FORENSIC SCIENCE SEMINAR

ISSN 2157-118X
Volume 4 Number 1 December 2014
Indexed by Google Scholar, OCLC, Ulrich

Published by ZolCat Academic House
Forensic Science Seminar

6

ISSN 2157-118X
December 2014 Volume 4 Number 1

Forensic Science Seminar

Single Issue Price $150.00 USD
Peer Reviewed | Limited Open Access

www.zolcat.com

fss.xxyy.info

ZolCat
www.zolcat.com
ISSN 2157-118X
9 772157 118006 >

Contents

Review Articles

Original Research Paper

Forensic Science Seminar (ISSN 2157-118X)
Editorial Department E-mail: fss@xxyy.info
Official Website http://fss.xxyy.info/
Publisher *ZolCat Academic House*, Brooklyn, NY, 11220.
Indexed by *Google Scholar, OCLC (659514459), Ulrich.*
Impact Factor = 0.1 (2013).
Sponsored by *the Library of Congress.*

Received 2 October 2014
Received in revised form 5 November 2014
Accepted 29 November 2014
Available online 16 December 2014
Peer Reviewed
2157118X-4.1.R1

The Confirmation on Method GA/T 1073-2013

XU Pian [*, a], ***MENG Qingfang*** [a], ***ZHENG Jing*** [a]

a Chongqing Engineering Research Center for Criminal Investigation Technology, Institute of Forensic Science, Chongqing Public Security Bureau, Chongqing 400021, China.

* Corresponding author:XU Pian, Ph.D. Tel: +86-023-63751658. E-mail: xupian555@126.com

Abstract *Objective* To confirm whether the examination method GA/T 1073-2013 for determining the content of ethanol in blood by headspace gas chromatography is feasible under the condition of our laboratory and can cope with large batches of cases in practise. *Methods* This paper set up experiments about precision, accuracy and repeatability to test some samples with high, medium and low concentration in the range of calibration curve. Through the results from the aboved experiments, the relative standard deviations were caculated. *Results* The relative standard deviations of the precision, the accuracy and the repeatability are all less than 5%. The average recovery of standard addition was founded to be in the range of 96.50%-99.42%. *Conclusion* All the results indicated that the method was precise, accurate and repeatable.

Keywords: Forensic science, Confirmation, Headspace gas chromatography, ISTD, Blood alcohol content.

1 Introduction

Ethanol (common name is alcohol and molecular formula is C_2H_6O) is widely used in daily life. Therefore, the determination of blood alcohol content is a routine test of forensic indetification and clinical diagnosis. Now the methods to detect the content of ethanol in blood are gas chromatography, enzyme and chemical process, etc[1-3]. This research center has adopted the analysis method for ethanol concentration in blood, which is the professional standard of public security of the PRC published and carried out in 2009[4]. The method, short for GA/T 842-2009, is just aimed at checking the content of ethanol in blood, which is reliable and can meet requirement in practical work. While in 2013, another professional standard namely GA/T 1073-2013 was issued and implemented by the department of public security of the PRC[5]. Its complete title is HS-GC examination methods for ethanol, methanol, n-propanol, aldehyde, acetone, isopropanol and n-butanol in blood and urine. It is clearly that GA/T 1073-2013 has a wide range of detection and can test several compounds which include ethanol. So this paper set up a series of experiments about specificity, precision, repeatability and accuracy to demonstrate whether GA/T 1073-2013 is applicable to determine the content of ethanol in blood under the current condition of our laboratory and can deal with large quantities of cases with results accurate and stable.

2 Materials and Methods

2.1 Instruments and Equipments

GC-2010 Plus; FID (Flame Ionization Detector); Guartz Capillary Column (Rtx-WAX, 30 m×0.25 mm×0.25μm); GCsolution Ver 2.32 workstation; DANI HSS 86.50 Headspace Autosampler; Pipettor; Headspace vials with matched caps; Sealing Plier.

2.2 Reagents

The standard substances: 99.9% anhydrous ethanol (chromatographically pure), 99.5% tertiary-butanol (analytically pure) as the internal standard. Make the preparation of 8000 mg/100 mL

ethanol as stock solution, store it in the refrigerator of 4°C. Dissolve appropriate tertiary-butanol into ultrapure water in the volumetric flask, preparing 4 mg/100 mL as working solution of the internal standard.

2.3 Instrumental Working Conditions

Headspace autosampler: oven temperature: 65°C; manifold temperature: 105°C; delivering tube temperature: 110°C; incubation time: 10.0 min.

Chromatographic column temperature: 40°C; flow velocity: 4.0 mL/min; purge flow velocity: 4.0 mL/min; H2 flow velocity: 40.0 mL/min; Air flow velocity: 400.0 mL/min; make-up gas flow velocity: 30.0 mL/min; injection port temperature: 150°C; detector temperature: 250°C; split ratio: 20.0.

2.4 Experimental Methods

2.4.1 Drawing of calibration curve

Make a preparation of 100 mg/100 mL, 200 mg/100 mL, 500 mg/100 mL, 800 mg/100 mL, 1000 mg/100 mL, 2000 mg/100 mL, 3000 mg/100 mL ethanol standard series. Add 0.09 mL blank blood respectively to 0.01 mL standard series solution and 0.5 mL internal standard working solution in the headspace vials, seal and blend them and wait for determination.

2.4.2 Sample preparation

Add 0.1 mL blood to be measured to 0.5 mL the internal standard working solution in the vial, seal and blend it and wait for detection. Fetch 0.1 mL blank blood and 0.5 mL

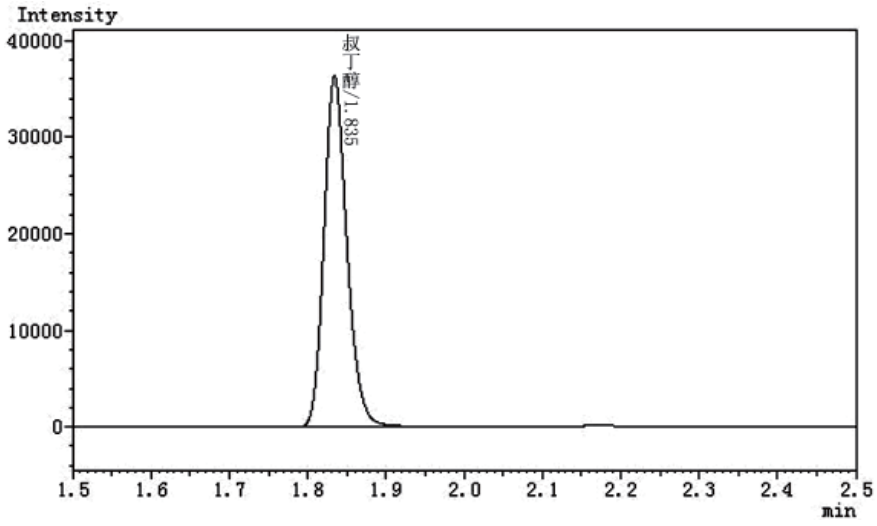


Fig 1. Chromatogram of Blank Control.

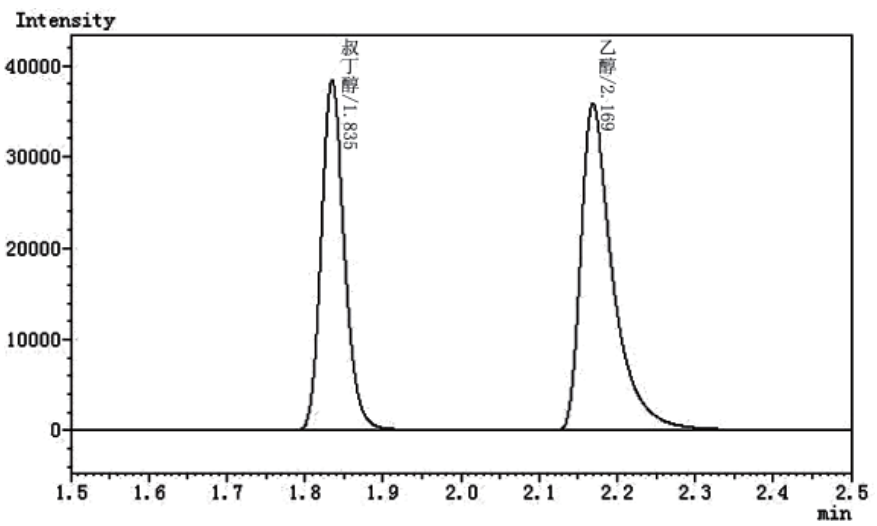


Fig 2. Chromatogram of Ethanol Standard.

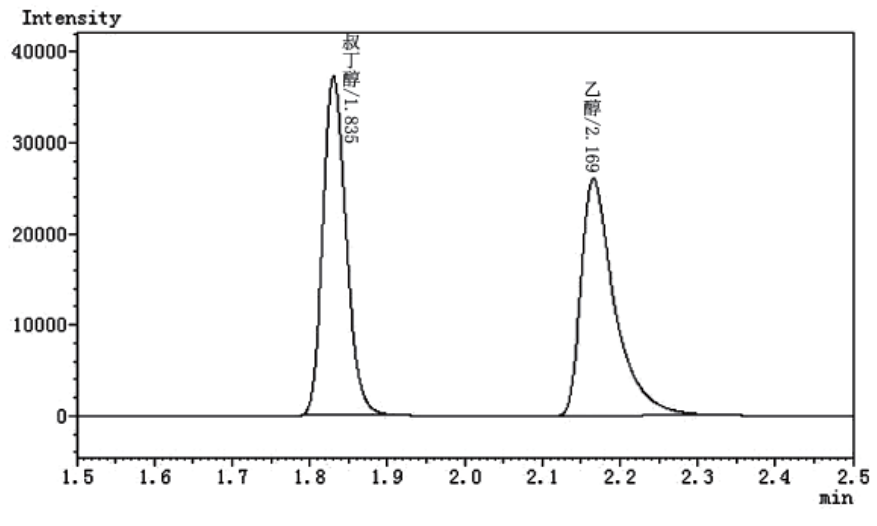


Fig 3. Chromatogram of Blood Sample.

internal standard working solution in the vial, as blank control.

2.4.3 Detection

Place vials of blank control, standard series and sample to be tested into HS autosampler, under the set condition of instrumental working conditions, get these samples injection. The retention times of ethanol and tertiary-butanol are 2.169 mins and 1.835 mins.

3 Results and Discussion

3.1 The Specificity of Method (Figs. 1-3)

Determine blank control, ethanol standard and blood sample successively, thus get chromatograms of them. From these graphs below, it is obvious that the qualification of the method is accurate.

3.2 Linear Regression Equation and Correlation Coefficient of Method

Use peak area ratio of ethanol to the internal standard as ordinate, the concentration of ethanol as abscissa, and get the linear regression equation $Y = 1.287\ X\text{-}1.061\times10^{-2}$. The correlation coefficient is $R^2 = 0.9997$, and the good linear range for ethanol concentration is 10 mg/100 mL-300 mg/100 mL. See ethanol standard series adjusting graph in Figure 4.

3.3 The Lowest Limit of Detection and Quantification Limit of Method

Make an attenuation of ethanol standard solution 100 mg/100 mL to 50 mg/100 mL, 25 mg/100 mL, 20 mg/100 mL, 15 mg/100 mL, 10 mg/100 mL and 5 mg/100 mL solutions. Add 0.09 mL blank blood respectively to 0.01 mL standard solution above and 0.5 mL internal standard working solution in the vials, seal and blend them and wait for determination. According to S/N≥ 3, the lowest minimum mass concentration of ethanol in blood is 1 mg/100 mL. According to S/N≥ 10, the limit of quantification is 5 mg/100 mL.

3.4 The Degree of Precision of Method

Prepare low, medium and high mass concentration of ethanol, and each concentration get 5 paralleled samples introduction. The results in Table 1, it is clear that the degree of precision is 1.96%-3.06%, which meets the requirement.

3.5 The Repeatability of Method

Prepare low, medium and high mass concentration of ethanol, and each concentration get 6 paralleled samples introduction. The results in Table 2 show the repeatability of method is favorable.

3.6 The Recovery Test of Method

Get blank blood prepared to 3 groups of solution in which ethanol concentration are 20 mg/100 mL, 80 mg/100 mL and 160 mg/100 mL, respectively. Each group has 3 paralleled samples tested to calculate the recovery rate of standard addition. The results in Table 3 obviously show that the recovery is 96.50%-99.42%, and the RSD is far lower than 2%, both which indicate the method is accurate and reliable.

3.7 Samples Stability Test

Prepare low, medium and high mass concentration of ethanol, store them in the refrigerator of 4°C. Each concentration check 6 paralleled samples on the first day, the third day, the seventh day and the fourteenth day. Separately compare the averages of 3rd day, 7th day and 14th day with that of 1st day. The results in Table 4 display the relative deviations are all less than 10%, which proves the blood samples can be preserved at least 2 weeks time in the refrigerator of 4°C.

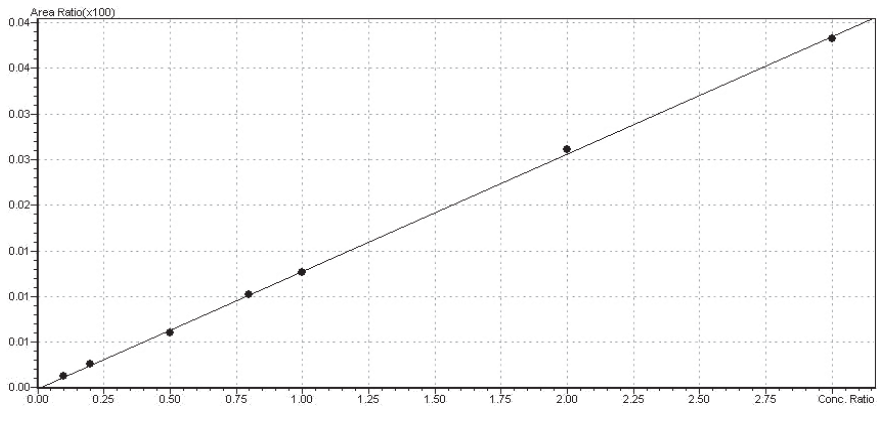


Fig 4. Cablibration Curve of Ethanol Standard Series.

4 Conclusion

It is feasible that the method can be used to detect the content of ethanol in blood, which has a high sensitivity, a good linearity of calibration curve. Moreover, the test results of precision, accuracy and repeatability show it meets requirement.

Acknowledgement

We appreciate the excellent technical assistance provided by SHI Yintao, DUAN Jie, SONG Shengli. We also thank TAO Daiqin for polishing the article.

References

[1] DAI Juan, CHEN Man, LAI Yi, et al. Study for Directly Injected Sample Gas Chromatography Capillary Column Determination of Ethanol Content in Whole Blood [J]. Chinese Journal of Health Laboratory Technology, 2010, 20(2): 283-284. (In Chinese)
[2] LUO Yun, YUE Yunyao, LI Dachun. A Study on Determination of Ethanol in Blood [J]. Chinese Journal of Health Laboratory Technology, 2010, 20(8): 1897-1898. (In Chinese)
[3] LIU Yingchun, CHEN Jie, JI Huagui, et al. Direct Dilution Gas Chromatographic Method for The Determination of Ethanol Content in Blood [J]. West China Medical Journal, 2011, 26(4): 572-575. (In Chinese)
[4] The Department of Public Security of The PRC. The Professional Standard of Public Security of The PRC GA / T 842-2009 [S]. Beijing: Standards Press of China. 2009. (In Chinese)
[5] The Department of Public Security of The PRC. The Professional Standard of Public Security of The PRC GA / T 1073-2013 [S]. Beijing: Standards Press of China. 2013. (In Chinese) ■

Table 1. Method Degree of Precision Test.

Sample Concentration	Measured quality (mg/100 mL)					Average (mg/100 mL)	RSD (%)
Low	16.30	15.70	15.60	15.80	16.20	15.92	1.96
Medium	74.90	72.00	72.70	74.90	76.30	74.16	2.38
High	150.0	142.9	142.5	147.4	152.9	147.14	3.06

Table 2. Method Repeatability Test.

Sample Concentration	Measured quality (mg/100 mL)						Average (mg/100 mL)	RSD (%)
Low	20.50	19.20	19.30	18.80	19.40	19.00	19.37	3.07
Medium	78.70	78.50	80.60	77.60	78.80	78.30	78.75	1.27
High	157.9	158.4	164.2	157.9	160.6	158.2	159.5	1.57

Table 3. Method Recovery Test.

Number	Addition (mg/100 mL)	Measured quality (mg/100 mL)	Recovery rate (%)	Average recovery rate (%)	RSD (%)
1	20	19.20	96.00		
2	20	19.30	95.00	96.50	0.518%
3	20	19.40	97.00		
4	80	78.70	98.38		
5	80	78.50	98.13	98.34	0.194%
6	80	78.80	98.50		
7	160	158.4	99.00		
8	160	160.6	100.38	99.42	0.837%
9	160	158.2	98.88		

Table 4. Method Degree of Precision Test.

Sample Concentration	Measured Average (n=6) (mg/100 mL)				Relative Deviation (%)		
	1st Day	3rd Day	7th Day	14th Day	3rd Day	7th Day	14th Day
Low	20.41	19.85	19.42	18.95	2.78	4.97	7.42
Medium	79.00	77.93	78.03	76.90	1.36	1.24	2.69
Medium	158.78	155.28	156.81	154.90	2.23	1.25	2.47

Received 3 June 2014
Received in revised form 30 August 2014
Accepted 11 October 2014
Available online 28 November 2014
Peer Reviewed
2157118X.4.1.R2

Proficiency Tests to Estimate Error Rates in the Forensic Sciences

Jonathan Koehler [*, a]

[a] Northwestern University School of Law, Chicago, Illinois, U.S.A.

[*] Corresponding Author: Jonathan Koehler, Ph.D. E-mail: jay.koehler@northwestern.edu

Abstract A proficiency test is an assessment of the performance of laboratory personnel using samples whose sources are known to the proficiency test administrator but unknown to the examinee. There are many purposes to such assessments: training personnel, ensuring that personnel achieve baseline competence levels, improving laboratory practices and procedures, and identifying future needs for a laboratory. Proficiency tests can also help identify reasonable first pass estimates for the rates at which various types of errors occur.

Keywords: Forensic science, Proficiency tests, Error rates.

A proficiency test is an assessment of the performance of laboratory personnel using samples whose sources are known to the proficiency test administrator but unknown to the examinee. There are many purposes to such assessments: training personnel, ensuring that personnel achieve baseline competence levels, improving laboratory practices and procedures, and identifying future needs for a laboratory. Proficiency tests can also help identify *reasonable first pass estimates for the rates at which various types of errors occur*.

It is crucial to obtain error rate estimates because the reliability and probative value of forensic science evidence is inextricably linked to the rates at which examiners make errors. Without such information, legal decision makers have no scientifically meaningful way of thinking about the risk of false identification and false non-identification associated with forensic reports.

When designing proficiency tests to estimate error rates, careful thought must be given to at least four issues:

(1) the composition of the test designers and administrators who oversee the testing process,

(2) the features of test and reference samples,

(3) the composition and selection of test participants, and

(4) the use of blind test protocols.

Issue #1: Test Designers and Administrators

The designers and administrators of proficiency tests should be qualified, disinterested parties. By "qualified," I mean people who have expertise in such areas as experimental design, testing, statistics, behavioral sciences, police investigation, and forensic science. It would be hard to overstate the importance of including statisticians, behavioral scientists and others who have training and experience in matters related to research methodology involved in this process. If the proficiency tests are not properly designed, then scientific inferences cannot be made. By "disinterested," I mean that proficiency test designers and administrators should not be affiliated with the examinees or the examinees' laboratories, nor should they stand to benefit from or be harmed by any particular outcome or set of outcomes on the proficiency tests..

Issue #2: Features of Test Samples

The samples used in proficiency tests should approximate a random sample of the types of evidentiary samples that arise in actual cases. This may be accomplished in different ways. One way is for test administrators to access a database of all cases in a county, state, country, or other population over some time period (e.g., five years), and to note which cases included forensic science evidence. A random sample of those cases might then be identified as prototypes for the manufacture of proficiency test samples. Samples

identified in this manner are likely to vary widely. Using a fingerprint example: one case might include two detailed latents plus rolled prints from one suspect and two innocents. Another case might include one badly smudged latent and rolled prints from each of ten suspects, including a pair of identical twins.

Once a random sample of cases has been identified, test administrators should write comparable cases and then manufacture forensic evidence that resembles the samples and cases chosen. Materials should not be reused across tests.

The newly created evidence should be rated for difficulty using an agreed-upon rating scheme to ensure that they parallel the sample of selected cases and to allow researchers to track the impact of sample difficulty on examiner accuracy. Likewise, administrators should track task features such as whether multiple samples are from a single common source, and whether the source of the print or marking is or is not present.

Issue #3: Test Participants

Test participants should be a random or otherwise representative sample of forensic scientists who testify in court. Pertinent background features of selected participants should be tracked. These features should include training, experience, and number of cases in which participants have testified. By tracking examiner characteristics, we will gain insight into the conditions under which performance varies.

All forensic scientists who testify in court must be part of the participant pool. Examiners cannot opt in or out. However, it is not important that all or even most forensic scientists be selected to participate in the proficiency tests. The idea is that participants are sampled using statistically sound methods. This method will allow for extrapolation of results to the broader forensic population.

The notion that forensic scientists should be selected at random rather than required to participate in any given test may come as a surprise. But it is consistent with the proficiency testing purpose described here: *to identify a reasonable first pass estimate for the rates at which various types of errors occur*. The purpose of the tests is not to identify lab-specific or examiner-specific error rates. Such data would be useful, but they are difficult to obtain, and likely to be misinterpreted or dismissed even if they were obtained. [1]

The notion that our focus should be on identifying general error rates rather than individual or situation-specific error rates is worth careful consideration because intuition suggests that the opposite is true. After all, why should a careful, well-trained, and experienced examiner be saddled with the same first pass error rate estimate as a careless, poorly trained, and inexperienced examiner?

The answer to this loaded question lies in the name of the error rate itself. It is a "first pass" error rate estimate. It is an estimated base rate for errors. Factfinders need such base rates, in combination with individuating information about a given examination and a given

Note

1 Some people favor proficiency tests that identify examiner-specific error rates rather than general error rates. Although such data are desirable in the abstract, they are unlikely to be helpful in practice. In any endeavor where errors are infrequent, too many tests are required under too many different conditions to identify person-specific error rates that apply to particular cases. Consider, for example, an examiner who makes 100 comparisons without error under controlled, rigorous and realistic test conditions. What does this test performance tell us about the examiner's error rate? From a statistical standpoint, zero errors out of 100 trials represents an underlying error rate anywhere from 0% (the examiner theoretically could be "perfect") to about 3%. And that 3% figure (which is the approximate upperbound of a 95% confidence interval) would be the cause of a lot of anger and misunderstanding ("How could the examiner have a 3% error rate when he/she didn't make any errors?"). The confusion can be resolved by a class in statistical inference. But judges, jurors and experts who lack statistical training will not appreciate this point. And even those who do appreciate the inferential uncertainty surrounding test performance are likely to dismiss the entire effort on grounds that the conditions in the focal case differ in significant ways from those in the test conditions. In the end, then, even if it were possible to get busy, backlogged examiner to sit for a lengthy set of proficiency tests designed to estimate individual rates of error, the data would cause confusion and prejudice at trial. ■

examiner, to make an informed judgment about the risk of error in any particular case. The job of the proficiency tests described here is to provide that base rate. It is not to provide the individuating information that the factfinder might use to adjust the base rate.

A sports analogy provides some clarification. In order to know the chance that a professional baseball player will get a hit in his next at bat, a forecaster needs to have a sense of the base rate for hits. Baseball players get hits about one time in four chances. A good player gets a hit one

time in 3.5 chances, and an outstanding player gets a hit one time in three chances. These are base rates. If individuating information is available – such as whether the player has an injury, is facing a tough pitcher, or is hitting with the wind blowing out – adjustments to the base rate should be made. But the base rate is the anchor. The base rate for a hit in baseball is very different than the base rate for a goal in professional hockey (about 10%) or for a successful free throw in professional basketball (about 75%). These first pass sports base rates are known. In contrast, the base rates for errors in the forensic sciences are unknown and cannot be discerned from case reports or legal outcomes: scientific study is needed.

Issue #4: Blind Tests

Ideally, proficiency tests should be blind in the sense that any party that has a direct interest in how the examiners perform should not be aware that the proficiency test materials are part of a test rather than part of actual casework. Behavior may change under observation and it is important to make test conditions as similar to casework conditions as possible. Part of that similarity means not telling examiners that they are being tested. This is a key feature in a scientifically valid proficiency test of human performance.

Some may take offense at the suggestion that forensic scientists' behavior may vary when they know they are being tested. After all, they are trained professionals and many have years of experience. But the notion that behavior changes under observation is well-documented across many domains for both experts and novices alike. It is simply part of the human condition.

The notion of blindness in proficiency tests is sometimes dismissed on grounds that it either cannot be done or would be too costly to implement. These criticisms should be rejected. Blind proficiency testing is already used in some forensic science areas (such as the DOD's forensic urine drug testing program and HIV testing program). Blind tests have also been used for DNA analyses. For example, a 2002 study in the Int'l J. of Legal Medicine reports the results of DNA blind trials across 129 laboratories in 28 European countries. Furthermore, Joe Peterson has conducted a detailed pilot investigation in the U.S. which showed that blind testing of DNA analysts can be done. His reports appear in two articles in the Journal of Forensic Sciences in 2003 and in detailed discussions in papers filed with the National Institute of Justice.

Conclusion

As noted above, the reliability and probative value of forensic science evidence is inextricably linked to the rates at which examiners make errors. Jurors and others cannot assess the significance of a reported forensic science match without having some sense of the rate at which false positive errors occur. Properly designed proficiency tests provide a necessary first step. The design and administration of these proficiency tests – i.e., tests that will provide reasonable first pass estimates for the rates at which various types of errors occur – is a major undertaking. And even a successful venture that identifies error rate estimates for different technologies undervarious conditions will not tell us everything we wish to know about the risk of error in specific cases. But if they provide jurors and others with a better sense of the probative value of a reported match, they will serve their purpose.

Of course, it would be easy to dismiss the entire enterprise described here by pointing out unanswered practical questions. Who will administer the tests? How much will they cost? How can participation be ensured? How will examiners create enough time to participate? These are important questions. But they are not the questions that the forensic science community in general, and the fingerprint community in particular, should be asking at this stage. Instead, the $64,000 question is this: *is the forensic science community prepared to accept the idea that error rates matter, and that the way to estimate those rates is through a carefully designed, rigorously scientific, testing program?* If the answer is "yes," then we are half way there. ■

2157118X.4.1.83 Peer Reviewed
Received 31 May 2014
Received in revised form 7 July 2014
Accepted 7 July 2014
Available online 28December 2014

On Limiting the Use of Bayes in Presenting Forensic Evidence

Norman Fenton [a], ***Martin Neil*** [b]

a Professor and Director of Risk and Information Management Research Group (Queen Mary University of London) and CEO Agena Ltd. Email: norman@eecs.qmul.ac.uk.

b Computer Science and Statistics (Queen Mary University of London) and CTO Agena Ltd. Email: martin@eecs.qmul.ac.uk.

Abstract A 2010 UK Court of Appeal Ruling (known as "R v T") asserted that Bayes theorem and likelihood ratios should not be used in evaluating forensic evidence, except for DNA and 'possibly other areas where there is a firm statistical base'. The potential impact of this ruling is enormous and it has drawn fierce criticism from expert witnesses, academics and lawyers, who have identified various weaknesses and fallacies in the ruling. This paper focuses on the strategic and cultural challenges that the ruling raises to ensure that the role of Bayes is better understood and exploited in the presentation of forensic evidence. We provide a simple unifying way of describing all probabilistic forensic 'match' evidence; this enables us to easily identify and avoid the kind of common misunderstandings and fallacies that have afflicted probabilistic reasoning about evidence, including especially why it is irrational to assume that some forensic evidence is 'statistically sound' whereas other less established forensic evidence is not. But these misunderstandings are not restricted to lawyers, since we show that both forensic scientists and even Bayesian experts have consistently failed to include all relevant information in their evidence, such as error probabilities, and this applies to DNA as much as any other forensic science. We also show that there are severe limits of the extent to which the results of Bayes can be presented in purely intuitive terms; we show that the scope in forensics is even narrower than previously assumed. Hence, there are two major challenges facing the opponents of the R v T ruling: First, there must be much greater awareness of the need to improve Bayesian forensic arguments (before they are even presented in court) in order to avoid the common errors and omissions that are made. Second, there must be a radical rethink on the strategy for presenting the results of Bayesian arguments in court. Resorting to the formulas and calculations in court is a dead end strategy since these will never be understood by most lawyers, judges and juries, but the intuitive presentations simply do not scale up. Ultimately this means getting the lay observers to 'accept' that they need only question the prior assumptions that go into the Bayesian calculations and not the accuracy or validity of the calculations given those assumptions. Bayesian networks may provide a suitable mechanism for performing these calculations.

Keywords: Forensic science, Bayes, Evidence.

1 Introduction

Proper use of probabilistic reasoning has the potential to improve dramatically the efficiency and quality of the entire criminal justice system. Bayes theorem is a basic rule, akin to any other proven maths theorem, for updating the probability of a hypothesis given evidence. Probabilities are either combined by this rule, or they are combined wrongly. Yet, the Court of Appeal in the case of R v T [1] ruled that the use of formulas to calculate probabilities and reason about the value of evidence was inappropriate in the area of footwear evidence. It regarded the forensics of footwear matching as 'unscientific' and not having a sufficiently 'firm statistical base' in contrast to DNA forensics. Specifically, Points 86 and 90 of the ruling respectively assert:

"*..We are satisfied that in the area of footwear evidence, no attempt can realistically be made in the generality of cases to use a formula to calculate the probabilities. The practice has no sound basis*".

