Qualitative Aspects in the Analysis of Pesticide Residues in Fruits and Vegetables Using Fast, Low-Pressure Gas Chromatography–Time-of-Flight Mass Spectrometry

ARTICLE *in* JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY · APRIL 2011

Impact Factor: 2.91 · DOI: 10.1021/jf104606j · Source: PubMed

5 AUTHORS, INCLUDING:

Steven J Lehotay
United States Department of Agriculture
118 PUBLICATIONS **6,026** CITATIONS
SEE PROFILE

H. G.J. Mol
Wageningen University
79 PUBLICATIONS **1,815** CITATIONS
SEE PROFILE

Natchanun Leepipatpiboon
Chulalongkorn University
59 PUBLICATIONS **1,266** CITATIONS
SEE PROFILE

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

ARTICLE

pubs.acs.org/JAFC

Qualitative Aspects in the Analysis of Pesticide Residues in Fruits and Vegetables Using Fast, Low-Pressure Gas Chromatography−Time-of-Flight Mass Spectrometry

Steven J. Lehotay,*[,†] Urairat Koesukwiwat,[†,§] Henk van der Kamp,[#] Hans G. J. Mol,[#] and Natchanun Leepipatpiboon[§]

[†]Eastern Regional Research Center, Agricultural Research Service, U.S. Department of Agriculture, 600 East Mermaid Lane, Wyndmoor, Pennsylvania 19038, United States

[§]Chromatography and Separation Research Unit, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand

[#]RIKILT Institute of Food Safety, Wageningen University and Research Centre, Akkermaalsbos 2, 6708 WB Wageningen, The Netherlands

Supporting Information

ABSTRACT: Quantitative method validation is a well-established process to demonstrate trueness and precision of the results with a given method. However, an assessment of qualitative results is also an important need to estimate selectivity and devise criteria for chemical identification when using the method, particularly for mass spectrometric analysis. For multianalyte analysis, automatic instrument software is commonly used to make initial qualitative identifications of the target analytes by comparison of their mass spectra against a database library. Especially at low residue levels in complex matrices, manual checking of results is typically needed to correct the peak assignments and integration errors, which is very time-consuming. Low-pressure gas chromatography−mass spectrometry (LP-GC-MS) has been demonstrated to increase the speed of analysis for GC-amenable residues in various foods and provide more advantages over the traditional GC-MS approach. LP-GC-MS on a time-of-flight (ToF) instrument was used, which provided high sample throughput with <10 min analysis time. The method had already been validated to be acceptable quantitatively for nearly 150 pesticides, and in this study of qualitative performance, 90 samples in total of strawberry, tomato, potato, orange, and lettuce extracts from the QuEChERS sample preparation approach were analyzed. The extracts were randomly spiked with different pesticides at different levels, both unknown to the analyst, in the different matrices. Automated software evaluation was compared with human assessments in terms of false-positive and -negative results. Among the 13590 possible permutations with 696 blind additions made, the automated software approach yielded 1.2% false presumptive positives with 23% false negatives, whereas the analyst achieved 0.8% false presumptive positives and 17% false negatives for the same analytical data files. False negatives frequently occurred due to challenges at the lowest concentrations, but 70% of them involved certain pesticides that degraded (e.g., captafol, folpet) or otherwise could not be detected. The false-negative rate was reduced to 5−10% if the problematic analytes were excluded. Despite its somewhat better performance in this study, the analyst approach was extremely time-consuming and would not be practical in high sample throughput applications for so many analytes in complicated matrices.

KEYWORDS: qualitative chemical identification, mass spectrometry, pesticide residues, analysis, low-pressure gas chromatography−time-of-flight mass spectrometry (GC-TOF), QuEChERS

INTRODUCTION

In routine analytical applications, sample throughput is an important issue to consider in the choice of an analytical method. Multiclass, multiresidue pesticide analysis in fruits, vegetables, and other commodities is a common application worldwide for the regulation of food safety, international trade, toxicological risk assessment, research investigations, and a host of other purposes. The "quick, easy, cheap, effective, rugged, and safe" (QuEChERS) approach for pesticide residues in foods[1−11] provides rapid sample preparation (high sample throughput) for subsequent analysis by liquid and gas chromatography (LC and GC) typically coupled to mass spectrometry (MS) detectors.

Two versions of the QuEChERS method have been successfully validated in multiple laboratories through the auspices of AOAC International (AOAC Official Method 2007.01)[3] and the European Committee for Standardization (CEN standard method 15662).[4] There are also many modified versions of the QuEChERS approach with small advantages and disadvantages mainly depending on personal preferences.[5−10] In the unified version employed in this study, the use of acetate buffering during extraction and dispersive solid-phase extraction (d-SPE) cleanup with a combination of 150 mg of anhydrous $MgSO_4$ plus 50 mg of primary secondary amino (PSA) plus 50 mg of C_{18} plus 7.5 mg of

Special Issue: Florida Pesticide Residue Workshop 2010

Received: November 30, 2010
Revised: March 21, 2011
Accepted: March 31, 2011
Published: April 01, 2011

Table 1. 151 Pesticide Analytes in the Study plus 2 Deuterated Internal Standards (IS) and Quality Control (QC) Compound

alachlor	*o,p′*-DDT	flucythrinate	phosalone
aldrin	*p,p′*-DDT	fluvalinate	phosmet
atrazine	deltamethrin	folpet	phosphamidon
azinphos-ethyl	demeton-*S*-methyl	fonofos	phthalimide
azinphos-methyl	demeton-*S*-methylsulfone	heptachlor	piperonyl butoxide
α-BHC	diazinon	heptachlor epoxide	pirimiphos-ethyl
β-BHC	dichlorfenthion	heptenophos	pirimiphos-methyl
δ-BHC	dicloran	hexachlorobenzene	procymidone
bifenthrin	4,4′-dichlorobenzophenone	iprodione	profenofos
bromophos-methyl	dicrotophos	isofenphos	propachlor
bromophos-ethyl	dieldrin	kepone (chlordecone)	propargite
bromopropylate	dimethoate	kresoxim-methyl	propazine
bupirimate	dioxathion	leptophos	propetamphos
buprofezin	diphenylamine	lindane (*γ*-BHC)	propham
cadusafos	disulfoton	malathion	propoxur
captafol	disulfoton sulfone	metalaxyl	propiconazole
captan	α-endosulfan	methacrifos	propyzamide
carbaryl	*β*-endosulfan	methidathion	pyrimethanil
carbofuran	endosulfan sulfate	methiocarb	quintozene (PCNB)
carbophenothion	endrin	methoxychlor	resmethrin
carfentrazone-ethyl	endrin ketone	metolachlor	simazine
chinomethionate	EPN	metribuzin	sulprofos
cis-chlordane	esfenvalerate	mevinphos	tebuconazole
trans-chlordane	ethafluralin	mirex	tecnazene
chlorfenvinphos	ethion	myclobutanil	terbufos
chlorothalonil	ethoprophos	*cis*-nonachlor	terbuthylazine
chlorpropham	ethoxyquin	*trans*-nonachlor	tetrachlorvinphos
chlorpyrifos	famphur	oxadixyl	tetraconazole
chlorpyrifos-methyl	fenamiphos	oxyfluorfen	tetradifon
coumaphos	fenarimol	parathion	tolclofos-methyl
cyanophos	fenchlorphos	parathion-methyl	triadimifon
cyfluthrin	fenitrothion	penconazole	triazophos
λ-cyhalothrin	fenoxycarb	pendimethalin	trifluralin
cypermethrin	fenpropathrin	pentachloroanisole	vinclozolin
cyprodinil	fensulfothion	pentachlorothioanisole	triphenylphosphate (QC)
o,p′-DDD	fenthion	*cis*-permethrin	atrazine-d_5 (IS)
p,p′-DDD	fenthion sulfone	*trans*-permethrin	fenthion-d_6 (IS)
o,p′-DDE	fenvalerate	*o*-phenylphenol	
p,p′-DDE	fipronil	phorate	

graphitized carbon black (GCB) sorbents per gram of matrix equivalent for all fruits and vegetables is presented as a more widely effective option. This simplifies the QuEChERS approach to involve fewer options for analysts and vendors of QuEChERS products. This QuEChERS version was used in a recent validation study for 150 pesticides and shown to give excellent results overall.[11] However, dry samples such as cereals, nuts, and grains give better cleanup without buffering and use of more PSA (150 mg/mL of extract) during cleanup in QuEChERS.[6,7]

QuEChERS is a high-throughput sample preparation method that allows a single analyst to prepare about 20 prehomogenized samples per hour for subsequent analysis, but it still relies on rapid LC-MS and GC-MS analytical methods to achieve wide analytical scope, good quantification ability, high precision, low detection limits, adequate robustness, acceptable sample throughput, and the degree of selectivity needed to make analyte identifications. The previous validation study[11] covered all of these listed aspects except the latter issue of qualitative analysis.

