Eurachem

CITAC
Cooperation on International Traceability in Analytical Chemistry

EURACHEM / CITAC Guide

ارزیابی عملکرد و عدم قطعیت در تجزیه شیمیایی کیفی

AQA 2021

من شخصاً چیزی دیدم شبیه به یک موجود زنده، اما نمی توانم کاملاً مطمئن باشم که یک ماده معدنی نبود. فکر می کنم آنچه بود انرژی اش از ماده بیشتر بود. اگر بخواهم به صورت نسبی بگویم، ساده‌تر است که همه چیز را به‌عنوان پدیده‌ای توصیف کنم که در مرزی با ابعاد و نام‌ها، روی تکیه‌گاه رنگ، شکل، بو، جرم، طول و عرض، خطوط، سایه‌ها، تاریکی و غیره معلق است و غیره ساده ترین چیز خواهد بود.

اسلاوومیر مروژک، "استریپ تیز"

ترجمه ادوارد روترت

ارزیابی عملکرد و عدم قطعیت در تجزیه وتحلیل شیمیایی کیفی

ویرایش اول (۲۰۲۱)

ویراستاران: ریکاردو بتنکورت داسیلوا (لیسبون) استفن ال آر الیسون (انگلستان)	
ترکیب گروه کاری:	
اعضای Eurachem آر. بتنکورت داسیلوا (دانشگاه لیسبون- پرتقال) – رئیس اس. آلیسون (منشی) – انگلستان الف. توگولا – فرانسه دی. ایوانووا – اوراشیم بلغارستان ای. تئودورسون – سوئد ای. توتو – دانشگاه پولیتنیکا بوخارست – رومانی اچ. ایمونز – کمیسیون اتحادیه اروپا	آی لیتو – دانشگاه تارتو – استونی ام. سگا – ایتالیا ا. لوبرگ – اوکراین ا. پلگرینو – پرتغال پی. پریرا – پرتقال آر. کائوس – آلمان اس. لاردی فونتان – فرانسه آ. وسچیدر – استرالیا
	تشکر و قدردانی این سند توسط کارگروه Eurachem/CITAC با ترکیب نشان داده شده در سمت راست تهیه و تدوین شده است. ویراستاران از همه سازمان ها و افراد قدردانی می نماید. همچنین افرادی که در ارائه نظرات، مشاوره و همکاری شرکت داشته اند. تولید این محتوا به وسیله سازمان علم و تکنولوژی کشور پرتغال مورد حمایت واقع گردید.
اعضای CITAC الف. بوتا – جنوب افریقا الف. لورنسو – برزیل	**ترجمه به فارسی:** دکتر مهدی نبی E-mail: mhdnabi@yahoo.com دکتر سارا براتی

توصیه می شود این سند به صورت زیر ذکر شود:

آر. بتنکورت داسیلوا و اس ال آر الیسون (ویرایشگران) راهنمای Eurachem/CITAC: ارزیابی عملکرد و عدم قطعیت در تجزیه وتحلیل شیمیایی کیفی. ویرایش اول، یوراکم (۲۰۲۱).ISBN 978-0-948926-39-6

سایت: http://www.eurachem.org

فهرست

پیش گفتار ۶

دامنه کاربرد ۷

۱ مقدمه ۱۰

۲ انواع تجزیه کیفی ۱۳

۳ ارزیابی عملکرد برای تجزیه کیفی ۱۵

۱.۳ اصطلاحات کلی ۱۵

۲.۳ کمی سازی عملکرد تجزیه کیفی ۱۶

۳.۳ ارزیابی نرخ های مثبت کاذب و منفی کاذب ۲۱

۴.۳ حد تشخیص و گزینش پذیری ۲۷

۴ عبارت های اطمینان در تجزیه کیفی ۲۹

۱.۴ ملاحظات کلی ۲۹

۲.۴ نسبت عدد احتمال ۲۹

۳.۴ احتمال پسین ۳۳

۴.۴ قابلیت اطمینان معیارها ۳۴

۵.۴ عدم قطعیت نسبت ها ۳۵

۵ گزارش نتیجه تجزیه ای کیفی ۳۸

۶ نتیجه گیری و توصیه ۳۹

۷ مثال ها ۴۰

۱.۷ م۱ شناسایی ترکیبات توسط طیف سنجی جرمی با تفکیک پذیری پایین با استفاده از جستجوی پایگاه داده یا وجود یونهای مشخصه ۴۰

۲.۷ م۲ شناسایی ترکیبات خالص شده توسط طیف سنجی مادون قرمز ۴۷

۳.۷ م۳ شناسایی دارو های غیر مجاز در ادرار توسط فن ایمونو اسی چندگانه آنزیمی (**EMIT**) ویک فن جایگزین ۵۰

۴.۷ م۴ شناسایی ژن **SRY** انسانی در موارد بیولوزیکی با**qPCR**: ۵۳

۵.۷ م۵ شناسایی بقایای آفت کش ها در مواد غذایی توسط **GC-MS/MS** بر اساس زمان بازداری و نسبت فراوانی یون ۵۵

۶.۷ م۶ شناسایی **RNA SARS-COV-2** با آزمایش تکثیر نوکلئیک اسید ۵۹

پیوست الف) تئوری بیز، احتمال نسبی و نسبت عدد احتمال ۶۲

الف.۱ تئوری بیز ۶۲

الف.۲ احتمال و احتمال نسبی ۶۳

الف.۳ شکل احتمالات نسبی تئوری بیز و نسبت عدد احتمال ۶۳

پیوست ب) تجزیه کیفی مرتبط با ارزیابی انطباق با یک حد کمی ۶۵
کتابشناسی ۷۰

پیش گفتار

در متون به موضوع ارزشیابی و بیان عدم قطعیت در تجزیه شیمیایی کیفیف بسیار کمتر از عدم قطعیت تجزیه کمی (یعنی اندازه گیری ها) پرداخته شده است [۱]. با این که بعضی از نویسندگان، به این حوزه پرداخته اند [۲]- [۱۱]، راهنمایی کلی برای ارزیابی عملکرد در تجزیه کیفی یا ارزشیابی و گزارش دهی عدم قطعیت تجزیه کیفی، کمیاب می باشند.

درحال حاضر انتظار نمی رود آزمایشگاه های اعتباردهی شده، عدم قطعیت های مربوط به نتایج تجزیه کیفی را ارزشیابی کنند یا گزارش دهند [۱۲]. با این حال، ISO/IEC 17025 [۱۳] و ISO 15189 [۱۴] هر دو نیازمند آزمایشگاه هایی هستند که مطمئن باشند می توانند نتایج تجزیه های کیفی و کمی معتبری را به دست آورند.آگاهی از قابلیت اطمینان نتایج تجزیه های کیفی، برای آزمایشگاه ها نیز بسیار مهم می باشد. این موضوع، آن ها را قادر می سازد، در صورت لزوم، محدودیت ها در تفسیر نتایج را هشدار دهند و پاسخ صحیح به سوالات مشتری ها در مورد قابلیت اطمینان، ارائه نمایند. وقتی که احتمال نتایج کاذب بیشتر است، ارزیابی کمی از قابلیت اطمینان نتایج تجزیه کیفی، بسیار مفید است. این راهنما برای استفاده در هنگام ارزیابی کمی قابلیت اطمینان در یک تحلیل کیفی، درنظر گرفته شده است.

این راهنما متکی به تجربیات چندین زمینه تجزیه ای است که تجزیه کیفی به صورت مکرر در آن استفاده می شود مثل پزشکی قانونی [۱۵] و زمینه های بالینی [۱۶]- [۱۸] و راهنمایی های عمومی گسترده [۷].

دامنه کاربرد

این راهنما برای کمک به آزمایشگاه ها، در تنظیم و اجرای روش های مناسب برای ارزیابی عملکرد روش های تجزیه کیفی و ارزیابی عدم قطعیت ها در تجزیه های شیمیایی کیفی در نظر گرفته شده است.

در این راهنما، تجزیه کیفی به عنوان «دسته بندی بر اساس معیارهای مشخص شده» تعریف شده است. برای شیمی تجزیه و رشته های مرتبط، به طور کلی معیارها، به اطلاعات جهت مشخص شدن ترکیبات شیمیایی، خواص و ساختارهای اقلام مورد تجزیه مرتبط هستند.

انواع معیارهای زیر، دراین راهنما در نظر گرفته شده است:

- معیارهای کمی که در آن نتایج عددی برای طبقه بندی یک قلم آزمون متعلق به یک دسته[1] از پیش تعیین شده، استفاده می شود.
- معیارهای کیفی مانند حضور یا عدم حضور یک ویژگی خاص، تغییر رنگ در یک آزمون و غیره.

این راهنما زمانی که به توصیف ابزارهای قابل دسترس برای ارزیابی عملکرد روش های تجزیه کیفی و ارزیابی نتایج کیفی عدم قطعیت می پردازد، جامع نمی باشد. ویژگی های عملکردی ارائه شده در این راهنما، بر اساس نرخ نتایج کاذب اندازه گیری شده یا تخمین زده شده، مبتنی می باشد و برای مثال، معیارهای توافقی بین روش های کیفی یا پردازش دسته بندی بر اساس مقیاس های ترتیبی[2]، غیر از دسته بندی درست یا نادرست، را در نظر نمی گیرد.

[1] Class

[2] Ordinal Scales

یک مقیاس ترتیبی، مقیاسی از مقوله های منظم طبیعی است که فاصله بین طبقه ها مشخص نیست. مقیاس موهس (Mohs) یک مقیاس ترتیبی برای سختی معدنی می باشد.

اختصارها و نمادها

اختصارها و علائم زیر در این راهنما وجود دارد. نمادهای استفاده شده در این سند در همه زمینه های علمی که در آن کاربرد دارد، هماهنگی ندارد. برای مثال در آزمایشگاه طبی، FN و FNR به ترتیب اختصار «تعداد نتایج منفی کاذب» و «نسبت منفی کاذب» می باشند.

فراوانی یونی طیف جرمی	A	مقدار پیش بینی کننده منفی	NPV
		احتمالات نسبی به نفع یک	O(.)
میانگین فراوانی یون طیف جرمی	$\bar{A}$	رویداد، برای مثال معنی (O(A	
		احتمال نسبی رویداد A را مشخص می کند	
نسبت فراوانی یون های طیف جرمی	AR	شانس به نفع رویداد	A
غلظت اندازه گیری شده (یا هر مقدار دیگر) از قلم تجزیه ای	C	تعداد نتایج مثبت	P
فاصله اطمینان	CI	احتمال وقوع یک رویداد، برای مثال P(A) احتمال رویداد A	P(.)
حداکثر غلظت قابل قبول	C_{max}	احتمال وقوع رویداد	A
حداقل غلظت قابل قبول	C_{min}	احتمال قبلی از مورد مثبت	P(+)
نسبت احتمال نسبی تشخیص	DOR	احتمال قبلی از مورد منفی	P(-)
بهره وری و کارایی	E	تعداد مورد مثبت	PC
تعداد نتایج منفی کاذب	fn	احتمال پسین مورد منفی (به پیوست A مراجعه کنید)	PN
نرخ منفی کاذب به موارد مثبت	FN	احتمال پسین مورد مثبت (به پیوست A مراجعه کنید)	PP
تعداد نتایج مثبت کاذب	fp	مقدار پیش بینی کننده مثبت	PPV
نرخ مثبت کاذب به موارد منفی	FP	واکنش زنجیره پلیمراز کمی	qPCR
کروماتوگرافی گازی طیف سنجی	GC-MS	رونویس معکوس واکنش زنجیره پلی مراز	RT-PCR
کروماتوگرافی طیف سنجی جرمی متوالی	GC-MS/MS	فراوانی نسبی	RA
راهنمای بیان عدم قطعیت در عدم اندازه گیری	GUM	گزارشگر نرمال شده	RN
حد بالای فاصله اطمینان ۹۵ درصد برای نرخ نتایج RR	$HL_{RR.95}$	نرخ نتایج	RR
کروماتوکرافی مایع-طیف سنجی جرمی	LC-MS	انحراف استاندارد فراوانی یون	S_A
حد اطمینان پایین ۹۵درصد فاصله نرخ اطمینان برای نتایج RR	$LL_{RR.95}$	ویژه بودن	SP
مقدار حداقل یا هدف برای $LL_{RR.95}$	$LL^{Tg}{}_{RR.95}$	حساسیت	SS
حد تشخیص قابل اندازه گیری	LOD	انحراف استاندارد زمان بازداری	S_{tRI}

حد تعیین مقدار (کمی سازی)	LOQ	تعداد نتایج منفی واقعی	tn
نسبت عدد احتمال	LR	نرخ منفی واقعی به موارد منفی	TN
نسبت عدد احتمال نتایج مثبت	LR(+)	تعداد نتایج مثبت واقعی	tp
نسبت عدد احتمال نتایج منفی	LR(-)	نرخ مثبت واقعی به موارد مثبت	TP
تعداد نتایج منفی	n	زمان بازداری	t_R
شاخص یودن	Y	میانیگن زمان بازداری	$\bar{t}_{Ri}$
میزان گزارشگر نرمال شده منهای پاسخ پایه	ΔRn	عدم قطعیت استاندارد C	u(c)
ضریب همبستگی اسپیرمن بین جفت هایی از فراوانی یون	ρ	کسر جرمی	W

۱) مقدمه

بسیاری از ارتباطات اجتماعی - اقتصادی یا منافع شخصی، مثل بهره وری صنعتی و وضعیت سلامتی به تجزیه شیمیایی بستگی دارد. تعدادی از این تجزیه ها منحصراً کیفی هستند یا شامل کمی سازی پس از شناسایی ماهیت شیمیایی می باشد.

منافعی که قرار است با این تجزیه‌ها محافظت شوند، فقط در صورتی حفظ می گردد که کیفیت تجزیه ای، برای استفاده مورد نظر مناسب باشد.

در بعضی انتشارات، اصطلاح «آزمایش[۳]» [۱] بررسی یک ویژگی و خصوصیت اسمی می باشد [۱۹] یا «آزمون» و «آزمون کردن» برای تجزیه کیفی استفاده می شود. در استاندارد بین المللی، برای اعتباربخشی به آزمایشگاه های طبی از عبارت «آزمایش» هم برای تجزیه کیفی هم کمی استفاده می شود [۶].

بنابراین از آنجایی که در مورد این عبارات، در میان جوامع بین المللی مرتبط و مختلف، هنوز اجماعی حاصل نشده است، این راهنما از عبارت «تجزیه کیفی» برای تعیین خصوصیات کیفی و اسمی در تجزیه های شیمیایی استفاده می کند.

به طور کلی نتیجه تجزیه کیفی، یک بیان ساده یا طبقه بندی شده از یک مورد یا موارد می باشد، یعنی یک دسته بندی[۴]. تصمیمات، همیشه بر اساس طبقه بندی[۵] گرفته می شود. برای مثال، برای صادر کردن یک بچ کود، آیا آب برای نوشیدن مناسب است؟ آیا یک شخص در تصرف یک ماده کنترل شده است یا خیر؟ آیا ماده ای که به تازگی سنتز شده است، ساختار درستی بر اساس الزامات دارد؟ در دسته بندی بندی نادرست مثل پذیرش یک محصول در حالی که آن ماده برای استفاده مناسب نمی باشد، ریسکهایی را برای همه افراد به دنبال دارد. برای کنترل این ریسکها افراد حرفه ای در بخش آنالیز به سختی کار می کنند تا مطمئن شوند که روش های آنها به طور قابل قبول منجر به ریسکهای کم دسته بندی نادرست می شود.

نتیجه اینکه، در توسعه هر روش آزمونی از این دسته در بعضی از نقاط، باید یک ارزشیابی از دسته بندی نادرست ریسک انجام شود. بنابراین برای اکثر این روش ها، منطقی است که انتظار تأسیس یک آزمایشگاه را داشت یا دسترسی به اطلاعات درمورد ریسکهای نتایج نادرست داشته باشیم. یک استثنای مهم، استفاده از روش های آزمون استاندارد شده می باشد که توسط گروه های خارج از آزمایشگاه، به طوری که مناسب با هدف مورد نظر باشد، ایجاد شده است [۲۰]- [۲۱-۲۲-۲۳]. آزمایشگاه ممکن است محدود باشد یا حتی به

[3] Examination
[4] Classification
[5] Categorization

داده های عملکردی چنین روش های آزمون دسترسی نداشته باشد. به هرحال، این روشها همیشه آزمایش را با جزئیات مربوطه، مشخص می کنند و از آزمایشگاه به طور کلی انتظار می رود نشان دهد که عوامل مربوطه در حین کنترل آن، واقعاً الزامات روش های آزمون را برآورده سازد، که به نوبه خود، ممکن است پارامترهای کنترل شده عدم قطعیت مربوطه را نشان دهد و آزمون کارآیی در ارتباط با هدف آزمون، کافی است.

ارزشیابی عدم قطعیت های مربوط به پارامترهای کمی یا نتایج تجزیه، موضوع تلاش های قابل توجهی از زمان انتشار «راهنمای بیان عدم قطعیت در اندازه گیری» (GUM)، که با عنوان ISO Guide 98 [۲۴] همچنین JCGM موجود است، بوده است [۲۵]. از طرفی دیگر، عدم قطعیت ها در تجزیه های کیفی، به مراتب کمتر مورد توجه قرار گرفته است. بعد از انتشار اولین ویرایش ISO/IEC 17025 [۲۶]، توجه به عدم قطعیت های تجزیه کیفی، افزایش یافته است. چالش ها در ایجاد عدم قطعیت مرتبط با تجزیه های کیفی مثل «تأیید/رد»، هویت یا هویت تطبیقی آنالیزهایی که بر این اساس به دست آمده اند، توجه بیشتری را به خود جلب کرده اند. به ویژه در زمینه هایی که تأثیر نتایج تجزیه کیفی، کاذب است، بسیار مرتبط هستند مثل پزشکی قانونی یا تجزیه دوپینگ.

برای بیان عدم قطعیت در نتایج کیفی، معیارهای متنوعی وجود دارد [۷]. به هر حال اجماع محدودی در مورد اینکه کدام معیارها استفاده شوند، وجود دارد. حوزه های اپیدمیولوژی و آزمایشگاه های بالینی مستثنی هستند،که مفاهیم «حساسیت بالینی» و «ویژگی بالینی» بطور مداوم به عنوان پارامترهای دقت بالینی، استفاده می شوند [۲۷].

تجزیه های کیفی و کمی، بطور قابل ملاحظه ای در چگونگی گزارش نتایج و عدم قطعیت مرتبط با آن، متفاوت هستند. نتایج کمی، به عنوان یک بازه گزارش می شوند که شامل مقدار واقعی اندازه گیری با سطح اطمینان تعریف شده می باشد، در حالی که ویژگی های اسمی به عنوان یک دسته بندی با معیارهایی که احتمال دسته بندی درست[۶] و نادرست[۷] را بیان می کند، گزارش می شود که شانس[۸] را می توان با یک احتمال[۹]، عدد احتمال[۱۰]، احتمالات نسبی[۱۱] یا دیگر معیارهای تخمین زده شده از تفسیر اطلاعات ورودی، توصیف کرد. کیفیت معیارهای گزارش شده، به تعداد و تنوع موارد مورد مطالعه بستگی دارد. تعیین این معیارها، امکان شناسایی موارد را فراهم می کند به طوری که روش ها باید بهبود یابند تا عدد احتمال تولید نتایج نادرست، کاهش یابد. این راهنما اصول کلی را برای ارزیابی عملکرد تجزیه کیفی، جهت گزارش دهی عدم قطعیت نتایج

[6] Correct
[7] Incorrect
[8] Chance
[9] Probability
[10] Likelihood
[11] Odds

تجزیه کیفی، شرح می دهد و نمونه های کاربردی نظریه توصیف شده را ارئه می دهد. این راهنما در مورد توانایی قلم آزمون برای نشان دادن گروهی از موارد یکسان یا بزرگتر، بحث نمی کند یعنی از تأثیر نمونه گیری در این ارزیابی ها بحث نمی کند. نتایج ترتیبی را می توان به خروجی دودویی (بله/خیر) کاهش داد و روش های استفاده شده در این راهنما، به وسیله تخصیص نتایج دسته بندی ترتیبی به عنوان «درست» یا «نادرست»، پردازش می شود. روش های دیگر برای پردازش مقیاس هاس ترتیبی، خارج از محدوده این راهنما هستند.

۲) انواع تجزیه کیفی

همان طور که در دامنه کاربرد این راهنما ذکر شد، تجزیه کیفی[۱۲] به عنوان «دسته بندی مطابق با معیارهای مشخص شده» تعریف شده است [۲۸]. چند مثال در جدول [۱] فهرست شده است. اگر چه همه این موارد بسیار متفاوت به نظر می رسند اما یک مشخصه مشترک دارند. پس از مشخص شدن معیارها، عملکرد روش دسته بندی از نظر میزان موفقیت یا شکست آن، نسبتاً ساده است. این نرخ های موفقیت و شکست، اساس اکثر معیارهای عملکرد تجزیه کیفی را تشکیل می دهد.

تجزیه های کیفی مطرح شده در متن اصلی، بر اساس انواع مختلف معیارهای طبقه بندی، یعنی کیفی یا کمی، به دو دسته تقسیم می شوند. جدول [۱] نمونه هایی از هریک را نشان می دهد. بخش ۳ استراتژی هایی را برای ارزیابی عملکرد، برای انواع مختلف معیارهای دسته بندی، توضیح می دهد. برای تجزیه کیفی، که در آن نرخ پاسخ درست یا نادرست به یک ویژگی کمی بستگی دارد، مانند وجود یک ماده ممنوعه که تشخیص آن به مقدار موجود بستگی دارد، همچنین حد تشخیص در نظر گرفته می شود (به بخش ۳.۴ مراجعه کنید).

گاهی اوقات، ارزیابی انطباق مقدار یک ویژگی کمی از یک قلم با مقدار یا بازه حدی، می تواند به عنوان تبدیل یک نتیجه اندازه گیری، به یک نتیجه کیفی (منطبق یا غیر منطبق) در نظر گرفته شود. استفاده از مقادیر اندازه گیری شده و عدم قطعیت های اندازه گیری آنها، برای ارزیابی انطباق، به طور مفصل در راهنمای دیگری از Eurachem/CITAC پوشش داده شده است [۲۹] و بر این اساس، در این راهنما، به تفضیل در نظر گرفته نشده است. با این حال، پیوست (ب) شرح می دهد چگونه برخی از معیارهای مورد استفاده برای ارزیابی عملکرد یا عدم قطعیت تجزیه کیفی می تواند برای ارزیابی انطباق کمی، تعیین شود.

جدول ۱ انواع تجزیه کیفی بر اساس انواع مختلف معیارهای دسته بندی

مثال تجزیه کیفی	معیار دسته بندی
۱) تشخیص آلدئیدهای آلیفاتیک در محلول با تغییر رنگ پس از افزودن معرف شیف[۱۳]	کیفی
۲) شناسایی شکل بلورین مواد با مشاهده	
۳) شناسایی نام تجاری و سن شراب با تجزیه و تجزیه حسی	
۴) شناسایی یک گونه بیولوژیکی با تعیین یا تشخیص یک توالی DNA خالص	
۵) شناسایی گروه خونی انسان با مشاهده آگلوتیناسیون	

[۱۲] آزمون کیفی در اصل، حوزه وسیع تری از تجزیه های کیفی است صرفاً به این دلیل که تجزیه شیمیایی، که اغلب به عنوان کار تحلیلی، به آن اشاره می شود، یکی از فعالیت های خاص در میان بسیاری از زمینه های آزمون است. به هر حال در این راهنما، برای شیمی دانان تجزیه و رشته های مرتبط با آنها، اصطلاحات مترادف استفاده می شود.