" *It is quite clear that outside the field of DNA (and possibly other areas where there is a firm statistical base)*

this court [1] *has made it clear that Bayes theorem and likelihood ratios should not be used*"

Given its potential to change the way forensic experts analyse and present evidence in court, experts have been understandably quick to publish articles criticising the ruling. At the time of writing there have already been at least four such excellent articles [12],[29],[31],[32] that provide a detailed analysis of the case and ruling. These papers recognise that there were weaknesses in the way the expert presented the probabilistic evidence (in particular not making clear that likelihood ratios for different aspects of the evidence were multiplied together to arrive at a composite likelihood ratio), but nevertheless express deep concern about the implications for the future presentation by experts of forensic evidence. The papers recognise positive features in the ruling (notably that experts should provide full transparency in their reports and calculations) but they provide compelling arguments as to why the main recommendations stated above are problematic. For example, [32] uses the following analogy of likelihood ratio calculations with area calculations:

Saying the expert should not use this 'mathematical formula' to assess the composite likelihood ratio is like saying that if one is just estimating by eye the area of a field, one is not allowed to multiply estimates of its width and length together. Clearly it is the correct procedure: there is no uncertainty in the relationship between length, width, and area, only in their values. If the Court were to say that the expert was not to use a logical procedure, rather than a 'mathematical formula', the flaw in its reasoning would be obvious.

1 The judge is actually referring to the Court of Appeal ruling in the case of Adams, which is mentioned in Point 89. ■

The authors in [32] also conclude that:

..the Court has not understood the difference between assessments of the probability of a proposition and of the strength of evidence for the proposition; the second is a confusion between uncertainty in the values of the variables and uncertainty in their relationship in a mathematical formula. The fact that variables cannot be precisely expressed does not affect the validity of the relationships described by the formula.

The authors in [31] highlighted the inconsistency in the ruling which, on the one hand rejects the use of Bayes and likelihood ratio calculations, while on the other hand insists on full transparency of all calculations. They ask:

..how could such an injunction ever be enforced on forensic scientists … The best that might be imagined would be a policy of "don't ask, don't tell", whereby experts formulated their conclusions according to their good faith understanding of scientific protocol but carefully concealed their "deviant" probabilistic reasoning from legal scrutiny.

On a similar theme the authors in [11] assert that:

…the evaluation of evidence for a court of law is not just a matter of "using likelihood ratios" but one of working to a set of principles that are founded on logic. To deny scientists the contemplation of the likelihood ratio – whether quantitative or qualitative – is to deny the central element of this logical structure

Clearly, as pointed out in [32], the ruling in [1] exhibits misunderstandings of some fundamental ideas of probabilistic reasoning and even includes instances of the fallacy of the transposed conditional, despite the dozens of papers and even rulings about it over many years. That such errors should continue to be made routinely by members of the legal profession (see also [19] for other recent examples) indicates that we (meaning the community of experts in probabilistic reasoning) have failed to communicate our arguments effectively where it matters most. In Section 2 we explain the challenges that this failing poses for expert witnesses and Bayesians. The rest of the paper addresses the challenge and is structured as follows:

- In Section 3 we introduce a hypothetical forensic 'science' in order to present the core ideas of forensic match evidence in a simple unifying way. This enables us to explain in very simple terms the Bayesian approach and to expose not just the fundamental misunderstandings in the R v T ruling, but also a number of key issues that have been missed in previous discussions.
- In Section 4 we use the generic example to highlight the irrationality of the core message in the R v T ruling (namely that there can be a clear distinction between forensic methods that are or are not 'statistically sound' and different allowed reasoning applied).
- While Sections 3 and 4 expose the weaknesses in the R v T ruling, Section 5 explains why, in many ways, the ruling is perfectly understandable, since we show that forensic probabilistic evidence is usually presented in a confusing -, and often incorrect - way. In particular, forensic scientists and even Bayesian experts typically ignore (or do not properly articulate) the potential for testing errors (false positives and false negatives).
- Hence in Section 6 we show that, when the potential for testing errors is included (as

it should be) this introduces significant complexity even in very simple cases. The key point is that, even in the simplest case, it is unrealistic ever to expect the associated Bayesian argument to be understood by lay people. We explain how the use of Bayesian network models may potentially address this problem.

- Finally, in Section 7 we present the grand challenge that Bayesians need to address before Bayes can ever take (what Bayesians feel should be) its rightful central position in legal reasoning.

2 The main challenges for expert witnesses and Bayesians

While the various papers on the ruling in [1] have done a fine job analysing in depth the weaknesses contained therein, there should be no doubt that the ruling is a damning indictment of the community of experts and academics who recognise the central importance of Bayesian reasoning for evidence evaluation. Despite some twenty-five years of work explaining the power and relevance of Bayes to the law, (resulting in several hundred academic publications and dozens of textbooks) the actual impact on legal practice has been minimal.

This failure must be attributed to our inability to communicate the core ideas in such way that they are accepted as a standard tool of the trade rather than as they are perceived now by much of the legal profession: an exotic, somewhat eccentric method to be wheeled out for occasional specialist appearances whereupon a judge or lawyer will cast doubts on, and even ridicule, its integrity (hence ensuring it is kept firmly locked in the cupboard for more years to come).

To address the problem we need to communicate the core ideas more effectively to both forensic scientists and lawyers. Specifically, we need to ensure that:

a. both the forensic scientists and lawyers know when Bayesian reasoning should be used.

b. the forensic scientists are able to properly articulate the assumptions required for a Bayesian analysis.

c. both the forensic scientists and lawyers know the difference between the assumptions required for the analysis (which will generally be disputed) and the Bayesian calculations that determine the conclusions based on the assumptions (which must not be disputed).

d. before evidence is used, the forensic scientists are able to perform the Bayesian calculations correctly and efficiently. The scale of this problem has been massively underestimated, and as we shall explain in this paper, can only be resolved by more widespread acceptance of the use of tools.

e. the forensic scientists (and ultimately the lawyers themselves) are able to present the results of Bayesian reasoning about evidence in a way that is understandable to jurors and other lawyers. This is the most difficult challenge of all since, ultimately it will only be achieved once it is accepted that we do not actually have to reason in court about the results of the Bayesian calculations themselves (i.e. the calculations are accepted in the same way as we might accept the results of using a calculator for long division[19]).

f. likelihood ratios (or some suitable graphical/verbal equivalent representation) are used as a standard means for stating the value of evidence (individually and in combination).

To see the extent of how and why we have failed to meet the above objectives we need only look at the range of relevant textbooks:

- There are two standard textbooks, [24] and [27], for forensic science training. Despite its apparently encyclopaedic coverage, [27] contains nothing at all on Bayes and only some basic high school material on statistics such as graphs and bar charts. The book [24] does contain a very brief introduction to Bayes and the likelihood ratio right at the end, but without attempting to link it in any way to the core material of the book (so that it appears as an afterthought, out of context).
- There is one standard book, [37], aimed at forensic scientists presenting evidence in court. Until its latest 2010 edition, this book did not contain any mention of Bayes, likelihood ratios, or even probability, and so failed to consider such basic issues as random match probability and the probability that tests may have less than perfect accuracy (more encouragingly, the new 2010 edition does contain a chapter on trace and contact evidence [14] that includes a discussion of the Bayesian approach).
- There are several excellent books that focus on the statistical and probabilistic aspects of forensic evidence. These include [8], [10], [11], [13], [18], [21], [28], [29]. These books cover exactly the right material in depth, and they also include introductory material on Bayes. However, they are most suited for people with a statistical or mathematical background (who wish to find out in detail how to properly reason with forensic evidence) rather than practicing forensic scientists lawyers. So, for example, even those that are considered the most accessible to non-

experts, namely [8], [18], [29], make extensive use of formulas and hence require a significant level of mathematical sophistication. The books also tend to focus on the details of specific types of forensics (especially DNA).

- There are no suitable relevant books we are aware of that are specifically targeted at lawyers. The closest would be populist books on probability and risk, such as [22] and [23], but these do not address the issue of evidence presentation.

In [19] [20] we argued that it was a mistake to assume that any kind of Bayesian formulas - such as those used in the case of R v Adams (and shown in Figure 1) could be presented to lawyers and juries no matter how 'simple' they appeared to statisticians.

In the relevant text books and papers discussed above the best approaches start with visual explanations of a very simply instance of Bayes (using, for example, tree diagrams with frequentist versions of the probabilities). However, for reasons we will explain in Section 3 below, these visual approaches do not scale up meaningfully in any realistic situation. It is at this point that the various authors normally resort to the formulas instead; hence, this is the point that most forensic scientists and lawyers never get beyond.

3 Clarifying the notions of 'forensic match' and common fallacies

To help readers understand that there is a simple unifying way to present any kind of forensic 'match' evidence we use a hypothetical (but not unreasonable) example of a completely new forensic science, which we call 'stature matching'. This avoids the problem of getting distracted by the details and biases of specific areas (such as shoeprint matching or DNA matching). This approach will enable us to expose numerous common misunderstandings about the meaning of match evidence and that, contrary to what the judge ruled in [1] (and indeed what forensic many experts assume), it is inappropriate to assume that certain methods are inherently 'scientifically sound' and others are not.

3.1 A new, but typical, forensic science: Stature matching

Our 'new' forensic science is called "stature matching". Stature matching assumes that, for any person, we can measure the following features:

- Sex (male, female)
- Height (in centimetres)
- Waistline ((in centimetres)

So each person has their own stature profile such as:

- (male, 131, 65)

The 'science' of stature matching is the ability to determine a person's stature profile accurately. They can do this either directly by observing and measuring the person or indirectly from an image of the person. If, for example, CCTV captures the image of a man at the scene of a crime (we can think of the image as a 'trace' left by the man) then stature matching scientists might determine that the trace has the following stature profile:

- (male, 132, 64)

A real person is said to be a 'match' to the stature profile of the trace if the following criteria are satisfied:

- Sex of the person = sex of the trace stature profile
- Height of the person differs from height of the trace stature profile by less than 2 centimetres
- Waistline of the person differs from waistline of the trace stature profile by less than 2 centimetres

So, for example, four different people with respective stature profiles

- (male, 132, 64)
- (male, 132, 64)
- (male, 131, 65)
- (male, 132, 65)

would all be considered to be a 'match' to the stature profile (male, 132, 64), whereas people with the following stature profiles would not be considered a match:

- (male, 135, 65) - this 'fails' on height
- (female, 132, 65) – this 'fails' on sex

Every branch of forensic matching that is based on some properties of people [2] (be it DNA, fingerprint, blood type, shoe-print, earprint, Gait, voice, ….and any other type not yet invented) is based on the same underlying principles as stature matching:

Specifically:

- Every person has a 'profile' (defined by the area of forensics) that can be
- measured by some defined procedure.
- In certain circumstances a person leaves a 'trace' (or 'print') of this profile
- In certain circumstances we can measure the profile of the trace that was left.
- There is a criterion for

$$V=\frac{Pr(H_p|E,I_1,I_2)}{Pr(H_d|E,I_1,I_2)}=\frac{Pr(E|H_p)}{Pr(E|H_d)}\times\frac{Pr(I_1|H_p)}{Pr(I_1|H_d)}\times\frac{Pr(I_2|H_p)}{Pr(I_2|H_d)}\times\frac{Pr(H_p)}{Pr(H_d)}$$

Fig 1. Typical Bayesian likelihood ratio calculation. Far too complex for lay people to understand

2 Other types of forensic match analysis, such as glass, fibres, pollen etc, are not concerned with attributes of people and do not exactly fit the same framework. ■

determining whether a trace profile matches the profile of a person.

The first simple (but extremely) important observation to make about forensic matching is that (in contrast to widely held assumptions) there is no definitive means for considering a forensic matching method to be 'scientific' or not. Most people assume that DNA is 'scientific' because the measurement and matching criteria and protocols are objective and reasonably standardised (in contrast to those that are widely assumed to be 'non-scientific' like gait analysis, face mapping, and fingerprinting). Yet, our new stature matching method is at least as scientific as DNA in this respect. For example, in stature matching we insist on always measuring the three specific values (sex, height, and waistline) and never any others; we can always assume that the height and waistlines are measured without clothes or shoes, and we always include the 2cm error margins for the match. There is no fundamental reason why any forensic method cannot in principle be made 'scientific'.

The second extremely important observation to make about forensic matching is the following (see [14], [26], [34] for a comprehensive discussion of this issue):

- A 'match' never means a unique identification of a person.

This is important because the assumption of uniqueness is a common fallacy arising in DNA, fingerprint, and many other areas of forensics. For example, in R v Kempster, EWCA Crim 975 [3] the ruling includes the following assertion about earprint evidence:

It is clear … that ear-print comparison is capable of providing information which could identify the person who has left an ear-print on a surface.

This assertion is highly misleading. In fact, when we find a 'match' (be it for stature matching, earprint matching, DNA or any of the areas of forensics discussed above) all we can conclude is that within the agreed criteria, the person's profile is the same as the profile of the trace. To equate this notion with 'identification' is always flawed.

An expert in stature matching could, in court, present the information about a match as follows:

"*I am absolutely certain that the stature profile of the trace found at the scene is a match of the defendant's stature profile.*"

Instead, the common error made by experts is to assert the following:

"*I am absolutely certain the stature profile trace found at the scene is that of the defendant*"

Indeed, this was exactly the error made by the expert witness on earprint evidence in R -v- Dallagher, EWCA Crim 1903 [4]. The judge consequently rejected the entire earprint evidence as inadmissible. While the judge's ruling was understandable in this particular case it would be extremely dangerous to interpret this as meaning that, unless a 'match' is the same as an 'identification', then match evidence can never be admissible. For not only would this rule all future earprint evidence as inadmissible, it would also rule as inadmissible every area of forensic match evidence.

3.2 Understanding the Bayesian approach to match evidence

In the simplest use of forensic match evidence in legal cases we assume that a person has left a trace at a particular location. Then we have the following (continuing with the stature matching example):

- Source profile: This is the stature profile of the trace found at the location.
- Target profile: This is the stature profile of a particular person believed (normally called the defendant) who some believe may have been the one who left the trace.

Let us, for the time being, make a massive simplification (it turns out that it is ONLY for this restrictive case that a simple explanation of Bayes is possible). We will assume that our stature testing is perfect. So, someone with type (male, 131, 65) will always be tested to be of type (male, 131, 65) and someone who is not type (male, 131, 65) will never be tested to be of type (male, 131, 65) .

With the above assumptions our typical simple forensic case amounts to the following:

- Prosecution hypothesis (H): "The target is the source" (i.e. the defendant is the person who left the trace at the scene).
- Defence hypothesis (not H): "The target is not the source" (i.e. a person other than the defendant left the trace at the scene).
- Evidence E1: The source profile type is known, say to be of type (male, 132, 64). For simplicity and generality we shall refer to a particular profile as type X.
- Evidence E2: Target profile matches the source profile (i.e. both have type X).

From an evidential perspective, the 'value' of the evidence is therefore completely determined by the following two pieces of (probabilistic) information:

1. 'Defence likelihood': How likely are we to see the evidence if the defence hypothesis is true. In other words how likely is it that the source and target (defendant) are both of type X, if the target was not the source.

   With the above simplistic assumptions, the defence likelihood is represented by the single branch (H false, E1 true, E2 true) in Figure 2. Suppose *m* is the proportion of people in the population who have type X. This is sometimes called the

frequency (of the particular type) or the *random match probability* (of the particular type). So, the defence likelihood is equal to m^2.

2. 'Prosecution likelihood': How likely are we to see the evidence if the prosecution hypothesis is true. In other words how likely is it that the source and target are both of type X if the target is the source.

With the above simplistic assumptions, the prosecution likelihood is represented by the single branch (H true, E1 true, E2 true) in Figure 2. Hence, the prosecution likelihood is simply equal to *m* (because our testing is perfect the target is certain to be of type X if the target is the source).

So, if the random match probability *m* is equal to 1 in a 100, then the prosecution likelihood is 100 times greater than the defence likelihood. In fact, we are 100 times more likely to observe the evidence if the prosecution hypothesis is true than if the defence hypothesis is true.

The likelihood ratio (the prosecution likelihood divided by the defence likelihood) is simply the mathematical formalism that expresses exactly this intuitive information.

The likelihood ratio is very well-suited to the legal context because it enables us to evaluate the impact of the evidence without having to specify what our prior belief is in the prosecution or defence hypothesis. What Bayes theorem additionally tells us is that, whatever our prior odds were for the prosecution hypothesis, the result of seeing the evidence is such that those odds are multiplied by the likelihood ratio[3]:

Posterior odds = Likelihood ratio × Prior odds

So, according to Bayes, if we started off assuming that the odds in favour of the defence hypothesis were 1000 to 1, then the 'correct' revised belief once we see the evidence is that the odds still favour the defence, but only by a factor of 10 to 1:

	Prior odds		*Likelihood ratio*		*Posterior odds*
Prosecutor	1	×	100	=	1
Defence	1000		1		10

And if we started off assuming that the odds in favour of the defence hypothesis were 4 to 1, then the 'correct' revised belief once we see the evidence is that the odds now favour the prosecution by a factor of 25 to 1:

	Prior odds		*Likelihood ratio*		*Posterior odds*
Prosecutor	1	×	100	=	25
Defence	4		1		1

3 Note the following (which we will assume later): If we assume that the prior odds are 'evens' i.e. 50:50 then the posterior odds will be the same as the likelihood ratio. Also odds can easily be transformed into probabilities: specifically, if the odd are x to y for hypothesis H over not H then the probability of H is x/(x+y) and the probability of not H is y/(x+y). So odds of 100 to 1 in favour of H means the probability of H is 100/101 and the probability of not H is 1/101. ■

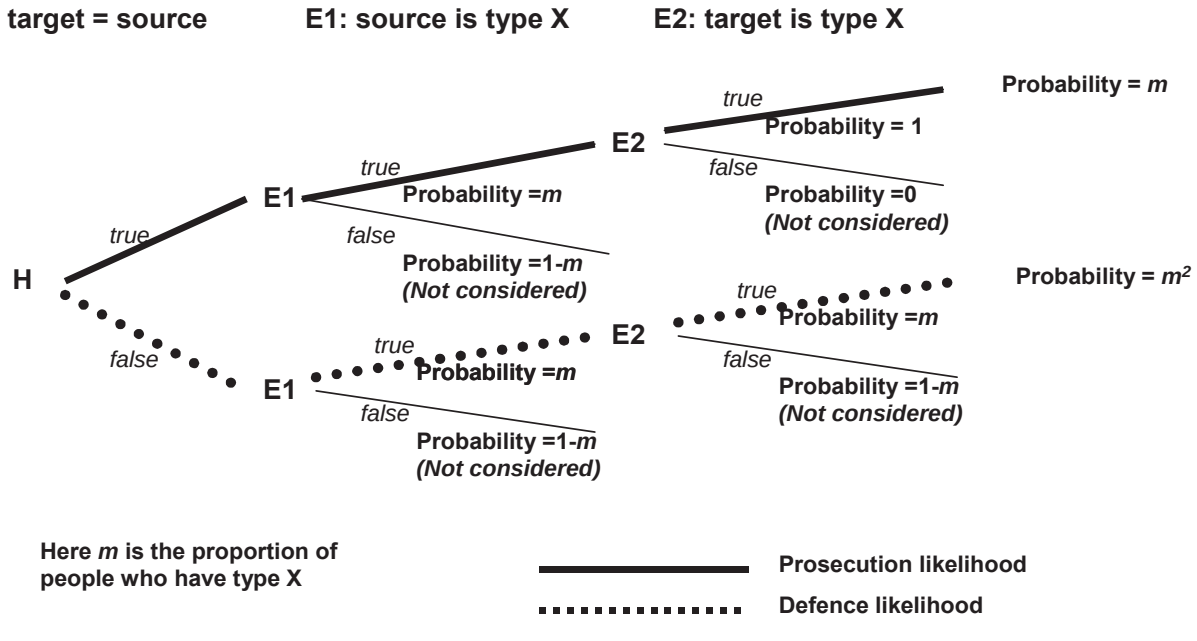


Fig 2. Determining the possible scenarios and likelihoods in simple case

But why should we accept that Bayes is the 'correct' interpretation? The standard way to convince lay people that Bayes is correct is to consider examples (often called the 'Island' example) like the following:

> ***Example 1:*** Suppose that, in addition to the defendant, it is known that another 1,000 other people were in the vicinity of the crime scene[4] – see Figure 3. Then our prior assumption, i.e. what we should assume before any evidence has been presented, is that any one of these other people is just as likely to be the person who left the trace as the defendant. So the prior odds are 1000 to 1 in favour of the defence hypothesis (or equivalently the probability that the defence hypothesis is true is 1000/1001). Since the random match probability is 1/100, we expect about 10 of the other 1000 people to have the type X. So, once we observe the evidence (defendant is type X) we can rule out all other people, except those 10, as having possibly left the trace. So, after observing the evidence the defendant and 10 others remain as possibilities. So the revised odds are now 10 to 1 in favour of the defendant (or equivalently the probability that the defence hypothesis is true is now 10/11). So, although the odds still favour the defence hypothesis the odds have swung by a factor of 100 (the likelihood ratio) towards the prosecution hypothesis.
>
> If we change the number of people we start with the odds still always swing by a factor of 100 (the likelihood ratio) towards the prosecution hypothesis. So, if there were 500 other people then we expect about 5 to have the same stature type. So the prior odds, which in the case are 500 to 1 in favour of the defence, drop to 5 to 1 after observing the evidence

If there were just 10 other people then the use of population diagrams such as in Figure 3 to represent Bayes becomes difficult because, in this case, the expected number of people who match is a fraction (one tenth) of a person. From a mathematical perspective this is not a problem: the prior odds are 10 to 1 in favour of the defence. After the evidence there is just 1/10 of another person other than the defendant. So the odds are now 10 to 1 in favour of the prosecution hypothesis. The swing is still a factor of 100 toward the prosecution. But this example shows that, even with the most simplistic assumptions we have made the standard explanation of Bayes and likelihood ratios may not be easily understandable to lay people. Because many types of forensic science (such as DNA) have very low match probabilities, it is inevitable that we have to consider 'fractions' of people if we adopt this approach. The trick to gaining acceptance from lay people is therefore to use hypothetical examples that do not involve fractions, and then explain that exactly the same method works no matter what the actual match probabilities are.

3.3 Exposing some common misunderstandings

Before tackling the core problem of what constitutes 'statistically sound' evidence it worth noting that the framework we have provided makes it easy to expose three common misunderstandings in probabilistic reasoning about evidence:

When likelihood ratios can and cannot be multiplied

The practice of multiplying likelihood ratios was explicitly criticised in [1]. The error in the ruling was the failure to understand and

4 In the standard 'Island problem' presentation it is assumed that the crime was committed on an island and that, in the absence of evidence, all residents are equally likely suspects. ■

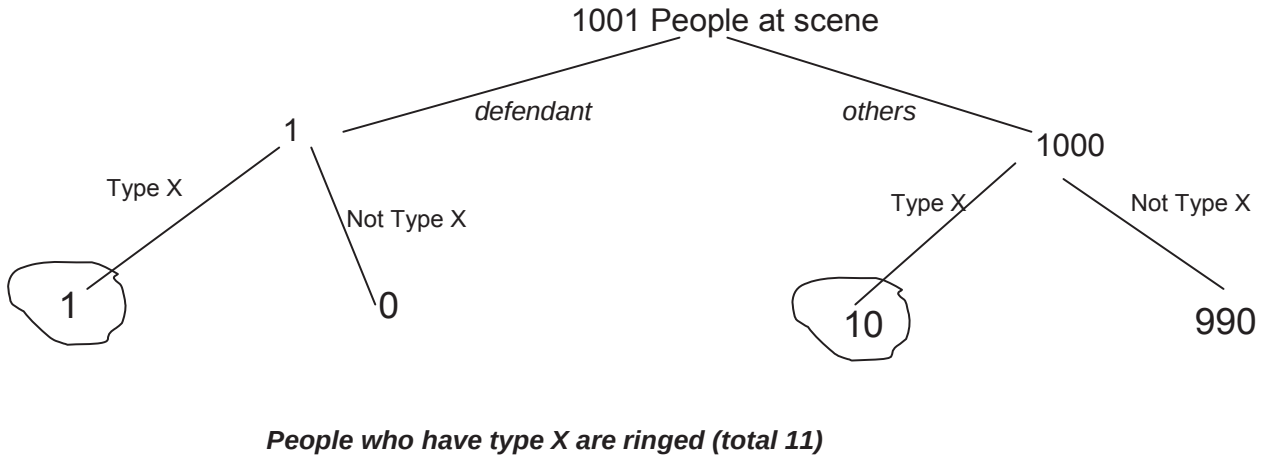


Fig 3. Bayes calculation explained visually

distinguish between the circumstances when multiplying likelihood ratios was and was not the correct thing to do.

When there are two pieces of independent evidence then multiplying likelihood ratios is the only correct way to reason about the impact of the combined evidence. To see why, suppose, that in addition to a match of the defendant's stature profile, we also discover a match of hair colour; the defendant and the person at the scene have brown hair. Suppose that the random match probability for brown hair is 1 in 5. Then the evidence in this case is that the stature profile and the hair profile of the defendant both match that of the person who left the trace (in the form of a CCTV image) at the scene. Since stature and hair colour can be considered independent, the probability of seeing both matches given that the defendant was not the person who left the print is the product of the two random match probabilities, i.e. 1/500. Hence the likelihood ratio is now 500. Assuming there are 1000 other people who were at the scene, it follows that 10 of these is likely to have the same stature profile as the defendant and of those 10 two are likely to have the same hair colour as the defendant. This means the odds in favour of the defence hypothesis have come down from 1000 to 1 to 2 to 1. That is a factor of 500, which is equal to the product of the two likelihood ratios (100 times 5).

So, when two pieces of evidence are genuinely independent it would, contrary to the ruling in [1], be irrational not to multiply the likelihoods - even for such 'unscientific' forensics as stature matching and hair colour.

However, the ruling against multiplying likelihood ratios is perfectly justified if the defence was unable to demonstrate that the underlying pieces of evidence were independent. If, for example, instead of hair colour we chose 'weight' it would certainly be wrong to conclude that weight was independent of stature. In such circumstances there are standard, but different, Bayesian calculations that need to be used (we have to consider explicitly the probability of one piece of evidence given the other). But such a scenario already puts us into the realms of problem complexity beyond which it is reasonable – or even possible – to perform manual calculations that lay people would be able to understand intuitively.

Fallacy of the transposed conditional.

This occurs when the defence likelihood, i.e. the probability of seeing the evidence given the defence hypothesis, is wrongly assumed to be equivalent to the probability of the hypothesis given the evidence.[5]

So, suppose we know that the defence likelihood is 1/100. By wrongly assuming this is the same as the probability of the hypothesis given the evidence, a prosecutor might state

- "The probability the defendant was not at the scene given this match evidence is 1 in 100"

In fact, if our prior was 1000 to 1 in favour of the defence hypothesis (as in Example 1 above) it turns out that what should have been stated was:

- "The probability the defendant was not at the scene given this match evidence is 10 in 11"

The danger of reading too much into very low match probabilities

For DNA the probability is normally presented as being so low (for example, 1 in 2 billion) that is it as 'good as' equal to zero[6] and hence a match is (wrongly) considered as a unique identification. In the case of fingerprints the situation is even worse, since there is still a strong assumption by many that a match is, by definition, a unique identification (i.e. the random match probability is assumed to be equal to zero).

Recent research, such as [16], has exposed this fallacy for fingerprint evidence and this was best exemplified by the dramatic Mayfield case [6] where a fingerprint match was subsequently discovered not to be that of the defendant. Primarily on the basis of this instance of a known match 'error', a State of Maryland Court subsequently ruled that fingerprint evidence was not admissible in a totally unrelated murder case [7]. If that way of thinking was applied to DNA or any other type of forensic evidence, then any example of a 'match' in which the person deemed matching was NOT the one who left the 'print', would be justification for rejecting as inadmissible the whole of that field of forensic evidence.

4 The irrational notion of 'statistically sound' evidence

Having dealt with some of the misunderstandings and fallacies in rulings such as [1] we now turn to the most critical and challenging misunderstanding that lies at the heart of the ruling: the assumption that the random match probability is 'statistically sound' for some areas of forensic science and not others. We again expose the weakness of this assumption by using our hypothetical stature matching example.

For any forensic science the match probability is based on some database of profiles. For our new science of stature matching we therefore need a database of people's stature profiles. For a particular

5 So, using the language of statisticians P(H|E) is wrongly assumed to be equal to P(E|H) hence why it is referred to as transposing the conditional.

6 This is especially true of the FBI in the US. In the UK the Forensic Science Service no longer assumes this, although lay people and many lawyers do. ■

profile, say (male, 132, 64), we simply count the frequency of profiles in the database that would be classified as a match to this profile. So this would include profiles like: (male, 132, 64), (male, 131, 65), (male, 132, 65), etc. If there are 1000 such matches in a database of 100,000 then we can say that the random match probability is 1 in a hundred, or equivalently 0.01.

The 'reliability' of the database could, of course, be questioned on numerous grounds such as the following:

- If the crime happened in the UK and the database comes from the USA then it may not be representative; perhaps people in the UK are smaller
- If it is known that the person who left the print was definitely a man then perhaps we should we consider only a database containing stature profiles of men.
- If it is known that the person who left the print was definitely Caucasian, then perhaps we should consider only a database containing stature profiles of Caucasians.
- If we are '90%' sure that the person who left the print walked with a limp, then perhaps we should consider only a database in which 90% of the stature profiles belong to people who walk with a limp
- Etc.