Highly evolved protocols and acceptance criteria have been devised for the quantitative validation of analytical methods,[12−15] but qualitative method validation (e.g., analyte identification) has not been the subject of as much scrutiny. Traditionally, qualitative method acceptance criteria for residue analysis, particularly for MS-based methods, have involved arbitrarily chosen parameters without validation,[12,14,16−20] but, recently, several proposals and studies have begun to assert that qualitative methods should be empirically validated in a similar way as quantitative methods.[13,16,20−24]

The current state-of-the-art LC-MS approach used in many routine pesticide monitoring laboratories is ultrahigh-performance liquid chromatography−tandem mass spectrometry (UHPLC-MS/MS), which commonly provides 2−4 times greater sample

throughput than traditional LC separations.[25] Typical analysis time for >200 analytes is ≈10 min with UHPLC-MS/MS,[6,26] but GC analysis is still required in a multiclass, multiresidue monitoring scheme to cover those pesticides that are not amenable to LC-MS methods. Fast GC-MS techniques have been available for many years,[27] but routine laboratories still commonly use methods that take 30−45 min in their applications. Among the current fast GC-MS options, the features of speed, sensitivity, ruggedness, and ease of low-pressure (LP)-GC-MS make it the most overall advantageous approach to match the sample throughput of UHPLC-MS/MS for concurrent analysis.[11,27−29] LP-GC-MS sacrifices a degree of separation efficiency compared to the use of traditional column dimensions, but the greater selectivity of MS can overcome that sole disadvantage.

LP-GC-MS works by attaching a short, uncoated, microbore restrictor capillary (e.g., 3 m, 0.15 mm i.d.) to the inlet, which maintains normal GC pressures and operation, and connecting it to a short, thick-filmed, megabore analytical column (e.g., 10 m, 0.53 mm i.d., 1 μm film thickness) that leads to the MS detector. In this way, the entire analytical column is under vacuum conditions, and the viscosity of the He carrier gas is reduced, thereby increasing the optimal flow velocity that helps offset the loss of separation efficiency when speed is increased in the GC analysis.[28]

Previously, we validated QuEChERS+LP-GC-MS with a time-of-flight (ToF) instrument for quantitative purposes and demonstrated sample throughput and ruggedness of the approach.[11] The aim of this complementary study was to qualitatively assess the LP-GC-ToF method by assessing the rates of false positives and negatives in the analysis of blind sample extracts. This is similar to what has been proposed and used previously.[13,16,22] An additional aspect of the study was to evaluate and compare human-based decision-making with an automated software process to distinguish presumptive positive findings from negatives.

EXPERIMENTAL PROCEDURES

Preparation of the Blind Spiked Extracts. A total of 151 pesticides as listed in Table 1 were included in the analytical method. All pesticide standards had purity ≥95% (typically >99%) and were obtained from the Environmental Protection Agency's National Pesticide Repository (Fort Meade, MD), Dr. Ehrenstorfer GmbH (Augsburg, Germany), and Chemservice (West Chester, PA). Isotopically labeled internal standards (IS), atrazine-d_5 (ethyl-d_5) and fenthion-d_6 (*o*,*o*-dimethyl-d_6), were obtained from C/D/N Isotopes (Pointe-Claire, PQ, Canada). Triphenylphosphate (TPP) was used as a quality control (QC) standard added to all final extracts, blanks, and calibration standards. HPLC-grade acetonitrile (MeCN) and analytical grade glacial acetic acid (HOAc) were obtained from J. T. Baker (Phillipsburg, NJ), and ACS-grade formic acid (88%) was from Spectrum (New Brunswick, NJ).

Organic-labeled strawberry, tomato, mixed lettuces, potato, and oranges were purchased from a local grocery store as the matrices for this experiment. These were the same samples as chosen in a previous study[11] due to their diversity and moderate degree of complexity to provide a reasonable challenge for the LP-GC-ToF method. The samples (including peels) were cut into small portions with a knife, placed in a freezer until frozen, and then comminuted in a 2 L chopper (Robotcoupe, Jackson, MS) with dry ice. The homogenized samples were stored in the freezer before being thawed just prior to extraction.

For extraction, we used the acetate-buffered QuEChERS method with a d-SPE cleanup step that was modified slightly from the AOAC Official Method 2007.01 version.[3] For the initial extraction step, we added 15 mL of MeCN containing 1% glacial HOAc to 50 mL polypropylene (PP) centrifuge tubes containing 6 g of anhydrous $MgSO_4$ and 1.5 g of anhydrous sodium acetate (NaOAc) obtained from Restek (Bellefonte, PA). Then, 15 g of chopped sample (strawberry, orange, potato, lettuces, or tomato) was weighed into four tubes for each matrix. The tubes were shaken vigorously by hand for 1 min and then centrifuged at 3000 rcf for 3 min at room temperature. The four initial extracts were combined for each matrix, and two 25 mL portions were added to 50 mL PP tubes containing 3.75 g of anhydrous $MgSO_4$ (99.5% purity; UCT, Bristol, PA), 1.25 g of PSA (40 μm particles, UCT), 1.25 g of C_{18} (40 μm particles; J. T. Baker), and 0.188 g of ENVI-Carb GCB (120/140 mesh; Supelco, Bellefonte, PA). This corresponded to 150 mg of anhydrous $MgSO_4$ plus 50 mg of PSA plus 50 mg of C_{18} plus 7.5 mg of GCB per gram of matrix or milliliter of extract, which was the ratio used in the previous quantitative validation study of this method.[11] The d-SPE tubes were shaken for 1 min by hand and centrifuged at 3000 rcf for 3 min at room temperature. The cleaned extracts (12.5 mL from each tube) were combined, and 1.5 mL of 3.33 ng/μL each of TPP, d_5-atrazine, and d_6-fenthion in MeCN containing 1.32% formic acid was added to yield 200 ng/g equivalent concentrations of the IS and QC compounds and ≈0.075% formic acid. The acidic conditions were intended to increase the stability of the base-sensitive pesticides.[30] Then, 1 mL of each final extract was transferred to 26 labeled autosampler vials for each matrix. Six of the vials (numbered 21−26) were used for matrix-matched calibration standards that included all 151 pesticide analytes at 25, 50, 100, 250, 500, and 1000 ng/g equivalents.

It was decided, mainly for logistical reasons, that 20 blind samples would be analyzed per commodity, plus the reagent blank, and sets of both matrix-matched and reagent-only standards were included in each sequence (33 injections/sequence times 5 sequences). To help produce the blind spiking scheme, an Excel spreadsheet was used in a similar manner as in a previous GC-MS qualitative assessment study.[22] This spreadsheet is provided as Supporting Information and may be referenced to help understand the following explanation. The first step entailed the selection of the number of pesticides to be spiked into each of the five matrices. A random number between 10 and 20 was produced by the spreadsheet function (11 for orange, 12 for tomato, 20 for potato, 18 for strawberry, and 16 for lettuces). Then, each pesticide was selected from the random number generator by its assigned number from 1 to 151 for each matrix. The number of blanks (between 5 and 15) per chosen pesticide was also decided from the spreadsheet, as well as which of the 1−20 vials would serve as the blanks for that analyte. Finally, the spiking levels for the given pesticides in the nonblank vials were assigned for that particular pesticide/matrix pair.

The spiking range was predetermined to be between 25 and 1000 ng/g equivalent concentrations in the samples, which was the validated quantitative range tested in the previous study.[11] Solutions A, 2.5 ng/μL, and B, 15 ng/μL, were prepared in MeCN with 0.05% formic acid for each chosen pesticide. The Excel random number function was used to assign 10−60 μL for solution A to yield between 25 and 150 ng/g spiking levels and 10−67 μL for solution B to yield between 150 and 1000 ng/g levels for the 1 g equivalent sample extracts. To better test the method at lower concentrations, solution A was entered into the spreadsheet twice as often as solution B for the vials assigned to be spiked. The total volume of spikes could not exceed 500 μL to yield 1.5 mL final extract volume in any particular vial; thus, some manual manipulations were needed to avoid this problem. The proposed added concentrations were plotted on the spreadsheet before the spikes were finalized, and some adjustments were made to ensure that the concentrations were reasonably distributed over the 25−1000 ng/g range. Finally, a few pesticide additions known to be difficult were made by the chemist to provide an extra challenge to the method and analyst.

Once the spreadsheet was finalized, an independent chemist in the laboratory added the required volumes of the assigned spiking solutions to the appropriate vials containing the blank extracts. When total spiking volume was <500 μL, MeCN with 0.05% formic acid was added to ensure that all vials contained 1.5 mL final volume, and the weights of the vials after the solutions were measured to help verify if the additions were made properly. All sample vials were stored at −20 °C when not in use.

LP-GC-ToF Analysis. The GC-MS analysis in this study was performed as before[11] on an Agilent (Little Falls, DE) 6890 GC coupled to a Leco (St. Joseph, MI) Pegasus 4D ToF instrument. Injection was conducted by a Combi-PAL autosampler (Leap Technologies, Carrboro, NC) in combination with an Optic-3 programmable temperature vaporizer (PTV) inlet (Atas-GL International, Veldhoven, The Netherlands). Ultrahigh-purity helium (Airgas, Radnor, PA) was used as carrier gas at 20 psi (138 kPa) constant inlet pressure. The injection volume was 10 μL into liners containing sintered glass on the walls (Atas-GL International part A100133). The liners were changed after injection of ≈75 samples.

