration increased 47.4% (1.9 μg/g to 2.8 μg/g) in the present study compared to our previously published values [15].

However, for zolpidem the median of group B increased by 11.5% (from 0.9 μg/g to 1 μg/g) and the number of cases increased from 148 to 205 compared with the previous study [16]. In carbamazepine group C the median of group C increased 11.1% (from 4.5 μg/g to 5.0 μg/g) and the number of included cases increased from 56 to 136 compared with the previous study [15]. These cases show that slight variation can be expected even with large samples, however changes on this small scale might also be explained by inherent measurement uncertainty.

With regard to quetiapine and olanzapine there are slight changes in the number of cases compared to our previous study. Quetiapine lost one case in group B (18 to 17) and group C (16 to 15) and olanzapine lost three cases in group C (94 to 91) [19], which has impacted the median concentrations of quetiapine in group B (from 3.85 μg/g to 5.60 μg/g) and group C (from 0.15 μg/g to 0.20 μg/g), even though the evaluated material is the same. With regard to quetiapine group A the median has decreased (from 8.1 μg/g to 2.3 μg/g) even though the 10th/90th percentiles and number of cases are unchanged. These variations are expected based on sample size. The variations in numbers and medians are probably due to cases being evaluated differently because of additional reference material (such as updated reference concentrations regarding co-detections) being available for the present study.

In general, the larger variations of the median (≥20%), with the exception of substances and groups with high standardized IQR at high sample sizes (alprazolam group B and citalopram group A), can be explained by the inherent variability at the lower sample sizes in the previous studies. However, care must be taken even at large sample sizes due to some substances having a large concentration spread.

4.3. Comparison with other reviews and compilations

In the scientific literature there are numerous publications in which postmortem concentrations of the included substances are presented. The following section will discuss previously published material subdivided by the type of study.

We have chosen not to review individual case reports or small case series since their individual value with regard to postmortem reference concentrations can be considered low [27]. The review below focuses on larger compilations.

Descriptive compilations - There are a few compilations based on a descriptive representation of detections of a substance, such as the compilations by Launiainen and Ojanperä [12], Ketola and Ojanperä [13] and Jones [11], with high concentration percentiles (90th, 95th and 97.5th) presented separately. One advantage of this approach is that there are often a high number of detections for each substance as no detections are excluded, which increases the reliability of the results. However, in this type of presentation there is no distinction between intoxication and non-intoxication cases, which is a weakness since it is difficult to know where one group ends and another begins. These compilations instead show if a given concentration can be considered high in relation to the group. It is important to note that this approach is dependent on the LOQ for each substance, a high LOQ can shift the population towards intoxication cases while a low LOQ can do the opposite. The concentrations seen in the high percentiles in all three studies align well with the concentrations seen in our intoxication groups, with the 90th percentile in the above studies being within the 10th/90th percentile range of our group A cases for about half the substances (except alprazolam, carbamazepine, nitrazepam, oxazepam, oxycodone, phenytoin and quetiapine) and our group B cases for most substances (except nitrazepam, oxazepam, phenytoin and quetiapine).

There are, in addition, a few compilations that present large tabulations of stated therapeutic, toxic and “lethal” concentrations [28,29]. In using these compilations a reviewer must proceed with caution. While the list of included substances is extensive, the postmortem reliability of the reference values are not always clear (as data regarding sample site and the number of detections on which the data is based is lacking), which limits their usefulness in postmortem casework.

Evaluated compilations - In forensic practice it is not only important to confirm intoxications but also to be able to rule them out. Thus, there need to be reference concentrations regarding both postmortem “normal” (i.e. which have not contributed to the cause of death) and “toxic” concentrations. Extracting these cases from a database and reviewing whether or not they are

intoxications can do this. Using a combination of details available in case reports and the extensive database search seen in the database compilations, aims to provide the best of both worlds. There are some publications that have taken some variant of this approach, in which a case series has been subdivided into natural deaths and intoxication groups, which will be discussed below. As with case reports, it is important to remember that studies with a low number of cases in each group has to be interpreted carefully.