Clearly even if we were able to change to a more 'representative' database or restrict the existing database to people with the relevant criteria (and normally this is not possible because the database will only contain the stature profiles and few other details) the random match probability will also change. Hence it is impossible to assume that there is a truly objective random match probability (what makes a measurement objective or subjective is the supposed level of rigour of the measurement instrument). But, all of these issues are inevitable for any database for any area of forensic science. In other words there are no objective criteria by which our stature matching database could be ruled as any less 'reliable' than the most sophisticated DNA database[7]. This fictional example exposes exactly the kind of questions that need to be asked about any forensic database (including DNA databases), but which rarely are. Indeed, as described in [13] and [17], because of very different databases and different assumptions about how to use them, DNA experts in the UK and the USA report very different random match probabilities for the same person (often many orders of magnitude different such as one in a billion compared to one in a trillion). These differences, even when the probabilities are so low, matter greatly as we have already shown (and matter even more when we factor in the possibility of testing errors as we show in the next section).

Contrary to what was argued in [1] the 'statistical base' for determining the defence likelihood in stature matching is no less well defined than it is it for DNA. In fact it is actually much easier to get a relevant database, easier to do the matching, and easier to explain to a jury precisely what the match probability means. The match probabilities are as well defined (in fact less subjective) than those in the 'mature' science of DNA.

The 'scientific' quality or maturity of the type of forensic science being considered is therefore irrelevant as far as the statistical argument is concerned. The level of 'scientific' or 'statistical' quality is certainly not synonymous with very low defence likelihood figures. This point is important because there is a misconception that DNA evidence is scientific because it produces very low defence likelihood figures, while earprint or footprint evidence is less scientific because it rarely produces very low defence likelihoods. The value for the defence likelihood actually has nothing to do with the reliability of the data.

What matters is that in all cases of a match (whether it be DNA, fingerprinting, footprints, earprints, stature matching or anything else) the expert should be obliged to present the random match probability (possibly as a range) along with a statement about the limitations of the underlying data. For example,

- "The probability of finding this match in a person who was NOT the one who left their stature print at the scene is between one in a thousand and one in two thousand. This figure is arrived at from a database of 100,000 stature profiles of which 150 match the print at the scene."

The defence likelihood is inevitably a statement of subjective probability, as is any statement involving uncertainty.

So, given that there is no rational basis for declaring DNA 'statistics' as more 'scientific' than any other type of forensic match evidence, the prohibition from using likelihood ratios and Bayes on all but "DNA (and possibly other areas where there is a firm statistical base)" [1] makes no sense. The only consistent strategy would be to either allow its use for all forensic match evidence or to ban it for all (including DNA).

Clearly our argument is that the former should apply. To support this

7 It is, of course, important to note that the databases that provide a basis for the frequency statistics for DNA cases are far more comprehensive than for most other areas of forensic science, and this is presumably what the Judge in R v T was recognising. However, that does not alter the fact that DNA is not inherently more or less scientific than other areas of forensics currently lacking extensive databases. ■

we can point to the examples we have already provided where Bayes provided the correct results that match our intuition. But an even more convincing argument is to show that banning it for all arguments would mean that we would have to reject all statistical analysis as the following example should make clear:

- ***Example Case 1:*** A man is charged with a gaming offence, specifically that he was using a rigged coin when taking bets on whether the coin he was tossing comes up Tails. The defence hypothesis is that it was a fair coin. The prosecution hypothesis is that the coin was double-headed (so the punters were always sure to lose). The evidence E is that the coin landed as Heads on 9 out of 9 plays.

The point about this example is that the evidence is not only purely statistical, but that the statistics involved – coin tossing – allow us to use classic frequentist analysis and hence avoid any debates between Bayesian and non-Bayesian statisticians. Thus, everybody will certainly agree on the following:

- The defence likelihood is 1/512 (a half to the power of 9) because that is the probability of seeing 9 out 9 Heads given that the coin is fair. This analogous to the random match probability in a forensic case.
- The prosecution likelihood is is equal to 1, because that is the probability of seeing 9 out 9 Heads given that the coin is double-headed.

It is clear that the evidence favours the prosecution hypothesis more than the defence hypothesis. Moreover, the likelihood ratio of 512 can be proved to be the 'correct' factor in favour of the prosecution hypothesis. For, suppose that before the game was played a double-headed coin was added to a bag of 1000 coins that were known to be fair. Suppose also that the coin played in the game was selected randomly from this bag. Then before we see the evidence the odds must favour the defence hypothesis by a factor of 1000 to 1 (these are just the odds of selecting the double-headed coin). We know that there is a 1 in 512 chance of tossing 9 out of 9 heads in a fair coin. So, having seen the evidence the odds are 1000 to 512 (i.e. about 2 to 1) that the coin chosen was a fair coin. So, the evidence increases the odds in favour of the prosecution hypothesis by a factor of 512, but the defence hypothesis is still more likely. Hence, any rational juror should not convict the defendant on the basis of this evidence alone. Think of it this way: The chance of getting 9 Heads in 9 tosses of a fair coin (defence hypothesis) is still higher than the chance of selecting the one double headed coin from a bag of 1001 coins (prosecution hypothesis).

There is no dispute, therefore, that in the above hypothetical legal case the use of likelihood ratios and Bayes leads to the undisputedly correct conclusion. There is no 'statistical doubt' at all. Why is this important if the case is purely hypothetical? The answer can be gleaned by changing the assumptions very slightly. The assumption that a 'fair' coin has a probability of ½ of landing on Heads is a simplification. Even if we have no reason to believe there are double-headed coins in circulation the actual frequency of heads tossed in all coins in circulation is not a number that can be practically determined, and even if we had a very large datatabase of coins and toss results on them, it would certainly not be exactly equal to ½. These minor additional assumptions of reality, already shift us out of the 'purely statistical' scenario. Do these changes mean that our approach to evaluating evidence using likelihood ratios is no longer valid? Of course not. The exact same methods apply. All that has changed is our confidence in the original assumptions. We counter this uncertainty not by declaring the calculus of probability as invalid but by either stating our uncertainty clearly up front or using ranges instead of exact values.

All evidence in any case ultimately has a 'statistical basis'. The 'soundness' of the statistical basis is a spectrum where examples like that of case 1 above just happen to sit firmly at the 'soundest' end. The rationale for the ruling in [1] is not just that that there is some point at the opposite end of the spectrum at which the use of likelihood ratios become inappropriate, but that most types of forensic match evidence are even further beyond this point of the spectrum. Readers may yet be unconvinced that the minor change in the example already discussed is insufficient to push the example beyond this point, but surely the following leaves no doubt.

- ***Example Case 2:*** This case is the same as case 1, except for the fact there is no possibility that the coin was double-headed because the defendant clearly showed the coin to have a head and tail before tossing it. The prosecution hypothesis here is simply that the coin is 'biased' – i.e. will in the long run produce a greater ratio of Heads than Tails. It is still an offence to knowingly use such a coin. It is not known exactly what this bias is, but it is known that a magic shop in the area was selling special coins that looked real but were biased. These coins were all made with a different weighting and all that can be said with reasonable certainty was that the range of Heads 'bias' in these coins was between 0.6 and 0.7. The prosecution hypothesis is that the defendant used one of the

coins from this magic shop.

The evidence of 9 Tails in this example case has less 'statistical soundness' than the evidence of a stature match (or indeed any type of forensic match) in the previous section. Yet, it is easy to see that the use of likelihood ratios can be applied just as rationally in this example as in example case 1. Specifically:

The defence likelihood is the probability of seeing 9 Heads in 9 tosses given that the coin is fair. We cannot assume that the probability of tossing a Head on a fair coin is exactly ½. If we have a database of what are believed to be fair coins in which the lowest frequency of heads is 0.495 and the highest frequency is 0.505 then we could consider a range for the defence likelihoods using these as assumptions that are respectively least and most favourable to the defence hypothesis. So the least favourable is 0.00178 (that is 0.495 to the power of 9) and the most favourable is 0.002136 ((that is 0.505 to the power of 9).

The prosecution likelihood is the probability of seeing 9 out of 9 Heads given that the coin is biased. Here we have an infinite number of different prosecution hypotheses corresponding to every potential number between 0.6 and 0.7. Taking just the two extremes as those being respectively least and most supportive of the prosecution hypothesis we end up with respective prosecution likelihoods of 0.01 (that is 0.6 to the power of 9) and 0.04 (that is 0.7 to the power of 9) .

Despite the 'unscientific' nature of the evidence, we can conclude that, with the assumptions that most favour the prosecution, the likelihood ratio is 22.5 (0.04 divided by 0.00178), while with the assumptions that most favour the defence the likelihood ratio is 4.7 (0.04 divided by 0.00178). So, despite the clear lack of 'statistically' sound evidence, we can rationally conclude that the odds in favour of the prosecution hypothesis have increased by a factor of between 4.7 to 22.5. Indeed that is the only rational conclusion to make.

If the evidence made by either an expert or a member of the jury does not lead to the conclusion that the evidence supports the prosecution hypothesis by a factor of at least 4.7 to 1, assuming the most optimistic defence assumptions, then such a conclusion is irrational. If, as the ruling in [1] suggests, the use of likelihood ratios to explain the impact of this kind of evidence was not allowed in court, then the jury would be expected to do their own reasoning. This would mean, for example, that it would be acceptable to conclude that the evidence actually supported the defence by a factor of 100 to 1 if that is what their own 'method' led them to conclude.

Having, hopefully, countered the argument against using Bayes for 'non-scientific' statistical evidence, we next return to the crucial issue of why Bayesian reasoning has failed to make an impact on 'non-scientific' forensic match evidence.

5 Moving to more realistic assumptions: why the R v T ruling was understandable

Recall that the assumption of perfect testing accuracy, used so far in our forensic match evidence examples, means that:

- Someone with type X will always be tested to be of type X. This means that there is *zero probability of false negatives*:
- Someone who is not type X will never be determined to be of type X. This means that there is *zero probability of false positives*:

In the case of stature matching neither of these assumptions is at all realistic, as they would require all of the following to hold:

- Stature traces (taken either from the crime scene or taken directly from the defendant) are always 'perfect' (so, for example, there is no possibility that distortion of the photographic/video evidence is such that the person's height could be determined to be 136 centimetres as opposed to 132 centimetres).
- The process of analysing the stature trace is infallible (so, for example, it is impossible for one stature expert to determine from a photo that the person is a man and for a different stature expert to determine from the same photo that the person has is a woman.
- Stature prints can never be tampered with before they are examined by the expert.
- A person's stature profile can never change (so, for example, if their waistline was 65 centimetres at the time they made the print, then when they are subsequently tested their waistline will inevitably be within 2 centimetres of 65 centimetres).

But these assumptions (especially the first three) are even more dubious in the case of DNA evidence than in the case of stature matching. If any of these statements is not true then neither the false negative probability nor false positive probability will be zero.

Yet, while it is accepted that random match probabilities need to be 'statistically sound' the same is never demanded of the probabilities of false positives and false negatives. Indeed, in many analyses they are simply (but wrongly) assumed to be zero, while in others (including DNA analyses) they are simply stated as subjective estimates. This prompted the authors in [36] to ask pointedly:

- "Why are the two possible sources of error in DNA testing treated so differently? In particular, why is it

considered essential to have valid, scientifically accepted estimates of the random match probability but not essential to have valid, scientifically accepted estimates of the false positive probability?"

The authors in [36] provide a strong argument on why it is just as critical to include the false positive probability as the random match probability. However, their omission of the case for the false negative probability (presumably because they only consider the scenario where there have been positive tests for both the source and target) is itself an oversight. Even assuming that both tests are positive, the Bayesian reasoning still requires us to know the probability of a true positive (which is equal to one minus the probability of a false negative, as shown in Table 1). The calculations in [36] assume that the true positive probability is 1 (and hence the false negative is 0). This is unrealistic. By assuming the more realistic assumption of non-zero false negative probability we allow for the scenario in which it is possible that some other suspect with profile type X was never considered because they were wrongly tested as not being type X.

It follows that, as soon as we drop our assumption about 'perfect testing' (as in practice we surely must), then the notion of a sound 'statistical base' for DNA compared with other types of forensic evidence becomes even more blurred than we previously explained, since there is no 'statistically sound' base for determining the error probabilities in DNA testing. If anything it will surely be easier and more objective to determine the values and exact causes of false positive and false negative errors for stature matching than it would be for DNA. It would also be easier to explain to a jury precisely what these errors are.

It should be clear now, conceptually, that there is no more justification for using the probabilities that arise from DNA as there is in using the probabilities that arise from just about any other type of forensic match evidence.

However, it turns out as we show in the next section, that as soon as we incorporate the potential for testing error in a Bayesian argument things become complex. It is not clear, for example, that these issues were properly addressed for the footwear evidence that was the subject of the R v T ruling, and this possibly makes the judge's lack of trust in the transparency and accuracy of the results of the Bayesian analysis more understandable.

6 The problem with scaling up Bayesian arguments

In [35] the authors state:

- "The best argument for the application of Bayesian theory in forensic science is to show that the theory agrees with personal intuitions, when inference problems are simple and intuitions are reliable, and that it helps to go beyond them, when problems become more complicated and intuitions are not so reliable."

This is exactly the strategy we have suggested. The problem with this strategy is that as soon as we recognise that the false positive and false negative probabilities may not be zero, the 'simple' problem actually becomes very difficult to explain using the intuitive, tree-diagram approach. In fact, although several authors have tried it, we are not aware of the problem being presented correctly in any way *other* than by using the formulaic approach. And, even then, the presentations fail to include the false negative probability. The net effect is that, unless people are prepared to understand the formulas they will ***not*** be able to see that the theory agrees with personal intuitions even in the 'simple' problem case. This goes some way to explaining why the basic misunderstandings discussed in Section 3 persist in the law.

To explain what the problem really is and how we might solve it, let us review the relevant information we have to consider for any forensic match case when the testing cannot be assumed to be perfect:

- ***Prosecution hypothesis (H1):*** "The target is the source" (unchanged)
- ***Defence hypothesis (not H1):*** "The target is not the source" (unchanged)
- ***Evidence E1:*** "The source profile is tested to be of type X" (note: we can no longer assume the source profile actually is type X)
- ***Evidence E2:*** "The target profile is tested to be of type X (note: we can no longer assume the target profile actually is type X)

Because of the probability of false positives we cannot assume from the above evidence that either the source or the target have type X. Instead these assertions are also unknown hypotheses:

- ***Source type hypothesis (H2):*** "The source profile really is type X" (true or false)
- ***Target type hypothesis (H3):*** "The target profile really is type X" (true or false)

What we have, therefore, is a problem involving five 'variables' H1,

Table 1. Error probabilities

Actual Type	Not X	Not X	X	X
Test result	Not X	X	X	Not X
	(True negative)	(False positive)	(True positive)	(False negative)
Probability	1-*u*	*u*	1-*v*	*v*

H2, H3, E1, E2 which can all be true or false (in order to do the necessary Bayesian reasoning). But this means there are 32 different scenarios representing the different possible true/false combinations (although some are 'impossible' and some are not observed, such as false values for the evidence). We can show this in a tree diagram - Figure 4 - but of course it is now far more complex than before; possibly too complex for lay people to understand.

Even when we ignore the impossible branches and all the scenarios in which the evidence E1 and E2 is false, we are left with six scenarios that need to be incorporated in the likelihood calculations:

- ***Scenario 1*** (this is the 'normal' prosecution scenario) in which H1, H2, H3, E1 and E2 are all true. This scenario has probability $m(1-v)^2$
- ***Scenario 2*** (this is an often ignored prosecution scenario) in which H1 is true (the target is the source) but the target is not actually type X. Both the test of the target and source, however, incorrectly result in an X. This scenario has probability $(1-m)u^2$.
- ***Scenario 3*** (this is the 'normal' defence scenario) in which the tests are correct but the match is coincidental. This scenario has probability $m^2(1-v)^2$.
- ***Scenario 4*** this is the defence scenario in which the target is incorrectly tested to be type X. This scenario has probability $m(1-m)(1-v)u$.
- ***Scenario 5*** this is the defence scenario in which the source is incorrectly tested as type X. This scenario has probability $(1-m)mu(1-v)$.
- ***Scenario 6*** this is an often ignored defence scenario in which both the source and target are wrongly tested to be X. This scenario has probability $(1-m)^2u^2$.

The prosecution likelihood is the sum of the probabilities for scenarios 1 and 2, while the defence likelihood is the sum of the probabilities for

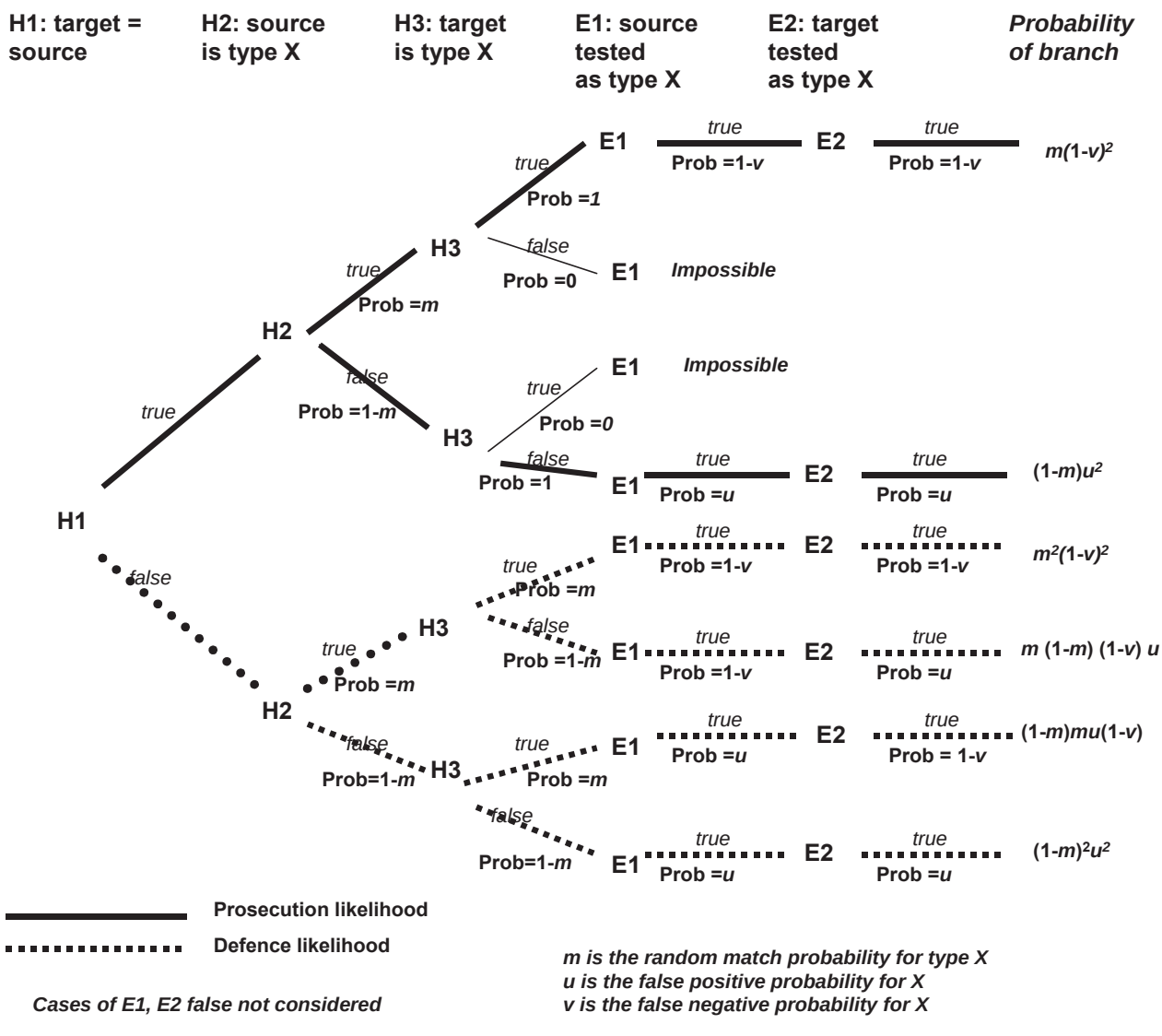


Fig 4. Bayes calculation explained visually (but this time possibly too complex to understand)

scenarios 3, 4, 5, and 6.

The problem is that the likelihoods, and hence also the resulting likelihood ratio, are not sufficiently 'simple and intuitive' to ensure that people can check they 'agree with personal intuition' (which is why it is not even worth the effort here of going through the motions). Resorting to the Bayes formulas, of course, only makes things much worse.

The example also shows that, even for experienced Bayesians, it can be difficult to model the problem in this way and difficult to perform the calculations (as we argued earlier, we have not previously seen a full solution of this problem taking into account both error probabilities). And this example still has many simplifications: it assumes that all three probabilities (random match, false positive, false negative) are all 'point' values, whereas in practice they would be uncertain distributions [11]; it assumes that all variables have just two possible values (true and false); it assumes that there is just one print; and it assumes the only evidence is the match evidence. When we include further aspects of reality (especially including multiple, related pieces of evidence) the possibility for producing the correct Bayesian calculations manually (with or without formulas) – let alone being able to explain them to a lay person – are non-existent.

In our view the best way to minimise this problem is to use Bayesian networks (as explained in [19] [25] [35]). By exploiting assumptions of independence between variables, a Bayesian network (BN) model is typically compact and efficient, since it avoids the problem we saw above whereby we had to consider all possible combinations of variable values (statisticians express this formally by saying that 'it is not necessary to consider the full joint probability distribution').

A BN (see Figure 5) is a graphical model that shows the dependency relationships between the unknown variables of interest (each variable is represented by a node in the graph).

In addition to the graphical structure we define, for each node, a probability table that defines the probability values for the node given the different combinations of parent states. For example, the probability table for node "target tested as X" simply encodes the error rates as shown in Table 2.

It is much easier to build and run this model with the relevant information than it is to either construct a tree as before or to produce the necessary formulas. Once built we can enter evidence and get the calculations immediately as shown in Figure 6 (this shows the results using a standard BN tool). Here we actually compare the results under two different sets of assumptions:

- In a) we encoded the assumption of perfect testing

Table 2. Probability table for node "target tested as X"

Target is type X	False	True
Target tested as X (False)	1-*u*	*v*
Target tested as X (True)	*u*	1-*v*

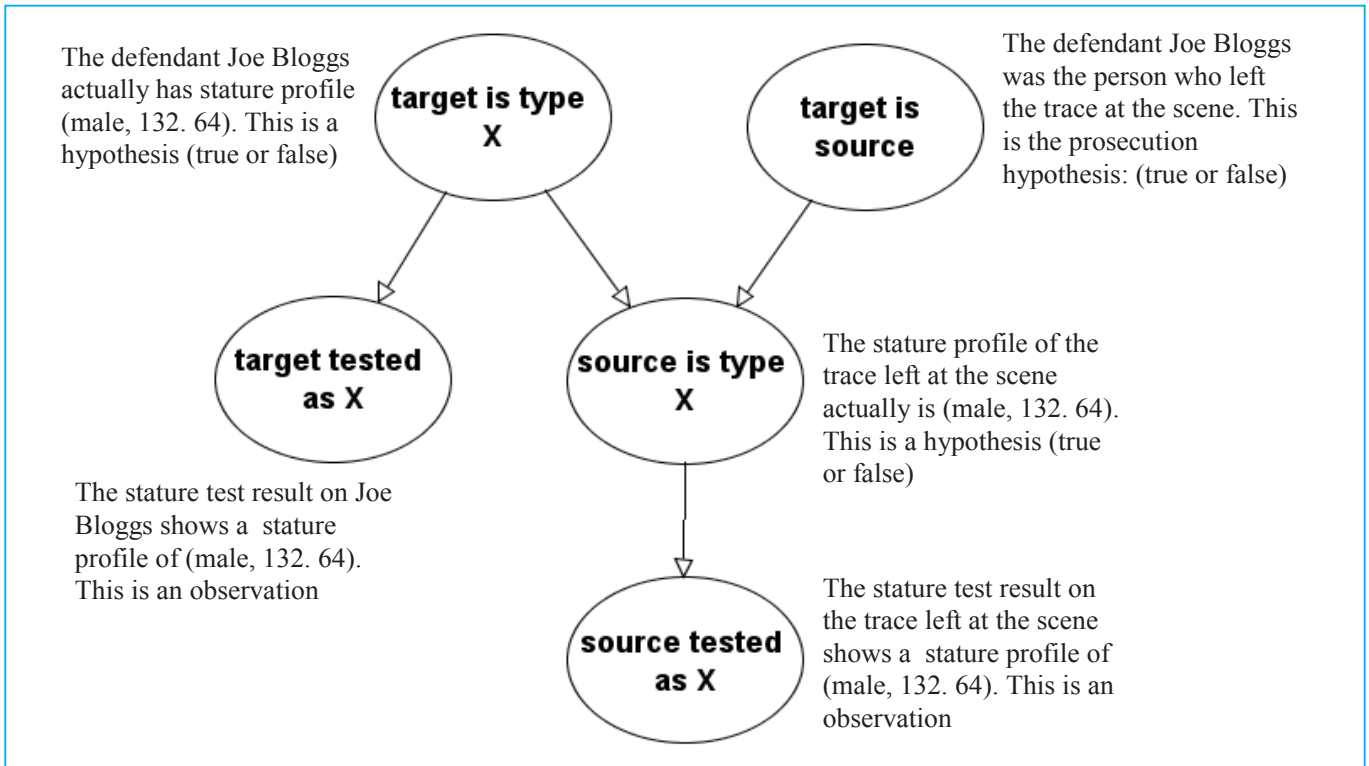


Fig 5. Bayesian network solution to the problem (with an example showing what the nodes would mean for a specific stature matching case)

accuracy (i.e. u and v are both set to zero).

- In b) we encoded the assumption that u (false positive) is 0.1 and v (false negative) is 0.01.

Although in both cases we assume the same match probability (1/100) and the same prior (50:50)[8] for the prosecution hypothesis ("target is source") the difference is quite dramatic. Although the evidence is identical in both cases, in the former t posterior odds[9] are 100 to 1 in favour of the prosecution hypothesis, whereas in the latter the posterior odds[10] are only 65 to 35 (i.e. about 2 to 1) in favour of the prosecution hypothesis.

Not only does the BN remove the need for performing the difficult Bayesian calculations manually, but its graphical representation is easy for a lay person to understand. We are not, however, suggesting that the BN model is what should be presented court. It should be used for pre-trial analysis of the evidence by forensic experts, preferably using different scenarios for the different ranges of match probabilities and error probabilities. The model structure should be agreed between legal teams and forensic experts on both sides. All that should be presented in court are clear statements of the prior assumptions being used (the match probabilitity, and error probabilities) and the results of the calculations under the different assumptions.

A detailed history of BNs in legal reasoning, along with proposed mechanisms for using them in practice can be found in [19] and [20].

7 Conclusions and recommendations

The ruling in R v T displayed some fundamental misunderstandings, including assertions that can be shown to be either illogical or irrational. However, the presentation of the Bayesian argument and likelihood ratios in the original case was both inadequate and inaccurate, as it has been in many similar cases. We have argued that this may be, in large part, due to the continued failure of the statistical community to provide the necessary support to forensic scientists and lawyers. That fundamental probabilistic reasoning should have therefore been discredited in the R v T ruling is hard for statisticians to take but, even in our view as Bayesians, was totally understandable.

If statisticians continue to believe that the way to explain their arguments in legal reasoning is by using first principle calculations and formulas, then the future for Bayes in the law is doomed.

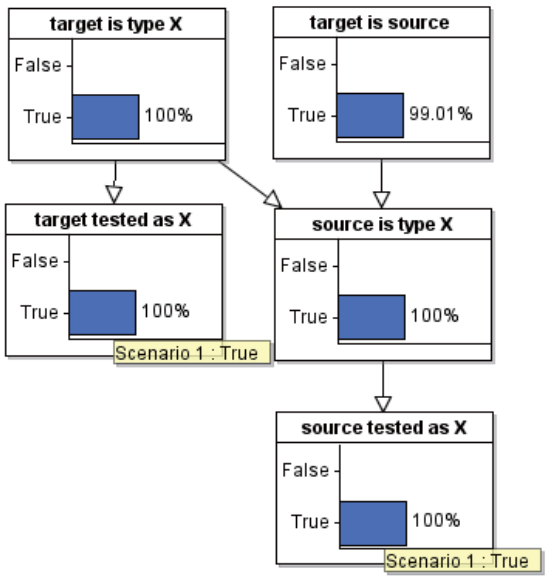
The challenge over the next few years is to get to the situation whereby everybody in the legal system understands the difference between

a. the genuinely disputable assumptions

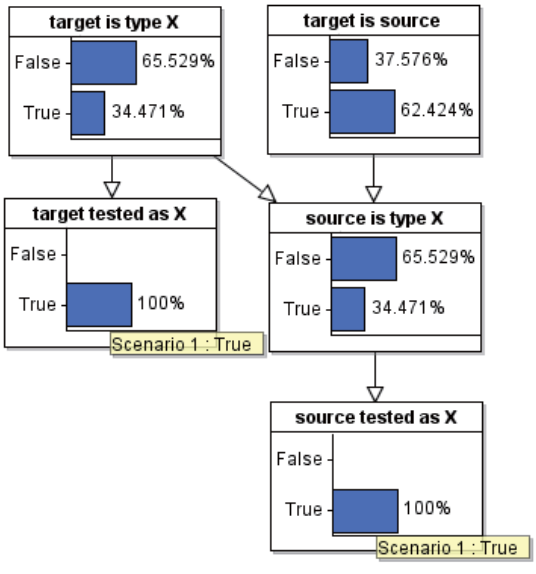
8 Recall that, by assuming a 50:50 prior, we know that the posterior odds are equal to the likelihood ratio..

