

Analysis of the Ex-vivo transformation of Semen, Saliva and Urine as they dry out using ATR-FTIR spectroscopy and chemometric Approach

Tanurup Das

Dr. A.P.J. Abdul Kalam Institute of Forensic Science & Criminology, Bundelkhand University, Jhansi

Abhimanyu Harshey

Dr. A.P.J. Abdul Kalam Institute of Forensic Science & Criminology, Bundelkhand University, Jhansi

Ankit Srivastava (✉ ankit_forensic81@rediffmail.com)

Dr. A.P.J. Abdul Kalam Institute of Forensic Science & Criminology, Bundelkhand University, Jhansi

Kriti Nigam

Dr. A.P.J. Abdul Kalam Institute of Forensic Science & Criminology, Bundelkhand University, Jhansi

Vijay Yadav

Dr. A.P.J. Abdul Kalam Institute of Forensic Science & Criminology, Bundelkhand University, Jhansi

Kapil Sharma

Directorate of Forensics Services, Junga, Shimla, Himachal Pradesh

Arun Sharma

Directorate of Forensics Services, Junga, Shimla, Himachal Pradesh

Research Article

Keywords: Body fluid, ATR-FTIR, Ex-vivo changes, Coefficient of determination, Chemometric, TSD

Posted Date: January 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-143925/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The ex-vivo biochemical changes of different body fluids are potential marker for the estimation of TSD. Infrared spectroscopy has great potential to reveal the biochemical changes in these fluids as previously reported by several researchers. Present study is focused to analyze the spectral changes in the ATR-FTIR spectra of three body fluids, commonly encountered in violent crimes i.e., Semen, Saliva and Urine as they dry out. The whole analytical timeline is divided into relatively slow phase I due to major contribution of water and faster Phase II due significant evaporation of water. Two spectral regions i.e., 3200-3400 cm^{-1} and 1600-1000 cm^{-1} are the major contributor to the spectra of these fluids. Several peaks in the spectral region between 1600-1000 cm^{-1} showed highly significant regression equation with higher coefficient of determination values in the Phase II in contrary to the slow passing Phase I. Principal component and Partial Least Square Regression analysis are the two chemometric tool used to estimate the TSD of the aforesaid fluids as they dry out. This study potentially estimates the TSD of an offence at the early stages after its occurrence as well as work as the precursor for further studies on extended timeframe.

Introduction

Body fluids are recurrently confronted as major evidence in violent crimes.¹ The potential application of different body fluids typically ranging from their identification to the successful extraction and DNA profiling.² All of these fluids experience instant biochemical change as they come out from the body. Some of these changes are rapid and remaining are gradual³, but both the changes are significant to reveal one of the most important aspects of forensic examination; 'time since the occurrence' of a crime.⁴ The estimation of 'time since deposition' (TSD) of a body fluid at the crime scene potentially solve the problem of 'situating the time of an offence'.³ The TSD of body fluids additionally solve a wide range of issues regarding the crime scene investigation.^{3,4} Most of the TSD studies are based on bloodstains²⁻⁴ except for a solitary publication on semen aging very recently.⁵ All these studies except one⁶ investigated the TSD for an extended period, escaping the initial changes in the fluid. With the inception of the TSD estimation study in the early twentieth century, various researchers have investigated this phenomenon on bloodstains.^{3,4} Since then the techniques used for TSD estimation went through a drastic evolutionary transformation. The analytical technology gradually transformed into nondestructive from destructive methods.^{3,4,9} A variety of techniques including Gas chromatography,^{10,11} Liquid chromatography,¹²⁻¹⁴ Oxygen electrode,^{3,15} microRNA based assay,^{3,16,17} Color transformation chart,¹⁸ Electron paramagnetic resonance,¹⁹⁻²¹ Reflectance spectroscopy in Ultraviolet, Visible and Infrared (IR) region,²²⁻²⁴ smartphone imaging,²⁵ IR absorption,^{1,3,4,5,6,26,27} and Raman scattering^{4,28} is explored around the world for the TSD estimation of blood and semen.⁵ As an age prediction tool, most of these methods realistically depict promising results. Yet, few methodological limitations hinder the universal acceptability of one or more of these techniques as the stand-alone tool for TSD studies except IR absorption and Raman scattering collectively known as vibrational spectroscopy.^{4,9,10} Leading researchers in the forensic sciences community around the world significantly attracted towards the vibrational spectroscopic technique due

to its several advantages over others. It reliably provides a rapid nondestructive analysis of evidence in forensic investigation.⁹ The last decade showed significant growth in the TSD estimation studies by using vibrational spectroscopic techniques.⁴

Despite the equal importance of semen, saliva, and urine as evidence in crime scene investigation, blood always received special attention in the research community.^{3,4,9,29,30} All the three aforesaid fluids including blood are consistently found in the broad range of homicides, suicide, rape, and other sexual assaults.^{2,31,32} Although, it is not very frequent that an investigating team can reach the crime scene as early as the fluid is still under drying process, but in several instances, investigators may come across a situation where the fluid is still drying. The earliest time after an offence has an enormous importance as this is the time when the most of the information in a crime scene remain unchanged and also during this timeframe the most significant information can be obtained from postmortem. Additionally, the empirical relation between the initial spectral change and TSD can help an investigator to differentiate between two or more overlapped older and fresh stains.⁶

Zhang et al.⁶ explored this specific and significant dimension of body fluid aging in 2016 and established a potential age estimation framework from the initial changes in the ATR-FTIR spectra of bloodstain collected from Sprague Dawley rats up to its drying under different conditions (temperature, humidity, and concentration) as the drying of fluid is affected by several factors like the component of the fluid, environmental conditions, and the supporting substrate.^{7,8} The present study is carefully focused to analyze the initial infrared spectral changes of semen, saliva, and urine as they dry out and estimate the TSD on the basis of the absorption changes in specific peaks. Although Zha et al.⁵ experimented with the change of semen stains up to one week, the initial change of this body fluid was not investigated. Hence, semen is also included in the present study. Attenuated total reflectance (ATR) Fourier transform infrared (FTIR) spectroscopy is combined with chemometric approach to investigate the spectral changes and estimate the TSD respectively.

Results And Discussion

Spectral change over the drying period

The study was based on three relevant body fluids i.e., semen, saliva, and urine except blood as a similar study is already reported on blood⁶. In this study, the major change is observed in two regions of the spectra i.e., 3300-2800 cm^{-1} (lipids are the major contributors) and 1700-950 cm^{-1} (proteins, nucleic acids and carbohydrates are the major contributors)^{33,34}. The list of peaks identified in the spectra of all three fluids is summarized in table 1 with their individual vibrational assignments. Although several peaks were identified in the spectra of each fluid during the study, only a few (age-linked peaks) showed linear changes in their absorption intensity with time. At the initial stage, a similar phenomenon has been observed in the drying of every body fluid. For the first several minutes, only two strong absorption peaks were visible (figure 1.a to 1.f). After a certain amount of time multiple significant peaks corresponding to

the biochemical profile of the fluids revealed throughout the fingerprint region of the IR spectra (figure 2.a to 2.f). Following the trend in each fluid, the whole drying time was divided into two phases.