Studies by Jones and Holmgren [30,31] provide reference values for non-intoxication causes of death for alprazolam (median 0.05 μg/g) and zolpidem (median 0.13 μg/g) that are well in line with our group C concentrations (median 0.05 μg/g and 0.16 μg/g respectively). However, the concentrations of alprazolam and zolpidem found in intoxication cases (median 0.06 μg/g and 0.3 μg/g respectively) are difficult to evaluate since it is not clear if the substances are assessed as contributing to the intoxication or not. Their lower median concentrations compared to the present study (group B median 0.2 μg/g and 1.0 μg/g for alprazolam and zolpidem respectively) could be explained by inclusion of intoxications in which the presence of alprazolam and zolpidem are merely incidental.

A study by Wolf et al. [32] concerning alprazolam related deaths between 2001 and 2003 compiled reference values regarding both combined drug intoxications (mean 0.10 μg/mL, range 0.01–1.20 μg/mL) and non-intoxication deaths (mean 0.04 μg/mL, range <0.002–0.08 μg/mL) which align well with our own (group B median 0.2 μg/g and group C median 0.05 μg/g). However, their single intoxications deaths present lower concentrations (0.01 μg/mL and 0.03 μg/mL respectively) than our single intoxications (median 0.3 μg/g). The number of cases in the single intoxication group are low (n = 2), and thus must be interpreted carefully, which also applies to our own data since the present study only includes six single intoxications. The deaths in their natural group (n = 12) and combined drug toxicity group (n = 87) are more robust.

With regard to citalopram, a study by Jonasson and Saldeen [33] presents higher concentrations in the cases in which citalopram had not contributed to death (range: 0.7–1.5 μg/g, median: 1.08 μg/g), than those found in our group C (median 0.3 μg/g). This is interesting since the study used the same database as the present study. However, they excluded cases in which the concentration of citalopram was below 0.7 μg/g, which explains the higher range and median concentration in their study. The concentrations in single citalopram intoxications (n = 5, median 11 μg/g) and contributing citalopram intoxications (n = 9, median 1 μg/g) does not align with our group A and group B concentrations (median 4.45 μg/g and 1.4 μg/g respectively) directly. However, the variation corresponds to that seen in low sample sizes in the present study.

In several publications regarding oxycodone [34–37] there are cases in which oxycodone has contributed to intoxication deaths with a concentration below our group B and group A concentrations. In the study by Cone et al. [35] they present medians in combined toxicity groups (0.19–0.58 μg/mL) which is lower than our own data (group B median 0.65 μg/g), however the blood in their study was collected from various anatomical sites which impacts interpretation. In the study by Wolf [37] the range in single oxycodone intoxications (0.21–4.71 mg/L) and combined drug toxicity (0.025–17.5 mg/L) include cases more in line with our group C median (0.1 μg/g). Baker and Jenkins [34] present a median concentration in intoxication deaths (0.56 mg/L) which are an agreement with our intoxication cases (group B median 0.65 μg/g and 10th percentile group A 0.5 μg/g). However, their range (0.01–36.54 mg/L) includes cases with low concentrations closer to our group C (median 0.1 μg/g) where circumstantial information probably has played a role for the diagnosis. Baker and Jenkins also present a large body of non-intoxications deaths (n = 135), which greatly exceeds our own (n = 4). In light of our results regarding variation in small samples their median (0.26 mg/L) is probably closer to the "truth" than our own (0.1 μg/g). Similarly the study by Ogle [36] presents large groups of deaths where oxycodone was a contributing factor (n = 117, mean 0.48 mg/L) and incidental finding (n = 38, mean 0.16 mg/L), which should be taken into account when evaluating the median and percentiles in our cases. As mentioned above, there are cases in which oxycodone has been classified as contributing which corresponds to our group C concentrations, which is probably due to circumstantial factors in the individual case.

A study by Pilgrim et al. [38] regarding quetiapine reports concentrations in quetiapine associated deaths (n = 114, median 0.7 mg/L, range 0.02–110 mg/L), which is well below our group B concentrations (median 5.6 μg/g). However, their material is not only comprised of intoxications in which quetiapine has played a key role, but also include mentions of quetiapine in autopsy- or police reports as well as in coroners findings which broadens their material and could explain their lower median concentration. With regard to natural deaths their data (n = 60, median 0.25 mg/L) aligns well with our own (median = 0.2 μg/g) and the variation is within expected limits.

