Microchemical Journal 207 (2024) 111578

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Reducing manual targeted LC-MS/MS peak integration using a supervised learning peak evaluation and automated review tool

Martin Alewijn *, Sjors Rasker, Dieke van Doorn, Marco Blokland

Wageningen Food Safety Research, WFSR, Akkermaalsbos 2, 6708 WB Wageningen, the Netherlands

ARTICLE INFO

Keywords:
Targeted mass spectrometry
Automation
Computer-assisted
Quality assurance
Peak integration

ABSTRACT

Background: Our laboratory analyses over 100,000 samples for multiple analytes yearly. Correctly annotating and quantifying analytes using targeted MS/MS analysis is crucial for most laboratories. Therefore, mass spectrometric instrument manufacturers include software capable of automatic chromatographic peak detection and integration, which is used for most routine applications. Although this generally works well, mistakes such as accidentally selecting a nearby matrix peak or drawing an incorrect baseline are relatively common. Especially when results are to be used for enforcement, a time-consuming manual review of each integrated peak is still required to obtain reliable results. This work aims to provide a tool that can significantly reduce the manual workload of reviewing peak integration, thereby reducing the time used for manual review while ensuring that errors made by automatic integration can be corrected by human experts.

Results: Peak Evaluation and Automated Review tool, or PEAR review, is a machine learning-type tool that can read automatic integrations from various brands of MS equipment and compare them with a set of examples stored in a database of correct peak integrations provided by analysts that are relevant for the type of analysis. Moreover, the automatic review process checks all available ion transitions for a target compound. With those ingredients, the tool can autonomously decide how a peak should be quantified or whether a human expert should review it. The developed tool was tested on routine data processed with a widely used vendor-specific software, and we found that 85% of all chromatograms were handled automatically by this tool. Only the remaining 15% needed a 'conventional' manual review. The qualitative and quantitative performance of the *PEAR* tool was found to be equivalent to that of expert human integration, underlining its reliability.

Significance: Our findings indicate that 85% of all manual integration checks can be skipped using *PEAR*. This reduces the often tedious workload of reviewing all peaks in multiple chromatograms while offering the same quality as full human intervention.

1. Introduction

Monitoring various analytes, such as veterinary drugs, growth promoters, pesticides, and mycotoxins in foods, is performed in food control laboratories within the European Union to maintain a safe food supply [1]. Liquid- and gas chromatography coupled to mass spectrometry (LC/GC–MS) are ubiquitous techniques for analysing these substances in complex matrices [2–4]. Through continuous development over time, these techniques can now detect many substances in a single sample extract at relevantly low concentration levels in a targeted manner. For example, detecting more than 100 analytes in a single chromatographic run is becoming increasingly common [e.g. 5,6,7]. In LC-MS/MS residue analysis for food control, two or more ion transitions for each analyte are usually monitored for identification purposes [8]. This means that more than 200 target peaks can be generated for a single sample, depending on the method. Other domains where targeted LC-MS/MS is used, like the forensic, doping, metabolomics and to some extent also clinical and pharmacological domains will encounter similar situations.

The introduction of high-throughput or even automated sample preparation methods has made it possible to routinely include up to around 100 samples, controls, and calibrators in a single analytical run when using, for example, 96-well plate solid phase extraction (SPE) for cleanup [e.g. 9]. This example of a manual, routine, targeted analysis of 100 analytes in 96 samples already results, in theory, more than 19,000 target peaks for a single analytical run. This number is only expected to increase now that high-resolution mass spectrometry (HR-MS) is

* Corresponding author.
E-mail address: martin.alewijn@wur.nl (M. Alewijn).

https://doi.org/10.1016/j.microc.2024.111578
Received 22 June 2024; Received in revised form 28 August 2024; Accepted 4 September 2024
Available online 7 September 2024

becoming more commonplace in food control laboratories since many more compounds can be analysed simultaneously compared to LC-MS/MS [10].

Instrument manufacturers ship their instruments with software capable of chromatographic peak detection and integration, which is used for most routine applications (e.g. TargetLynx [11], Masshunter [12]). For optimal peak detection, integration parameters are optimised during method development. In residue analysis, they are optimised to detect each (small) peak present, even those close to a signal-to-noise ratio of 3 and more difficult baseline drift, shoulders, fronting, and tailing peaks. Although automatic integration works well for most chromatograms, it is still required for analysts to manually check each chromatogram for peaks and their integration because mistakes such as accidental selection of a nearby matrix peak or drawing an incorrect baseline are common enough that the included peak integration software cannot be trusted blindly.

Therefore, in fields where the occurrence of both false positives and false negatives should be minimised and accurate quantification is essential, for example, in residue analysis in the food safety field, the most common approach currently relies on human review. Due to the enormous number of chromatograms generated with modern analytical techniques, manual peak review is taking up an increasingly large amount of time.

Multiple data processing platforms have been published as alternatives to the software packages included by instrument manufacturers and to provide access to more advanced or specialised data processing workflows. Kensert et al. [13] trained a convolutional neural network (CNN) using simulated chromatograms that outputs locations, probabilities, and areas of multiple peaks in a single chromatogram. Melnikov et al. [14] used an algorithm to detect the region of interest (ROI) and trained two CNNs to classify the ROI and perform peak integration. Rupprecht et al. [15] developed an algorithm capable of peak detection, integration, and quantification using internal standards and calibration samples. Other platforms, such as XCMS, perform peak detection and integration before further processing the data for metabolomics [16–18]. In these examples, data conversion is often needed before processing as these platforms cannot always read the data in vendor-specific formats [14,15]. Others have been developed with metabolomics in mind [16–18], meaning that they are designed around feature detection in HR-MS data and subsequent data clustering steps. Using summation peak integration has also been proposed [19]; however, this approach still requires human review of all chromatograms. Especially when running methods on multiple instruments, columns, etc. where retention times are known to shift, and interfering peaks can occur within the set integration window.