Table 1: Strong, medium and weak peaks observed in all the three body fluids during their ATR-FTIR spectral analysis with their vibrational modes and spectral assignments.

Body fluid	Wavenumber (Cm ⁻¹)	Spectral assignment
Semen	3270.84	Symmetric N-H stretching (Amide A) ^{34,35,39}
	1636.97	C=O stretching (Amide I) ^{34,35,39}
	2967.98	Asymmetric CH ₃ stretching ^{34,35,39}
	1546.31	N-H bending and C-N stretching (Amide II) ^{34,35,39,40}
	1446	C-H bending of CH ₂ and CH ₃ ^{27,34,40}
	1395.85	C=O stretching in COO ⁻ (Fatty acids and polysaccharides) ³⁵
	1316.76 and 1243.46	N-H bending and C-N stretching (Amide III) ^{35,40}
	1065.99 and 969.541	CH ₂ OH groups, C-O stretching and COH groups, symmetric Glycosylated proteins, PO ₂ ⁻ stretching (Glycosylated proteins) ^{34,35,39,40}
Saliva	3272.77 and 3286.27	Symmetric N-H stretching (Amide A) ^{34,35,39}
	3076.01	Amide B ³⁵
	2958.34	Asymmetric CH ₃ stretching ^{34,35,39}
	2881.18	Asymmetric CH ₂ (methylene) stretching ⁴²
	1636.97 and 1644.69	C=O stretching (Amide I) ^{34,35,42}
	1546.31	N-H bending and C-N stretching (Amide II) ^{34,35,40-42}
	1447.93	C-H bending of CH ₂ and CH ₃ ^{27,35}
	1403.57	Symmetric CH ₃ ³⁹
	1316.76 and 1243.46	N-H bending and C-N stretching (Amide III) ^{35,40}
	1077.56 and 1042.84	CH ₂ OH groups, C-O stretching and COH groups, symmetric Glycosylated proteins, PO ₂ ⁻ stretching (Glycosylated proteins) ^{34,35,39,40}
Urine	3346.07	Asymmetric stretching of H-O-H ^{35,39}
	3272.77	Symmetric N-H stretching (Amide A) ^{34,35,39}
	3205.25	O-H stretching ^{26,34,35}
	1658.19; 1636.97 and 1623.47	C=O stretching (Amide I) ^{34,35,42}
	1455.65	Asymmetric C-N stretching (Urea) ^{38,39,43}
	1156.65	NH ₂ deformation (Urea) ^{38,39,43}
	1081.42	CH ₂ OH groups, C-O stretching and COH groups, symmetric Glycosylated proteins, PO ₂ ⁻ stretching (Glycosylated proteins) ^{34,35,39,40}

Note*: '**Bold**' marks denote the peaks changed during the transition from phase I to phase II and '**Bold**'+ '*Italic*' mark denotes the age-linked peaks.

The dilution of the fluids showed an extended (2-4 minutes) phase I due to the excess amount of water in the diluted sample. The longest phase I observed in the Semen samples and shortest in the saliva samples. Phase II was relatively similar in the spectra of both raw and diluted samples. The duration of phase II of three body fluids was relatively the same (10-12 minutes). Table 2 demonstrates the minimum,

maximum, and mean values of both phases. The difference in the drying time of all three fluids is potentially a result of the qualitative and quantitative variability in their biochemical components.

Table 2: Summarized the minimum, maximum and mean values of two phases of body fluid drying

Fluids	concentration	Phase I (Minutes)			Phase II (Minutes)			Total (Minutes)		
		Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
Semen	100%	28	22	26	14	10	10	42	32	36
	40%	34	28	30	12	8	10	46	34	40
Saliva	100%	14	10	12	14	10	12	28	20	24
	40%	20	14	16	12	8	10	32	22	26
Urine	100%	24	18	20	12	6	10	36	24	30
	40%	26	18	22	14	8	12	40	26	34

Few researchers reported the correlation between the evaporation of distilled water and time.^{35,36} Although, their experimental parameters were different. Phase I included two strong peaks of amide A ($3270\text{-}3273\text{ cm}^{-1}$) and amide I (1636.97 cm^{-1}) is dominated by the water content of these body fluids (figure 1.a to 1.f). Except for similar height, the peak corresponding to amide A was broader than the peak due to amide I. Water is the major component of all the secreted fluids in the human body. The O-H stretching of water masks the absorption of IR radiation by other minor but significant components of the fluid.⁶ Except for urine, the absorption intensity of these two peaks showed insignificant change throughout phase I in the spectra of semen and saliva (figure 1.a to 1.f and figure 3). In the contrary, Zhang et al.⁶ found a different result for blood as the peak at 3308 cm^{-1} showed very weak but linear absorption change during the early stage. Only the spectra of urine (100% and 40%), showed analogous results with the study by Zhang et al.⁶ as the peak at 3272.77 cm^{-1} showed a significant decline in the mean absorbance with time during the phase I (figure 2.c, 2.f and 3.c).

Phase II is the fast declination stage where a significant amount of water evaporates rapidly and reveal the other peaks and their intensity changes with time in each body fluid. The peaks at 3270.84 and 1636.97 cm^{-1} showed no shift during the whole drying (phases I and II) process of semen stain but the former one sharpens with time and rapidly declined during the phase II (figure 1.a, 1.d, 2.a and 2.d). In the phase II drying of semen droplets, one strong (1546.31 cm^{-1} : Amide II), two medium (1446 cm^{-1} : methylene; CH_2 and CH_3 and 1065.99 cm^{-1} : Glycosylated proteins: probably prostate-specific antigen) and three weak (2967.98 cm^{-1} : CH_3 stretching, 1395.85 cm^{-1} : fatty acids and polysaccharides and 1243.46 cm^{-1} : Amide III) significant age-linked peaks were observed (figure 2.a and 2.d). Zha et al.⁵ investigated the changes in few similar peaks at marginally different positions i.e., 1539 cm^{-1} (Amide II), 1448 cm^{-1} (Methylene: CH_2 and CH_3), 1392 cm^{-1} (Fatty acids and polysaccharides), 1059 cm^{-1} (Prostate-specific antigen). Few more researchers reported the IR spectra of semen in several body fluid identification research articles.³⁷⁻³⁹ Phase II spectra of saliva showed the shift of strong peaks at 3272.77 cm^{-1} (amide A) to 3286.27 cm^{-1} (amide A) and 1636.97 cm^{-1} (amide I) to 1644.69 cm^{-1} (amide I) that