Numerous publications from Denmark [39–42] use a similar classification system as ours, with slight differences in how the control group is selected. In a study concerning olanzapine [39] their control group median is lower than ours (0.064 μg/g compared to 0.1 μg/g) and slightly outside the variation we would expect based on our data, which is probably due to our LOQ of 0.05 μg/g pushing our median upwards. With regard to multi-drug intoxications our results align well, and while their study includes few single drug intoxication cases (n = 2) their results (1.8 μg/g and 3.6 μg/g) are comparable to the 10th/90th percentile range of our group A cases (0.29–2.71 μg/g). In the study concerning citalopram [40] their results align well with our own with regard to controls (0.31 μg/g and 0.3 μg/g respectively) and mixed intoxications (1.3 μg/g and 1.4 μg/g respectively), and their single drug intoxication case (3.0 μg/g) is well within the 10th/90th percentile range of our group A cases (1.09–21.23 μg/g). The variations being within the expected range based on the sample size data of the present study. In the study by Skov et al. concerning quetiapine [41] the median of their control group is lower than ours (0.15 μg/g compared to 0.2 μg/g), which, again, is probably because of our higher LOQ (0.05 μg/g). The median of their mixed intoxication group (n = 5, 3.19 μg/g) is well below ours (n = 17, 5.6 μg/g) but within the variation based on sample size we see in our population with a similar sample size. Their single intoxication case (8.99 μg/g) is well within the 10th/90th percentile range of our single intoxication group (0.7–22.0 μg/g). In a second study featuring, among others, quetiapine [42] a noteworthy detail is that the median in the mixed intoxication group (n = 7, 1.3 μg/g) is lower than in the previous study [41] and diverging more than expected from the one in the present study based on the sample size.

In general, this type of approach provides reference values which can both aid in confirming and excluding intoxications. However, it is important to know how intoxications and non-intoxications have been classified when using the reference data and comparing it with other studies. Also, as mentioned above, the number of cases in each group can limit the applicability of the reference value to be generalized. As always, reference values must be used with caution and individual case circumstances must be taken into account.

4.4. The importance of control cases

The strengths of our strategy is that it produces not only toxic concentrations but also "normal" (i.e. which have not contributed

to the cause of death) postmortem reference concentrations. In our experience having a reference of which concentrations can be considered "normal" greatly aids in ruling out intoxications. Our control group (group C) comprises, as mentioned in section 2.4, cases were the death was unrelated to intake of drugs. In practice this means cases were the cause of death cannot have been influenced by drugs and that the deceased should have been alive and fully capable to act, had not something external (such as a suicidal hanging or gunshot) occurred. The control group is further selected by removing cases in which damage to a drug reservoir organ [7,43] might contaminate the femoral blood used in testing.

As has been mentioned in section 4.1 the selection process used in this study discards a large amount of cases. Thus, group C is not merely cases with another cause of death than intoxication, but rather a highly selected group in which as many uncertainties as possible has been ruled out. The advantage of this approach is that the resulting group C can be used with high confidence.

4.5. Statistical considerations

In this study we used several statistical approaches in order to analyze the reliability of obtained reference values in regards to the sample size. The main objective was to be able to detect trends among the studied substances, for example how well a small sample size corresponds to the "true" distribution of a given substance and case type. By performing replicative sampling it was possible to study how the concentrations in small sample sets varied between different samples. Based on this it was possible to analyze how representative a smaller sample set is in relation to the full data set, and to see possible differences among different classes of substances. It should be noted that we only used non-parametric statistical methods (due to the non-normal distribution of the concentration data), which have the general disadvantage of a lower power compared with parametric statistical tests. This shortage was observed for the estimated power when N = 1–3 (Fig. 4). In future studies it would be valuable to perform more "in depth" studies for specific substances. For example by analyzing and fitting more detailed distributions in order to apply alternative statistical tests to further increase the specificity.