All the examples start with raw data, performing peak detection, peak integration and further processing steps using different approaches, often completely circumventing the vendor-specific software. In contrast, for the routine targeted analysis workflows considered in this work, the software should be as user-friendly and streamlined as possible and fit within often pre-existing workflows and quality assurance approaches. Most crucially, error-free performance is not guaranteed for methods relying on parametric integration (standard in current vendor software packages) or entirely on machine learning.

Inspired by image recognition software, we imagined that a software tool could partly replace the manual inspection by analysts of all automatically integrated peaks. The software tool should partly take over the analyst's role during manual review of integrated peaks, thereby reducing the workload and saving time. The portion of peaks that is processed automatically should be equally or even more reliable than manual inspection. A portion of the peaks that cannot be processed automatically could still need human review, which admittedly [20] are not free from errors. At least the reduced workload should keep concentration up.

For this purpose, in this study a novel software tool was developed based on machine learning: *PEAR* (Peak Evaluation and Automated Review), which assists the current human review process of automatically integrated peak by automatically reviewing series by only asking for human intervention for peaks in which it is not absolutely sure of the quality of the integration. The developed *PEAR* tool was evaluated using integrated peaks of widely used vendor-specific software packages designed around LC-MS/MS data processing (e.g., MassLynx (Waters), TraceFinder (ThermoFisher Scientific)) for the initial data processing, followed by semi-automatic peak review using *PEAR*.

2. Development of *PEAR-review*

PEAR-review was designed based on the following five key elements:

(1) *Operates non-vendor specific*. Different vendors have software suites integrated tightly into the current routine workflow, including quality control checks, traceability, and accreditation. They take different approaches to integrate peaks, but none of them uses a tool that reviews the automatically integrated peaks with a database of company-specific correctly integrated peaks to assess the quality of the integrated peaks. We decided to build such a comparison tool called *PEAR*, which can work on exported data, including integrated peaks from each vendor's software suite. The portable document format (pdf) export function was used, as this format is universal and sufficiently open for further processing. It only requires that the vendor software can export the relevant chromatogram sections with sufficient resolution.

(2) *Learn correct integration from human expert examples.* The idea behind this thought is that MS chromatography is a repeatable technique. If the sample, target compound, and analytical conditions are identical, the same chromatogram around the target-specific retention time will be generated with very little variation. Such a similar chromatogram segment should be integrated similarly. So, we developed *PEAR* that scans the relevant peak time segments and evaluates the section around that peak using a database of known peak integrations. Suppose that segment looks sufficiently similar to a previous case in a company-specific database that was integrated by a human expert. In that case, that segment can be assessed and, if needed, re-integrated without any human intervention following the example in the database.

(3) *Combine multiple ion transition chromatograms per component during evaluation.* Targeted (MS/MS) analyses typically use at least two ion transitions per compound to confirm identity. An ion ratio is calculated after integration, and identity is only confirmed if that ratio is within a specific range. However, the shape of the separated peaks is not automatically compared, and therefore, in the current manual review process, the ion transitions need to be manually reviewed separately even if the ion criteria are fulfilled, which is relatively inefficient. Therefore, *PEAR* checks if the peak segments and the integrated peaks for all ion transitions for a certain compound match sufficiently on retention time and ion-ratio but also peak shape and rejects all chromatogram segments without a sufficient match, as that proves that the compound of interest is not present and needs no further (human) attention.

(4) *Identify if possible conflicting (neighbouring) peaks are present.* One of the biggest challenges for correct integration by vendor software is cases where peaks elute very closely together. In targeted analyses, a retention time window is pre-defined, and in most cases, the peak closest or most intense to the median time window is selected and integrated. If the nearest observed peak is not the target compound, there is a risk that the target is missed entirely or attributed to the wrong signal and incorrectly quantified. This phenomenon challenges human integration and automated tools even more. Therefore, *PEAR* identifies possible peaks within a pre-defined time window, enabling the human expert to check

compounds with more than one possible peak within the range. Moreover, the retention time deviation per compound across positively identified peaks in the analytical series is added, allowing clear clues in which signals to pay extra attention are offered for manual review by the analyst.

(5) *Allow human integration. PEAR* intends to automate what can be automated and present any automatic integrated peaks that are not automatically interpreted with 100 % certainty by *PEAR* to an analyst. By doing so, the number of components to review manually should be drastically reduced, but the remainder should still be reviewed in a user-friendly manner.

3. Implementation of *PEAR-review*

The workflow of *PEAR* is presented in Fig. 1. First, data is imported from the pdf file generated by the vendor software. Then, each chromatogram section, one of each monitored transition for each target compound in each sample, is subjected to four subsequent questions by *PEAR*: is a peak found → is the result above a pre-defined intensity threshold → do the different ion transition and peak shape for the same component match sufficiently → is there a match within the existing database (DB)? The integration result will be generated automatically in all cases, except if the answer to the first three questions is "yes", and "DB match?" is "no". In that case, the segment will only be presented to the user for manual review. After all automatic and manual reviews have been performed, a numerical and graphical report is generated as the final output. *PEAR* is written in R (3.6.4, [21]), with a.pdf conversion step performed in Python (3.9.13) using PyMuPDF [22].