indicated the initiation of this phase (figure 2.b and 2.e). The shifted peak of amide A sharpens following the trend of semen samples. Among others, one strong and sharp (1546.31 cm^{-1} : amide II), four weak (1447.93 and 1403.07 cm^{-1} : Methylene, 1077.56 , and 1042.84 cm^{-1} : glycosylated proteins) significant age-linked peaks were found (figure 2.b and 2.e). Including the peak corresponding to amide II, two peaks at 1447.93 cm^{-1} and 1077.56 cm^{-1} are similar to the peaks at 1446 and 1065.99 cm^{-1} in the spectra of semen and placed at marginally different positions. But the intensity of the peak corresponding to glycosylated protein is relatively weaker in the spectra of saliva. The peaks in phase I spectra of urine bifurcated in phase II. The peak at 3272.77 cm^{-1} divided into 3346.07 and 3205.25 cm^{-1} and 1636.97 cm^{-1} divided into 1623.47 and 1658.19 cm^{-1} (figure 2.c and 2.f). One strong (1658.19 cm^{-1} : amide I), one medium (3346.07 cm^{-1} : H-O-H stretching), and two weak (1156.65 cm^{-1} : urea and 1081.42 cm^{-1} : Glycosylated proteins) peaks were observed in the phase II spectra of urine samples (figure 2.c and 2.f) that significantly changes during the drying process. The peak at 1081.42 cm^{-1} is similar to the peaks at 1065.99 and 1077.56 cm^{-1} of semen and saliva, respectively. Due to the presence of a significant quantity of prostate-specific antigen in semen, the peak corresponding to glycosylated protein is stronger in its spectra than saliva and urine.³³ In several previous literatures on the IR signature of saliva and urine, the above-mentioned peaks were reported by researchers.^{37,38,40-42} Amide A, I, II, and glycosylated proteins are the common biochemical components found in all the three body fluids. Elkins, Orphanou, and Takamura et al.^{35,39,40} previously reported the presence of the common biochemicals in all these body fluids on their article of body fluid identification by ATR-FTIR. In the Phase II, all the age-linked peaks of each fluid showed a linear relationship between the mean absorbance at each time point and TSD (figure 2.a to 2.f). The calculated

Statistical results

Coefficient of determination (R^2) measures that whether a regression line can explain most of the variance of the representing data. Hence, higher R^2 value for a particular peak indicates it as a good predictor. R^2 values were calculated for significantly transforming peaks of semen, saliva and urine and most of them except the peak at 3346.07 cm^{-1} corresponding to H-O-H stretching (R^2 value: 0.8863), have the R^2 value above 0.9 (max: 0.9994; min: 0.9016) (Table 3a. to 3.c). The ' p ' values of all these correlation equations are <0.05 , which is also statistically significant. The mean absorbance of the selected peaks at each time point increased with time for both raw and diluted fluid samples. Only the peak at 3346.07 cm^{-1} in the spectra of urine decreased with time, potentially due to the faster evaporation of water.

Table 3: Correlation equation of changes in the absorption intensity and TSD with R^2 and ' p ' values of age-linked peaks in the phase II drying of raw and diluted a. semen. b. saliva and c. urine.

Table 3. a						
Peak	Semen (100%)			Semen (40%)		
	Correlation equation	Coefficient of determination (R^2)	Significance (p)	Correlation equation	Coefficient of determination (R^2)	Significance (p)
2967.98	$y = 0.017x - 0.321$	0.9569	<0.05	$y = 0.0036x + 0.1727$	0.9383	<0.05
1546.31	$y = 0.0764x - 1.7746$	0.9522	<0.05	$y = 0.0207x + 0.0742$	0.9016	<0.05
1446	$y = 0.0385x - 0.899$	0.9751	<0.05	$y = 0.0071x + 0.1173$	0.9085	<0.05
1395.85	$y = 0.0596x - 1.4397$	0.9724	<0.05	$y = 0.0152x - 0.0399$	0.938	<0.05
1243.46	$y = 0.0378x - 0.8978$	0.9823	<0.05	$y = 0.0058x + 0.1105$	0.9788	<0.05
1065.99	$y = 0.0153x - 0.0763$	0.963	<0.05	$y = 0.0551x - 1.3207$	0.9767	<0.05
Table 3. b						
Peak	Saliva (100%)			Saliva (40%)		
	Correlation equation	Coefficient of determination (R^2)	Significance (p)	Correlation equation	Coefficient of determination (R^2)	Significance (p)
1546.31	$y = 0.0017x + 0.5852$	0.9775	<0.05	$y = 0.0124x + 0.4735$	0.9919	<0.05
1447.93	$y = 0.002x + 0.3325$	0.9618	<0.05	$y = 0.0124x + 0.2091$	0.9975	<0.05
1403.07	$y = -0.0025x + 0.284$	0.9898	<0.05	$y = 0.0119x + 0.1669$	0.9981	<0.05
1077.56	$y = 0.0021x + 0.3391$	0.994	<0.05	$y = 0.0112x + 0.1958$	0.9963	<0.05
1042.84	$y = 0.0025x + 0.2755$	0.9869	<0.05	$y = 0.0096x + 0.1179$	0.9994	<0.05
Table 3. b						
Peak	Urine (100%)			Urine (40%)		
	Correlation equation	Coefficient of determination (R^2)	Significance (p)	Correlation equation	Coefficient of determination (R^2)	Significance (p)
3272.77	$y = -0.0029x + 0.6719$	0.9089	<0.05	$y = -0.0005x + 0.6785$	0.9179	<0.05
3346.07	$y = -0.0108x + 0.7233$	0.8863	<0.05	$y = -0.0138x + 1.139$	0.871	<0.05
1658.19	$y = 0.0034x + 0.8304$	0.9269	<0.05	$y = 0.0017x + 0.8538$	0.9227	<0.05
1156.65	$y = 0.0028x + 0.1758$	0.9323	<0.05	$y = 0.0016x + 0.2008$	0.95	<0.05
1081.42	$y = 0.003x + 0.1597$	0.9223	<0.05	$y = 0.0011x + 0.2161$	0.924	<0.05

Note*: '**Bold**' + '*Italic*' marked peak belongs to the phase I with a significant R^2 and ' p ' values.

Estimation of TSD in Phase II of drying

Body fluids show significant ex-vivo degradation when exposed for a long period. While, early changes are very limited as only a few relevant peaks for aging are visible. Hence, the TSD estimation was performed only on the age-linked peaks of three body fluids. Both Principal Component Regression (PCR) Analysis and Partial Least Square Regression (PLSR) are strong chemometric tool for the estimation of TSD of body fluids as reported in previous studies.^{1,4,5,10,29} The calculated R^2 values for the calibration and prediction of both models are more than 0.9. While the root mean square error of cross validation (RMSECV) and prediction (RMSEP) values in both the models for all the three fluids except diluted urine, showed low values (Table 4.a and 4.b). Although the initial changes in the IR spectra of body fluids are relatively less distinguishable in comparison to the samples exposed for longer period, the High R^2 and low RMSE values indicate a good prediction of TSD during this timeframe. In diluted urine samples, the RMSECV (PCRA: 0.77; PLSR: 0.76) and RMSEP (PCRA:0.74; PLSR: 0.73) (Table 4.a and 4.b) in both the models are relatively higher, that potentially interfere in accurate age estimation. While, diluted semen also showed higher RMSECV (0.72) and RMSEP (0.68) in PCRA regression model, while PLSR predicts

the age for the same condition with significantly higher accuracy with RMSECV and RMSEP values of 0.19 and 0.15, (Table 4.a and 4.b) respectively. Comparatively, PLSR showed better efficiency of prediction than PCRA as for every fluid it records relatively higher R^2 and lower RMSE values than in PCRA. Figure 4 and 5 depicts the PCR and PLS plots of actual vs predicted regression lines for age linked peaks of three body fluids.