4.6. Limitations

It is well known that certain substances, such as opioids and benzodiazepines, are subject to the development of tolerance in an individual [44]. Our method has not differentiated tolerant and non-tolerant individuals. In order to evaluate toxicological findings were tolerance might be a factor concurrent hair analysis can provide additional information. Absence of a drug in the inner segment(s) of a hair sample suggests a lack of tolerance [45–47], which is of importance when interpreting the toxicological result in blood.

Many substances have active metabolites (in the present study this is the case for amitriptyline, olanzapine, oxazepam, oxycodone and tramadol), which may have a direct pharmacological effect or induce other toxicity [48,49]. Chiral structures can impact the effect of a drug [50] and genetic mutations affect both the metabolism and effect of drugs [51,52]. In our studies we have only evaluated parent compounds without regard to active metabolites, genetics and chiral structures. Thus, these factors must be evaluated on a case-by-case basis and taken into account when using reference values.

With regard to multi-drug intoxications there is also the complicated problem of to which extent drugs interact with one another in a detrimental manner, producing intoxication greater than the sum of its parts. This has to be kept in mind when there are multiple drugs of the same class present, or multiple drugs that act on the same physiologic system (such as respiration or consciousness) [44]. In our material it is difficult to evaluate the extent of these interactions.

Table 4
Key numerical factors.

Numerical factors	
Average number of detections needed to provide a single case in group A, B or C after evaluation	5.3
Number of cases in each group needed to reach a power of >0.95 when separating intoxications from controls	10[a]
Normalized interquartile range (IQR) of medians across all groups with a sample size of:	
5	75.3%
10	47.0%
20	32.3%
30	27.7%
40	21.8%
50	17.5%

[a] Olanzapine group B compared to group C has been excluded as an outlier, which require 17 cases to reach a power of >0.95.

4.7. Conclusions and recommendations

Postmortem toxicology is a complex matter with multiple sources of possible error and often a paucity of information regarding any intake of drugs. In this environment it is important to provide the best tools for evaluation possible, while at same time being mindful of the pitfalls and limitations of the postmortem setting.

The reference values presented in this study, and our previous publications, along with the statistical information fill an important role in this regard. Table 4 present a summary of numerical benchmarks for using our method.

In conclusion, while we believe we are being as diligent as possible in the compilation of our results, the variation in small and medium sample sizes urges caution when using the current and other reference compilations. Case circumstances and postmortem factors must thus always be taken into account. Future research should strive to use databases with as many detections as possible, in order to minimize the impact of variation of small sample sizes.

CRediT authorship contribution statement

C. Söderberg: Conceptualization, Investigation, Writing - original draft, Writing - review & editing. **A. Tillmar:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - review & editing. **A. Johansson:** Investigation, Writing - review & editing. **E. Wernvik:** Investigation, Writing - review & editing. **A.K. Jönsson:** Conceptualization, Investigation, Data curation, Writing - review & editing. **H. Druid:** Supervision, Conceptualization, Methodology, Investigation, Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgements

This study was supported by the Swedish National Board of Forensic Medicine.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at https://doi.org/10.1016/j.forsciint.2020.110292.