4. Detailed process description

4.1. Extraction and preparing data from source files

The first step is to read all required integrated peaks and related data. In that manner, it is convenient and fitting to the existing routine workflow to extract the already integrated chromatographic trace after an export step that each vendor suite offers. By doing so, all the required metadata is also linked to chromatographic data, and the correct time segment and mass traces are selected. Fig. 2 shows an example of the pdf file that serves as the input, within both side-lines are the descriptions of the data required for further processing. Apart from the.pdf file name and corresponding page numbers, the other required traceability data are the compound, ion transition, and sample identification. The ion transition is needed to identify the trace and link the different ion transitions for each compound and sample. The intensity scale is required to convert pixel dimensions into counts (arbitrary units), like the retention time on the horizontal axis, which is also needed to find the actual retention time for each compound trace. Page S12 of the supplementary material provides some more detail how that information was extracted from the pdf file the chromatogram segment, with sufficient resolution, is the core of the data to be processed. The area proposed by the vendor software is not used. The pdf files are first converted into text format (PyMuPDF), which are parsed into the relevant pieces of information and stored in the R environment. Supplementary Fig. S1-2 show an example of an actual pdf file that *PEAR* can process.

4.2. Detection in case peaks are found

PEAR is intended to facilitate the review process, so (automated) peak integration by the software provided by the MS instrument vendor should be performed first. For targeted MS/MS analysis, the user should provide an expected retention time and integration window (margin), smoothing parameters, and threshold parameters to separate real peaks from noise. To capture all targeted compounds accurately, it is recommended to set the integration parameters in the vendor software such that false negatives are extremely unlikely, which is generally the case in residue analysis. *PEAR* currently will process only those chromatograms for which the vendor software indicated the (possible) presence of a chromatographic peak, which the software plots as a shaded area and is easily identified by the script.

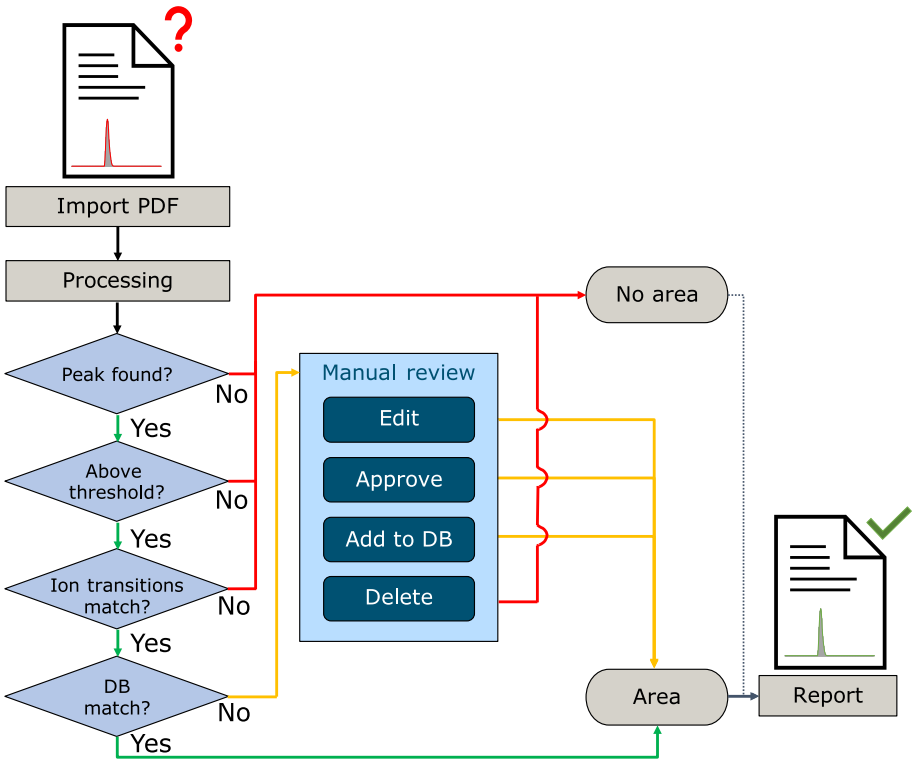

Fig. 1. Schematic overview of the workflow of *PEAR-review*. 'DB': database, a collection of manually selected examples of chromatogram segments and their correct way of integration, with a selection of metadata to ensure the traceability of these traces.