Table 4: a. Validation results of PCR models for multiple age-linked peaks of three body fluids b. Validation results of PLSR models for multiple age-linked peaks of three body fluids

4.a Principal Component Regression (PCR) Analysis Table				
Fluids	Cross-validation		External Prediction	
	R²	RMSECV	R²	RMSEP
Semen (100%)	0.9934612	0.2345791	0.9939425	0.2201363
Semen (40%)	0.9389567	0.7167367	0.9415222	0.6839755
Saliva (100%)	0.9884937	0.374184	0.9890894	0.3567773
Saliva (40%)	0.9973184	0.1517414	0.9973434	0.1457813
Urine (100%)	0.9948414	0.2083559	0.9950851	0.1982911
Urine (40%)	0.9517956	0.7658792	0.9532829	0.7382633
4.b Partial Least Square Regression (PLSR) Analysis Table				
Semen (100%)	0.9935297	0.2333477	0.9940089	0.2189272
Semen (40%)	0.9957254	0.1896652	0.9970123	0.1546034
Saliva (100%)	0.9889287	0.3736748	0.9891293	0.3561241
Saliva (40%)	0.9968555	0.1617815	0.9973462	0.1457073
Urine (100%)	0.9949841	0.205454	0.995233	0.1952845
Urine (40%)	0.9522571	0.7622038	0.9539837	0.7327054

Conclusions

Ex-vivo changes in the biochemical components of human body fluids are potentially the best estimator of time since an offence occurred. Up to the present scenario, most of the age estimation studies on body fluids are based on bloodstains and a solitary study reported on semen, but those are insufficient for the above cause. In the present study three forensically significant body fluids other than blood i.e., semen, saliva, and urine were considered to explore the instantaneous changes in their ATR-FTIR spectra up to their drying. This study revealed that all the body fluids undergo significant water loss at the initial stage of their ex-vivo degradation. This rapid loss of water significantly divided the drying process into two phases as the first phase consists of slow evaporation with minimal spectral change and second phase depicts relatively faster evaporation. These two phases can be clearly distinguishable from the ATR-FTIR spectra of each fluids. This study also revealed the spectral regions of interest for the TSD estimation of these fluids as saliva and urine are not explored previously for this purpose. This study also reveals that, if a body fluid accidentally diluted during the deposition, the accuracy of the estimation of

TSD can be significantly alter. Hence, further experiments with different concentration of these fluids should be performed. Despite, relatively low changes in the IR absorption, chemometric tools like; PCR and PLSR successfully estimate the TSD for each fluid during the initial spectral changes. Hence, it is potentially used for saliva and urine based TSD estimation studies with longer timeframe.

Methods

Sampling

Samples of saliva, semen, and urine were collected from randomly selected eight healthy volunteers (28-40 years). All the donors were informed about the nature and procedure of the work. Each body fluid from a donor was collected separately in a glass test tube(s) just prior to the spectral analysis to avoid any loss of time after its release from the human body. All the fluids were collected by voluntary secretion, ejaculation, and excretion process without using any invasive technique. Spectra of fluids were obtained in two different concentration i.e., 100% and 40% (2:3) to investigate the effect of water content on the drying time. The dilution concentration was randomly selected for this study to investigate the difference in drying time between concentrated and dilute fluid samples. 40% solutions of all the fluids were prepared by mixing 2 parts of the fluid and 3 parts of distilled water in a test tube. For experimental purpose saliva and urine were repeatedly taken for three consecutive days and semen were collected three times with an interval of three days to observe any intra-variations of their composition.

Ethical Approval

Written informed consent was obtained from all the participants. The study was conducted under the ethical guidelines and it was approved by the Institute Ethics Committee (Human Studies) of Maharani Laxmi Bai Medical College, Jhansi, vide letter no. 4647 dated 05.09.2020.

Collection and pre-treatment of FTIR spectra

IR spectra of all the body fluid samples were collected by a 'Spectrum Two' FTIR spectrophotometer manufactured by Perkin Elmer corporation, equipped with a 2 mm diameter diamond crystal ATR accessory and spectrum software (Version 10.0). The spectrum software used for the collection of spectra. The crystal face was cleaned with a 70% methanol solution before drop the fluid sample for each spectral measurement. One drop (approximately 50 μL) of each body fluid was separately added onto the crystal from a constant height of 4 cm. Throughout the whole study, the volumes of fluids were kept relatively constant. The experiment was performed in the month of January. The approximate temperature and relative humidity during the complete spectral collection varied between 13-17⁰C and 56-75%, respectively. As the ambient conditions during the study did not vary significantly, the effect of variable temperature and humidity were not considered in this study. The spectra of each fluid were collected with an interval of two minutes from immediately after placing the droplet on the crystal within the range of 4000 cm^{-1} to 500 cm^{-1} with 32 scans and resolution of 4/cm. In this study, the term 'drying out' is the time point at which there was a negligible change between the three consecutive scans taken

at a 2-minute interval.²⁶ At this point, the droplet transformed into a dried stain. Before analyzing the obtained data, all the spectra of body fluids were preprocessed by using Unscrambler X software with several spectral corrections i.e., baseline offset, spectral smoothing with Savitzky-Golay algorithm including 13 smoothing points and 3 polynomial orders in a symmetric kernel and range normalization [8,34].

Statistical and chemometric analysis

The mean values of the absorbance of all the significant peaks and the drying time were calculated. Peak identification, fitting, and statistical analysis were carried out by using Origin Pro 2016 and Microsoft excel 2019 software. The correlation coefficient (R) and coefficient of determination (R^2) between the TSD and changes in the absorbance values for each body fluid was established with a fitting correlation equation. All the equations showed a ' p ' value of <0.05 which is statistically significant.

PCR and PLSR are the two most frequently used tools for prediction model creation and estimation studies. Both of these tools decompose the multiple X variables with respect to the values of Y variables and generate single values for each sample and establish correlation between X and Y.¹ In the present study, the time has been considered as Y and the different age-linked peaks with absorbance values are considered as X. All the chemometric analyses were performed by using 'Unscrambler X' (CAMO Analytics) software. RMSECV and RMSEP were calculated to check the consistency and predictive ability of the regression model. Higher R^2 values and lower RMSE values are the indicative of good prediction model. Full cross validation was performed for each fluid. Eight samples randomly for each fluid were selected for the model creation and two randomly left out samples were applied in external validation purpose.

Declarations

Acknowledgement

The authors are thankful to all the volunteers for providing body fluid samples and The Narcotic Division, Directorate of Forensic Science, Himachal Pradesh, for providing analytical facilities and technical support.

Author contributions

T.D., A.H. and A.S.* (Corresponding Author) conceived the idea of this experiment. T.D. collected the samples. K.S. and A.S provided the instrumental and technical support regarding the analysis of the samples. T.D., A.H. and A.S.* (Corresponding Author), K.N. and V.K.Y. performed the statistical analysis and interpretation of the data. T.D. wrote the draft of the manuscript. A.H., A.S.* (Corresponding Author), K.N. and V.K.Y reviewed, corrected and approved the final version of the manuscript.

Additional Information

Competing of interests

The authors declare that they have no competing interests.

Data Availability

All data generated or analysed during this study are available from the corresponding author on reasonable request.

