scientific reports

OPEN

Analysis of the ex-vivo transformation of semen, saliva and urine as they dry out using ATR-FTIR spectroscopy and chemometric approach

Tanurup Das[1], Abhimanyu Harshey[1], Ankit Srivastava[1]✉, Kriti Nigam[1], Vijay Kumar Yadav[1], Kapil Sharma[2] & Arun Sharma[2]

The ex-vivo biochemical changes of different body fluids also referred as aging of fluids are potential marker for the estimation of Time since deposition. Infrared spectroscopy has great potential to reveal the biochemical changes in these fluids as previously reported by several researchers. The 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 the major contribution of water and faster Phase II due to significant evaporation of water. Two spectral regions i.e., 3200–3400 cm⁻¹ and 1600–1000 cm⁻¹ are the major contributors to the spectra of these fluids. Several peaks in the spectral region between 1600 and 1000 cm⁻¹ showed highly significant regression equation with a higher coefficient of determination values in 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 time since deposition of the aforesaid fluids as they dry out. Additionally, this study potentially estimates the time since deposition of an offense from the aging of the body fluids at the early stages after its occurrence as well as works as the precursor for further studies on an extended timeframe.

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 of DNA and its profiling[2]. All of these fluids experience instant biochemical change as they come out from the body. This change can be referred as aging[1–4]. Some of these changes are rapid and the remaining are gradual[3], but both the changes are significant to reveal one of the most important aspects of forensic examination; 'time since the deposition' of a crime[4]. The estimation of 'time since deposition (TSD) of a body fluid at the crime scene potentially solves the problem of situating the time of an offense[3]. The TSD of body fluids additionally counter a wide range of issues regarding the crime scene investigation[3,4]. Most of the TSD studies are based on bloodstain aging[2–7] except for a solitary publication on semen very recently[8]. All these studies except one[9] 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[3,4,10]. A variety of techniques including Gas chromatography[11], Liquid chromatography[12–14], Oxygen electrode[3,15], microRNA based assay[3,16–19], Color transformation chart[20], Electron paramagnetic resonance[21–23], Reflectance spectroscopy in Ultraviolet, Visible and Infrared (IR) region[24–26], smartphone imaging[27], IR absorption[1,3,4,8,9,28,29], Raman scattering[4,6,7,11,30] and fluorescence spectroscopy is explored around the world for the TSD estimation of blood (peripheral and menstrual)[31] and semen[8]. 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

[1]Dr. A.P.J. Abdul Kalam Institute of Forensic Science & Criminology, Bundelkhand University, Jhansi 284128, Uttar Pradesh, India. [2]DFS, Himachal Pradesh, Junga, Shimla 171218, Himachal Pradesh, India. ✉email: ankit_forensic81@rediffmail.com

Body fluid	Wave number (cm⁻¹) (approx.)	Spectral assignment
Semen	3271	O–H stretching (Phase I) and Symmetric N–H stretching of Amide A (Phase II)[37,38,39]
	1637	H–O–H scissoring (Phase I) and C=O stretching (Amide I)[37,38,39]
	2968	Asymmetric CH_3 stretching[37,38,39]
	1546	N–H bending and C–N stretching (Amide II)[37,38,39]
	1446	C–H bending of CH_2 and CH_3[37,38,39,41]
	1396	C=O stretching in COO^- (Fatty acids and polysaccharides)[39]
	1317 and *1243*	N–H bending and C–N stretching (Amide III)[37,39]
	1066 and 970	CH_2OH groups, C–O stretching and COH groups, symmetric Glycosylated proteins, PO_2—stretching (Glycosylated proteins)[37,38,39]
Saliva	*3273* and *3286*	O–H stretching (Phase I) and Symmetric N–H stretching of Amide A (Phase II)[37,38,39]
	3076	Amide B[37,46,39]
	2958	Asymmetric CH_3 stretching[37,38,39]
	2881	Asymmetric CH_2 (methylene) stretching[38,39,CR40CR42]
	1637 and *1645*	H–O–H scissoring (Phase I) and C=O stretching (Amide I)[38,42]
	1546	N–H bending and C-N stretching (Amide II)[38,39,46]
	1448	C–H bending of CH_2 and CH_3[39,40,42,47]
	1404	Symmetric CH[42,47]
	1317 and 1243	N–H bending and C–N stretching (Amide III)[37,39, 47]
	1078 and *1043*	CH_2OH groups, C–O stretching and COH groups, symmetric Glycosylated proteins, PO_2—stretching (Glycosylated proteins)[37–39]
Urine	*3346*	Asymmetric stretching of H–O–H[39,40]
	3273	O–H stretching (Phase I) and Symmetric N–H stretching of Amide A (Phase II)[37,38,46]
	3205	O–H stretching[38–40]
	1658; *1637* and *1623*	H–O–H scissoring (Phase I) and C=O stretching (Amide I)[37–39]
	1456	Asymmetric C–N stretching (Urea)[38–40]
	1157	NH_2 deformation (Urea)[39,40]
	1081	CH_2OH groups, C–O stretching and COH groups, symmetric Glycosylated proteins, PO_2—stretching (Glycosylated proteins)[37–39]

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

stand-alone tool for TSD studies except IR absorption and Raman scattering collectively known as vibrational spectroscopy[4,10,11]. Leading researchers in the forensic sciences community around the world are significantly attracted towards the vibrational spectroscopic technique due to its several advantages over others. It reliably provides a rapid non-destructive analysis of evidence in forensic investigation[4,10,11]. 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 investigations, blood always received special attention in the research community[3,4,10,11,32,33]. All the three aforesaid fluids including blood are consistently found in the broad range of homicides, suicides, rapes, and other sexual assaults[2,34,35]. 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 offense has enormous importance as this is the time when most of the information in a crime scene remains unchanged and also during this timeframe, the most significant information can be obtained from post-mortem. 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[9]. Zhang et al.[9] 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[9,11,36]. 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 based on the absorption changes in specific peaks. Since, Zha et al.[8] experimented with the change of semen stains up to 1 week, in which the first spectra were acquired 12 hours after the deposition of semen ignoring the initial changes. Hence, semen is also included in the present study. In Attenuated total reflectance (ATR) Fourier transforms infrared (FTIR) spectroscopy is combined with the 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 for blood as a similar study is already reported on it[9]. 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)[37–39]. The list of peaks identified in the spectra of all three fluids is summarized in Table 1 with their vibrational assignments. All the acquired
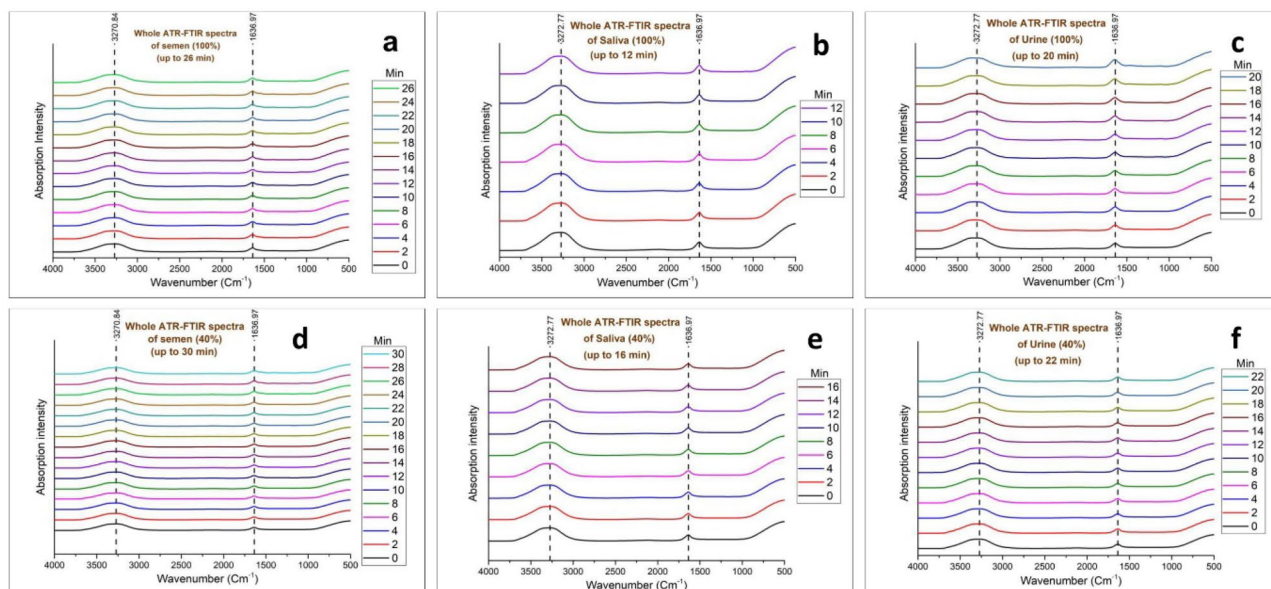
Figure 1. Whole ATR FTIR Phase I spectra for (**a**) 100% semen, (**b**) 100% saliva, (**c**) 100% urine, (**d**) 40% semen, (**e**) 40% saliva, (**f**) 40% urine.
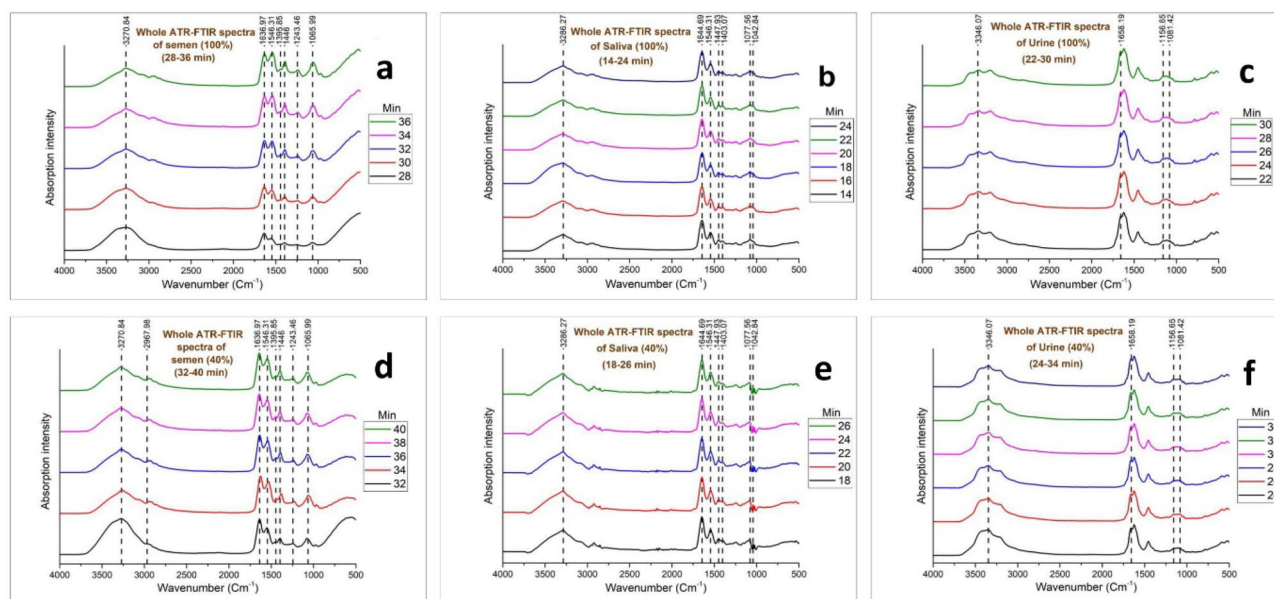

Figure 2. Whole ATR FTIR Phase II spectra for (**a**) 100% semen, (**b**) 100% saliva, (**c**) 100% urine, (**d**) 40% semen, (**e**) 40% saliva, (**f**) 40% urine.

spectra in the present study showed similarity with previously reported spectra by several researchers. Although, some minor spectral bands are not visible. It would probably occur due to the spectral acquisition timing, as in this study, all the spectra were collected at the earliest stage of the ex-vivo degradation process up to their drying. The minor bands may have occurred at the later stages in the ex-vivo environment[8,39–42]. The spectra in other reported articles were collected at least 4–6 h from the time of deposition[39,40]. 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 (Fig. 1a–f). These two peaks at 3270–3273 cm^{-1} (O–H stretching) and 1637 cm^{-1} (approx.) (scissoring of two H atoms bonded with O molecule) appeared due to the high amount of water in all the fresh fluid samples[43–45]. Similar results were obtained in the study by Zhang et al., on blood[9]. After a certain amount of time multiple significant peaks corresponding to the biochemical profile of the fluids are revealed throughout the fingerprint region of the IR spectra (Fig. 2a–f). Following the trend in each fluid, the whole drying time was divided into two phases.

Fluids	Concentration (%)	Phase I (min)			Phase II (min)			Total (min)		
		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

Table 2. Summarized the minimum, maximum and mean values of two phases of body fluid drying.
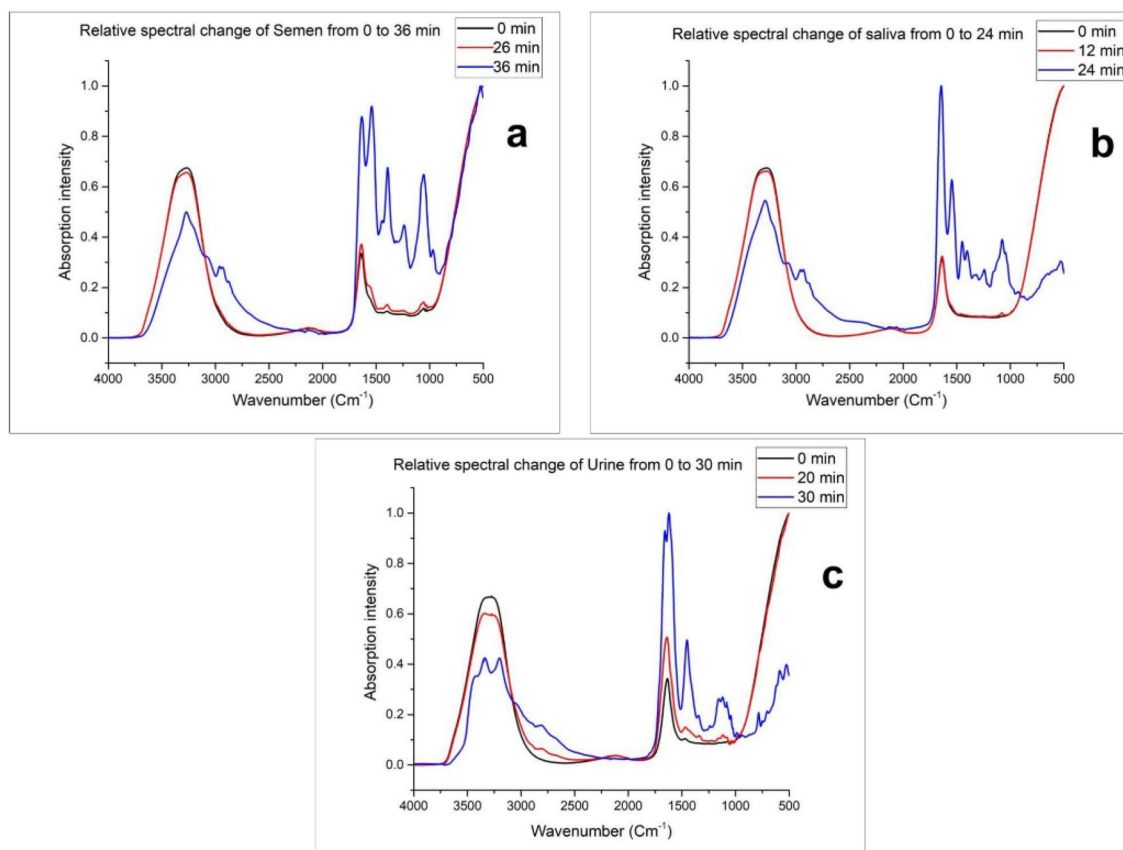
Figure 3. Relative changes in the ATR-FTIR spectra of (**a**) semen, (**b**) saliva and (**c**) urine until drying.

Except for the concentrated samples of fluids, diluted samples of 2:3 ratio (40%) were also prepared to investigate the changes in the drying of the fluids. The dilution was kept constant for all the fluids and a similar extended drying time had been recorded. Multiple dilutions can also alter the drying time as these factors can be studied in future studies on these fluids separately with other factors as experimented by Zhang et al., on blood[9]. The dilution of the fluids showed an extended (2–4 min) phase I due to the excess amount of water in the diluted sample. The longest phase I observed in the Semen samples and the 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 min). 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.

Few researchers reported the correlation between the evaporation of distilled water and time[48,49]. Except for similar height, the peak corresponding to O–H stretching was broader than the peak due to H–O–H scissoring. Except for urine, the absorption intensity of these two peaks showed insignificant change throughout phase I in the spectra of semen and saliva (Figs. 1a–f, 3). On the contrary, Zhang et al.[9] 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.[9] as the peak at 3273 cm^{-1} showed a significant decline in the mean absorbance with time during phase I (Figs. 2c,f, 3c).

(a) Peak	Semen (100%)			Semen (40%)		
	Slopes (a.u/min)	Intercepts (a.u)	Significance (95%) (p value)	Slopes (a.u/min)	Intercepts (a.u)	Significance (95%) (p value)
2967.98	0.017	− 0.32	< 0.05	0.004	0.17	< 0.05
1546.31	0.076	− 1.77	< 0.05	0.021	0.07	< 0.05
1446	0.039	− 0.90	< 0.05	0.007	0.12	< 0.05
1395.85	0.072	− 1.78	< 0.05	0.015	− 0.04	< 0.05
1243.46	0.038	− 0.90	< 0.05	0.006	0.11	< 0.05
1065.99	0.055	− 1.32	< 0.05	0.015	− 0.08	< 0.05
(b) Peak	Saliva (100%)			Saliva (40%)		
1546.31	− 0.002	0.65	< 0.05	− 0.012	1.02	< 0.05
1447.93	− 0.002	0.40	< 0.05	− 0.012	0.74	< 0.05
1403.07	− 0.003	0.38	< 0.05	− 0.012	0.69	< 0.05
1077.56	− 0.002	0.42	< 0.05	− 0.011	0.69	< 0.05
1042.84	− 0.003	0.37	< 0.05	− 0.010	0.54	< 0.05
(c) Peak	Urine (100%)			Urine (40%)		
3346.07	− 0.011	0.72	< 0.05	− 0.014	1.14	< 0.05
1658.19	0.003	0.83	< 0.05	0.002	0.85	< 0.05
1156.65	0.003	0.18	< 0.05	0.002	0.20	< 0.05
1081.42	0.003	0.16	< 0.05	0.001	0.22	< 0.05

Table 3. Slopes and intercepts of age-linked peaks of (a) semen (b) saliva (c) urine at 95% level of significance during the phase II drying. P value is the measure of the statistical significance of the observed difference. In this table it indicates the significance of the slope and intercepts, which are less than 0.05 at 95% level of significance.

Phase II is the fast declination stage where a significant amount of water evaporates rapidly and reveals the other peaks and their intensity changes with time in each body fluid. The peaks at 3271 cm^{-1} (Amide A) and 1637 cm^{-1} (Amide I) 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 phase II (Figs. 1a,d, 2a,d) as the peak in Phase II appeared due to the N–H stretching of Amide instead of O–H stretching of water[8,43–45]. In the phase II drying of semen droplets, one strong (1546 cm^{-1}: Amide II), two medium (1446 cm^{-1}: methylene; CH$_2$ and CH$_3$ and 1066 cm^{-1}: Glycosylated proteins: probably prostate-specific antigen) and three weak (2968 cm^{-1}: CH$_3$ stretching, 1396 cm^{-1}: fatty acids and polysaccharides and 1243 cm^{-1}: Amide III) significant age-linked peaks were observed (Fig. 2a,d). Zha et al.[8] 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[38–40]. Phase II spectra of saliva showed the shift of strong peaks at 3273 cm^{-1} (O–H stretching) to 3286 cm^{-1} (amide A) and 1637 cm^{-1} (H–O–H scissoring) to 1645 cm^{-1} (amide I) that indicated the initiation of this phase (Fig. 2b,e). The shifted peak of amide A sharpens following the trend of semen samples. Among others, one strong and sharp (1546 cm^{-1}: amide II), four weak (1448 and 1403 cm^{-1}: Methylene, 1078 cm^{-1}, and 1043 cm^{-1}: glycosylated proteins) significant age-linked peaks were found (Fig. 2b,e). Including the peak corresponding to amide II, two peaks at 1448 cm^{-1} and 1078 cm^{-1} are similar to the peaks at 1446 and 1066 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 3273 cm^{-1} divided into 3346 cm^{-1} and 3205 cm^{-1} and 1637 cm^{-1} divided into 1623 and 1658 cm^{-1} (Fig. 2c,f). One strong (1658 cm^{-1}: amide I), one medium (3346 cm^{-1}: H–O–H stretching), and 2 weak (1156 cm^{-1}: urea and 1081 cm^{-1}: Glycosylated proteins) peaks were observed in the phase II spectra of urine samples (Fig. 2c,f) that significantly changes during the drying process. The peak at 1081 cm^{-1} is similar to the peaks at 1066 cm^{-1} and 1078 cm^{-1} of semen and saliva, respectively. Due to the presence of a significant quantity of prostate-specific antigens in semen, the peak corresponding to glycosylated protein is stronger in its spectra than saliva and urine[37–39]. In several previous literatures on the IR signature of saliva and urine, the above-mentioned peaks were reported by researchers[37–42]. Amide A, I, II, and glycosylated proteins are the common biochemical components found in all the 3 body fluids. Elkins, Orphanou, and Takamura et al.[37–39] previously reported the presence of the common biochemicals in all these body fluids in their article on body fluid identification by ATR-FTIR. In Phase II, all the age-linked peaks of each fluid showed a linear relationship between the mean absorbance at each time point and TSD (Fig. 2a–f).

Statistical results. The regression equation of a line is the representation of a prediction model. Table 3 depicts the slopes and intercepts calculated for all the age-linked peaks of three body fluids with a 95% level of significance.

Body fluids show significant ex-vivo degradation when exposed for a relatively longer period. While early changes are very limited as only a few relevant peaks due to aging are visible. Hence, the TSD estimation was

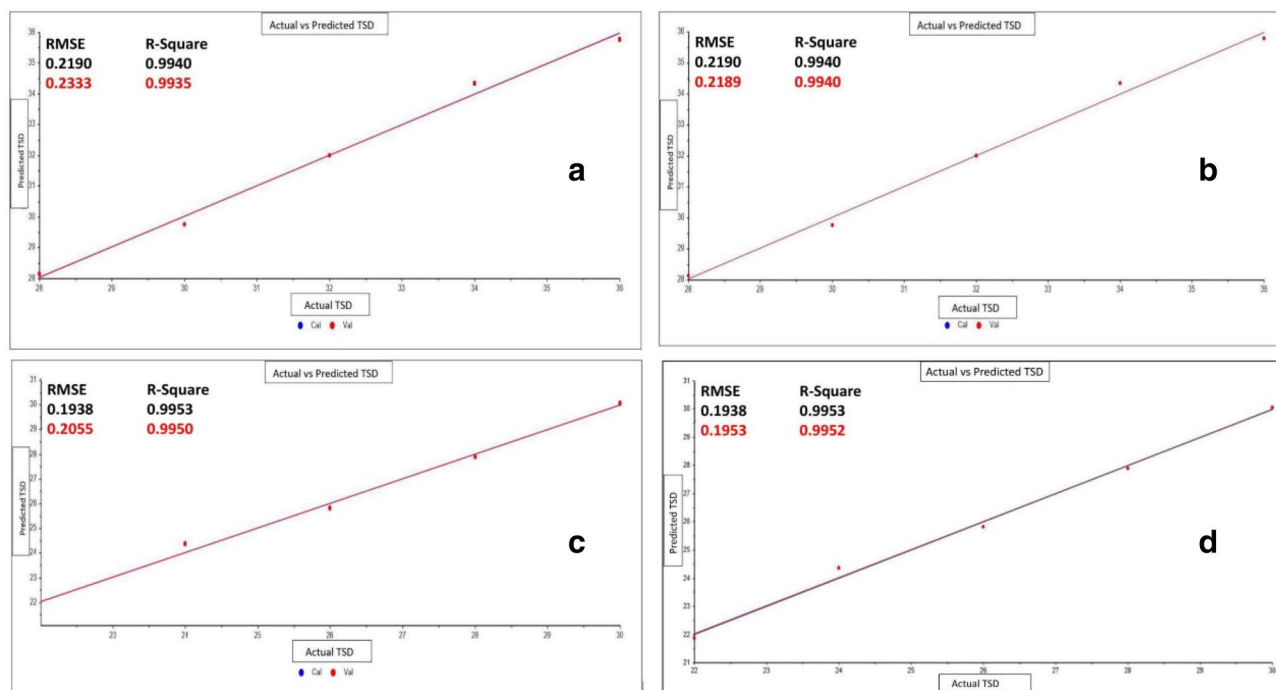| Fluids	Cross-validation		External Prediction		RPD
	R^2	RMSECV (min)	R^2	RMSEP (min)	
(a) Principal component regression (PCR) analysis table					
Semen (100%)	0.9935	0.2346	0.9939	0.2201	13.50
Semen (40%)	0.9390	0.7167	0.9415	0.6840	20.53
Saliva (100%)	0.9885	0.3742	0.9891	0.3568	9.64
Saliva (40%)	0.9973	0.1517	0.9973	0.1458	19.53
Urine (100%)	0.9948	0.2084	0.9951	0.1983	15.01
Urine (40%)	0.9518	0.7659	0.9532	0.7383	4.73
(b) Partial least square regression (PLSR) analysis table					
Semen (100%)	0.9935	0.2333	0.9940	0.2189	13.57
Semen (40%)	0.9957	0.1897	0.9970	0.1546	19.26
Saliva (100%)	0.9889	0.3737	0.9891	0.3561	9.67
Saliva (40%)	0.9969	0.1618	0.9973	0.1457	20.44
Urine (100%)	0.9950	0.2055	0.9952	0.1953	15.24
Urine (40%)	0.9523	0.7622	0.9540	0.7327	4.78

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.

performed only on the age-linked peaks of three body fluids. Among several strong and medium peaks 6 (semen), 5 (saliva), and 4 (urine) peaks were selected to calculate the TSD of the body fluids. Irrespective of the variation in the numbers of age-linked peaks for the three body fluids, it was evident that the regression models successfully estimated the TSD with very high accuracy. Both Principal Component Regression (PCR) Analysis and Partial Least Square Regression (PLSR) are strong chemometric tools for the estimation of TSD of body fluids as reported in previous studies[1,4–11]. 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 and Prediction (RMSECV and RMSEP) values in both the models for all the three fluids except diluted urine, showed low values (Table 4a,b). Although the initial changes in the IR spectra of body fluids are relatively less distinguishable in comparison to the samples exposed for a 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: 0.7659; PLSR: 0.7622) and RMSEP (PCRA: 0.7383; PLSR: 0.7327) (Table 4a,b) in both the models are relatively higher, that potentially interfere in the accurate age estimation. While, diluted semen also showed higher RMSECV (0.7167) and RMSEP (0.6840) in the PCRA regression model, while PLSR predicts the age for the same condition with significantly higher accuracy with RMSECV and RMSEP values of 0.1897 and 0.1546, (Table 4a,b) respectively. The lower RMSE value also depicts that there is very minute inter-donor variation and low standard deviation values (0.00002–0.00004) of the spectral data obtained from the repeated sampling showed minimal intra-donor variation. Additionally, the RPD (Residual Predictive Deviation) values for each model were also calculated and it has been found that all the values are above 3 which indicates an excellent prediction model accuracy (Fig. 4)[50,51]. 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. Figures 4 and 5 depicts the PCR and PLS plots of actual vs predicted regression lines for age-linked peaks of three body fluids. Finally, one-way ANOVA has also been applied to the age-linked peaks separately. The F-statistic for all the age-linked peaks showed a significant difference between the time intervals (Table 5).

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 with a major contribution of water and the second phase depicts relatively faster evaporation. These two phases can be distinguishable from the ATR-FTIR spectra of each fluid which is a significant marker for estimating the time since deposition of the fluid(s). 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. Additionally, if a body fluid is accidentally diluted during the deposition on a wet nonporous substrate during the earliest phase post-deposition, the accuracy of the estimation of TSD can be significantly altered as the water evaporation potentially take more time. In the practical scenario, fresh body fluid samples in liquid and semi-liquid (fluid samples with loss of a certain quantity of water since deposition) conditions from any non-porous surface (e.g., glass, metal, tile, etc.) can be easily collected through a pipette-like apparatus. We can acquire the spectra of the freshly collected fluid with a portable IR instrument containing an ATR-FTIR crystal face exclusively dedicated to forensic crime scene investigation purposes. Although the drying process of body fluids is different on porous substrates like cloth fabrics, carpets, etc. Hence, further experiments on this subject can be framed based on several factors like, different concentrations, quantities, and interference of porous substrates. 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 with a very low RMSECV and high R^2 values. Hence, the results (spectral and statistical) of this study potentially be used as a reference for the further TSD studies on these three fluids with a longer timeframe including different factors.
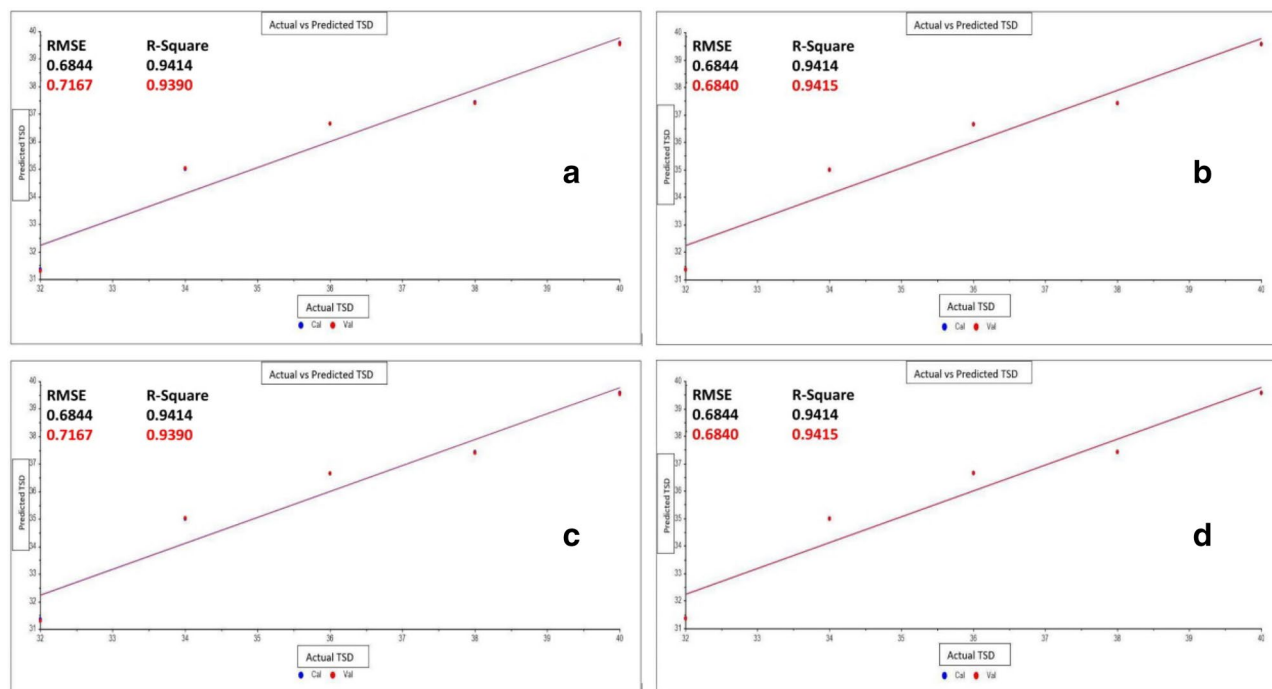
Figure 4. PCR plots of (**a**) actual vs cross validated prediction of TSD for semen (40%), (**b**) actual vs externally validated prediction of TSD for semen (40%), (**c**) actual vs cross validated prediction of TSD for saliva (100%), (**d**) actual vs externally validated prediction of TSD for saliva (100%).

Figure 5. PLSR 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%).

Fluid	F-statistic (95% level of significance)					
Semen (100%)	2968 cm^{-1}	1546 cm^{-1}	1446 cm^{-1}	1396 cm^{-1}	1243 cm^{-1}	1066 cm^{-1}
	2.87	5.44	1.09	2.83	1.26	2.58
Semen (40%)	1.11	6.68	6.28	2.82	3.83	3.17
Saliva (100%)	1546 cm^{-1}	1448 cm^{-1}	1404 cm^{-1}	1078 cm^{-1}	1043 cm^{-1}	
	4.72	6.57	8.80	5.24	9.56	
Saliva (40%)	1.65	1.37	1.46	9.33	7.47	
Urine (100%)	3346 cm^{-1}	1658 cm^{-1}	1156 cm^{-1}	1081 cm^{-1}		
	8.84	1.04	6.84	1.21		
Urine (40%)	2.89	3.95	2.59	1.61		

Table 5. Calculated F-statistics by one way ANOVA for all the age-linked peaks of three body fluids.

Materials and methods

Sampling. Samples of saliva, semen, and urine were collected from randomly selected eight healthy volunteers (28–40 years). All the methods of this study were carried out in accordance with the World Medical Association Declaration of Helsinki. The experimental protocols were approved by the Ethical Committee of Maharani Laxmi Bai Medical College (4647/IEC/2020/SC-1), Bundelkhand University, Jhansi, Uttar Pradesh, India. All the donors were informed about the nature and procedure of the work and written informed consent was taken from each donor. Each body fluid from a donor was collected separately in a glass test tube(s) just before 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 concentrations 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 instantly after the collection of the fluid. For experimental purposes, saliva and urine samples were repeatedly taken for three consecutive days, and semen was collected three times with an interval of 3 days to observe any intra-variations of their composition. All the repeated samples were collected from the same individuals.

Collection and pretreatment 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 is 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. The crystal was used as the drying surface to reduce any loss of water from the fluids during the transfer of the samples from the substrates. 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 and 17 °C 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 immediately after placing the droplet on the crystal within the range of 4000 cm^{-1} to 500 cm^{-1} with 12 scans and a 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-min 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[52,53].

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 a 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 indicative of a 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 for external validation purposes.

Data availability

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

References

1. Lin, H. *et al.* Estimation of the age of human bloodstains under the simulated indoor and outdoor crime scene conditions by ATR-FTIR spectroscopy. *Sci. Rep.* **7**, 13254 (2017).
2. Zapata, F. Emerging spectrometric techniques for the forensic analysis of body fluids. *Trends Anal. Chem.* **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 Sci. Int.* **216**, 1–11 (2012).
4. Zadora, G. & Menżyk, A. In the pursuit of the holy grail of forensic science—Spectroscopic studies on the estimation of time since deposition of bloodstains. *Trends Anal. Chem.* **105**, 137–165 (2018).
5. Weber, A. R. & Lednev, I. K. Crime clock—Analytical studies for approximating time since deposition of bloodstains. *Forensic Chem.* **19**, 100248 (2020).
6. Doty, K. C., Muro, C. K. & Lednev, I. K. Predicting the time of the crime: Bloodstain aging estimation for up to two years. *Forensic Chem.* **5**, 1–7 (2017).
7. Doty, K. C., McLaughlin, G. & Lednev, I. K. A Raman, "spectroscopic clock" for bloodstain age determination: The first week after deposition. *Anal. Bioanal. Chem.* **408**(15), 3993–4001 (2016).
8. Zha, S. *et al.* Estimation of the age of human semen stains by attenuated total reflection Fourier transform infrared spectroscopy: A preliminary study. *Forensic Sci. Res.* **5**, 119 (2019).
9. Zhang, Y. *et al.* Changes in attenuated total reflection Fourier transform infrared spectra as blood dries out. *J. Forensic Sci.* **62**, 761 (2016).
10. Muro, C. K., Doty, K. C., Bueno, J., Halamkova, L. & Lednev, I. K. Vibrational spectroscopy: Recent developments to revolutionize forensic science. *Anal. Chem.* **87**, 306–327 (2015).
11. 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. *Microchem. J.* **158**, 105278 (2020).
12. Inoue, H., Takabe, F., Iwasa, M. & Maeno, Y. Identification of fetal hemoglobin and simulated estimation of bloodstain age by high-performance liquid chromatography. *Int. J. Legal Med.* **104**, 127–131. https://doi.org/10.1007/BF01369715 (1991).
13. Andrasko, J. The estimation of age of bloodstains by HPLC analysis. *J. Forensic Sci.* **42**(4), 601–607 (1997).
14. Seok, A. E. *et al.* Estimation of age of bloodstains by mass-spectrometry: A metabolomic approach. *Anal. Chem.* **90**(21), 12431–12441. https://doi.org/10.1021/acs.analchem.8b01367 (2018).
15. Matsuoka, T., Taguchi, T. & Okuda, J. Estimation of bloodstain age by rapid determinations of oxyhemoglobin by use of oxygen electrode and total hemoglobin. *Biol. Pharm. Bull.* **18**(8), 1031–1035. https://doi.org/10.1248/bpb.18.1031 (1995).
16. Arany, S. & Ohtani, S. Age estimation of bloodstains: A preliminary report based on aspartic acid racemization. *Forensic Sci. Int.* **212**(1–3), e36. https://doi.org/10.1016/j.forsciint.2011.05.015 (2011).
17. Nakao, K.-I., Shimada, R., Hara, K. & Kibayashi, K. Experimental study on age estimation of bloodstains based on biological and toxicological analysis. *Open Forensic Sci. J.* **6**, 6–11. https://doi.org/10.2174/1874402801306010006 (2013).
18. 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 Sci. Int.* **138**(1–3), 94–103. https://doi.org/10.1016/j.forsciint.2003.09.008 (2003).
19. Anderson, S., Howard, B., Hobbs, G. R. & Bishop, C. P. A method for determining the age of bloodstain. *Forensic Sci. Int.* **148**(1), 37–45. https://doi.org/10.1016/j.forsciint.2004.04.071 (2005).
20. Tomellini, L. De l'emploi d'une table chromatique pour les taches du sang. *Arch. d'Antropologie criminelle de Criminol* **14** (1907).
21. 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**(3), 934–939. https://doi.org/10.1073/pnas.63.3.934 (1969).
22. Svistunenko, D. A. *et al.* The pH dependence of naturally occurring low-spin forms of methaemoglobin and metmyoglobin: An EPR study. *Biochem. J.* **351**(3), 595–605. https://doi.org/10.1042/bj3510595 (2000).
23. Miki, T., Kai, A. & Ikeya, M. Electron spin resonance of bloodstains and its application to the estimation of time after bleeding. *Forensic Sci. Int.* **35**(2–3), 149–158. https://doi.org/10.1016/0379-0738(87)90051-X (1987).
24. Li, B., Beveridge, P., O'Hare, W. & Islam, M. The estimation of the age of a blood stain using reflectance spectroscopy with a micro-spectrophotometer, spectral pre-processing and linear discriminant analysis. *Forensic Sci. Int.*. https://doi.org/10.1016/j.forsciint.2011.05.031 (2011).
25. 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 Sci. Int.* **206**(1–3), 166–171. https://doi.org/10.1016/j.forsciint.2010.07.034 (2011).
26. Sun, H. *et al.* Non-invasive prediction of bloodstain age using the principal component and a back propagation artificial neural network. *Laser Phys. Lett.* **14**(9), 095601. https://doi.org/10.1088/1612-202X/aa7c48 (2017).
27. Thanakiatkrai, P., Yaodam, A. & Kitpipit, T. Age estimation of bloodstains using smartphones and digital image analysis. *Forensic Sci. Int.* **233**(1–3), 288. https://doi.org/10.1016/j.forsciint.2013.09.027 (2013).
28. Lamontagne, M., Brown, C. W., Botenjic-Sehic, E. & Tsaparikos, M. Forensic application of near-infrared spectroscopy: Aging of bloodstains. *Spectroscopy* **24**(2) (2009).
29. Lu, Z. *Applications of Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy for Forensic Analysis.* (Doctoral dissertation). Retrieved from https://scholarcommons.sc.edu/etd/3234 (2015).
30. Menzyk, A. *et al.* Toward a novel framework for bloodstains dating by Raman spectroscopy: How to avoid sample photodamage and subsampling errors. *Talanta* **209**, 120565. https://doi.org/10.1016/j.talanta.2019.120565 (2020).
31. Wójtowicz, A., Weber, A., Wietecha-Posłuszny, R. & Lednev, I. K. Probing menstrual bloodstain aging with fluorescence spectroscopy. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **248**, 119172. https://doi.org/10.1016/j.saa.2020.119172 (2020).
32. Li, R. *Forensic Biology* (CRC Press, Taylor & Francis Group, Boca Raton, 2011).
33. Gunn, A. *Essentials Forensic Biology* 2nd edn. (Wiley, Chichester, 2009).
34. Jackson, A. R. W. & Jackson, J. M. *Forensic Science* 115–136 (Pearson Prentice Hall, Harlow, England, 2008).
35. Li, R. *Forensic Chemistry Handbook* 69–290 (Wiley, New Jersey, 2012).
36. 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. *Adv. Condens. Matter Phys.* **2018**, 1–24. https://doi.org/10.1155/2018/5214924 (2018).