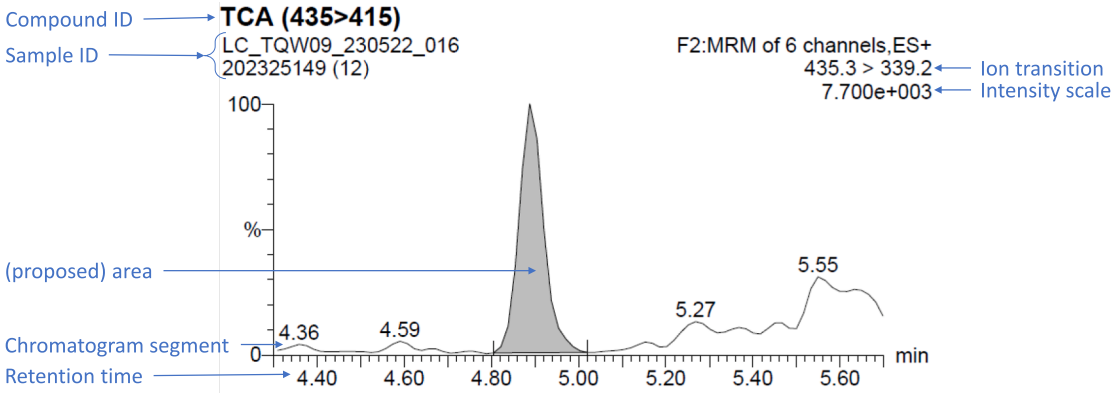

Fig. 2. Example input from.pdf file for one specific compound. The descriptions indicated in blue represent the required information for further processing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Assessment of chromatogram segment similarity with database

A database of reference chromatograms using previously acquired and already checked by analysts of all peaks with corresponding integrations has been iteratively created using the graphical user interface of *PEAR*. For each new chromatogram segment, the peak width is approximated by a simple Gaussian fit, minimising the residuals using a constrained optimisation of parameters scale and standard deviation. The fitted standard deviation value is used to create a standardised segment full chromatogram segment. The portion that equals four times the standard deviation (full width at half peak height, FWHM) before the observed peak apex and six times FWHM after this peak apex is selected and this section is interpolated and set to exactly 32 points using a spline function to form the 'signature section' of this chromatogram segment. This standardised portion of the chromatogram section enables easy comparisons of lines, and the asymmetric approach reflects the area of interest for chromatographic peaks. The relevant scaled sections are subsequently transformed using a periodised, orthogonal forward wavelet transform [23] using a symmlet-6 filter. This step was introduced as a feature enhancer rather than as a noise filter. A selection of the wavelet transformed section was used as input for the comparison with database signals based on a simple Euclidean distance calculation. Figs. S5-6 in the supplementary material provide a visualisation and some background on this step. Unknown signals were matched to database signals based on the smallest distance. A threshold value of 6 was applied to identify no match with existing database signals. Segments are flagged either as matched to a certain database entry, needing no further human attention, or as not matching a database entry moving forward in the process.

4.4. Scale thresholding

An optional step in which, based on analysts' experience, peaks lower than an (adjustable) threshold of 2000 counts are considered to be too noisy, and are automatically flagged as too low and hidden from manual review.

4.5. Ion transition filtering

In targeted MS/MS analyses, components are often determined using two MS/MS (product-) ion transitions. More than two can be used, and there are cases where only one transition is used. When two or more ions are present, *PEAR* will use the standardised sections described above to determine whether the ion transition shape matches sufficiently. For two ions, the Pearson correlation between the two segments is calculated. For three or more ions, the correlation of each signal with the average is calculated. A correlation <0.9 was empirically found to be a suitable threshold value to conclude that ion transitions belong to different compounds. In those cases, it is concluded that the target compound is not present, and these segments are flagged as such and removed from the list of peaks to be reviewed manually. Figs. S7.5 and S7.6 in the supplementary material illustrate this concept.

4.6. Detecting possible conflicts

Integration should enable the user to detect possible problems easily. Apart from problems with how the baseline is drawn automatically, the (automatic) selection of background peaks near the target compound is a significant risk. This problem is especially valid for analytes with only one ion transition, although there could be circumstances where neighbouring background compounds share ion transitions, too. Although *PEAR* is currently incapable of fully automating checks in this respect, two diagnostic numbers allow the analyst to spot potential problems. *PEAR* checks the 'signature' section of the chromatogram segment for possible peaks, as defined by three consecutive ups and three consecutive downs, using the findpeaks function of pracma [24]. If the maximum intensity for candidate peaks exceeds 3x the observed noise level, it is reported as an additional (potential) neighbouring peak for the analyst to check during the evaluation of the data. Moreover, for each compound in the analytical series, which typically consists of multiple samples for which the same series of target compounds are determined, the average retention time for each compound is calculated based on the automatic and manually confirmed detected peaks and the retention time deviation is reported for each chromatogram segment. This can be quickly manually checked after the series is processed, and although the tolerable value for RT might vary, it is easy to spot excessive numbers where a wrongly identified compound is likely, and a check is warranted.

4.7. Graphical user interface

The manual interaction interface of *PEAR* is a Shiny-webpage [25] that displays the entire chromatogram segment, including all required metadata (Fig. 3). Mouse-based zooming is available, and a proposed baseline is drawn and user-adjustable. If available, corresponding ion transitions are drawn in the background. The user can keep the proposed baseline, adjust or delete it, and add a chromatogram to the database for future use. In principle, only the peaks that require human attention are shown in the order as reported, but the user can also review all other available peaks.