References

[1] M. Eddleston, Patterns and problems of deliberate self-poisoning in the developing world, QJM 93 (11) (2000) 715–731.
[2] V. Ajdacic-Gross, et al., Methods of suicide: international suicide patterns derived from the WHO mortality database, Bull. World Health Organ. 86 (9) (2008) 726–732.
[3] K.W. Simonsen, et al., Fatal poisoning in drug addicts in the nordic countries in 2012, Forensic Sci. Int. 248 (2015) 172–180.
[4] M. Hickman, et al., London audit of drug-related overdose deaths: characteristics and typology, and implications for prevention and monitoring, Addiction 102 (2) (2007) 317–323.
[5] E. Groot, et al., Drug toxicity deaths after release from incarceration in Ontario, 2006-2013: review of coroner's cases, PLoS One 11 (7) (2016) e0157512.
[6] D.J. Pounder, G.R. Jones, Post-mortem drug redistribution—a toxicological nightmare, Forensic Sci. Int. 45 (3) (1990) 253–263.
[7] A.L. Pelissier-Alicot, et al., Mechanisms underlying postmortem redistribution of drugs: a review, J. Anal. Toxicol. 27 (8) (2003) 533–544.
[8] G. Skopp, Postmortem toxicology, Forensic Sci. Med. Pathol. 6 (4) (2010) 314–325.
[9] G. Skopp, Preanalytic aspects in postmortem toxicology, Forensic Sci. Int. 142 (2-3) (2004) 75–100.
[10] E. Saar, et al., Assessment of the stability of 30 antipsychotic drugs in stored blood specimens, Forensic Sci. Int. 215 (1-3) (2012) 152–158.
[11] A.W. Jones, A. Holmgren, Concentration distributions of the drugs most frequently identified in post-mortem femoral blood representing all causes of death, Med. Sci. Law 49 (4) (2009) 257–273.
[12] T. Launiainen, I. Ojanpera, Drug concentrations in post-mortem femoral blood compared with therapeutic concentrations in plasma, Drug Test. Anal. 6 (4) (2014) 308–316.
[13] R.A. Ketola, I. Ojanpera, Summary statistics for drug concentrations in post-mortem femoral blood representing all causes of death, Drug Test. Anal. (2019).
[14] O.H. Drummer, Post-mortem toxicology, Forensic Sci. Int. 165 (2-3) (2007) 199–203.
[15] H. Druid, P. Holmgren, A compilation of fatal and control concentrations of drugs in postmortem femoral blood, J. Forensic Sci. 42 (1) (1997) 79–87.
[16] A.K. Jonsson, et al., Sedative and hypnotic drugs—fatal and non-fatal reference blood concentrations, Forensic Sci. Int. 236 (2014) 138–145.
[17] M. Reis, et al., Reference concentrations of antidepressants. A compilation of postmortem and therapeutic levels, J. Anal. Toxicol. 31 (5) (2007) 254–264.
[18] C. Soderberg, et al., Reference values of lithium in postmortem femoral blood, Forensic Sci. Int. 277 (2017) 207–214.
[19] C. Soderberg, et al., Antipsychotics - postmortem fatal and non-fatal reference concentrations, Forensic Sci. Int. 266 (2016) 91–101.
[20] H. Druid, P. Holmgren, P. Löwenhielm, Computer-assisted systems for forensic pathology and forensic toxicology, J. Forensic Sci. 41 (5) (1996) 830–836.
[21] F.C. Kugelberg, et al., Influence of blood loss on the pharmacokinetics of citalopram, Forensic Sci. Int. 161 (2–3) (2006) 163–168.
[22] F.C. Kugelberg, P. Holmgren, H. Druid, Codeine and morphine blood concentrations increase during blood loss, J. Forensic Sci. 48 (3) (2003) 664–667.
[23] A.K. Jonsson, et al., Antipsychotics associated with pulmonary embolism in a Swedish medicolegal autopsy series, Int. Clin. Psychopharmacol. 23 (5) (2008) 263–268.
[24] R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 2012.
[25] G.L. Horowitz, Reference intervals: practical aspects, EJIFCC 19 (2) (2008) 95–105.
[26] G.L. Horowitz, et al., Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline, third edition, Clinical Laboratory and Standards Institute, 2010.
[27] F. Musshoff, Case reports in forensic toxicology, Drug Test. Anal. 5 (9-10) (2013) 723–724.