[13] Schiff

کمی	۱) شناسایی باقی مانده آفت کش در میوه با استفاده از جرم شکستهای[۱۴] اندازه گیری شده و فراوانی نسبی شکستها در GC-MS ۲) تعیین هم ارزی طیفی مادون قرمز بین یک ماده خام صنعتی جدید و قبلاً پذیرفته شده با استفاده از معیارهای طول موج و شدت ۳) شناسایی یک ماده ادرار آور[۱۵] در ادرار از یک ورزشکار با استفاده از زمان بازداری جرم شکستهای اندازه گیری شده در GC-MS ۴) تشخیص یک دارو در خون با استفاده از زمان بازداری و جرم شکستهای اندازه گیری شده در LC-MS ۵) شناسایی یک ویروس در نمونه بالینی بر اساس شدت فلورسانس در واکنش کمی زنجیره پلی مراز (qPCR)

[14] Fragments
[15] Diuretic

۳) ارزیابی عملکرد برای تجزیه کیفی

۱.۳ اصطلاحات کلی

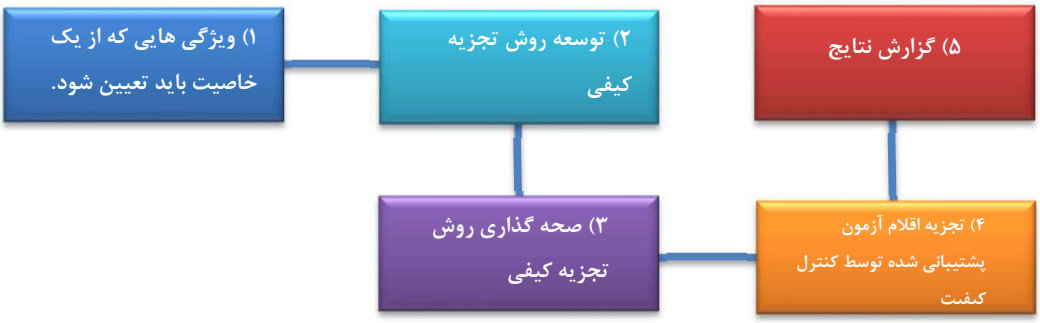
این بخش، راهنمایی در مورد ارزیابی و بیان عملکرد، برای روش هایی را فراهم می کند که به منظور دسته بندی ساده در دو دسته ارائه شده اند (دسته بندی دودویی). در اینجا، برای نشان دادن این که دسته ها، «عضو دسته بندی هستند یا نیستند»، به صورت «مثبت» یا «منفی» نشانه گذاری می شوند. این دسته بندی، بیشتر موقعیت های عملی، شامل «بالای یک حد»، «قابل قبول»، «غیر قابل قبول»، «شناسایی به عنوان» یا «حضور یک گونه خاص» را پوشش می دهد.

دسته ها جامع و انحصاری فرض می شوند تا امکان محاسبه نرخ های پاسخ کاذب بدون ابهام را فراهم کنند. این بدان معناست که هیچ قلم آزمونی را نمی توان به عنوان عضوی از دسته سوم، دسته بندی کرد که به طور کلی می تواند با مشخصات دقیق معیارهای دسته بندی، به دست بیاید. با این حال ممکن است یک نتیجه اعتماد کافی را در دسته بندی فراهم نکند.

تحت این شرایط، برای تحلیلگر، کاملاً منطقی است که یک نتیجه آزمایش را به صورت «غیر منطقی» به معنای قطعیت ناکافی گزارش کند. نتایج غیرمنطقی نیازمند مطالعه بیشتر، برای گزارش نتایج به صورت «قطعی» می باشد. این نتایج در آزمایشگاههای طبی، به صورت «منطقه خاکستری» یا «منطقه مبهم» شناخته می شوند.

در اصل برخی از مفاهیم توصیف شده، می توانند به دسته های بیشتری، مانند دسته بندی مقیاس ترتیبی با ارزیابی نرخ های دسته بندی درست یا نادرست، برای همه دسته ها بسط داده شوند. یک توسعه مفید این است که شناسایی ساختار یا هویت (به طور رسمی یک مسئله چند دسته ای) به عنوان «درست» یا «نادرست» پردازش شود و این رویکرد، در اینجا فرض می شود. با این حال پردازش مفصل مسئله چند دسته ای، که ممکن است شامل تخصیص همزمان یا تخصیص مربوط به چند دسته باشد، فراتر از محدوده این راهنماست.

تجزیه کیفی شامل مراحل مختلفی است که عبارتند از: ۱) شرح مسئله، ۲) توسعه روش، ۳) صحه گذاری، ۴) آزمون ها روی اقلام ناشناخته که از طریق کنترل کیفیت بررسی می شوند، ۵) گزارش نتایج (شکل ۱). مشخصات بدون ابهام خاصیت، که باید تعیین شوند و ارزیابی تناسب تجزیه، برای استفاده مورد نظر بحرانی است. گزارش یک نتیجه تجزیه کیفی، باید توسط روش های اجرایی معتبر و کنترل کیفیت کافی آزمون، پشتیبانی شود. نحوه گزارش نتایج بستگی به هدف تجزیه و گیرنده گزارش بستگی دارد. این راهنما به جزئیات نحوه توسعه روش یا نحوه طراحی کنترل کیفیت، اشاره نمی کند.



شکل ۱ فرآیند تجزیه کیفی از توصیف مسئله تا گزارش نتایج

۲.۳ کمی سازی عملکرد تجزیه کیفی

۱.۲.۳ تعریف مبنای ارزیابی عملکرد

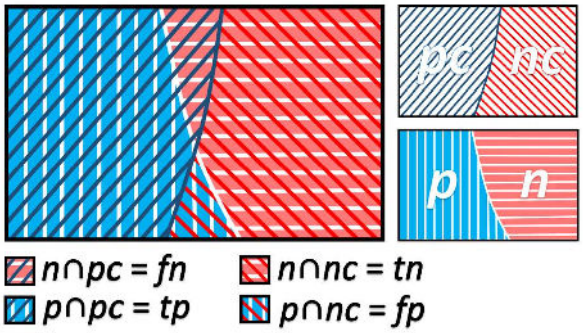
اساسی ترین راه برای کمی کردن عملکرد یک روش تجزیه کیفی، محاسبه نرخ نتایج کاذب است. بهتر است با نتایج «مثبت» یا «منفی»، به ترتیب «مثبت واقعی» و «مثبت کاذب» یا نرخ های «منفی واقعی» و «منفی کاذب» گزارش شود. با این حال این نرخ ها را می توان به تعداد کل نوع خاصی از یک مورد یا تعداد کل موارد یا نتایج ممکن، ارجاع داد.

به عنوان مثال، نرخ مثبت کاذب را می توان به صورت زیر تعریف کرد:

۱) کسری از موارد منفی، که به اشتباه، مثبت گزارش شده اند (fp/nc)، که در آن fp و nc به ترتیب تعداد نتایج مثبت کاذب و موارد منفی هستند. شکل (۲) به صورت گرافیکی همپوشانی انواع مختلف موارد و نتایج را نشان می دهد. fp از جدول (۲) این مورد را نشان می دهد. fp/nc به وسیله نسبت نواحی متقاطع (∩) نتایج مثبت (p)، با nc ($p \cap nc = fp$) و نواحی (nc) نشان داده می شود. FP از جدول (۲) این مورد را نشان می دهد.

۲) کسری از نتایج مثبت، که به اشتباه، مثبت گزارش شده اند (fp/p) که p تعداد نتایج مثبت می باشد. در شکل (۲) این نرخ به وسیله نسبت بین نواحی ($p \cap nc = fp$) و p نشان داده شده است.

۳) کسری از تعداد کل موارد یا نتایجی که به اشتباه مثبت گزاش شده است، ($fp/(pc + nc) = fp\ (p + n)$)، که pc و n، به ترتیب تعداد موارد مثبت و نتایج منفی را نشان می دهد. در شکل (۲) این نرخ، به وسیله نسبت بین مساحت مشخص شده ($p \cap nc = fp$) و مساحت کل شکل، نشان داده شده است.



شکل ۲ نمایش گرافیکی مثالی از همپوشانی تعداد موارد مثبت، *pc*، یا موارد منفی، *nc*، با تعداد نتایج مثبت، *p*، یا منفی، *n*، می باشد. علامت ∩، نشان دهنده تقاطع گروه ها می باشد. برای مثال، *n*∩*pc* در اینجا مجموعه ای از نتایج منفی از موارد مثبت را نشان می دهد. *n*∩*pc*، *p*∩*pc*، *n*∩*nc*، *p*∩*nc* به ترتیب *fn*، *tp*، *tn* و *fp* تعریف می شوند.

تفاوت بین این تعاریف، بسیار مهم است. در مورد (۱)، نرخ، با نسبت موارد منفی در جمعیت، تفاوتی ندارد، یعنی $nc/(nc+pc)$ ، زیرا $FP = fp/nc$ می باشد. با این حال، برای مورد (۲) و (۳)، نرخ مثبت کاذب، به $nc/(nc+pc)$ بستگی دارد، چون *FP* بیشتر در جمعیت های حاوی *nc* بیشتر مشاهده می شود. بنابراین، این تعاریف عملکرد تجزیه کیفی را به روش های مختلف مشخص می کنند و از این رو شامل تفاسیر متفاوتی از مقادیر آنهاست.

نرخ های مثبت واقعی، *TP*، (tp/pc) و منفی واقعی، *TN*، (tn/nc) به یک تعداد مشخص از مواردی، که در شیمی بالینی به ترتیب به عنوان «حساسیت» و «ویژه بودن» تجزیه های کیفی شناخته شده اند، ارجاع داده می شوند [۷] (جدول۲). تعیین حساسیت و ویژه بودن بالینی، مستلزم تعیین مناسب موارد مورد مطالعه با تشخیص بالینی قطعی می باشد. برای تجزیه کمی عبارت «حساسیت» [۱] یا «حساسیت تجزیه ای» [۳۰] مفهوم متفاوتی دارد.

نرخ مثبت واقعی، به موارد مثبتی (tp/p) که به عنوان تجزیه کیفی «دقت» یا «مقدار پیش بینی کننده مثبت»، *PPV*، شناخته شده اند، ارجاع داده می شوند [۳۰]. عبارت «مقدار پیش بینی کننده منفی»، *NPV*، برای نرخ منفی واقعی که به تعداد کل نتایج منفی ارجاع داده می شود، استفاده می شود (یعنی tn/n).

کارآیی تجزیه کیفی، به عنوان کسری از هر نوع نتایج درست، با توجه به همه نتایج تعریف می شود. (یعنی $(tp+tn)/(p+n)$). شاخص یودن عبارت است از یک روش جایگزین، برای کمی سازی موفق تجزیه های کیفی (جدول۲) [۳۱].

اگر چه معیارهای ارجاع شده، به تعداد موارد مثبت یا منفی به شیوع انواع موارد بستگی ندارد، این اعداد به تنهایی نمی توانند احتمال درستی یک نتیجه خاص را ارائه دهند. برای تخمین احتمال درست بودن یک نتیجه، نرخ نتیجه مرتبط و شیوع موارد نیز، باید در نظرگرفته شود. این مورد و سایر معیارهای اطمینان در نتایج کیفی، در بخش ۴ مورد بحث قرار گرفته است.

جدول ۲ مشخصات عملکردی دیگر برای بیان کیفیت نتایج تجزیه کیفی

مشخصات عملکردی	عبارت
نرخ مثبت واقعی، *TP*، (حساسیت، *SS*)	*tp/pc = tp/(tp + fn) = 1 - FN*
نرخ مثبت کاذب، *FP*	*fp/nc = fp/(tn + fp) = 1 - TN*
نرخ منفی واقعی، *TN*، (ویژه بودن، *SP*)	*tn/nc = tn/(tn + fp) = 1 - FP*
نرخ منفی کاذب، *FN*	*fn/pc = fn/(tp + fn) = 1 - TP*
دقت یا مقدار پیش بینی کننده مثبت، *PPV*	*tp/p = tp/(tp + fp)*
مقدار پیش بینی کننده منفی، *NPV*	*tn/n = tn/(tn + fn)*
کارآیی، *E*	*(tp+tn)/(p+n)*
شاخص یودن، *Y*	*SS(%) + SP(%) - 100*
نسبت عدد احتمال نتایج مثبت، *LR*(+)	*TP/FP*
نسبت عدد احتمال نتایج منفی، *LR*(-)	*TN/FN*
احتمال پیشین	به پیوست الف مراجعه کنید.

tp: تعداد نتایج مثبت واقعی، fp: تعداد نتایج مثبت کاذب، tn: تعداد نتایج منفی واقعی، fn: تعداد نتایج منفی کاذب، p: تعداد نتایج مثبت (tp+fp)، n: تعداد نتایج منفی (tn+fn)، pc: تعداد موادر مثبت و nc: تعداد موارد منفی

۲.۲.۳ تعریف مرجع ارزیابی عملکرد

معیارهای مورد استفاده برای کمی کردن عملکرد تجزیه کیفی، می تواند ویژگی های دیگری داشته باشد. موارد مثبت و منفی را می توان به روش های مختلفی تعیین کرد. برخی موارد یا نمونه هایی را که به عنوان مرجع استفاده می شوند، می توان به دلیل منشأ آنها یا از طریق فرمولاسیون، برای یک مشخصه «مثبت» دانست. از موارد دیگری که ممکن است وجود داشته باشد، به نحوی که در AOAC تعریف شده است، مواردی است که از نتایج یک فن تأییدی و دیگر فنون تجزیه ای هر دو مثبت باشند [۳۲]. برخی از نمونه ها با منشأ کافی از موارد مثبت، می توانند بیمارانی باشند که مبتلا به یک بیماری خاص هستند یا خاکی که مشخص شده آلوده می باشد. اقلام آزمون مثبت را می توان با افزودن گونه هایی که در یک بافت معادل اقلام تجزیه شده شناسایی شوند، فرمول بندی کرد، مانند یک آفت کش موجود در یک محصول غذایی که سطوح بومی آن آفت کش را تایید کرده یا تایید نکرده است. اگر عملکرد شناسایی با یک ویژگی کمی تغییر کند (به عنوان مثال غلظت ماده ای که شناسایی یا تعیین می شود) ضروری است فرمولاسیون، امکان تعیین آن سطح را فراهم کند. یک مورد منفی همچنین می تواند به عنوان موردی که از منشأ، فرمولاسیون یا به عنوان منفی

تعریف شده، منفی شناخته شده است، تفسیر شود. از آنجایی که «فن تأییدی و فن تحلیلی دیگر هر دو منفی هستند». تعاریف بین المللی AOAC، از موارد مثبت و منفی کاربرد جامع تری دارند، زیرا این تنها رویکرد قابل اجرا برای تجزیه اقلام پیچیده است که تولید مجدد آن از فرمول، چالش برانگیز است. با این حال، بر کیفیت خروجی فنون تجزیه ای استفاده شده، متکی است. در برخی زمینه ها، آماده سازی مصنوعی اقلام با آنالیت مورد مطالعه و مزاحمت های احتمالی برای آزمون عملکرد تجزیه، دشوار است. زیرا بافت های اقلام، ناشناخته و غیرقابل پیش بینی هستند.

موارد مثبت و منفی نیز می توانند به عنوان داده های مرجع، ارائه شوند، مانند طیف هایی که از یک ترکیب خاص، شناخته می شود. پس از تعیین معیارهای شناسایی، احتمال گزارش تطابق ترکیب به درستی یا نادرست، در مقایسه با این معیارها قابل تعیین است. به عنوان مثال، در طیف سنجی جرمی، شناسایی می تواند بر اساس ارزیابی حضور یا حضور و فراوانی یون های مشخصه، باشد. شانس تطابق طیف سنجی را می توان با آماره های دو جمله ای یا فرا هندسی، همانطور که در مثال های E1 و E2 مورد بحث قرار گرفت، پیش بینی کرد.

۳.۲.۳ گزارش عملکرد روش

۱.۳.۲.۳ جداول احتمالی

یک روش بسیار راحت، برای گزارش عملکرد یک روش کیفی تجزیه که در دامنه کار تجزیه ای تغییر قابل توجهی ندارد، از طریق جدول احتمالی است. جدول ۳ نمونه ای از چنین جداول را نشان می دهد. در این مثال، PP، FP، TN و FN به ترتیب (۲۲۸/۲۳۳) ۹۷٫۸٪، (۱/۳۰۱) ۰٫۳۳٪، (۳۰۰/۳۰۱) ۹۹٫۷٪ و (۵/۲۳۳) ۲٫۱٪ هستند.

به طور معمول، دامنه کار تجزیه ای می تواند سطوح مختلف گونه یا ویژگی مورد مطالعه و بافتهای مختلف مورد تجزیه را شامل شود. این می تواند به جداول احتمالی جداگانه برای بخشهای مختلف دامنه کار تجزیه ای، نیاز داشته باشد.

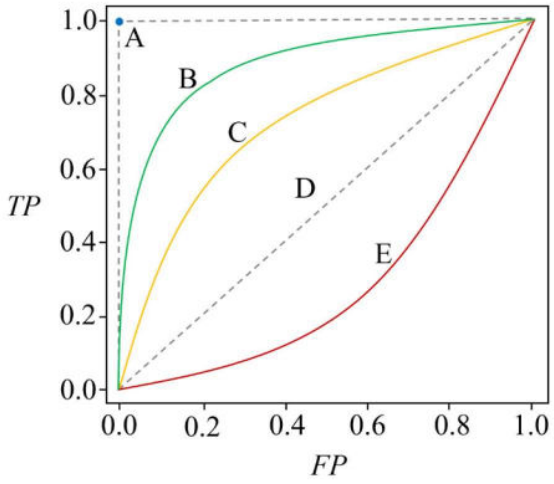
۲.۳.۲.۳ منحنی ویژگی عملکردی دریافت کننده

برای تجزیه بر اساس ارزیابی یک ویژگی کمی، انتخاب معیارهای دسته بندی که نرخ نتایج درست و کاذب را متعادل می کند، معمولاً TP و FP را می توان با استفاده از منحنی های ویژگی عملکردی دریافت کننده (ROC) انجام داد که جفت (TP،FP) را به عنوان یک معیار دسته بندی (به عنوان مثال آستانه تمایز) ترسیم می کند، متفاوت است. این منحنی ها همچنین می توانند برای مقایسه روش های مختلف تجزیه کیفی استفاده شوند [۳۱]. اگرچه شرح دقیق این منحنی ها خارج از محدوده این راهنماست.

جدول ۳ یک مثال خاص از یک جدول احتمالی که عملکرد یک روش تجزیه ای کیفی را توصیف می کند که توصیه می شود تقریباً در دامنه کار تجزیه ای ثابت باشد.

		مورد		
		مثبت (*pc*)	**منفی (*nc*)**	**کل نتایج**
نتیجه	**مثبت(*p*)**	$tp = ۲۲۸$	$fp = ۱$	$p = ۲۲۹$
	منفی(*n*)	$fn = ۵$	$tn = ۳۰۰$	$n = ۳۰۵$
	کل موارد	$pc = ۲۳۳$	$nc = ۳۰۱$	

شکل ۳ پنج نمونه شماتیک از منحنی های ROC را نشان می دهد.



شکل ۳ پنج نمونه از منحنی های ROC که در آن تغییرات *TP* و *FP* با تغییر معیارهای شناسایی کمی نشان داده شده است. منحنی A (نقطه آبی) یک آزمون کامل را ارائه می دهد که در آن معیارهای شناسایی بر نتایج تأثیر نمی گذارد و *TP* و *FP* به ترتیب ۱۰۰٪ و ۰٪ هستند . منحنی B و C نشان دهنده روش های مناسب هستند، که در آن $TP \geq FP$ می باشد. در بین این سه روش، روش B بر روش C ارجحیت دارد. منحنی D نشان دهنده قطر شانس است که در آن برای تمام آستانه های تصمیم گیری $TP = FP$ می باشد. این یک طبقه بندی مفید نخواهد بود. منحنی E در نگاه اول طبقه بندی کننده بسیار ضعیفی به نظر می رسد. به طور مداوم نرخ مثبت کاذب را بزرگتر از نرخ مثبت واقعی نشان می دهد. با این حال یک تغییر ساده از نتیجه گزارش شده منحنی ROC نزدیک به منحنی C ایجاد می کند. ممکن است طبقه بندی کننده مفید واقع شود.

هر منحنی نشان می دهد که چگونه نرخ مثبت واقعی و نرخ مثبت کاذب تغییر می کند وقتی که معیار شناسایی از شناسایی سخت گیرانه تا شناسایی کمتر سخت گیرانه موارد مثبت در ارتباط با *TP* کم یا زیاد به ترتیب تغییر می کند.

علاوه بر ارائه یک تصویر بصری از عملکرد سطح زیر منحنی (اغلب به اختصار AUC) می تواند به عنوان خلاصه‌ای از عملکرد دسته بندی کننده استفاده شود.

۳.۳ ارزیابی نرخ های مثبت کاذب و منفی کاذب

۱.۳.۳ دامنه روش و جزئیات صحه گذاری

صحه گذاری یک روش تجزیه کیفی شامل تعریف الزامات عملکرد و بررسی اینکه آیا آنها برآورده می شوند یا خیر، می باشد.

قبل از این ارزیابی عملکرد دامنه ی تجزیه باید به وضوح از نظر نوع دسته بندی تعریف شود (به عنوان مثال وجود پنتاکلروفنل بالای ۱ میلی گرم در کیلوگرم) و اقلام تجزیه شده (به عنوان مثال محصولات چرمی) توصیه می شود روش دسته بندی نیز مشخص شود برای مثال فن تجزیه ایی (مثل GC-MS/MS)، نحوه استفاده از این تکنیک (به عنوان مثال آماده سازی نمونه و شرایط دستگاهی) و معیارهای دسته بندی می باشد. معیارهای دسته بندی باید به وضوح توصیف شوند تا تضمین شود که داده های عملکرد جمع آوری شده برای تجزیه های بعدی اعمال شود.

در برخی از تجزیه های کیفی با توجه به ملاحظات کارایی، روش تجزیه ایی به دو مرحله تقسیم می شود.

یک روش غربالگری اولیه سریعتر و ارزانتر مقدماتی که هر زمان که لازم باشد با استفاده از یک روش وقت گیر و گرانتر تأیید می شود.

تأیید زمانی انجام می شود که اولین نتایج ارزیابی نتایج بر خلاف آنچه مورد انتظار است ایجاد می نماید یا بتواند بر منافع فردی یا جمعی تأثیر بگذارد. با این حال ارزیابی نرخ منفی و مثبت کاذب ضروری می باشد که شامل آزمون های غربالگری و تأییدی می باشد. به عنوان مثال، زمانی که فقط نتایج مثبت مشمول تأیید هستند. بررسی اینکه آیا نرخ منفی کاذب مرحله غربالگری به اندازه کافی پایین است ضروری می باشد.

باتوجه به جزئیات صحه گذاری روش، برای روش های قابل استفاده برای اقلام مختلف (مانند محصولات غذایی مختلف) توصیه می شود عملکرد برای مجموعه ای نماینده از انواع اقلام آزمون شود. انواع و تعداد اقلام آزمایش شده به تأثیر بافت تجزیه شده بر عملکرد بستگی دارد.

در برخی موارد درک اصول دسته بندی می تواند امکان پیش بینی گروه هایی از اقلام مرتبط با عملکرد تجزیه کیفی معادل را فراهم کند که از آن می توان یک قلم نماینده انتخاب و بررسی کرد. سپس عملکرد تجزیه قلم نماینده را می توان به گروهی از اقلام مرتبط با عملکرد تجزیه کیفی معادل تعمیم داد.

اگر عملکرد تکنیک دسته بندی اجازه دهد، برای مطالعه عملکرد تجزیه اقلام یا ویژگی های مقادیری که در آن نرخ نتایج کاذب به بالاترین مقدارها می رسد می توان تصمیم گرفت. آزمایشگاه باید ضمن در نظر گرفتن زمان و منابع موجود برای این ارزیابی، کامل بودن ارزیابی عملکرد را مدیریت کند.

دربرخی موارد ممکن است اجرای یک استراتژی صحه گذاری قابل قبول باشد که در آن هر بار یک قلم آزمون جدید برای آزمایشگاه، آزمون می شود ، کنترل های اضافی و خاص کیفیت تجزیه انجام می شود.

۲.۳.۳ استفاده از اطلاعات متون

برای روش های اجرایی تجزیه که معمولاً مورد استفاده قرار می گیرد، انتظار می رود اطلاعات عملکرد در حوزه عمومی باشد. قبل از شروع مطالعه عملکرد یک روش تجزیه ای به خوبی تثبیت شده، توصیه می شود مطالعه مناسبی از متن مربوطه در این زمینه انجام شود تا اطلاعات مستقلی در مورد تناسب آن برای استفاده مورد نظر، جمع آوری شود.