9 The likelihood ratio is 100, meaning equivalently the probability the prosecution hypothesis is true is 100/101 = 99.01%)

10 The likelihood ratio is 65/35, meaning equivalently the probability the prosecution hypothesis is true is 65%). ■



a) Impact of evidence when error probabilities are assumed to be zero



b) Impact of evidence when false positive rate is 0.1 and false negative is 0.01

Fig 6. Comparing the different impact of the evidence when we assume different error rates (in both cases the match probability is 1/100 and the prior probability for "target is source" is ½)

that go into a probabilistic argument; and

b. the Bayesian calculations required to compute the conclusions based on the different disputed assumptions.

Crucially, there should be no more need to explain the Bayesian calculations in a complex argument than there should be any need to explain the thousands of circuit level calculations used by a calculator to compute a long division. Lay people do not need to understand how the calculator works in order to accept the results of the calculations as being correct to a sufficient level of accuracy. The same must eventually apply to the results of calculations from a Bayesian analysis. The more widespread use of tools such as Bayesian networks makes this a feasible target.

However, ensuring that the distinction between a) and b) is firmly understood by lawyers is only a necessary requirement for the more widespread takeup of Bayes. There is, as yet, no significant understanding among lawyers that any legal argument can ever be couched in Bayesian terms. The challenge for statisticians is to break down this significant cultural barrier. In this challenge we also propose that the use of Bayesian network models will be useful, but any progress requires a major educational effort aimed at all levels of the criminal justice system. It requires 'buyin' from senior members of the legal profession and politicians, as well as a united front presented by the community of statisticians.

If we can meet these challenges then there is no reason why Bayes should not become a standard (possibly even the central) method for evaluating evidence in every aspect of legal reasoning.

Acknowledgements

We are indebted to the following for providing comments, corrections, relevant information, and contacts: David Balding, Sheila Bird, Tiernan Coyle, Ian Evett, Joseph Kadane, Jay Koehler, Margarita Kotti, Amber Marks, William Marsh, Geoff Morrison, Richard Nobles, David Ormerod, Mike Redmayne, David Schiff, Bill Thompson, Patricia Wiltshire.

References

[1] (2010). R v T. EWCA Crim 2439 http://www.bailii.org/ew/cases/EWCA/Crim/2010/2439.pdf
[2] R v Adams [1996] 2 Cr App R 467, [1996] Crim LR 898, CA and R v Adams [1998] 1 Cr App R 377
[3] (2008). R v Kempster, EWCA Crim 975
[4] (2002). R -v- Dallagher, EWCA Crim 1903
[5] (2007). R. v. George, EWCA Crim 2722
[6] Brandon Mayfield v. U.S.A (04cv1427)
[7] (2007). "Ruling Casts Doubts On Infallibility Of Fingerprint Evidence." from http://www.wbaltv.com/news/14433107/detail.html.
[8] Adam, C. (2010). Essential Mathematics and Statistics for Forensic Science. Chichester, John Wiley & Sons.
[9] Aitken, S. G. G. (2000). Interpretation of evidence, and sample size determination. Statistical Science in the Courtroom. J. L. Gastwirth. New York, Springer: 1-24.
[10] Aitken, C. G. G. and F. Taroni (2004). Statistics and the evaluation of evidence for forensic scientists (2nd Edition), John Wiley & Sons, Ltd.
[11] Balding, D. J. (2005). Weight-of-Evidence for Forensic DNA Profiles, Wiley.
[12] Berger, C. E. H., J. Buckleton, C. Champod, I. Evett and G. Jackson (2011). "Evidence evaluation: A response to the court of appeal judgement in R v T." Science and Justice 51: 43-49.
[13] Buckleton, J., C. M. Triggs and S. J. Walsh (2005). Forensic DNA Evidence Interpretation, CRC Press.
[14] Broeders, T. (2009). Decision-Making in the Forensic Arena. In "Legal Evidence and Proof: Statistics, Stories and Logic". (Eds H. Kaptein, H. Prakken and B. Verheij, Ashgate) 71-92.
[15] Coyle, T. (2010). Trace and Contact Evidence. Crime Scene to Court: Essentials of Forensic Science. P. C. White, Royal Society of Chemistry.
[16] Dror, I. E. and R. Rosenthal (2008). "Meta-analytically quantifying the reliability and biasability of forensic experts." Journal of Forensic Sciences 53(4): 900-903.
[17] Evett, I. W., L. A. Foreman, G. Jackson and J. A. Lambert (2000). "DNA profiling: a discussion of issues relating to the reporting of very small match probabilities." Criminal Law Review (May) 341-355.
[18] Evett, I. W. and B. S. Weir (1998). Interpreting DNA evidence : statistical genetics for forensic scientists, Sinauer Associates.
[19] Fenton, N. E. and M. Neil (2008). Avoiding Legal Fallacies in Practice Using Bayesian Networks. Seventh International Conference on Forensic Inference and Statistics. Lausanne, Switzerland
[20] Fenton, N. and M. Neil (2010). "Comparing risks of alternative medical diagnosis using Bayesian arguments." Journal of Biomedical Informatics 43: 485-495.
[21] Gastwirth, J. L., Ed. (2000). Statistical Science in the Courtroom. New York, Springer-Verlag.
[22] Gigerenzer, G. (2002). Reckoning with Risk: Learning to Live with Uncertainty. London, Penguin Books.
[23] Haigh, J. (2003). Taking Chances: Winning with Probability, Oxford University Press.
[24] Jackson, A. R. W. and J. M. Jackson (2004). Forensic Science. Harlow, Pearson.
[25] Kadane, J. B. and D. A. Schum (1996). A Probabilistic Analysis of the Sacco and Vanzetti Evidence, John Wiley & Sons.
[26] Kaye, D. H. (2009). "Identification, Individualization, Uniqueness." Law, Probability & Risk 8(2): 85-94.
[27] Langford, A., J. Dean, et al. (2005). Practical Skills in Forensic Science. Harlow, Pearson.
[28] Lucy, D. (2006). Introduction to Statistics for Forensic Scientists. Chichester, UK, John Wiley & Sons Ltd.
[29] Morrison, G. M. (2012). "The likelihood ratio framework and forensic evidence in court: a response to RvT." International Journal of Evidence and Proof 16(1).
[30] Redmayne, M. (2001). Expert Evidence and Criminal Justice, Oxford University Press.
[31] Redmayne, M., P. Roberts, C. Aitken and G. Jackson (2011). "Forensic Science Evidence in Question." Criminal Law Review (5): 347-356.
[32] Robertson, B., G. A. Vignaux and C. E. H. Berger (2011). "Extending the confusion about Bayes." The Modern Law Review 74(3): 444-455.
[33] Robertson, B. and T. Vignaux (1995). Interpreting Evidence: Evaluating Forensic Science in the Courtroom, John Wiley and Son Ltd.
[34] Saks, M. J. and J. J. Koehler (2007). "The Individualization Fallacy in Forensic Science Evidence." http://works.bepress.com/michael_saks/1
[35] Taroni, F., C. Aitken, P. Garbolino and A. Biedermann (2006). Bayesian Networks and Probabilistic Inference in Forensic Science, John Wiley & Sons.
[36] Thompson, W. C., F. Taroni and C. G. G. Aitken (2003). "How the probability of a false positive affects the value of DNA evidence." Journal of Forensic Sciences 48(1): 47-54.
[37] White, P. C., Ed. (2004). Crime Scene to Court: Essentials of Forensic Science, Royal Society of Chemistry. ■

2157118X.4.1.R4 Peer Reviewed
Received 5 November 2014
Received in revised form 10 December 2014
Accepted 15 December 2014
Available online 28 December 2014

Pattern Exploration and Construction Process of the System of Judicial Authentication Aid in China

JIN Xin [*, a] ***ZHAO Jie*** [*, b]

a Nanjing normal university, Nanjing 210046. *JIN Xin*, from Anji, Zhejiang, lecturer of Nanjing Normal University Law School, judicial authentication engineers, appraiser of judicial authentication center of Nanjing Normal University.
b Nanjing normal university, Nanjing 210046. Prof. *ZHAO Jie*, from Huang county in Shandong, professor of Nanjing Normal University Law School, director of the center for law experiment, director of judicial authentication center.

* Corresponding author. Tel : 025-83249948; E-mail: jasonjx1983@163.com.

Abstract Judicial authentication assistance is mainly to maintain the legitimate rights and interests of the vulnerable groups by helping them to solve the problems when they cannot afford the appraisal charge to enter the proceedings or when they have difficulties in proofing in the lawsuit. The system of judicial authentication aid can effectively protect the judicial authentication equality of the eligible recipients in the judicial process and can secure judicial justice and the implementation of social equity and justice. We have not formulated the national unified judicial authentication legal aid measures in our country at present, but many provinces, cities and counties have developed provisions to specify judicial authentication assistance norms. However, the makers of the local specifications are various and the contents are different, the patterns of judicial authentication aid are not clear, and the programs of judicial authentication aid are still worth trying and exploring. This paper will make an analysis of the two aspects: the choice of a pattern and the construction of a program.

Keywords: Forensic science, Judicial authentication, Legal aid, Pattern, Program.

1 The present situation of judicial authentication aid in China

The system of judicial authentication is an important part of the lawsuit system and the system of judicial authentication aid is also an important part of legal aid system. Legal aid system, also called legal aid, is a kind of judicial relief system generally used many countries in the world. Specifically speaking, it is a legal security system provided by the country in each link and each level of the implementation of the legal system to the disadvantageous of a society who cannot typically use legal relief means to maintain their basic social rights because of economic hardship and other factors through breaking charge. As a state act to implement social justice and judicial justice and to maintain basic rights of the citizens, it plays a very important role in the legal system of a country. Judicial authentication assistance system, also called judicial authentication aid system, its meaning is that to secure the practical implementation of all rights given by law to the citizens in real life in the process of litigation activities, the country provides measures to avoid, reduce charge for the parties concerned who need judicial authentication proof to protect us from illegal infringement but cannot afford the appraisal charge to secure their legitimate rights and interests and also to maintain judicial justice.

China enacted "The Legal

Project Fund: This paper is the class achievement of the research project “Research on the system of judicial authentication aid” (Project grant number: 2013SJD820005) of Philosophy and social sciences in colleges and universities of Education Department of Jiangsu Province hosted by the author. ■

Aid Ordinance" in 2003, which didn't make it clear that judicial authentication aid is included in it. At present, the legal aid obligations are mainly undertaken by the legal aid centers at all levels. The participants Mainly includes lawyers, notaries and grass-roots legal services staff. They provide legal services that can break the corresponding law service fee for the parties concerned who have economic difficulties and cannot afford legal service fees or those who are involved in special cases by using their own legal professional knowledge. In our country, there are not clear requirements that the judicial authentication institutions and the judicial appraisers should bear the corresponding legal aid obligations.

In January 1997, the department of justice legal aid center was set up. Then, China's legal aid foundation approved by the state council was set up. On July 16, 2003, the state council published "Legal Aid Ordinance", which was to be implemented from September 1, 2003. "Legal Aid Ordinance" makes a comprehensive and specific provisions for the basic problems such as the nature, task, organization, scope, program, implementation and the legal responsibility of the legal aid in our country. Its implementation marked that the legal aid work in our country entered a new stage of legalization and standardization and it provided the necessary security of laws and regulations for the further promotion and specifications of the legal aid work. It played an important role in ensuring the citizens who had economic difficulties to obtain the necessary legal services and in promoting social fairness and justice and the construction of a harmonious socialist society. On February 21, 2012, the justice department ministerial meetings reviewed "Procedures for the Legal Aid Cases Handled by Rules", which was to implemented from July 1, 2012. the department of justice put forward that efforts should be taken to deal with over millions of legal aid cases throughout the year.

We have not formulated the national unified judicial authentication legal aid measures in our country at present. So far, judicial authentication aid hasn't hit as a headline and it took time and efforts to find in "The Central Subsidies to Local Legal Aid Case Special Fund Management Interim Measures" and files such as the judicial appraisal aid rules of some provinces, cities and counties. Since July 21, 2006, when the Justice Department of Sichuan Province published the more systematic "Interim Measures for Judicial Authentication Assistance in Sichuan Province", many provinces, cities and counties have formulated provisions to standardize judicial identification aid. Most of the provisions are from the local judicial administrative department, a few from the courts and the commodity price departments. Among those provisions, some systematically and clearly indicate the problems referring to the purpose of the judicial authentication relief, the scope and conditions of judicial authentication aid object, operating procedures, the rights and responsibilities of the authentication institutions and the appraisers that undertake the obligation to help. However, some of the local provisions are difficult and others are simple. Besides, some of the problems are not paid full attention to.

Take Jiangsu Province for example, Nanjing City Judicial Bureau issued "Opinions about Judicial Authentication Assistance Work" which indicated that judicial authentication aid is included in legal aid since June 2007. Nanjing Province has determined the scope of judicial authentication assistance, which is the forensic clinical identification, forensic pathology identification, forensic psychiatric assessment, forensic toxicology evaluation, paternity test, documents identified, trace identification etc. of the recipients of legal aid cases approved by the legal aid centers of the two levels of the municipals, districts (counties). In the aid process, the legal aid center uniformly accept and approve them first and then the management office of the judicial authentication assigned relevant judicial authentication institutions to take turns to deal with the judicial appraisal aid issues. In March 2012, the Kunshan City Court and Kunshan City Price Bureau assembled "The Measures for the Implementation of the Judicial Identification of Legal Aid" on the basis of coordination and communication of many sides. It provided free price verification aid for the parties concerned who cannot afford the appraisal charge because of economic difficulties. The "Measures for the Implementation" had a total of seven items. From the aspects such as the connotation of the judicial identification of legal aid, the applicable objects, the examination and approval procedures, matters that should be avoided, it indicated specific operation scheme of judicial

identification of legal aid, severed the responsibilities that shall be investigated for. In 2012, the judicial identification of legal aid was adopted in the project plans promoting the innovation in social management of the provincial courts by Jiangsu provincial high court.

2 Pattern exploration of the system of judicial authentication aid

The pattern selection of the system of judicial authentication legal aid is crucial, which determines whether the system can be established and carried out smoothly in the judicial practice in our country. The patterns of legal aid can be roughly divided into two kinds: society-dominant pattern and state-welfare pattern.

2.1 Society-dominant pattern

Society-dominant pattern of legal aid can be seen in Anglo-American law system most of the time. In this mode, social groups, the lawyer organizations, law offices and the lawyers themselves are dominant in legal aid. They are responsible for the specific implementation of legal aid. The country plays a secondary or supporting role in the whole mechanism. It only guides and specifies social organizations and personal legal aid in such two aspects as Policy guidance and legislation. Case guidance is not included.

2.2 State-welfare pattern

The representative countries of state-welfare pattern are Canada and Australia and some European countries. The country has the responsibility to ensure every citizen to resort to the law and to seek the judicial relief and to get the fair trial rights. The country should provide necessary legal assistance for the citizens who are in poor economic conditions or adverse situations. Legal aid is mainly state or government behaviors and it gradually begins to develop towards social direction. State-welfare pattern of legal aid makes the country dominant in legal aid system, in which legal aid matters are to be dealt with by specialized agencies set up by the country. The participants of the main body of society-dominant pattern are only in a subordinate position.

We lack developed social assistance organizations in our country at present, therefore it is not suitable for society-dominant pattern to set up the system of judicial authentication legal aid. The characteristics of state-welfare pattern are their universality and sociability. It enables most of the people to enjoy legal aid services but its disadvantage is obvious, which is that the state financial burden is heavy. Therefore, the author thinks that We should explore a legal aid pattern that conforms to our national conditions. Such kind of pattern should give priority to state-welfare pattern and take society-dominant pattern as a supplement.

3 Construction process of the system of judicial authentication aid

3.1 The object of judicial authentication aid

Judicial authentication aid ensures the right of vulnerable groups to appeal, so the qualifications of the object of judicial authentication aid must be limited. What kind of group should be given aid to is the primary issue of the system of judicial authentication aid. What is ”vulnerable groups”? Vulnerable groups, a new term of political economy, refers to those people who obtain social wealth less and more difficult in the social production and living and therefore are in poorer condition because of their relatively weak strength and power of the group. Vulnerable groups are defined according to the social status, living conditions, physical characteristics and physical state of the people. It in name is a virtual community. It is generally referred to as the people who have difficulties in living, or people whose abilities are not enough, or people who are marginalized and scattered around the society.

3.2 The process of the system of judicial authentication aid

3.2.1 Application for judicial authentication aid

The parties concerned who are eligible for judicial authentication aid apply to the judicial administrative department that has jurisdiction. It should be a written application and the the application form should be filled out. At the same time, the following materials should be submitted: 1) ID card, residential certificate or temporary residence permit of the applicant. 2) Documents of the applicant’s economic condition issued by the street agency, the township people's government and the civil affairs department at or above the county level that are in the place of the applicant’s residency or habitual residence. 3)The case materials related to appraisal.

When the applicant has no capacity for civil conduct or limited

capacity for civil conduct, his or her guardian or legal representative can apply instead. At the same time, the proof of identity and the agency agreement of the guardian or legal representative should be submitted either.

3.2.2 Review of judicial authentication aid

The judicial administrative department should review the application submitted by the applicant. Ways of the review include examination as to substance and a written review. In order to save the judicial resources and improve the efficiency of lawsuit, the written review is mainly used. Make sure that the economic condition of the applicant cannot afford the appraisal charge. The applications submitted by the eligible applicants are determined to accepted and dealt with. The applicants should be informed of the right to the administrative reconsideration at the same time.

3.2.3 Implementation of judicial authentication aid

If the judicial administrative department determined to offer aid, the court and Judicial authentication institutions where the applicant and case exist. If the case is directly entrusted by the recipients, the recipients and the authentication institution will sign a judicial authentication assistance agreement. If the case is entrusted by the court, the recipients, the court and the authentication institution will sign a judicial authentication assistance agreement, which indicates the rights and obligations of the three sides and the specific items of the assistance. After the assistance agreement is signed, the recipients' appraisal charge can be delayed. After the case is over, the appraisal charge is paid by the losing party. If the losing party is the recipient, the appraisal charge can be reduced or freed according to specific situation. If both parties bear responsibilities, the appraisal charge is shared by both. If the recipient still have difficulties, the appraisal charge can be reduced or freed.

3.3 Security of judicial authentication aid

3.3.1 The government grant

The source of the funds of judicial authentication aid is mainly the government grant at present. The government should include the special funds of judicial authentication aid in the annual government budget. The judicial administrative departments at all levels should control the spending.

3.3.2 The fund of judicial authentication aid

At present, the grant of Chinese government given to legal aid is far from sufficient to meet the demand of the assistance in the judicial practice. Besides waiting for the rise of the government grant year after year, other sustainable Sources of aid should be developed.

The country should mobilize all social resources, raise funds variously, establish judicial authentication aid funds. The funds should b operated by special funds operating companies to ensure the benign development of the fund. The country should take effective measures to supervise the operating of the foundation.

3.3.3 Implementation of tax preferential policies

In the aspect of security of the system of judicial identification aid, it is necessary to take tax preferential treatment measures. The social judicial authentication institutions of our country are self-sustaining on the fiscal. Accepting the entrust of aid identification is bound to increase the financial burden, so we can provide paid services or income for the society to support the daily operations. We can implement preferential measures such as the appropriate tax breaks to the authentication institutions that accept the entrust of the aid identification.

References

[1] NING Song. A legal aid system of forensic authentication should be established. Chinese Journal of Forensic Sciences, 2004, (2): 27-29. (In Chinese)
[2] HAN Wei, DONG Xin-wei. Problems and suggestions of aid of judicial expertise system. Chinese Journal of Forensic Sciences, 2009, (2): 75-76. (In Chinese)
[3] WANG Su-fang, LI Guo-qin. On China's judicial system identification assistance. Journal of Lanzhou, 2011, (11): 220-221. (In Chinese)
[4] XU Si. On the Urgency and Necessity of the legal aid relating to forensic authentication.Chinese Journal of Forensic Sciences, 2003, (2): 50-52. (In Chinese)
[5] CHEN Shang-kun, CAI Yi-sheng. On public-good nature of expert testimony. Changbai Journal, 2007, (2): 72-75. (In Chinese)
[6] WANG Yue, YI Min, CHEN Ru-chao. Problems and analysis of the legal aid system in forensic appraisal. Chinese Journal of Forensic Sciences, 2011, (3): 14-20. (In Chinese)
[7] JIA Zhi-hui, KE Chang-lin. On China's judicial system identification assistance. Journal of Jiangxi Public Security College, 2006, (3): 54-56, 108. (In Chinese)
[8] JIANG Kui. On construction of aid of judicial expertise system in China. Chinese Journal of Forensic Sciences, 2006, (2): 16-18. (In Chinese) ■

2157118X-4-4.RS Peer Reviewed
Received 4 June 2014
Received in revised form 28 June 2014
Accepted 30 June 2014
Available online 30 December 2014

An Evidence Based Approach to Curriculum Development in Forensic Medicine -- Point of View

Edussuriya D H [a], ***Marambe K N*** [b], ***Abeyasinghe N*** [c], ***Jayawickramarajah P T*** [d]

a Department of Forensic Medicine, University of Peradeniya, Sri Lanka.
b Medical Education Unit, University of Peradeniya, Sri Lanka.
c Department of Forensic Medicine and Toxicology, University of Colombo, Sri Lanka.
d Medical Education and Research Unit, Eastern University, Sri Lanka.

In recent years, educational institutions around the world have been increasingly confronted with the challenge of making curricula relevant to the needs of the time with regular up-dating in terms of actual necessity [1]. Financial constraints and increasing accountability for the use of public funds further highlights the need for universities to clarify their goals based on concrete data.

Undergraduate medical curricula cannot include all major specialties and the criteria for including content into the undergraduate curricula should not be based purely on the enthusiasm of teachers. It is clear that the more common strategy of offering a course in the hope that at some future time it may serve the physician graduate is not satisfactory. Thus, what should be taught at undergraduate level should be based on the needs and expectations of society from medical graduates. Rapid changes in medical information technology and the growth of government and patient influences (satisfaction rates, legal rights, etc.) should be factors which direct the behaviour of a medical professional and should be reflected in the content of medical programs. Since the requirements of patients, society and peers make a professional career much more complex now than it has been in the past, it emphasizes the need for medical curricula to be focused on outcomes supported by a strong foundation of educational theory and research. It is therefore necessary to develop an evidence based approach to making decisions on the content that is to be included in undergraduate curricula.

The importance of determining what needs to be included in curricula is clearly illustrated in the statement "effective teaching can be more dangerous than no teaching at all if it is not really relevant" [2]. One way of determining what needs to be taught to students is to rely on the judgments of experts to determine what a neophyte in the profession ought to know and ought to be able to do. However, reliance on expert opinion per se to determine educational goals result in "curricula being crammed with an ever burgeoning quantity of new and highly specialized knowledge…." [3]. Alternately gathering evidence about what a competent medical officer needs to know and needs to be able to do by feedback from stakeholders, critical incident technique, task analysis or by epidemiological studies would render more relevant scientific evidence regarding curriculum development.

Over the past few decades, the emphasis in medical education has been on methods of teaching, learning and assessment and on instructional strategies and tactics. More recently, attention has shifted to some extent from an emphasis on the education process to a consideration of the product and the expected learning outcomes. In short it is now appropriate to ask ourselves the questions "What sort of doctor are we trying to train? Have the needs and expectations of the society in which they will be practicing been taken into consideration?" [3]. These questions become even more relevant in the field of forensic medicine where the extent, duration and pure existence of a forensic medicine training program in undergraduate medical curricula have now become controversial. In the Sri Lankan context, while some feel that Forensic Medicine should be a postgraduate subject others justify its existence in undergraduate curricular based on the fact that all medical officers[1], on graduation, are expected to perform medico-legal duties [4 - 7]. The changing face of medico-legal practice in Sri Lanka is evident by the increase in the number of board certified medico-legal specialists/consultant judicial medical

1 PART V of the Medical Ordinance (1988). Available at http://www.saitm.edu.lk/fac_of_medicine/MED_files/MO_1988.pdf. Accessed on 10/07/2012. ■

officers and by the introduction of short duration informal training programs for those medical officers who request such training. Therefore it may be assumed that the 'actual' medico-legal requirements of a non-specialist medical officer are diminishing. However the fact that Forensic Medicine is not a popular branch of medicine for specialisation, lack of updating of the circular of the ministry of health and the informal nature of the short duration training programs make it necessary to ensure adequate undergraduate training in Forensic Medicine. Furthermore, the fact that Forensic medicine is not merely the conduct of autopsies or the examination of clinical medico-legal cases and that it encompasses many other aspects at the inter-phase of medicine and law (eg., certification of death, documentation, maintaining records, ethical behaviour), justifies the inclusion of Forensic medicine as an undergraduate subject in the medical program in Sri Lanka.

The ill-defined 'medico-legal role' of medical officers, concerns expressed by interested parties that; there is a reluctance and lack of confidence among medical graduates to perform medico-legal duties, dissatisfaction among stakeholders about the performance of medical officers and concerns that too much time in the undergraduate curriculum is being used for Forensic medicine highlights the necessity to define these so called professional competencies that should be aquired at the end of undergraduate medical education. These should be defined even in embryonic form with provision for further development during the course of their careers [8].

Review of Sri Lankan literature revealed that a discrepancy does exist between private sector employer needs and graduate skills of those passing out from the state universities[9]. Tharmaseelan (2007) noted that universities have a danger of offering programs that are not relevant or do not match the needs or demands of the work world [10]. Subjects are introduced without prior consideration of future markets, review of needs or long term vision but merely because there are specialists in a subject area in the department who have an interest to promote their own favourite subject. However no studies have been published with reference to state sector graduate competencies and the state sector employer needs of Sri Lanka. It appears that there is a presumption that since the supply and demand are both related to the state sector that a needs analysis is irrelevant. The high expectations of the medico-legal system in Sri Lanka from a medical officer who has limited training in medico-legal work (purely undergraduate) lead to the hypothesis that a gap exists between stakeholder expectations and graduate competencies with regard to medico-legal work.

It is seen that Forensic Medicine training at undergraduate level is not based on a formal process of needs assessment. It is not linked to the key priorities of the the ministry of justice which in turn should be communicated to the ministry of health and the ministry of higher education. In short there is a fairly loose relationship between the service provider (ministry of health), service recipient and the pre-service training institutions (ministry of higher education) which is not used in any formal way for workforce training and development [11]. Unless this task is approached thoughtfully and systematically the curriculum would merely be a reflection of faculty interest rather than of stakeholder, student or public needs.

References

[1] Majumder, MAA. A review of the undergraduate medical curriculum in Bangladesh. Bangladesh Medical Journal, 2002, (31): 47-49.
[2] WHO. Educational handbook for health personel. (1998). Available at http://whqlibdoc.who.int/ publications / 1987/924170635X_eng.pdf. Accessed on 22/07/2012
[3] Harden, RM. Learning outcomes and instructional objectives: is there a difference? Medical Teacher, 2002, 24(2): 151-155.
[4] PART V of the Medical Ordinance (1988). Available at http://www.saitm.edu.lk / fac_of_medicine / MED_ files / MO_1988.pdf. Accessed on 10/07/2012
[5] Section 45 of the evidence ordinance of Sri Lanka
[6] The Manual of management of district Hospitals, peripheral units and rural hospitals. (1995). Ministry of health and social welfare
[7] The manual of management of teaching, provincial, base and special hospitals. (1995). Ministry of health and social welfare.
[8] Simpson, JG, Furnace, J, Crosby, J, Cumming, AD, Evans, PA, Friedman, M, Harden, R M, Lloyd, D, McKenzie, H, Mclachlan, JC, Mcphate, GF, Percyrobb, IW, Macpherson, SG. The Scottish doctor-learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners. Medical Teacher, 2002, 24(2): 136-143.
[9] Weligamage, S, Siengthai, S. (2003). Employer Needs and Graduate Skills: The Gap between Employer Expectations and Job Expectations of Sri Lankan University Graduates, 9th International conference on Sri Lanka Studies. Matara, Sri Lanka.
[10] Tharmaseelan, N. Tertiary education in Sri Lanka:Issues and challenges. Bulgarian Journal of science and education policy, 2007, 1(1): 173-188.
[11] Situation Analysis:Human Resources for Health: Strategic Plan (2009 -2018); Supplementary Document, Ministry of Healthcare And Nutrition, Sri Lanka, 2009. ■

2157118X-4.1.P1 Peer Reviewed
Received 20 April 2014
Received in revised form 3 September 2014
Accepted 2 October 2014
Available online 28 December 2014

Effect of Four Latent Blood Visualization Products on DNA

Céline Nicloux [a], *Jessica Bressler* [a]

a Institut de Recherche Criminelle de la Gendarmerie Nationale (IRCGN) Rosny-Sous-Bois, France.

Abstract Bloodstain pattern analysts first document visible stains and patterns at a scene. Then, in order to obtain as much information as possible from the scene, they search for latent bloodstains and patterns. Latent bloodstains may be the result of old scenes, scenes altered by weather or fire or clean up attempts to destroy evidence. Due to the existence of false positives, analysts must determine if those latent patterns are blood, human blood and then have DNA analysis performed.

Several chemical products are universally utilized to visualize latent bloodstains. Some have a direct and known negative effect on DNA such as o-tolidine, benzidine and TMB (tetramethylbenzidine) while some others can improve both visualization and DNA analysis after sampling. But how "DNA friendly" they are?