The PTV was programmed as follows: the initial injector temperature was 75 °C for 18 s (vent time, 15 s) with split flow of 50 mL/min, followed by splitless transfer of analytes to the column for 2 min while the injector was ramped to 280 at 8 °C/s; then the split flow was reduced to 20 mL/min, and the injector temperature was decreased to 250 °C until the end of the run.

The analytes were separated on a 10 m × 0.53 mm i.d. × 1 μm film thickness Rti-5 ms analytical column coupled to a 3 m × 0.15 mm i.d. noncoated restriction capillary at the inlet (Restek). A GC column connector was used to couple the columns in which the restriction capillary fit inside the megabore column to make a zero dead volume connection. The GC oven was set at an initial temperature of 90 °C (held for 1 min), ramped to 180 °C at 80 °C/min, then at 40 °C/min to 250 °C, and ramped to 290 °C at 70 °C/min, and held for 4 min. An oven insert pad was used to reduce the oven size, which enabled slightly faster heating and cooling. The transfer line and ion source temperature were set at 280 and 250 °C, respectively; the electron ionization energy was −70 eV; the detector voltage was 1800 V; and a 130 s filament and multiplier delay was used. The spectral data acquisition rate was 10 spectra/s for collection of m/z 70−600.

Leco ChromaTOF software version 3.22 was employed for the instrument control and data acquisition. Two approaches for data evaluation were employed. In the first approach the data were processed by the software, which involved deconvolution and library searching, but a full manual data re-evaluation and verification in each chromatogram was performed after that. For the latter, extracted ion chromatograms of diagnostic ions for target pesticides in all samples were manually retrieved from the raw data to find the compounds of interest and to verify peak assignments and/or integrations. For qualitative assessment, the deconvoluted spectra of the compounds found were compared with the reference spectra from the library. NIST 2005 mass spectral library software and Agilent's pesticide and endocrine disruptor database were used to help the analyst with mass spectral matching and peak identification, but human judgment was used to make the final decisions. For quantification, the matrix-matched standards normalized to d_5-atrazine were used, and the reagent-only calibration standards were used to assess matrix effects. Due to carry-over and possible cross-contamination issues (see Results and Discussion), we dropped two samples from the study and set a reporting limit of ≥10 ng/g for determined concentrations (and ≥20 ng/g for comparison purposes).

In the second approach for analyte identification, the Leco LP-GC-ToF raw data files were processed by the coauthors at RIKILT as reported elsewhere.[31,32] The procedure for automated detection and the parameter settings and criteria first involved preprocessing of the raw data files (involving deconvolution), which was done by ChromaTOF 3.26 software (Leco). To this end, the "Baseline", "Peak Find", and "Library Search" tasks were used. For Baseline and Peak Find, the parameters were set as follows: S/N threshold = 10; number of apexing masses = 2; baseline offset = 0.5; data points averaging for smoothing = "auto" setting; and peak width = 3 s. For library searching, the following parameters were used: library identity search mode = "normal"; library search mode = "forward"; minimum and maximum molecular weight allowed = 50 and 600, respectively; mass threshold = 5; and minimum similarity match before name is assigned = 600. The library used for searching was a custom library of the 151 target pesticides from this work. All spectra included had previously been acquired by injection of the pesticide reference standards on a Pegasus 4D ToF-MS instrument. The in-house library was created by injecting 5 ng equivalent of each pesticide in solution. In the case of low response, higher concentrations were injected. The resulting compound list from the ChromaTOF software was automatically further processed by an in-house developed Excel macro. Using this macro, retention times of the pesticides in sample peak lists were checked against reference retention times, which were collected from the data files of reference standards and included in the macro. A tolerance of ±9 s was set, and any pesticides not meeting this criterion were removed from the list. In the remaining peak list, for pesticides occurring more than once, only the one with the highest similarity value was kept (this feature was designed for GC×GC data and not relevant for the current work). The final resulting output was an Excel list of pesticides automatically detected on the basis of the preset criteria in both ChromaTOF and the Excel macro. An additional evaluation was done with enhanced thresholds for S/N (100) and similarity (≥700).

RESULTS AND DISCUSSION

Carry-Over, Mistakes, and/or Cross-Contamination. An issue that the chemists encountered when reviewing the results from the extracts was a preponderance of findings for certain pesticides at ultratrace levels in many extracts for all of the matrices. These pesticides consisted of those that yielded the lowest limits of detection (estimated to be ≈1 ng/g), such as DD/D/E/T, mirex, metolachlor, nonachlor, and chlordane. Carry-over was isolated to be the main cause of the problem because the number of low-level findings was particularly high for the two samples in each sequence that followed the high calibration standards. The problem was not observed to this extent in the previous validation study on quantification.[11] We believe that the inlet conditions or sintered-glass liner was the cause of the carry-over but did not have the opportunity to investigate the source before moving on to other projects. Rather than try to sort out tainted results due to carry-over from those that were legitimate, we decided to drop all results from the two blind samples that followed the high calibration standards; thus, we report the results as if 18 samples were prepared, not 20.

In addition to carry-over, it is possible that a degree of cross-contamination of pesticides occurred due to impurities in the standards (particularly for isomers) and ultratrace laboratory contaminations. Hundreds of solution additions were made to prepare the samples, and although great care was taken to avoid pipetting errors and cross-contamination problems, it is unavoidable that some mistakes would be made. When a mistake was recognized during the pipetting process (e.g., wrong solution or amount added to the wrong vial), the answer key was adjusted to reflect the correct information. In all, the chemist made eight notations about possible errors during pipetting. Residue laboratory personnel also have to be cognizant that high concentration standards and solutions should not be stored or used near samples in chemical residue analysis applications. Working

conditions entailed dozens of 15 ng/μL solutions being kept near the sample vials, and it is not inconceivable that some ultratrace cross-contamination occurred despite the best intentions of the chemist.

For these reasons, we set as one of the analyte identification criteria that calculated concentrations had to be ≥10 ng/g for the identification to be reported. This is not unrealistic because real-world pesticide residue analysis for regulatory purposes does not typically report findings <10 ng/g. Moreover, our lowest calibration and spiking levels were 25 ng/g, and the validation experiments[11] showed pesticide recoveries were typically near 100% with the method (in any case, we spiked the extracts in this study to eliminate recoveries as a factor). Even so, some false-positive identifications in this study almost certainly arose from pesticides that actually appeared in the chromatograms due to carry-over, cross-contamination, or mistakes during preparation of the extracts. This could not be verified; therefore, all results are reported, but some of them remain suspicious to the authors. Similarly, some false negatives were probably a result of vial mix-ups or mistaken nonadditions of the pesticides, but we have this suspicion in few cases based on the results.

Quantification. Although quantification for the method was already validated via recovery experiments,[11] we were able to assess quantitative aspects for the added pesticides in this study, too. Figure S1 (Supporting Information) shows the comparison of added versus determined concentrations for atrazine in the extracts. In this case, d_5-atrazine was used as the internal standard in the analysis, which compensated for possible differences in matrix effects. Ideally, the slope of the line relating added versus determined concentrations in the samples would be 1.0, but a −12.5% difference occurred in this case. This was probably the result of two main factors: (1) a different set of solutions was used to prepare the calibration standards than was used for making the sample additions; and (2) the calibration standards contained 151 analytes, all at the same known concentrations, whereas the sample solutions contained far fewer analytes at widely different concentrations unknown to the analyst. Fillion et al.[33] demonstrated that the analytes themselves can cause the matrix-induced response enhancement effect,[34] and we believe that this was occurring to different degrees depending on the susceptibility of the pesticide analyte to this effect. The trace amounts of toluene and/or ethyl acetate from the initial stock solutions, particularly for the high-concentration calibration standards, may also have played a small role in different slopes in the plots. Overall, the average difference was −15 ± 44% (n = 576) for the calculated concentrations versus the expected added concentrations based on the six-point matrix-matched calibration curves using d_5-atrazine as the internal standard in all cases.

Degradation. The factors listed above are the main reason for biases in the quantitative results, because all of the standards were prepared separately in the sample additions, but another factor relates to degradation of some compounds. This was the most important aspect for many of the "false negatives" in this study. Should we really count an unidentified pesticide that was added to the extracts, but which probably had degraded prior to analysis, as a false negative? The more conservative opinion is "yes" because "the overall method" failed to identify it, but the reality is that no method would detect the parent analyte if it had fully degraded. However, we could not be sure that the analyte had degraded, so we followed the conservative approach in reporting the overall results, but we also considered the results without the questionable pesticides included.

For example, kepone (chlordecone) in tomato gave a −85% difference between the expected and determined slope in the concentrations. Also, it gave false negatives in all three cases in the experiment when spiking levels were <60 ng/g (and calculated concentrations were presumably <25 ng/g). Degradation is the most likely cause of the large discrepancy, but another possibility is that an error could have been made in the solution preparation. Degradation surely occurred for captafol and folpet in the extracts, which are known to be unstable in MeCN,[30] and we also believe that carbaryl, which is more suitable for LC analysis, had degraded.