References

1. Lin, H., Zhang, Y., Wang, Q., Li, B., Huang, P. & Wang, Z. Estimation of the age of human bloodstains under the simulated indoor and outdoor crime scene conditions by ATR-FTIR spectroscopy. *Scientific Reports* **7**:13254 (2017).
2. Zapata, F. Emerging spectrometric techniques for the forensic analysis of body fluids. *Trends in Analytical Chemistry* **64**, 53-63 (2015).
3. Bremmer, R.H., De Bruin, K.G., Van Gemert, M.J.C., Van Leeuwen, T.G. & Alders, M.C.G. Forensic Quest for Age Determination of Bloodstains. *Forensic Science International* **216** (2012).
4. Zadora, G. & Menzyk, A. In the pursuit of the holy grail of forensic science – Spectroscopic studies on the estimation of time since deposition of bloodstains. *Trends in Analytical Chemistry* **105**, 137-165 (2018).
5. Zha, S., Wei, X., Fang, R., Wang, Q., Lin, H., Zhang, K., Zhang, H., Liu, R., Li, Z., Huang, P. & Wang, Z. Estimation of the age of human semen stains by attenuated total reflection Fourier transform infrared spectroscopy: a preliminary study. *Forensic Science Research* (2019).
6. Zhang, Y., Wang, Q., Li, B., Wang, Z., Li, C., Yao, Y., Huang, P. & Wang, Z. Changes in attenuated total reflection Fourier transform infrared spectra as blood dries out. *Journal of Forensic Sciences* (2016).
7. Tarafdar, S., Tarasevich, Y.Y., Dutta Choudhury, M., Dutta, T. & Zang, D. Droplet drying patterns on solid substrates: from hydrophilic to superhydrophobic contact to levitating drops. *Advances in Condensed Matter Physics* (2018).
8. Gautam, R., Vanga, S., Ariese, F. & Umapathy, S. Review of multidimensional data processing approaches for Raman and Infrared spectroscopy. *EPJ Tech Instrum* **2**, 8-38 (2015).
9. Muro, C.K., Doty, K.C., Bueno, J., Halamkova L. & Lednev, I. K. Vibrational spectroscopy: recent developments to revolutionize forensic science. *Analytical chemistry* **87**, 306-327 (2015).
10. Das, T., Harshey, A., Nigam, K., Yadav, V.K., Srivastava, A. K. Analytical approaches for bloodstain aging by vibrational spectroscopy: Current trends and future perspectives. *Microchemical Journal*. **158** (2020) 105278.
11. Arany, S. & Ohtani, S. Age estimation of bloodstains: A preliminary report based on aspartic acid racemization. *Forensic Science International* **212**, (1-3) (2011).

12. Nakao, K., Shimada, R., Hara K. & Kibayashi, K. Experimental study on age estimation of bloodstains based on biological and toxicological analysis. *The Open Forensic Science Journal* **6**, 6-11 (2013).
13. Inoue, H., Takabe, F., Iwasa, M. & Maeno, Y. Identification of fetal hemoglobin and simulated estimation of bloodstain age by high-performance liquid chromatography. *International Journal of Legal Medicine* **104**, 127-131 (1991).
14. Andrasko, J. The estimation of age of bloodstains by HPLC analysis. *Journal of Forensic Sciences* **42**(4), 601-607 (1997).
15. Seok, A. E., Lee, J., Lee, Y.-R., Lee, Y.-J., Kim, H.-J., Ihm, C., Sung, H.J., Sung, H.H. & Kang, H.-G. Estimation of Age of Bloodstains by Mass-spectrometry: A Metabolomic Approach. *Analytical Chemistry* **90**, 21, 12431-12441 (2018).
16. Matsuoka, T., Taguchi, T., Okuda, Estimation of bloodstain age by rapid determinations of oxyhemoglobin by use of oxygen electrode and total hemoglobin. *J. Biological & Pharmaceutical Bulletin* **18**(8), 1031-1035 (1995).
17. Bauer, M., Polzin, S. & Patzelt, D. Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: a possible indicator of the age of bloodstains. *Forensic Science International* **138**, 94– 103 (2003).
18. Anderson, S., Howard, B., Hobbs, G. R. & Bishop, C. P. A method for determining the age of bloodstain. *Forensic Science International* **148**, 37–45 (2005).
19. Tomellini, L. De l'emploi d'une table chromatique pour les taches du sang. *Arch. d'Antropologie criminelle de Criminol* **14**, (1907).
20. Peisach, J., Blumberg, W.E., Wittenberg, B.A., Wittenberg, J.B. & Kampa, L. Hemoglobin A: an electron paramagnetic resonance study of the effects of interchain contacts on the heme symmetry of high-spin and low-spin derivatives of ferric alpha chains. *Proc. Natl. Acad. Sci. U. S. A.* **63**, 934–939 (1969).
21. Svistunenko, D.A., Sharpe, M.A., Nicholls, P., Blenkinsop, C., Davies, N.A. & Dunne, J. The pH dependence of naturally occurring low-spin forms of methaemoglobin and metmyoglobin: an EPR study. *Biochemical Journal* **351**, 595–605 (2000).
22. Miki, T., Kai, A. & Ikeya, M. Electron spin resonance of bloodstains and its application to the estimation of time after bleeding. *Forensic Science International* **35**, 149–158 (1987).
23. Li, B., Beveridge, P., O'Hare, W. & Islam, M. The estimation of the age of a blood stain using reflectance spectroscopy with a microspectrophotometer, spectral pre-processing and linear discriminant analysis. *Forensic Science International* **212**(1-3), 198-204 (2011).
24. Bremmer, R. H., Nadort, A., Van Leeuwen, T.G., Van Gemert, M.J.C. & Aalders, M.C.G. Age estimation of blood stains by hemoglobin derivative determination using reflectance spectroscopy. *Forensic Science International* **206**, 166-171 (2011).
25. Sun, H., Meng, Y., Zhang, P., Li, Y., Li, N., Li C. & Guo, Z. Non-invasive prediction of bloodstain age using the principal component and a back propagation artificial neural network. *Laser Physics Letters* **14**(9) (2017).