با این حال، توصیه می شود نرخهای پاسخ کاذب منتشر شده، با احتیاط استفاده شود. آنها را می توان با استفاده از تجهیزات، معرف ها و کارکنان خاص به دست آورد و به بافت های نمونه خاص و سطوح مشخصه اشاره کرد، بنابراین تحلیل گر باید بررسی کند که آیا وضعیت او هم ارز است یا خیر؟ به عنوان مثال اگر اقلام مورد مطالعه در متون، دارای سطوح مشخصه دور از آستانه های مورد استفاده برای تمایز بین دسته ها باشند و اگر بافت های آنها نسبتاً عاری از مزاحمت باشند، عملکرد شناسایی تعیین شده می تواند در مقایسه با مسائل تجزیه ای «واقعی» که توسط آزمایشگاه تجربه می شود، بسیار خوش بینانه باشد. بنابراین نرخ نتایج درست و کاذب، به شدت به داده های موجود بستگی دارد. در برخی موارد، می توان پیش بینی کرد که آیا عملکرد مشاهده شده در متون، بهتر یا بدتر از عملکرد مشاهده شده برای تجزیه کیفی در آزمایشگاه خواهد بود.

اگر نتیجه گیری شود که روش اجرایی تجزیه کیفی، برای بدترین سناریوها معتبر است یعنی اینکه می تواند نتایج درخور هدف مورد نظر خود را تولید کند، این روش اجرایی می تواند برای تجزیه اقلام ناشناخته، بدون محدودیت استفاده شود.

بخش ۴، به این موضوع می پردازد که چگونه می توان معیارهایی برای تصمیم گیری در مورد درخور بودن تجزیه برای استفاده مورد نظر تعیین کرد.

۳.۳.۳ ارزیابی انحصاری از آزمون

صرف نظر از نوع تجزیه کیفی ذکر شده در بخش ۲، FP و FN را می توان مستقیماً از تعداد نتایج کاذب از مجموعه ای از تجزیه تخمین زد. در تجزیه کیفی که منحصراً براساس ورودی های کیفی استوار است (جدول

۱)، این، تنها راه تخمین عدم قطعیت تجزیه کیفی است. با این حال، اگر پاسخ های کاذب بعید باشد، این رویکرد، به تعداد زیادی آزمایش نیاز دارد.

با توجه به اینکه توصیه می شود، تعداد پاسخ های کاذب، در حالت ایده آل کم باشد، این مسئله پیش می آید که چند نمونه باید آزمایش شود تا به طور منطقی از یافتن تعداد غیر صفر پاسخ های کاذب، مطمئن شویم.

از اطلاعات منتشر شده (به عنوان مثال فرارا و همکاران [۳۳] را ببینید.) بدیهی است که نرخ های مثبت یا منفی کاذب، می تواند تا ۰/۵٪ و در برخی موارد حتی کمتر باشد [۶،۸،۹].

برای محدوده ای از احتمالات نتایج کاذب، جدول ۴، تعداد نمونه هایی را نشان می دهد که لازم است تجزیه شوند تا حداقل در سطوح اطمینان نشان داده شده، از یافتن یک یا چند نتیجه کاذب مطمئن شویم.

جدول ۴، ازتوزیع دو جمله ای استفاده می کند و نشان می دهد که برای ۹۵٪ شانس تشخیص یک یا چند نتیجه کاذب، تعداد آزمایشهایی که لازم است انجام شود، سه برابر بیشتر از تعداد آزمایشهایی خواهد بود که به طور متوسط یک نتیجه کاذب ایجاد می کند.

به عنوان مثال، برای روشی با نرخ مثبت کاذب ۱٪، مشخص شده است که (به طور متوسط) برای هر ۱۰۰ تجزیه موارد منفی، یک نتیجه مثبت مشاهده می شود. با این حال، برای اطمینان ۹۵٪ که یک نتیجه مثبت کاذب مشاهده شود، باید ۲۹۹ (در حدود ۱۰۰×۳) آزمون روی موارد منفی انجام شود.

جدول ۴ حداقل تعداد تجزیه برای یافتن یک یا چند نتیجه کاذب (مثبت یا منفی)

سطح اطمینان		
نرخ نتایج کاذب	٪۹۵	٪۹۹
٪۰/۵	۵۹۸	۹۱۹
٪۱	۲۹۹	۴۵۹
٪۵	۵۹	۹۰

مقادیر جدول ۴، برای برآورد خوب نرخ نتایج کاذب یا مقایسه روش های مختلف، کافی نیست. حتی یک تخمین تقریبی، معمولاً به پنج تا ده برابر حداقل تعداد مشاهدات ارائه شده در جدول ۴، نیاز دارد. این جدول همچنین می تواند به عنوان حداقل تعداد تجزیه لازم، برای بررسی انطباق با نرخ های مختلف نتایج کاذب قابل قبول، بررسی شود، همانطور که در ادامه مورد بحث قرار می گیرد.

در تلاش برای تعیین نرخ نتایج کاذب، به طور مستقیم از آزمایش برای یک روش اجرایی جدید، تحلیلگر اغلب با یک دو راهی مواجه می شود. از طرف دیگر، برای یک روش اجرایی معین، نرخ پاسخ کاذب مورد نظر، ناشناخته است و بنابراین دسته بندیهای انجام شده، می تواند غیر قابل اعتماد باشد. از سوی دیگر، صرفاً تجزیه

تا زمانی که اولین پاسخ کاذب رخ ندهد، لزوماً تصویری واقعی از نرخ پاسخ کاذب، ارائه نمی دهد. برای بررسی این مسئله، پیشنهاد می شود که تحلیلگر، از قبل در مورد سطوح قابل تحمل برای دو نرخ پاسخ کاذب، تصمیم گیری کند. برای یک سطح اطمینان انتخاب شده، توزیع دوجمله ای ممکن است برای تخمین تعداد آزمایش های مورد نیاز برای یافتن یک یا چند پاسخ کاذب با اطمینان کافی، استفاده شود.

این رویکرد، تضمینی برای ایجاد یک رقم دقیق، برای نرخ پاسخ کاذب، نیست اما محدودیتی برای آن ایجاد می کند. برای مثال فرض کنید تحلیل گر تصمیم بگیرد که FP ۵٪ قابل قبول است و پس ازانجام ۵۹ آزمایش (جدول۴) با پوشش محدوده احتمالی بافتها ، هیچ مثبت کاذبی یافت نمی شود. درآن صورت می توان نتیجه گرفت که FP بیشتر از ۵٪ نیست. به عنوان معیار کنترل کیفیت روش اجرایی صحه گزاری شده، توصیه می شود که بین شاهدها[16] و مواد مرجع حاوی مشخصه هدف (به عنوان مثال آنالیت)، در سطوح مشخصه مربوطه قرار داده شود.

همیشه به یاد داشته باشید که نرخ نتایج کاذب، بسیار به تغییرات یا ویژگی های جمعیت نمونه گیری شده و به استراتژی نمونه گیری این جمعیت، بستگی دارد.

جدول ۴، نشان می دهد که برای نرخ پاسخ کاذب کم، ممکن است تجزیه تعداد کافی نمونه برای تشخیص پاسخ کاذب، غیر عملی باشد.

براین اساس، اگر آزمایشی ارزان باشد و یا به طور گسترده مورد استفاده قرار گیرد، به عنوان مثال آزمایش غربالگری دارو، می توان ابتدا ثابت کرد که نرخ پاسخ کاذب از حد بالا، مثلاً ۵ درصد، با آزمایش تجاوز نمی کند. و سپس این عدد را در راستای تجربه با نمونه های بیشتر، اصلاح می کنیم. در مواردی که تعداد نمونه احتمالاً نسبتاً کم است و یا اجرای آزمون گران است، توصیه می شود همه آزمون ها همراه با یک آزمون تأییدی اجرا شوند و توصیه می شود هر از چند گاهی، نرخ پاسخ کاذب، مجدداً محاسبه شود.

پردازش ریاضی اطلاعات موجود، می تواند برای غلبه بر برخی محدودیت های تعیین تجربی نرخ پاسخ کاذب، استفاده شود .(بخش ۳۰۳۰۴ و ۳۰۳۰۵ را ببینید).

۴.۳.۳ ارزیابی از یک پایگاه اطلاعاتی

یک جایگزین برای تعیین نرخ نتایج کاذب از آزمایش، مطالعات عدم تطابق تصادفی شانسی در پایگاه های داده مرجع است. مانند پایگاه داده طیف های جرمی یا طیف های مادون قرمز. در برخی موارد، این امکان، هم ارز هزاران آزمایش می باشد. با این حال هرچند آموزنده و قدرتمند، محدودیت فعلی این است که چنین پایگاه

[16] Blank

های اطلاعاتی اغلب کاملاً نماینده جمعیت آزمایش کننده نیستند. به عنوان مثال در حالی که شیوع مواد مختلف در استفاده عمومی، بسیار متفاوت است، یک پایگاه اطلاعاتی مرجع، تنها حاوی یکی از هر یک خواهد بود. این ممکن است به برآوردهای احتمالی دارای انحراف مهم، منجر شود. مجدداً مقادیر به دست آمده، بعید است بهتر از برآوردهای بزرگتر باشند. مثال های E1 و E2 استفاده از این روش را برای ارزیابی عملکرد تجزیه کیفی نشان می دهد.

۵.۳.۳ ارزیابی از مدل سازی داده های کمی

ارزیابی عملکرد تجزیه های کیفی بسیار گزینش پذیر و وقت گیر و یا گران قیمت، منحصراً از آزمایش های انجام شده در یک آزمایشگاه، امکان پذیر نیست.

در تجزیه کیفی بر اساس معیار دسته بندی کمی برای نتایج کمی (مانند روش دستگاهی تجزیه ای)، مدلهای پراکندگی نتایج را می توان برای تخمین نرخ نتایج درست و کاذب استفاده کرد. پیوست ب جزئیات بیشتری را ارائه می دهد. به عنوان مثال اگر علامت دستگاهی مربوطه، مانند زمان بازداری آنالیت در روش کروماتوگرافی، به طور نرمال توزیع شده باشد، شانس اینکه یک جزء مزاحم، دارای زمان بازداری در بازه زمانی بازداری قابل قبول برای آنالیت باشد، قابل پیش بینی است (به مرجع سریع ۱، صفحه ۲۵ مراجعه کنید).

با این حال، مدل سازی بر اعتبار فرض مدل و مقادیر متغیر ورودی، تکیه دارد. به عنوان مثال، از آنجایی که زمان های بازداری نسبی، به طور نرمال توزیع نمی شود، فرض نرمال بودن می تواند نرخ نتایج کاذب را، کم تخمین زند. شبیه سازی علامت دستگاهی با روش مونت کارلو، یک راه مناسب برای تخمین FP و FN از پارامتر هایی که توزیع غیر نرمال دارند، می باشد [۹و۸]. مثال E5، مدل سازی پراکندگی علامت دستگاهی را برای تخمین نرخ نتیجه کاذب شناسایی های بسیار گزینش پذیر GC-MS/MS را نشان می دهد.

مرجع سریع ۱- مثال مدل سازی علامت

اگر برای شناسایی دلتامترین در روغن زیتون، توسط GC-MS، انحراف استاندارد تخمینی تکرارپذیری زمان بازداری، S_{tRi}، ۰/۰۲۲ min با $v = 32$ درجه آزادی، رواداری زمان بازداری برای شناسایی این ترکیب در یک نمونه، می تواند به صورت زیر باشد:

$$(t_R \pm t \times s_{tRi}) = (t_R \pm 2.04 \times 0.022) = (t_R \pm 0.045)\ min$$

که t_R زمان بازداری مشاهده شده از تزریق روزانه یک محلول استاندارد است و مقدار بحرانی ۹۵٪ دو طرفه برای توزیع t با درجه آزای ۳۲ می باشد. بنابراین، برای t_R برابر با ۳۶/۰۵۵ min، فاصله پذیرش برای یک پیک نمونه، (۳۶/۰۵۵±۰/۰۰۴۵) دقیقه خواهد بود. با فرض اینکه یک مزاحمت، دارای زمان بازداری ۰/۰۵ دقیقه کمتر از دلتامترین باشد و دقت هر دو زمان بازداری، معادل باشد، احتمال اینکه مزاحم، در بازه پذیرش، زمان بازداری داشته باشد، ۱/۵٪ خواهد بود.

این مقدار با توزیع t تجمعی برای مقدار t از (۰/۰۲۲÷۰/۰۵-) و v تخمین زده می شود.
فرمول اکسل: T.DIS(-0.05/0.022;32;TRUE)

۶.۳.۳ ارزیابی عملکرد آزمون کیفی وابسته به یک متغییر پیوسته :

بسیاری از آزمونهای تأییدی[17] یا تشخیص، وابستگی شدیدی به احتمال تشخیص یا نرخ پاسخ کاذب برای برخی از متغیرهای پیوسته، نشان می دهند. برای مثال، نرخهای تشخیص اغلب به غلظت یا تعداد ذرات ماده مورد جستجو، بستگی دارد. ممکن است مدل سازی وابستگی نرخ پاسخ کاذب به متغیرهای پیوسته (یا دیگر متغیرها) ارزشمند باشد.

رگرسیون لجستیک[18] و رگرسیون پروبیت[19] [۳۵ و۳۴] معمولاً برای چنین مشکلاتی به کار برده می شوند و (با مثالهایی) برای ارزیابی عملکرد روش های تجزیه کیفی پیشنهاد شده اند [۳۶]. رگرسیون لجستیک در تشخیص DNA با تعداد نسخه کم، نشان داده شده است [۳۷]. این روش، به خوبی در کتاب های درسی، مستند شده است و اساساً در تمام بسته های نرم افزاری آماری موجود است. بنابراین، دراینجا به تفصیل ارائه نشده است. رگرسیون لجستیک ساده، احتمال یک پاسخ دودویی را به عنوان تابعی از برخی متغیرهای پیوسته، مدل می کند. مدل به صورت زیر است:

$$p = \frac{\exp(b_0+b_1x)}{1+\exp(b_0+b_1x)} \qquad (۱-۱)$$

یا معادل آن:

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1x \qquad (۱-۲)$$

که در آن p احتمال مورد نظر است (برای مثال احتمال یک نتیجه مثبت). x متغیر پیوسته (معمولاً غلظت آنالیت) و b_0 و b_1 ضرایب رگرسیون هستند. اکثر بسته های آماری، یک روش مناسب از داده های خام (زوجهای غلظت/ نتیجه کیفی) یا از نسبت های محاسبه شده از تعداد نتایج را ارائه می دهند. توجه داشته باشید که اولی فقط به توالی بله/خیر (یا۱/۰) نیاز دارد و به تناسب نیاز ندارد. این، امکان اعمال روش را برای طیف وسیعی از نمونه های آزمایشی با غلظتهای متفاوت (معلوم یا با اندازه گیری مستقل)، که هر کدام فقط یک بار تحت آزمایش کیفی قرار می گیرند، فراهم می کند.

هنگامی که یک رابطه بر قرار می شود، تخمین حدود تشخیص (به زیر مراجعه کنید) از رابطه برازش شده بین غلظت و احتمال تشخیص، به سادگی با انتخاب یک حد مناسب، برای احتمال تشخیص، که مطابق با

[17] Confirmatory
[18] Logistic
[19] Probit

تعریف توانایی تشخیص در حال استفاده می باشد، امکان پذیر می شود. مثال E4 یک مثال عملی از رگرسیون لجستیک، ارائه می دهد.

۷.۳.۳ قضاوت کارشناسی

وقتی که هیچ داده ای برای عملکرد روش تجزیه ای، از مرجع ثالث در دسترس نمی باشد و ارزیابی عملکرد، بر اساس تجربه (بخش ۳.۳.۳) یا مدل سازی (۳.۳.۴ و ۳.۳.۵) امکان پذیر نمی باشد، تحلیل گر می تواند از تجربه عملیاتی خود در فن دسته بندی، برای اقلام بررسی شده یا مشابه، استفاده نماید، تا تصمیم بگیرد آیا روش برای کاربرد مورد نظر در خور هدف می باشد یا خیر.

جایی که امکان پذیر است، توصیه می شود، تصمیم در مورد درخور هدف بودن یک روش، برای کاربرد مورد نظر، توسط شواهد عینی تأیید شود.

فرآیند فرمول بندی قضاوت کارشناسی، موضوع چندین مطالعه می باشد. قضاوتها، تحت تأثیر عوامل بسیاری که منجر به تخمین عدم قطعیت نتیجه می شود، می باشد [۳۸].

۴.۳ حد تشخیص و گزینش پذیری

۱.۴.۳ حد تشخیص

حد تشخیص (LOD)، معمولاً کمترین غلظت یک ماده را که منجر به تشخیص قابل اعتماد می شود، شرح می دهد. برای آزمون هایی که دسته بندی، مستلزم ارزیابی یک مشخصه کمی است و مقدار این مشخصه بر نتایج کیفی تأثیر می گذارد، توصیه می شود «حد تشخیص» (LOD) و یا «حد کمی سازی» (LOQ) در نظر گرفته شده در تجزیه کیفی و یا کمی، در رابطه با عملکرد تجزیه کیفی بررسی شود [۳۰]. توصیه می شود نتیجه تجزیه کیفی، برای استفاده مورد نظر در آن سطوح، مناسب باشد.

توجه: مقررات کمیسیون (EU) شماره ۵۸۹/۲۰۱۴ [۳۹] و شماره ۱۵۲/۲۰۰۹ [۴۰] یک LOQ را به عنوان «کمترین محتوای آنالیت که می تواند با قطعیت آماری معقول اندازه گیری شود، بارعایت معیارهای شناسایی، همانطور که در استانداردهای بین المللی شناخته شده است »، تعریف می کند [۴۱].

برای تجزیه صرفاً کیفی، LOD را می توان با اعمال روش بر روی اقلامی که حاوی سطوح به طور تصاعدی کوچکتر از مشخصه، هستند، پیدا کرد تا زمانی که عدد احتمال تولید نتایج کاذب، به یک معیار از پیش تعیین شده برسد. در این نوع ارزیابی، می توان از رگرسیون لجستیک و پروبیت[۲۰] نیز استفاده کرد (بخش ۳.۳.۴).

۲.۴.۳ گزینش پذیری

[20] Logistic and probit regression

گزینش پذیری[21]، به معنایی که این عبارت معمولاً در شیمی تجزیه به کار می رود، به: «میزان استفاده از یک روش خاص، برای تعیین آنالیت ها، درشرایط معین، در حضور سایر اجزای با رفتار مشابه» اشاره دارد [۴۲]. واژگان بین المللی اندازه شناسی (VIM) این عبارت را به عنوان یک ویژگی سیستم اندازه گیری تعریف می کند [۱].

توجه: عبارت ویژه بودن[22] در زمینه تجزیه کمی، برای تجزیه کاملاً گزینش پذیر استفاده می شود [۴۳و۴۲] که در شیمی، فقط در موارد بسیار نادر می توان ادعا کرد. با این حال، استفاده جایگزین واضحی از اصطلاح «ویژه بودن» در زمینه تجزیه کیفی وجود دارد. (جدول ۲ را ببینید). در این راهنما، اصطلاح «گزینش پذیری» به معنای کلی استفاده می شود و اصطلاح «ویژه بودن» برای اهداف ذکر شده در جدول (۲) استفاده می شود.

گزینش پذیری را می توان با تجزیه یک یا چند قلم آزمون، که دارای ویژگی های شناخته شده یا مزاحم هستند، ارزیابی کرد، یعنی ویژگی هایی که هدف تجزیه نیستند؛ اما ممکن است به احتمال زیاد، پاسخ آزمون را ایجاد کنند.

گاهی اوقات، می توان گونه ها یا عوامل مزاحم را شناسایی کرد، که به ویژه احتمال دارد، نتایج مثبت کاذب، ایجاد کند. به عنوان مثال، انتظار می رود که آزمونهای آمونیاک، به طور منطقی به آمین های اولیه پاسخ دهند و آزمونهای برای سویه های باکتریایی خاص، ممکن است انتظار رود که به هر باکتری از گونه های عمومی یکسان، پاسخ دهد.

اگر تجزیه کیفی، در بدترین شرایط، عملکرد نسبتاً خوبی داشته باشد، می توان نتیجه گرفت که این روش برای همه انواع اقلام، معتبر است.

اگر چه نرخ پاسخ کاذب را می توان برای هر ماده مختلف یا هر گونه مزاحم موجود، اندازگیری کرد، بعید است مطالعات گزینش پذیری، یک مقدار قطعی برای گزینش پذیری ایجاد کنند. این به این دلیل است که پاسخ به گونه های بالقوه، واکنش متقابل موجود در مطالعه و سطح این گونه ها بستگی دارد. بنابراین، مطالعات گزینش پذیری، به بهترین وجه، به عنوان فراهم کننده یک نشانه گسترده از کفایت روش تجزیه کیفی در مواجهه با چالش های مختلف، درنظر گرفته می شوند.

[21] Selectivity
[22] Specificity

۴) عبارت های اطمینان در تجزیه کیفی

۱.۴ ملاحظات کلی

در حالی که بیانیه های عدم قطعیت اندازه گیری در تجزیه کمی، معمولاً منجر به محدوده ای از مقادیر می شوند (مانند فاصله عدم قطعیت گسترده یا خلوص حداقل)، بیانیه های دسته بندی معمولاً نمی توانند با یک محدوده مرتبط شوند. به طور کلی نمی توان گزارش داد که یک ماده، ۹۰٪ تأیید است یا یک آنالیت، به احتمال ۹۹٪ وجود دارد یا یک گونه شیمیایی، در یک توالی به هم پیوسته است. در عوض، شکل نوعی اطلاعات عدم قطعیت، ماهیت احتمالی دارد. به این معنی که نشانه ای از احتمال صحیح بودن یک دسته بندی داده شده یا احتمالات معمول دسته بندی نادرست اقلامی که کلاس صحیح آنها مشخص است، را نشان می دهد.

ارقام عملکردی را که می توان از مطالعات صحه گذاری به دست آورد، می توان با نتیجه آزمون کیفی، گزارش کرد. با این حال، به طور کلی، آنها به ندرت اطلاعات مستقیمی درمورد (مثلاً) احتمال درست بودن یک نتیجه کیفی می دهند. دراین بخش، دو معیاری که برای این منظور پیشنهاد شده اند، با هدف کمک به درک و بهبود وضعیت حرفه در بیان عدم قطعیت نتایج تجزیه کیفی، تشریح می شوند.

معیار های ارائه شده در اینجا از انواع قاعده بیز استفاده می کنند [۴] (به پیوست A مراجعه کنید). این را می توان، استفاده کرد برای:

الف) نشان دادن قدرت شواهد ارائه شده توسط یک یا چند نتیجه کیفی، به نفع یک دسته بندی ممکن بر دیگری.

ب) در ارتباط با اطلاعات مناسب در مورد احتمال مواجهه با مقادیر مختلف (واقعی) ویژگی های کیفی در یک جمعیت، نشانه ای ازاحتمال واقعی بودن یک دسته بندی خاص با توجه به یک نتیجه تجزیه کیفی خاص.

۲.۴ نسبت عدد احتمال

مشهورترین و پرکاربرد ترین گزارش عملکرد تجزیه کیفی، نرخ نتایج کاذب است، به ویژه FP و FN یا نرخهای مکمل شان، TN و TP، به ترتیب (به عنوان مثال، $TN = 1\text{-}FP$). دو مورد ازاین نرخ ها را می توان در یک مشخصه عملکرد، ترکیب کرد: نسبت عدد احتمال، LR.

اگر یک نتیجه مثبت، گزارش شود، LR(+) به وسیله معادله (۲) تخمین زده می شود:

$$LR(+) = TP/FP \qquad (۲)$$

$LR(+)$ نسبتی از دو احتمال است. احتمال گزارش یک نتیجه مثبت، در صورت مثبت بودن مورد، تقسیم بر احتمال گزارش نتیجه مثبت، در صورت منفی بودن مورد. به طور کلی، نسبت عدد احتمال، معیاری از تغییر در احتمال مثبت بودن واقعی نمونه را پس از مشاهده نتیجه آزمایش مثبت، نشان می دهد. از نظر ریاضی، نسبت عدد احتمال، تغییر در احتمال است که به صورت احتمال نسبی بیان می شود (به پیوست A مراجعه کنید). نسبت عدد احتمال بالا از یک آزمایش، نشان می دهد که احتمال مثبت بودن قلم آزمون، بیش از آنچه که قبل از انجام آزمایش می توان گفت، وجود دارد. گاهی اوقات، این، به عنوان وزن شاهد[۲۳] تعبیر می شود که با نتیجه مثبت آزمایش ارائه می شود، با توجه به اینکه قلم آرمون، واقعاً مثبت باشد.