Isomers. There were several structural isomers, enantiomers, and cis/trans-diastereomers included among the possible analytes as listed in Table 1. For the sake of this qualitative assessment, we treated each isomer as an individual analyte, even though, in reality, the applied pesticides or environmental contamination typically contains multiple forms of these types of compounds. The presence of multiple species of the same chemical actually serves as a very helpful factor when identifications are made in real-world applications. For example, if a positive finding of both *cis*- and *trans*-permethrin is made in the correct ratio of typical pesticide formulations, then chances are very good that this is a real result, despite the relatively nonselective nature of the mass spectrum for permethrin, which mainly consists of a base peak at m/z 183. Conversely, if only one of the spatial isomers is detected, or if the peak area ratio between the permethrins is not typical, the analyst should investigate further, or discard, the result depending on the importance of the analysis.

In this study, we chose to treat the isomers as separate analytes because we wanted to create a greater challenge to the method, analyst, and software used to make identifications. The chemist reserved some special additions to the extracts for this very purpose, as if a teacher was devising trick questions for the pupils. In most cases, such as *p,p′*- and *o,p′*- DD/D/E/T, α- and *β*-endosulfan, and *trans*- and *cis*-chlordane and -nonachlor, the t_R differences were great enough that there were no problems in distinguishing the different analytes from each other. However, the enantiomers of fenvalerate and esfenvalerate more closely eluted and proved to be more challenging. Their analysis was further complicated by either the presence of the other enantiomer in the individual standards or conversions in solutions or during chromatography.

Figure 1 shows this relationship between fenvalerate and esfenvalerate in terms of added concentrations for individual enantiomers and determined concentrations in the samples found to contain them. If the analytes are treated separately, 36 additional false presumptive positives occurred for the qualitative human decision-making approach and 23 for the automated software approach. Because both enantiomers have the same mass spectrum, the ≈7 s difference in t_R is the only distinguishing aspect between the two compounds. The automated instrument data processing software and postrun Excel program frequently gave mismatches of fenvalerate as esfenvalerate and vice versa. In both cases, human intervention was needed to assign the analytes to the proper pesticide name based on the t_R.

Even after the correct analyte assignments were made, esfenvalerate was nearly always identified in extracts to which only fenvalerate was thought to be added, and vice versa. This relationship is displayed in Figure 1. Only a single instance occurs (lettuces sample 2) in the analyst identification approach when both fenvalerate and esfenvalerate were identified (at ≈23

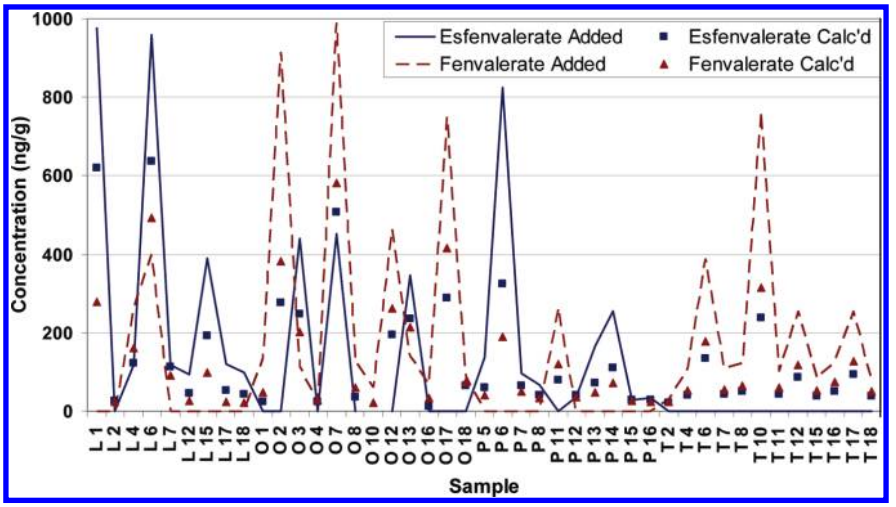

Figure 1. Differences between added and determined concentrations for fenvalerate and esfenvalerate. The attempt to treat them individually did not lead to acceptable quantitative or qualitative results (L, lettuces; O, orange; P, potato; and T, tomato).

ng/g) but were not expected to appear in the extract. Quantitatively, when the analytes were treated individually, a −40% bias occurred in the slopes of the determined versus expected concentrations, but when both enantiomer concentrations were summed, the bias was reduced to −23%. In retrospect, it would have been better to treat fenvalerate plus esfenvalerate as a single analyte because either cross-contamination or chemical conversion was occurring. Therefore, false presumptive positives were not counted in cases when the other isomer was added to the given sample.

Unlike es/fenvalerate, *cis*- and *trans*-permethrins could be treated as individual analytes in the study, even though they gave the same mass spectrum with only a ≈3 s t_R difference between them. Only a few instances occurred in which both permethrin species were identified in the same extract, and these were expected to be the case. Additionally, *cis*-permethrin was not regularly identified when *trans*-permethrin was present at high concentrations, and vice versa. As reflected in Figure 2, the 12 false-positive identifications for both forms of permethrin (6 each) all occurred at relatively low levels (≤60 ng/g). The reason for the larger number of false positives for permethrins than for other analytes almost certainly is a result of its simplistic mass spectrum. The full mass spectrum of permethrin essentially yields only the base peak of m/z 183, which makes it more difficult to distinguish from chemical noise in the chromatograms.[22]

The plot in Figure 2 also shows the only false negative to occur for permethrins. Both the manual human-based and automated software approaches gave the same result, which more than likely means that the chemist neglected to add the 70 ng/g *cis*-permethrin as had been intended. In terms of quantification, Figure 2 also indicates a difference in the calculated concentrations for *cis*- and *trans*-permethrins versus the expected values. The most probable reason for this comes from the use of both forms of permethrin in the calibration standards at equal concentrations, but the samples contained either one distereoisomer or both at unequal levels. The peaks coeluted to an extent, which led to integration differences and inaccuracies in the quantification of the individual permethrin analytes.

Qualitative Criteria. In this paper, we use the qualitative analysis definitions described in the literature.[16] In this respect,

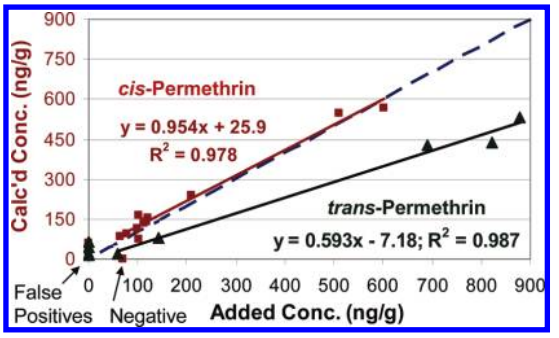

Figure 2. Plot of the analytical results of permethrins in lettuces, orange, and strawberry showing the false presumptive positives and false negatives, which are compiled in Table 3. The false negative for *cis*-permethrin at 70 ng/g is very likely a pipetting mistake during sample preparation. Similarly, some of the "false positives" may actually have been present by mistake. The presence of both closely eluting compounds in calibration standards (whereas only one of them was typically present in the samples) was probably the main cause of the quantitative differences in the results.

the quantitative results are called "determinations", and the reported qualitative MS results are called "identifications", which must meet the MS identification criteria that we chose for this purpose (human judgment in one case and fixed criteria in the other). By itself, this rapid GC-ToF method does not perform "confirmation" of the results because this would entail agreement of findings in a second analysis, potentially involving sample preparation of a second sample portion and use of an orthogonally selective analytical method (depending on the stakes of the results). Thus, we often refer to the positive identifications in this study as "presumptive positives" because they have not been confirmed.

In theory, the goal of qualitative identification is to achieve 100% true results, in which case the rates of false identifications and false negatives each would be 0%. However, this ideal situation cannot be realized in practice near the concentration limit of identification due to measurement uncertainties. These variabilities dictate that the percentage of false positives increases in this region of uncertainty as the rate of false negatives decreases, and vice versa.[19] Identification criteria should be chosen on the

basis of the need for the results to find the most acceptable balance between false positives and false negatives. In legal applications, false-positive findings can lead to improper enforcement actions; thus, a small increase in the number of false negatives is tolerable to eliminate false positives. In screening applications, the rates of false negatives at the reporting limit should be minimized, and a reasonable percentage of false indications are acceptable because these should be eliminated by the confirmatory method. Our pesticide residue analysis application is a mixture of the two extremes, because it involves screening, determination, and identification of pesticide residues for regulatory purposes; thus, our overall goal was to find the optimal balance in the rates of false positives and negatives to maximize the percentage of correct qualitative findings overall for the targeted list of pesticides. This optimized balance is achieved by choosing the most appropriate MS identification criteria.

Previously, Lehotay and Gates[22] conducted a GC-MS study in which selected ion monitoring (SIM) mode with a quadrupole instrument was compared in the analysis of shared blind samples with MS/MS using an ion trap instrument. In that case, the number of ions and their relative abundance ratios were used to set the identification criteria (as well as t_R, S/N, peak shape, and other factors).

In the case of ToF, "full scan" mass spectra are provided; thus, MS matching factors with the library reference spectrum (ideally contemporaneously generated within the same analytical sequence) replace the ion abundance ratio criterion in SIM and MS/MS for identification purposes. We have a targeted list of known analytes in this application (as in typical pesticide residue monitoring analyses), which allowed us to use a mass spectral library consisting of only the analytes on our list in the automated software approach. It is also very important, however, to exclude possible chemicals that could yield a similar spectral fit, and thus in the human-based approach, the analyst frequently cross-checked that possibility using the NIST 2005 mass spectral library and Agilent's pesticide and endocrine disruptor database. In both cases, the instrument software's mass spectral deconvolution feature provides additional help to isolate the mass spectrum of each chromatographic peak.

In the case of the human assessment, no specific identification criteria were given except that the analyst was to use his/her best judgment for the purposes of publishing the results. The analyst was trained in the use of the instrument and method and had conducted the previous validation study.[11] The chemist had read the literature on chemical residue identifications[16−22] and was familiar with the aim of this study. The instrument software with deconvolution and mass spectral searching was used to process the chromatograms in the sequences, and the analyst specifically checked manually every chromatogram for every analyte to see if the finding and integrated peak were correct or not according to knowledgeable personal judgment. The deconvolution program tends to increase the limit of identification,[35] and the software missed many analyte peaks that the analyst identified and integrated manually. The only caveat was the reporting limit was set to ≥10 ng/g for reasons already described. We also evaluated a reporting limit ≥20 ng/g to see its effect.

In the automated Excel-based approach, the instrument software report files containing t_R, peak width, S/N, matching factor (similarity value), peak area, and other information from the analyses were downloaded into the spreadsheet. The spreadsheet contained macros in the cells that automatically generated a report for those pesticide analytes that met the identification criteria specified in the program (which could be adjusted). The identification criteria we used were $t_R \pm 9$ s of expected t_R for the analyte, S/N ≥ 10 (or 100), and matching factor ≥600 (or 700) versus the library reference spectrum generated on the GC-ToF instrument

Table 2. Qualitative Results from the Automated Software Approach with Respect to Concentrations of the Pesticide Analytes in the Study Using the Matrix-Matched Calibration Standards from 25 to 1000 ng/g

Analytes with a Matching Factor ≥600 for 25 ng/g Calibration Standards in All Matrices
alachlor, aldrin, α-BHC, bifenthrin, bromophos-ethyl, bromopropylate, cadusafos, carbophenothion, chlorpyrifos, chlorpyrifos-methyl, λ-cyhalothrin, *p,p′*-DDD, *o,p′*-DDE, *p,p′*-DDE, *p,p′*-DDT, dichlofenthion, 4,4′-dichlorobenzophenone, dieldrin, disulfoton, endrin, ethion, fenarimol, fenvalerate/esfenvalerate, flucythrinate, fonophos, heptachlor, heptachlor epoxide, hexachlorobenzene, isofenphos, methacrifos, mirex, *cis*-nonachlor, pentachloroanisole, *trans*-permethrin, phorate, pirimiphos-ethyl, pirimiphos-methyl, procymidone, sulprofos, trifluralin, vinclozolin
Analytes with a Matching Factor ≥600 for Calibration Standards 25−50 ng/g
δ-BHC, bromophos, buprofezin, carbofuran, carfentrazone-ethyl, chlorfenvinphos, cyprodinil, diazinon, α-endosulfan, endosulfan sulfate, endrin ketone, ethalfluralin, fenchlorphos, fenpropathrin, fenthion, fluvalinate, leptophos, malathion, metolachlor, myclobutanil, *trans*-nonachlor, pentachlorothioanisole, *cis*-permethrin, profenofos, propachlor, propargite, propetamphos, propoxur, resmethrin, tebuconazole, tecnazene, terbufos, tetraconazole, tetradifon, tolclofos-methyl, triadimefon
Analytes with a Matching Factor ≥600 for Calibration Standards 50−100 ng/g
cis-chlordane, *trans*-chlordane, cypermethrin, deltamethrin, diphenylamine, ethoprophos, ethoxyquin, fipronil, heptenophos, oxyfluorfen, parathion, *o*-phenylphenol, phosalone, propiconazole, pyrimethanil, quintozene (PCNB)
Analytes with a Matching Factor ≥600 for Calibration Standards 100−250 ng/g
atrazine, azinphos-ethyl, β-BHC + γ-BHC (lindane), bupirimate, chlorothalonil, cyfluthrin, fenitrothion, kresoxim-methyl, methoxychlor, parathion-methyl, phosphamidon, propazine, propham, propyzamide, tetrachlorvinphos, triazophos
Analytes with a Matching Factor ≥600 for Calibration Standards 250−500 ng/g
captan, carbaryl, chlorpropham, coumaphos, cyanophos, *o,p′*-DDD, demeton-*S*-methyl, dicrotophos, β-endosulfan, famphur, fenamiphos, iprodione, kepone (chlordecone), metalaxyl, methiocarb, metribuzin, mevinphos, oxadixyl, penconazole, pendimethalin, phthalimide, piperonyl butoxide
Analytes without a Matching Factor ≥600 in the Calibration Standards up to 1000 ng/g
azinphos-methyl, captafol, *o,p′*-DDT, demeton-*S*-methylsulfone, dicloran, dimethoate, dioxathion, disulfoton sulfone, EPN, fenoxycarb, fensulfothion, fenthion sulfone, folpet, methidathion, phosmet, quinomethionate, simazine, terbuthylazine

at RIKILT. Ideally, the reference spectra for each analyte in this study would have been generated on the same instrument (the tuning parameters and temperature settings had slight differences that undoubtedly led to slight spectral differences).

Qualitative Results for Calibration Standards. In the case of the human judgment approach, the analyst already knew that all of the pesticides had been added to each calibration standard at the stated concentrations; thus, no judgment was needed for identification purposes. In this situation, identification was a foregone conclusion and the analyst needed only to integrate the chromatographic peaks for the quantitation ion of each analyte (when a peak was present at the proper t_R). However, when so many analytes were present in a compressed chromatographic run of complex matrices, this is not necessarily a simple task, and some analytes were not detected, as described previously.[11]

Unlike the human-based approach, the automated software approach could not be forced to "over-ride" the identification criteria, which becomes a critical issue as demonstrated below. The software-only approach was first tested in this study using the reagent-only and matrix-matched calibration standards from 25 to 1000 ng/g. Table 2 lists the pesticides that could meet identification criteria with respect to the lowest concentration range that could be achieved in the five different matrices. Figure 3 graphically shows the percentage of the 151 analytes that could be identified at the given concentration standards in the different matrices. Certainly, a few specific examples occurred in which a particular pesticide/matrix pair made a difference in the lowest concentration of identification, but in general, the matrices did not make much of a difference, even compared to reagent-only standards.

Even at the highest concentrations, only 110−115 of the 151 pesticides could meet the identification criteria in the calibration standards using the automated software approach. Those pesticides consistently detected even at low levels appear at the top of Table 2, and the problematic ones appear at the bottom. The ones in the middle rows yield variable concentrations for identification because their spectral fit values tend to hover in the 600−700 region even for the higher calibration standards, often due to analyte−analyte coelutions and somewhat mismatched reference spectra. Degradation is one reason for the problems with some of the analytes listed at the bottom of Table 2, but in most of the cases for unidentified analytes, the pesticides gave poor chromatography, few MS ions (poor mass spectra), or peculiarities in the analysis or were buried by other coeluting analytes. This led to poor mass spectral deconvolution in the complex spectral chromatograms, particularly for those difficult analytes, and it was against the identification rules to use human intervention in this case.

Another factor as described in ref 11 is that we used a data acquisition rate of 10 spectra/s to lower detection limits, but higher acquisition rates (e.g., 40 spectra/s) tend to help the deconvolution process. Even at 10 spectra/s, we were not able to reach 10 ng/g limits of quantification for some analytes; thus, we could not increase the data acquisition rate to improve automatic deconvolution. Moreover, the mass spectra used as the references in the matching factor result were analyzed contemporaneously with the sample sequences in the study. In practice, the spectra from the calibration standards themselves could be used as the references (if they are free from interferences) to yield better matches. In many cases, the maximum matching factor achieved even for pesticides detected at levels <25 ng/g was a value of 800 rather than the ideal match of 1000.

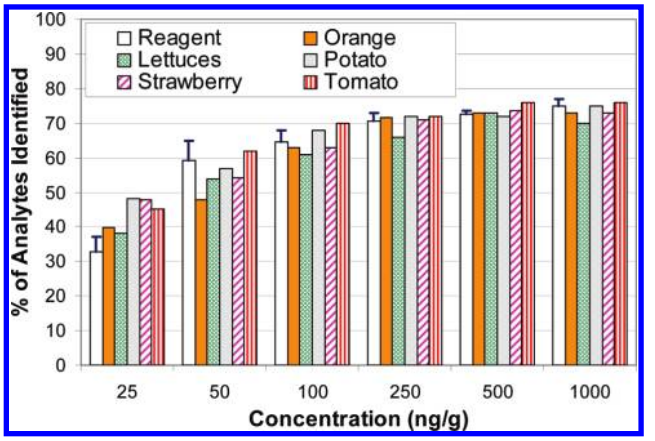

Figure 3. Automated software identification results for calibration standards containing the 151 analytes in the different matrices. The error bars for the reagent-only standards represent standard deviation from $n = 5$ sequences.

