VALIDATION AND QUALITY ASSURANCE OF A BROAD SCALE GAS CHROMATOGRAPHIC SCREENING METHOD FOR DRUGS

Merja GERGOV, Jari NOKUA, Ilpo RASANEN, Ilkka OJANPERÄ
Department of Forensic Medicine, University of Helsinki, Helsinki, Finland

ABSTRACT: Validation and quality assurance procedures are described for a gas chromatographic screening method for 104 basic drugs in blood. The instrument validation included repeatability of injection, retention times, retention indices and peak areas. The validation of quantitative measurements consisted of linearity, accuracy, intra-assay precision, limits of quantitation and estimation of the uncertainty of measurement. In quality assurance, preventive maintenance of the instrument, calibration and internal and external quality control were performed regularly. The method was accredited by Finnish Accreditation Service (FINAS) in 1997.

KEY WORDS: Validation; Quality assurance; Gas chromatography; Drug screening.

Problems of Forensic Sciences, vol. XLIII, 2000, 70–78
Received 9 September 1999; accepted 16 May 2000

INTRODUCTION

The demands of reliability and tracebility of analytical results in forensic toxicology set high requirements to the methods used. The suitability and limits of the used methods and instruments must be known for the correct interpretation of the acquired data. Validation and proper maintenance of the system is a key to high quality and accreditation. This paper describes an example of validation and quality assurance of a GC screening and quantitation method for 104 basic drugs in autopsy blood.

The extent of validation and quality assurance for a method depends on the analytical technique and the particular application. Thus, all of the various procedures are usually not needed, but the researchers can focus on those, which have the greatest effect on the final analysis results [1].

EXPERIMENTAL

Materials

The retention index standards (N,N-dialkyl-4-fluoroanilines) were synthesized in our laboratory. The internal standard (dibenzepin) and basic drug substances were obtained from various pharmaceutical companies and they were of pharmaceutical purity.

Instrumentation

The gas chromatograph was a Micromat HRGC 412 (HNU-Nordion, Helsinki, Finland) equipped with two nitrogen/phosphorus detectors and a CTC A2000S autosampler. DB-5 and DB-1701 columns (15 m x 0.32 mm x 0.25 μm) were used as analytical columns and deactivated fused silica (10 m x 0.32 mm) was used as pre-columns. The retention index standard solution was taken into the injection needle prior to the sample by the autosampler. The data handling and reporting were performed with SC-Workstation 3.0 software supplied by Sunicom (Helsinki, Finland).

Analytical methods

1 μg of internal standard (dibenzepin) was added to 1 g of blood samples, which were extracted with 0.3 ml of butyl acetate at pH 9.3 according to the previously described method [2]. During the gas chromatographic runs, the injector and detector temperatures were 270°C and 290°C, respectively. The oven temperature program was as follows: 70°C (0.7 min) → 20°C/min → 140 °C (0 min) → 10°C/min → 290°C (9.5 min). The carrier gas (He) flow rate was 2 ml/min (70°C, DB-5) and hydrogen and air flow rates were 1 ml/min and 80 ml/min, respectively.

In qualitative analysis, seven retention index standards were co-injected with each sample (Figure 1) and the retention index standard values were calculated for every compound for both columns. Calculated index values from samples were compared with values from calibration standards in the library, and a summary report was created based on hits on both columns (Table I).

Quantitative analysis was performed by determining the relative response factors related to dibenzepin for all compounds. In the summary report, the concentrations for all identified compounds were listed separately for both columns.

Fig. 1. Chromatograms from a control sample containing chlorprothixene, mirtazapine, tramadol and doxapram. Retention index stardards are marked with numbers 1–7.

RESULTS AND DISCUSSION

Validation

To describe the repeatability of injection of the instrument, 10 serial injections were performed. The repeatability of absolute retention times, retention indices and peak areas were measured for amitriptyline, doxepin, clozapine, thioridazine, strychnine and dibenzepin (internal standard). Injection reproducibility for the instrument was very good. The coefficient of variation for the retention indices was lower than 0.01% and for the absolute peak areas lower than 9%.

TABLE I. SUMMARY OF THE CONTROL SAMPLE RESULTS FROM COLUMNS DB-5 (Ch. 1) AND DB-1701 (Ch. 2).

*** SC-Compare Report (Version 1.30) ***

Data file: 1ek0308x.dta
Method: basic.MTD
Date created: Wed Aug 4 1999 at 03:19:44
Date analyzed: Thu Aug 19 1999 at 10:31:14

Compound	Ch.	Peak	AbsRT	IdPara	Diff	Area	µg/g
Tramadol	1	5	9.845	623.49	0.399	2863	0.623
	2	5	10.238	652.59	0.434	3920	0.642
Mirtazapin	1	6	12.813	784.37	0.759	2185	0.237
	2	8	13.267	824.64	0.866	3301	0.248
Chlorprothixene	1	9	15.117	920.78	0.719	836	0.223
	2	9	15.403	956.80	0.628	1398	0.281
Dibenzepin	1	8	14.725	897.30	0.351	7356	1.000
(internal standard)	2	10	15.687	974.33	1.017	10700	1.000
Doxapram	1	11	18.310	1122.13	1.594	6689	0.996
	2	13	19.408	1220.43	0.105	9485	0.975

Target values: tramadol 0.5 µg/g, mirtazapin 0.2 µg/g, chlorprothixene 0,2 µg/g and doxapram 1,0 µg/g. Internal standard: dibenzepin – 1.0 µg/g.

Linearity was demonstrated by spiking blood samples with five concentration levels and four parallels. The correlation coefficients (R^2) were calculated for all compounds. The calibration curves showed satisfactory fit to linear regression, and the correlation coefficients (R^2) were mainly better than 0.98 (Figure 2). In routine use, the calibration was done with one point only. The stability of quantitative calibration had to be tested as it would be too time-consuming to perform in every sequence. The test was done by analyzing calibration standards as samples one month after calibration, and the inaccuracy was calculated. For most of the compounds, the calibration was proved to be stable for at least one month. For those compounds, like chloroquine, maprotiline and fluvoxamine, which were not reproducible, the calibration was in routine use always done immediately before quantitation.

The obtained data was also used for measuring intra-assay precision and for establishing the quantitation limits. In accuracy determination, results from calibration samples and from the inter-laboratory studies were also used, as well as the data from the sets of spiked blood samples used for the linearity studies. We established the limits of quantitation using as the criteria a 20% maximum acceptable standard deviation and a signal to noise better than 3. The limits for quantitation varied depending on the compound and were mainly between 0.1 and 0.2 µg/g. The basic validation results for the most commonly found drugs are listed in Table II.

TABLE II. VALIDATION RESULTS FOR THE MOST COMMON DRUG FINDINGS IN AUTOPSY BLOOD SAMPLES IN FINLAND

Drug	LOQ [μg/g]	STDEV [%]	Bias [%]	Correlation coefficient DB-5	Uncertainty [%]
Dextropropoxyphen	0.1	8	9	0.998	13
Amitriptyline	0.1	7	5	0.992	14
Levomepromazine	0.1	10	16	0.999	19
Doxepin	0.1	11	13	0.989	17
Promazine	0.1	8	7	0.996	11
Zopiclone*	0.1	9	38	0.974	41
Chlorprothixene	0.1	5	2	0.988	10
Thioridazine	0.1	9	13	0.997	20
Diltiazem	0.1	10	2	0.994	11
Citalopram	0.1	7	3	0.996	12
Melperone	0.2	9	8	0.961	23
Clozapine	0.1	5	11	0.991	13
Trimipramine	0.1	8	5	0.989	11
Verapamile	0.1	5	13	0.993	18
Chlorpromazine	0.1	7	8	0.997	15
Mianserine	0.1	8	7	0.986	15
Chloroquine	0.2	14	11	0.982	19
Clomipramine	0.1	7	9	0.988	15
Fluoxetine	0.2	8	11	0.972	18

* Zopiclone is routinely analyzed by a separate method for benzodiazepines and related drugs.

Accredited laboratories are expected to know the uncertainty of measurement (U) for each quantitatively measured substances. U consists of two elements, systematic and random errors. There are many approaches for determining U, and the estimation has been considered difficult, because of a large number of variables included. However, according to the recommendations given by the Finas, U can be approximated even by "using the judgement of an experienced chemist".

An adequate and practical approach was to measure the systematic and random errors in a very simple way from the existing data. In estimating the systematic errors, the results of calibration standards (analyzed as samples) were used. In addition, a mixture of 36 most common drugs in two concentration levels was analyzed with two parallels and the deviation from the theoretical values were calculated (inaccuracy). The inaccuracy was sometimes positive and sometimes negative and therefore, the average bias of absolute values described the method best. Random error was estimated from the standard deviations of calibration runs, spiked samples in linearity studies, and of separate mixtures of 25 common compounds at two concentration levels. The judgement of our experienced chemist was also used.

The uncertainty of measurement was calculated from the following equation: U = SQRT (inaccuracy^2 + random error^2). We obtained very different values for separate drugs, varying from 10% to 25%. In our investigations of autopsy samples, the obtained accuracy is good enough, while the demand is to distinguish therapeutic, toxic and lethal concentrations of toxicants.

QUALITY ASSURANCE

The quality assurance consists of four elements: preventive maintenance of the instrument, calibration, internal and external quality controls. To ensure the performance of the gas chromatograph, an instrument check sample was run daily containing amitriptyline, dextropropoxyphen and venlafaxine at detection limit levels. Peak areas must exceed the previously set criteria. A contamination check, which is the solvent

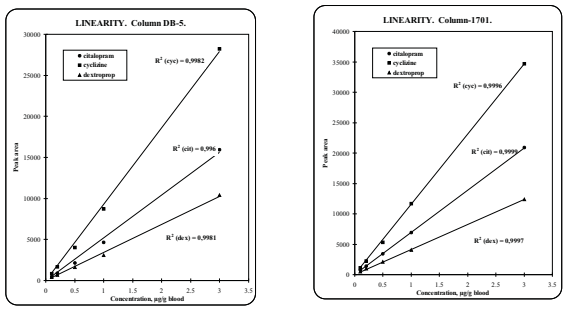

Fig. 2. Calibration curves and correlation coefficients for citalopram, cyclizine and dextropropoxyphen.

(butyl acetate) with retention indices, was also run. The liner, septum and silanized glass wool were changed once a week. The use of pre-columns diminished the contamination of the columns and therefore extended the life of the column, and weekly cutting of 40 cm from the front of the pre-columns weekly improved the peak shapes.

For calibration, fresh stock solutions of all the 104 drugs were made once a year. From these stock solutions, working solutions were diluted every three months. Twenty five calibration samples were prepared by spiking the zero blood with 3–5 working solutions and carrying out the sample preparation as for samples. Calibration was performed every month using one point calibration with two parallels.

The internal quality control included a control sample, which was run once a week. The contents of the control sample mixture was varied every year. The intention was not only to go through the most common drugs in a reasonable period of time, but also to get a picture of the performance of the total analysis. Therefore, compounds are selected to be representatives of different drug groups and of different chromatographic behavior, being thus susceptible to disturbances in various stages of the analysis. At present the

control sample contains chlorprothixene, mirtazapine, tramadol and doxapram at two concentration levels. The alarm and action limits are 20% and 50% of the theoretical value (Figure 3). The long-term (one year) standard deviation for chlorprothixene was 18%, and for mirtatzapine, tramadol and doxapram it was 12%, 16%, and 11%, respectively.

Running these samples only once a week is a question of limited time of the instrument and of the chemist, who is to interpret the results. The concentration limits of the alarm and action limit were set quite high, compared to the limits of targeted compound analysis, keeping in mind that very accurate results with several decimals are not needed.

For external quality control, the laboratory takes part in two proficiency testings organized by the National Institute of Forensic Toxicology in Oslo, Norway. Nordquant is a quantitative test, which is organised twice a year with 12 participants from forensic laboratories in Nordic countries and England, Ireland and the Netherlands. The whole blood sample is spiked with 13–15 therapeutic drugs and drugs of abuse. Nordscreen is a qualitative test, which is organised twice a year with 8 participants from forensic laboratories in Nordic countries only. The sample is autopsy blood from an authentic case containing varying drugs and drugs of abuse.

In the quantitative test (Nordquant) There are only a few compounds included and they have not been changed over the years. During 1997 to 1999, the results of five Nordquant tests have been reported. Methadone, dextropropoxyphen, amitriptyline, diazepam and levomepromazine were analyzed with the present method. The z-scores (observed value – median value of the participants/standard deviation) obtained were generally very good (Table III).

Neither comprehensive interlaboratory studies nor certified reference standards were available for all compounds. Therefore, a lot of effort is required from internal control. This includes e.g. changing the composition of the control sample regularly, and a careful preventive maintenance of the instrumentation, because the instrument appears to be the most probable cause for interruptions in the operation.

TABLE III. Z-SCORES IN THE NORDSCREEN INTERLABORATORY STUDIES IN 1997–1999

Compound	Z-scores for basic drugs, n = 12				
	1/99	2/98	1/98	2/97	1/97
Methadone	–0.11	0.00	1.84	1.22	–0.26
Dextropropoxyphen	0.28	0.43	0.35	0.89	–0.19
Amitriptyline	0.92	0.63	0.16	1.04	0.54
Diazepam	–0.39	1.28	–0.12	0.18	0.45
Levomepromazine	ND	0.52	ND	ND	ND

In accreditation, the criteria to accept the performance of the method is set by the laboratory itself and not by the assessors. Thus, for example quite high uncertainty of

quantitative measurement can be acceptable, when very accurate results are not required. Because of the great number of compounds and the fact that more than twenty

Fig. 3. Control chart for mirtazapine, target value 1.0 μg/g blood.

samples must be analyzed every day, the assessors accepted the schedule including a monthly calibration and a weekly control sample, and the limitations in the interlaboratory tests. The method has been in successful use for several years and was accredited in 1997 by Finas.

References:

1. Green J. M., A practical guide to analytical method validation, *Analytical Chemistry, News and Features* 1996, pp. 305–309.
2. Rasanen I., Ojanperä I., Vartiovaara J,. Vuori E., The advantage of dual-column approach and retention indices combined with refined reporting in gas chromatographic drug screening, *Journal of High Resolution Chromatography* 1996, vol. 19, pp. 313–322.