[28] C.L. Winek, et al., Drug and chemical blood-level data 2001, Forensic Sci. Int. 122 (2-3) (2001) 107–123.
[29] M. Schulz, et al., Therapeutic and toxic blood concentrations of nearly 1,000 drugs and other xenobiotics, Crit. Care 16 (4) (2012) R136.
[30] A.W. Jones, A. Holmgren, Concentrations of alprazolam in blood from impaired drivers and forensic autopsies were not much different but showed a high prevalence of co-ingested illicit drugs, J. Psychopharmacol. 27 (3) (2013) 276–281.
[31] A.W. Jones, A. Holmgren, Concentrations of zolpidem and zopiclone in venous blood samples from impaired drivers compared with femoral blood from forensic autopsies, Forensic Sci. Int. 222 (1-3) (2012) 118–123.
[32] B.C. Wolf, et al., Alprazolam-related deaths in Palm Beach County, Am. J. Forensic Med. Pathol. 26 (1) (2005) 24–27.
[33] B. Jonasson, T. Saldeen, Citalopram in fatal poisoning cases, Forensic Sci. Int. 126 (1) (2002) 1–6.
[34] D.D. Baker, A.J. Jenkins, A comparison of methadone, oxycodone, and hydrocodone related deaths in northeast Ohio, J. Anal. Toxicol. 32 (2) (2008) 165–171.
[35] E.J. Cone, et al., Oxycodone involvement in drug abuse deaths. II. Evidence for toxic multiple drug-drug interactions, J. Anal. Toxicol. 28 (7) (2004) 616–624.
[36] A. Ogle, et al., Clinical history and characteristics of persons with oxycodone-related deaths in Hillsborough County, Florida in 2009, Forensic Sci. Int. 223 (1-3) (2012) 47–52.
[37] B.C. Wolf, et al., One hundred seventy two deaths involving the use of oxycodone in Palm Beach County, J. Forensic Sci. 50 (1) (2005) 192–195.
[38] J.L. Pilgrim, O.H. Drummer, The toxicology and comorbidities of fatal cases involving quetiapine, Forensic Sci. Med. Pathol. 9 (2) (2013) 170–176.
[39] M. Nedahl, S. Johansen, K. Linnet, Reference brain and blood concentrations of olanzapine in postmortem cases, J. Anal. Toxicol. (2018).
[40] M. Nedahl, S.S. Johansen, K. Linnet, Reference Brain/Blood concentrations of citalopram, duloxetine, mirtazapine and sertraline, J. Anal. Toxicol. 42 (3) (2018) 149–156.
[41] L. Skov, S.S. Johansen, K. Linnet, Postmortem quetiapine reference concentrations in brain and blood, J. Anal. Toxicol. 39 (7) (2015) 557–561.
[42] L. Skov, S.S. Johansen, K. Linnet, Postmortem femoral blood reference concentrations of aripiprazole, chlorprothixene, and quetiapine, J. Anal. Toxicol. 39 (1) (2015) 41–44.
[43] C. Sastre, et al., Post mortem redistribution of drugs: current State of knowledge, Curr. Pharm. Des. 23 (36) (2017) 5530–5541.
[44] T. Richardson, Pitfalls in forensic toxicology, Ann. Clin. Biochem. 37 (Pt 1) (2000) 20–44.
[45] H. Druid, et al., Evaluation of the role of abstinence in heroin overdose deaths using segmental hair analysis, Forensic Sci. Int. 168 (2-3) (2007) 223–226.
[46] G. Hoiseth, et al., Is hair analysis useful in postmortem cases? J. Anal. Toxicol. 42 (1) (2018) 49–54.
[47] G.A. Cooper, et al., Society of hair testing guidelines for drug testing in hair, Forensic Sci. Int. 218 (1-3) (2012) 20–24.
[48] D.E. Drayer, Pharmacologically active metabolites of drugs and other foreign compounds. Clinical, pharmacological, therapeutic and toxicological considerations, Drugs 24 (6) (1982) 519–542.
[49] M. Pirmohamed, N.R. Kitteringham, B.K. Park, The role of active metabolites in drug toxicity, Drug Saf. 11 (2) (1994) 114–144.
[50] S.W. Smith, Chiral toxicology: it's the same thing . . . only different, Toxicol. Sci. 110 (1) (2009) 4–30.
[51] U.A. Meyer, U.M. Zanger, Molecular mechanisms of genetic polymorphisms of drug metabolism, Annu. Rev. Pharmacol. Toxicol. 37 (1997) 269–296.
[52] F. Musshoff, U.M. Stamer, B. Madea, Pharmacogenetics and forensic toxicology, Forensic Sci. Int. 203 (1-3) (2010) 53–62.