The analyte−analyte coelutions in the standards were found to be the most severe problem because, ultimately, the randomly added pesticides could be identified using the automated software approach at much lower concentrations in the samples than in the standards. Fewer pesticides were added in the blind samples; thus, there were fewer coelutions.

Qualitative Results for Blind Samples. In Table 3, we summarize the bottom-line factor in qualitative analysis, which is the rate of false presumptive positives and false negatives for all added and/or identified pesticides in the study. Those analytes in Tables 1 and 2 that do not appear in Table 3 were still monitored and quantitatively validated,[11] and they did not yield any false positives. To save space, the results for the pesticides that were added and/or identified in multiple matrices were combined, and the findings are summed for each column in the last row of the table. Altogether, 696 pesticide additions should have been made, which should have resulted in 12894 pesticide/matrix blank pairs for the 151 analytes in the 90 samples.

Table 4 summarizes the overall results using the identification criteria listed in Table 3 and additionally when the reporting limit was changed to ≥20 ng/g for the analyst review results and S/N ≥ 100 for the automated software results with forward-fit matching factors ≥600 and ≥700. These automated software results demonstrate the principle described previously in which the rate of false negatives increases as the rate of false positives decreases. The human analyst results did not show this relationship when the reporting limit was increased because one of the parameters in the study was that added concentrations were ≥25 ng/g; thus, only low-level (<10 or <20 ng/g) false presumptive positives were eliminated, not false negatives.

In all cases, the human analyst was able to correctly recognize through a host of objective factors in a similar way as the software approach (e.g., t_R, signal intensity, spectral matching factor, and peak width) but was also able to process subjective factors for each pesticide/matrix pair that the software was not programmed to take into account. The advantage of the heuristic approach of the analyst is that the rates of false positives and negatives were generally lower than the automated software-based detection, which used fixed identification criteria for all pesticides. This is demonstrated in Table 3 for many pesticides, including atrazine, bifenthrin, bromopropylate, famphur, fenarimol, iprodione,

Table 3. Results from the Study Comparing Human Decision-Making (≥10 ng/g Determinations) with the Excel-Based Software Using Forward-Fit Matching Factors of ≥600 or ≥700 and S/N ≥ 10 as Threshold Values

			true			false presumptives			false negatives		
				software			software			software	
analyte	matrices[a]	added	human	≥600	≥700	human	≥600	≥700	human	≥600	≥700
aldrin	O, S					2	3	2			
atrazine	P, S	18	18	12	10					6	8
δ-BHC	L, P, S	8	8	8	8						
bifenthrin	L, P, O, S						16	12			
bromophos	P	6	6	6	5						1
bromopropylate	all	13	13	13	13	5	11	8			
buprofezin	O						1				
cadusafos	O, T	9	9	8	6	1	1	1		1	3
captafol	P	11							11	11	11
carbaryl	L, P, T	33	3	7	7				30	26	26
cis-chlordane	L, P, S	8	8	8	8	2	5	1			
trans-chlordane	L, O, S					2	3				
chlorpropham	P						5				
chlorpyrifos	L	6	6	6	6						
coumaphos	S, T	16	15	12	8				1	4	8
λ-cyhalothrin	O, T	10	10	10	10		1				
cypermethrin	L	5	5	3	1					2	4
o,p′-DDD	P					3					
o,p′-DDE	O, S					3					
p,p′-DDE	L, O, S					1	7	2			
p,p′-DDT	S					4					
diazinon	T	11	11	11	11						
dichlorobenzophenone	P	11	10	11	11				1		
dicloran	L	12		2					12	10	12
dicrotophos	T						1				
dieldrin	P, S	5	5	5	5	3	4	2			
disulfoton	P	8	8	8	8						
α-endosulfan	O, P, S	19	18	19	19	5	4	2	1		
β-endosulfan	O, P					2	3	2			
endosulfan sulfate	O, S	16	16	16	14		1				2
endrin	L, O, S						6	2			
EPN	S	7	7	7	7						
esfenvalerate	all	20	20	20	20	1	5	2			
famphur	L, P	22	21	10	7				1	12	15
fenamiphos	S, T	23	23	21	16	1				2	7
fenarimol	L, O, S	7	7	7	7	2	12	6			
fenchlorphos	L	10	10	10	10						
fenpropathrin	O						1				
fenthion sulfone	O	12		1					12	11	12
fenvalerate	all	27	27	26	26	1	4			1	1
flucythrinate	S						1				
folpet	T	9		3	2				9	6	7
heptachlor	P	9	9	9	9						
heptachlor epoxide	O, S					2	5				
iprodione	L, P	12	11	1					1	11	12
isofenphos	L, O						3	1			
kepone	T	12	9	5	3				3	7	9
kresoxim-methyl	S	10	10	10	10		1				
leptophos	L	9	9	8	7					1	2

dx.doi.org/10.1021/jf104606j |*J. Agric. Food Chem.* 2011, 59, 7544–7556

Table 3. Continued

			true			false presumptives			false negatives		
				software			software			software	
analyte	matrices[a]	added	human	≥600	≥700	human	≥600	≥700	human	≥600	≥700
lindane (β+γ-BHC)	L	8	8	8	8						
metalaxyl	L, O					4	2	1			
methacrifos	S	7	7	7	6						1
methiocarb	L	3	3	1	1					2	2
metolachlor	L, P, S, T	5	5	5	5	7	5				
mevinphos	P,S	17	17	5	4					12	13
mirex	O, P, S, T	11	11	11	11	16	10	7			
myclobutanil	S	10	10	9	9	3	3	2		1	1
pendimethalin	P	9	9	9	9		2				
o-phenylphenol	P	11	11	10	8					1	3
phthalimide	T	6	3	2					6	3	4
cis-nonachlor	all	9	9	9	9	1	10	2			
trans-nonachlor	L, O, S					2	6	1			
oxadixyl	O, S	21	10	19	18	1	1		11	2	3
oxyfluorfen	L, O	17	17	17	14	1	1				3
penconazole	O	5	5	5	4	2	1	1			1
cis-permethrin	L, O, S	12	11	11	11	6	6	5	1	1	1
trans- permethrin	L, O, S	5	5	5	5	6	5	3			
phosmet	P	10	8	2	1				2	8	9
piperonyl butoxide	O					2					
procymidone	T	6	6	6	6						
propachlor	S	9	9	9	7						2
propoxur	L	11	11	11	10						1
propyzamide	P, S	21	21	21	21	2	1				
quintozene	T	6	6	6	6						
resmethrin	L, T	17	17	17	16						1
simazine	O, S	16		4	2				16	12	14
tebuconazole	O, S						3				
tetraconazole	L, S					6					
tetradifon	O, P, S	14	14	14	14	1	4	1			
tolclofos-methyl	O, P	16	16	12	12					4	4
totals		696	578	539	493	105	158	66	118	157	203

[a] L, lettuces; O, orange; P, potato; S, strawberry; T, tomato.

mevinphos, and phosmet. The analyst was able to achieve a slightly higher overall percentage of correct results of 98.4−98.8 versys 97.7−98.0% in the software approach, as shown in Table 4.

Despite the ability of the analyst to outperform the automated software approach, there are major practical drawbacks associated with the analyst review of results. For one, it took more than 20 working days for the analyst to review all of the chromatograms for all of the analytes and make the final decisions, whereas the software review took a matter of hours to process and compile. Another problem with analyst decision-making is that the decisions are quite variable from one person to another, and all humans make mistakes. Some mistakes can be discerned in Table 3 in which the software reports some identifications for certain difficult pesticides (folpet and carbaryl) that were missed by the analyst. The chemist knew that the method did not work well for these analytes and did not spend the time needed to find them in the blind samples. In quantitative validation studies, chemists know when the pesticides have been added and will integrate questionable peaks at the correct t_R to yield a quantitative result.

As emphasized in a previous paper,[16] human error is often to blame when clear misidentifications are made. One example concerns oxadixyl identified by the analyst in the orange extracts with no false negatives and a single false presumptive positive, but the analyst failed to report the positive results in the strawberry extracts. Figure 4 shows an example of the human error in blind strawberry extract 1. When informed of the false negatives, the analyst was able to review the chromatograms and make the identifications, but we decided to report the initial blind results because it serves as a real-world example, such as those listed in the previous paper,[16] that demonstrates human fallibility. We also learned after the initial report that the analyst had used the reference spectrum for atrazine at the retention time of simazine, which explains the false negatives for simazine in orange. Otherwise, the human with help from the computer outperformed the

Table 4. Rates of False Identifications Depending on Criteria Used for Identification in Each Approach[a]

			automated software			
	human judgment		forward fit ≥ 600		forward fit ≥ 700	
factor	≥10 ng/g	≥20 ng/g	S/N ≥ 10	S/N ≥ 100	S/N ≥ 10	S/N ≥ 100
false positives (%)	0.8	0.3	1.3	1.0	0.5	0.5
false negatives (%)	17.1	17.1	22.6	23.1	29.0	29.4
true (added) (%)	82.9	82.9	77.4	76.9	71.0	70.6
true (overall) (%)	98.4	98.8	97.7	97.8	98.0	98.0

[a] The elimination of captafol, carbaryl, folpet, fenthion sulfone, phthalimide, and simazine lowers the false negative rates to 5−10% in the approaches.