4.8. Reporting

After review, all data is exported in a table format, including traceability data, area, flag label on how *PEAR* judged each peak, matching scores, reference to best matching database entry, including a column on

Fig. 3. Example of graphical user interface (GUI). This shows a part of the chromatogram for the 405.2 → 329.1 transition of prednisolone (blue line), with the other ion transition (402.5 → 280.1) for this compound plotted in grey, confirming the green shaded area belonging to this compound, and not the smaller peak with a max around 38 (CX, arbitrary time units). The red dots indicate the start/stop positions of this peak, which can be freely modified. Meta information on the current trace is given in the table at the top. The bottom buttons allow to keep or save the current baseline, delete the peak from the results, or add the current chromatogram section and its start/stop points to the database. The GUI is rendered in a web browser. The review process can be saved at any time, and when finishing the manual review process, *PEAR* is triggered to update all results and save them appropriately. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the number of possible conflicting peaks. Moreover, a graphical (pdf) report is generated with the chromatogram segment line shapes and final baselines. Examples of the exported table (Table S1) and the graphical report (Fig. S3-4) are shown in the supplementary material.

5. Data used for performance assessment

To develop and test *PEAR-review,* a curated data set was needed. However, in many of our analytical series for residue screening, most samples appear to be free from contamination with prohibited substances, which is the preferred and expected result. Therefore, to be able to test *PEAR* we selected the following data sets containing samples that were spiked with different analytes at different concentration levels:

- *Tranq1:* 1080 chromatograms of kidney samples spiked with the tranquillisers carazolol, propionylpromazine, and xylazine. The spiked concentrations are relatively high, and separation is relatively good, so this series was subjectively classified as 'easy' to integrate and review. Moreover, each sample was analysed 20 times to determine the homogeneity, which allowed a comparison of the repeatability of quantitative integration between human experts and *PEAR*.
- *Tranq2:* was selected for the same reasons as *Tranq1* and contained 792 chromatograms of kidney samples spiked with the tranquilisers azaperol, azaperone, chlorpromazine, and haloperidol.
- *Tranq3:* 1924 chromatograms of beef and pork meat and egg samples spiked with the tranquillisers acetopromazine, azaperol, azaperone, carazolol, chlorpromazine, haloperidol, propionylpromazine and xylazine. These samples were used to validate the cleanup method for different matrices. Although all targeted compounds are spiked, the concentrations were relatively close to the limit of detection, making correct integration a little more challenging.
- *Cortico1:* 820 chromatograms of egg samples spiked with 14 different corticosteroids meant for method development. 20α-dihydro prednisolone, 20β-dihydro prednisolone, beclomethasone, betamethasone, clobetasol, cortisol, cortisone, dexamethasone, flumethasone, isoflupredon, methylprednisolone, prednisolone, prednisone, triamcinolone acetonide. This set was chosen as some compounds (isomers) elute very close to each other, making the integration extra challenging.
- *Cortico2:* 1968 chromatograms, a routine set of corticosteroids in egg samples. This contained (mostly without interfering peaks) routine samples and the internal standards and calibration samples for the same 14 corticosteroids as in *Cortico 1*. This routine-representative set was selected to detect how *PEAR* deals with blank samples.

The vendor software integration settings like retention times, S/N, automatic peak detection windows, baseline, smoothing, and threshold parameters are chosen during method development and validation and differ for various series and analysed components. For the export of the chromatograms, eight chromatograms were plotted on a page, and a rather large time window of 1.5 min was selected, 0.75 min before and after the expected retention time of the peaks. This time window is larger than needed since a drift of 0.75 min during an analytical run is normally not acceptable. However, using this large time window demonstrates that *PEAR* can assess if a peak was correctly integrated. These 1.5-min windows provide useful pictures for *PEAR*; peaks have sufficient graphical resolution with enough baseline before and after the peak.

The (startup) database with previously acquired reference chromatogram segments contained 524 segments and corresponding integration start/stop positions. It originated from 11 analytical series from different compounds, matrices, and machines, which were different from the sets used for the performance evaluation below. For use in routine situations, it is recommended that database entries be checked and discussed by experienced technicians if needed.

6. Results and performance assessment

6.1. Efficiency

Based on the five pilot data sets, a total of 6436 targeted segments were subjected to automated analyses using *PEAR*. A performance

Table 1
Overview of the number of chromatogram segments processed and their distribution across the different process steps.

	Tranq1	Tranq2	Tranq3	Cortico1	Cortico2	Total
Total	1080	792	1776	820	1968	6436
No peak found	4	1	17	0	18	40
	(0.4 %)	(0.1 %)	(1 %)	(0 %)	(0.9 %)	(0.6 %)
Match with database	711	524	1030	382	545	3192
	(65.8 %)	(66.2 %)	(58 %)	(46.6 %)	(27.7 %)	(49.6 %)
Invalid ion transitions	292	236	494	164	938	2124
	(27 %)	(29.8 %)	(27.8 %)	(20 %)	(47.7 %)	(33 %)
Below threshold	0	0	13	29	90	132
	(0 %)	(0 %)	(0.7 %)	(3.5 %)	(4.6 %)	(2.1 %)
Automatically processed	1007	761	1554	575	1591	5488
	(93.2 %)	(96.1 %)	(87.5 %)	(70.1 %)	(80.8 %)	(85.3 %)
For manual review	73	31	222	245	377	948
	(6.8 %)	(3.9 %)	(12.5 %)	(29.9 %)	(19.2 %)	(14.7 %)