مرجع سریع ۲ - تفسیر نسبت عدد احتمال

اگر نتیجه مثبت، گزارش شود، احتمال اینکه مورد واقعاً مثبت باشد، PP، به وسیله معادله (س۱.۲) محاسبه می شود. (در زیر مشاهده کنید). این معادله، تئوری معروف بیز (ضمیمه A) است که نرخ های مثبت واقعی و کاذب را جایگزین احتمالات شرطی می کند.

$$PP = \frac{P(+)TP}{P(+)TP+P(-)FP} \quad \text{(س۱.۲)}$$

که در آن P(+) احتمال مثبت بودن مورد، قبل از آزمایش است. این را می توان به شکل «احتمالات نسبی» نیز بیان کرد (به پیوست A مراجعه کنید).

$$\frac{PP}{1-PP} = \frac{P(+)TP}{P(-)FP} \quad \text{(س۲.۲)}$$

در س۲.۲، نسبت P(+)/P(-) نشان دهنده احتمالات نسبی محاسبه یک مورد مثبت، قبل از اعمال آزمون کیفی است یعنی احتمالات نسبی پیشین. نسبت TP/FP، نسبت عدد احتمال محاسبه شده LR(+) می باشد.

بنابراین نسبت عدد احتمال، توضیح می دهد که چگونه احتمال (که با احتمال نسبی ارائه می شود) پس از یک نتیجه آزمایش مثبت، تغییر می کند. این، می تواند به عنوان معیاری از اطلاعات اضافی ارائه شده توسط آزمون در نظر گرفته شود.

در حالت خاصی که P(+) = P(-) = 0.5، به طوری که احتمال نسبی پیشین ۱/۰ باشد؛ پس LR(+) نسبت احتمالات پسین مثبت یا منفی بودن مورد را نشان می دهد. به عنوان مثال، با شیوع یکسان (یا شیوع فرضی) موارد مثبت و منفی، یک نتیجه مثبت با LR(+) ۱۰۰۰ نشان می دهد که احتمال پسین موردی که واقعاً مثبت است، ۱۰۰۰ برابر بیشتر از احتمال منفی بودن مورد است.

درمورد خاص که هم موارد مثبت و هم موارد منفی، به یک اندازه محتمل است، P(+) = P(-) = 0.5، که درآن P(+) و P(-) به ترتیب شیوع موارد مثبت و منفی می باشند. می توان درک کرد که LR(+) نشان می دهد نتیجه مثبت گزارش شده، چقدر بیشتر احتمال دارد واقعی باشد تا کاذب (به مرجع سریع ۲ مراجعه کنید).

[23] Weight of evidence

به عنوان مثال اگر موارد مثبت و منفی، قبل از آزمایش به یک اندازه محتمل در نظر گرفته شوند، یک نتیجه مثبت همراه با LR(+) ۷۳۰۰، به این معنی است که نتیجه مثبت، ۷۳۰۰ برابر بیشتر از اینکه کاذب باشد، واقعی است.

اگر نتیجه منفی گزارش شود، LR(-) عبارت است از:

$$LR(-) = \frac{TN}{FN} \qquad (۳)$$

برای مورد خاص که موارد مثبت یا منفی به یک اندازه محتمل باشد قبل از انجام آزمایش، LR(-) نشان می دهد که یک نتیجه منفی چقدر بیشتر احتمال دارد که واقعی باشد، تا اینکه کاذب باشد.

برخی از نویسندگان هر دو نسبت عدد احتمال را در پارامتر «نسبت احتمال نسبی تشخیص[۲۴]» ترکیب می کنند [۳۰]:

$$DOR = LR(+)/LR(-))$$

یکی از مفیدترین ویژگی های نسبت عدد احتمال (LR(+) یا LR(-))، این است که اگر دسته بندی به دو بخش مستقل شواهد بستگی داشته باشد (یعنی نتیجه تنها زمانی گزارش می شود که دو تجزیه مستقل از روش های مستقل آن را تایید کنند)، $LR_{(1\&2)}$ از نتیجه هر دو تجزیه، با ضرب LR که عدم قطعیت هر بخش شواهد را کمی می کند، تخمین زده می شود ($LR_{(1)}$, $LR_{(2)}$):

$$LR_{(1\&2)} = LR_{(1)} \times LR_{(2)} \qquad (۴)$$

به عنوان مثال اگر وجود یک آلاینده در یک محصول غذایی، که به وسیله GC-MS تعیین شده، بر اساس زمان بازداری با یک LR(+) برابر با ۹۹/۹ و داده های طیف جرمی با یک LR(+) برابر با ۴۹۰، LR(+) از شناسایی های مبتنی بر هر دو ابزار، 4.9×10^4 می شود (یعنی ۴۹۰×۹۹/۹). معادله (۴) از این واقعیت نشأت می گیرد که احتمال همگرایی دو نتیجه مستقل، با ضرب احتمالات منفرد مربوطه، تخمین زده می شود. اگر m بخش شواهد مستقل (مدرک مستقل) در نظر گرفته شود (m تا i =1)، برای گزارش یک نتیجه مثبت یا منفی، به عنوان مثال، یک نتیجه تنها درصورتی گزارش می شود که با m بخش شواهد نشان داده شود، LR از بخشهای شواهد مرکب با معادله (۵) تخمین زده می شود.

$$LR = \prod_{i=1}^{m} LR_{(i)} \qquad (۵)$$

[24] Diagnostic odds ratio (DOR)

که $\prod$ ضرب توالی متغیرها را نشان می دهد و $LR_{(i)}$ نسبت عدد احتمال از تجزیه کیفی *i*ام است ($LR_{(i)}(+)$ or $LR_{(i)}(-)$).

وقتی بخشهای شواهد، مستقل نیستند، معادله (۵)، احتمال مشترک را خیلی کم یا خیلی زیاد، برآورد می کند. مرجع سریع ۳، نشان می دهد که چگونه احتمالات غیر مستقل را می توان ترکیب کرد.

تفسیر نسبت های عدد احتمال، به ویژه برای افراد غیر متخصص، می تواند چالش برانگیز باشد. برای کاربردهای پزشکی قانونی، مقیاس جدول ۵ [۴۴]، برای ارائه یک نشانه شفاهی از قدرت شواهد، توصیه شده است. طبق این جدول، شواهد جمع آوری شده، تنها در صورتی بسیار قوی در نظر گرفته می شوند که LR بزرگتر از 10^6 باشد. در اصل، این نوع رویکرد را می توان برای شرایط دیگر تطبیق داد، اگر یک نشانه کلی از قدرت شواهد، مورد نیاز باشد. به عنوان مثال، برای شناسایی نوع پلیمری میکروپلاستیک های جمع آوری شده از رسوبات، در پایش محیطی، معیارهای ارائه شده درجدول ۵، بسیار سختگیرانه هستند. برای این تجزیه ها، نتایج مرتبط با LR(+) بیشتر از ۱۹، کافی می باشد (یعنی با یک $TP \geq ٪۹۵$ و $FP \leq ٪۵$)، از آنجایی که آلودگی پس از شناسایی بسیاری از ذرات از چندین نمونه، مشخص می شود [۴۵].

مرجع سریع ۳- احتمال بخشهای شواهد غیر مستقل

احتمال دو رویداد مستقل A و B، $P(A \cap B)$ ، با معادله س۱.۳ تخمین زده می شود.

$$P(A \cap B) = P(A)P(B) \quad \text{(س۱.۳)}$$

که در آن P(A) و P(B) احتمالات رخدادهای B و A هستند. به عنوان مثال، ایجاد یک نتیجه مثبت از تجزیه یک مورد مثبت (یعنی یک TP).

با این حال اگر P(A) و P(B)، مرتبط باشند، احتمال همزمانی هر دو رویداد، با معادله س۲.۳ تعیین می شود. س۲.۳، که شامل احتمال شرطی رویداد B با توجه به وقوع رویداد A می باشد.

$$P(A \cap B) = P(A)P(B|A) \quad \text{(س۲.۳الف)}$$

یا به صورت معادل:

$$P(A \cap B) = P(B)P(A|B) \quad \text{(س۲.۳ب)}$$

برای همبستگی مستقیم یا معکوس، مرتبط با به ترتیب $r_{AB} > 0$ یا $r_{AB} < 0$، $P(A \cap B)$ به ترتیب بزرگتر یا کوچکتر از مواردی خواهد بود که A و B، مستقل هستند.

اگر چه تعیین یک ویژگی دودویی، تنها می تواند یکی از دو نتیجه را ایجاد کند، اگر محتمل ترین نتیجه (به عنوان مثال، بله یا خیر) با LR پایین همراه باشد، می توان تصمیم گرفت که به جای گزارش معادل شفاهی جدول ۵، نتیجه غیر قطعی گزارش شود. به عنوان مثال، اگر LR مربوطه بزرگتر از (مثلاً) 10^5 باشد و LR پایین تر، غیرقطعی گزارش شود، می توان تصمیم گرفت که یک نتیجه، مثبت یا منفی گزارش شود. این

منطقه خاکستری برای مقدار LR را می توان بسته به هدف تجزیه، زیر $۱۰^{۵}$، ۱۹ یا هر مقدار دیگر تنظیم کرد. گزارش یک نتیجه به عنوان غیرقطعی، در صورتی مفید است که نتایج مثبت کاذب و منفی کاذب، تأثیر مرتبط داشته باشند. هنگام آزمایش مواد دوپینگ در ادرار ورزشکار، مثبتهای کاذب، بسیار جدی تر از نتایج منفی کاذب است که پیشنهاد می کند اگر هیچ شواهدی از دوپینگ مشاهده نشد، نتیجه می تواند منفی گزارش شود (یعنی هیچ شواهدی از دوپینگ وجود ندارد) [۴۶ و ۹]. با این حال، هم نتایج مثبت کاذب و هم نتایج منفی کاذب، می توانند برای شناسایی پدر و مادران مسئله ساز باشند که پیشنهاد می کند یک تطابق مثبت با LR پایین، به صورت «بدون تطابق» قطعی گزارش نشود [۴۷].

جدول ۵ تفسیر نسبتهای عدد احتمال پیشنهاد شده برای علوم پزشکی جنایی، توسط شبکه اروپایی مؤسسه علوم پزشکی قانونی

هم ارزی شفاهی	مقدار نسبت عدد احتمال
یافته ها از یک پیشنهاد نسبت به سایر پیشنهادها پشتیبانب نمی کند.	۱
پشتیبانی ضعیف از یک پیشنهاد نسبت به پیشنهاد دیگر.	۲-۱۰
پشتیبانی متوسط از یک پیشنهاد نسبت به پیشنهاد دیگر.	۱۰-۱۰۰
پشتیبانی نسبتاً قوی از یک پیشنهاد نسبت به پیشنهاد دیگر.	۱۰۰-۱۰۰۰
پشتیبانی قوی از یک پیشنهاد نسبت به پیشنهاد دیگر.	۱۰۰۰-۱۰۰۰۰
پشتیبانی بسیار قوی از یک پیشنهاد نسبت به پیشنهاد دیگر.	۱۰۰۰۰-۱۰۰۰۰۰۰
قویترین پشتیبانی از یک پیشنهاد نسبت به پیشنهاد دیگر.	>۱۰۰۰۰۰۰

۳.۴ احتمال پسین

اگر اطلاعات قابل اعتمادی در مورد شیوع یک ویژگی خاص وجود داشته باشد (به عنوان مثال جمعیتی با شیوع کاملاً مستند از یک بیماری خاص)، LR(+) مرتبط با نتیجه آزمایش را می توان به احتمال PP مبنی بر مثبت بودن قلم آزمون شده با توجه به نتیجه آزمایش مثبت، تبدیل کرد. این به عنوان یک احتمال پسین شناخته می شود و با استفاده از تئوری بیز (پیوست الف) تخمین زده می شود. یک شکل از این تئوری، با استفاده از نسبت عدد احتمال به صورت زیر است:

$$PP = \frac{\frac{P(+)}{P(-)}LR(+)}{\frac{P(+)}{P(-)}LR(+)+1} \qquad (۶)$$

در اینجا P(+) و P(-) احتمالات پیشین (قبلی) هستند، یعنی اطلاعات موجود قبل از آزمون، PP و PN اطلاعات پسین هستند. با در نظر گرفتن مثال قبلی از آنالیز یک آلودگی در یک محصول مواد غذایی توسط GC-MS که در آن نتیجه مثبت با یک LR(+) از $۴٫۹\times۱۰^{۴}$، با فرض اینکه P(+)=P(-) = ۰/۵، PP، ٪۹۹٫۹۹۸ می شود.

($PP=4.9\times10^4/(4.9\times10^4+1)$)

اگر یک نتیجه منفی گزارش شود، احتمال پسین نمونه ای که در واقع منفی بوده، PN، به وسیله معادله زیر تخمین زده می شود:

$$PN = \frac{\frac{P(-)}{P(+)}LR(-)}{\frac{P(-)}{P(+)}LR(-)+1} \qquad (۷)$$

این معادله شبیه معادله (۶) می باشد. P(+) و P(-) شیوع موارد مثبت و منفی را بیان می کند. به طور کلی، از انجایی که احتمال پسین به یک دسته بندی گزارش شده مربوط می شود، احتمال پسین را می توان به عنوان معیاری از احتمال درست بودن مقدار گزارش شده، در نظر گرفت. استفاده از احتمالات پسین در عمل دشوار است. گاهی اوقات احتمالات قبلی به اندازه کافی مناسب و قابل اعتماد، در دسترس نیستند. اگر چه برخی از نویسندگان پیشنهاد کرده اند که می توان با فرض اینکه نتایج مثبت و منفی به یک اندازه محتمل هستند، بر این نگرانی غلبه کرد به طوری که P(+)/P(-) =1 همیشه معقول نیست. گاهی اوقات به ویژه در کادر پزشکی قانونی، ممکن است استنتاج احتمالات قبلی برای یک مورد خاص از دانش موارد غیر مرتبط نامناسب باشد. درچنین مواردی، یک نسبت عدد احتمال (بخش ۴.۲) می تواند خلاصه مفیدی از اطمینان ارائه شده توسط یک نتیجه آزمایش، بدون نیاز به تعیین احتمالات قبلی ارائه دهد.

در برخی از زمینه ها مانند علوم پزشکی، در نظر گرفتن شیوع یک بیماری یا ویژگی در تصمیم گیری در مورد نتایج تجزیه ای کیفی می تواند برای تشخیص بسیار مهم باشد. تشخیص بیماری یا وضعیت بالینی بر اساس نتایج تجزیه ای بالینی نیز بر اطلاعات اضافی مانند رنگ مخاط، محل و شدت درد، سن و جنسیت، رفتارهای پرخطر و غیره تکیه خواهد داشت.

سهم این اطلاعات در تصمیم گیری نهایی در مورد نتیجه مشاهده نشده را می توان با محاسبه PP یا PN نشان داد، اگر چه پزشکان به طور معمول این محاسبات را انجام نمی دهند، نسبتاً انتظار می رود هنگام تشخیص بر اساس نتیجه آزمایش از اهمیت کلی شیوع آگاه باشند. جزئیات بیشتر در مورد این معیارها درفهرست کتاب ارائه شده است [۴،۸،۹].

۴.۴ قابلیت اطمینان معیارها

قابلیت اطمینان PP، LR یا PN محاسبه شده به قابلیت اطمینان نرخ نتیجه در نظر گرفته شده و برای احتمالات پسین به قابلیت اطمینان هر احتمال قبلی استفاده شده، بستگی دارد. جدول ۴ تعداد آزمون های مورد نیاز برای تشخیص قابل اعتماد یک یا چند پاسخ کاذب در احتمالات مختلف پاسخ کاذب را نشان می دهد. کمی سازی قابل اعتماد چنین نرخی معمولاً به موارد بیشتری نیاز دارد (به بخش ۳.۳.۳ مراجعه کنید).

تعداد موارد مورد مطالعه ممکن است نیاز به افزایش بیشتر داشته باشد تا دامنه کامل روش آزمون را پوشش دهد، به عنوان مثال، در آزمایش مواد غذایی ممکن است لازم باشد چندین بافت غذایی مختلف، بررسی شود. مدل سازی یک علامت دستگاهی که در تجزیه کیفی در نظر گرفته می شود، می تواند کمی کردن نرخ های نتایج کاذب پایین را امکان پذیر کند اما به کفایت داده های ورودی و الگوریتم مدل سازی، بستگی دارد.

برای تعیین PP یا PN از P(+) و P(-) بسیار متناقض، می توان از جدول ۴ برای تعریف تعداد مواردی (از جمعیت هدف) استفاده کرد که باید مطالعه شوند. کیفیت داده های ورودی برای تخمین این معیارها زمانی که قسمت های مختلف شواهد با هم ترکیب می شوند و معیارهایی که قدرت اطلاعات محاسبه شده را کمی می کنند، حیاتی تر است.

بنابراین توصیه می شود، معیارهای ارائه شده با احتیاط مورد استفاده قرار گیرند و جزئیات مربوط به داده های ورودی، نحوه استفاده از نتیجه گزارش شده و پیامدهای مربوطه را در نظر داشته باشید. تفسیر بیش از حد داده های عملکرد تجزیه کیفی، می تواند به اندازه نادیده گرفتن محدودیت های یک تجزیه کیفی خاص، مفید باشد.

۵.۴ عدم قطعیت نسبت ها

کیفیت آماری نرخ نتیجه تخمینی، که به تعداد آزمایش های مورد استفاده برای تعیین آنها بستگی دارد، می تواند به صورت فاصله اطمینان، CI، برای نرخ محاسبه شده، بیان شود. این فاصله اطمینان، همچنین به عنوان عدم قطعیت شرطی شناخته می شود، (۴.۴.۶ از [۱۸])، که معمولاً برای سطح اطمینان ۹۵٪ (۹۵٪ CI) محاسبه می شود. به عنوان مثال، فاصله اطمینان ۹۵٪ گسترده برای حساسیت SS نشان می دهد که مقدار واقعی SS می تواند بسیار متفاوت از تخمین باشد. همین منطق را می توان برای سایر نرخ های نتیجه مانند SP نیز اعمال کرد. از آنجایی که نرخ های نتیجه، از روی هر نوع دانش قبلی از جمعیت موارد، تخمین زده نمی شوند، این فواصل فقط کیفیت عملکرد تجزیه ای برآورد شده را مشخص می کند.

تفسیر فاصله اطمینان ۹۵٪ تا حدی مشابه آنچه در اندازه گیری عدم قطعیت بسط یافته اتفاق می افتد [۱]، می باشد. برای فاصله اطمینان ۹۵٪ (CI) احتمال ۵٪ وجود دارد که مقدار «واقعی» نرخ نتیجه، خارج از محدوده فاصله اطمینان باشد. به طور مشابه فاصله اطمینان ۹۵٪، برای نرخ تعیین شده تجربی، عدم قطعیت آماری را برای نرخ محاسبه شده فراهم می کند.

به عنوان مثال اگر توانایی یک روش آنالیز کیفی، برای شناسایی درست موارد مثبت، از تجزیه ۴۰۰ مورد از این اقلام آزمایش شود، هر ۴۰۰ نتیجه مثبت باشد، SS تخمینی ۱۰۰٪ مربوط به فاصله اطمینان ۹۵٪ می باشد که بین ۹۹٪ و ۱۰۰٪ قرار گرفته است. به عنوان مثال مقدار واقعی SS بین ۹۹٪ و ۱۰۰٪ با اطمینان

۹۵٪ متغیر است. اگر روش، تنها با ۵ مورد مثبت آزمایش شود، فاصله اطمینان ۹۵٪ از SS به وسیله ۵۷٪ و ۱۰۰٪ محدود خواهد شد . فاصله اطمینان ۹۵٪ اجازه می دهد تا کیفیت پارامترهای عملکردی روش تجزیه ای، بیان شود که برای تفسیر مناسب آنها، لازم است. در مثال های بالا هر دو SS، ۱۰۰٪ گزارش شده اند اما برآورد SS در مورد اول، بسیار قابل اعتماد تر است. محاسبه فاصله اطمینان ۹۵٪ برای SS و SP یک کار استاندارد در آزمایشگاه بالینی است (۱۰.۱.۳ از [۲۷]).

چندین مدل برای محاسبه فاصله اطمینان، منتشر شده است [۴۸]-[۴۹،۵۰،۵۱،۵۲،۵۳،۵۴]. فاصله نمره ویلسون [۵۴] برای سادگی و کاربردی بودن شمارش های کوچک، مورد استفاده قرار می گیرد. معادلات ۸ و ۹ برای محاسبه محدوده های مقادیر کم، $LL_{ss.95}$، و بالا، $HL_{ss.95}$، از فاصله اطمینان ۹۵٪ برای SS یا TP استفاده می شود.

$$LL_{ss.95} = \frac{A_1 - A_2}{A_3} \times 100 \quad (۸)$$

$$LL_{ss.95} = \frac{A_1 + A_2}{A_3} \times 100 \quad (۹)$$

که:

$$A_1 = 2tp + 1.96^2$$

$$A_2 = 1.96(1.96^2 + 4tp.\frac{fn}{tp + fn})^{1/2}$$

$$A_3 = 2(tp + fn + 1.96^2$$

از معادلات ۱۰ و ۱۱ برای محاسبه محدوده پایین، $LL_{sp.95}$، و محدوده بالای، $HL_{sp.95}$، فاصله اطمینان ۹۵٪ برای SP یا TN استفاده می شود.

$$HL_{sp.95} = \frac{B_1 - B_2}{B_3} \times 100 \quad (۱۰)$$

$$HL_{sp.95} = \frac{B_1 + B_2}{B_3} \times 100 \quad (۱۱)$$

که:

$$B_1 = 2tn + 1.96^2$$

$$B_2 = 1.96(1.96^2 + 4fp.\frac{tn}{fp + tn})^{1/2}$$

$$B_3 = 2(fp + tn + 1.96^2)$$

توصیه می شود یک مقدار حداقل یا هدف از $LL_{ss.95}$ یا $LL_{sp.95}$ (به عنوان مثال $LL^{tg}_{ss.95}$ یا $LL^{tg}_{sp.95}$) بر اساس هدف تجزیه، تعریف شود. هدف، به ویژه زمانی که تأثیر نتایج کاذب حیاتی است، بسیار مهم است. به عنوان مثال برای اجزای خونی که در انتقال خون استفاده می شود، توصیه می شود غربالگری بیماری های عفونی با آزمایش هایی مربوط به یک $LL_{ss.95}$ نزدیک به ۱۰۰٪، که می تواند تنها در صورتی تأیید شود که بسیاری از موارد مثبت در طول صحه گذاری، آزمایش شوند.

هنگامی که نرخ نتیجه با یک مقدار هدف حداقلی مقایسه می شود یا زمانی که افزایش یا کاهش در پارامتر بررسی می شود، توصیه می شود یک ارزیابی یک طرفه انجام شود. برای یک آزمون اطمینان ۹۵٪، توصیه می شود، ضریب ۱٫۹۶ به ۱٫۶۴ تغییر کند.

۵) گزارش نتیجه تجزیه ای کیفی

در حال حاضر آزمایشگاه های اعتباردهی شده نیازی به گزارش نتایج تجزیه کیفی با عدم قطعیت ندارند. بر این اساس، مثال های این بخش رویکردهای گزارش دهی ممکن را پیشنهاد می دهد، زمانی که آزمایشگاه برای کمک به مشتری این کار را انجام می دهد.

یک نتیجه مثبت را می توان با *TP* و *FP*، *LR*(+) یا *PP* و یک نتیجه منفی را می توان با *TN* و *FN*، *LR*(-) یا *PN* گزارش کرد. سایر معیارهای ارائه شده در جدول ۲ نیز می توانند برای گزارش اطمینان به نتیجه، استفاده شوند.

این معیارها معمولاً اطلاعاتی در مورد یک نتیجه آزمایش ارائه می دهند. به هر حال برای مواردی که میزان یک معیار برای دامنه تجزیه ای، ثابت است، چنین پارامترهایی را می توان به عنوان مشخص کننده روش تجزیه ای، تفسیر کرد.

چهار مثال زیر نشان می دهد که چگونه می توان نتایج کیفی را با عملکرد یا عدم قطعیت مربوطه، گزارش کرد.

مثال ۱: متن مورب، به عدم قطعیت تجزیه کیفی اشاره می کند.
خانم .A.B با ویروس سارس کووید (SARS-COV 2) آلوده شده است.
(آزمون با حساسیت ۹۰٪ و ویژه بودن ۹۹٪)

مثال ۲: متن مورب، به عدم قطعیت تجزیه کیفی اشاره می کند.
ادرار آقای .C.D حاوی بقایای کانرنون است.
(شناسایی با نسبت عدد احتمال $۴٫۹ \times ۱۰^{۴}$)

مثال ۳: متن مورب، به عدم قطعیت تجزیه کیفی اشاره می کند.
کوکائین در نمونه ۱۲۳ موجود است.
(شناسایی با نسبت عدد احتمال $۴٫۹ \times ۱۰^{۴}$ و شواهد بسیارقوی از حضور آنالیت در نظر گرفته شده است.)