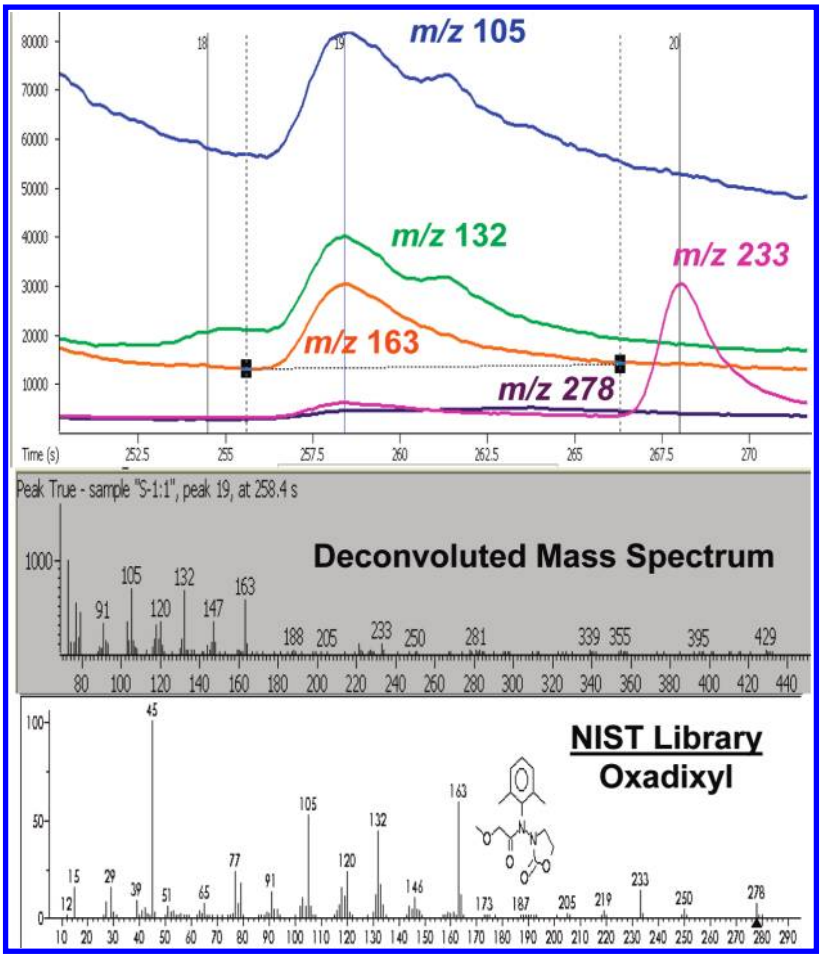

Figure 4. Example of a correct identification of oxadixyl added at 115 ng/g equivalent concentration in strawberry extract, but which was a false negative in the human assessment due to a simple mistake.

computer-only approach. Figure 5 further demonstrates how the analyst was better able to make identifications at lower analyte concentrations in the complex extracts, but some oversights were made at higher concentrations.

The false-negative rates were felt to be too high for regulatory qualitative screening purposes in both identification approaches, but this was a result of the randomly chosen analytes as much as the identification methods per se. The list of analytes giving false negatives in Table 3 consists mainly of known problematic analytes in GC (captafol, carbaryl, folpet, fenthion sulfone, and phthalimide). Dicloran in lettuces is the only pesticide that was surprisingly missed in the qualitative identifications (along with oxadixyl in strawberries and simazine in orange as stated above). All of these chemicals except folpet and captafol, which also degrade, are better determined and identified using LC-MS/MS methods. In the case of human judgment, the 84 false negatives for carbaryl, captafol, folpet, fenthion sulfone, phthalimide, and simazine make up 71% of the 118 false negatives altogether. If these are eliminated from the study, the 17% rate of false negatives becomes a much more respectable 5%, and nearly half of that if the mistake of oxadixyl in strawberry is allowed to be corrected.

However, software-based decision-making is the only practical option for routine analyses of so many analytes in complex

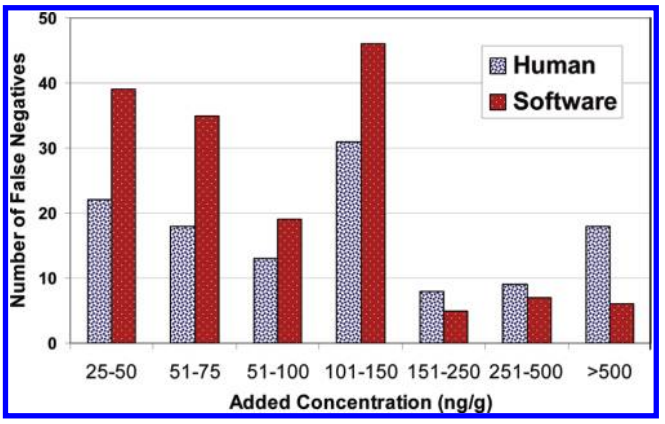

Figure 5. Comparison of human decision-making (calculated concentrations ≥ 10 ng/g) and automated software results (forward-fit matching factor ≥ 600) in the number of false negatives with respect to added analyte concentrations in the study.

matrices. The key to its successful implementation is to improve its performance. This can be done in the same way that human judgment gives better results, which is to incorporate greater flexibility in the identification criteria used. When the same fixed criteria are used for all analytes in all matrices, some criteria will be too strict for some pesticides and too loose for others, depending on the matrix, which leads to nonoptimal rates of false positives and negatives. The use of a higher S/N threshold of 100 rather than 10 would also help in this case (mainly due to carry-over problems), depending on the desired reporting limit. In practice, an empirical assessment of the matching factors and S/N results can be determined from the analyses of ≈20 true blanks from different sources compared with the measurements from the same number of analyte additions made at the reporting level. The thresholds could be set based on a statistical evaluation to minimize the rates of false results at the reporting level.

The 2002/657/EC rules for chemical residues in foods of animal origin use this concept in estimating the regulatory enforcement action limits for a determination,[12] but it does not use this concept for devising flexible identification criteria for qualitative assessments. Instead, it sets arbitrary, fixed criteria that do not necessarily best meet the analytical needs.[12,16] Although qualitative method validation requires additional work, we believe that a reasonable empirical validation is worth the effort to optimize the validation criteria and thereby reduce the rates of false positives and negatives in chemical residue monitoring programs. Modern MS instruments and software programs are able to handle and process a great deal of data to calculate the optimal identification criteria based on predetermined acceptable rates of false positives and negatives if a fixed empirical validation protocol is conducted. The SANCO guidelines[13] for pesticide residue analysis in foods permit this type of approach, and we intend to continue the pursuit to devise empirical validation protocols using modern instruments and software.

In conclusion, the combination of QuEChERS sample preparation and LP-GC-ToF analysis was shown to provide high sample throughput to yield acceptable quantitative and qualitative results for nearly 150 pesticide residues in fruits and vegetables. The use of more selective and sensitive GC×GC-ToF[31,32] would provide better performance, but at the cost of 4-fold greater time per sample. The fast GC-MS method attains the same chromatographic analysis times as UHPLC-MS/MS to enable concurrent analysis covering hundreds of GC- and LC-amenable pesticides with typical detection limits near 10 ng/g or less. Although the pesticide detection limits of the ToF instrument are not as low as in MS/MS, it provides full mass spectral data acquisition, which yields greater confidence in identifications than SIM methods, and permits postrun searching for nontargeted chemicals in the sample extracts if so desired. The analyst review of chromatograms was shown to yield lower rates of false positives and false negatives in this study than the automated software, but it is impractical for high-throughput monitoring applications of so many analytes in complex matrices. A dedicated software approach is the faster alternative more appropriate for this application, and its performance can be improved by setting specific, empirically determined identification criteria for each analyte/matrix pair.

■ ASSOCIATED CONTENT

Supporting Information. Figure S1 and spreadsheets for the blind spiking scheme. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: steven.lehotay@ars.usda.gov. Phone: (215) 233-6433. Fax: (215) 233-6642.

Funding Sources

This work was supported by U.S.−Israel Binational Agricultural Research and Development Fund Grant US 4273-09. This research was partially supported by The Royal Golden Jubilee Ph.D. Program (PHD0036/2550), Chulalongkorn University Graduate Scholarship to Commemorate the 72nd Anniversary of his Majesty King Bhumibol Adulyadej, and the Thailand Research Fund and Commission on Higher Education, Research Grant for Mid-Career University Faculty (TRF-CHE-RES-MR) (RMU518009). Mention of brand or firm names does not constitute an endorsement by the U.S. Department of Agriculture above others of a similar nature not mentioned.