26. Thanakiatkrai, P., Yaodam, A. & Kitpipit, T. Age estimation of bloodstains using smartphones and digital image analysis. *Forensic Science International* **233** (1-3), 288 (2013).
27. Lamontagne, M., Brown, C.W., Botenjc-Sehic, E. & Tsaparikos, M. Forensic Application of Near-Infrared Spectroscopy: Aging of Bloodstains. *Spectroscopy* **24**(2) (2009).
28. Lu, Z. Application of Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy for Forensic Analysis. University of South Carolina, Columbia (2015).
29. Menzyk, A., Damin, A., Martyna, A., Alladio, E., Vincenti, M., Marta, G. & Zadora, G. Toward a novel framework for bloodstains dating by Raman spectroscopy: How to avoid sample photodamage and subsampling errors. *Talanta* **209** (2020).
30. Li, R. *Forensic Biology*, CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton FL 33487-2742 (2011).
31. Gunn, A. *Essentials Forensic Biology*, Second edition, Wiley-Blackwell, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, UK, (2009).
32. Jackson, A.R.W. & Jackson, J.M. *Forensic Science*, Pearson Prentice Hall, Harlow, England 115- 136 (2008).
33. Li, R. *Forensic Chemistry Handbook*, Wiley, New Jersey 69-290 (2012).
34. Lee, L.C., Liong, C-Y, Jemain, A. A. [A contemporary review on Data Preprocessing \(DP\) practice strategy in ATR-FTIR spectrum](#). *Chemom Intell Lab Syst.* **163** (2017) 64-75.
35. Orphanou, C-M. The Detection and Discrimination of Human Body Fluids Using ATR FT-IR Spectroscopy. *Forensic Science International* **252** (2015).
36. Garidel, P. & Schott, H. Fourier-Transform Midinfrared Spectroscopy for Analysis and Screening of Liquid Protein Formulations. *Bio Process Technical* (2006).
37. Al-Shammiri, M. Evaporation rate as a function of water salinity. *Desalination* **150**, 189–203 (2002).
38. Jodat, A., Moghiman, M. & Rad, E.Y. An experimental study of the ability of similarity theory to predict water evaporation rate for different convection regimes. *Arab J Sci Eng* **38**(12), 3505–3513 (2013).
39. Takamura, A., Watanabe, K., Akutsu, T. & Ozawa, T. Soft and robust identification of body fluid using Fourier transform infrared spectroscopy and chemometric strategies for forensic analysis. *Scientific reports* (2018).
40. Elkins, K.M. Rapid Presumptive “Fingerprinting” of Body Fluids and Materials by ATR FT-IR Spectroscopy. *Journal of forensic sciences* **56**(6) (2011).
41. Barċot, O., Balarin, M., Gamulin, O., Jeżek, D., Romac, P. & Brbjas-Kraljević, J. Investigation of spermatozoa and seminal plasma by fourier transform infrared spectroscopy. *Applied Spectroscopy* **61**(3), 309–313 (2007).
42. Scott, D.A., Renaud, D.E., Krishnasamy, S., Meriç, P., Buduneli, N., Cetinkalp, S. & Liu, K-Z. Diabetes-related Molecular Signature in Infrared Spectra of Human Saliva. *Diabetology & Metabolic Syndrome* **2**(48), 1–9 (2010).

43. Khaustova, S., Shkurnikov, M., Tonevitsky, E., Artyushenko, V. & Tonevitskya, A. Noninvasive biochemical monitoring of physiological stress by Fourier transform infrared saliva spectroscopy. *Analyst* **135**, 3183–3192 (2010).

Figures

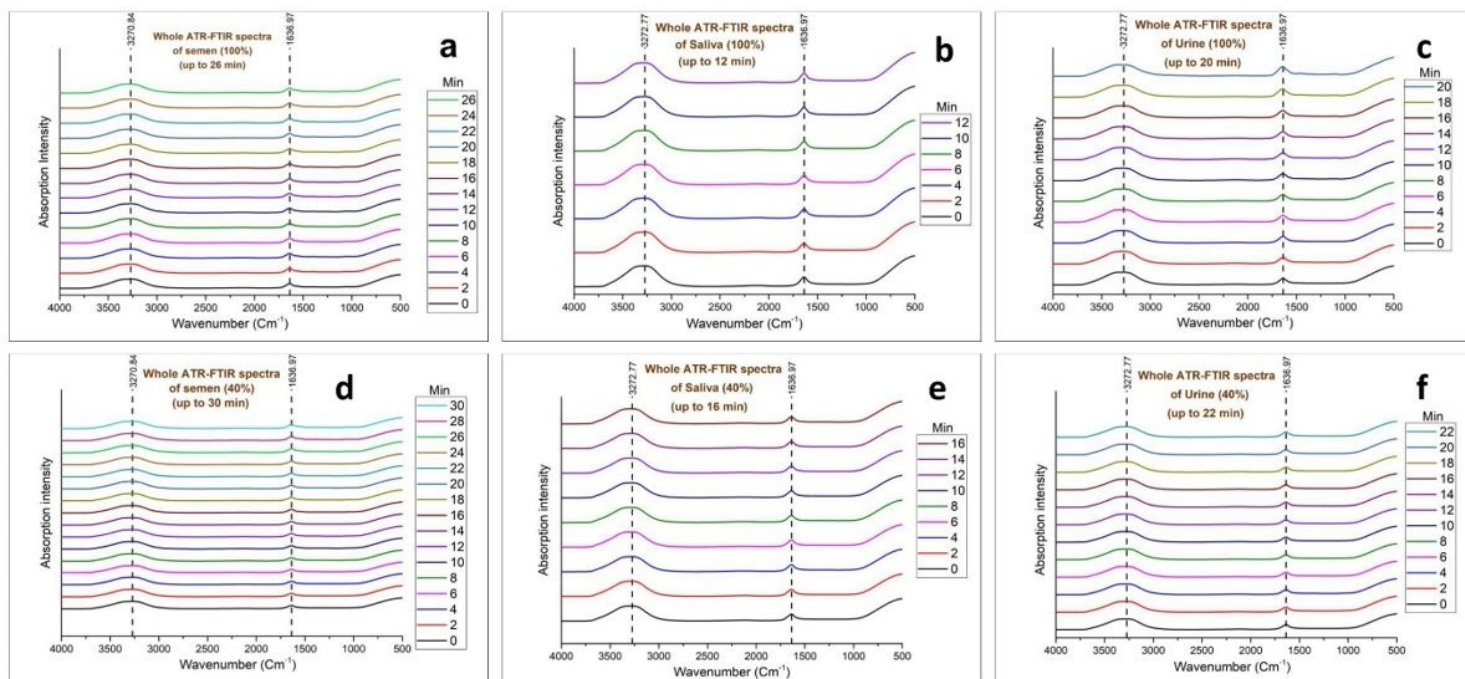


Figure 1

Whole ATR FTIR Phase I spectra for a. 100% semen, b. 100% saliva, c. 100% urine, d. 40% semen e. 40% saliva f. 40% urine

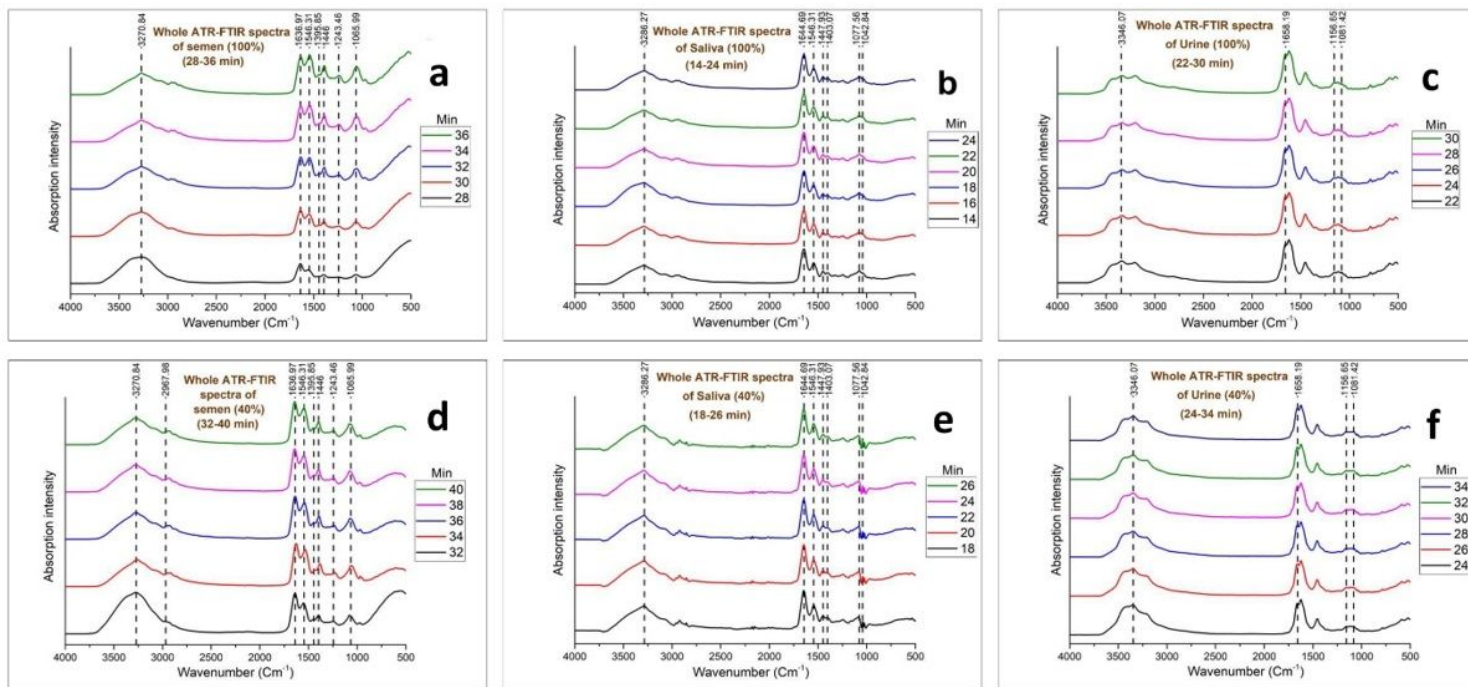


Figure 2

Whole ATR FTIR Phase II spectra for a. 100% semen, b. 100% saliva, c. 100% urine, d. 40% semen e. 40% saliva f. 40% urine

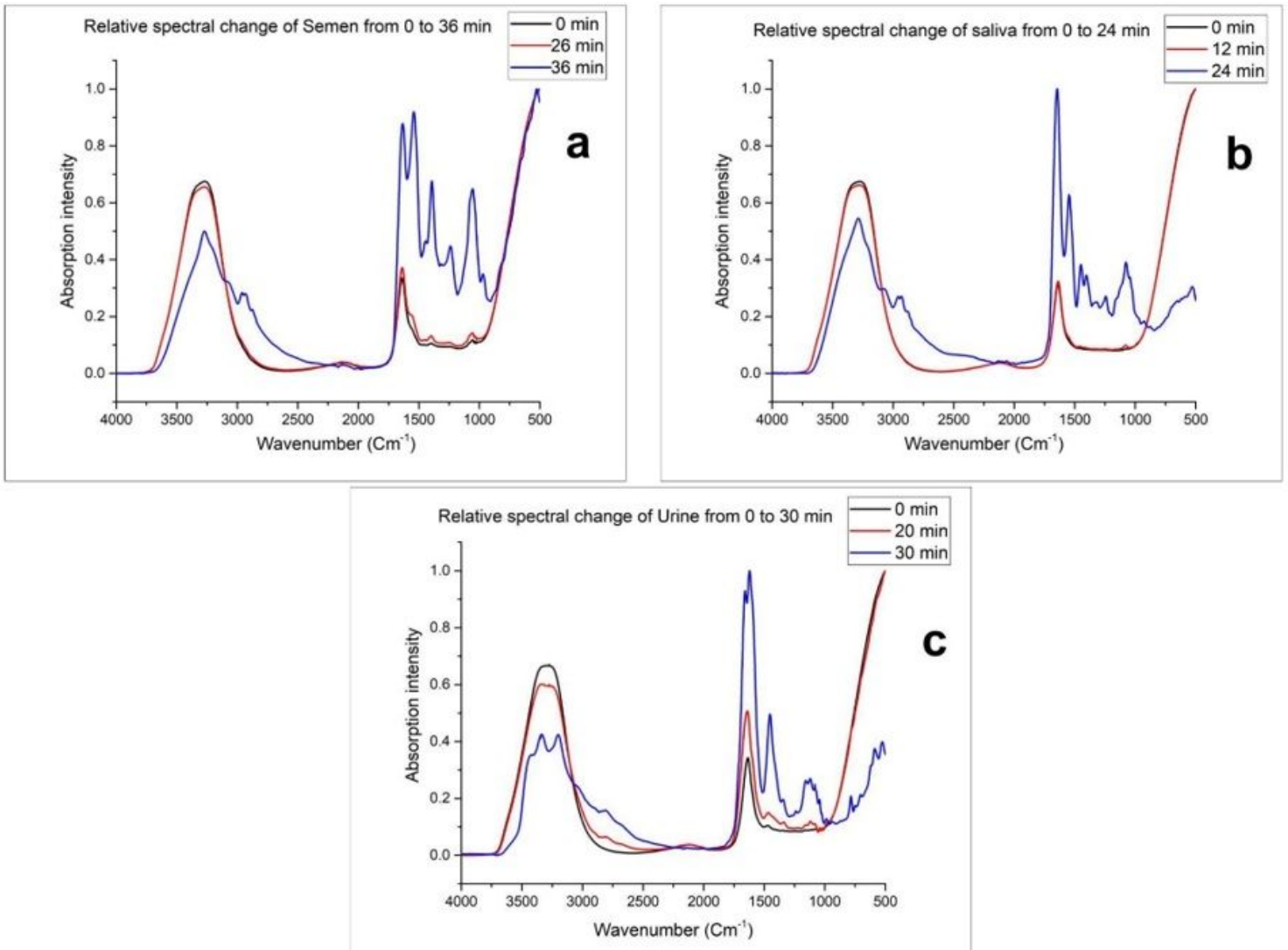


Figure 3

Relative changes in the ATR-FTIR spectra of a. semen, b. saliva and c. urine until drying

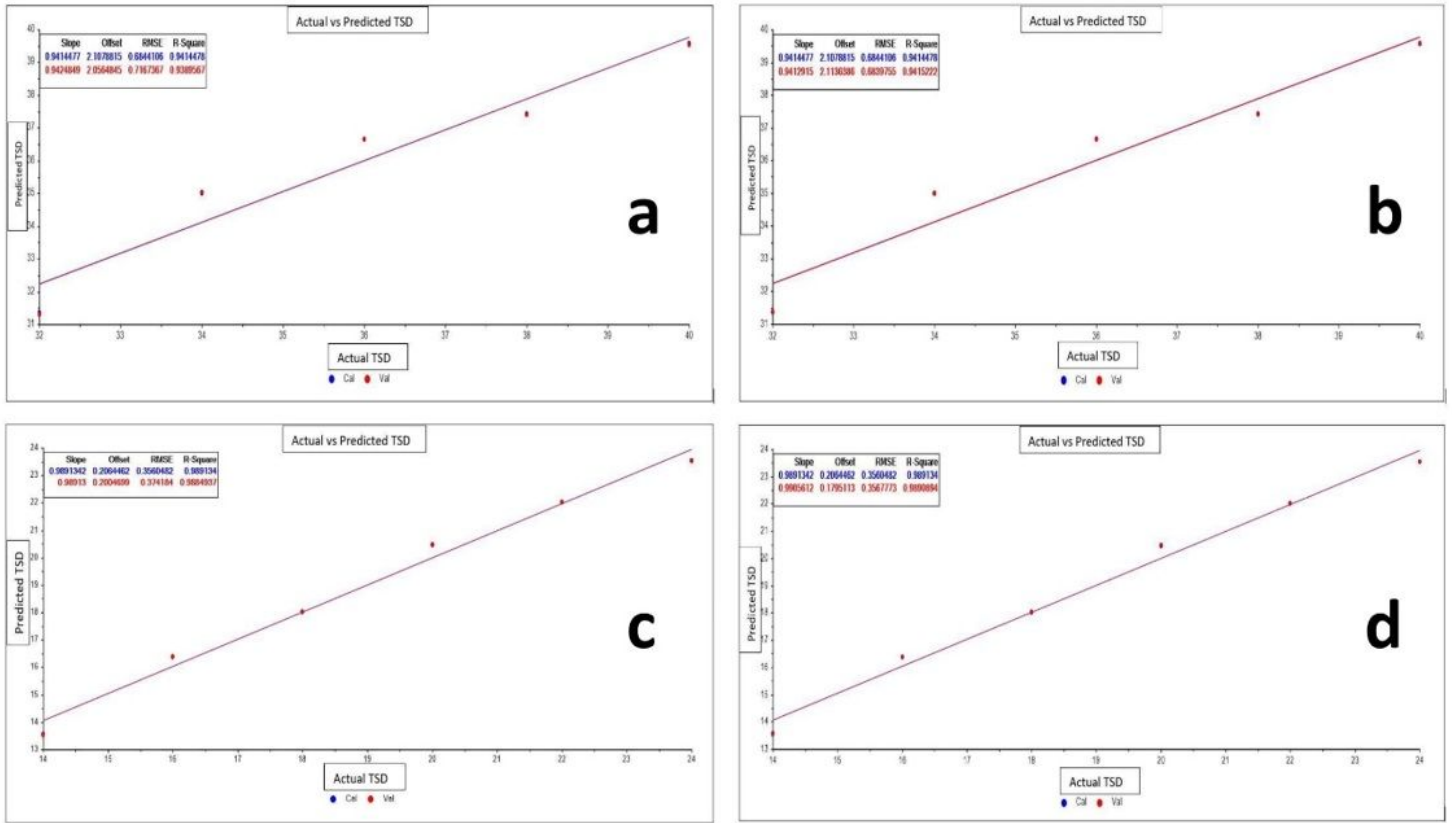


Figure 4

PCR plots of a. actual vs cross validated prediction of TSD for semen (40%) b. actual vs externally validated prediction of TSD for semen (40%) c. actual vs cross validated prediction of TSD for saliva (100%) d. actual vs externally validated prediction of TSD for saliva (100%)

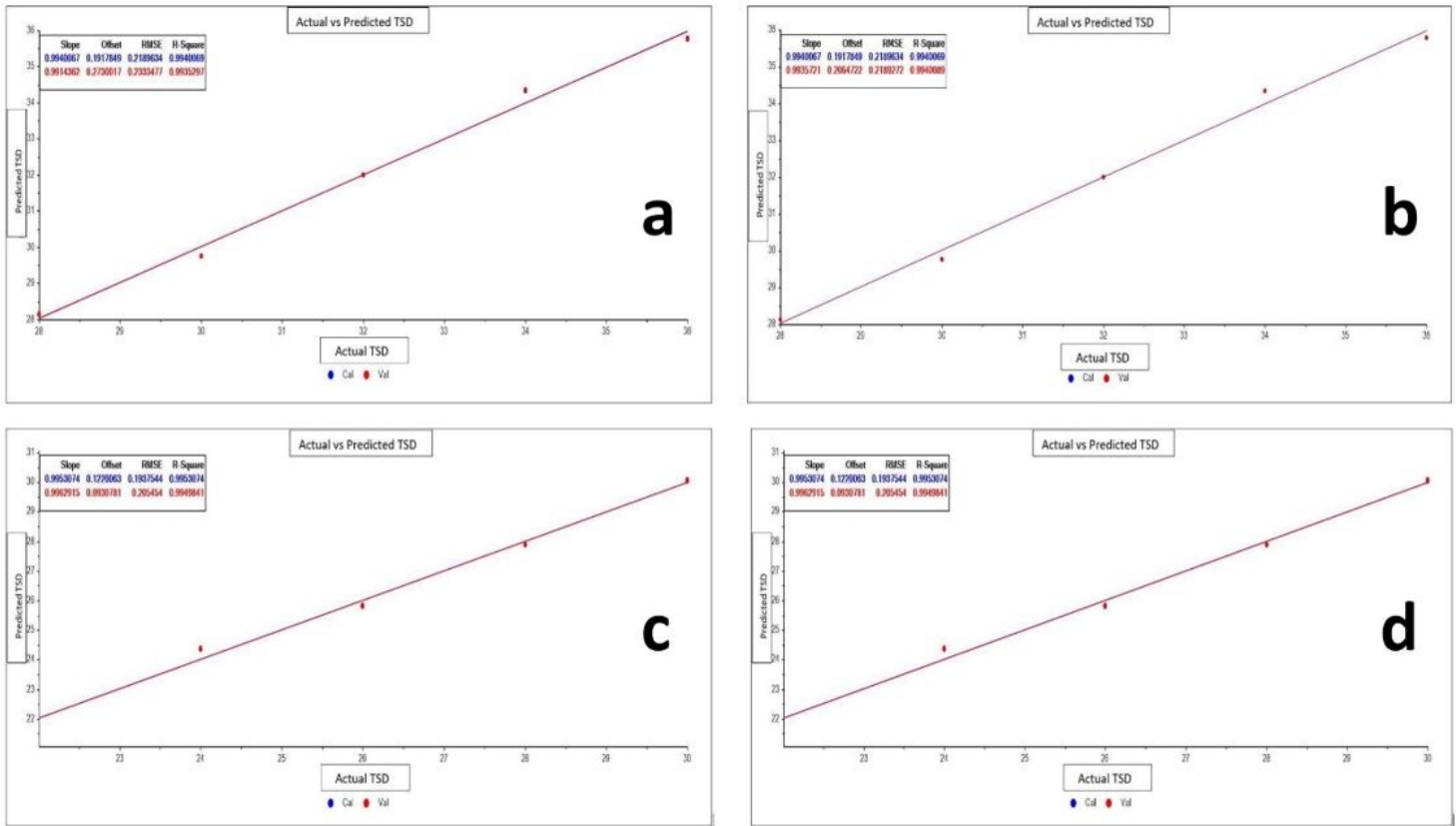


Figure 5

PLS plots of a. actual vs cross validated prediction of TSD for semen (100%) b. actual vs externally validated prediction of TSD for semen (100%) c. actual vs cross validated prediction of TSD for urine (100%) d. actual vs externally validated prediction of TSD for urine (100%)