In this article, four different products used to visualize latent bloodstains will be compared. The products tested all have luminol-based formulations that contain varying amounts of hydrogen peroxide. Two of the reagents contain fluorescein.

Keywords: Forensic science, Blood pattern analysis, Latent blood, Visualization, DNA.

1 Introduction

The goal of this experiment is to compare the effect/no effect over time of Bluestar® Forensic, Bluestar® Forensic Magnum, Lumiscene and Lumiscene Ultra on DNA analysis. In order to simulate a crime scene in which blood was subjected to clean-up attempts or otherwise altered, blood dilutions from 1 to 1:2000 were tested. As all analysts use a variety of chemical methods and solution strength to visualize latent bloodstains, three different volumes of products (20, 60 and 100 µl) were added to each dilution. For a statistic goal, three samples were prepared for each volume and dilution. There is usually a delay between the collection of the blood sample and the actual DNA analysis. Therefore, DNA analyses were made at D0, D30 and latter at D60. At least 924 swabs prepared and analyzed.

- D0 = DNA analysis on first day with chemicals in contact with diluted blood
- D30 = DNA analysis at 30 days with chemicals in contact with diluted blood
- D60 = DNA analysis at 60 days with chemicals in contact with diluted blood

In order to complete the study, a second experiment was done to compare two types of immunochromatographic tests Hemoglobin (Hb) test and Rapid Stain Identification of Human Blood (RSID™) were used to confirm the presence of human blood.

The second test revealed human blood without any false positive results. The tests were performed at the Biology department of the Institut de Recherche Criminelle de la Gendarmerie Nationale (IRCGN) in Rosny-Sous-Bois France.

2 Materials and Methods

2.1 Blood Collection

Blood was drawn from a known donor into a collection tube containing EDTA. For all experiments, a total of eight dilutions were done. Dilutions

This research was first reported on The 2014 IABPA Annual Training Conference (organized by Herb Leighton, Brandi Caron and Alison Gingras of the Maine State Police), Portland, Maine September 30-October 3, 2014. ■

were the following: 1:1, 1:10, 1:20, 1:100, 1:200, 1:500, 1:1000 and 1:2000.

2.2 Chemical Reagents for Experiments 1 and 2

- Bluestar®
- Forensic, Bluestar® Forensic Magnum
- Lumiscene
- Lumiscene Ultra

2.3 Procedure

2.3.1 Experiment 1: DNA Sensitivity

- Each experiment was conducted at D0, D+30 and D+60 periods.
- The samples were prepared in individual Eppendorf® tubes. Three volumes of chemicals were tested (20, 60 and 100µL).
- A volume of 20µL of diluted blood was added to each tube.
- Three swabs were tested for each volume.
- For each period of the experiment, negative and positive controls were performed
- At least 924 swabs were prepared and analysed

The sequential steps are presented in figure 1.

2.3.2 Experiment 2: Human Blood Testing

- Each experiment was performed at D0 and D+60 periods
- For each dilution, one volume of 50µL of chemicals was tested and one volume of 20µL of blood was used.
- For each volume, three spots were put on a mat paper. Each paper was dried at room temperature for 0 and 60 day periods.
- For each period of experiment, some negative and positive controls were performed
- Each sample was prepared by placing the blood volume on

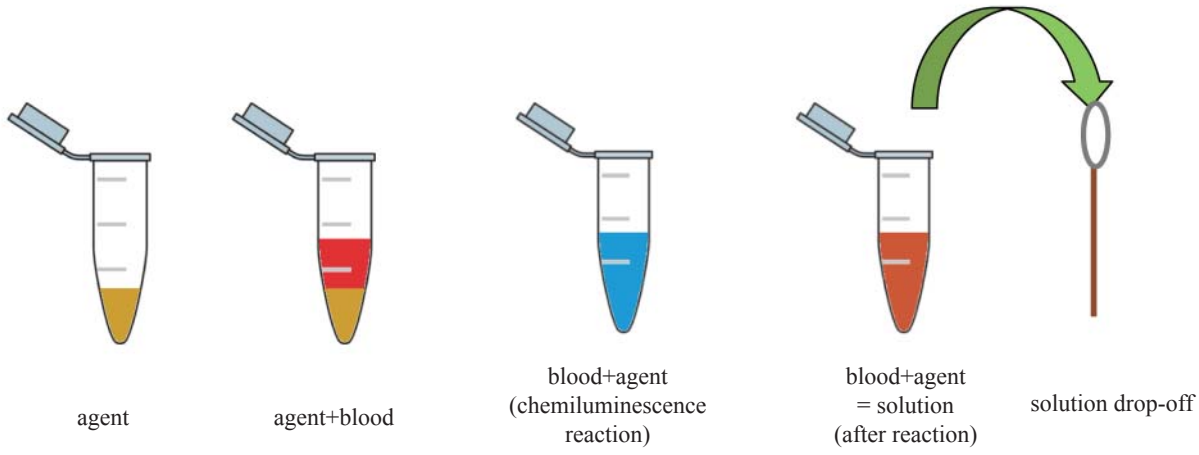


Fig 1. Diagram of sequential steps for experiment 1: DNA sensitivity.

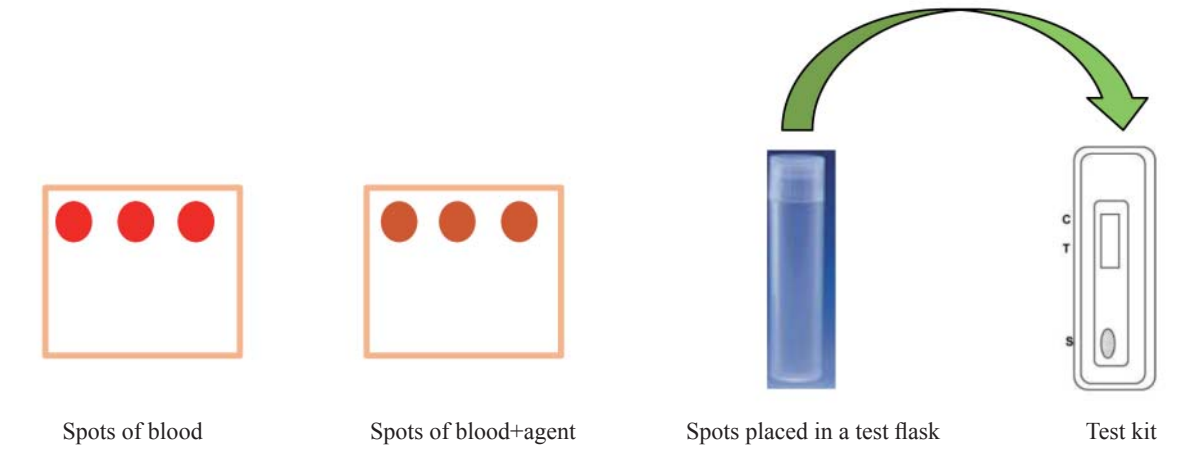


Fig 2. Sequential steps for experiment 2: Human blood testing.

a mat followed by the blood chemical volume.

- A total of 328 spots for both tests (2 × 164) were prepared and analyzed for all periods.
- Spots were cut and placed in a specific reagent (specific for the test studied) according to Hemoglobin (Hb) and RSID™ blood testing protocols.
- For each period of experiment, negative and positive controls were performed.

The sequential steps are presented in Figure 2.

3 Results

3.1 DNA Sensitivity

Each profile was analysed according to their quality based on a score as shown in Table 1.

3.1.1 Results at D0

Negative controls did not reveal any positive results and positive controls revealed positive results for all swab from undiluted blood to 1:500 dilution. The results for Lumiscene (L), Bluestar® Forensic (BS) Bluestar® Forensic Magnum (BM) and Lumiscene Ultra (LU) are presented in Tables 2-6.

As demonstrated in tables 2 and 3, Bluestar® Forensic produced better results according to the quality of the DNA profile at the 1:100 dilution than did Lumiscene Ultra. Results were quite similar for the remainder of the dilutions. Bluestar® Forensic did not affect the ability to obtain the known donor's DNA profile at 1:10, 1:20 and 1:100 dilutions of blood. Similar results were obtained with Bluestar® Forensic Magnum with a slight variation at the 1:100 dilutions.

One characteristic of the Lumiscene Ultra chemical is that it is sold only for reconstruction. That is why complete profiles were given only with the 1:10 dilutions. Good results were obtained with a majority of complete profiles for the 1:20 dilutions. Beginning with the 1:100 dilutions, the results showed a strong DNA alteration.

3.1.2 Results at D30

Negative controls did not reveal any positive results and positive controls revealed positive results for swabs from undiluted blood to a 1:500 dilution. The results for Lumiscene (L), Bluestar® Forensic (BS) Bluestar® Forensic Magnum (BM) and Lumiscene Ultra (LU) are presented in Table 6.

For Lumiscene, from undiluted blood to the 1:20 dilutions, DNA profiles qualities were the same. However, from 1:100 dilutions, light damage. Indeed, there were more scores of 2 than of 0 to the 1:200 dilutions. Scores of 3 appeared at the same dilutions than at D0.

The results for Bluestar® Forensic, as seen in experiments at D0, it is a satisfactory chemical to obtain a DNA profile at 1:10 and 1:20 dilutions. Nevertheless from the 1:100 to the 1:2000 dilutions, results were worse with a majority of scores of 2 for the 1:100 dilutions and a score of 3 for all tests from the 1:500 dilutions. The same trend can be noticed with Bluestar® Forensic Magnum with incomplete profiles obtained from the 1:200 dilutions.

Overall, profiles showed the same quality for Lumiscene Ultra at D0 and D30 with better results at D30 from the 1:100 dilutions with a majority of scores of 2 to the 1:200 dilutions.

3.2 Human Blood Identification Tests

3.2.1 Human Blood Testing

Results at D0

Each test was analysed based the on the scoring system as presented in table 7.

Hb tests

Negatives controls did not reveal any positive results. Positive controls revealed positive results for undiluted blood and for the following dilutions: 1:10, 1:20, and 1:100 and weaker for 1:500 and 1:5000. The results are presented in table 8.

As noted in table 8 positive results were obtained with undiluted and diluted blood (1:10 and 1:20). Results for the 1:100 dilution could be acceptable but they were better with Lumiscene. It was observed that for undiluted blood, results were somewhat worse than for the 1/10 dilution. This could be due to the "Hook Effect" which is a false negative result with certain immunoassays due to very high concentration of a particular reagent or blood.

RSID

Negative controls did not reveal any positive results. Positive controls revealed positive results for undiluted blood and for the 1:10 dilutions as presented in table 9.

Good results were obtained with RSID for undiluted blood and for the 1:10 dilutions as shown in table 9. The difference between these tests and Hb tests may be due to the sought element (hemoglobin for Hb tests and glycophorin A for RSID tests). Detection of glycophorin A seems to be more specific. This can explain why it could be more difficult to obtain a good result with a diluted blood than with haemoglobin detection.

Table 2. Results for Lumiscene (L)

			Lumiscene							
			Blood dilutions							
			Pure	1/10	1/20	1/100	1/200	1/500	1/1000	1/2000
Chemical volumes (µL)	20	Test 1	0	0	0	0	2	3	2	3
		Test 2	0	0	0	0	1	2	3	3
		Test 3	0	0	0	0	1	2	2	3
	60	Test 1	0	0	0	0	2	3	3	3
		Test 2	0	0	0	0	0	2	3	3
		Test 3	0	0	0	0	2	2	3	3
	100	Test 1	0	0	0	0	2	2	3	3
		Test 2	0	0	0	2	2	2	3	3
		Test 3	0	0	0	0	2	3	3	3

Table 3. Results for Bluestar® Forensic (BS)

			Bluestar							
			Blood dilutions							
			Pure	1/10	1/20	1/100	1/200	1/500	1/1000	1/2000
Chemical volumes (µL)	20	Test 1	0	0	0	1	2	2	2	3
		Test 2	0	0	0	2	1	2	2	3
		Test 3	0	0	0	0	2	2	2	3
	60	Test 1	0	0	0	0	2	2	3	3
		Test 2	0	0	0	0	0	3	3	3
		Test 3	0	0	0	0	1	2	3	3
	100	Test 1	0	0	0	0	2	3	3	3
		Test 2	0	0	0	0	2	2	3	3
		Test 3	0	0	0	0	2	2	2	3

Table 4. Results for Bluestar® Forensic Magnum (BM)

			Bluestar Magnum							
			Blood dilutions							
			Pure	1/10	1/20	1/100	1/200	1/500	1/1000	1/2000
Chemical volumes (µL)	20	Test 1	0	0	0	0	2	2	3	3
		Test 2	0	0	0	0	2	2	3	3
		Test 3	0	0	0	0	2	2	3	3
	60	Test 1	0	0	0	0	2	3	3	3
		Test 2	0	0	0	0	1	3	3	3
		Test 3	0	0	0	1	2	3	3	3
	100	Test 1	0	0	0	2	2	3	3	3
		Test 2	0	0	0	2	2	3	3	3
		Test 3	0	0	0	2	2	3	3	3

Table 5. Results for Lumiscene Ultra (LU)

			Lumiscene Ultra							
			Blood dilutions							
			Pure	1/10	1/20	1/100	1/200	1/500	1/1000	1/2000
Chemical volumes (µL)	20	Test 1	0	0	0	2	2	2	3	3
		Test 2	0	0	0	1	2	3	3	3
		Test 3	0	0	0	2	2	3	3	3
	60	Test 1	0	0	0	2	2	3	3	3
		Test 2	0	0	0	2	3	3	3	3
		Test 3	0	0	1	2	3	3	3	3
	100	Test 1	0	0	0	3	3	3	3	3
		Test 2	0	0	0	3	3	3	3	3
		Test 3	0	0	0	3	3	3	3	3

A unique test with the 1:20 blood dilutions was performed for each chemical. Indeed, only the 1:10 and the 1:100 were tested in the

Table 1. Scoring system for each DNA profile

0	Complete profile (0 or 1 locus affected)
1	Profile with 1 anomaly: Partial or unbalanced (from 2 loci affected)
2	Profile with 2 anomalies: Partial or unbalanced (from 2 loci affected)
3	Non explicable profile

first phase. The results for the 1:100 dilutions were significantly lower or absent. That is why it became interesting to test the 1:20 blood dilutions.

4 Discussion

Lumiscene provided better results for the 1:100 dilutions than the other chemicals. Moreover, most of the results obtained for D0 and D30 show that the quality of genetic profiles are more similar with Lumiscene than with Bluestar®. Genetics profiles are the same for light dilutions for all chemicals.

Additional results for D60 will be obtained soon. Nevertheless, variations according to scores assigned to each profile were noted for D0 and D30. The effect of time was observed. The only variable parameter was the time of contact between blood and chemicals. However, chemicals did not seem to have an effect on time for undiluted blood. Hence, this trend that was observed during one month indicated that the quantity of those products has an effect on the quality of the genetic profiles of diluted blood. This may be due to the period during which chemicals were in contact with blood and/or the quantity of chemicals placed on each swab.

In this experiment, the exact number of leukocytes was unknown. It was not possible to count precisely the number of cells with a specific counter. Blood pipetting was done to allow a homogeneous distribution of cells for each dilution, hence the reason that three samples were tested for each dilution and for each volume of chemicals added.

Only 20 µL of blood dilutions were used on each swab. This parameter may explain some poor results concerning the quality of genetic profiles. We suggest that a larger quantity of blood would provide better genetic profiles.

It is important to underline the fact that no surfaces were tested in this experiment. Bloodstains and patterns are influenced by surface texture and composition upon which they were deposited or come into contact. Indeed, this can have an effect on the absorption, the size and shape of the bloodstains and may also influence the quality of samples.

5 Conclusions

Results obtained for D0 and D30 seem indicate that Lumiscene is slightly better than Bluestar® Forensic especially at the 1:100 dilutions at the 30 day period. Lumiscene Ultra was tested with the understanding that it is used for reconstruction and not for DNA analysis. Accordingly, poorer results were obtained with a period of one month during which chemicals were in contact with blood. This result appears to be linked with the quantity of chemicals put on swab and the time of contact.

At D0, the quantity of chemicals did not significantly affect the quality of the profile. The tests also revealed that an increase in time and chemical volume more greatly affected the ability to obtain a DNA profile. This trend will be analyzed and perhaps certified with D60 results. With all results, an analysis will be done on 924 data with a statistical study. We hope to publish a scientific article in the near future with all the results and more details will be presented.

With regard to human blood detection, without results for D60, RSID provided worse results than the Hb test despite its specificity. This trend has to be confirmed with experiments at D60 which will be performed in the near future.

References

[1] Gnyaneshwari Patel, Andy Hopwood. An evaluation of luminol formulations and their effect on DNA profiling. Int. J. Legal Med. 2013; 127: 723-729.

[2] BaseClear B.V., Einsteinweg 5, 2333 CC Leiden, The Netherlands. Forensic DNA analysis (STR-testing) of human blood treated with Lumiscene. BASECLEAR. 22 August 2012. 52613 report.

[3] Juliana Piva de Almeida, Nadine Glesse, Cristina Bonorino. Effect of presumptive tests reagents on human blood confirmatory tests and DNA analysis using real time polymerase chain reaction. Forensic Science International. 2011; 206: 58-61.

[4] Filippo Barni, Simon W. Lewis, Andrea Berti, Gordon M. Miskelly, Giampietro Lago. Forensic application of the luminol reaction as a presumptive test for latent blood detection. Talanta. 2007. 72: 896-913.

[5] Cathy J. Jakovich. STR Analysis Following Latent Blood Detection by Luminol, Fluorescein, and Bluestar. Journal of Forensic Identification. 2007. 57(2)193-198.

[6] Rapid Stain Identification of Human Blood (RSID) Technical Information and Protocol Sheet, European Distributor: Galantos Genetics, GmbH, 55128, Mainz, Germany.

[7] Hemoglobin (Hb) Test Technical Information and Protocol Sheet, Veda Lab, Alencon, France. ■

Received 3 June 2014
Received in revised form 30 August 2014
Accepted 11 October 2014
Available online 28 December 2014
Peer Reviewed
2157118X.4.1.P2

The Reliability of Current Methods of Sequencing Bloodstain

Terry L Laber [a], ***Michael C Taylor*** [*, b], ***Paul E Kish*** [c]

a Minnesota Bureau of Criminal Apprehension (BCA), 1430 Maryland Ave East, St Paul, MN 55106.
b Institute of Environmental Science and Research Limited (ESR), P O Box 29-181, Christchurch, New Zealand.
c Paul Erwin Kish Forensic Consultant & Associates, P.O. Box 814, Corning, NY 14830.

* Corresponding Author: Michael C Taylor. E-mail: michael.taylor@esr.cri.nz

Abstract Despite the potential value associated with determining the sequence of events from superimposed bloodstain patterns, no formal assessment of the reliability of current methods was found in the published literature. We present here a study of superimposed spatter/transfer patterns on three different substrate surfaces under conditions where the spatter pattern component consisted of a small, medium and large number of stains. This test was done in the absence of perimeter stain effects.

Keywords: Forensic science, Bloodstain pattern, Spatter, transfer, Methods, Classification, Sequence, Error rate.

1 Introduction

It is common for bloodstain patterns at crime scenes to be superimposed [1]. The order in which such patterns are deposited can sometimes be valuable evidence of the timing of the events that took place[1]. The observation, for example, that a bloodied shoeprint impression has spatter stains from a beaten victim on top of the impression indicates that those stains occurred after the shoeprint impression was made. This information would be highly probative if, the defendant claimed he arrived at the scene after the victim had been beaten. Despite the value of this type of evidence, there have been few published studies made of bloodstain pattern sequencing and no standardized methods have emerged.

Determining the sequence of events from bloodstain patterns frequently involves altered patterns. For example, the presence of perimeter stains [2] in a pattern is evidence that more than one event took place with a lapse of time between events [1]. If blood is initially dripped onto a surface and subsequently wiped prior to complete drying, a perimeter ring of staining remains, providing evidence of the sequence of events. If one pattern dries before the second is superimposed, however, sequencing becomes more difficult.

Hurley and Pex [2] concluded that it was difficult to distinguish a dried spatter pattern overlaid by a bloodied shoe impression from a combination of patterns in the reverse sequence. They recommended particular caution when attempting to determine such a sequence from photographs. While Hurley and Pex produced photographs to illustrate their conclusions no controlled experiments were presented. The objective of this study was to formally assess the reliability of current methods for establishing the sequence of superimposed patterns where the first pattern deposited has completely dried.

2 Methods

Two pattern types were selected for this study, namely spatter [3] and transfer [4]. These patterns are commonly encountered at bloodied crime scenes and can often be superimposed. A total of 112 bloodstain patterns comprising superimposed transfer and spatter

1 In the context of this paper a superimposed pattern is the deposition of that pattern on top of an existing pattern.
2 A perimeter stain is an altered stain that consists of the peripheral characteristics of the original stain.
3 A spatter stain is a bloodstain resulting from a blood drop dispersed through the air due to an external force applied to a source of liquid blood.
4 A transfer stain is a bloodstain resulting from contact between a blood-bearing surface and another surface. ■

stains were prepared, half of which were spatter stains superimposed on transfer stains and half were transfer stains on spatter stains. Fresh human blood, donated by project volunteers, and containing EDTA anticoagulant was used within seven days of drawing. Patterns were created in a controlled laboratory setting at the Minnesota BCA Laboratory. They were created on 16 inch x 16 inch (40 cm x 40 cm) wooden targets. Completed targets were coated with a clear lacquer to prevent deterioration and to assist with biohazard safety. It has been assumed that the clear lacquer coating had no significant effect on the analyst's conclusions.

Transfer stains were created by drawing a blood-soaked cotton glove across the target surface, producing a swipe [5] pattern showing four fingers. Excess blood was removed from the glove before swiping. Spatter stains were created by using a hammer to strike one drop of blood placed on a wooden block in the center of the striking zone. The hammer was propelled by rubber bands and gravity. Bloodstains forming the first applied pattern were allowed to dry thoroughly before the second pattern was superimposed. There were two manipulated variables relating to pattern construction, namely pattern extent (amount of spatter) and target substrate. There were three levels of pattern extent; minimum, medium, and maximum. Category membership was determined by an approximation of the total number of stains in the pattern and the number of stains larger than 1 mm in diameter (Table 1).

Table 1. Scoring system for each DNA profile

Extent	Total Number of Stains	Number of Stains > 1 mm
Minimum	< 50	< 10
Medium	< 50	> 10
Medium	50 – 500	< 50
Maximum	50 – 500	> 50
Maximum	> 500	no criterion

The second controlled variable was the substrate that the pattern was created on. The substrates used were three different hard surfaces, representing varying levels of anticipated identification difficulty, namely: paint (A), wallpaper (B), and chipboard (C). Two coats of white Zinsser 1-2-3 primer were applied to a smooth wooden surface for the paint surface. The wallpaper surface was white Brewster Easy Texture paintable wallpaper (STRIA Pattern 99417F) glued to a smooth wooden surface with one coat of Zinsser 1-2-3 primer applied on top of the wallpaper. The target was rotated prior to pattern creation so that the wallpaper texture ran vertically. The chipboard surface was made from oriented strand board (OSB), which comprises wood fragments bonded in a resin and oriented in random directions. Examples of these target surfaces are shown in Figures 1a and 1b.

Participants chosen for this study were 27 experienced bloodstain pattern analysts with at least 80 hours of training in bloodstain pattern analysis, a minimum of 5 years of BPA experience and had been qualified in court as BPA experts. Analysts were individually invited to participate and were informed that the aim of the study was not to test analyst competency, but rather the reliability of current BPA methodology. They were also informed that their participation and responses would remain anonymous. Materials were only sent to analysts after they had indicated a willingness to participate.

Each analyst received a unique set of 3 or 4 sequencing targets and a response sheet for each target. A number was placed at the top of each target that identified which target corresponded to each response sheet and also indicated the pattern alignment during pattern construction. The response sheet stated: "This sample has both a transfer and a spatter pattern on it. You are asked to determine the sequence in which these two patterns have been applied. Please choose ONE of the following:

- Spatter first followed by transfer
- Transfer first followed by spatter
- I cannot determine which pattern occurred first

After completing the survey, analysts emailed or posted their responses to an independent third party. All materials and responses were returned to experimenters via the third party to ensure they remained anonymous. Each possible variable combination (sequence order X substrate X extent) was replicated a minimum of 5 times.

3 Results

Of the 112 samples distributed to participants, responses to 104 were received. These comprised 50 combinations of spatter stains superimposed on transfer stains and 54 instances of transfer stains on

5 A swipe pattern is a bloodstain pattern resulting from the transfer of blood from a blood bearing surface onto another surface, with characteristics that indicate relative motion between the two surfaces. ■

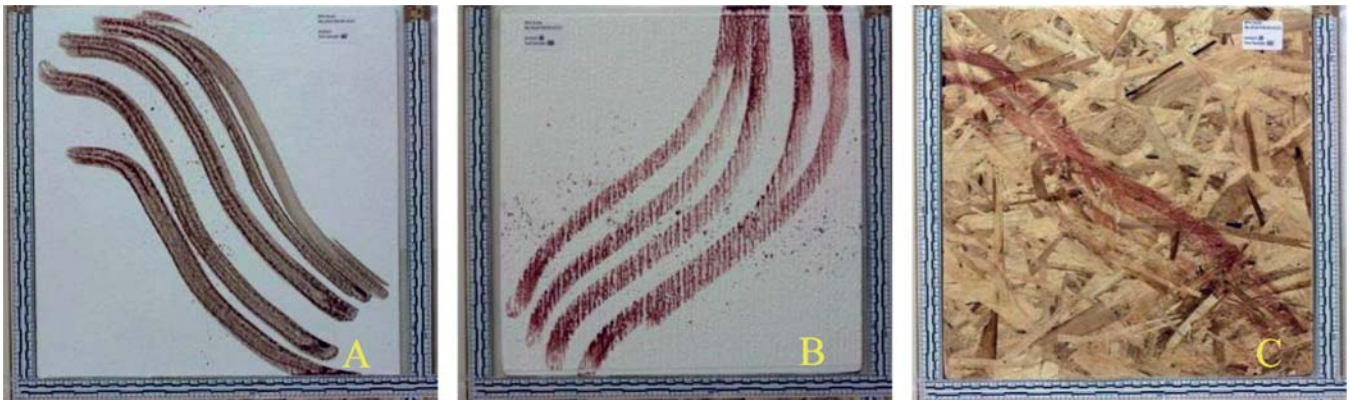


Fig 1a. Example of patterns created on paint (A), wallpaper (B), and chipboard (C) surfaces.

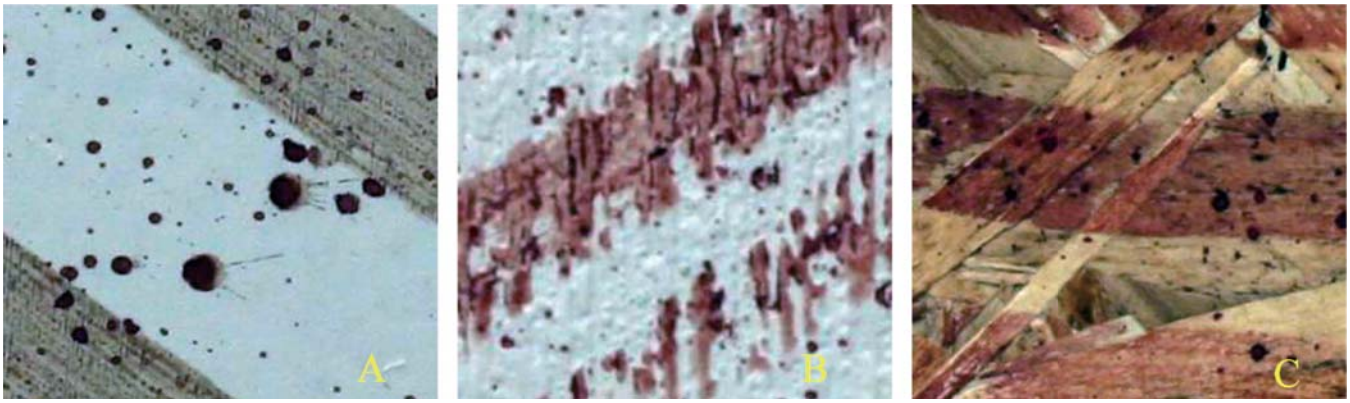


Fig 1b. Close-up images of superimposed patterns on paint (A), wallpaper (B), and chipboard (C) surfaces.

spatter. Of the 104 conclusions given, over half (52.9%) were recorded as inconclusive meaning they could not determine which pattern occurred first, 32.7% correctly assigned the sequence and 14.4% gave an incorrect interpretation.

Figure 2 shows a breakdown according to the original pattern combination presented. It is apparent from these results that there was a marked difference in the response of analysts to the two pattern sequences they were presented with. Where spatter stains were deposited on top of transfer stains, 48% of the patterns were correctly sequenced, whereas for the reverse sequence this figure dropped to 19%. There was a corresponding increase in the proportion of inconclusive responses from 40% to 65%.