breakdown is given in Table 1, following the steps described in Fig. 1 – without the manual review step. As the settings for peak detection in the vendor's MS software were set to detect every peak present to minimise the chance of false-negative results, only in very few segments was no peak found, and the first step only removed less than one per cent of all segments to review. The second step, the matching of chromatogram shapes with those in the database, held the largest reduction in workload, with 49.6 % of the chromatograms being recognised as sufficiently close to one of the entries in the reference database. As expected, this number varies by the type of study and, thus, the type of peak shapes present. Note that an unknown and irrelevant number of database matches have subsequently been overruled by the next steps. Next, *PEAR* found that about one-third of the chromatogram segments have an insufficient match between the two ion transition shapes. Therefore, they are judged not to be the compound of interest and need no further human attention. Lastly, a few compounds (2.1 % of the total) are considered too low to be relevant. All these compounds are integrated using only one ion transition, so the preceding selection step could not be performed on these chromatograms. These steps together show that, on average, 85 % of chromatogram segments can reliably be dealt with automatically, while the remaining 15 % is considered to need manual attention. It is noted that this number varies between the different cases and depends on the cleanliness of the peak shapes within the study and the composition of the database. Moreover, the remaining 15 % does not necessarily need manual adjustment of the proposed integration. In fact, in many cases, the integration proposed by the vendor's software is sufficient, but certainly, there are some chromatograms that arguably had integration errors and benefited from human interaction.

6.2. Trueness

To compare automatic quantification with manual (expert) integration, the peak areas from the five sets that were automatically assigned to a database case were compared with those obtained by conventional human-based integration (Fig. 4). With an average quantitative difference of less than 0.01 % for the 3192 automatically quantified compounds, the two methods are in excellent general agreement, but the spread is relevant. The absolute area deviation between automatic and manual integration was smaller than 4.4 % in 95 % of all cases. Higher deviations are due to the integration of tailed peaks, as illustrated in Fig. S7.4.

6.3. Precision

In addition to overall agreement, it is relevant to check if automatic integration yields at least similar precision when samples are repeated. Therefore, we used the results of the Tranq1 and Tranq2 sets to check the performance in terms of relative coefficient of variance (*r*CV) for 20 repeats for five compounds, each integrated for two transitions and a corresponding deuterated standard. That is, the *r*CV for each of the 22 compound-transitions was calculated based on the 20 values obtained from manual integration and on the 20 values obtained by *PEAR*. Note that the actual value of the observed *r*CV is not what is important in this study: the homogeneity of the sample and the analytical repeatability for each compound are the main drivers for that value. Given the 20 chromatograms per sample, we observe the difference between the obtained *r*CV for automatic and manual integration, which indicates the repeatability of the integration effort. The results are shown in Fig. 5. The average performance based on an average observed difference of −0.008 % for these 22 compounds is essentially identical. The *r*CVs for azaperone (Azn-t1 and Azn-t2) and the second (less intense) t2 transition of carozolol (Crz-t2) for with the automatic *r*CV is about 0.8 % higher than for the manual integration. This is explained by a fairly noisy baseline and a substantial amount of peak tailing, which the current (limited) database with integration examples is not optimised for. Fig. S7.4 in the supplementary material illustrates this issue. However, the performance loss is small and can likely be improved by adding more examples of this compound to the database. For the other compounds, both integration modes perform comparably, or the automatic integration performs marginally better than the manual integration regarding repeatability.

6.4. Computational performance

Processing time is important if *PEAR* is to be used routinely. Programming is done in R and mostly depends on R native libraries, and the code has not been optimised for speed. On a standard 6-core 2.7 GHz laptop with 16 GB of internal memory, processing a targeted pdf-export file with 100 pages and eight chromatogram sections per page, the process of reading, processing, and writing the results to disk takes about 51 s, 45 sec of which are needed to import the pdf file.

7. Conclusions and outlook

Manually checking integrated peaks of large triple quad datasets is time-consuming and tedious. To reduce this workload, we developed the *PEAR-review* tool, which automates these checks, taking over the task from analysts. *PEAR* can autonomously decide if a peak was "good" or "bad" integrated on most automatically integrated peaks, mimicking the exact integration checks of skilled analysts. It flags only the undecided cases for analyst review, significantly reducing the workload while maintaining the quality of full human intervention. Additionally, building a database of reference integration examples to improve the interpretation by *PEAR* is straightforward, allowing the creation of either general or specific databases tailored to analyses involving coelution or imperfect peak shapes.

While it is possible to run the current script as written in R and Python and maintain it in response to inevitable changes in MS vendor's export formats, we encourage other parties to integrate this approach more tightly into software that can be used routinely to evaluate