■ REFERENCES

(1) Anastassiades, M.; Lehotay, S. J.; Štajnbaher, D.; Schenck, F. J. Fast and easy multiresidue method employing acetonitrile extraction/partitioning and "dispersive solid-phase extraction" for the determination of pesticide residues in produce. *J. AOAC Int.* **2003**, *86*, 412–231.

(2) Li, L.; Li, W.; Qin, D.; Jiang, S.; Liu, F. Application of graphitized carbon black to the QuEChERS method for pesticide multiresidue analysis in spinach. *J. AOAC Int.* **2009**, *92*, 538–547.

(3) Lehotay, S. J. Determination of pesticide residues in foods by acetonitrile extraction and partitioning with magnesium sulfate: collaborative study. *J. AOAC Int.* **2007**, *90*, 485–520.

(4) European Committee for Standardization (CEN) Standard Method EN 15662; Food of Plant Origin − Determination of Pesticide Residues using GC-MS and/or LC-MS/MS following Acetonitrile Extraction/Partitioning and Clean-Up by Dispersive SPE − QuEChERS Method, 2008; www.cen.eu and www.quechers.com.

(5) Lehotay, S. J.; Son, A. K.; Kwon, H.; Koesukwiwat, U.; Fu, W.; Mastovska, K.; Hoh, E.; Leepipatpiboon, N. Comparison of QuEChERS sample preparation methods for the analysis of pesticide residues in fruits and vegetables. *J. Chromatogr., A* **2010**, *1217*, 2548–2560.

(6) Mastovska, K.; Dorweiler, K. J.; Lehotay, S. J.; Wegscheid, J. S.; Szpylka, K. A. Pesticide multiresidue analysis in cereal grains using modified QuEChERS method combined with automated direct sample

introduction GC-TOFMS and UPLC-MS/MS techniques. *J. Agric. Food Chem.* **2010**, *58*, 5959–5972.

(7) Koesukwiwat, U.; Lehotay, S. J.; Mastovska, K.; Dorweiler, K. J.; Leepipatpiboon, N. Extension of the QuEChERS method for pesticide residues in cereals to flaxseeds, peanuts, and doughs. *J. Agric. Food Chem.* **2010**, *58*, 5950–5958.

(8) Payá, P.; Anastassiades, M.; Mack, D.; Sigalova, I.; Tasdelen, B.; Oliva, J.; Barba, A. Analysis of pesticide residues using the quick easy cheap effective rugged and safe (QuEChERS) pesticide multiresidue method in combination with gas and liquid chromatography and tandem mass spectrometric detection. *Anal. Bioanal. Chem.* **2007**, *389*, 1697–1714.

(9) Lehotay, S. J.; Anastassiades, M.; Majors, R. E. QuEChERS, a sample preparation technique that is "catching on": an up-to-date interview with its inventors. *LC GC North Am.* **2010**, *28*, 504–516.

(10) Wilkowska A.; Biziuk, M. Determination of pesticide residues in food matrices using the QuEChERS methodology. *Food Chem.* **2011**, *125*, 803−812.

(11) Koesukwiwat, U.; Lehotay, S. J.; Miao, S.; Leepipatpiboon, N. High throughput analysis of 150 pesticides in fruits and vegetables using QuEChERS and low-pressure gas chromatography−time-of-flight mass spectrometry. *J. Chromatogr., A* **2010**, *1217*, 6692–6703.

(12) European Commission. Commission Decision 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. *Off. J. Eur. Communities* **2002**, *L221*, 8–36.

(13) Method Validation and Quality Control Procedures for Pesticide Residues Analysis in Food and Feed, Document No. SANCO/10684/2009.

(14) U.S. Dept. of Agriculture Agricultural Marketing Service, Pesticide Data Program, SOP No.: PDP-QC - Chemical Compounds, PDP Commodity Grouping, Method Validation and Quality Control, 2010; www.ams.usda.gov/science/pdp/.

(15) Fajgelj, Á., Ambrus, A., Eds. *Principles and Practices of Method Validation*; Royal Society of Chemistry: Cambridge, U.K., 2000.

(16) Lehotay, S. J.; Mastovska, K.; Amirav, A.; Fialkov, A. B.; Martos, P. A.; de Kok, A.; Fernández-Alba, A. R. Identification and confirmation of chemical residues in food by chromatography−mass spectrometry and other techniques. *Trends Anal. Chem.* **2008**, *27*, 1070–1090.

(17) U.S. Food and Drug Administration, Center for Veterinary Medicine, Guidance for Industry 118, Mass Spectrometry for Confirmation of the Identity of Animal Drug Residues, 2003.

(18) U.S. Food and Drug Administration, Office of Regulatory Affairs, Guidance for the Analysis and Documentation to Support Regulatory Action On Pesticide Residues, Document ORA-LAB.10 version 1.0, Nov 6, 2009.

(19) Bethem, R. A.; Boison, J.; Gale, J.; Heller, D.; Lehotay, S.; Loo, J.; Musser, S.; Price, P.; Stein, S. Establishing the fitness for purpose of mass spectrometric methods. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 528–541.

(20) Heller, D. N.; Lehotay, S. J.; Martos, P. A.; Hammack, W.; Fernandez-Alba, A. R. Issues in mass spectrometry between bench chemists and regulatory laboratory managers. *J. AOAC Int.* **2010**, *93*, 1625–1632.

(21) Milman, B. L. Identification of chemical compounds. *Trends Anal. Chem.* **2005**, *24*, 493–508.

(22) Lehotay, S. J.; Gates, R. A. Blind analysis of fortified pesticide residues in carrot extracts using GC-MS to evaluate qualitative and quantitative performance. *J. Sep. Sci.* **2009**, *32*, 3706–3719.

(23) Meyer, M. R.; Peters, F. T.; Maurer, H. H. Automated mass spectral deconvolution and identification system for GC-MS screening for drugs, poisons, and metabolites in urine. *Clin. Chem.* **2010**, *56*, 575–584.

(24) Norli, H. R.; Christiansen, A.; Holen, B. Independent evaluation of a commercial deconvolution reporting software for gas chromatography−mass spectrometry analysis of pesticide residues in fruits and vegetables. *J. Chromatogr., A* **2010**, *1217*, 2056–2064.

(25) Guillarme, D.; Schappler, J.; Rudaz, S.; Veuthey, J.-L. Coupling ultra-high-pressure liquid chromatography with mass spectrometry. *Trends Anal. Chem.* **2010**, *29*, 15–27.

(26) Grimalt, S.; Sancho, J. V.; Pozo, O. J.; Hernández, F. Quantification, confirmation and screening capability of UHPLC coupled to triple quadrupole and hybrid quadrupole time-of-flight mass spectrometry in pesticide residue analysis. *J. Mass Spectrom.* **2010**, *45*, 421–436.

(27) Maštovská, K.; Lehotay, S. J. Practical approaches to fast gas chromatography−mass spectrometry. *J. Chromatogr., A* **2003**, *1000*, 153–180.

(28) de Zeeuw, J.; Peene, J.; Janssen, H. G.; Lou, X. A simple way to speed up separations by GC-MS using short 0.53 mm columns and vacuum outlet conditions. *J. High Resolut. Chromatogr.* **2000**, *23*, 677–680.

(29) Maštovská, K.; Lehotay, S. J.; Hajšlová, J. Optimization and evaluation of low-pressure gas chromatography−mass spectrometry for the fast analysis of multiple pesticide residues in a food commodity. *J. Chromatogr., A* **2001**, *926*, 291–308.

(30) Mastovska, K.; Lehotay, S. J. Evaluation of common organic solvents for gas chromatographic analysis and stability of multiclass pesticide residues. *J. Chromatogr., A* **2004**, *1040*, 259–272.

(31) van der Lee, M. K.; van der Weg, G.; Traag, W. A.; Mol, H. G. Qualitative screening and quantitative determination of pesticides and contaminants in animal feed using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry. *J. Chromatogr., A* **2008**, *1186*, 325–339.

(32) Mol, H. G. J.; van der Kamp, H.; van der Weg, G; van der Lee, M. K.; Punt, A.; de Rijk, T. C. Validation of automated library-based qualitative screening of pesticides by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry. *J. AOAC Int.* **2011**, in press.

(33) Fillion, J.; Sauvé, F.; Selwyn, J. Multiresidue method for the determination of residues of 251 pesticides in fruits and vegetables by gas chromatography/mass spectrometry and liquid chromatography with fluorescence detection. *J. AOAC Int.* **2000**, *83*, 698–713.

(34) Erney, D. R.; Pawlowski, T. M.; Poole, C. F. Matrix-induced peak enhancement of pesticides in gas chromatography: is there a solution? *J. High Resolut. Chromatogr.* **1997**, *20*, 375–378.

(35) Zhang, W.; Wu, P.; Li, C. Study of automated mass spectral deconvolution and identification system (AMDIS) in pesticide residue analysis. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 1563–1568.