These results appear to show that when spatter stains are deposited on transfer stains, analysts were more willing to give an interpretation and those interpretations are more likely to be correct. For those targets that analysts were prepared to make an interpretation, 80% were correct when the pattern was spatter on transfer, but only 53% were correct if the transfer followed the spatter. The difference in response between the two pattern combinations was statistically significant, X^2 (2, N=104)=5.26, p=0.004. Overall, the effect of substrate on correct responses was not significant, although Figure 3 shows a higher proportion of correct

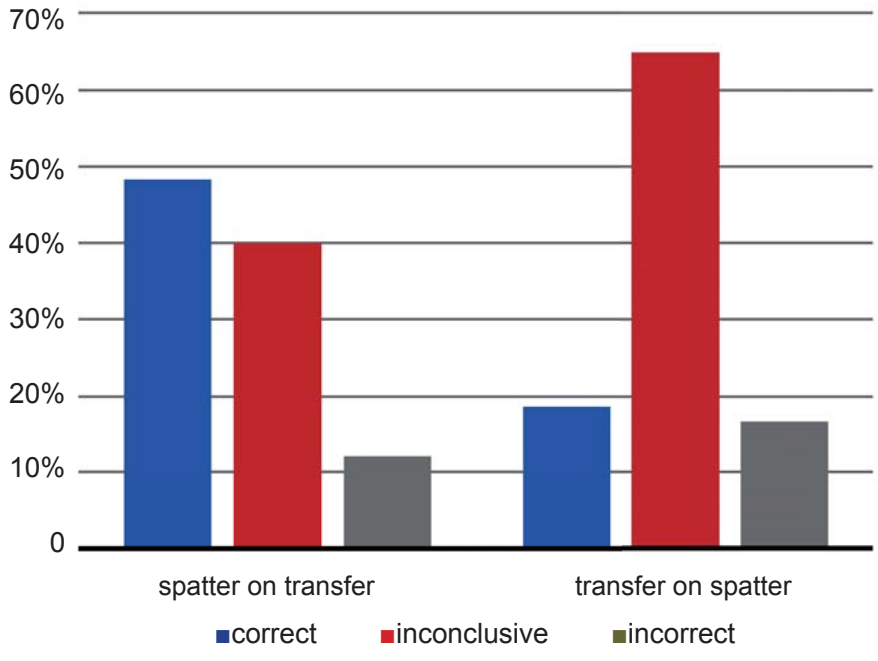


Fig 2. Accuracy in the determination of pattern deposition sequence.

interpretations for painted surfaces when the pattern combination was spatter on transfer.

Figure 4 shows that the number of incorrect interpretations increased and the number of inconclusive responses decreased slightly, as the extent of spatter increased in both spatter on transfer and transfer on spatter combinations. Those targets that had a spatter pattern with many stains (i.e., maximum extent) overlaid with a transfer pattern gave the highest number of incorrect interpretations. No errors were made, when the spatter pattern was at its minimum extent although 70% of the responses were inconclusive. However, the overall effect of pattern extent was significant, $X^2(4, N=104)=9.71$, $p=0.046$.

4 Discussion

Because the bloodstains in this study were allowed to dry completely between the two depositions, there were no perimeter stain effects to give clues as to the order of deposition. In the absence of this, it is possible that the analysts' attention was drawn to the intensity of the individual stains, with the more intense stains reckoned to be the more recent of the two depositions. Spatter stains deposited on transfer stains will generally

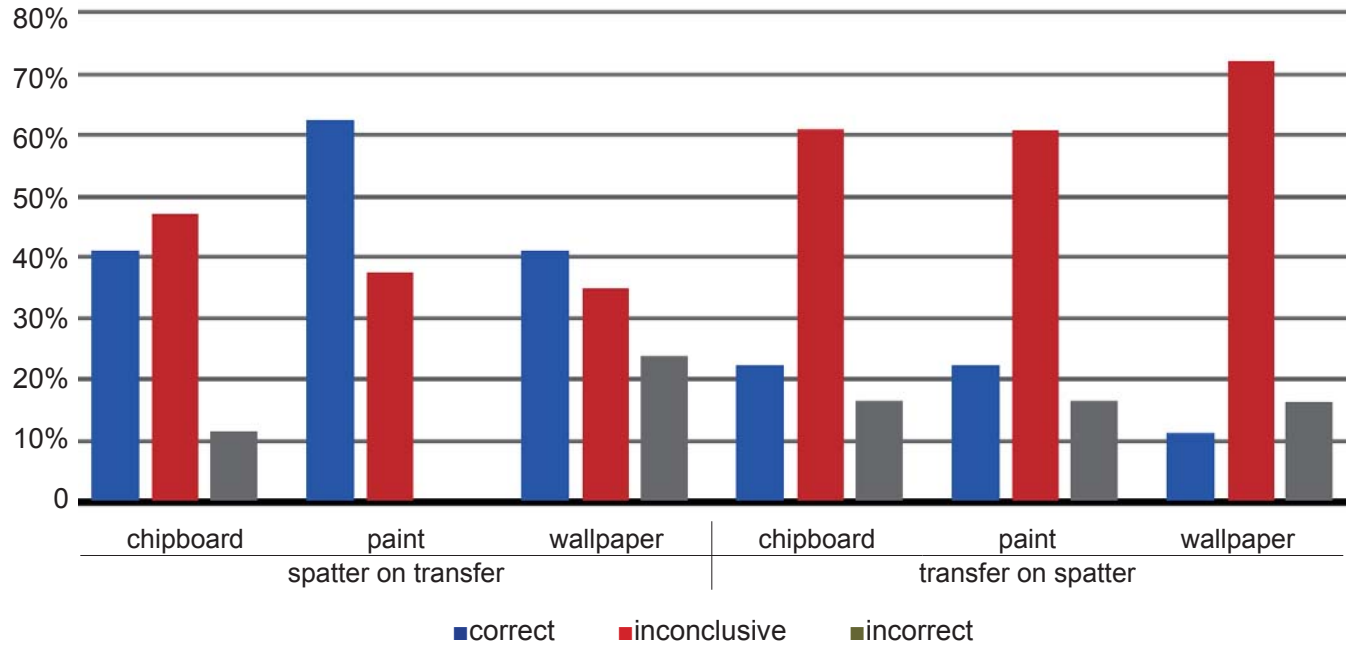


Fig 3. Effect of substrate on sequencing conclusions.

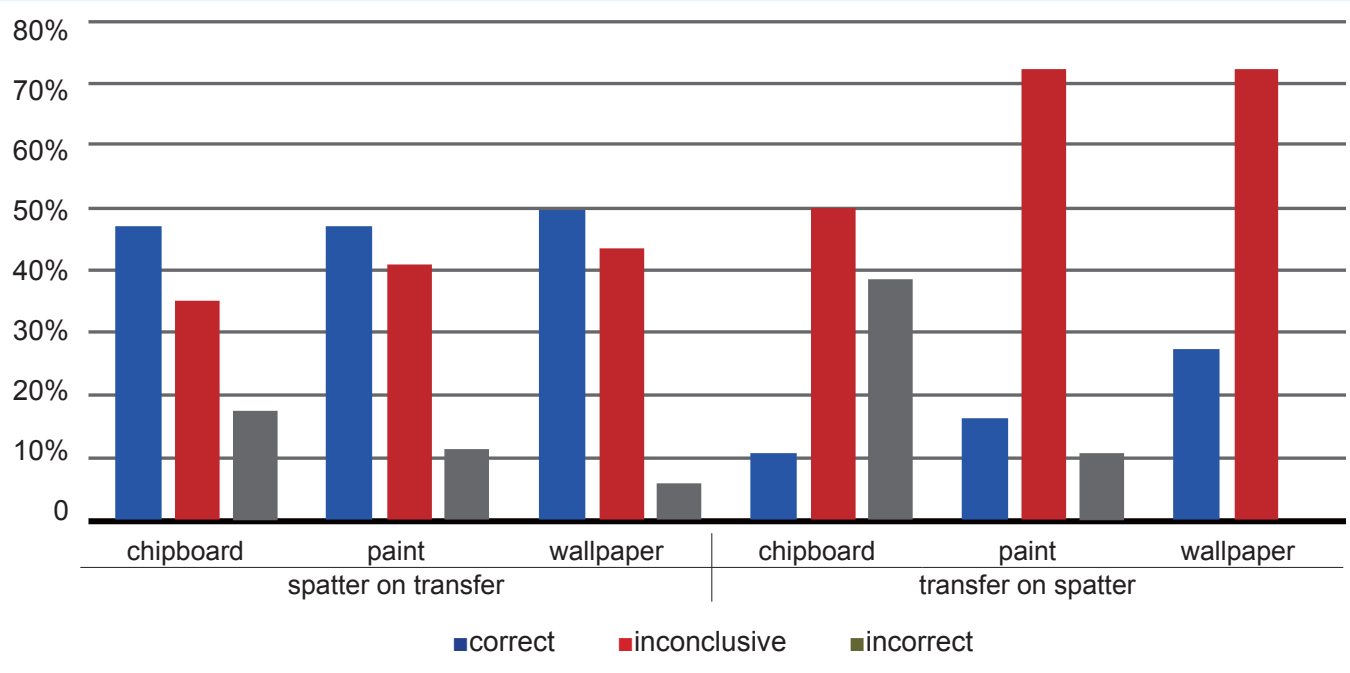


Fig 4. Effect of pattern extent on sequencing conclusions.

appear darker in color, suggesting they have been deposited last (e.g. Figure 5). However spatter stains deposited under transfer stains may also appear darker in color, especially if the transfer stain is a thin smear (e.g. Figure 6).

Under these circumstances the spatter stains may also give the impression they have been deposited on top of the transfer stain. This may also account for the fact that patterns with welldefined, maximum extent spatter overlaid by transfer stains had the highest proportion of incorrect conclusions. Where the spatter stains were fewer in number these incorrect interpretations were not evident. The number of incorrect conclusions increased as the extent of spatter increased in both spatter on transfer and transfer on spatter, while the number of inconclusive responses decreased slightly. This suggests that an increase in the number of datum points in the pattern is giving an increasingly false sense of confidence for an analyst when making a judgment about sequencing.

5 Conclusions

This study on superimposed patterns showed that, for the current sequencing methods and in the absence of perimeter stain effects, the chances of incorrectly concluding the order of deposition in a spatter/transfer pattern combination is approximately 12% where spatter stains are deposited on top of transfer stains and 17% for the reverse sequence. These results demonstrate the need for more reliable methods for bloodstain pattern sequencing and suggest that extreme caution should be exercised in making such determinations.

In general, analysts were reluctant to draw a firm conclusion in approximately half the samples and incorrectly concluded the sequence of patterns in approximately 12% of pattern combinations where spatter stains were deposited on top of transfer stains and 17% for the reverse sequence. The pattern substrate was not found to be a significant factor in the accuracy of sequencing. There was limited evidence to suggest that a more extensively spattered pattern, in combination with a transfer pattern may increase the incidence of misinterpretation of sequence.

Acknowledgements

The authors gratefully acknowledge the support of the National Institute of Justice through project grant 2010-DN-BX-K213.

References

[1] Kish, P.E., T.P. Sutton, and S.H. James, Principles of Bloodstain Pattern Analysis: Theory and Practice 3ed. Boca Raton, 2005: CRC Press: 576.
[2] Hurley, N.M. and J.O. Pex, Sequencing of Bloody Shoe Impressions by Blood Spatter and Blood Droplet Drying Times. International Association of Bloodstain Pattern Analysts News, 1990, 6(4): 1-8. ■

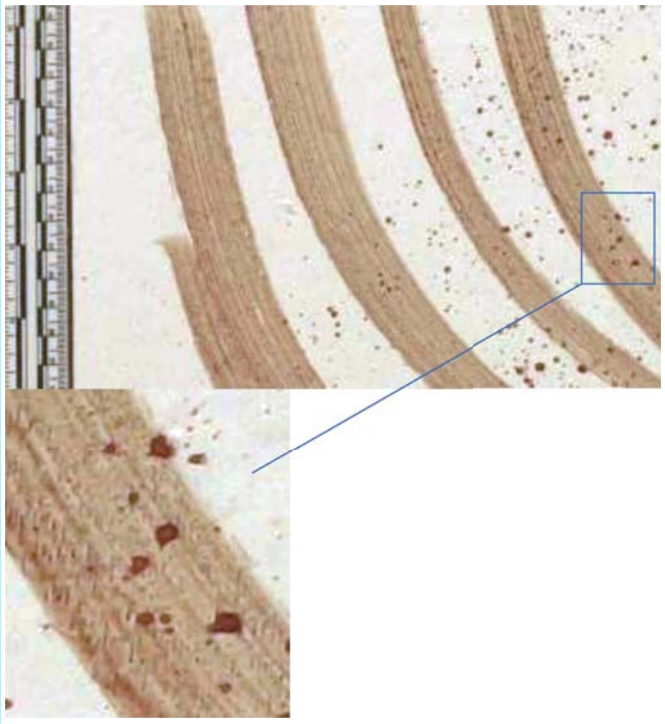


Fig 5. Spatter stains deposited onto a transfer pattern.

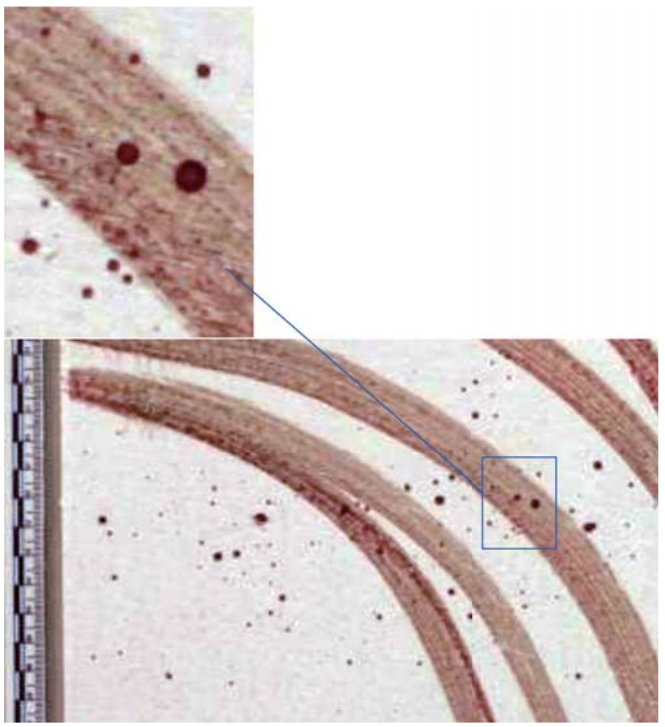


Fig 6. Transfer pattern deposited onto spatter stains.

2157118X,4,1,P3 Peer Reviewed
Received 29 September 2014
Received in revised form 21 October 2014
Accepted 27 October 2014
Available online 10 December 2014

Amido Black 10B in the Forensic Investigation - Comparative Studies of Forensic Stain Material

P J Laberke [a], *S Ilg* [a], *H-P Bieri* [b], *R Hausmann* [a], *B Balitzki* [*, a]

a Institute of Legal Medicine, University of Basel, Switzerland.
b Forensic Technical Department of the Police Basel-Stadt, Basel, Switzerland.

* Corresponding Author: Dr. P. J. Laberke. Institute of Forensic Medicine, University of Basel, Pestalozzistr. 22, 4056 Basel, Switzerland. E-mail: patrick.laberke@bs.ch

Abstract Bloodstains are among the most meaningful traces in forensic case analysis, which can offer extensive information in forensic case analysis. On the one hand evaluation of the structure, amount, shape and distribution at a crime scene may provide evidence for a reconstruction of events, on the other hand it is now possible to generate a complete DNA profile out of smallest amounts of blood, which can also be used to reconstruct an event and allow conclusions on participation of persons in a crime by molecular genetic analysis. In this study whole blood samples were applied on different substrates and then analyzed for blood using Hexagon OBTI®. Incubation with two different mixtures of Amido Black 10B showed strong adverse effects and false negative results with the OBTI. In contrast no influence of Amido Black 10B was seen in the subsequent DNA analysis.

Keywords: Forensic science, Blood, Indicators and reagents, Sensitivity and specificity, Reliability of results, DNA.

In recent years, a new method has been established in forensics for the visualization and contrast of blood traces, based on the reaction of the blood with Amido Black (Amidoschwarz, AS) 10B. The immunological test principle of Hexagon OBTI® based on the detection of a plasma protein that is also stained by AS 10B. The influence of this staining of blood proteins on the result of the immunological detection of blood and subsequent molecular genetic DNA typing has not been considered.

1 Basics

Traces of blood can be saved in diverse designs at crime scenes [9, 10]. From large pools of blood to micro traces of blood they find themselves in response to the offense at crime scenes in their environment or accomplices to track sources. Now there is a plethora of methods which are available [1, 2, 3, 4, 5, 6, 7, 11, 13, 14, 15] for the visualization and contrast, as well as for the qualitative detection of traces of blood. They are usually used directly at the crime scene and especially for the smallest amounts of blood or visually bad undiscovered tracks, especially in cases where the scene of conspicuous traces of blood already be cleaned .

It can be minimized the amount of blood visible. For daktylo-scopic investigations, the gain in visual perception of tracks may be necessary to document them photographically can. Many methods use the Pseudoper-oxidase reaction of hemoglobin and its derivatives, [2, 11, 15]. In part, this chemiluminescence production increases the visual perception of the trail of blood [14]. The visualization of blood traces by excitation with different light sources is used at crime scenes [13].

After visualizing a trail of blood and its fingerprint analysis, it follows by the actual blood detection using commercial, immunological detection methods, which are characterized at high specificity and sensitivity as well as quick and easy to use[1,3]. The positive detection of blood leads mostly for molecular genetic analysis of the track in order to make a personal assignment.

I has established the individual elements of the blood trail analysis of the visualization, documentation, the blood evidence to DNA typing in forensic trace analysis [2, 11]. Less known are the influences of the methods used in each subsequent analysis. For luminol reagent often-used for the visualization of blood, impairment of subsequent blood detection with the Hexagon OBTI® could be detected [4] as a function of blood volume and application time.

2 Methods

Amido black 10B

Amido black 10B, also known as naphthol blue black B, is an azo dye, which was used as one of the first dyes in gel electrophoresis for visualization of proteins. It has been increasingly used [1] in the fingerprint traces of dye since the beginning of the 90s.

"BDA" -recipe the AS-10B dyeing and rinsing solution

0.4 g AS 10B are dissolved in 180 ml of methanol and 20 ml of glacial acetic acid staining solution for the "BKA" -Recipe. After spraying the dye solution on the track first supports the excess ink with a methanol-glacial acetic acid solution (9: 1), followed by a solution of distilled water and glacial acetic acid (19: 1) and finally rinsed with distilled water.

"The FBI" -recipe the AS 10B dyeing and rinsing solution

For the staining solution according to the "FBI" -Recipe 8.0 g sulfosalicylic acid, 1.2 g and 1.2 g of sodium carbonate AS 10B are successively dissolved in 200 ml of distilled water and then with 20 ml of formic acid, 20 ml of glacial acetic acid and 5 ml Kodak Photo Flo 600® added. The staining solution is sprayed onto the track carrier and then rinsed off the excess paint with distilled water.

Immunological detection of blood

Principle of Hexagon OBTI®

The Occult Blood Test Immunological (OBTI) is a clinical rapid test for the detection of occult blood in stool samples using an immunochromatographic test procedure. According to the manufacturer is the detection limit, based on the transport medium, at 0.1 µg / ml hemoglobin (Hb) and the maximum Hb concentration at 2 mg / ml. The test results can be read after 2-3 min. In the case of a negative result must be a second reading after 10 minutes.

Determination of the optimal blood concentration

To exclude effects of blood concentration of the test result, 0.1, 1, 10 and 100 µl of the whole blood samples were pipetted directly and shaken in 2000 µl OBTI-test-buffer. After 5 s 2 TrpF., placed in the sample window of OBTI Test Cassette. The time from dropping into the sample window to safe reading of a positive test result was documented.

Blood Testing in the presence of amido black 10B

Each 1 µl of the male blood was pipetted on 30 white cotton swabs. Each of these 10 samples were treated with the "BKA" -Recipe and the "FBI" -Recipe; the remaining 10 lobules were left untreated. Samples from each batch were then placed in the OBTI buffer. After 5 s 2 TrpF., added dropwise to the mixture in the sample window until the appearance of the positive result.

In a second step, each 1 ul of the whole blood was pipetted into a total of 30 reaction vessels made of plastic. For each of these 10 samples, the components of the AS 10B staining methods "BKA" and "FBI" were successively added and removed. The remaining 10 samples in the reaction vessels were left untreated. After addition of the OBTI-Buffers in the reaction vessels and 5-second incubation, 2 TrpF. were added to the test cassette.

DNA testing

Extraction

Analogous to the sample preparation for the proof of blood in the presence of AS 10B 60 samples were created and performed DNA extraction with the Investigator® kit (Qiagen). To the manufacturer's protocol was used for isolation of DNA from bodily fluids. The analysis period was 1 h.

Quantification

The PLEXOR®-technology is a "real-time polymerase chain reaction" - (Real-time PCR "-) test procedureallows [8] the quantification of total male DNA at the same time. The DNA-content of the samples was measured with the RotorGene®-system. For this purpose, each 2 µl of each sample were indicated used with the PLEXOR®-HY-system according to the manufacturer's protocol in duplicate and the mean. It was also performed a melting curve determination for the products of autosomal, gonosomal and internal PCR-Control (IPC). The analysis of the raw data was performed using the PLEXOR® HY Analysis Software and forensic analysis settings.

DNA profile

For each series of tests of different pretreatments and carriers of a median of DNA-Concentration sample was exemplary, from a DNA-Profile in 11 STR systems and the gender Amelogeninsystem was created.

3 Results

Determination of the optimal blood concentration

In determining the optimal blood concentration for the OBTI using the investigation of a serial dilution of whole blood from a healthy male subjects the control line appeared immediately in all activities undertaken OBTI trials after contact with the test medium and thus indicated a regular course of the test.

After 30 min, the lowest dilution of 1 : 20 could also be detected. The test was therefore seen as negative. At dilutions of 1: 200 and 1: 20.000 after 13 resp. 18 s only a weak test line was visible; this resulted in the assignment to "weakly positive" result. At a dilution of 1: 2000, a significantly positive test result was already evident after 12 s.

Blood Testing using the OBTI

The comparison of the results in OBTI the untreated and the AS 10B ("BKA" or "FBI") stained blood samples shows in Fig. 1. The results were futher classeified based on their underlying Träger material cotton) Trägermaterialien Baumwolle, BW) or Reaction-vessel (Reaktionsgefäß, RG). Because of the support materials used cotton. A total of 60 blood samples were investigated. For the 20 untreated blood samples for BW and RG, clearly positive test results could be read within 10-22.

After treatment with AS 10B "BKA", it could only have noted 9 BW weak positive and a negative test results; this also had to be considered as negative due to exceeding the maximum test time of 5 weakly positive samples.

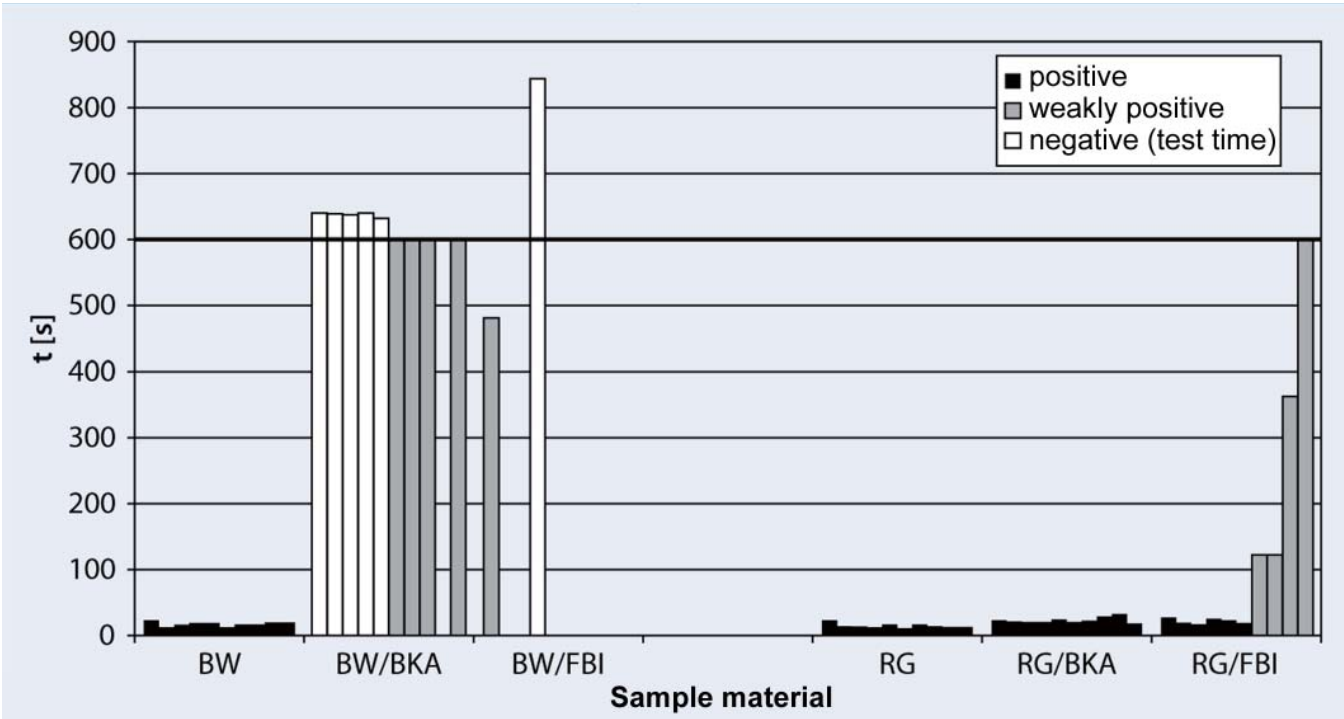


Fig 1. Blood Testing with the OBTI. BKA "BKA" -Recipe for amido black 10B, BW cotton, FBI "FBI" -Recipe for Amido 10B, RG reaction vessel.

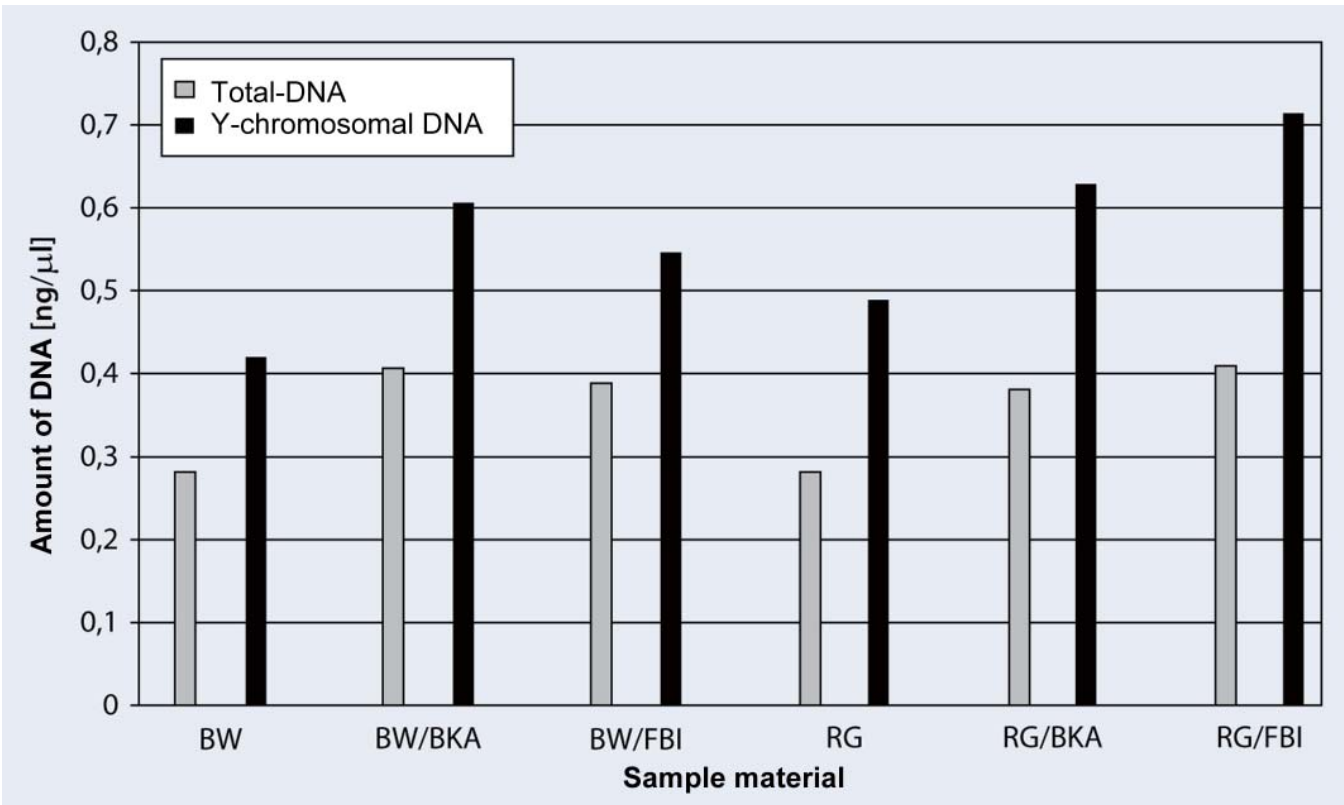


Fig 1. Blood Testing with the OBTI. BKA "BKA" -Recipe for amido black 10B, BW cotton, FBI "FBI" -Recipe for Amido 10B, RG reaction vessel.

These samples are in Fig. 1 shows in white. The 10 samples in the RG gave, after treatment with AS 10B "BKA" with an average response time of 23 s consistently positive results.

Of the 10 blood samples for BW, who had been following the "FBI" -Recipe with AS 10B be-, it is showed only a sample after 8 min of a positive result in OBTI. The weak test line to another sample could be read after 14 min. Because of this reaction time in the test, the negative results has been assigned as all other samples of this series of tests. The blood samples RG from this series of tests were positive with 6 and 4 weakly positive results as well as response times between 18 s and up to 10 minutes of a rather heterogeneous overall result.