مثال ۴: متن مورب، به عدم قطعیت تجزیه کیفی اشاره می کند.
بقایای بنزین در آوار آتش سوزی با کد نمونه ۴۵۶ شناسایی شد.
(شناسایی با احتمال پسین ۹۹/۹۹۸٪، تخمین زده شده از شبیه سازی مدل علامت و با فرض وجود یا عدم وحود آنالیت با احتمال یکسان است.)

۶) نتیجه گیری و توصیه

برای آزمایشگاه ها مهم است که حداقل، بحرانی ترین نرخ پاسخ کاذب را بررسی کنند برای برخی از معیارها هم نرخ مثبت کاذب و هم نرخ منفی کاذب، باید تعیین شود.

واقع بینانه است که انتظار داشته باشیم اکثر آزمایشگاه ها پارامترهای مربوط به روشهای تجزیه کیفی خود (یعنی شرایط تجزیه) را تحت کنترل کافی داشته باشند. شواهد آن معمولاً شامل موارد زیر است:

- شواهد روشنی از ردیابی اندازه شناسی کافی مقادیر پارامترهای تحت کنترل، به دلیل ارتباط آنها با آزمون
- شواهدی مبنی بر اینکه عدم قطعیت این پارامترها، برای این اهداف به اندازه کافی کوچک است.

منطقی است که انتظار داشته باشیم آزمایشگاه ها، از کدهای بهترین عملکرد منتشر شده در تجزیه کیفی در جایی که در دسترس هستند، از جمله استفاده از داده ها و مواد مرجع مناسب، پیروی کنند.

توصیه می شود گزارشهای کمی (یعنی عددی) از عدم قطعیت در نتایج آزمون کیفی به گونه ای ارائه شود که از تفسیر نادرست، جلوگیری شود. هر گاه به این نتیجه رسیدیم که نتیجه تجزیه ای به دست آمده، به ترتیب با نرخهای نتیجه واقعی کاذب بسیار پایین و خیلی زیاد همراه است، کاملاً منطقی است که نتیجه آزمایش را به عنوان غیرمنطقی، به معنای قطعیت ناکافی، گزارش کنیم.

۷) مثال ها

مثالها، پس از فهرست شدن دامنه کار آنها تشریح می شوند.

۱.۷ م۱ شناسایی ترکیبات توسط طیف سنجی جرمی با تفکیک پذیری پایین با استفاده از جستجوی پایگاه داده یا وجود یونهای مشخصه

۱.۱.۷ مقدمه

این مثال به حالت A یا B تقسیم می شود، جایی که روشهای مختلفی برای شناسایی ترکیبات بافتهای پیچیده توسط طیف سنجی جرمی با تفکیک پذیری پایین، استفاده می شود، ارائه همزمان و موازی این دو مورد، ماهیت جایگزین گزینه های شناسایی را برجسته می کند.

توجه داشته باشید که در شناسایی عملی، معمولاً از معیارهای متعددی مانند ترکیبی از تطابق طیفهای جرمی و زمان بازداری کروماتوگرافی استفاده می شود. این مثال فقط بر مؤلفه طیف سنجی جرمی، تمرکز دارد. مثال م۵، نمونه ای از معیارهای چندگانه را ارائه می دهد.

دامنه کاری

نوع تجزیه کیفی: تجزیه براساس معیارهای کمی
قلم/ بافت: الف) فرآورده های گوشتی ب) نمونه های پزشکی قانونی یا محیطی
پارامتر/ آنالیت: الف) دی اتیل استیل بو استرول، DES (هورمون رشد ممنوع برای گوشت گاو وطیور) یا ب) هروئین، DES و دی کلرو دی فنیل تری کلرو اتان (DDT)
نوع معیار طبقه بندی: ۱) شناسایی براساس رواداری برای فراوانی های نسبی (RA) یون های خاص طیف جرمی ۲) شناسایی طیف جرمی براساس حضور یونهای خاص طیف جرمی بدون توجه به مقادیر (RA)
فن/دستگاهوری: کروماتوگرافی گازی با طیف سنجی جرمی با تفکیک پذیری پایین، با استفاده از یونیزاسیون برخورد الکترون (GC-MS)
نوع گزارش نتایج: نسبت عدد احتمال

این مثال ارزیابی عدم قطعیت برای شناسایی ترکیبات توسط GC-MS را با استفاده از معیارهای شناسایی مختلف، توصیف می کند (بخش های ۷.۱.۱ و ۷.۱.۲). مثال ها نتایجی را برای شناسایی سه ترکیب (یعنی DES، هروئین و DDT) در دو نوع نمونه (فرآورده های گوشتی و نمونه های پزشکی قانونی یا محیطی) ارائه می دهند.

طیف سنجی جرمی، به ویژه در ترکیب با مرحله جداسازی کروماتوگرافی، ابزار قدرتمندی است که می تواند به شناسایی ترکیبات ناشناخته کمک کند. برای اکثر اهداف، طیف سنجی جرمی با تفکیک پذیری پایین با

استفاده از یونیزاسیون برخورد الکترون (EI) زمانی که شناسایی به جای کمی سازی مورد نیاز است، روش انتخابی است. یک طیف جرمی می تواند حاوی یونهای زیادی باشد که همه آنها برای اهداف تشخیصی، مفید نیستند. این سؤال مطرح می شود که آیا حداقل تعداد یون ها برای کفایت اطمینان از شناسایی صریح، وجود دارد؟

نکته: در برخی از زمینه های تجزیه ای، حداقل تعداد یون های مورد نیاز برای شناسایی ترکیبات، تعریف شده است [۲۳-۲۲-۲۱-۲۰].

۲.۱.۷ شناسایی بر اساس فراوانی نسبی یونهای مشخصه

اسفن[۲۵] (۵۵) حداقل تعداد یون هایی را که نیاز به پایش دارند تا شناسایی واضحی از دی اتیل استیل بواسترول[۲۶] (DES) در محصولات گوشتی ایجاد شود، بررسی کرد. داده های مربوط به مطالعه بعدی [۵۶]، بر اساس یک کتابخانه طیفی جرمی تجاری، حاوی حدود ۲۷۰۰۰۰ ورودی در جدول م۱.۱ ارائه شده است.

جدول م۱.۱، تعداد طیف های موجود در کتابخانه مورد استفاده را نشان می دهد که با معیارهای مشخص شده برای فراوانی نسبی، RA، یک یا چند یون مطابقت دارد. RA با تقسیم فراوانی یون مورد مطالعه بر فراوانی فراوان ترین یون (یعنی پیک پایه)، تخمین زده می شود. هدف این نرمال سازی، تولید یک پارامتر شناسایی است که کمتر به سطح آنالیت (به عنوان مثال، غلظت آنالیت) وابسته است. در جدول م۱.۱، می توان مشاهده کرد که وقتی یون های بیشتر و محدوده های فراوانی باریک تری در نظر گرفته شوند، تعداد تطابق هایی که در پایگاه داده اتفاق می افتد، به شدت کاهش می یابد.

جدول م۱.۱ تعداد طیف های یک کتابخانه وایلی با ۲۷۰۰۰۰ ورودی مطابق با معیارهای خاص، برای فراوانی نسبی برخی یون ها

معیار های شناسایی			
#	یون (m/z)	RA (%) بازه پذیرش	تطابق
۱	۲۶۸	۱-۱۰۰	۹۹۹۵
۲	۲۶۸ ۲۳۹	۱-۱۰۰ ۱-۱۰۰	۵۵۳۶
۳	۲۶۸ ۲۳۹	۹۰-۱۰۰ ۱۰-۹۰	۴۶
۴	۲۶۸ ۲۳۹	۹۰-۱۰۰ ۵۰-۷۰	۹

[25] Sphon

[۲۶] DES به عنوان هورمون رشد برای گاو و طیور مورد استفاده قرار گرفت و پس از اثبات خواص سرطان زایی آن، ممنوع شد.

۵	۲۶۸ ۲۳۹ ۱۴۵	۹۰-۱۰۰ ۵۰-۹۰ ۵-۹۰	۱۵
۶	۲۶۸ ۲۳۹ ۱۴۵	۹۰-۱۰۰ ۵۰-۷۰ ۴۵-۶۵	1(DES)[a]

RA: فراوانی نسبی (درصد فراوان ترین یون، پیک مبنا)
[a] تطابق منفرد، با طیف جرمی DES مطابقت دارد.

مجموعه معیارهای شناسایی شماره ۶، طیف جرمی DES را جدا می کند و منجر به یک تطابق منفرد می شود.

برای مقایسه با یک کتابخانه جایگزین، جدول م۲.۱ تعداد منطبق ها را از یک کتابخانه مرجع عمومی در دسترس، ارائه می کند، که شامل ۶۲۲۳۵ طیف، با در نظر گرفتن رواداری برای فراوانی نسبی (RA) یک یا چند یون [۵۷] می باشد. همانطور که رواداری های مرتبط با RA یون های بیشتر، باریک تر می شود، طیف جرمی از ترکیبات کمتری جدا می شود. جدول م۲.۱ تعداد طیف های منطبق بر سه ترکیب هدف مختلف یعنی DES، هروئین و DDT را نشان می دهد. هروئین وDDT به ترتیب برای تجزیه برخی از نمونه های پزشکی قانونی و محیطی مرتبط هستند.

مقایسه جداسازی طیف جرمی DES در هردو کتابخانه، که در جداول م۱.۱ و م۲.۱ شرح داده شده است، به این نتیجه می رسد که همانطور که انتظار می رود، تعداد تطابق ها به تعداد طیف های موجود در کتابخانه بستگی دارد. (معیار های شناسایی #۱ و#۲ را در جداول م۱.۱ و م۲.۱ ببینید). اگر تعداد تطابق ها برتعداد کل ورودی ها تقسیم شود، تفاوت های مشاهده شده در جداول م۱.۱ و م۲.۱ کاهش می یابد.

جدول م۳.۱ خلاصه ای از اطلاعات جمع آوری شده در جداول م۱.۱ و م۲.۱ از انتخابی ترین شناسایی هاست. جدول م۳.۱ اطلاعات جمع آوری شده را به FP و TP با ترکیب بیشتر به یک (+) LR تبدیل می کند، که عدم قطعیت یک نتیجه مثبت (یعنی گزارش حضور آنالیت) را تخمین می زند.

TP تخمین زده شده (یعنی تقریباً ٪۱۰۰) فرض می کند که رواداری های تعریف شده برای RA یون ها، تنوع آنها را درنظر می گیرد. در حالت ایده آل توصیه می شود، رواداری ها از مدل های تغییر پذیری علامت ساخته شده از طیف های تکراری محلول های نمونه با غلظت آنالیت مربوطه، تنظیم شود [۹و۸] (مثال م۵).

FP ارائه شده در جدول م۳.۱ فرض می کند که مزاحمت های احتمالی، همه ترکیباتی هستند که طیف جرمی آنها در کتابخانه مورد استفاده موجود است. بسیاری از ترکیبات موجود در محلول های نمونه، توسط GC-MS قابل تشخیص نیستند یا در آماده سازی نمونه، حذف می شوند. از سوی دیگر، بسیاری از ترکیباتی که طیف آنها در کتابخانه است، به دلیل ناسازگاری شیمیایی یا مستقل از منابع یا منشأ، احتمالاً در نمونه

های آنالیز شده وجود ندارند. بدترین حالت FP معادل یک در تعداد کل طیف ها، N، منهای یک (FP=1/(N-1)) تخمین زده می شود، زیرا مشخص است که FP، صفر نخواهد بود. توصیه می شود، محدودیت های توصیف شده در مورد نحوه ارزیابی FP، هنگام استفاده از این مقدار، در نظر گرفته شود. FP را می توان به طور متناوب، از مدل های نویز علامت، تخمین زد، همانطور که در مثال E5 مورد بحث قرار گرفت [۹ و ۸].

از داده های جدول م۳.۱ می توان دید که با استفاده از همان کتابخانه، همه آنالیتها با یک تطابق، تنها بسته به تعداد ورودیهای کتابخانه، TP، FP و LR(+) یکسانی دارند. با توجه به معیار های تعریف شده توسط شبکه اروپایی مؤسسات علوم پزشکی (جدول ۵) شناسایی یک آنالیت به وسیله طیف سنجی جرمی، با استفاده از یک روش شناسایی توصیف شده، شواهد «بسیار قوی» از حضور آنالیت ایجاد می کند (LR (+) بین $۱۰^{۴}$ و $۱۰^{۶}$). فرض کنید شناسایی توسط زمان بازداری آنالیت در سیستم کروماتوگرافی (یعنی در GC) و پنجره زمان بازداری، کافی است. در این مورد LR(+) شناسایی می تواند افزایش یابد (مثال E5).

جدول م۲.۱ تعداد طیف ها در یک کتابخانه در دسترس عموم، با ۶۲۲۳۵ ورودی، مطابقت دادن معیارهای خاص برای فراوانی نسبی برخی یون ها

معیار های شناسایی			
تطابق ها	بازه پذیرش(٪)RA	یون (m/z)	#
۳۵۹۷	۱-۱۰۰	۲۶۸	۱
۱۵۹۷	۱-۱۰۰	۲۶۸	۲
	۱-۱۰۰	۲۳۹	
۸۳	۵۵-۹۵	۲۶۸	۳
۴	۵۵-۹۵	۲۶۸	۴
	۳۰-۷۰	۲۳۹	
۱ (DES)	۵۵-۹۵	۲۶۸	۵
	۳۰-۷۰	۲۳۹	
	۶۰-۱۰۰	۱۴۵	
۱۶۲۷	۱-۱۰۰	۳۶۹	۶
۵۲۶	۱-۱۰۰	۳۶۹	۷
	۱-۱۰۰	۳۲۷	
۴۳	۴۵-۸۵	۳۶۹	۸
۱ (هروئین)	۴۵-۸۵	۳۶۹	۹
	۶۰-۱۰۰	۳۲۷	
۱۲۴۲	۱-۱۰۰	۳۵۲	۱۰
۲۳۴	۱-۱۰۰	۳۵۲	۱۱
	۱-۱۰۰	۲۳۵	

۱۲	۳۵۲	۱-۴۰	۱۱۴۰
۱۳	۳۵۲	۱-۴۰	۱(DDT)
	۲۳۵	۶۰-۱۰۰	
۱۴	۳۵۲	۱-۴۰	۷
	۲۳۵	۱-۱۰۰	
	۲۳۷	۴۸-۸۸	
۱۵	۳۵۲	۱-۴۰	۱(DDT)
	۲۳۵	۶۰-۱۰۰	
	۲۳۷	۴۸-۸۸	

RA: فراوانی نسبی (درصد فراوان ترین یون ، پیک مبنا)

[a] تطابق منفرد مربوط به طیف جرمی DES، هروئین یا DDT است.

۳.۱.۷ شناسایی مبتنی بر حضور یونهای مشخصه – احتمالات تطابق شانس تخمینی

فرض کنید به جای شناسایی آنالیت با استفاده از رواداریها برای RA یون های خاص، شناسایی بر اساس حضور ساده سه یون انتخابی است. در آن صورت از ریاضیات زیر می توان برای ارزیابی تقریبی عدم قطعیت شناسایی، استفاده کرد. فرض کنید از طیف سنج جرمی با تفکیک پذیری پایین استفاده می شود که فقط می تواند واحد های m/z (یعنی واحد های نسبت جرم به بار) را تشخیص دهد و یون ها باید مقادیر بین m/z ۱۸۰ و m/z ۴۸۰ را داشته باشد. در این مورد تقریباً ۳۰۰ مقدار ممکن m/z در یک طیف، قابل مشاهده است. از این رو، از آنجایی که تعداد ترکیب های ۳۰۰ شیء در یک زمان، $300!/[3! \times (300-3)!] = 4455100$ می شود برای سه پیک، احتمال تطبیق سه پیک تصادفی با سه یون انتخابی به طور تصادفی، با فرض اینکه همه نسبت های m/z به یک اندازه محتمل هستند، $۱ \div ۴٫۶ \times ۱۰^{۶}$ یا تقریباً $۲٫۲ \times ۱۰^{-۷}$ می باشد. با این حال، این اجازه نمی دهد که طیف های جرمی معمولاً بیش از سه یون در ناحیه مورد نظر داشته باشند. این موضوع، احتمال اینکه یک طیف جرمی حاوی n یون، ممکن است با سه عامل انتخاب شده با فاکتور $n!/[3! \times (n-3)!]$ مطابقت داشته باشد را افزایش می دهد. با درنظرگرفتن ۱۰ به عنوان تعداد نوعی پیک ها، احتمال تطابق شانسی به میزان $10!/3! \times 7! = 120$ افزایش می یابد. . بر این اساس، احتمال تطابق شانسی تخمین زده شده، $۱۲۰ \times ۲٫۲ \times ۱۰^{-۷}$ یا تقریباً $۲٫۶ \times ۱۰^{-۵}$ می باشد. این، یک نرخ مثبت کاذب تقریبی را ارائه می دهد، FP.

با فرض اینکه سطوح آلاینده به اندازه کافی بالا باشد تا طیف جرمی قابل اعتمادی ارائه دهد، می توان در نظر گرفت که TP تقریباً ۱ یا ۱۰۰٪ است.

جدول م۳.۱ عدم قطعیت شناسایی چندین آنالیت در بافت های مختلف، توسط MS-GC، از تعداد مطابقت طیف ها از کتابخانه مورد استفاده تخمین زده شده، با در نظر گرفتن رواداری برای فراوانی نسبی دو یا سه یون.

آنالیت	قلم تجزیه شده	تعداد طیف های کتابخانه، N	تعداد تطابق ها	TP (%)[b]	FP (%)[c]	LR(+) (TP/FP)
DES	محصولات گوشتی	۲۷۰۰۰۰	یک در N	~۱۰۰	$۳/۷\times۱۰^{-۴}$	$۲/۷\times۱۰^{۵}$
DES	F,E	۶۲۲۳۵	یک در N	~۱۰۰	$۱/۶\times۱۰^{-۳}$	$۶/۲\times۱۰^{۴}$
هروئین	F,E	۶۲۲۳۵	یک در N	~۱۰۰	$۱/۶\times۱۰^{-۳}$	$۶/۲\times۱۰^{۴}$
DDT	F,E	۶۲۲۳۵	یک در N	~۱۰۰	$۱/۶\times۱۰^{-۳}$	$۶/۲\times۱۰^{۴}$

DES: دی اتیل استیل بو استرول – FوE : نمونه های پزشکی قانونی و محیطی – TP = نرخ مثبت واقعی

FP: نرخ مثبت کاذب – LR(+) : نسبت عدد احتمال (TP/FP) – (درجدول ۲ ببینید)

[a] تعداد تطابق با درنظر گرفتن رواداری های تعریف شده برای RA از یون های خاص

[b] تخمین خوش بینانه TP (تقریباً ۱۰۰٪)

[c] تخمین زده شده براساس 1/(N-1) (بد ترین حالت از اطلاعات جمع آوری شده)

FP وTP تخمین زده شده در بالا را می توان با هم ترکیب کرد تا یک LR(+) $۱\div۲/۶\times۱۰^{-۵}$برابر با $۳/۸\times۱۰^{۵}$ به دست آورد که یک حد بالایی را برای نسبت عدد احتمال برای شناسایی ها بر اساس روش توصیف شده، نشان می دهد.

توجه: این تخمین فرض می کند همه ترکیب های یونی یک اندازه محتمل و ممکن می باشد و اینکه ظاهر آنها، مستقل است. این به عنوان یک تقریب، شناخته شده است (به مثال ۷.۲ مراجعه کنید). بنابراین احتمال تطابق شانس مؤثر، بسیار بیشتر از رقم محاسبه شده است که منجر به نسبت عدد احتمال پایین تر می شود.

۴.۱.۷ ملاحظات پایانی

روش شناسی ها برای تخمین LR(+) شناسایی توسط GC-MS ارائه شده دراین مثال، نسبت به اعتبار نتایج تجزیه ای کیفی، بیش از حد خوش بینانه است (برای مقایسه با احتمالات تطابق واقعی به مثال ۷.۲ مراجعه کنید). بنابراین توصیه می شود این محاسبات، فقط به عنوان ارزیابی اولیه اعتبار شناسی استفاده شوند. مثال م۵ درمورد تعیین های جایگزین و واقعی تر عدم قطعیت شناسایی انجام شده توسط GC MS /MS بحث می کند.

اگر چه روش شناسی برای تخمین احتمال های تطابق شناسی و نسبت های عدد احتمال ارائه شده در اینجا، ممکن است خوش بینانه باشد، برای مثال شناسایی هروئین در نمونه های صحنه جرم، صرفاً بر اساس وجود یون هایی با m/z ۳۶۹ و ۳۲۷ کافی نیست، زیرا مرتبط با LR(+) از ۱۱۸ ($\{100\times[526/(62235-1)]\}/\{FP\}=$ ۱۰۰/ TP = 118) (روش شناسی بخش ۷.۱.۱ مورد ۷ جدول E1.2) یا 4.5×10^4 که $(4.5\times10^4=100/\{[1/(44850-1)]\times100\})$ (44850 = 300!/2!(300-2)!) (روش شناسی بخش ۷.۱.۲) بسته به رویکرد مورد استفاده برای ارزیابی

اطمینان در نتایج می باشد. درعمل این نشان می دهد که معیارهای اضافی یا آزمایش های تأییدی بیشتر، برای ایجاد اطمینان کافی، لازم است.

به طور مشابه اگر چه اسفن [۵۵] و دیگران پیشنهاد کردند که بقایای داروهای دامپزشکی درگاو را می توان با درنظر گرفتن سه یون طیف جرمی، برای پایش رسمی مواد غیر مجاز، شناسایی کرد. اتحادیه اروپا (Eu) نیاز به جمع آوری شواهد اضافی از وجود این ترکیبات دارد [۲۰]. به عنوان مثال، اگر فقط امکان پایش دو یون مشخصه توسط GC-MS در سطوح مناسب آنالیت، وجود داشته باشد، باید دو مرحله کروماتوگرافی مستقل براساس یونیزاسیون برخورد الکترون یا یونیزاسیون شیمیایی برای تأیید وجود آنالیت، در نظر گرفته شود [۲۰].

۲.۷ م۲ شناسایی ترکیبات خالص شده توسط طیف سنجی مادون قرمز

دامنه کاری

نوع تجزیه کیفی: تجزیه بر اساس معیارهای کمی
قلم/ بافت: ترکیب شیمیایی خالص شده
پارامتر/تجزیه: ترکیبی از پایگاه داده طیف مادون قرمز موجود
نوع معیار دسته بندی: مطابقت عدد موجی سه یا شش نوار طیف مادون قرمز در فاصله $[500,1800]\ cm^{-1}$
فن/دستگاهوری: طیف سنجی مادون قرمز
نوع گزارش نتایج: نسبت عدد احتمال

چندین نویسنده، استفاده از آماره های پایگاه داده در ارزیابی معیارهای تجزیه کیفی را بررسی کرده اند. دی روگ و همکاران [۵۸]، معیارهای برای شناسایی بقایای دارو های دامپزشکی در محصولات گوشتی (به بخش ۷.۱.۲ مراجعه کنید) پیشنهاد کردند. نویسندگان مقادیر نشان دهنده احتمال تطابق تصادفی را بر اساس یک مدل دو جمله ای، ارائه می کنند. السیون و همکاران نشان داده اند که توزیع فراهندسی، مدل مناسب تری برای تطابق تصادفی در طیف هاست. زیرا امکان تعداد کمی از پیک های تطبیق بین دو طیف را فراهم می کند که هر دو دارای تعداد بیشتری از پیک ها هستند [۵]. نویسندگان اخیر، هنگامی که طیف مادون قرمز با کتابخانه طیفی مقایسه می شود، بر روی تطابق های شانسی، تمرکز کردند. الیسون و همکاران [۵]، قابلیت اطمینان شناسایی ترکیبات خالص شده را با مقایسه طیف مادون قرمز به دست آمده با طیف های یک کتابخانه مورد مطالعه، قرار داد.

کتابخانه مورد استفاده توسط الیسون و همکاران، کتابخانه سدتلر[۲۷] حاوی طیف هایی از ۵۹۶۲۶ مواد مختلف بود. یک زیر مجموعه تصادفی از ۳۰ ترکیب از این کتابخانه انتخاب شد و تعداد نوار ها، m، در فاصله cm^{-1} (۵۰۰-۱۸۰۰) برای هر ترکیب ذکر شد. مشخص شد که میانگین تعداد نوارها در هر طیف در بازه cm^{-1} (۵۰۰-۱۸۰۰)، M، ۱۶ بود. تفکیک پذیری طیفی موجود cm^{-1} ۴ بود و این دلالت بر وجود ۳۲۵ = ۴÷۱۳۰۰ موقعیت های پیک مجزا، N، در بازه cm^{-1} (۵۰۰-۱۸۰۰) داشت. برای هر طیف مختلف در زیر مجموعه انتخاب شده، کل پایگاه داده، دو بار، ابتدا برای حداقل n=۳ پیک منطبق و بار دوم برای حداقل n=۶ پیک منطبق، جستجو شد.

برای $n \geq 3$ ۳ تعداد مطابقتهای مشاهده شده، تقریباً دو برابر تعداد پیش بینی شده توسط توزیع فراهندسی[۲۸] بود. برای $n \geq 6$ ۶، اگرچه تعداد مطابقت ها به طور قابل توجهی کمتر بود، همانطور که انتظار می رود، مطابقت های مشاهده شده به میزان ۱۰ برابر مطابقت پیش بینی شده، بیشتر بود. بخشی از داده ها، برای شش پیک منطبق در جدول م۱.۲ ارائه شده است.

[27] Sadtler
[28] Hypergeametric

جدول م۱.۲ تطابق های شانسی در برابر شش نوار در پایگاه داده مادون قرمز

ترکیب	تعداد نوارهای m در محدوده	احتمال تطابق شانسی[a]	تطابق های پیش بینی شده[b]	تطابق های مشاهده شده	LR(+)
۱-کلرو ۳-(۱-نفتیلوکسی)-۲-پروپانول	۲۳	$۳/۱۹\times۱۰^{-۴}$	۱۹	۱۹۲	۳۱۱
α-سیانو سینامیک اسید، متیل استر	۱۷	$۵/۰۳\times۱۰^{-۵}$	۳	۲۹	۵۰۵۶
فنیل v-تری آزولو(۱-۵-α) پیریدین-۳- یل کتون	۲۴	$۴/۱۹\times۱۰^{-۴}$	۲۵	۱۹۰	۳۱۴
بنزو-β-تیوفنل-۶-اکریلیک اسید	۲۰	$۱/۳۴\times۱۰^{-۴}$	۸	۵۲	۱۱۴۷
۳-((دی پروپیل آمینو)متیل) ۱-۵ نیتروایندول	۱۷	$۵/۰۳\times۱۰^{-۵}$	۳	۲۹	۲۰۵۶
۲-مسیتیل-۵-فنیل-اکسازول	۲۲	$۲/۵۲\times۱۰^{-۴}$	۱۵	۹۹	۶۰۲
p-هیدروکسی-بنزوئیک اسید	۱۸	$۶/۷۱\times۱۰^{-۴}$	۴	۴۴	۱۳۵۵
کاپروئیک اسید، ایزوبوتیل استر	۸	$۱/۳۶\times۱۰^{-۷}$	۰	۱	۵۹۶۲۶
۱ -بروموآدامنتان	۱۰	$۹/۶۴\times۱۰^{-۷}$	۰	۱	۵۹۶۲۶
فنیل پروپیل اتر	۱۷	$۵/۰۳\times۱۰^{-۵}$	۳	۴۷	۱۲۶۹

[a] احتمال تطابق شانسی، احتمالی است که n=۶ پیک که به طور تصادفی در دو طیف از نوار m با فرض وقوع تصادفی نوارها در سراسر طیف مطابقت دارند. احتمال با استفاده از توزیع فراهندسی محاسبه شد [۵].

[b] تعدادپیش بینی شده تطابق ها، احتمال تطابق شانسی ضرب در تعداد طیف های موجود در پایگاه داده است.

احتمال های تطابق شانسی محاسبه شده، برای تطابق های شش پیک، در بازه ($۱۰^{-۵}$ و $۱۰^{-۷}$) بود. احتمال تطابق شانسی برای یک ترکیب، وقتی در تعداد ورودی های پایگاه داده ضرب می شود، تعداد ترکیبات مناسب با معیارهای جستجو را تخمین می زند. در مورد دو ترکیب از ترکیبات جدول م۱.۲، یعنی کاپروئیک اسید ایزو بوتیل استر و برومو آدامانتان، معیارهای جستجو، یک تطابق واحد ایجاد می کند و از این رو در صورت مشکوک شدن به این ترکیبات، کافی به نظر می رسند. تعداد زیادی تطابق بیشتر برای ترکیبات باقی مانده تولید می شود که نشان دهنده نیاز به معیار های دقیق تر است.

با فرض اینکه TP، تقریباً ۱۰۰٪ می باشد (چون طیف IR ترکیبات خالص به طور قابل اعتمادی با طیف مرجع خود مطابقت دارد) و با درنظر گرفتن FP به عنوان نسبت بین تعداد تطابق های مشاهده شده و تعداد کل طیف های کتابخانه (یعنی ۵۹۶۲۶)، این امکان وجود دارد که LR(+) شناسایی را تخمین زد. آخرین ستون جدول م۱.۲، LR(+) تخمین زده شده را نشان می دهد (محاسبه شده به صورت TP/FP، با TP=۱/۰). LR(+) گزارش شده، بسیار پایین تر از حداقل میزان $۱۰^{۶}$ در نظر گرفته شده برای دسته بندی شواهد به عنوان «بسیار

قوی» (جدول ۵) نشان می دهد یک تطابق شش پیکی ساده در طول موج، به تنهایی، ممکن است بدون معیارهای اضافی، اطمینان کافی را ایجاد نکند. معیارهای اضافی (مانند تطبیق شدت پیک های اضافی، عدم وجود پیک هایی که در ترکیب هدف وجود ندارند، تطبیق زمان بازداری کروماتوگرافی یا نیاز به تطابق بصری نزدیک با طیف کامل) ممکن است برای ارائه شناسایی به اندازه کافی بدون ابهام، مورد نیاز باشد.

این مثال تأکید می کند که پایگاه های داده مرجع، که کتابخانه های طیفی یکی از انواع آنها هستند، فقط می توانند اطلاعات نشان دهنده ای در مورد نرخ پاسخ کاذب، به دست آورند. نرخ پاسخ فقط برای جمعیت های مشابه با جمعیت نمونه های مورد انتظار، کاملاً قابل اعتماد است. همچنین این مسؤلیت تحلیل گر است که تصمیم بگیرد، در صورت وجود، کدام یک از مجموعه ای از تطابقات، با یک مجهول مطابقت دارد.

۳.۷ م۳ شناسایی دارو های غیر مجاز در ادرار توسط فن ایمونو اسی چندگانه آنزیمی (EMIT) ویک فن جایگزین:

دامنه کاری

نوع تجزیه کیفی: تجزیه بر اساس معیارهای کمی (مطالعه با استفاده از اطلاعات کیفی)
قلم/بافت: ادرار
پارامتر/آنالیت: کوکائین، متادون یا مواد افیونی
نوع معیار دسته بندی: مشخص نشده
فن/دستگاهوری: فن ایمونو اسی چندگانه آنزیمی، EMIT و یک فن جایگزین اختصاصی
نوع گزارش نتایج: نسبت عدد احتمال و احتمال صحیح بودن نتیجه مثبت

اگرچه تعیین های EMIT، شامل پردازش یک علامت دستگاهی است، این مثال، عملکرد این تجزیه کیفی را از روی نرخ های تعیین شده نتایج مثبت و منفی کاذب به طور تجربی، ارزیابی می کند. بنابراین این مثال، تعیین کیفیت یک تجزیه کیفی براساس اطلاعات کیفی را نشان می دهد.

استفاده از پایگاه داده های نتایج نمونه، برای بدست آوردن احتمالات مربوطه برای ارزیابی بیزین عملکرد تجزیه کیفی، در متون گزارش شده است. برای آزمایش داروهای مضر در ادرار، فرارا و همکاران [۳۳]، پایگاه داده هایی را جمع آوری کردند که حاوی اطلاعاتی در مورد انواع داروها، روش های تجزیه ای، نرخ پاسخ منفی برای فنون و دارو های رایج می باشند. برای آزمایشگاه نویسندگان ذکر شده، جدول م۱.۳ بخشی ازاین داده ها را برای EMIT خلاصه می کند. این جدول همچنین احتمال پیشین مثبت تخمینی نمونه آزمایشی که واقعاً مثبت می باشد، PP، را همانطور که در معادله (۶) توضیح داده شده است، نشان می دهد. برای محاسبه، شیوع نتایج منفی، به صورت $\big(1 - P(+)\big)$، $LR(+) = \frac{TP}{FP}$ و $TP = 1 - FN$ به دست می آید.

جدول م۱.۳ احتمالات برای تشخیص EMIT دارو های مضر در ادرار [۳۳]

مقادیر ویژگی های عملکردی برای داروهای مختلف یا کلاس های دارویی			
کوکائین	متادون	مواد مخدر	احتمالات
۰/۲۰	۰/۲۶	۰/۴۴	P (+)
۰/۰۰۹	۰/۰۰۴	۰/۰۲۸	FP
۰/۰۵۶	۰/۰۱۸	۰/۰۶۹	FN
۰/۹۶۳	۰/۹۸۸	۰/۹۶۳	PP
P (+) = شیوع نتایج مثبت			

به عنوان مثال، برای شناسایی متادون، PP با معادله م۱.۳، ۰/۹۸۸ ارزیابی می شود (PP احتمال پیشین است که از تئوری بیز تعیین می شود). به عبارت دیگر، تحلیل گر می تواند بیش از ۹۸ درصد مطمئن باشد که

پاسخ مثبت به متادون، واقعاً نشان دهنده وجود این دارو است. با این حال، توجه داشته باشید که این تا حدی به شیوع بالای مشاهده شده در جمعیت خاص نمونه گیری شده، بستگی دارد.

PP ممکن است درجمعیت عمومی بسیار کمتر باشد، به عنوان مثال در جمعیتی که مصرف دارو چندان مکرر نیست.

$$PP = \frac{\frac{P(+)}{P(-)}LR(+)}{\frac{P(+)}{P(-)}LR(+)+1} = \frac{(\frac{0.26}{-0.26})(\frac{1-0.018}{0.004})}{(\frac{0.26}{1-0.26})(\frac{1-0.018}{0.004})+1} = 0.98 \qquad \text{م}۱.۳$$

جدول E3.2 داده های مشابهی را برای یک فن غیر ایمونو شیمیایی متفاوت نشان می دهد. توجه داشته باشید که FP برای کوکائین با این فن (روش)، صفر گزارش شده است. با این حال قابل بحث است که آیا نرخ پاسخ کاذب، برای چنین آزمایش های غربالگری واقعاً می تواند صفر باشد. دراین مورد، هیچ مثبت کاذبی یافت نشد، اما اگر نمونه های بیشتری آنالیز شده بود، ممکن بود یک یا چند مثبت کاذب، ظاهر شود. بر این اساس، یک نرخ پاسخ کاذب، در محاسبه احتمال پیشین استفاده شده است.

جدول م۱.۳ احتمالات تشخیص داروهای مضر در ادرار، با فن اختصاصی [۳۳]

مقادیر ویژگی های عملکردی برای دارو های مختلف یا کلاس های دارویی			
کوکائین	متادون	مواد مخدر	احتمالات
۰/۲	۰/۲۶	۰/۴۴	P(+)
۰/۰۰۰	۰/۰۱۲	۰/۰۳۸	FP
۰/۲۴۷	۰/۱۷۹	۰/۲۷۶	FN
۰/۹۹۵*	۰/۹۶۰	۰/۹۳۷	PP
* PP با استفاده از بدترین تخمین مقدار FP برابر با ۰/۰۰۱ (یک در ۱۰۰۰ آزمایش) محاسبه شد.			

با بررسی مجدد متادون، ۰/۹۶ = PP می باشد. به طور منطقی، این یک احتمال نسبتاً بالاست، اگر چه کمی کمتر از آنچه توسط آزمون EMIT تولید شده است، متقاعد کننده است.

برای نشان دان اثر ترکیب داده ها، فرض کنید که هردو تست غربالگری انجام شده است. اگر در هر مورد، پاسخ مثبت به دست آید، آنگاه بر اساس ترکیب معادله ۵ و ۶، PP مرکب، ۰/۹۹۹۸ می شود (به معادله م۲.۳ مراجعه کنید:

$$PP = \frac{\frac{P(+)}{P(-)}LR(+)}{\frac{P(+)}{P(-)}LR(+)+1} = \frac{(\frac{0.26}{1-0.26})(\frac{1-0.018}{0.004})(\frac{1-0.179}{0.012})}{(\frac{0.26}{1-0.26})(\frac{1-0.018}{0.004})(\frac{1-0.179}{0.012})+1} = 0.9998 \qquad \text{م}۲.۳$$

در این مثال، مقادیر شیوع قابل اعتماد (یعنی احتمالات قبلی) در دسترس هستند. اگر اینها در دسترس نبودند یا اگر تحلیل گر، ترجیح می داد از آنها استفاده نکند، می توان به جای آن، از نسبت های عدد احتمال استفاده نمود. مقادیر متناظر ۲۴۶ (EMIT) و ۶۸ ~ (اختصاصی) است که یک نسبت عدد احتمال مرکب تقریباً

۱۷۰۰۰ (معادله(E3.3)) را ارائه می دهد که مطابق با جدول ۵، با شواهد بسیار قوی وجود متادون مطابقت دارد.

$$LR(+) = \left(\frac{1-0.018}{0.004}\right) + \left(\frac{1-0.179}{0.012}\right) = 1.7 \times 10^4 \qquad \text{م۳.۳}$$

درهمه موارد، GC-MS به عنوان یک فن مرجع برای ایجاد نرخ نتایج کاذب، استفاده شد. پایگاه داده های خاص، که در اینجا به آن اشاره شده است، برای آنالیت های مورد مطالعه، کاملاً جامع می باشد و برای ارائه مجموعه ای نماینده، که تحلیل بیزین داده ها را میسر می کند، طراحی شده است. وقتی داده های بیشتر اضافه می شود، به ناچار برخی مقادیر از دست می رود و تعداد و صحت پیش بینی ها، بیشتر بهبود می یابد. مزیت دیگر پایگاه داده ای که برای ثبت داده های نماینده، برای چندین فن مختلف تنظیم شده است، اطلاعاتی است که برای بهینه سازی عملکرد تجزیه ای، فراهم می کند.برای مثال انتخاب یک روش غربالگری با نرخ مثبت کاذب پایین، هزینه های تجزیه های تأییدی گران قیمت را به حداقل می رساند. با این حال، لازم است عوامل دیگری مثل حد تشخیص فن، نرخ منفی کاذب آن و سرعت و هزینه تجزیه، در نظر گرفته شوند.

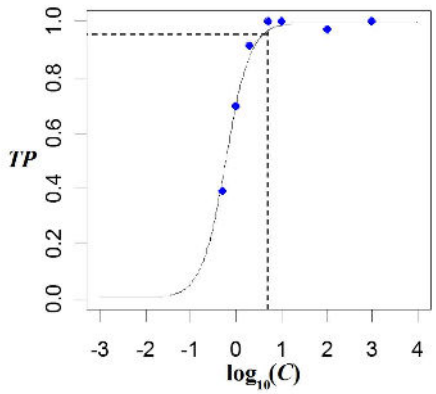
۴.۷ م۴ شناسایی ژن SRY انسانی در موارد بیولوزیکی باqPCR:

دامنه کاری

نوع تجزیه کیفی: تجزیه براساس معیارهای کمی (فلورسانس بیشتر از آستانه)
قلم/بافت: مواد بیولوژیکی
پارامتر/تجزیه: ژن SRY (ناحیه تعیین کننده جنسیت)
نوع معیار دسته بندی: مشخص نشده
فن/ابزار: واکنش زنجیره ای پلی مراز کمی (qPCR)
نوع گزارش نتایج: نرخ مثبت واقعی

اگرچه تفسیر qPCR، شامل پردازش یک علامت دستگاهی است، در این مثال، عملکرد با تعیین TP در غلظت های مختلف DNA ارزیابی می شود.

شکل م۱.۴ برخی از داده های تجربی از مطالعه تشخیص ژن SRY انسانی را به وسیله PCR کمی (qPCR) در موارد بیولوژیکی [۳۷] نشان می دهد. داده ها از یک سنجش ۳ صفحه ای با استفاده از سنجش ۵- نوکلئاز با یک پروب فلوروژنیک برچسب دار دوگانه (TaqMan)، که به سمت ناحیه SRY اختصاصی مردانه ژنوم انسان هدایت می شود، تولید می شود. اگر فلورسانس مشاهده شده از یک آستانه از پیش تعیین شده در ۵۵ سیکل تقویت فراتر رود، یک نتیجه مثبت در نظر گرفته می شود. اگرچه روش های مختلف آماده سازی، به عنوان بخشی از مطالعه صحه گذاری استفاده می شود، آزمون های آماری، تفاوت معنی داری نشان ندادند، بنابراین، داده ها به عنوان یک مجموعه داده واحد، در نظر گرفته شدند.



شکل م۱.۴ داده های تشخیص DNA پردازش شده توسط رگرسیون لجستیک. TP به عنوان تابعی از $\log_{10}(C)$ ترسیم می شود که در مطالعه ای در مورد تشخیص DNA و قابلیت دسته بندی، C، شماره نسخه ژن است. نقاط، نسبت نتایج مثبت را از ۳۶ تکرار در هر شماره نسخه، نشان می دهد. خط توپر، تناسب رگرسیون لجستیک را با b_0= 0.85 و b_1=3.75 نشان می دهد. (در بخش ۳.۳.۶ ببینید). خط چین TP=0.95 و غلظت لگاریتمی مربوطه از $\log_{10}(C)$ = 0.56 برای C = 3.6 نسخه را نشان می دهد.

توجه داشته باشید که نقطه در $\log_{10}(C) = 2$، TP برابر با ۱۰۰٪ را نشان نمی دهد، هر چند نقاط هر دو طرف این کار را انجام می دهند و ارزیابی صحیح قابلیت دسته بندی را دشوار می کند. با این حال، منحنی رگرسیون اجازه می دهد تا یک مکان دقیق معقول از حد مؤثر تشخیص، همانطور که در شکل نشان داده شده است، با استفاده از یک مدل پیوسته به طور مؤثر، داده های شمارش تصادفی را هموار می نماید. این یک مزیت قابل توجه است. این به تحلیل گر اجازه می دهد تا به جای نیاز به تکرارهای زیاد در سطوح معدود، تعداد زیادی از غلظت ها با تکرارهای نسبتاً کمی در هر سطح، مطالعه کند و همچنان تخمین های احتمال نسبتاً قابل اعتمادی را به دست آورد.

این مجموعه داده، یک هشدار مهم در مدل سازی را نشان می دهد. داده های غلظت، رسم شده و در دامنه لگاریتمی، مدل سازی می شود. یک روش معمول در کار با غلظت های DNA تجربی یا شمارش های میکروبی برای یک متغیر وابسته، تبدیل اغلب توسط توزیع خطا دیکته می شود. با این حال، هیچ دلیل قانع کننده ای برای انتخاب تبدیل لگاریتمی برای متغیر مستقل در این مثال، وجود ندارد. انتخاب، اساساً دلخواه است. انتخاب مدل لجیستیک (منطقی) نیز همینطور است. مدل های دیگر نیز ممکن است به خوبی با داده ها، مطابقت داشته باشند. در جاهایی که مدل های متفاوت، اما صاف به طور مشابه با داده ها مطابقت دارند، درون یابی، به شدت به انتخاب مدل، حساس نیست. با این حال، احتمالات انتهایی می توانند در واقع نسبت به انتخاب مدل، بسیار حساس باشند. نتیجه این است که حتی زمانی که یک مدل توصیف خوبی از داده ها و احتمالات و حدود تشخیص نسبتاً قابل اعتماد در آن محدوده ارائه می دهد، برون یابی تخمین های احتمال، بسیار فراتر از محدوده مورد مطالعه، بدون شواهد قابل توجهی از اعتبار مدل بسیار نا امن است.

۵.۷ م۵ شناسایی بقایای آفت کش ها در مواد غذایی توسط GC-MS/MS بر اساس زمان بازداری و نسبت فراوانی یون

دامنه کاری

نوع تجزیه کیفی: تجزیه بر اساس معیارهای کمی
قلم/بافت: مواد غذایی با منشأ گیاهی
پارامتر/تجزیه: کلرپیریفوس متیل و مالاتیون
نوع معیار/دسته بندی: فواصل زمان بازداری و نسبت های فراوانی یون های مشخصه
فن/دستگاهوری: GC-MS/MS
نوع گزارش نتایج: نسبت عدد احتمال و احتمال صحیح بودن نتیجه مثبت

این مثال، تخمین FP تعیین های بسیار انتخابی باقیمانده آفت کش ها در مواد غذایی توسط GC-MS/MS، از طریق مدل سازی علامت دستگاهی با استفاده از روش مونت کارلو را مورد بحث قرار می دهد. شبیه سازیهای روش مونت کارلو در صفحه گسترده اکسل انجام شد.

شناسایی آنالیت ها بر اساس زمان بازداری، t_R، در سیستم کروماتوگرافی و نسبت فراوانی، $AR=A_1/A_2$ دو یون مشخصه طیف جرمی آنالیت است. اگرچه t_R تقریباً به طور نرمال توزیع شده است، AR می تواند به طور قابل توجهی از حالت نرمال، منحرف شود. مشخص شده است نسبت متغیرهای همبسته، به طور نرمال توزیع نمی شوند، به ویژه اگر متغیر با دقت کمتر (یعنی پراکندگی بیشتر مقادیر) در مخرج باشد [۸].

توسعه و صحه گذاری روش، برای شناسایی سطوح ناچیز آنالیت ها در مواد غذایی توسط GC MS/MS با تعریف روش تجزیه کیفی آغاز می شود، از جمله آماده سازی نمونه و شرایط GC MS/MS. مشخصات شرایط GC MS/MS، شامل انتخاب یون های مشخصه طیف جرمی آنالیت است (یون های کلرپیریفوس متیل: 208 m/z و 271 m/z؛ یون های مالاتیون: 99 m/z و 127 m/z). پس از آن تزریق های تکراری محلول های مادر آنالیت و استخراج شده مواد غذایی، انجام می شود. تزریق محلول های مادر آنالیت، برای مطالعه پراکندگی t_R و فراوانی یون های A_1 و A_2 و همبستگی فراوانی های یون هر آنالیت، استفاده می شود. داده های عملکرد در غلظت های مختلف آنالیت، جمع آوری شد، زیرا مقدار و پراکندگی A_1 و A_2 با غلظت تغییر می کند. جدول م۱.۵ خلاصه ای از پارامتر های عملکردی t_R و A_1 و A_2 را نشان می دهد. تجزیه تکراری استخراج شده ها بدون سطوح آنالیت قابل تشخیص، برای تعریف مدل های پراکندگی نویز علامت در پنجره زمان بازداری استفاده شد (جدول م۱.۵). علامت های استخراج شده های نماینده محصولات غذایی الگوهای تغذیه ای مواد غذایی با منشأ گیاهی با محتوای آب بالا، مورد بررسی قرارگرفت.

از داده ها در جدول م۱.۵ ، مدل های تغییر پذیری t_R و AR ایجاد شد. مدل های t_R، از فاصله های اطمینان براساس توزیع t ساخته شدند ($\bar{t}_{Ri} \pm ts_{tRi}$ که $\bar{t}_{Ri}$ و ts_{tRi}، میانگین و انحراف استاندارد t_R هستند و t از مقدار t درتوزیع t برای سطح اطمینان مشخص و درجه آزادی $\bar{t}_{Ri}$ وS_{tRi} تعریف شده است. مدل های AR از شبیه سازی های روش مونت کارلو، ایجاد شدند. از پراکندگی مشاهده شده انواع A_1 و A_2 آنالیت، نسبت های همبسته A_1 و A_2 (یعنی AR) شبیه سازی شدند. از استخراج شده های شاهد، نویز علامت و در نتیجه AR در شاهد ها، شبیه سازی شد. نویز علامت، به عنوان یک توزیع نرمال کوتاه شده در صفر، مدل شد، زیرا پیکهای کروماتوگرافی، سطوح منفی ندارند. جدول م۲.۵، پراکندگی تخمین زده شده t_R و AR را نشان می دهد. این جدول همچنین فرمول های اکسل را که درشبیه سازی A_1 و A_2 استفاده شدند، ارائه می دهد. حدود اطمینان برای t_R وAR به ترتیب برای سطوح اطمینان ۹۹٫۹٪ یا ۹۸٪ تنظیم شدند که مربوط به TP می باشند.

FP از شناسایی های مبتنی بر t_R براساس تجربه تحلیل گر ۱۰٪ تعیین شد. این FP، نشان دهنده احتمال عدم تأیید پیکی از آنالیت است که در پنجره زمان بازداری تعریف شده برای آنالیت، مشاهده می شود. FP از AR از شبیه سازی های مونت کارلو نویز علامت و از تعیین چند بار نویز شبیه سازی شده AR در بازه پذیرش برای این پارامتر تخمین زده شد. از آنجایی که FP می تواند برای سطوح آنالیت کم، بسیار بزرگ باشد، در کسر های جرمی آنالیت مختلف با تعریف حداقل فراوانی هر یون، تعیین شد. جدول م۳.۵ برآورد TP وFP و ترکیب آنها را در LR(+) نشان می دهد. در ستون آخر جدول، عملکرد شناسایی ها بر اساس هم AR هم t_R گزارش می شود. جدول م۳.۵ همچنین احتمال پسین PP را نشان می دهد که یک قلم آزمون مثبت است با این فرض که نتایج مثبت یا منفی به یک اندازه، محتمل هستند یعنی P(+) = P(-) = 0.5.

جدول م۱.۵ پارامترهای عملکرد مربوط به شناسایی کلرپیریفوس متیل و مالاتیون دراستخراج شده های منشأ گیاهی توسطGCMS/MS. تمامی پارامترها با ۱۱ درجه آزادی برآورد شدند.

آنالیت: کلرپیریفوس متیل								
				فراوانی				
		زمان بازداری، t_R		یون: ۲۰۸ m/z		یون: ۲۷۱ m/z		
استخراج شده	w (mg kg^{-1})	$\bar{t}_{Ri}$ (min)	S_{tRi} (min)	$\bar{A}$ (a.u.)	S_A (a.u.)	$\bar{A}$ (a.u.)	S_A (a.u.)	ρ
E§	۳/۳۳	۱۷٫۲۴	۰٫۰۲۴	۱۰۵۶۶۸	۱۳/۳	۱۳۸۶۷۸	۶۰۱۴	۰٫۹۹۵۶
E§	۰٫۳۳	۱۷٫۲۴	۰٫۰۲۴	۱۰۱۶۳	۱۰/۵	۱۵۰۲۵	۸۰۱۰	۰٫۶۱۵۱
E§	۰٫۰۸۳	۱۷٫۲۴	۰٫۰۲۴	۴۳۶۶	۲۱/۴	۵۷۹۰	۱۵۰۱	۰٫۳۹۶۵
G	<۰٫۰۴	-	-	۳۷۲	۸۹۲	۳۷۲	۸۹۲	-
O	<۰٫۰۴	-	-	۳۷۲	۸۹۲	۳۷۲	۸۹۲	-

I	<۰/۰۴	-	-	۳۷۲	۸۹۲	۳۷۲	۸۹۲	-

آنالیت: مالاتیون

				فراوانی				
		زمان بازداری، t_R		یون: ۹۹ m/z		یون: ۱۲۷ m/z		
استخراج شده	w (mg kg^{-1})	$\bar{t}_{Ri}$ (min)	S_{tRi} (min)	$\bar{A}$ (a.u.)	S_A (a.u.)	$\bar{A}$ (a.u.)	S_A (a.u.)	ρ
E§	۳/۳۳	۱۹/۴۵	۰/۰۷۰	۲۲۶۵۹۲	۷/۸۵	۲۲۶۷۶۵	۱۰/۳	۰/۹۸۸
E§	۰/۳۳	۱۹/۴۵	۰/۰۷۰	۲۲۳۵۴	۱۷/۴	۲۲۹۶۹	۱۵/۶	۰/۹۶۷۲
E§	۰/۰۸۳	۱۹/۴۵	۰/۰۷۰	۵۸۸۲	۳۰/۷	۶۳۴۵	۲۸/۰	۰/۷۶۷۷
G	<۰/۱۱	-	-	۳۷۲	۸۹۲	۳۷۲	۸۹۲	-
O	<۰/۱۱	-	-	۳۷۲	۸۹۲	۳۷۲	۸۹۲	-
I	<۰/۱۱	-	-	۳۷۲	۸۹۲	۳۷۲	۸۹۲	-

§= پراکندگی فراوانی های یونی با ترکیب علامت های آنالیت در یک حلال خالص با علامت های استخراج شده های گیاهی برآورد شد. بافت استخراج شده : G – زنجبیل ، O – پیاز بهاری ، I – جلبک دریایی ، خزه ایرلندی ، E- بافت نامشخص، W =کسر جرمی آنالیت (میلی گرم بر کیلوگرم)؛ $\bar{t}_{Ri}$ – میانگین زمان بازداری (دقیقه) (این عامل می تواند با روز تزریق تغییر کند)، S_{tRi} - انحراف استاندارد زمان بازداری تخمین زده شده تحت شرایط تکرارپذیری (دقیقه)؛ $\bar{A}$ – میانگین فراوانی های یون (واحد های دلخواه، a.u.)؛ S_A – انحراف استاندارد فراوانی یون؛ ρ – ضریب همبستگی اسپیرمن

جدول م۲.۵ فواصل پذیرش برای زمان بازداری و نسبت فراوانی یون

آنالیت	استخراج شده	بازه کسر جرمی w (mg kg^{-1})	حداکثر اختلاف t_{Ri} (دقیقه) (*c.l.* 99.9%)§	بازه AR (*c.l.* 98%) †
کلرپیریفوس متیل	E	۰/۰۴-۳/۳۳	۰/۱۸	۰/۴۳۹-۱/۱۸
مالاتیون	E	۰/۱۱-۳/۳۳	۰/۵۴	۰/۴۶۷-۱/۵۴

E – بافت نامشخص ؛ *c.l.* سطح اطمینان

§ - حداکثر تفاوت بین زمان بازداری آنالیت در محلول استاندارد و در نمونه تجزیه شده ($\sqrt{2}\ t_{StRi}$) ؛

† - فرمول اکسل مورد استفاده در شبیه سازی = یون اول: $A_1 = \bar{A}_1 + S_{A_1} \times TINV\ (R_1, v_1)$ و یون دوم

$A_2 = \bar{A}_2 + S_{A_1} \times \text{TINV}\left(R_1 \text{ و } v_2\right) * \rho + \text{TINA}\left(R_2 \text{ و } V_2\right) \times (1-\rho^2)\text{^}0.5)$ که v_i درجه آزادی مربوطه به A_i و S_{A_i} و R_1 و R_2 دو مولد مقادیر تصادفی مستقل U(۰و۱) (فرمول اکسل RAND()).

بر اساس جدول م۳.۵، فقط زمانی که شناسایی ها در حد کمی سازی یا بالا تر ازحد کمی سازی (قابل اندازه گیری) انجام شود (به ترتیب ۰/۱۴ میلی گرم در کیلو گرم یا ۰/۳۸ میلی گرم در کیلو گرم برای کلرپیریفوس متیل و مالاتیون)، شناسایی ها توسط قسمت های شاهد «بسیار قوی» پشتیبانی می شوند (یعنی $۱۰^{۶} > LR(+) > ۱۰^{۵}$). همه ویژگی های عملکردی که در جدول م۳.۵ ارائه شده است (TP و FP، LR(+) و PP)، جایگزینهای معتبری برای گزارش عملکرد یا عدم قطعیت تجزیه کیفی در کسر های جرمی مختلف

آنالیت هستند. گزارش LR(+) دارای مزیت ترکیب TP و FP در یک معیار است و نیازی به فرض شیوع آفت کش درنمونه های تجزیه شده ندارد.

جدول م.۳.۵ ویژگی های عملکردی شناسایی کلرپیریفوس متیل و مالاتیون توسط GS MS /MS

ویژگی های عملکردی در سطوح مختلف آنالیت، w					
آنالیت		w (mg kg^{-1})	t_R	AR	t_R & AR
کلرپیریفوس متیل	TP (%)	≥۰/۰۴	۹۹/۹	۹۸	۹۷/۹
	FP (%)	۰/۰۴§	۱۰	۳۰/۲	۳/۰۲
		۰/۰۸	۱۰	۰/۲	۰/۰۲
		۰/۱۴§	۱۰	۰/۰۰۱	۰/۰۰۰۱
		۰/۲۷	۱۰	۰/۰۰۱	۰/۰۰۰۱
	LR (+)	۰/۰۴§	۹/۹۹	۳/۲۴	۴/۳۲
		۰/۰۸	۹/۹۹	۴۹۰	۴۸۹۵
		۰/۱۴§	۹/۹۹	۹۸۰۰۰	۹۷۹۰۲۰
		۰/۲۷	۹/۹۹	۹۸۰۰۰	۹۲۹۰۲۰
	PP (%)	۰/۰۴§	۹۰/۹	۷۶/۴	۹۷/۰
		۰/۰۸§	۹۰/۹	۹۷/۸	۹۹/۹۸
		۰/۱۴	۹۰/۹	۹۹/۹۹۹	۹۹/۹۹۹۹
		۰/۲۷	۹۰/۹	۹۹/۹۹۹	۹۹/۹۹۹۹
مالاتیون	TP (%)	≥ ۰/۱۱	۹۹/۹	۹۸	۹۷/۹
	FP (%)	۰/۱۱§	۱۰	۲۹/۸	۲/۹۸
		۰/۲۳	۱۰	۰/۰۰۱	۰/۰۰۰۱
		۰/۳۸	۱۰	۰/۰۰۱	۰/۰۰۰۱
		۰/۷۷	۱۰	۰/۰۰۱	۰/۰۰۰۱
					۰/۰۰۰۱
	LR (%)	۰/۱۱§	۹/۹۹	۳/۲۹	۳۲/۹
		۰/۲۳	۹/۹۹	۹۸۰۰۰	۹۷۹۰۲۰
		۰/۳۸	۹/۹۹	۹۸۰۰۰	۹۷۹۰۲۰
		۰/۷۷	۹/۹۹	۹۸۰۰۰	۹۷۹۰۲۰
	PP (%)	۰/۱۱§	۹۰/۹	۷۶/۶۸	۹۷/۰
		۰/۲۳	۹۰/۹	۹/۹۹۹	۹۹/۹۹۹
		۰/۳۸	۹۰/۹	۹/۹۹۹	۹۹/۹۹۹
		۰/۷۷	۹۰/۹	۹/۹۹۹	۹۹/۹۹۹

§ - حد تشخیص ؛ § حد کمی سازی ؛ t_R- زمان بازداری آنالیت؛ AR- نسبت فراوانی دو یون مشخصه از طیف جرمی آنالیت.

این مثال نشان می دهد که چگونه شبیه سازی مونت کارلو برای علامتها می تواند بر دشواری تعیین تجربی نرخ های مثبت کاذب شناسایی های بسیار گزینش پذیر، غلبه کند.

۶.۷ م۶ شناسایی RNA SARS-COV-2 با آزمایش تکثیر نوکلئیک اسید

دامنه کاری

نوع تجزیه کیفی: تجزیه بر اساس معیار های کمی
قلم/بافت: سوآب بینی - سوآب نازوفارنکس[۲۹] (بینی و حلق) و اوروفارنکس[۳۰] (دهان و حلق)
پارامتر/آنالیت: سارس - کووید - RNA ۲
نوع معیار دسته بندی: آستانه چرخه، Ct، مقادیر برابر با یا کمتر از Ct، به عنوان مثبت، دسته بندی می شوند، مقادیر بالاتر به عنوان منفی، دسته بندی می شوند.
فن/دستگاهوری: واکنش زنجیره ای پلی مراز رونویسی معکوس (RT-PCR)
نوع گزارش نتایج: نرخ مثبت واقعی و نرخ منفی واقعی

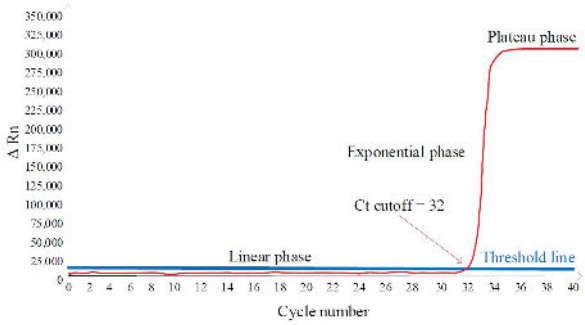
سندرم حاد تنفسی ویروس کرونا ۲ (SARS-COV-2) ویروسی است که باعث بیماری کووید ۱۹، بیماری تنفسی مسئول همه گیری کووید ۱۹ می شود. یکی آزمایش های غربالگری برای حضور این ویروس در سوآب های بینی، شامل واکنش زنجیره ای پلی مراز رونویسی معکوس (RT-PCR)، نوعی آزمایش تکثیر نوکلئیک اسید (NAAT) است. صحه گذاری «داخلی[۳۱]» این آزمون شامل تعیین برش[۳۲] آستانه سیکل، LOD، حساسیت (SS) و ویژگی (SP) بالینی (جدول ۲) با سایر پارامترهای عملکرد (که در این مثال ارائه نشده است) می باشد. در تجزیه بالینی، SS و SP به عنوان تخمین زننده صحت بالینی، شناخته می شوند. صحت بالینی توسط FN و FP و شیوع اپیدمیولوژیک، انواع و زیر انواع[۳۳] عوامل جهش ها و سایر عوامل بیولوژیکی محدود می شود.

۱.۶.۷ آستانه چرخه

تعداد چرخه های مورد نیاز برای قابل تشخیص شدن یک آمپلیکون (تکثیر تعداد نسخه[۳۴] های ژن) در بالای زمینه، به عنوان آستانه چرخه (Ct) تعریف می گردد [۵۹] - تعدا چرخه های مورد نیاز برای تکثیر RAN ویروسی، برای رسیدن به سطح قابل تشخیص. برای درک کاربرد برش چرخه (Ct- Cut Off)، شناخت برخی از متغیرها توصیه می شود. Rn (مقدار گزارشگر نرمال شده)، بزرگی سیگنال تولید شده، توسط مجموعه ای مشخص از شرایط PCR می باشد. ΔRn (شکل م۱.۶)، مقدار گزارشگر نرمال شده منهای پاسخ خط پایه می باشد. آستانه، سطحی از علامت است که افزایش آماری قابل توجهی را نسبت به علامت پایه محاسبه شده، منعکس می کند (شکل م۱.۶ را ببینید). این خط تصمیم، برای تشخیص علامت های تکثیر مربوطه، از زمینه

[29] Nasopharyngeal
[30] Oropharyngeal
[31] In-house
[32] Cut-off
[33] Subtype
[34] Copy

ایجاد می شود. در مثال، نرم افزار، آستانه را ۱۰ برابر انحراف استاندارد مقدار فلورسانس پایه، تنظیم می کند. حد[۳۵]، درناحیه مربوط به رشد نمایی محصول PCR تعریف می شود. شکل م۱.۶ دسته بندی مثبت (کمتر یا مساوی برش آستانه چرخه ۳۲) یک نمونه انسانی را نشان می دهد. نتایج فلورانس بالاتر از ۳۲، به عنوان منفی دسته بندی می شوند. توجه داشته باشید که تنها منحنی های تکثیر S شکل، نشان دهنده تکثیر واقعی هستند (شکل م۱.۶ ببینید).



شکل م۱.۶ تشخیص ار ان ای SARS – COV – 2 از یک نمونه مثبت واقعی

۲.۶.۷ تخیمن LOD

توانایی تشخیص نزدیک به برش آستانه چرخه، با استفاده از LOD ارزیابی می شود. LOD به عنوان غلظت چندگانه (به عنوان مثال بیان تعداد کپی در میلی لیتر) مرتبط با یک TP ٪۹۵ (LOD 95%) تعریف می شود. LOD 95% به وسیله مدل سازی تغییرات TP باغلظت و تخمین غلظت، برآورد می شود، که درآنTP با استفاده از رگرسیون پروبیت ۹۵٪ می باشد. [۶۳ و ۶۰]. در مثال، LOD 95% دریک سری هفت رقتی از یک نمونه با غلظت ۵۰۰ نسخه در میلی لیتر، ۱۱۴ نسخه در میلی لیتر می باشد. جزئیات بیشتر درمورد روش ارائه شده برای تعیین LOD 95%، در فهرست (۵.۵ از [۶۳]) وجود دارد.

۳.۶.۷ صحت بالینی

ارزیابی صحت بالینی، شامل تعیین مقادیر هدف یا حداقل، برای حدود پایین تر از فاصله اطمینان ۹۵٪ از SS وSP و $LL_{sp.95}^{tg}$ و $LL_{ss.95}^{tg}$ و بررسی اینکه آیا حدود پایین تخمین زده شده، مساوی با یا بیشتر از مقدار هدف مربوطه است، می باشد (یعنی اینکه اگر $LL_{ss.95} \geq L_{ss.95}^{tg}$ و $LL_{sp.95} \geq LL_{sp.95}^{tg}$). برای این هدف آزمون $LL_{ss.95}^{tg} = 95\%$ و $LL_{sp.95}^{tg} = 90\%$، یعنی توصیه می شود حد پایین حساسیت، بیشتر از ۹۵٪ و حد پایین ویژه بودن، بیشتر از ۹۰٪ باشد.

برای صحه گذاری روش، ۲۰۰ نمونه سوآب بینی، نمونه های سوآب حلق و بینی و دهان و بینی، مورد تجزیه قرار گرفت. مشخص شده که ۱۰۰ نفر از افراد آلوده به

[35] Limit

SARS-COV-2 و تأیید شده ۱۰۰ نفر از افراد، به این ویروس آلوده نیستند. جدول م۱.۶ جدول احتمال به دست آمده از ۲۰۰ آزمون را نشان می دهد. این جدول نشان می دهد هیچ منفی کاذب و ۳ مثبت کاذب، گزارش شده است، بنابراین SS وSP بالینی، به ترتیب، ۱۰۰٪ و ۹۷٪ می باشند. از معادله (۸) تا (۱۱) حدود فواصل اطمینان ۹۵٪ برای SS و SP به ترتیب (۹۶/۳، ۱۰۰) و (۹۱/۶ و ۹۹/۰) هستند (جدول م۲.۶). از آنجایی که $LL_{ss.95}$ و $LL_{sp.95}$، به ترتیب، بالاتر از ۹۵٪ و ۹۰٪ هستند، روش تجزیه، معتبر تلقی می شود. SS به وسیله حساسیت تبدیل سری، تکمیل می شود [۶۴].

جدول م۱.۶ جدول احتمالی که عملکرد روش تشخیص SARS COV-2 RAN را در نمونه های سوآب بینی، نازوفارنکس و سوآب اورو فارنکس خلاصه می کند.

	مورد		
کل نتایج	منفی(nc)	مثبت(*pc*)	
$p = ۱۰۳$	$fp = ۳$	$tp = ۱۰۰$	مثبت(p)
$n = ۹۷$	$tn = ۹۷$	$fn = ۰$	منفی(n)
	$nc = ۱۰۰$	$pc = ۱۰۰$	کل موارد

جدول م۲.۶ صحت بالینی روش تشخیص SARS-COV-2 RNA در سرم یا پلاسمای انسان

حساسیت بالینی		
$HL_{ss.95}$ = ٪۱۰۰	$LL_{ss.95}$ = ٪۹۶/۳	SS = ٪۱۰۰
ویژگی بالینی		
$HL_{sp.95}$ = ٪۹۹/۰	$LL_{sp.95}$ = ٪۹۱/۶	SP = ٪۹۷

پیوست الف) تئوری بیز، احتمال نسبی و نسبت عدد احتمال

الف.۱ تئوری بیز

تئوری بیز[۳۶] توضیح می دهد که چگونه احتمال یک رویداد A (مثلاً مثبت بودن یک قلم آزمون) با اطلاعات جدید E، مانند نتیجه آزمایش مثبت، تغییر می کند. تئوری بیز، اغلب برای دو رویداد نوشته می شود، A و E به طوری که:

$$P(A|\mathrm{E}) = \frac{\mathrm{P(E|A)P(A)}}{P(E)} \qquad \text{الف.۱}$$

دراینجا، P(A) و P(E)، احتمال وقوع رویدادهای A و E به تنهایی هستند، P(A|E)، احتمال وقوع رویداد A می باشد، وقتی که رویداد E اتفاق افتاده باشد و P(E|A)، احتمال وقوع رویداد E هست، وقتی که رویداد A اتفاق افتاده باشد. در آمار، P(A|E) و P(E|A) معمولاً به عنوان احتمالات مشروط شناخته می شوند، مثلاً P(E|A) می تواند اشاره داشته باشد به احتمال مشروط رویداد E، به شرطی که A اتفاق افتاده باشد.

در زمینه تجزیه کیفی، با در نظر گرفتن مورد مثبت به عنوان مثال، P(A) را می توان به عنوان احتمال این که یک قلم آزمون انتخاب شده به طور تصادفی، قبل از انجام هر آزمونی واقعاً مثبت است، درک کرد. P(E|A) احتمال این است که یک قلم آزمون، واقعاً یک نتیجه آزمایش مثبت ایجاد کند - نرخ مثبت واقعی TP در جدول ۲. P(E)، احتمال یک نتیجه آزمون مثبت، صرف نظر از وضعیت قلم آزمون است. در نهایت، P(A|E)، عبارت است از احتمال اینکه قلم آزمون، واقعاً مثبت است، بعد از در نظر گرفتن اطلاعات اضافه شده توسط نتیجه آزمون مثبت. زیرا پس از در دسترس قرار گرفتن شواهد E، محاسبه می شود، P(A|E) معمولاً «احتمال پسین» برای A نامیده می شود. یک احتمال پسین تخمینی، نشان دهنده مستقیم اطمینانی است که می تواند در یک دسته بندی، قرار داده شود.

یادآوری این نکته مهم است که (در ادامه حالت مثبت)، P(E) هم نتایج مثبت واقعی و هم نتایج مثبت کاذب را شامل می شود و همچنین، P(E) برای جمعیت کامل اقلام آزمون، اعمال می شود. این بدان معنی است که P(E) هم به نرخ مثبت واقعی و کاذب و هم به نسبت اقلام آزمون واقعاً مثبت و منفی، حساس است. از نظر کمی، برای دو مورد A و ¬A (که نشان دهنده غیر A، یک قلم آزمون واقعاً منفی است)، P(E) را می توان به صورت جمع وزنی نوشت.

$$\mathrm{P(E) = P(A)\ P(E|A) + P(\neg A)P(E|\neg A)} \qquad \text{الف.۲}$$

[36] Bayes' theorem

با در نظر گرفتن E به عنوان یک نتیجه آزمایش مثبت، معادله الف.۲ می گوید که احتمال مرکب E، نرخ مثبت واقعی ضربدر نسبت نمونه های مثبت واقعی، به علاوه نرخ مثبت کاذب ضربدر نسبت نمونه های منفی واقعی است. به همین دلیل است که نرخ مثبت کاذب بالا، اعتماد به نتایج آزمایش مثبت را کاهش می دهد. با ارجاع به معادله الف.۱، اگر احتمال بالای نتایج مثبت از اقلام آزمایشی واقعاً منفی، وجود داشته باشد، P(A|E) کاهش می یابد زیرا P(E) افزایش می یابد. این، با شهود مطابقت دارد، هر چقدر هم نرخ مثبت واقعی بالا باشد، توصیه می شود شانس تعداد زیادی از مثبتهای کاذب، اطمینان کمتری به ما می دهد که یک نتیجه مثبت، نشان دهنده یک قلم آزمون واقعاً مثبت است.

الف.۲ احتمال و احتمال نسبی

یک احتمال P معمولاً به صورت عددی بین ۰ و ۱ بیان می شود. با این حال می توان آن را به شکل «احتمال نسبی»[۳۷] نیز بیان کرد، اصطلاحی که شاید در شرط بندی ورزشی بسیار آشنا باشد. اگر احتمال یک رویداد A، P(A) باشد و احتمال متناوب، به سادگی غیر A باشد، احتمال نسبی O(A) به نفع A را می توان با استفاده از معادله زیر محاسبه کرد:

$$O(A)=\frac{P(A)}{1-P(A)} \qquad \text{الف.۳}$$

بر خلاف احتمالات، احتمال نسبی می تواند هر مقدار غیر منفی را بگیرد. احتمال نسبی ۱۰^۶ یا یک میلیون به یک، امکانپذیر است. با مرتب کردن مجدد الف.۳ می توان احتمالات نسبی را با نوآرایی معادله الف.۳ به احتمالات تبدیل کرد:

$$P(A)=\frac{O(A)}{O(A)+1} \qquad \text{الف.۴}$$

الف.۳ شکل احتمالات نسبی تئوری بیز و نسبت عدد احتمال

اگر فقط دو فرضیه جایگزین و مکمل وجود داشته باشد، A و ¬A (یعنی غیر A) و برخی شواهد E (مثل یک نتیجه آزمون که برای A مثبت است) برای به روز رسانی احتمالات هر کدام، استفاده می شود، تئوری بیز، احتمالات پسین را به صورت زیر می دهد:

$$P(A|E)=\frac{P(E|A)\,P(A)}{P(E)} \qquad \text{الف.۱.۵}$$

[۳۷] معمولاً احتمالات نسبی به صورت جمع در نظر گرفته می شود.

$$P(\neg A|E) = \frac{P(E|\neg A)P(\neg A)}{P(E)}$$ الف.۲.۵

نسبت احتمالات آنها برابر است با:

$$\frac{P(A|E)}{P(\neg A|E)} = \frac{P(E|A)P(A)}{P(E|\neg A)P(\neg A)}$$ الف.۶

یا تفکیک اصطلاحات برای وضوح:

$$\frac{P(A|E)}{P(\neg A|E)} = \frac{P(E|A)}{P(E|\neg A)} \times \frac{P(A)}{P(\neg A)}$$

از آنجایی که فقط دو فرضیه وجود دارد، A و ¬A، مجموع احتمالات پیشین و احتمالات پسین باید برابر با یک باشد. به این معنا که $P(\neg A) = 1 - P(A)$ و $P\big((\neg A|E)\big) = 1 - P(A|E)$. یعنی سمت چپ الف.۶ برابر است با $P(A|E)/(1 - P(A|E))$. در مقایسه با الف.۳، این هست فقط احتمالات نسبی به نفع A به شرط E یا O(A|E)، احتمالات نسبی پسین، به نفع فرضیه A. به صورت مشابه، احتمالات نسبی پیشین O(A) در سمت راست الف.۶ ظاهر می شود به طوری که $\frac{P(A)}{P(\neg A)} = P(A)/(1 - P(A)) = \mathrm{O(A)}$ نسبت باقیمانده P(E|A)/P(E|¬A)، به عنوان نسبت عدد احتمال شناخته می شود. برای مورد تجزیه کیفی، جایی که ¬A مربوط به یک قلم آزمون واقعاً منفی است، جدول ۵ نسبت عدد احتمال (تخمینی) را به صورت TP/FP نشان می دهد. بنابراین شکل احتمالات نسبی تئوری بیز را می توان به صورت زیر نوشت:

$$O(A|E) = O(A) \times \frac{P(E|A)}{P(E|\neg A)}$$ الف.۷

یا به صورت شماتیک:

احتمالات نسبی پسین= احتمالات نسبی پیشین × نسبت عدد احتمال

بنابراین نسبت عدد احتمال را می توان به صورت کمی، به عنوان تغییر در احتمالات نسبی به نفع یک فرضیه خاص، تفسیر کرد.

پیوست ب) تجزیه کیفی مرتبط با ارزیابی انطباق با یک حد کمی

ب.۱: ارزیابی انطباق به عنوان یک تجزیه کیفی

ارزیابی انطباق مقدار یک پارامتر کمی از قلم مورد تجزیه با یک مقدار یا بازه حدی را می توان یک تجزیه کیفی، با استفاده از یک معیار کمی واحد (بخش۲)، با نتایج «منطبق» یا «غیرمنطبق» در نظر گرفته می شود. جدول ب.۱ نمونه هایی از این نوع تجزیه ها را ارائه می دهد.

جدول ب.۱ نمونه هایی از تجزیه کیفی بر اساس ارزیابی انطباق مقدار یک پارامتر کمی از قلم مورد تجزیه با مقدار یا بازه حدی.

۱) ارزیابی رنگ ماده خام با مقایسه اندازه گیری جذب با آستانه
۲) ارزیابی انطباق یک آلیاژ با حداقل محتوا برای جزء اصلی آن
۳) ارزیابی انطباق دارو با بازه مشخصات برای غلظت ماده فعال
۴) ارزیابی انطباق باقیمانده آفت کش در میوه با توجه به حداکثر سطح باقیمانده
۵) ارزیابی وضعیت سلامتی یک فرد با مقایسه یک جزء خون اندازه گیری شده با یک بازه از مقادیر افراد سالم

استفاده از قواعد تصمیم و عدم قطعیت اندازه گیری در ارزیابی انطباق به طور مفصل در راهنمای Eurachem/CITAC در مورد «استفاده از اطلاعات عدم قطعیت در ارزیابی انطباق» مورد بحث قرار گرفته است [۲۹] (راهنمای انطباق). با این حال، برای کامل بودن، این ضمیمه چگونگی استفاده از اطلاعات عدم قطعیت یا عملکرد برای تجزیه کمی را برای استخراج برخی از معیارهای جدول ۲ مورد بحث قرار می دهد. سپس می توان از اینها برای توصیف عملکرد روش های تجزیه کیفی مبتنی بر مقایسه نتایج اندازه گیری با یک حد یا مشخصات، به طور کامل یا جزئی استفاده کرد.

هنگامی که تجزیه شامل ارزیابی این است که آیا یک ویژگی اندازه گیری شده بالاتر، پایین تر یا در حد یا بازه مشخصات است، عدم قطعیت اندازه گیری می تواند برای کمی کردن قابلیت اطمینان ارزیابی انطباق، استفاده شود.

یادآوری: راهنمای حاضر در مورد چگونگی ارزشیابی عدم قطعیت اندازه گیری بحث نمی کند. ارزشیابی عدم قطعیت اندازه گیری به طور مفصل در راهنمای Eurachem/CITAC، کمی کردن عدم قطعیت در اندازه گیری تجزیه ای [۶۵] توضیح داده شده است.

استفاده از عدم قطعیت اندازه گیری برای تصمیمهای انطباق شرح داده شده در راهنمای انطباق Eurachem/CITAC [۲۹]، شامل تنظیم معیاری برای تصمیم گیری در مورد انطباق یا عدم انطباق یک قلم آزمون با حداکثر احتمال تصمیمات انطباق کاذب $x\%$ است. راهنمای انطباق، ریسک های «خاص» و «عمومی»

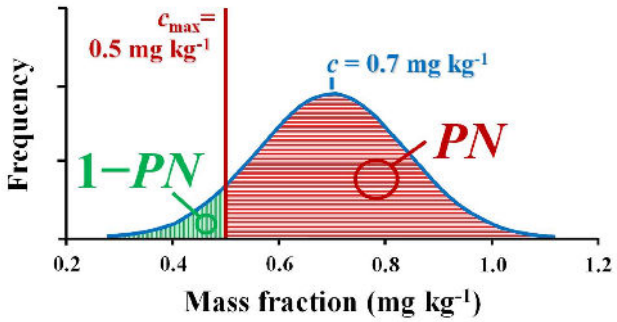
را متمایز می کند. «ریسک خاص»، احتمال تصمیم کاذب در مورد انطباق یک قلم خاص را کمی سازی می کند. این تنها بر اساس توزیع مرتبط با نتیجه اندازه گیری برای آن قلم است. در مقابل، «ریسک عمومی» احتمال تصمیمهای کاذب در مورد انطباق یک قلم آزمون را که در آینده به طور تصادفی انتخاب می شود، کمی می کند [۶۶]. ریسک عمومی، توزیع مقادیر احتمال اقلام اندازه گیری شده را مانند توزیع مقادیر اقلام از یک خط تولید یا یک منطقه محیطی را در نظر می گیرد. به عنوان مثال، محاسبه ریسک تولید کننده عمومی، مستلزم احتمال این است که یک خط تولید، محصولاتی با مقداری نزدیک به مقدار حد، تولید کند، به طوری که می توان آن را به اشتباه به عنوان نامنطبق در نظر گرفت. بنابراین برای تعیین ریسک عمومی، توزیع مقادیر برای جمعیت اقلام باید به خوبی مشخص شود.

اغلب، تحلیلگران، علاقمند به ارزیابی انطباق یک قلم خاص تجزیه شده هستند. در چنین مواردی، چگونه می توان معیارهای مورد استفاده برای تعیین قابلیت اطمینان انواع دیگر تجزیه های کیفی را تعیین کرد؟ یک مطالعه موردی و فرمول های مورد استفاده برای تعیین این پارامترها، در زیر ارائه شده است.

ب.۲ نتایج ارزیابی انطباق مثبت و منفی

اگر قلم از یک «نتیجه مثبت منطبق درنظر گرفته شود، توزیعی که عدم قطعیت اندازه گیری را توصیف می کند، می تواند برای ارائه یک نسبت عدد احتمال LR(+)، به نفع انطباق یا تحت برخی شرایط احتمال پسین، PP از موردی که واقعاً مطابقت دارد (یک مورد مثبت) و احتمال قلمی که، در واقع، نامنطبق می باشد (یک مورد منفی) (1- PP) استفاده شود. به همین ترتیب، اگر یک نتیجه با مقایسه با یک حد یا بازه به عنوان «نامنطبق» یا «منفی» گزارش شود می توان از عدم قطعیت اندازه گیری برای به دست آوردن نسبت عدد احتمال مربوطه LR(-) یا احتمالات پسین استفاده کرد PN و (1-PN) مثال زیر احتمالات پسین را در نظر می گیرد.

یادآوری: تعریف «نتایج مثبت» و «نتایج منفی» به ترتیب به عنوان «منطبق» و «نامنطبق» اختیاری است. می توان از قرارداد مخالف پیروی کرد.



شکل ب.۱ نمایش گرافیکی احتمال PN و اینکه یک مورد تجزیه وتحلیل شده نامنطبق با یک حد حداکثر، c_{MAX}، ازیک مقدار اندازه گیری شده،C ، باعدم قطعیت استاندارد مرتبط، $u(c) = 0.14$ mg kg^{-1} و احتمال مربوطه 1-PN که مطابقت ندارد.

ب.۳ مثال - ارزیابی انطباق برای باقی مانده آفت کش در میوه

فرض کنید انطباق یک نمونه از انگور در برابر حداکثر سطح باقی مانده c_{max} ، ۰/۵ میلی گرم درکیلوگرم برای استامیپراید ارزیابی می شود [۶۷] و کسر جرمی اندازه گیری شده در نمونه ۰/۷۰ میلی گرم درکیلوگرم است[۳۸]، c.

نتیجه اندازه گیری دارای توزیع نرمال با عدم قطعیت استاندارد u(C)، ۰/۱۴ میلی گرم برکیلوگرم با تعداد درجات آزادی بسیار بالا می باشد. چون $c > c_{max}$، محتمل ترین نتیجه درمورد انطباق، عدم انطباق انگور است (نتیجه منفی).

به طور رسمی یک احتمال پسین مانند PP یا PN مستلزم یک احتمال قبلی است. دراین مورد هیچ اطلاعاتی درمورد توزیع عمومی استامیپراید درانگور دردسترس نیست.

با این حال، درموارد مشکوک به آلودگی، اغلب منطقی است که فرض کنیم توزیع آنقدر گسترده است که اساساً در ناحیه نتیجه اندازه گیری اطلاعاتی نداشته باشد. درجایی که چنین است، عدم قطعیت اندازه گیری را می توان به عنوان تقریبی برای توزیع پسین درنظر گرفت. با در نظر گرفتن این رویکرد در اینجا، احتمال پسین یک نمونه آزمایش منفی (نامنطبق) PN، مساحت زیر تابع چگالی احتمال اندازه گیری در سمت راست c_{max}، درشکل ب.۱ به رنگ قرمز نشان داده شده است.

مساحت احتمال دم بالایی یک توزیع نرمال با میانگین c وانحراف معیار برابر با u(c) است. این را می توان از یک صفحه گسترده یا بسته آماری محاسبه کرد. برای مثال در مایکروسافت اکسل فرمول مورد نیاز:

$$1 - \text{NORM.DIST}\,(c_{max} = 0.5\,, \text{c} = 0.7, \text{u(c)} = 0.14, \text{TRUE})$$

[۳۸] در این راهنما برای موارد قابل کاربرد برای انواع مختلف کمیتها نظیر غلظتهای جرمی، کسر جرمی و pH، از نماد c برای غلظت استفاده می شود.

(جدول ب.۲ را ببینید) برای این مثال، ناحیه ۰٫۹۲۳ یا ۹۲٫۳٪ می باشد. احتمال متناظر PP که نمونه مثبت (منطبق) (1-PN) یا ۷٫۶٪ می باشد.

درجایی که اطلاعات قبلی آموزنده در دسترس است و استفاده از آن مناسب است، محاسبات شامل انتگرال گیری در توزیع قبلی است. انتگرال ها برای عدم قطعیت قبلی که دارای توزیع نرمال می باشد و عدم قطعیت اندازه گیری برای مثال JCGM106 [۶۶] همراه با راهنمایی در مورد توزیع های دیگر داده می شود.

یادآوری: ایجاد یک توزیع قبلی واقعاً غیر اطلاعاتی می تواند به طرز شگفت آوری سخت باشد. به عنوان مثال، یک توزیع یکنواخت ساده در محدوده شکل ب.۱ (۰/۲ - ۱/۲ $mg\ kg^{-1}$) با احتمال ۳۰٪ پیشین مطابقت دارد که یک قلم آزمون با حد مطابقت دارد، زیرا فقط ۳۰٪ از آن زیر حد ۰/۵ میلی گرم در کیلوگرم است. بنابراین، یک تحلیل بیزی کامل، شامل یک بررسی می‌شود که معمولاً شامل انتخاب‌های جایگزین توزیع قبلی می‌شود، تا مطمئن شود که نتیجه‌گیری به‌طور غیرمنطقی نسبت به توزیع قبلی مفروض حساس نیست.

ب.۴ فرمول های صفحه گسترده برای احتمالات ارزیابی انطباق

جدول ب.۲ فرمول های اکسل را ارائه می کند که توصیه می شود در زمانی که حدود یا فواصل انطباق متفاوت در نظر گرفته می شوند و مقدار اندازه گیری شده پایین - بالا - داخل یا خارج از حدود است، استفاده شود. اگر عدم قطعیت استاندارد ، u(c)، با تعداد کمی درجات آزادی، υ، تخمین زده شود به جای توصیف پراکندگی تخمین اندازه‌ده با توزیع نرمال، توصیه می شود از توزیع t-Student درنظر گرفته شود. درآن صورت توصیه می شود فرمول کلی استفاده شده درجدول B.2 NORM.DIST (C, c, u(c), TRUE)، با TDIST(ABS(C-c))/ u(C), υ , TRUE جایگزین شود.

جدول ب.۲ فرمول های اکسل که برای تخمین احتمال انطباق استفاده می شود، PP، یا عدم انطباق، PN، تصمیم گیری روی قلم خاص مورد تجزیه از فرمول ها می توان برای محاسبه LR(+)=PP/(1-PP) و LR(-) = PN /(1-PN) استفاده کرد.

S	**محدوده**	**انطباق قلم** (نوع نتیجه) سناریو	**قابلیت اطمینان انطباق**	فرمول MS-Excel (براساس توزیع نرمال تجمعی)
۱	حداکثر	**انطباق** (مثبت) $c \leq c_{max}$	PP	NORM. DIST(C_{max}, c, u(c), TRUE)
۲	حداکثر	**نامنطبق** (منفی) $c > c_{max}$	PN	1-NORM.DIST (c_{max}, c, u(c), TRUE)
۳	حداقل	**انطباق** (مثبت) $c \geq c_{min}$	PP	1-NORM. DIST (c_{min}, c, u(c), TRUE)
۴	حداقل	(منفی) نامنطبق $c < c_{min}$	PN	NORM. DIST (c_{min}, c, u(c), TRUE)
۵	بازه	(مثبت) انطباق	PP	NORM.DIST (c_{max}, c, u(c), TRUE)

		$c_{min} \leq c \leq c_{\max}$		NORM.DIST (c_{min}, c, u(c), TRUE)
۶	بازه	(منفی) نامنطبق $c > c_{max}$ یا $c < c_{min}$	PN	1-NORM.DIST (c_{max}, c, u(c), TRUE) + NORM.DIST (c_{min}, c, u(c), TRUE)
S - سناریو؛ بازه: حداکثر یا حداقل - بازه ، حد حداقل یا حداکثر ، نتایج مثبت یا منفی، یک نتیجه منطبق یا نامنطبق ؛ C و$u(c)$ – غلظت اندازه گیری شده و عدم قطعیت استاندارد مربوطه، c_{max} یا c_{min} حداکثر یا حداقل غلظت قابل قبول.				

Bibliography

[1] JCGM, International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (3rd edn.) (JCGM 200:2012), Sevres: BIPM, 2012.

[2] W. G. D. Ruig, G. Dijkstra and R. W. Stephany, "Chemometric criteria for assessing the certainty of qualitative analytical methods," Anal. Chim. Acta, pp. 277-282, 1989.

[3] B. L. Milman and L. A. Konopelko, "Identification of Chemical Substances by Testing and Screening of Hypotheses. I. General," Fresenius. J. Anal. Chem., vol. 367, pp. 621-628, 2000.

[4] S. L. R. Ellison and S. Gregory, "Quantifying uncertainty in qualitative analysis," Analyst, vol. 123, pp. 1155-1161,1998.

[5] S. L. R. Ellison and S. L. Gregory, "Predicting chance infrared spectroscopic matching frequencies", Anal. Chim. Acta, vol. 370, pp. 181-190, 1998.

[6] S. L. R. Ellison, "Uncertainties in qualitative testing and analysis," Accred. Qual. Assur., vol. 5, pp. 346–348, 2000.

[7] A. Ríos, D. Barceló, L. Buydens, S. Cárdenas, K. Heydorn, B. Karlberg, K. Klemm, B. Lendl, B. Milman, B. Neidhart, R. W. Stephany, A. Townshend, A. Zschunke and M. Valcárcel, "Quality assurance of qualitative analysis in the framework of the European project 'MEQUALAN'," Accred. Qual. Assur., vol. 8, pp. 68-77, 2003.

[8] R. B. Silva, "Evaluation of trace analyte identification in complex matrices by low-resolution gas chromatography - mass spectrometry through signal simulation," Talanta, vol. 150, pp. 553-567, 2016.

[9] J. Narciso, C. Luz and R. B. d. Silva, "Assessment of the Quality of Doping Substances Identification in Urine by GC/MS/MS," Anal. Chem., vol. 91, no. 10, pp. 6638-6644, 2019.

[10] ISO, Reference materials - Examples of reference materials for qualitative properties (ISO/TR 79:2015), Geneva: ISO, 2015.

[11] P. Pereira, B. Magnusson, E. Theodorsson, J. O. Westgard and P. Encarnação, "Measurement uncertainty as a tool for evaluating the 'grey zone' to reduce the false negatives in immunochemical screening of blood donors for infectious diseases," Accred. Qual. Assur., vol. 21, pp. 25-32, 2016.

[12] ILAC, ILAC Guidelines for Measurement Uncertainty in Testing (ILAC G17:01), Silverwater: ILAC, 2021.

[13] ISO, IEC, General requirements for the competence of testing and calibration laboratories (ISO/IEC 17025 ,(Geneva: ISO, 2017.

[14] ISO, Medical laboratories – Requirements for quality and competence (ISO 15189), Geneva: ISO, 2012.

[15] R. Bramley, A. Brown, S. Ellison, W. Hardcastle and A. Martin, "Qualitative analysis: A guide to best practice - forensic science extension," Sci. Justice, vol. 3, no. 40, pp. 163-170, 2000.

[16] L. Wide and C. A. Gemzell, "An immunological pregnancy test," Acta Endocrinol. (Copenh.), no. 35, pp. 261-267, 1960.

[17] U. Forsum, H. O. Hallander, A. Kallner and D. Karlsson, "The impact of qualitative analysis in laboratory medicine," Trends Anal. Chem., vol. 6, no. 24, pp. 546-555, 2005.

[18] P. Pereira, Quality control of qualitative tests for medical laboratories, Lisbon: Author-edition, 2019.

[19] G. Nordin, R. Dybkaer, U. Forsum, X. Fuentes-Arderiu and F. Pontet, "Vocabulary on nominal property, examination, and related concepts for clinical laboratory sciences (IFCC-IUPAC Recommendations 2017 ",(Pure Appl. Chem., vol. 5, no. 90, pp. 913-935, 2018.

[20] EU, Commission decision implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (2002/657/EC), EU, 2002.

[21] EU, Commission Regulation No 2017/644 of 5 April 2017 laying down methods of sampling and analysis for the control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation (EU) No 589/2014, EU, 2017.

[22] SANTE, Guidance document on analytical quality control and method validation procedures for pesticide residues and analysis in food and feed (SANTE/12682/2019), DG SANTE, 2019.

[23] WADA, Technical Document – TD2010IDCR, Identification Criteria for qualitative assays incorporating column chromatography and mass spectrometry, WADA, 2010.

[24] JCGM, Evaluation of measurement data – Guide to the expression of uncertainty in measurement (JCGM 100:2008), Sèvres: BIPM, 2008.

[25] ISO/IEC, Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)(ISO/IEC Guide 98-3), Geneva: ISO, 2008.

[26] ISO, IEC, General requirements for the competence of testing and calibration laboratories (ISO/IEC 17025)(superseded), Geneva: ISO, 1999.

[27] Clinical Laboratory and Standards Institute, EP12-A2 User protocol for evaluation of qualitative test performance, 2nd ed., Wayne (PA): CLSI, 2008.

[28] N. R. Campbell, Physics, the elements, Cambridge: Cambridge University Press, 1920.

[29] A. Williams and B. Magnusson, (Eds.), Eurachem/CITAC Guide: Use of uncertainty information in compliance assessment, 2nd ed., Eurachem, 2021.

[30] B. Magnusson and U. Örnemark, (Eds.), Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics (2nd Edn.), Eurachem, 2014.

[31] A. Agresti, Categorical Data Analysis (3rd Ed.), New Jersey: Wiley, 2012.

[32] AOAC International, Official Methods of Analysis of AOAC International (20th Edn.), Maryland: AOAC International, 2016.

[33] S. D. Ferrara, L. Tedeschi, G. Frison, G. Brusini and F. Castagna, "Drugs-of-Abuse Testing in Urine: Statistical Approach and Experimental Comparison of Immunochemical and Chromatographic Techniques," J. Anal. Toxicol., vol. 18, no. 5, pp. 278-291, 1994.

[34] R. J. Freund and W. J. Wilson, Regression Analysis, San Diego, CA: Academic Press, 1998.

[35] J. Fox, An R and S-Plus companion to applied regression, Thousand Oaks, CA.: Sage Publications Inc,. 2002.

[36] S. L. R. Ellison and T. Fearn, "Characterising the performance of qualitative analytical methods: Statistics and terminology," Trends Anal. Chem. , vol. 24, pp. 468-476, 2005.

[37] S. L. R. Ellison, C. A. English, M. J. Burns and J. T. Keer, "Routes to improving the reliability of low level DNA analysis using real-time PCR," BMC Biotechnology, vol. 6, pp. 33 (1-11), 2006.

[38] I. Kuselman and F. Pennecchi, “IUPAC/CITAC Guide: Classification, modelling and quantification of human errors in chemical analytical laboratory (IUPAC Technical Report),” Pure Appl. Chem., vol. 88, pp. 477-515, 2016.

[39] EU, Commission Regulation No 589/2014 of 2 June 2014 laying down methods of sampling and analysis for the control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation No 252/2012, EU, 2014.

[40] EU, Commission Regulation No 152/2009 of 27 January 2009 laying down the methods of sampling and analysis for the official control of feed, EU, 2009.

[41] T. Wenzl, J. Haedrich, Schaechtele, Alexander, P. Robouch and J. Stroka, Guidance Document on the Estimation of LOD and LOQ for Measurements in the Field of Contaminants in Feed and Food, JRC, 2016.

[42] J. Vessman, R. I. Stefan, J. F. Van Staden, K. Danzer, W. Lindner, D. T. Burns, A. Fajgelj and H. Muller, “Selectivity in Analytical Chemistry (IUPAC Recomendation 2001),” Pure Appl. