Qualitative evaluation of chromatographic data from quality control schemes using a support vector machine

M. Ventura,*[ab] A. Sanchez-Niubo,[c] F. Ruiz,[d] N. Agell,[e] R. Ventura,[af] C. Angulo,[d] A. Domingo-Salvany,[c] J. Segura[af] and R. de la Torre[af]

Received 30th July 2007, Accepted 4th October 2007
First published as an Advance Article on the web 25th October 2007
DOI: 10.1039/b711653p

The qualitative evaluation of chromatographic data in the framework of external quality assurance schemes is considered in this paper. The homogeneity in the evaluation of chromatographic data among human experts in samples with analytes close to the limit of detection of analytical methods was examined and also a Support Vector Machine (SVM) was developed as an alternative to experts for a more homogeneous and automatic evaluation. A set of 105 ion chromatograms obtained by anti-doping control laboratories was used in this study. The quality of the ion chromatograms was evaluated qualitatively by nine independent experts (associating a score from 0 to 4) and also more objectively taking into account chromatographic parameters (peak width, asymmetry, resolution and S/N ratio). Results obtained showed a high degree of variability among experts when judging ion chromatograms. Experts applying extremely outlying evaluation criteria were identified and excluded from the data used to develop the SVM. This machine was built providing the system with qualitative information (scores assigned by experts) and with objective data (parameters) of the ion chromatograms. A seven-fold cross-validation approach was used to train and to evaluate the predictive ability of the machine. According to the results obtained, the SVM developed was found to be close to the reasoning process followed by the homogeneous human expert group. This machine also could provide a scoring system to sort laboratories according to the quality of their results. The qualitative evaluation of analytical records using a scoring system allowed the identification of the main factors affecting the quality of chromatographic analytical data, such as the specific analytical technique applied and the adherence to guidelines for reporting positive results.

1. Introduction

Results provided by analytical laboratories are routinely used to take administrative, legal and/or clinical decisions. Analytical testing laboratories are requested to perform their activities under strict quality standards in order to assure the reliability of their results. Requirements of these standards include, in addition to method validation and internal quality control procedures, the participation of laboratories in external quality assessment schemes (EQAS).[1]

Quite often, laboratories are asked to report results close to threshold concentrations (either administrative or close to limits of detection of analytical methods) for reporting positive analytical findings. Drug testing laboratories in the area of analytical toxicology (*i.e.* horse testing, doping control, forensic toxicology) are representative of this situation. Regardless of whether the results are reported on a quantitative basis or a qualitative one, the evaluation of chromatographic data supporting analytical findings should be of relevance particularly when substances detected are close to threshold concentrations. The quality of chromatographic data at these concentrations, where the signal-to-noise ratio is low and the contribution of interfering peaks may be high, should be checked. EQAS are the tools to verify the performance of laboratories in these demanding conditions; nevertheless, most of them evaluate results on the basis of qualitative findings (the presence or absence of a given drug/metabolite) or quantitative results, but little attention is paid to the quality of analytical records. Efforts to develop an objective and reliable evaluation system of chromatographic data appear as a relevant aspect in many areas of analytical chemistry.

Doping laboratories are accredited following the ISO17025 regulation,[1] one of the most demanding quality standards for laboratories, and their compliance is evaluated by national accreditation bodies and by an EQAS, run until the year 2003 by the International Olympic Committee (IOC) and at present by the World Anti-Doping Agency (WADA). The proficiency of these laboratories is evaluated through the analytical data obtained in the analysis of five samples four times per year. These data consist of qualitative information on the substances

[a] *Unitat de Recerca en Farmacologia, Institut Municipal d'Investigació Mèdica-Hospital del Mar, Parc de Recerca Biomèdica de Barcelona, Doctor Aiguader 88, 08003 Barcelona, Spain. E-mail: mventura@imim.es; Fax: +34 933 160 499; Tel: +34 933 160 472*
[b] *Universitat Autònoma de Barcelona, UDIMAS, Barcelona, Spain*
[c] *Health Services Research Unit, Institut Municipal d'Investigació Mèdica-Hospital del Mar, Parc de Recerca Biomèdica de Barcelona, Barcelona, Spain*
[d] *Research Group on Knowledge Engineering (GREC), Universitat Politècnica de Catalunya, Barcelona, Spain*
[e] *Departament de Mètodes Quantitatius, ESADE, Barcelona, Spain*
[f] *Universitat Pompeu Fabra, CEXS, Barcelona, Spain*

found in the samples, quantitative determinations for those substances where a threshold concentration has been established as well as chromatograms and methodological information. The evaluation of results in qualitative (*i.e.* sample compatible with the intake of a given drug) as well as quantitative terms is straightforward and benefits from well-defined evaluation procedures. One of the limitations of the system lays with the difficulties of an objective evaluation of the quality of raw chromatographic data provided by laboratories. However, the workload is often such that evaluation of these data suffers from the subjective appreciation of human evaluators (EQAS Steering Committee). Consequently, the feedback from them on the quality of analytical records is usually minimal and provides little information to laboratories on their performance when compared to the rest of the laboratories.

The European Union funded the ALADIN 2002 project (Analytical Laboratories for AntiDoping control: International Network for external quality assessment) with the objective to optimize several aspects of the EQAS.[2] This manuscript will address two major issues concerning the qualitative evaluation of analytical data from EQAS: (i) how experts evaluate chromatographic analytical data (GC–MS) in samples with analytes' concentrations close to the detection limit and (ii) the application of artificial intelligence (AI) techniques as an alternative to the experts' EQAS Steering Committee for a more homogenous and automatic evaluation of analytical records.

The technique selected in this study was a Support Vector Machine (SVM) which is a machine-learning method developed by Vapnik and co-workers in the mid 1990s.[3,4] SVM is built upon a theory about learning with limited samples, called statistical learning theory.[3–5] To the authors' knowledge, SVM has never been applied to evaluate chromatographic data for identification of drugs in biological fluids. However, the SVM has been extensively studied[6–8] and applied in several areas, including chemometrics.[9] The SVM has been used to screen and identify potential biomarkers and to establish patterns for the detection of different type of diseases (cancer,[10–15] tuberculosis,[16] heart failure[17] and cerebral accidents),[18] to compare gene expression patterns induced by xenobiotics in toxicogenomics[19–21] and to characterize the effects of sample handling[22] on mass spectrometric data from proteins' analysis. SVM has also been used to classify plastics by means of their mid-infrared spectra,[23] to detect meat and bone meal in compound feeds using near-infrared imaging spectroscopy[24] and to characterize materials using pyrolysis-gas chromatography–mass spectroscopy data.[25] SVM has been also applied to ion-exchange chromatography, used for the purification of biomolecules from complex biological mixtures, to predict the retention times of proteins[26] as well as to study the factors that could improve the separation between proteins and impurities.[27]

The results presented in this work, in spite of being exemplified in the doping control area, are of general interest for most analytical laboratories using chromatographic data.

2. Experimental

2.1. Description of data

Nine independent experts, belonging to the managing staff of four European IOC/WADA accredited anti-doping control laboratories, were selected. These experts had excellent skills in the analysis and evaluation of doping control samples and had more than ten years of experience in this analytical field.

A set of 105 ion chromatograms was compiled to be evaluated by the experts. These chromatograms were obtained from analytical reports of laboratories accredited by the IOC participating in several re-accreditation exercises (IOC EQAS) from 1997 to 2001.[28] Data selected were ion chromatograms obtained after the analysis of samples containing low concentrations of 19-norandrosterone (around 2 ng ml^{-1}), the main metabolite of the anabolic agent nandrolone,[29,30] by different GC–MS techniques: GC coupled to low resolution MS (LRMS), GC coupled to high resolution MS (HRMS) and GC coupled to tandem MS (MS-MS). For LRMS and HRMS, a chromatographic peak at *m*/*z* 405 (see Fig. 1) corresponding to the bis-*O*-trimethylsilyl (bis-*O*-TMS) derivative of the compound was evaluated; for MS-MS, the chromatograms of daughter ions arising from *m*/*z* 405 (*m*/*z* 315; *m*/*z* 225)[31–33] were considered. The ion chromatograms studied, sorted by year and by technique applied, are listed in Table 1 with the *m*/*z* values of the ions selected for the evaluation.

Experts were provided with both the printed ion chromatograms and a spreadsheet for recording the evaluation. No additional information (*e.g.* the analytical technique used) was provided. Experts evaluated the quality of 105 ion chromatograms, assigning to each of them a score from 0 to 4: (0) very bad, (1) bad, (2) sufficient, (3) good and (4) very good.

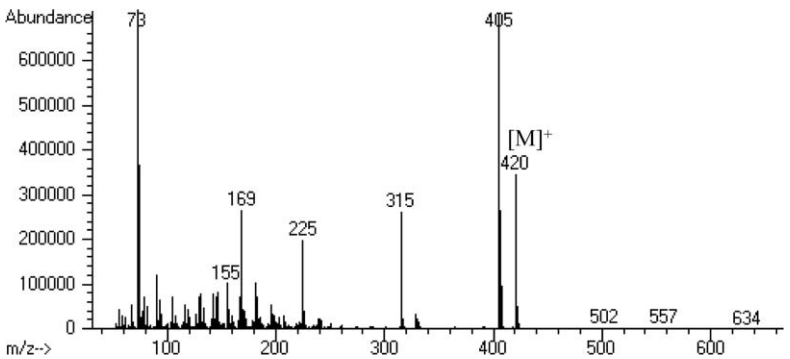

Fig. 1 Electron impact mass spectrum of 19-norandrosterone-bis-*O*-TMS.

Table 1 Chromatograms studied by year, by technique and *m/z* values of the ions selected for the evaluation

Year	Analytical technique			Total
	LRMS[a]	HRMS[a]	MS-MS[b]	
1997	7	4	4	15
1998	10	6	4	20
1999	10	8	4	22
2000	14	6	3	23
2001	13	7	5	25
Total	**54**	**31**	**20**	**105**

[a] *m/z* 405. [b] *m/z* 405 (n = 2), *m/z* 315 (n = 13), *m/z* 225 (n = 2) and *m/z* of the ion is unknown (n = 3).

2.2. Development of the support vector machine (SVM)

Six sets of 15 ion chromatograms (training chromatograms) were used to develop the SVM and a seventh set was used to evaluate the predictive ability of the SVM developed. This procedure was applied seven times, as follows. The approach randomly divided the 105 ion chromatograms in seven subsets of equal size. The SVM was trained seven times, each time using six out of seven subsets as a training set and leaving out one of the subsets (1 out of 7). The ion chromatograms of the seventh subset were used to evaluate the predictive ability of the SVM.

To develop the SVM [see Fig. 2(A)], the system was provided with objective data and with qualitative information for each training ion chromatogram. The objective data were four parameters measured for each chromatographic peak:

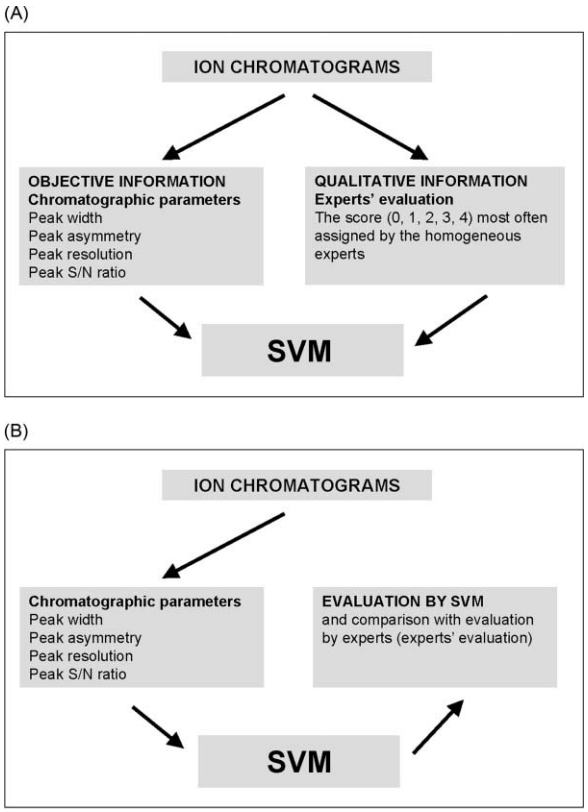

Fig. 2 Description of the process of training (A) and testing (B) of the SVM.

peak width, peak asymmetry, peak resolution and signal-to-noise ratio.[34–36] Peak width was calculated by using the number of theoretical plates: $N = 5.54 \times (t_R/W_{1/2})^2$, where t_R is the retention time and $W_{1/2}$ is the width at 50% peak height. Peak asymmetry was measured calculating the asymmetry factor $A_S = b/a$ at 50% peak height, where b is the distance from the center line of the peak to the back slope and a is the distance from the center line of the peak to the front slope. Peak resolution was measured taking into account the separation between two peaks in terms of their average peak width at base: $R_S = 2\Delta t/(W_{b_1} + W_{b_2})$. And, the signal-to-noise ratio was calculated as the height of the peak divided by the height of the baseline noise. These parameters were measured by an independent expert not belonging to the experts' team. Data were normalized to mean zero and standard deviation one by using the training ion chromatograms and this pre-processing was extended to the ion chromatograms for evaluation. The qualitative information was the evaluation score most often assigned to each ion chromatogram by the experts (from now "experts' evaluation").

The objective data and the qualitative information for each training ion chromatogram were introduced in a software program. Gaussian functions were used as kernel and the hyperparameters (regularization parameter and kernel width) were optimized on a grid, as is usual in the literature.[37] Employed software was Matlab R14® (the Mathworks, Inc., Natick, MA, USA) with the exact QP solver from the Optimization Toolbox® (The Mathworks).[38] As a result, an algorithm was generated; the algorithm generalized the information contained in the training ion chromatograms and designed a function able to assign the score for new ion chromatograms using the associated objective data.

The predictive capacity of the SVM [see Fig. 2(B)] was evaluated by introducing the objective data of new ion chromatograms (chromatograms of each seventh subset). The evaluation score assigned by the SVM was compared to the "experts' evaluation". The performance of the SVM was considered successful when it assigned the same score "experts' evaluation". No intermediate scores were considered. Then, the five scores were taken as five classes and the capacity of the SVM to classify the ion chromatogram in the same score (class) as the experts' evaluation was also evaluated.

2.3. Calculations

Homogeneity between experts was assessed using four statistical tests on the scores assigned by them to the ion chromatograms. Variability in the experts' evaluation was assessed using two descriptive analyses: the mean of the scores assigned by each expert and the inter–expert variance by way of ANOVA. Concordance between experts was evaluated assessing the absence or presence of inter-expert bias using the Cochran's Q-test[39] and measuring the proportion of agreement.[40,41] The two last analyses were also used as criteria to exclude those experts with outlying evaluation patterns in order to obtain a homogeneous experts' team. To apply these two tests it was necessary to define a dichotomous variable; for this reason ion chromatograms with scores 2, 3 and 4 (sufficient, good and very good, respectively) were labeled as

'acceptable' and those with scores 1 and 0 (bad and very bad, respectively) as 'unacceptable'.

To measure how much agreement was actually present between the SVM and the homogeneous experts compared to how much agreement would be expected to be present by chance alone, the kappa coefficient was calculated.[42,43]

To assess the influence of the GC–MS technique used and the year of reporting (from 1997 to 2001) on the quality of ion chromatograms, an additional statistical analysis was applied on the scores assigned by experts. A linear mixed model using the SPSS 12.0 Statistical Software Package (SPSS, Inc., Chicago, IL) was considered including as fixed effects the technique applied and the year of reporting, and, as random effects, the expert and the ion chromatogram.

3. Results and discussion

3.1. Experts' evaluation

One of the objectives of the present work was to assess to what extent experts were homogeneous when evaluating ion chromatograms. This exercise is presumed to be representative of how ion chromatograms are evaluated by EQAS members of Steering Committees. On the other hand, when developing an artificial intelligence application, the machine has to be trained before generating results, and a maximum degree of agreement between experts contributing with their know-how to the training of the machine is needed.

The agreement between different experts was not always good. In Fig. 3, ion chromatograms with a high degree of agreement between experts (examples A and B) and one ion chromatogram in which experts were not concordant (example C), are presented. Ion chromatograms A and B were evaluated as very good (score 4) and as very bad (score 0), respectively, by all experts. A lack of agreement was observed in the evaluation of ion chromatogram C to which different experts assigned scores from 2 to 4.

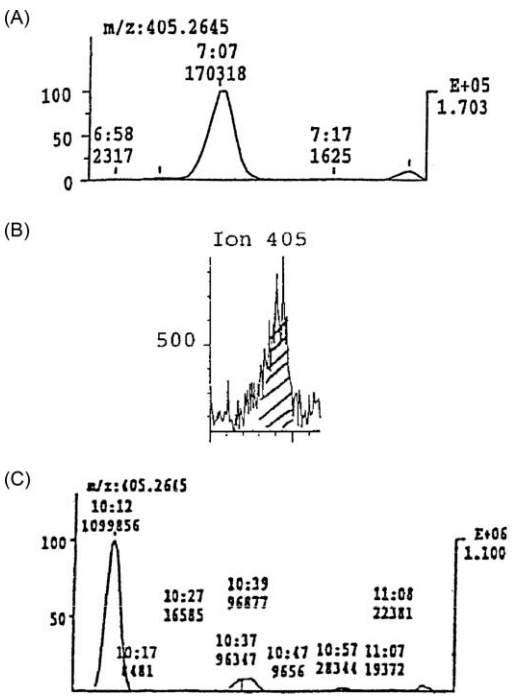

Fig. 3 Examples of three ion chromatograms corresponding to ion at *m/z* 405: (A) ion chromatogram evaluated as very good (score 4) by all experts; (B) ion chromatogram evaluated as very bad (score 0) by all experts; and (C) ion chromatogram in which a lack of agreement was observed in the evaluation of experts (scores assigned from 2 to 4).

The calculation of the mean of scores assigned by each expert showed that ion chromatograms were evaluated as sufficient and good (mean of scores between 2 and 3). This is logical taking into account that laboratories providing ion chromatograms had an extensive experience in the analysis of prohibited substances in sports, such as 19-norandrosterone. Expert number 6 (mean of scores of 1.8) assigned lower scores than the rest of the experts. This expert also presented the highest variance (1.2 *vs.* 0.3–0.5 by the rest of the experts), thus his evaluation differed from the rest of the group. Bias was detected in the whole group of experts that persisted after eliminating expert number 6 ($p < 0.05$), by applying the Cochran's Q-test. Eliminating the next expert presenting the highest variance and so on, a group of experts, constituted by experts 1, 3, 4, 5 and 7, was obtained in which no bias was detected ($p > 0.05$). These results show that evaluation performed by experts 1, 3, 4, 5 and 7 can be considered homogeneous.

Measuring the proportion of agreement for the nine experts (Table 2), concordance was observed (the proportion of agreement higher than 0.8) when considering 'acceptable' ion chromatograms; however, discrepancies (the proportion of agreement around 0.4) were seen when considering 'unacceptable' ones. Eliminating experts according to the results of the Cochran's Q-test (experts 2, 6, 8 and 9), the concordance increased for 'acceptable' ion chromatograms to higher than 0.9 and, for 'unacceptable' to higher than 0.6.

Therefore, all statistical approaches (ANOVA, Cochran's Q-test and the proportion agreement of acceptable records *vs.* unacceptable) showed that experts were quite heterogeneous in evaluating qualitatively analytical records in particular for those which quality is more doubtful. The results obtained allowed the identification of experts 1, 3, 4, 5 and 7 as homogeneous experts in the evaluation of chromatographic data. Only the scores assigned by these experts were used to develop the SVM.

3.2. The use of SVM in the automatic evaluation of ion chromatograms

The second objective of the present work was to develop a SVM able to evaluate ion chromatograms as a human expert but providing a more homogeneous and automatic evaluation.

Table 2 Agreement between experts in the evaluation of ion chromatograms as acceptable or unacceptable

	Proportion of agreement (± 95% confidence interval)	
Experts	Acceptable	Unacceptable
1, 2, 3, 4, 5, 6, 7, 8, 9	0.81 (± 0.01)	0.42 (± 0.05)
1, 2, 3, 4, 5, 7, 8, 9	0.87 (± 0.01)	0.51 (± 0.05)
1, 3, 4, 5, 7	0.92 (± 0.02)	0.64 (± 0.08)

Table 3 Results obtained when assessing the predictive ability of the SVM to classify ion chromatograms according to score (0: very bad, 1: bad, 2: sufficient, 3: good, 4: very good)

	SVM					
Experts' evaluation	0	1	2	3	4	Total
0	0	2	2	0	0	4
1	1	4	5	1	0	11
2	1	3	20	7	1	32
3	0	2	11	20	3	36
4	0	0	3	6	13	22
Total	**2**	**11**	**41**	**34**	**17**	**105**

The predictive ability of the developed SVM was evaluated according to its capacity to classify an ion chromatogram with reference to the experts' evaluation (0, 1, 2, 3, or 4, considering each score as one class) and, also according to its ability to correctly classify an ion chromatogram as either 'acceptable' or 'unacceptable', again using the experts' evaluation as the point of reference. In the first case, the predictive ability of the machine was around 54% (see Table 3) and, in the second case, the predictive ability of the SVM increased to 87% (see Table 4) and was similar to the percentage of agreement obtained between homogeneous experts (93%). Nevertheless, the weighted kappa coefficients obtained when considering five categories (value = 0.49) and two categories (value = 0.42) indicated that there was a similar and moderate agreement in both cases.[44] The highest weighted kappa obtained in the case of five categories was due to the weighted kappa assigning less weight to the agreement as the categories are further apart. A paradox was found when considering two categories (acceptable and unacceptable): 87% of agreement between SVM and the homogeneous experts but, according to the kappa value (0.42), a moderate agreement only existed. A possible explanation to this paradox might be the lower number of unacceptable ion chromatograms when comparing to the number of acceptable ones.[44] As, previously stated, all laboratories providing ion chromatograms had extensive experience in the analysis of prohibited substances in sports, such as 19-norandrosterone. In order to increase the level of agreement between experts and the SVM, future training of the machine with a larger number of unacceptable ion chromatograms should be considered.

To assess the absence of bias between the homogeneous group of experts and the SVM developed when evaluating the acceptability or unacceptability of the ion chromatograms, the Cochran's Q-test was applied. A $p > 0.05$ was obtained, indicating that no bias existed. This result corroborated the level of concordance existing between the SVM developed and the homogeneous group of experts.

Table 4 Results obtained when assessing the predictive ability of the SVM to classify ion chromatograms according to class ('acceptable' or 'unacceptable')

	SVM		
Experts' evaluation	Unacceptable	Acceptable	Total
Unacceptable	7	8	15
Acceptable	6	84	90
Total	**13**	**92**	**105**

According to these results, it was concluded that the evaluation performed by the SVM was close to the reasoning process followed by a human expert evaluating chromatographic data. Then, it would be difficult to elucidate if a given result was obtained by a human expert or by the SVM.

The SVM developed had the potential of providing a scoring system for sorting laboratories according to the quality of their ion chromatograms. This could be very useful for participating laboratories in EQAS in order to compare the quality of their results to the ones reported by others, to detect and define problems affecting the quality of analytical results.

3.3. Factors influencing the quality of ion chromatograms

According to the scores assigned by the homogeneous experts, differences in the quality of the ion chromatograms were observed depending on the GC–MS technique used and depending on the year of reporting the ion chromatogram ($p < 0.05$, by applying a linear mixed model; please refer to Calculations, section 2.3). A distribution of the mean of scores assigned by the experts according to the technique used and the year of reporting is presented in Fig. 4.

Ion chromatograms obtained using HRMS (mean of 3.4) had better scores than the ones corresponding to MS-MS (mean of 2.4) and the ones corresponding to LRMS (mean of 2.0). Although these differences were expected taking into account the characteristics of the three analytical techniques and the low concentrations of 19-norandrosterone present in the samples, this is an example on how this type of analysis can help to elucidate other factors for improving the quality of results.

With reference to the distribution of the mean of scores according to the year of reporting (referring to the annual re-accreditation test usually performed by the end of the year), there are some trends showing small but significant differences. The most important result was the change in the mean score values from 1998 to 1999 (2.0 *vs.* 2.9). Scores assigned to ion chromatograms belonging to years 1997 and 1998 were statistically different when compared to those assigned to ion chromatograms belonging to 1999, 2000 and 2001 ($p < 0.05$,

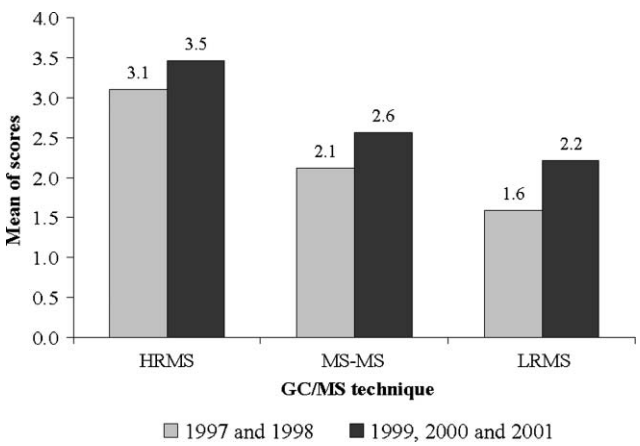

Fig. 4 Mean of scores assigned by experts to the set of ion chromatograms according to the GC–MS technique used and the year of reporting ($p < 0.05$, statistical differences between years for each technique).

by applying a linear mixed model). Concentrations of samples analyzed in these five years were similar. In August 1998, the IOC Medical Commission proposed analytical criteria for reporting low concentrations of five anabolic steroids, including 19-norandrosterone,[45] and constrained laboratories to detect these five anabolic steroids at low concentrations. This guideline could explain the improvement of the analytical methodology used to detect 19-norandrosterone at low concentrations and then the improvement in the quality of results.

4. Conclusion

In the evaluation of qualitative analytical data, experts show a high degree of variability when judging ion chromatograms, especially for poor-quality data. Despite this heterogeneity, there is general agreement in that data provided by laboratories are also highly variable in terms of quality. Independent of the fact that all laboratories were able to find substances in control urines submitted (*i.e.* 19-norandrosterone), they would further benefit from the evaluation of qualitative data, thus providing a scoring system allowing comparisons among laboratories. These two objectives can be attained using machine-learning algorithms.

A SVM has been developed for the evaluation of ion chromatograms. It has been shown that it is close to the reasoning process followed by a human expert. It has the potential of providing more homogeneous evaluation results than randomly selected experts and to perform evaluations automatically. The SVM could provide a scoring system for sorting laboratories as a function of the quality of their ion chromatograms. The qualitative evaluation of analytical records using a scoring system allows the identification of the main factors affecting the quality of chromatographic analytical data, such as the analytical technique applied and the availability of guidelines for reporting positive results.

Acknowledgements

This research has been partially supported by European Commission grant ALADIN 2002: Analytical Laboratories for AntiDoping control: International Network for External Quality Assessment (contract number G7RT-CT-2000-05022). The authors would like to thank Dr David Cowan and his collaborators (King's College, London) for the compilation of data used in this study, the doping experts that reviewed the data from the Aker Hospital (Oslo), the German Sports University (Cologne), King's College, London and the Institut Municipal d'Investigació Mèdica (Barcelona) and also Dr Klaus Langohr for his statistical assistance. The authors also wish to acknowledge their colleagues from GREC (Research Group on Knowledge Engineering) for helpful discussions and suggestions and also to the members of the ALADIN consortium: David Cowan (Drug Control Center, King's College, London); Wilhelm Schänzer, Georg Opfermann and Gerd Sigmund (Deutsche Sporthochschule Köln, Institut für Biochemie, Cologne); Peter Hemmersbach and Hanne Lund (Hormone Laboratory, Aker Hospital, Oslo); John Williams (Cardiff Bioanalytical Services, Cardiff); Ferran Sanz, Carlos Diaz and Juan Antonio de los Cobos (Unitat de Recerca en Informàtica Biomèdica, Institut Municipal d'Investigació Mèdica, Barcelona); Magi Farre (Unitat de Recerca en Farmacologia, Institut Municipal d'Investigació Mèdica, Barcelona).

References

1 ISO 17025 International Organization for Standardization, ISO/IEC 17025: *General Requirements for the Competence of Testing and Calibration Laboratories*, ISO, Gèneve, 2005.
2 C. Jimenez, R. Ventura, J. Williams, J. Segura and R. de la Torre, *Analyst*, 2004, **129**, 449–455.
3 V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
4 C. Cortes and V. Vapnik, *Machine Learning*, 1995, **20**, 273–297.
5 V. N. Vapnik, *Statistical Learning Theory*, Wiley, New York, 1998.
6 C. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, New York, 2000.
7 C. Angulo and A. Català, *Lect. Notes Comput. Sci.*, 2000, **1810**, 55–64.
8 C. Angulo and A. Català, in *Connectionist Models of Neurons, Learning Processes and Artificial Intelligence. Proceedings of the 6th International Work-Conference on Artificial and Natural Neural Networks IWANN 2001, Part I*, ed. J. Mira and A. Prieto, Springer-Verlag, London, 2001, Lecture Notes in Computer Science, vol. 2084, pp. 661–668.
9 Y. Xu, S. Zomer and R. G. Brereton, *Crit. Rev. Anal. Chem.*, 2006, **36**, 177–188.
10 J. H. Oh, A. Nandi, P. Gurnani, L. Knowles, J. Schorge, K. P. Rosenblatt and J. X. Gao, *J. Bioinform. Comput. Biol.*, 2006, **4**, 1159–1179.
11 F. M. Smith, W. M. Gallagher, E. Fox, R. B. Stephens, E. Rexhepaj, E. F. Petricoin, L. Liotta, M. J. Kennedy and J. V. Reynolds, *Ann. Surg.*, 2007, **245**, 259–266.
12 J. K. Yu, S. Zheng, Y. Tang and L. Li, *J. Zhejiang Univ. Sci.*, 2005, **6**, 227–231.
13 J. K. Yu, Y. D. Chen and S. Zheng, *World J. Gastroenterol.*, 2004, **10**, 3127–3131.
14 L. Li, H. Tang, Z. Wu, J. Gong, M. Gruidl, J. Zou, M. Tockman and R. A. Clark, *Artif. Intell. Med.*, 2004, **32**, 71–83.
15 M. Wagner, D. Naik and A. Pothen, *Proteomics*, 2003, **3**, 1692–1698.
16 D. Agranoff, D. Fernandez-Reyes, M. C. Papadopoulos, S. A. Rojas, M. Herbster, A. Loosemore, E. Tarelli, J. Sheldon, A. Schwenk, R. Pollok, C. F. Rayner and S. Krishna, *Lancet*, 2006, **368**, 1012–1021.
17 R. Willingale, D. J. Jones, J. H. Lamb, P. Quinn, P. B. Farmer and L. L. Ng, *Proteomics*, 2006, **6**, 5903–5914.
18 J. Prados, A. Kalousis, J. C. Sanchez, L. Allard, O. Carrette and M. Hilario, *Proteomics*, 2004, **4**, 2320–2332.
19 S. Ruepp, F. Boess, L. Suter, M. C. de Vera, G. Steiner, T. Steele, T. Weiser and S. Albertini, *Toxicol. Appl. Pharmacol.*, 2005, **207**, 161–170.
20 S. K. Thukral, P. J. Nordone, R. Hu, L. Sullivan, E. Galambos, V. D. Fitzpatrick, L. Healy, M. B. Bass, M. E. Cosenza and C. A. Afshari, *Toxicol. Pathol.*, 2005, **33**, 343–355.
21 G. Steiner, L. Suter, F. Boess, R. Gasser, M. C. de Vera, S. Albertini and S. Ruepp, *Environ. Health Perspect.*, 2004, **112**, 1236–1248.
22 M. West-Nielsen, E. V. Hogdall, E. Marchiori, C. K. Hogdall, C. Schou and N. H. Heegaard, *Anal. Chem.*, 2005, **77**, 5114–5123.
23 A. I. Belousov, S. A. Verzakov and J. Von Frese, *J. Chemom.*, 2002, **16**, 482–489.
24 J. A. Fernández Pierna, V. Baeten, A. Michotte Renier, R. P. Cogdill and P. Dardenne, *J. Chemom.*, 2004, **18**, 341–349.
25 S. Zomer, R. G. Brereton, J. F. Carter and C. Eckers, *Analyst*, 2004, **129**, 175–181.
26 M. Song, C. M. Breneman, J. Bi, N. Sukumar, K. P. Bennett, S. Cramer and N. Tugcu, *J. Chem. Inf. Comput. Sci.*, 2002, **42**, 1347–1357.
27 N. Tugcu, M. Song, C. M. Breneman, N. Sukumar, K. P. Bennett and S. M. Cramer, *Anal. Chem.*, 2003, **75**, 3563–3572.

28 M. Donike, in *First International Symposium on Current Issues of Drug Abuse Testing*, ed. J. Segura and R. de la Torre, CRC Press, Boca Raton, 1992, ch. 34, pp. 225–237.
29 L. L. Engel, J. Alexander and M. Wheeler, *J. Biol. Chem.*, 1958, **231**, 159–165.
30 W. Schänzer, *Clin. Chem. (Washington, D. C.)*, 1996, **42**, 1001–1020.
31 W. Schänzer, in *Drug Abuse Handbook*, ed. S. B. Karch, CRC Press, Boca Raton, 1998, ch. 9.4, pp. 671–689.
32 J. Marcos, J. A. Pascual, X. de la Torre and J. Segura, *J. Mass Spectrom.*, 2002, **37**, 1059–1073.
33 C. Ayotte, D. Goudreault and A. Charlebois, *J. Chromatogr., B: Biomed. Appl.*, 1996, **687**, 3–25.
34 L. D. Bowers, D. A. Armbruster, T. Cairns, J. T. Cody, R. Fitzgerald, B. A. Goldberger, D. Lewis and L. M. Shaw, *Gas Chromatography/Mass Spectrometry (GC/MS) Confirmation of Drugs; Approved Guideline*, NCCLS document C43-A, NCCLS, Pennsylvania, 2002.
35 H. Du and M. J. Stillman, *Anal. Chim. Acta*, 1997, **354**, 65–76.
36 A. Skoog, F. J. Holler and T. A. Nieman, *Principios de Análisis Instrumental*, McGraw-Hill Interamericana de España, S.A.U., Madrid, 5th edn, 2001, pp. 737–752.
37 C. W. Hsu and C. J. Lin, *IEEE Trans. Neural Netw.*, 2002, **13**, 415–425.
38 Optimization Toolbox User's Guide – Version 3, The MathWorks, Inc., Natick, MA, USA, 2004.
39 G. Upton and I. Cook, *A Dictionary of Statistics*, Oxford University Press, Oxford, 2002.
40 R. de la Torre, A. Domingo-Salvany, R. Badia, G. Gonzalez, D. McFarlane, L. San and M. Torrens, *Clin. Chem. (Washington, D. C.)*, 1996, **42**, 1433–1438.
41 J. M. Grant, *Lancet*, 1991, **337**, 215–218.
42 J. Cohen, *Ed. Psychol. Meas.*, 1960, **20**, 37–46.
43 J. Cohen, *Pshycol. Bull.*, 1968, **70**, 213–220.
44 A. J. Viera and J. M. Garrett, *Family Med.*, 2005, **37**, 360–363.
45 *Analytical criteria for reporting low concentrations of anabolic steroids*, International Olympic Committee, Lausanne, August 1998.