DNA quantification

Extraction of DNA was accomplished for all 60 blood samples studied. In addition to the amount autosomal DNA (total DNA) and the amount of Y-chromosomal DNA for each sample and a value for the IPC were determined. The median of the measured amounts of DNA provides (ng / ul extracted DNA) and its dependence on the applied staining method and the different carriers as Fig. 2 illustrated.

Melting curve provisions and total DNA

All completed melting curve provisions confirmed the specificity of the amplification. In quantification of total DNA from whole blood was applied to BW in the native sample from 0.11 to 0.63 ng / µl DNA (median 0.28 ng / µl) were identified. For the "BKA" samples coloring on BW, it was measuable between 0.32 and 0.71 ng / µl of total DNA (median 0.41 ng / µl), and for using the "FBI" samples, between 0.18 and 0.0.66 ng / µl of total DNA (median 0.39 ng / µl).

The untreated samples reported in RG between 0.17 and 0.64 ng / µl of total DNA (median 0.28 ng / µl). After staining using the "BKA" - and "FBI"- method, here could a median of 0.38 ng / µl of a total DNA (variation from 0.10 to 1.12 ng / µl), respectively, a total DNA of 0.41 ng / µl in (variation from 0.15 to 0.88 ng / µl) are detected.

Y-chromosomal DNA

To the native BW sample, the evidence left between 0.17 and 1.41 ng / µl Y-chromosomal DNA (Median 0.42 ng / µl). The sample treated with AS 10B "BKA" showed 0.39 to 0.95 ng / µl (median 0.61 ng / µl), while the sample with the "FBI" mixture from 0.17 to 0.83 ng / µl (median 0.55 ng / µl) Y-chromosomal DNA.

The native sample in the RG had a content of Y-chromosomal DNA from 0.31 to 0.98 ng / µl (median 0.49 ng / µl). With the addition of AS 10B "BKA", values were 0.26 to 1.31 ng / µl (median 0.63 ng / µl), with the addition of AS 10B "FBI" 0.32 to 1.20 ng / µl (median 0.72).

DNA profiles

For all selected samples a complete male profile could not be created.

4 Discussion

Determination of the optimal blood concentration

False-negative or weakly positive results of OBTI are already described in the literature in both large and too low blood levels [3, 4]. This observation was confirmed when determining the optimal blood concentration for the implementation of further experiments. At higher concentrations of human hemoglobin (hHb) in the sample of immobilized anti-hHb antibody unbound hHb can be occupied, making a false-negative result is due (Hook effect). This effect was at 1: 20 dilution observed. A negative test result must therefore be confirmed with a dilution of the sample. The dilution of 1: 200 initially showed no effect on the reaction time (13 s). The color intensity of the forming test line was, however, significantly affected, so the result was "weak positive" classified in the category. The result for the 1: 20.000 dilution also showed the effect of the decreasing intensity of the test line and was therefore assigned to the category "weakly positive".

The optimal dilution proved to be a mixture ratio of 1: 2000, in which, after a few seconds a positive result could be read. This value varies depending on the study - partly were still at dilutions of 1 : 100.000 described clearly positive results [3, 4]. One possible explanation for this could be the lack of information provided by the manufacturer for optimum working temperature of OBTI. The storage of the kit components is recommended at 2-25°C. Whether the test after previous cooling should only obtain room temperature, is not specified. An influence of the temperature of the kit when applied to the sensitivity is conceivable in principle in immunological methods.

Blood Testing using the OBTI

The results for the untreated blood sample in the ratio 1: 2000 revealed no influence of the substrate on the functionality of the OBTI. For all samples within a few seconds clear positive test result was read.

During a meeting of experts "crime scene" in 2007, different formulations of the dye in securing traces of blood were tested and evaluated in AS 10B as a result of especially suitable dye for various track carriers [6]. The experts of the forensic favor an investigation to make the "BKA" -Recipe providing for the solution of the dye in methanol and glacial acetic acid. As an alternative method, especially for working without methanol, the "FBI" -Recipe is given [6]. In addition, there is a wealth of other provisions to prepare the working solution and rinse solutions to be used, the excess,

unbound to proteins dye should be removed from the track carriers. In the application of the dye solution, there are a number of techniques - from the spraying of the track support to the use of immersion baths [1, 6].

The staining with AS 10B "BKA" - and "FBI" -recipe and the use of various trace carriers have partly clear influence on the OBTI blood evidence. Using the track girder RG and consistently positive results for the "BKA" coloring no influence on the sensitivity of the OBTI could be derived. With the "FBI" mixture at least 60% of the samples analyzed in the RG rendered still significantly positive test result.

In contrast, it has been demonstrated that the use of AS-10B recipes with traces of blood on BW predominantly leads to false-negative or only weakly single-positive results, thus limiting the value of OBTI severely limited in terms of pre-tests or no longer exists.

Which chemicals used in the different staining techniques responsible for this result could not be verified on the basis of this study. The effect of different track support is readily explicable that the solutions used can be absorbed by the cotton and washed out with the bad rinsing solutions than be possible in the reaction vessel.

DNA analysis

With regard to the determination of the amount of total and Y-chromosomal DNA in the studies wide variations were between samples of the sub-groups, which, however, showed (Fig. 2) in the overall view of a similar median. These fluctuations can be explained informally by the different cell number of the applied whole blood samples in which no previous cell count was performed. Striking was the over content of total DNA median amount of Y-chromosomal DNA at all quantifications. This could be interpreted as the variance of the PCR. The values are within the natural biological variation between 0.4 and 2.0 for the ratio of autosomal for Y chromosomal DNA of male subjects.

Ultimately, however, in each case, "short tandem repeat" - (STR) analysis was carried out in which no influence could be observed in other studies [12] and other chemicals to the blood evidence seen for AS 10B.

5 Conclusions

For the practical use of AS 10B in the forensic analysis, i is important to note that, in previous application on alleged bloodstains, strong limitations or false-negative results of OBTI may occur. The use of test methods in which the Pseudoperoxidas Reaktion of Hb results in a color change of an initially colorless substance (benzidine, luminol, Leukomalachit etc.) by transfer of liberated oxygen, could offer an alternative. On the possibility to create a DNA profile that blends used "BKA" and "FBI", however, it seems to have no influence. In this respect, the DNA analysis still runs samples in cases of a negative OBTI to relevant results and should always be carried out in case of doubt.

References

[1] Blumenberg C. Daktyloskopische Spurensuche – Amido-Schwarz für blutige Fingerspuren. Kriminalistik, 1991, 8-9: 547-566.
[2] Gaensslen RE. Identification – catalytic tests. In: Gaensslen RE (Hrsg) Source book in forensic serology, immunology and biochemistry. U.S. Department of Justice, United States Government Printing Office, Washington DC, 1983, S: 101-145.
[3] Hermon D, Shpitzen M, Oz C et al. The use of Hexagon OBTI test for detection of human blood at crime scenes and on items of evidence. J Forensic Identif, 2003, 53: 566-575.
[4] Hochmeister MN, Budowle B, Sparkes R et al . Validation studies of an immunochromatographic 1-step test for the forensic identification of human blood. Forensic Sci, 1999, 5: 597-602.
[5] Klein A, Feudel E, Türk E et al. Lumineszenz nach Luminolanwendung. Rechtsmedizin, 2007, 17: 146-152.
[6] Lehmann P . Sichtbarmachung blutiger Fingerspuren mittels unterschiedlicher Amidoschwarz-Ansätze und Acid-Yellow. Expertentagung Tatortarbeit – Spurensicherung LKA, Baden-Württemberg, 2007.
[7] Nielsen K, Mogensen HS, Hedman J et al. Comparison of five DNA quantification methods. Forensic Sci Int, 2008, 2: 226-230.
[8] Krenke BE, Nassif N, Sprecher CJ et al. Developmental validation of a real-time PCR assay for the simultaneous quantification of total human and male DNA. Forensic Sci Int, 2008, 3: 14-21.
[9] Peschel O, Mützel E, Rothschild MA (2008) Blutspurenmuster-Verteilungsanalyse. Rechtsmedizin 18: 131-146.
[10] Peschel O, Rothschild MA, Mützel E. Blutspuren bei Schussverletzungen. Rechtsmedizin, 2010, 20: 91-97.
[11] Ponce AC, Pascual FAV. Critical revision of presumptive tests for bloodstains. Forensic Sci Commun, 1999, 1: 1-5.
[12] Poon H, Elliott J, Modler J, Frégeau C. The use of Hemastix® and the subsequent lack of DNA recovery using the Promega DNA IQ TM System. J Forensic Sci, 2009, 54: 1278-1286.
[13] Seidl S, Hausmann R, Betz P. Comparison of laser and mercury-arc lamp for the detection of body fluids on different substrates. Int J Legal Med, 2008, 122: 241-244.
[14] Thorogate R, Moreira J, Jickells S et al. A novel fluorescence-based method in forensic science for the detection of blood in situ. Forensic Sci Int, 2008, 2: 363-371.
[15] Tobe SS, Watson N, Daeid NN. Evaluation of six presumptive tests for blood, their specificity, sensitivity, and effect on high molecular-weight DNA. J Forensic Sci, 2007, 52: 102-109. ■

Received 22 November 2014
Received in revised form 23 December 2014
Accepted 24 December 2014
Available online 28 December 2014
Peer Reviewed
2157118X.3.3.P4

Qualitative and Quantitative Analysis of Clonazepam and its Metabolite 7-aminoclonazepam in Blood by LC-tandem QTOF/MS and LC-MS/MS

DING Jing [a], ***SHI Yintao*** [a], ***WANG Junwei*** [*, a]

[a] Chongqing Engineering Researh Center for Criminal Investigation Technology, Institute of Forensic Science, Chongqing Public Security Bureau, Chongqing 400021, China.

[*] Corresponding author: Senior Coroner *WANG Junwei*. E-mail: dingjing_111@126.com

Abstract *Objective* A highly sensitive and specific analytical method was developed for the qualitative and quantitative analysis of clonazepam (CLOZ) and its metabolite 7-aminoclonazepam (7-AC) in blood. The analytes were identified by automated solid phase extraction-liquid chromatography-tandem quadrupole time-of-flight mass spectrometry (ASPE-LC-QTOF/MS) and quantitatively determined by automated solid phase extraction-liquid chromatography-tandem triple quadrupole mass spectrometry (ASPE-LC-MS/MS).

Methods After purified by automated solid-phase extraction (ASPE), the qualitative analysis of CLOZ and 7-AC in blood was performed by separation on an Agilent Eclipse Plus C_{18} HPLC column (50 mm×2.1 mm, 1.8 μm) using methanol-0.1%(v/v) formic acid aqueous solution. The detection was performed on a quadrupole time-of-flight mass spectrometer. The quantitative analysis of CLOZ and 7-AC in blood was carried out on LC-MS/MS. The chromatography separation was performed on a Waters Atlantis TM dC18 column (150mm×3.9mm,5.0μm) using methanol-water with 0.1%(v/v) formic acid as mobile phase. The detection was carried out on a triple quadrupole mass spectrometer under electrospray ion source (ESI) in positive ionization mode and multiple reaction monitoring (MRM).

Results The result screening database of antidepressants was built using Agilent Mass Hunter Personal Compound Database Library (PCDL) Manager software and then used for the analysis of spiked samples. The results showed that both CLOZ and 7-AC could be correctly identified with low deviation of retention time (<0.1 min), mass (<1 mDa) and MS matching scores, isotopic abundance matching scores, isotope spacing matching scores (all> 95 points). Furthermore, good linear responses were obtained between the concentration range, 2-1000 ng/mL with a regression coefficient (r^2) higher than 0.99 for both CLOZ and 7-AC. The LOD values for CLOZ and 7-AC were 0.5 ng/mL and 0.2 ng/mL, respectively. The LOQ values for CLOZ and 7-AC were 2 ng/mL and 1 ng/mL, respectively. The RSD values obtained from intra-day and inter-day experiments ranged from 4.2 to 10.3%. The recovery ranges for CLOZ and 7-AC were all above 70% for three spiked levels.

Conclusion The developed method was further applied for the analysis of poisoning case, both CLOZ and 7-AC in blood were confirmed and quantified. The combination of ASPE-LC-QTOF/MS and ASPE-LC-MS/MS is suitable for the identification and quantification of benzodiazepines and their metabolites in forensic toxicology.

Keywords: Forensic science, Toxicology, Clonazepam, 7-aminoclonazepam, Metabolite, Automated solid phase extraction, Liquid chromatography, Tandem quadrupole time-of-flight mass spectrometry, Triple quadrupole mass spectrometry, Blood, ASPE-LC-QTOF/MS, ASPE-LC-MS/MS.

1 Introduction

Clonazepam (CLOZ) is a benzodiazepine derivative that was approved for use as an anticonvulsant in the US in 1975[1]. At present, CLOZ has been one of the most frequently prescribed psychoactive drugs world wide due to its hypnotic, anxiolytic, anticonvulsant and muscle-relaxant properties[2-3]. However, CLOZ, because of its pharmacological effects has been identified as a compound frequently used in drug-facilitated crimes (DFC) such as robberies and sexual assaults in recent years[4-6].

CLOZ has a plasma half-life varying from 19 to 60 h, where the mean value is 40 h[7]. It is extensively metabolized in the liver, primarily by CYP3A4, to its major metabolite 7-aminoclonazepam (7-AC)[8]. 7-AC is excreted mainly in urine. Only about 0.5% of parent drug is

excreted unchanged in the urine[6]. In clinical studies oral administration of a single 2 mg dose resulted in an average plasma concentration of 17 ng/mL (range: 7-24 ng/mL) of CLOZ between 1 and 4 h after ingestion[9]. In studies of patients receiving 6 mg/day chronic therapy, the plasma concentration of CLOZ and 7-AC were reported as 29-75 ng/mL and 23-137 ng/mL, respectively[10].

Because of above metabolic characteristics, sometimes parent drug CLOZ cannot be directly detected while only its major metabolite 7-AC can be detected[7-8]. This indicated the detection of 7-AC can be regarded as evidence of CLOZ intake. It has been reported that postmortem bioconversion to 7-amino metabolites may also occur, to the extent that little or no parent drug may be present; the identification and quantification of the 7-amino metabolites are of importance toxicologically as they are often the only indication of nitrobenzodiazepine use prior to death[11]. Consequently the simultaneous determination of CLOZ and its metabolite 7-AC is very important in the fields of forensic toxicology.

On the other hand, in most cases, victims don't often report until after some time; therefore, blood and urine samples are often detected more than 24-72 h after ingestion[6]. Moreover, only immunoassay screening and sometimes liquid chromatography-diode array detector (HPLC-DAD) and gas chromatography-mass spectrometry (GC-MS) analysis are performed on blood and urine at the hospital laboratory where the victim is admitted[12]. Since most of the drugs involved in DFC are not detectable at low levels by these techniques, resulting false negative results may lead to an inaccurate conclusion and to premature destruction of samples[12]. For this reason it is important to develop highly sensitive and specific analytical methods for the simultaneous determination of CLOZ and its metabolite 7-AC in biological fluids.

Numerous analytical methods including spectrophotometry[13-14], gas chromatography (GC)[15-19], gas chromatography-mass spectrometry (GC-MS)[20-24] and high-performance liquid chromatography (HPLC)[3, 4, 9, 25-26] have been presented for testing CLOZ in biological fluids. However, when these methods are used, some problems of low sensitivity, specificity and reproducibility are encountered[20]. Due to its ability to analyze thermally unstable or polar compounds, such as 7-AC, liquid chromatography (LC) based methods offer a distinct advantage over GC related techniques. Recently, detection of CLOZ and its metabolite 7-AC in biological fluids has become available with the development of more and more sensitive LC-MS apparatuses[27]. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) has emerged as the most suitable tool to test for low concentrations of CLOZ[2, 28-31] and its metabolite 7-AC[5, 7, 12, 27, 32].

At present, LC-MS/MS is widely used in the determination of benzodiazepines and their related drugs due to its strong specificity, high sensitivity, good reproducibility, wide linear dynamic range and other characteristics[30]. But LC-MS/MS can only analyze the target compound which is included in established database, while it can't screen the unknown compound which is not included in established database no matter how much concentration of the unknown compound is. These lacks of LC-MS/MS assays limit qualitative analysis and rapid screening of target compound containing multiple components at trace levels (e.g., ng/L)[29].

Recent advances in LC-MS technology have lead to the availability of time-of-flight (TOF) LC/MS systems. These provide a greater level of the analyte information as a result of high resolution and the ability to collect accurate mass information to the sub ppm levels[33]. This significantly increases the confidence in the analyte identification by limiting the possible number of candidate compounds. The coupling of a quadrupole and a collision cell to the TOF analyzer, the fragmentation of pre-selected ions and the identification of compounds based on their product ion spectra. Accurate mass determination of both the precursor and product ions is therefore possible[34-35]. Rapid screening and confirmation of the analyte can be achieved with MS matching scores, deviation of retention time, measured mass, isotopic abundance matching scores and isotope spacing match scores and MS/MS matching scores[36-37].

With the exception of some immunoassays[1], liquid liquid extraction (LLE)[3, 9, 20, 21, 25] and solid phase extraction (SPE)[28, 31, 38-40] are commonly included in sample pretreatment methods for the determination of CLOZ and 7-AC in biological fluids. Compared with LLE, SPE can improve the extraction recovery of CLOZ and 7-AC in biological fluids due to its advantages such as less use of organic solvents, stronger selectivity, shorter separation time, better reproducibility and better the clean-up effect. Furthermore, a large-diameter extraction disk is equipped with fully automatic solid phase extraction system (ASPE) for SPE, and the sample pretreatment procedure becomes simple and efficient[29].

The aim of this report is to develop a highly sensitive and specific analytical method for qualitative and quantitative analysis of CLOZ and its metabolite 7-AC in blood by using ASPE-LC-QTOF/MS combined with ASPE-LC-MS/MS. This study details the use of a ASPE coupled to LC-QTOF/MS system for the qualitative determination of CLOZ and its metabolite 7-AC in blood. The screening database of antidepressants is built using Agilent Mass Hunter

Personal Compound Database Library (PCDL) Manager software and then used to screen and confirm CLOZ and its metabolite 7-AC in blood. In addition, a ASPE coupled to LC-MS/MS system is used for the quantitative determination of CLOZ and 7-AC in blood. In order to evaluate the reliability of the developed method, it was applied for the simultaneous determination of CLOZ and 7-AC in case of poisoning.

2 Materials and Experiments

2.1 Chemicals and reagents

CLOZ and 7-AC were purchased from Sigma-Aldrich (St Louis, USA). HPLC-grade methanol and acetonitrile were obtained from TEDIA (Ohio, USA). All other chemicals were of analytical-grade in the highest purity available. Water was distilled and purified using a Millipore Milli-Q Plus system (Bedford, USA).

2.2 Preparation of stock solutions and working solutions

Stock solutions of CLOZ and 7-AC were prepared by dissolution of each drug in methanol to obtain a concentration of 1 mg/mL. Working solutions of CLOZ and 7-AC were prepared respectively each day by making appropriate dilutions of the stock solutions in methanol. All these solutions were stored at -20 °C in the absence of light.

2.3 Sample pretreatment

1 mL blood sample was added a 16 mm × 125 mm screw top test tube and diluted with 2 mL of Milli-Q water and then buffered with 2 mL of saturated sodium carbonate (Na_2CO_3) solution to give a final pH of 9. The tube was capped and submitted to vortex mixing for 5 min.

2.3.1 Liquid-liquid extraction (LLE)

The resultant mixture was shaken for 5min with 5 mL ethyl acetate. After centrifugation at 9500 r/min for 5 min, the upper organic phases were transferred to conical tubes and separated. The aqueous phase was re-extracted with 5 mL ethyl acetate. The organic phases were evaporated at 40 °C under flow of nitrogen to obtain good results. The residue was dissolved in 1 mL of the mobile phase and filtered with a Nylon filter (0.22 μm).

2.3.2 Automatable solid-phase extraction (ASPE)

A C18 cartridge (500 mg/3 mL, Waters technologies, USA) was activated with 3 mL of methanol and washed with 3 mL of Milli-Q water. After that, the resultant mixture was loaded and washed with 3 mL of Milli-Q water. The elution was performed with 3 mL of methanol. The whole SPE procedure was performed on a Gilson GX-274 fully automatable solid phase extraction apparatus (Wisconsin, USA). The collected solvent was evaporated at 40 °C under flow of nitrogen and the residue was dissolved in 1 mL of the mobile phase and filtered with a Nylon filter (0.22 μm).

2.4 Instrumentation

The qualitative detection of analyses was performed on the Agilent 1290-6550 ultra-high performance liquid chromatography-tandem quadrupole time-of-flight mass spectrometer equipped with Dual Agilent Jet Stream Electrospray Ionization (Dual AJS ESI) (Agilent Technologies, Santa Clara, USA). The quantitative detection of analyses was carried out using the LC-20A liquid chromatography (Shimadzu, Japan)-API2000 tandem triple quadrupole mass spectrometer (AB Technologies, USA).

2.5 Analytical conditions

2.5.1 Chromatographic conditions

Qualitative conditions: the chromatographic separation was performed on Agilent Eclipse Plus C18 (50 mm × 2.1 mm, 1.8 μm; Agilent Technologies, Santa Clara, USA) at 35 °C with the eluents A = methanol and B = water (containing 0.1% formic acid) and the following time program of the gradient: 0.0-2.0 min 20% A, linear to 95% A at 7.0 min, const. 95% A to 9.0 min, back to 20% A at 9.1 min and equilibration for 3 min. The flow rate was 0.3 mL/min.

Quantitative conditions: gradient elution was performed on separation column (Waters Atlantis TM dC18, 50 mm × 3.9 mm, 5.0 μm; Waters technologies, USA). The mobile phase consisted of methanol (eluent A) and water containing 0.1% formic acid (eluent B). The gradient was programmed as follows: 0.0-5.0 min: gradient increases from 20% to 100% eluent A; 5.0-12.0 min: gradient 100% eluent A; 12.0-12.1 min: gradient decreases to 20% eluent A; 12.1-16.0 min: gradient 20% eluent A. The column oven was set at 35 °C and the flow rate was 1.0 mL/min.

2.5.2 MS conditions

Qualitative conditions: The QTOF-MS instrument was operated under electrospray ion source (ESI) in positive ionization mode with full scanning range of *m/z* 50-1000 Da. The source parameters were: dry gas temperature 200 °C, dry gas flow rate 16.0 L/min, nebulizer pressure 35 psi, sheath gas temperature 350°C, sheath gas flow rate 12.0 L/min, VCap voltage 4000 V and fragmentor voltage 135 V, reference ions for mass calibration: purine 121.050873 $[M+H]^+$, HP-921 = hexakis (1H,1H,3H-tetrafluoropropoxy) phosphazine 922.009798 $[M+H]^+$.

Data acquisition and qualitative management were performed by Agilent Mass Hunter Workstation Data Acquisition and Agilent Mass Hunter Qualitative Analysis (B.06.00) software, respectively. The screening database of antidepressants was established by Agilent Mass Hunter Personal Compound Database Library (PCDL) Manager software and then used for the analysis of spiked samples.

Quantitative conditions: For detection and quantification the following ESI inlet conditions were

applied: gas 1, nitrogen (35 psi); gas 2, nitrogen (40 psi); ion spray voltage, 5200 V; ion source temperature, 450 °C; curtain gas, nitrogen (20 psi). The mass spectrometer was operated in the information dependent acquisition (IDA) mode. As survey scan the multi-reaction monitoring mode (MRM) was used, followed by the dependent scan, which was an enhanced product ion scan (EPI). The settings for the MRM mode were as follows: collision gas was set at 4, the dwell time was set at 20 ms. All other settings were analyte-specific and were determined using Analyst software in the quantitative optimization mode. The MRM transitions per analyte and respective settings for both LC-MS/MS systems are summarized in Table 1.

2.6 Method validation

The method was validated for the following parameters: sensitivity (limit of detection LODs and limit of quantification LOQs), linearity, intra- and inter-day precision, recovery and matrix effect.

2.6.1 Linearity

Analytes were quantified by means of calibration curves constructed from seven known concentrations of analyte standards in pure solvent (2, 5,10, 50, 200, 500 and 1000 ng/mL). Linearity of calibration curves based on peak area as function of the nominal concentration was assessed by least square regression[25].

2.6.2 Sensitivity

The limit of detection (LOD) was considered the lowest concentration of these analytes corresponding to three times the background noise or relationship signal to noise ratio 3:1. The limit of quantification (LOQ) was defined as the lowest point of the calibration curve and fulfilled the requirement of LOQ signal-to-noise ratio of 10:1[39].

2.6.3 Precision and recovery

Intra- and inter-day precision of the analytical method were shown by following the analysis of three different spiked blood samples (50, 200 and 800 ng/mL). Precision was calculated as relative standard deviation (RSD) of the experimental concentrations[33].

The recoveries of COLZ and 7-AC were determined by comparing the response of the analyte from pure standard (n = 6) with the response of the analyte from extracted spiked blood sample at equivalent quantity[27]. Recovery was determined at low, mid, and high quality control concentrations, 50, 200 and 800 ng/mL respectively.

2.6.4 Matrix effect

The matrix effect of biological samples can affect the analyte signals, enhance background noises or suppress the analyte responses[41]. This matrix effects may result as positive or negative responses depending on the level of ion suppression and can greatly affect the method reproducibility and accuracy[42].

3 Results and Discussion

3.1 Optimization of sample pretreatment

3.1.1 Liquid-liquid extraction (LLE) and Automated solid phase extraction (ASPE)

The impact of the extraction pretreatment (LLE or ASPE) of blood sample on extraction recoveries of analyte in blood sample was investigated in LC-QTOF/MS system. Extraction recoveries were calculated from peak area ratios between extracted blood samples spiked at three concentrations (50, 200 and 800 ng/mL) before and after the extraction pretreatment[45].

It was shown in Table 2 that extraction recoveries of CLOZ and 7-AC in blood sample by LLE ranged between 56.4% and 82.5%, while extraction recoveries of CLOZ and 7-AC in blood sample by ASPE ranged between 68.4% and 91.8%. The results suggested that the extraction efficiency of CLOZ and 7-AC in blood sample by ASPE were better than that by LLE. On the other hand, the RSD values of CLOZ and 7-AC in blood sample by ASPE were obviously lower than those by LLE. The results indicated that compared with LLE, ASPE provided superior stability, better repeatability and lower susceptibility to matrix interferences for the extraction pretreatment of CLOZ and 7-AC in blood sample.

3.1.2 pH values of blood sample

In the present study the effect of different pH values of blood sample extracted by ASPE on extraction recoveries of the analyte at medium

Table 1. The MRM transitions per analyte and respective settings for LC-MS/MS system (CE: collision energy [V], DP: declustering potential [V], EP: entrance potential [V], CEP: collision cell exit potential [V]).

Analyte	Ion pairs (m/z)	CE	DP	EP	CEP
CLOZ	316.1/270.0* 316.1/241.1	26,36	20	10	8
7-AC	286.2/222.2* 286.2/250.1	36,28	20	10	8

* stands for the quantitative ion pair.

Table 2. Extraction recoveries of analytes in blood sample by liquid-liquid extraction(LLE) and automated solid phase extraction (ASPE) (n = 6, LC-QTOF/MS system)

Analyte	Concentration added(ng/mL)	LLE		ASPE	
		Recovery (%)	RSD(%)	Recovery(%)	RSD(%)
CLOZ	50	56.4	8.5	68.4	6.8
	200	70.3	7.9	78.7	5.9
	800	78.9	5.4	87.3	5.4
7-AC	50	61.7	8.9	73.7	7.1
	200	75.3	9.2	86.9	5.4
	800	82.5	5.8	91.8	4.4

concentration was also investigated. The results were shown in Table 3. It was seen that increasing sample pH value provided higher extraction recoveries for CLOZ and 7-AC in blood sample. Nevertheless, in order to avoid analyte decomposition in strong alkaline solutions (pH ≥11) and blood solidification resulting in low extraction recovery, pH 9 was chosen as final pH value of blood sample extracted by ASPE.

3.2 Establishment and optimization of qualitative method in LC-QTOF/MS system (*Database building and screening)*

A large number of antidepressants, which included CLOZ and 7-AC, were selected as target compounds and the screening database of them was built using Agilent Mass Hunter PCDL Manager software. The target compound name, molecular formula, retention time and extract mass for (de)protonated compounds, and information on the elemental compositions for their main fragment ions were included in the database[43]. The resulting screening database was built using Agilent Mass Hunter PCDL Manager software and then used to screen unknown samples based on a combination of retention time, observed spectral molecular weight and isotope ratio, generating a match score for any identified compounds[35]. Further confidence in analyte identification can then be obtained through the selected fragmentation of any identified compounds, and comparing the MS/MS data with that stored as library spectra[36].

Thus, the built screening database was used for the analysis of spiked samples. MS matching scores, isotopic abundance matching scores, isotope spacing matching scores (all> 95 points) and MS/MS matching scores (> 70 points)[3] were applied to identify the analytes (CLOZ and 7-AC). The results were shown in Table 4. As it can be seen that both CLOZ and 7-AC could be correctly identified with low deviation of retention time (<0.1 min), mass (<1 mDa) and MS matching scores, isotopic abundance matching scores, isotope spacing matching scores (all> 95 points).

Moreover, MS/MS matching scores of 7-AC were 73.87 points (> 70 points) while MS/MS matching scores of CLOZ were 65.42 points (< 70 points). The main reason was that this drug CLOZ might produce the low signal response (e.g. in sensitivity) and might be highly affected by matrix effect[44]. As shown in Fig. 1, the agreement of the spectra (similarity index and specific fragments m/z of identity) as well as of the retention time were used as criteria for identification. This indicated that CLOZ was screened and identified.

