Possible False Positive Identifications in Gas Chromatography / Mass Spectrometry (GC/MS) Analysis during Chemical Weapons Convention (CWC) Inspections

W. G. Mallard
National Institute of Standards and Technology (NIST)
United States of America

In this work the approved OPCW Analytical Database (340 spectra) is compared to the NIST/EPA/NIH Mass Spectral Database (108,000 chemicals, 129,000 spectra) using identification algorithms developed at NIST. A very small number of possible false positive identifications are found. The use of retention index data to eliminate these false positive identifications is discussed.

.Mallard, W. G. ,"Possible False Positive Identification in Gas Chromatography / Mass Spectrometry (GC/MS) Analysis in Chemical Weapons Convention (CWC) Inspections", *Sixth International Symposium on Protection Against Chemical and Biological Warfare Agents*, Defense Research Establishment, Stockholm Sweden, 1998, p 266.

False Positive Risks in GC/MS Analysis under the CWC

Gary Mallard
NIST

Algorithmic Comparison of Mass Spectra

The Chemical Weapons Convention (CWC) calls for on-site analytical inspections to occur in a manner that does not compromise confidential business or national security information unrelated to the treaty. To do this it is essential to automate the analysis of the GC/MS data during an inspection.

The algorithm to compare spectra has been previously developed. A match factor (MF) is defined to provide a quantitative measure of the similarity between two spectra. Since high mass peaks are the most characteristic peaks for a mass spectrum, some additional weight is given to these peaks. An optimized function for compound identification is a scaled dot product. (Stein & Scott, 1994)

$$MF = \frac{\left(\sum_i L_i D_i\right)^2}{\left(\sum_i L_i^2\right)\left(\sum_i D_i^2\right)}$$

L_i and D_i are the scaled intensities at mass i for the Library and Database spectra.

Scaled Intensity = $(Intensity)^{1/2}(Mass)$

For the "normal" match factor the summation includes all masses. For the "reverse" match factor the summation includes only the masses in the library spectrum.

The net match factor is defined as:
Net = 0.7(Normal) + 0.3(Reverse)

A net match factor greater than 80 is considered an identification.

The use of the reverse match factor provides a method of eliminating spurious peaks, much as is done by a mass spectrometrist in comparing the data to the library spectrum.

Methodology for analysis of OPCW Analytical Database for False Positive Risks

- The NIST database is primarily composed of complete spectra with no artificial peak additions. The database is designed to represent the compounds found in commerce and the environment. It does not contain a high proportion of data that have only been used to prove synthesis. As a result, data taken from the NIST database will represent spectra found in normal analysis using GC/MS instruments. The current version NIST 98 contains spectra on about 108,000 compounds with an additional 21,000 replicate spectra.

- The Organization for the Prohibition of Chemical Weapons (OPCW) is responsible for implementing the CWC. The current OPCW official analytical database has 340 mass spectra that have been very thoroughly evaluated by an international group of mass spectrometrists. All of these spectra are from chemicals that are in some way regulated under the terms of the CWC.

- Each of these spectra was compared to the NIST database by treating the spectrum as arising from an unknown. The resulting hit lists were analyzed and all scheduled chemicals were deleted.

- As noted above, both the forward and reverse match factors are used in developing a net match factor. The use of the reverse match factor provides a method for discounting spurious peaks that arise due to incomplete separation in a GC column.

- Because the hit list for each spectrum was analyzed independently, there are cases where a single compound appeared on more than one hit list - this is especially true for cases where the OPCW database had multiple spectra for a single compound. No attempt has been made to eliminate compounds that matched more than one scheduled compound; therefore:

- **The data shown represent an upper limit to the number of compounds that may give possible false positives.**

False Positive Identifications

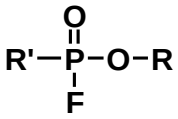
Number of non-scheduled chemicals matching scheduled spectra			
OPCW Spectra		Match Factor > 80	Additional Matches due to "Reverse"
All Schedules	340	59	35
Schedule 1	164	33	22
Schedule 2	157	21	13
Schedule 3	19	5	0

Schedule 1 chemicals are all very toxic and contain all of the chemicals used in weapons. Schedule 2 contains a few less toxic chemicals and a large number of chemicals that are used in the production of Schedule 1 chemicals. Schedule 3 contains chemicals that are used extensively in industry but are either toxic or widely used in weapons production.

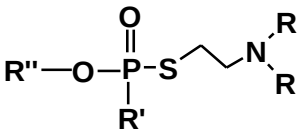
The number of false positive matches is about 0.1% of the total NIST database of 108K compounds. The overall rate is very low. However a few of the chemicals that have high match factors with spectra of scheduled chemicals are common.

No known weaponized chemical produced a false positive identification.

Schedule 1 chemicals producing false positives



Schedule 1.A.01 chemicals have R' = methyl, ethyl, propyl or isopropyl and R=alkyl or cycloalkyl groups with less than or equal to 10 carbon atoms. Sarin and Soman are members of this group.



Schedule 1.A.03 chemicals have R and R'= methyl, ethyl, propyl or isopropyl, R" = alkyl or cycloalkyl groups with less than or equal to 10 carbon atoms. VX is a member of this group.

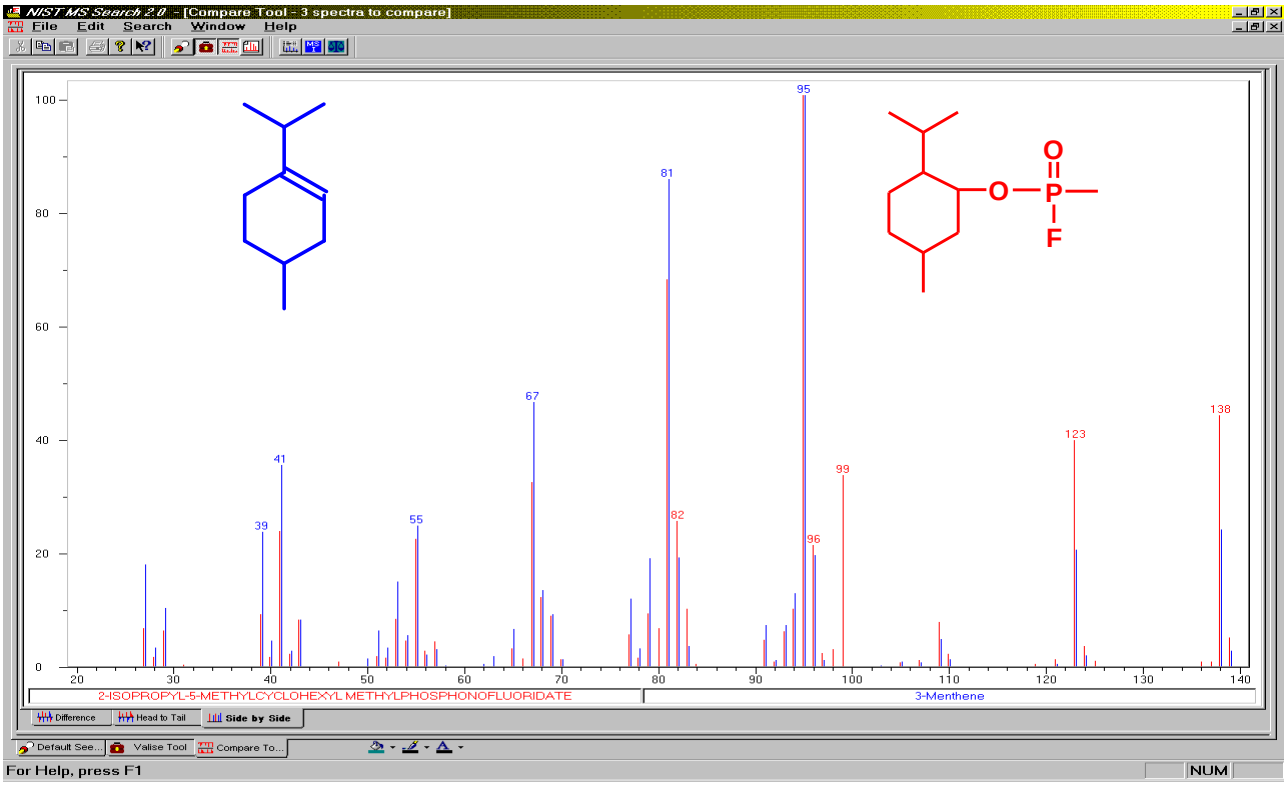
The schedule 1 false positive identifications all come from the 1.A.01 and 1.A.03 schedules. In the case of the O-alkyl phosphonofluoridates (1.A.01) as the O-alkyl side chain became larger and more branched, the chance of false positives increased. The side chain comes to dominate the mass spectrum in contrast to soman, sarin and GF where the $CH_3PF(OH)_2$ at m/z 99 dominates the spectrum. Most of the false positives arise from the 1.A.01 chemicals. The alkyl S-2-(dialkyly)ethyl alkyl phosphonothiolates (1.A.03), the dominance of the dialkyl amino ethyl group will cause a very small number of false positives.

No false positives were found for schedule 1.A.02 (Tabun family), the 1.A.04 or 1.A.06 (the sulfur and nitrogen mustards) or the 1.A.05 (lewisites).

Shown below are scheduled chemicals with false positives.

Schedule 1 Chemicals with Spectra that Have High False Identification Risk		
Schedule	**Name of scheduled chemical**	**Matching unscheduled chemicals**
1.A.01	2-isopropyl-5-methylcyclohexyl methyl phosphonofluoridate	31
1.A.01	2,6-dimethylcyclohexyl methylphosphonofluoridate	8
1.A.01	4-t-butylcyclohexyl methyl phosphonofluoridate	2
1.A.01	2-ethylcyclohexyl methyl phosphonofluoridate	2
1.A.01	1-ethylhexyl methylphosphonofluoridate	4
1.A.03	ethyl S-2-(diethylamino)ethyl ethylphosphonothiolate	3
1.A.03	ethyl S-2-(diethylamino)ethyl ethylphosphonothiolate	1
1.A.03	ethyl S-2-(dimethylamino)ethyl methylphosphonothiolate	1
1.A.03	cyclohexyl s-2-diethylaminoethyl methylphosphonothioate	1
1.A.03	isopropyl S-2-(dimethylamino)ethyl ethylphosphonothiolate	1

Spectral similarity between scheduled and non-scheduled chemicals

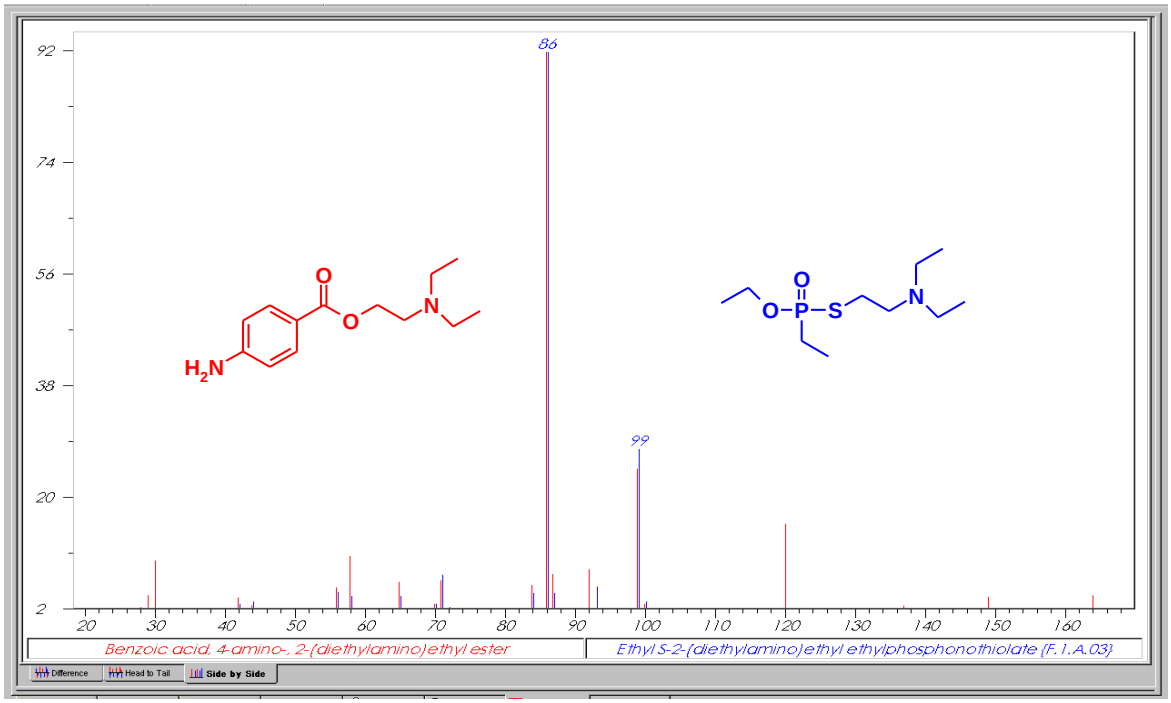


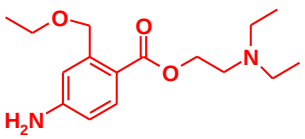
The two spectra- shown above are 3-menthene and a Schedule 1 compound. Only the characteristic 99 peak is absent from the 3-menthene library spectrum. The use of this characteristic peak can be made a part of the algorithm, but the overall match is very good - even for the high mass peaks. Menthene is a naturally occurring terpene.

The common schedule 1.A.01 chemicals used in weapons such as sarin, soman and GF all have strong characteristic peaks at 99 and few other major peaks. As the hydrocarbon fragment becomes larger and more highly branched, the relative intensity of the side chain becomes greater. Thus the scheduled chemical looks like the corresponding alkane, alkene or alcohol.

False Positives due to Spectral Similarity

For the schedule 1.A.03 and some schedule 2 compounds, the spectral similarity occurs because in tertiary amine the $N(R_2R')^+$ ion is very stable and it and its decomposition products dominate the mass spectrum. The result is shown below. There is very little (only the 120) that would differentiate Novocain from the agent. The net match factor between these two compounds is 80 and so right at the border.





The spectrum of Pravocaine, another anesthetic also has a high match factor with the scheduled compound above (net match factor = 82).

Retention Index Data Measured and Estimated

While the overall risk of false positive identification is low, it is possible to reduce it even further. The analysis of the sample with GC/MS not only gives the mass spectrum of the compound, but also the retention time. This retention time, normalized into a retention index, can also be used in identification - and especially in denying identification.

Comparison of Retention Indices for Phosphonofluoridates with Related Chemicals							
	RI for Phosphonofluoridate		Alcohol RI	Alkane RI	Alkene RI	Difference between phosphonofluoridate and	
	Methyl[1]	Ethyl				Alcohol	Alkane
ethyl		866	500				
2-propyl	820	907	500			407	
1-propyl	869	967	535	300	293	432	667
1-propyl-2-methyl	*936*	1028	616	362	390	412	666
2-butyl	*923*	1015	591			424	
1-butyl	971	1067	655	400	391	412	667
1-pentyl	1073	1165	756	500	593	409	665
2-butyl-3,3-dimethyl	1046	1141		567			574
1-hexyl	1172	1265	858	600	591	407	665
cyclohexyl	1208	1307	880	658		427	649
1-Heptyl	1272	1364	957	700		407	664
1-octyl	*1372*	1464	1061	800		403	664
1-nonyl	*1472*	1564	1161	900		403	664
		Number Points				11	10
		Standard Deviation[2]				10.0	28.8
		Average				413	655

[1] *Estimates in italics*
[2] *Estimated values not included in statistical analysis*

As can be seen from the data shown above, the difference in the retention index of the scheduled compound and the hydrocarbon is large. Using the retention index to eliminate chemicals that are spectrally similar would reduce the false positives due to the schedule 1.A.01 compounds to zero.

Furthermore, the very small standard deviation of the difference between the alcohol and the corresponding phosphonofluoridate means that the retention index of the alcohol could be used to predict the retention index of the schedule compound with high confidence.

Schedule 2 and 3 False Positive Identifications

While the problems of schedule 2 and 3 chemicals are not as severe, a few points need to be made.

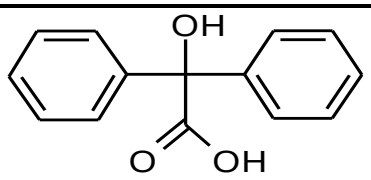
Most classification of schedule 2 chemicals are used in the production of common chemical agents in wide spread military use. There is an important group that has a chemical structure that is similar to the agents - these are the schedule 2.B.04. All of the accepted spectra in 2B.04 as well as the other schedule 2 compounds were examined. Only a very small set of compounds gave false positives. Some of these are given below.

The high usage of C6 alcohols in industry has caused some concern with possible false positives due the presence of pinacolyl alcohol in schedule 2. Pinacolyl alcohol (3,3-dimethylbutan-2-ol) is the alcohol used in the production of soman (GD) and has essentially no commercial uses. This study showed that the only chemical that gave a false positive was 3,3-dimethylbutan-1-ol. The net match factor here was in the range of 75-83. Retention time data for these two species will be fairly close together and so it will be essential to have measured data.

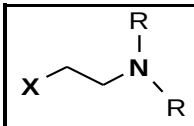
There is also a well known similarity between chloropicrin and carbon tetrachloride. This is a clear case where the use of retention time data will reduce the false positives.

Schedule 2 Problems

The table below illustrates a consequence of the way in which the OPCW database was built. Mass spectral data for all compounds that were scheduled are included. Benzilic acid is used in the manufacture of BZ - a hallucinogenic agent. Benzilic acid will not elute from a GC column, but the presence of the benzilic acid spectrum in the database means that it may provide false positives for compounds which can be detected by GC/MS.

Benzilic Acid Schedule 2.B.08	Possible False Positives	CAS Number
	1,2-ethanediol, 1,1,2,2-tetraphenyl methyl benzilate phenyl 4-pyridyl ketone 2-t-butyl-5-(hydroxydiphenylmethyl)-5-methyl[1,3]dionolan-4-one t-butyldiphenylmethanol	464-72-2 76-89-1 14548-46-0 81286-89-7 1657-60-9

Similar problems occur with the schedule 2.B.10 and 2.B.12 compounds which either do not elute from a GC column or elute with very broad peaks. But, many of the compounds which are identified as the scheduled compounds can be detected with GC/MS.

	**Scheduled Compound CAS Number**	**Number of nonscheduled compounds with net match factors > 80**
R = ethyl X = SH Schedule 2.B.12	**100-38-9**	**11**
R = isopropyl X = SH Schedule 2.B.12	**5842-07-9**	**2**
R = ethyl X = Cl Schedule 2.B.10	**100-35-6**	**1**

Conclusions and Recommendations

The spectral similarity between some scheduled chemicals and commonly occurring chemicals may lead to false positives in on-site inspections.

The vast majority of these possible false positives may be eliminated by the use of retention index data. For most scheduled chemicals the estimated retention index will be sufficient.

A limited experimental program should be carried out to measure, using standard on-site techniques, retention indices of a series of alcohols and the corresponding methyl, ethyl, propyl and iso-propyl phosphonofluoridates. These would then be the basis of a complete estimation scheme for schedule 1.A.01. Additional measurements on schedule 1.A.03 will also be essential.

The full automated evaluation of the GC/MS data file requires not only the comparison of spectra, but more crucially the extraction of the spectra from a potentially complex matrix. To evaluate the risk of false positives from the entire system, a large set of GC/MS data files have been examined with the on-site software and the official OPCW database. The goal is to provide another method of accessing the false positive risk during an inspection. The analysis of 50,000 GC/MS data files will be reported in the near future.

Testing for False Positives

30,000 GC/MS files
None expected to have CW agents
1,000,000 Components

False Positives		
Agent	MF = 70-79	MF >80
Sarin	1*	0
Soman,Mustard, VX	0	0
* MF = 72		

In this test a very large number of data files were processed using the blinded software that is used by the OPCW. No true false positives were found. The question that remains is how typical of all cases was this. The data source was largely environmental samples that have been taken in EPA analysis. These have a very large number of compounds in them, many that are not in commerce, but by no means all that are in commerce were necessarily in this set. This is a different method of looking at the problem and in many ways duplicates the situation in the field better than testing against all of the data in the MS database. Both methods give results that indicate the risk of false positives is very low.