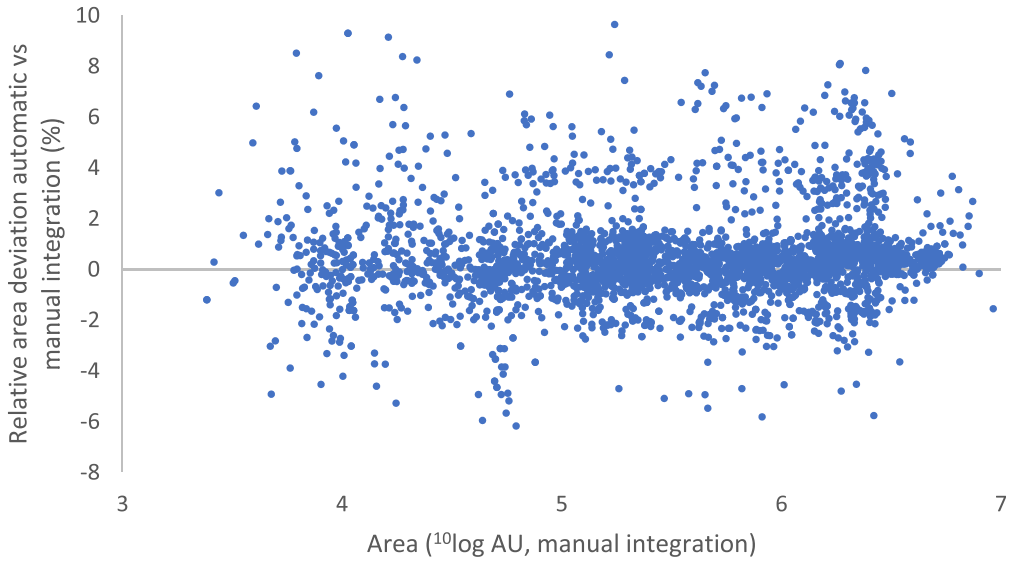

Fig. 4. The relative deviation of areas (AU) between automatic and human integration for 3192 observations in which *PEAR* identified a database match in the current database.

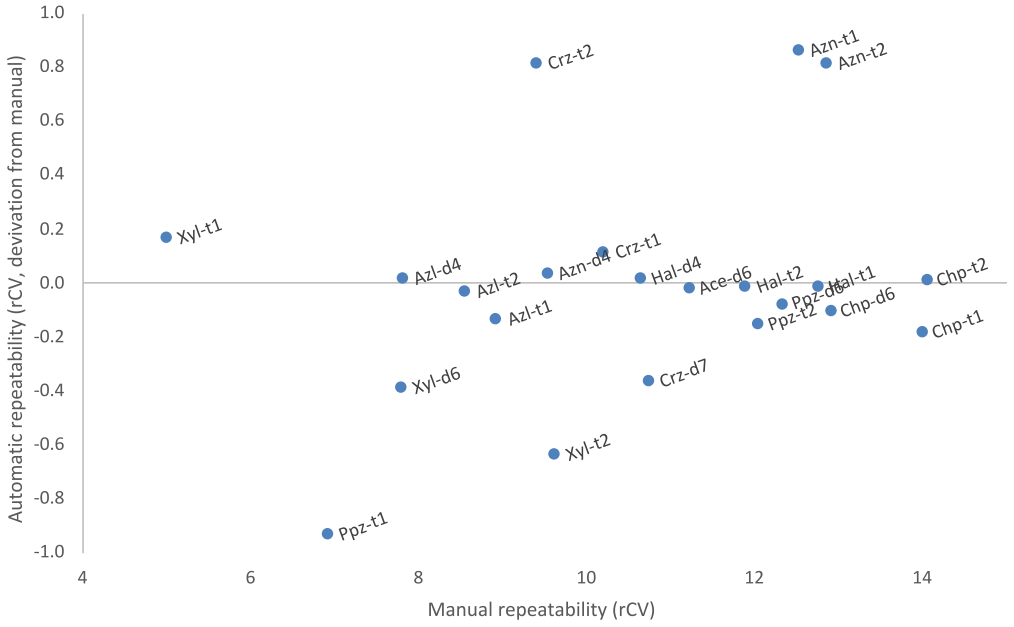

Fig. 5. Repeatability performance for automatic integration – for automatically matched peak shapes. Repeatability was calculated as relative coefficient of variance (*rCV*) for 22 compounds that were repeated 20 times. The automatic *rCV* was subtracted from the manual values, so negative values mean a more precise performance for the automatic integration. Ace: Acetopromazin; Azl: Azaperol; Azn: Azaperone; Crz: Carazolol; Chp: Chloropromazine; Hal: Haloperidol; Ppz: Propionylpromazine; Xyl: Xylazine; -t: ion transition number; -d: deuterated standard (number of deuterium atoms).

targeted MS analyses. Moreover, there is potential to develop the algorithm further to handle larger non-targeted datasets as well correctly.

8. Data statement

The analytical data used in this paper to develop and train the tool is based on real data reflecting real-life situations. Due to confidentiality restrictions, this data cannot be shared, but the *PEAR-review* code will be shared on request.

Funding

This research was funded by the Dutch Ministry of Agriculture, Nature, and Food Quality under their knowledge development program "Healthy and Safe Food Systems" (project KB 37-002-037).

CRediT authorship contribution statement

Martin Alewijn: Writing – original draft, Visualization, Validation, Software, Conceptualization. **Sjors Rasker:** Writing – original draft, Visualization, Investigation. **Dieke van Doorn:** . **Marco Blokland:**

Writing – review & editing, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2024.111578.

References

[1] European Commision, REGULATION (EU) 2017/625 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 15 March 2017 on official controls and other official activities performed to ensure the application of food and feed law, rules on animal health and welfare, plant health and plant protection products, Official Journal of the European Union (L 95/1), 2017.

[2] S.C. Barros, A.S. Silva, D. Torres, Multiresidues multiclass analytical methods for determination of antibiotics in animal origin food: a critical analysis, Antibiotics 12 (2) (2023).

[3] D.A. Vargas Medina, J.V. Bassolli Borsatto, E.V.S. Maciel, F.M. Lanças, Current role of modern chromatography and mass spectrometry in the analysis of mycotoxins in food, TrAC - Trends Anal. Chem. 135 (2021).

[4] D.A. Varela-Martínez, M.Á. González-Curbelo, J. González-Sálamo, J. Hernández-Borges, Analysis of multiclass pesticides in dried fruits using QuEChERS-gas chromatography tandem mass spectrometry, Food Chem. 297 (2019).

[5] S.H. Monteiro, S.J. Lehotay, Y. Sapozhnikova, E. Ninga, A.R. Lightfield, High-throughput mega-method for the analysis of pesticides, veterinary drugs, and environmental contaminants by ultra-high-performance liquid chromatography−tandem mass spectrometry and robotic mini-solid-phase extraction cleanup + low-pressure gas chromatography−tandem mass spectrometry, Part 1: Beef, J. Agric. Food Chem. 69 (4) (2021) 1159–1168.

[6] Y.S. Jung, D.-B. Kim, T.G. Nam, D. Seo, M. Yoo, Identification and quantification of multi-class veterinary drugs and their metabolites in beef using LC–MS/MS, Food Chem. 382 (2022) 132313.

[7] M. Qie, Y. Zhao, S. Yang, W. Wang, Z. Xu, Rapid simultaneous determination of 160 drugs in urine and blood of livestock and poultry by ultra-high-performance liquid chromatography-tandem mass spectrometry, J. Chromatogr. A 1608 (2019).

[8] European Commision, COMMISSION IMPLEMENTING REGULATION (EU) 2021/808 of 22 March 2021 on the performance of analytical methods for residues of pharmacologically active substances used in food-producing animals and on the interpretation of results as well as on the methods to be used for sampling, Official Journal of the European Union (L 180/84), 2021.

[9] F. van Tricht, M. Essers, M. Groot, S. Sterk, M. Blokland, L. van Ginkel, A Fast quantitative multi-analyte method for growth promoters in bovine meat using bead-disruption, 96-well SPE clean-up and narrow-bore UHPLC-MS/MS analysis, Food Anal. Methods 11 (8) (2018) 2206–2217.

[10] E. Jongedijk, M. Fifeik, A. Arrizabalaga-Larrañaga, J. Polzer, M. Blokland, S. Sterk, Use of high-resolution mass spectrometry for veterinary drug multi-residue analysis, Food Control 145 (2023) 109488.

[11] Waters, TargetLynx, 2022. <https://www.waters.com/waters/en_US/TargetLynx-/nav.htm?cid=513791&locale=en_US> (Accessed 17/10/2023).

[12] Agilent, Acquisition for LC/MS, 2024. <https://www.agilent.com/en/product/software-informatics/mass-spectrometry-software/data-acquisition/acquisition-for-lc-ms> (Accessed 12/02/2024).

[13] A. Kensert, E. Bosten, G. Collaerts, K. Efthymiadis, P. Van Broeck, G. Desmet, D. Cabooter, Convolutional neural network for automated peak detection in reversed-phase liquid chromatography, J. Chromatogr. A 1672 (2022) 463005.

[14] A.D. Melnikov, Y.P. Tsentalovich, V.V. Yanshole, Deep learning for the precise peak detection in high-resolution LC–MS data, Anal. Chem. 92 (1) (2020) 588–592.

[15] F. Rupprecht, S. Enge, K. Schmidt, W. Gao, R. Miller, Automating LC–MS/MS mass chromatogram quantification: Wavelet transform based peak detection and automated estimation of peak boundaries and signal-to-noise ratio using signal processing methods, Biomed. Signal Process. Control 71 (2022) 103211.

[16] H. Gowda, J. Ivanisevic, C.H. Johnson, M.E. Kurczy, H.P. Benton, D. Rinehart, T. Nguyen, J. Ray, J. Kuehl, B. Arevalo, P.D. Westenskow, J. Wang, A.P. Arkin, A. M. Deutschbauer, G.J. Patti, G. Siuzdak, Interactive XCMS Online: simplifying advanced metabolomic data processing and subsequent statistical analyses, Anal. Chem. 86 (14) (2014) 6931–6939.

[17] T. Pluskal, S. Castillo, A. Villar-Briones, M. Orešič, MZmine 2: Modular framework for processing, visualising, and analysing mass spectrometry-based molecular profile data, BMC Bioinform. 11 (1) (2010) 395.

[18] A. Lommen, MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing, Anal. Chem. 81 (8) (2009) 3079–3086.

[19] S. Lehotay, Utility of the summation chromatographic peak integration function to avoid manual reintegrations in the analysis of targeted analytes, LCGC North America 35 (2017) 391–402.

[20] S.J. Lehotay, K. Mastovska, A. Amirav, A.B. Fialkov, P.A. Martos, A.D. Kok, A. R. Fernández-Alba, Identification and confirmation of chemical residues in food by chromatography-mass spectrometry and other techniques, TrAC Trends Anal. Chem. 27 (11) (2008) 1070–1090.

[21] R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2019. <https://www.r-project.org/>.

[22] Artifex, PyMuPDF 1.23.8, 2023. <https://pypi.org/project/PyMuPDF/>.

[23] F. Navarro, C. Chesneau, R package rwavelet: Wavelet Analysis (Version 0.4.1), 2020. <https://github.com/fabnavarro/rwavelet>.

[24] H.W. Borchers, pracma: Practical Numerical Math Functions. R package version 2.4.2, 2022. <https://CRAN.R-project.org/package=pracma>.

[25] W. Chang, J. Cheng, J. Allarie, Y. Xie, J. McPherson, shiny: Web Application Framework for R. R package version 1.3.2, 2019. <https://shiny.posit.co/>.