3.3 Establishment and optimization of quantitative method in LC-MS/MS system

3.3.1 Optimization of chromatographic Separation

Table 3. Extraction recoveries of C18 disk under different pH values (n = 3, LC-QTOF/MS system)

Analyte	Extraction recoveries (%)				
	pH 3.0	pH 5.0	pH7.0	pH9.0	pH 11.0
CLOZ	42.5±8.4	58.2±5.9	72.5±6.3	80.3±5.6	78.6±6.4
7-AC	51.8±7.8	64.2±8.1	80.4±6.0	85.9±5.7	87.2±7.1

Table 4. Screening results of analytes from Agilent MassHunter PCDL Manager software (LC-QTOF/MS system)

Name	MS matching scores	Theoretical Mass/Da	Measured Mass/Da	Deviation of Mass/mDa	Deviation of retention time/min	Isotopic abundance matching scores	Isotope spacing matching scores	MS/MS matching scores
CLOZ	91.68	315.0411	316.0499	-1.58	-0.031	95.85	99.18	65.42
7-AC	97.68	285.0675	286.0748	-0.63	-0.064	96.76	99.38	73.87

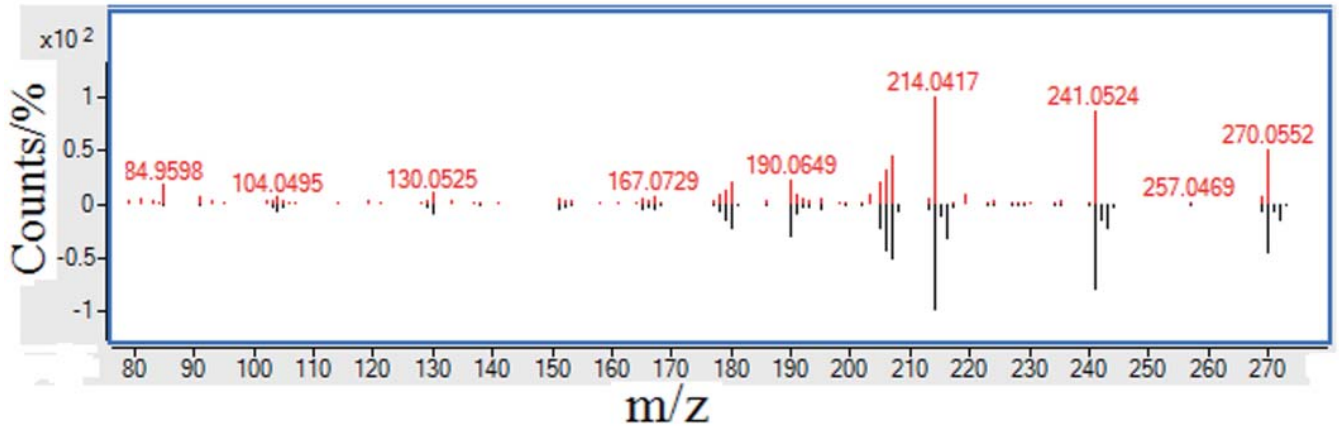


Fig 1. A chromatogram identifying CLOZ in blood sample using the custom PCDL(retention time < 0.1 min, mass < 1 mDa).

The optimization of LC conditions was focused on improving the shape of peaks in order to achieve good separation among major substances from complicated matrices. Thus, the effects of two different mobile phases on ionization degree were compared in the present study. Two different mobile phases were (1) acetonitrile -water with 0.1%(v/v) formic acid and (2) methanol-water with 0.1%(v/v) formic acid. The result suggested a better chromatographic resolution and sensitivity for blood samples using methanol-water with 0.1%(v/v) formic acid as mobile phase.

Furthermore, several other chromatographic conditions, such as type of column and its length, flow rate, elution gradient program and column temperature were also optimized to obtain a satisfactory chromatographic separation (good resolution and efficiency) for target analytes. These optimized chromatographic conditions were described in Section 2.5.1 and produced the results presented in Fig. 2. It can be observed that CLOZ had a retention time of 5.82 min and 7-AC of 4.03 min.

3.3.2 Optimization of MS conditions

The MS spectra were obtained in positive mode using LC-MS/MS, and the experimental conditions were optimized. Results were described in Section 2.5.2 (Table 1). As shown in Table 1, the protonated molecule $[M+H]^+$ were observed at *m/z* 316.1 and 286.2 in Q1 full scan for CLOZ and 7-AC, respectively. The major fragments observed in MS-MS spectrum of CLOZ and 7-AC were at *m/z* 270.0 and 241.1, 222.2 and 250.1, respectively. Precursor-product ion transitions were selected according to the highest sensitivity, optimal selectivity and reproducibility of the ion ratio. Confirmation of the identity of CLOZ and 7-AC was based on retention time, transitions selected and ion ratio statistics for the

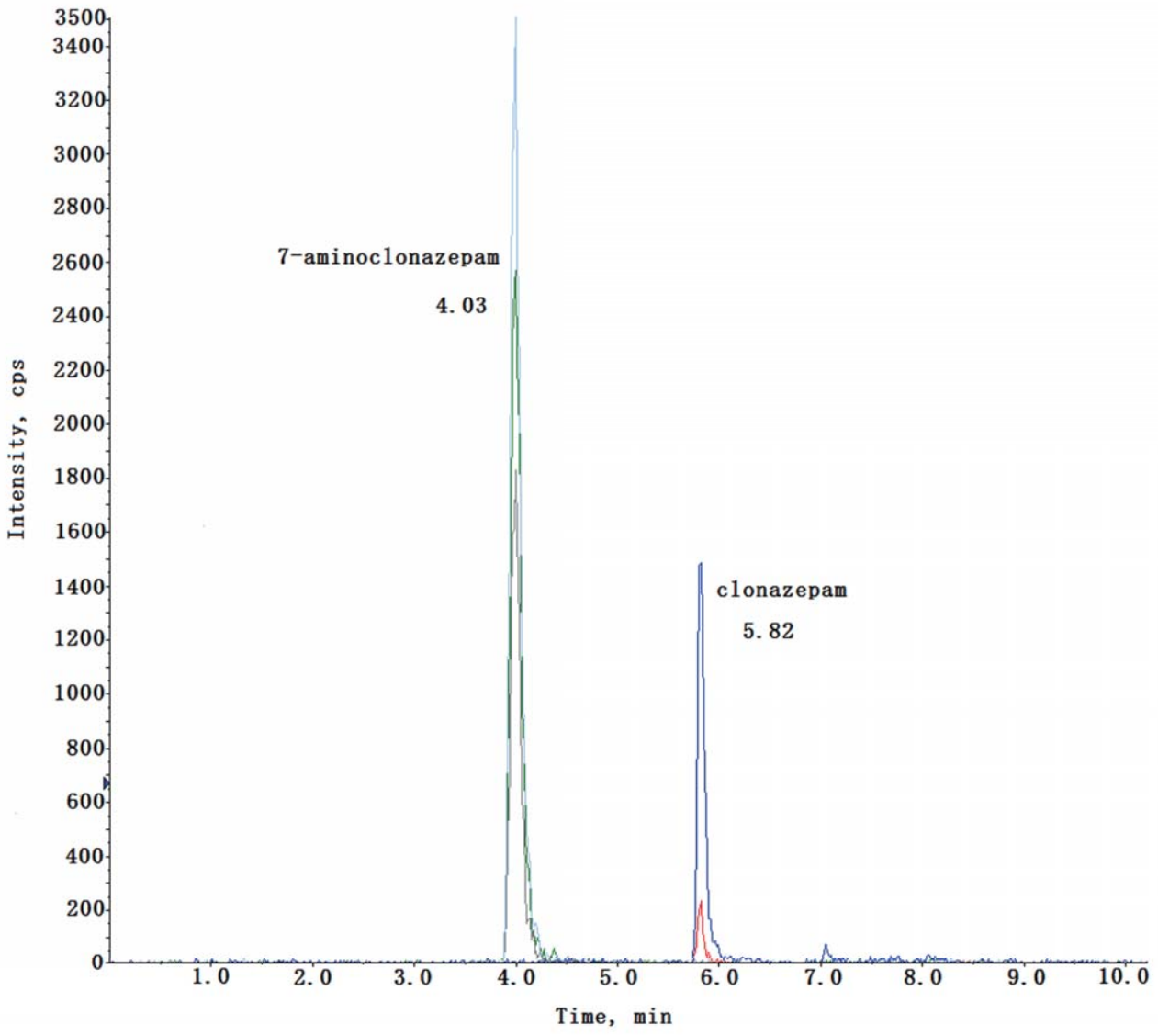


Fig 2. MRM chromatogram (LC-MS/MS system) of CLOZ(clonazepam) and 7-AC(7-aminoclonazepam) (100 ng/mL) in methanol.

transition monitored. The *m/z* ions 270.0 and 222.2 were used for the quantification of CLOZ and 7-AC, respectively, with 241.1 and 250.1 as qualifier ions. An external standard method was developed for qualifying CLOZ and 7-AC over a level of 2-1000 ng/mL. Replicate calibration standards (n=6) for each level were analyzed, and peak areas were used for quantification.

3.4 Quantitative method validation in LC-MS/MS system

3.4.1 Linearity, LOD and LOQ

Good linear responses were obtained between the concentration range, 2-1000 ng/mL with a regression coefficient (r^2) higher than 0.99 for both CLOZ and 7-AC. Each calibration equation was fitted by the linear regression equation $y=mx+c$ where y was the signal peak area of the spiked analyte and x is the corresponding concentration of the spiked analyte. Slopes, intercepts, and regression coefficients were summarized in Table 5.

LOD and LOQ for all analyzing CLOZ and 7-AC were calculated in the range of 2-1000 ng/mL and were also presented in Table 5. It can be seen that the LOD values for CLOZ and 7-AC were 0.5 ng/mL and 0.2 ng/mL, respectively. The LOQ values for CLOZ and 7-AC were 2 ng/mL and 1 ng/mL, respectively. This indicated that the proposed method was suitable for the high sensitive determination of CLOZ and 7-AC in blood samples.

3.4.2 Precision and recovery

In this work, the precision of intra- and inter-day of low, medium and high (50, 200 and 800ng/mL) spiked blood samples were determined. The RSD values obtained from intra-day (five successive injections) and inter-day (three successive days) experiments ranged from 4.2 to 10.3%. Recovery study was carried out by preparing the spiked blood samples at three concentrations of low, intermediate and high spiked levels as mentioned in precision. The recovery ranges for CLOZ and 7-AC were all above 70% for three spiked levels. Results of both precision and recovery tests were shown in Table 6.

3.4.3 Evaluation of matrix effect

Due to the complexity and high inter-subject variability of blood, ESI-MS signals may seriously impair the reliability of a method[41]. The assessment of matrix effects in different batches of blood has been included as part of the validation procedure. In the present study, the matrix effect was calculated by comparing peak areas of each analyte standard in pure solvent (S) with the those of corresponding peak areas of spiked blood extracted by ASPE, according to the following equation[42].

Ion suppression or enhancement (%) = (Y - X) / X ×100

Where, X is the peak area of an analyte standard in S and Y is the peak area of the analyte standard in blood extracted by ASPE. The results showed that the values of matrix effects for CLOZ and 7-AC in extracted blood samples spiked at three concentrations (50, 200 and 800 ng/mL) were 6-14% and -14-1%, respectively. The ion suppression or enhancement by blood was within ±15% for all for both CLOZ and 7-AC and was within acceptable experimental errors. It was confirmed that after spiked blood sample extracted by ASPE, no serious matrix interferences with the analyte's signal occurred.

3.5 Case examples

The practical application of the developed method using ASPE-LC-QTOF/MS combined with ASPE-LC-MS/MS was demonstrated at two examples.

3.5.1 Case #1

This is a sample of blood taken from a 58-year-old male. He was robbed at home after having had a cup of coffee offered by two men who spent the night at his home. Blood was collected at the hospital 24h after the crime.

3.5.2 Case #2

This is a sample of blood taken from a victim of drug-facilitated sexual assault. The 23-year-old female suspected that she had been victim of a sexual assault during her stay at a retiring room of a nightclub after having had two drinks in the nightclub. Blood was collected at the hospital 48h after the crime.

The blood samples of two case examples were analyzed by the present methodology. The CLOZ and 7-AC levels of Case #1 were 24.5 and 30.8 ng/mL, respectively. The CLOZ and 7-AC levels of Case #2 were 12.5 and 20.2 ng/mL, respectively. The results indicated that this

Table 5. Quantitative performance of the analytes in LC-MS/MS system

Analyte	Calibration range (ng/mL)	Calibration curve	Regression coefficient (r^2)	LOD (ng/mL)	LOQ (ng/mL)
CLOZ	2-1000	$y = 90.2x+11.6$	0.9964	0.5	2
7-AC	2-1000	$y = 167x+1050$	0.9918	0.2	1

Table 6. Recoveries and RSDs of the analytes in LC-MS/MS system (n = 6)

Analyte	Concentration added(ng/mL)	Recovery (%)	RSD (%)	
			Intra-day	Inter-day
CLOZ	50	72.6	5.6	10.3
	200	78.6	5.2	8.7
	800	85.6	4.8	7.6
7-AC	50	79.1	6.2	9.4
	200	84.2	5.4	8.2
	800	96.3	4.2	6.8

Recovery = peak area of the analyte spiked in blood after ASPE/peak area of the analyte spiked with the same concentration in blank matrix extract×100%

metabolite 7-AC can be regarded as the target analyte to prove ingestion of clonazepam. Furthermore, CLOZ and 7-AC were judged to be the significant contributors in the two poisoning cases.

4 Conclusion

In the present work, the combination of ASPE-LC-QTOF/MS and ASPE-LC-MS/MS proved to be a practicable and efficient way for the simultaneous identification and quantification of CLOZ and 7-AC in blood. The developed method using ASPE-LC-QTOF/MS is regarded as a sensitive and specific analytical method successfully applied to the simultaneous screening and confirmation of CLOZ and 7-AC in blood due to QTOF analyzer's higher sensitivity, its much higher separation power by retention time as well as by accurate mass. Additionally, by using ASPE-LC-MS/MS, a sensitive and robust analytical method for the simultaneous quantification of CLOZ and 7-AC in blood has also been developed with good linearity, LOD, LOQ, precision and recovery as well as its low susceptibility to matrix interferences.

The above developed method was further applied for the analysis of poisoning cases, both CLOZ and 7-AC in blood were confirmed and quantified. In conclusion, the developed method using ASPE-LC-QTOF/MS combined with ASPE-LC-MS/MS can be used in future studies for the simultaneous identification and quantification of benzodiazepines and their metabolites in forensic toxicology.

Acknowledgements

The authors would like to thank Agilent Technologies Inc. for generous technical support of these investigations and in particular Dr. JIANG Chang (Chengdu, China), Mr. GUO Jingqi and Mr. WANG Huijun(Chongqing, China) for helpful cooperation and discussions.

This paper is supported by the Natural Science Foundation of Chongqing of China (cstc,2011pt-gc00004), and the Key Science and Technology Program of Chongqing Public Security Bureau (G2013-04).

Reference

[1] A A Elian. ELISA detection of clonazepam and 7-aminoclonazepam in whole blood and urine. Forensic Sci Int, 2003, 134: 54-56.
[2] E N Sauve, M Langødegård, D Ekeberg, Å M L Øiestad. Determination of benzodiazepines in ante-mortem and post-mortem whole blood by solid-supported liquid–liquid extraction and UPLC–MS/MS. J Chromatogr B: Biomed Appl, 2012, 883: 177-188.
[3] A E Mahjoub, C Staub. Simultaneous determination of benzodiazepines in whole blood or serum by HPLC/DAD with a semi-micro column. J Pharm Biomed Anal, 2000, 23: 447-458.
[4] W M Mullett, J Pawliszyn. Direct LC analysis of five benzodiazepines in human urine and plasma using an ADS restricted access extraction column. J Pharm Biomed Anal, 2001, 26: 899-908.
[5] M Chèze, M Villain, G Pépin. Determination of bromazepam, clonazepam and metabolites after a single intake in urine and hair by LC-MS/MS application to forensic cases of drug facilitated crimes. Forensic Sci Int, 2004, 145: 123-130.
[6] L W Storhaug, A Enger, K Hjelmeland, E L Øiestad, V Vindenes. Prolonged excretion of 7-aminoclonazepam in urine after repeated ingestion of clonazepam: A case report. Forensic Sci Int, 2012, 222: 33-35.
[7] A Steentoft, K Linnet. Blood concentrations of clonazepam and 7-aminoclonazepam in forensic cases in Denmark for the period 2002-2007. Forensic Sci Int, 2009, 184: 74-79.
[8] A A Salem, B N Barsoum, E L Izake. Spectrophotometric and fluorimetric determination of diazepam, bromazepam and clonazepam in pharmaceutical and urine samples. Spectrochim Acta Part A, 2004, 60: 771-780.
[9] I F Bares, F Pehourcq, C Jarry. Development of a rapid RP-HPLC method for the determination of clonazepam in human plasma. J Pharm Biomed Anal, 2004, 36: 865-869.
[10] J Hackett, A A Elian. Extraction and analysis of clonazepam and 7-aminoclonazepam in whole blood using a dual internal standard methodology. Forensic Sci Int, 2007, 166: 209-217.
[11] M D Robertson, O H Drummer. High-performance liquid chromatographic procedure for the measurement of nitrobenzodiazepines and their 7-amino metabolites in blood. J Chromatogr B: Biomed Appl, 1995, 667: 179-184.
[12] M Chèze, G Duffort, M Deveaux, G Pépin. Hair analysis by liquid chromatography-tandem mass spectrometry in toxicological investigation of drug-facilitated crimes: Report of 128 cases over the period June 2003-May 2004 in metropolitan Paris. Forensic Sci Int, 2005, 153: 3-10.
[13] F R Gil, J A Daros, A Salvador, M de la Guardia. Direct derivative spectrophotometric determination of nitrazepam and clonazepam in biological fluids. J Pharm Biomed Anal, 1991, 9: 539-545.
[14] F R Gil, A Salvador, M de la Guardia. Influence of the differentiation system on the analytical parameters for the spectrophotometric determination of clonazepam in urine. Microchem J, 1991, 44: 249-257.
[15] A G de Boer, J R Kaiser, H Bracht, D D Breimer. Assay of underivatized nitrazepam and clonazepam in plasma by capillary gas chromatography applied to pharmacokinetic and bioavailability studies in humans. J Chromatogr B: Biomed Appl, 1978, 145: 105-114.
[16] N R Badcock, A C Pollard. Micro-determination of clonazepam in plasma or serum by electron-capture gas-liquid chromatography. J Chromatogr B: Biomed Appl, 1982, 230: 353-361.
[17] P Lillsunde, T Seppälä. Simultaneous screening and quantitative analysis of benzodiazepines by dual-channel gas chromatography using electron-capture and nitrogen-phosphorus detection. J Chromatogr B: Biomed Appl, 1990, 533: 97-110.
[18] J Naestoft, N Larsen. Quantitative determination of clonazepam and its metabolites in human plasma by gas chromatography. J Chromarogr, 1974, 93: 113-122.
[19] D Song, S Zhang, K Kohlhof. Quantitative determinationof clonazepam in plasma by gas chromatography-negative ion chemical ionization mass spectrometry. J Chromatogr B: Biomed Appl, 1996, 686: 199-204.
[20] M Augsburger, L Rivier, P Mangin. Comparison of different immunoassays and GC-MS screening of benzodiazepines in urine. J Pharm Biomed Anal, 1998, 18: 681-

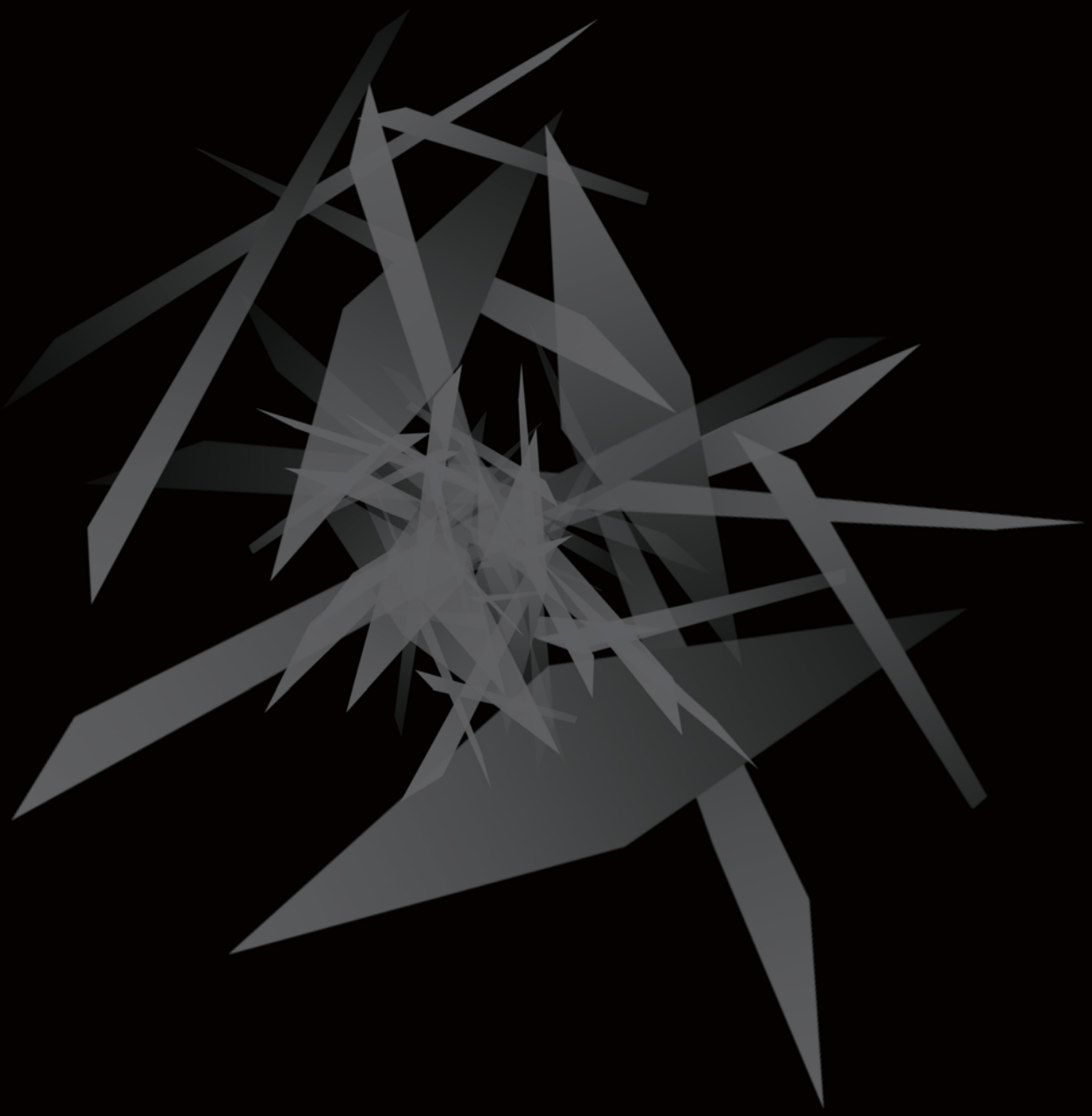
687.
[21] I I Papoutsis, S A Athanaselis, P D Nikolaou, C M Pistos, C A Spiliopoulou, C P Maravelias. Development and validation of an EI-GC-MS method for the determination of benzodiazepine drugs and their metabolites in blood: Applications in clinical and forensic toxicology. J Pharm Biomed Anal, 2010, 52: 609-614.
[22] H Maurer, K Pfleger. Identification and differentiation of benzodiazepines and their metabolites in urine by computerized gas chromatography-mass spectrometry. J Chromatogr B: Biomed Appl, 1987, 422: 85-101.
[23] M Yegles, F Mersch, R Wennig. Detection of benzodiazepines and other psychotropic drugs in human hair by GC/MS. Forensic Sci Int, 1997, 84: 211-218.
[24] A Negrusz, C M Moore, J L Kern, P G Janicak, M J Strong, N A Levy. Quantitation of clonazepam and its major metabolite 7-aminoclonazepam in hair. J Anal Toxicol, 2006, 24: 614-620.
[25] C L Guellec, M L Gaudet, M Breteau. Improved selectivity for high-performance liquid chromatographic determination of clonazepam in plasma of epileptic patients. J Chromatogr B: Biomed Appl, 1998, 719: 227-233.
[26] J C Spell, J T Stewart. Analysis of clonazepam in a tablet dosage form using smallbore HPLC. J Pharm Biomed Anal, 1998, 18: 453-460.
[27] K Y Rust, M R Baumgartner, N Meggiolaro, T Kraemer. Detection and validated quantification of 21 benzodiazepines and 3 “z-drugs” in human hair by LC-MS/MS. Forensic Sci Int, 2012, 215: 64-72.
[28] C A A de Almeida, C G B Brenner, L Minetto, C A Mallmann, A F Martins. Determination of anti-anxiety and anti-epileptic drugs in hospital effluent and a preliminary risk assessment. Chemosphere, 2013, 93: 2349-2355.
[29] S L Yuan, X F Li, X M Jiang, H X Zhang, S K Zheng. Simultaneous determination of 13 psychiatric pharmaceuticals in sewage by automated solid phase extraction and liquid chromatography-mass spectrometry. Chin J Anal Chem, 2013, 41: 49-56.
[30] M Jang, H Chang, W Yang, H Choi, E Kim, B H Yu, Y Oh, H Chung. Development of an LC-MS/MS method for the simultaneous determination of 25 benzodiazepines and zolpidem in oral fluid and its application to authentic samples from regular drug users. J Pharm Biomed Anal, 2013, 74: 213-222.
[31] R Verplaetse, E Cuypers, J Tytgat. The evaluation of the applicability of a high pH mobile phase in ultrahigh performance liquid chromatography tandem mass spectrometry analysis of benzodiazepines and benzodiazepine-like hypnotics in urine and blood. J Chromatogr A, 2012, 1249: 147-154.
[32] R C Irving, S J Dickson. The detection of sedatives in hair and nail samples using tandem LC-MS-MS. Forensic Sci Int, 2007, 166: 58-67.
[33] S Broecker, S Herre, F Pragst. General unknown screening in hair by liquid chromatography-hybrid quadrupole time-of-flight mass spectrometry (LC-QTOF-MS). Forensic Sci Int, 2012, 218: 68-81.
[34] D DeTata, P Collins, A M Kinley. A fast liquid chromatography quadrupole time-of-flight mass spectrometry (LC-QToF-MS) method for the identification of organic explosives and propellants. Forensic Sci Int, 2013, 233: 63-74.
[35] C Boix, M Ibáñez, J V Sancho, N León, V Yusá, F Hernández. Qualitative screening of 116 veterinary drugs in feed by liquid chromatography-high resolution mass spectrometry: Potential application to quantitative analysis. Food Chem, 2014, 160: 313-320.
[36] S Broecker, F Pragst, A Bakdash, S Herre, M Tsokos. Combined use of liquid chromatography–hybrid quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) and high performance liquid chromatography with photodiode array detector (HPLC-DAD) in systematic toxicological analysis. Forensic Sci Int, 2011, 212: 215-226.
[37] A E Mahjoub, C Staub. Simultaneous determination of benzodiazepines in whole blood or serum by HPLC/DAD with a semi-micro column. J Pharm Biomed Anal, 2000, 23: 447-458.
[38] P M Kabra, E U Nzekwe. Liquid chromatographic analysis of clonazepam in human serum with solid-phase (Bond-Elut®) extraction. J Chromatogr B: Biomed Appl, 1985, 341: 383-390.
[39] K B Borges, E F Freire, I Martins, M E P B (de) Siqueira. Simultaneous determination of multibenzodiazepines by HPLC/UV : Investigation of liquid-liquid and solid-phase extractions in human plasma. Talanda, 2009, 78: 233-241.
[40] N Badawi, K W Simonsen, A Steentoft, I M Bernhoft, K Linnet. Simultaneous screening and quantification of 29 drugs of abuse in oral fuild by solid-phase extraction and ultraperformance LC-MS/MS. Clin Chem, 2009, 55: 2004-2018.
[41] The fitness for purpose of analytical methods, a laboratory guide to method validation and related topics. EURACHEM guide, Teddington, UK, 1998.
[42] H Jin, A P Kumar, D H Paik, K C Ha, Y J Yoo, Y I Lee. Trace analysis of tetracycline antibiotics in human urine using UPLC-QToF mass spectrometry. Microchem J, 2010, 94: 139-147.
[43] M C Y Wong, W T K Lee, J S Y Wong, G Frost, J Lodge. An approach towards method development for untargeted urinary metabolite profiling in metabonomic research using UPLC-QToF/MS. J Chromatogr B: Biomed Appl, 2008, 871: 341-348.
[44] F Hernández, L Bijlsma, J V Sancho, R Díaz, M Ibáñez. Rapid wide-scope screening of drugs of abuse, prescription drugs with potential for abuse and their metabolites in influent and effluent urban wastewater by ultrahigh pressure liquid chromatography-quadrupole-time-of-flight-mass spectrometry. Anal Chim Acta, 2011, 684: 96-106.
[45] F Badoud, E Grata, L Perrenoud, M Saugy, S Rudaz, J L Veuthey. Fast analysis of doping agents in urine by ultra-high-pressure liquid chromatography–quadrupole time-of-flight mass spectrometry. II: Confirmatory analysis. J Chromatogr A, 2010, 1217: 4109-4119. ■

ISSN 2157-118X

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1 0

Cover Designed by Mirror

FORENSIC SCIENCE SEMINAR



ZolCat
www.zolcat.com

ISSN 2157-118X

9 772157 118006 >

Forensic Science Seminar

Volume 4 Number 1 December 2014
Single issue price $150USD (4.72%)