www.nature.com/scientificreports/

37. Orphanou, C.-M. The detection and discrimination of human body fluids using ATR FT-IR spectroscopy. *Forensic Sci. Int.* **252**, e10–e16. https://doi.org/10.1016/j.forsciint.2015.04.020 (2015).
38. 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. *Sci. Rep.* https://doi.org/10.1038/s41598-018-26873-9 (2018).
39. Elkins, K. M. Rapid presumptive "fingerprinting" of body fluids and materials by ATR FT-IR spectroscopy. *J. Forensic Sci.* **56**(6), 1580–1587. https://doi.org/10.1111/j.1556-4029.2011.01870.x (2011).
40. Takamura, A., Halamkova, L., Ozawa, T. & Lednev, I. K. Phenotype profiling for forensic purposes: determining donor sex based on fourier transform infrared spectroscopy of urine traces. *Anal. Chem.* **91**, 6288–6295. https://doi.org/10.1021/acs.analchem.9b01058 (2019).
41. Barcˇot, O. *et al.* Investigation of spermatozoa and seminal plasma by fourier transform infrared spectroscopy. *Appl. Spectrosc.* **61**(3), 309–313. https://doi.org/10.1366/000370207780220804 (2007).
42. Scott, D. A. *et al.* Diabetes-related molecular signatures in infrared spectra of human saliva. *Diabetol. Metab. Syndr.* **2**, 48. https://doi.org/10.1186/1758-5996-2-48 (2010).
43. Chwastowski, J. *et al.* Structure and physiochemical properties of water treated under nitrogen with low-temperature glow plasma. *Waters* **12**, 1314. https://doi.org/10.3390/w12051314 (2020).
44. Rios-Reina, R., Callejon, R. M., Oliver-Pozo, C., Amigo, J. M. & Garcia-Gonzalez, D. ATR-FTIR as a potential tool for controlling high quality vinegar categories. *Food Control* **78**, 230–237. https://doi.org/10.1016/j.foodcont.2017.02.065 (2017).
45. Popa, A. M., Vleugels, J., Vermant, J. & Van der Biest, O. Influence of surfactant addition sequence on the suspension properties and electrophoretic deposition behaviour of alumina and zirconia. *J. Eur. Ceram. Soc.* **26**(6), 933–939. https://doi.org/10.1016/j.jeurceramsoc.2004.12.023 (2004).
46. Garidel, P. & Schott, H. Fourier-Transform Midinfrared Spectrscopy for Analysis and Screening of Liquid Protein Formulations. *Bio Process Int.* (2006).
47. 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. https://doi.org/10.1039/c0an00529k (2010).
48. Al-Shammiri, M. Evaporation rate as a function of water salinity. *Desalination* **150**, 189–203. https://doi.org/10.1016/S0011-9164(02)00943-8 (2002).
49. 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. https://doi.org/10.1007/s13369-013-0653-7 (2013).
50. Cozzolino, D., Cynkar, W. U., Shah, N. & Smith, P. Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality. *Food Res. Int.* **44**(7), 1888–1896. https://doi.org/10.1016/j.foodres.2011.01.041 (2011).
51. Cozzolino, D. *et al.* Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Anal. Chim. Acta* **513**(1), 73–80. https://doi.org/10.1016/j.aca.2003.08.066 (2004).
52. Lee, L. C., Liong, C.-Y. & Jemain, A. A. A contemporary review on data preprocessing (DP) practice strategy in ATR-FTIR spectrum. *Chemometr. Intell. Lab. Syst.* **163**, 64–75 (2017).
53. Gautam, R., Vanga, S., Ariese, F. & Umapathy, S. Review of multidimensional data processing approaches for Raman and Infrared spectroscopy. *EPJ Tech. Instrum.* **2**(1). https://doi.org/10.1140/epjti/s40485-015-0018-6 (2015).

Acknowledgements

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.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.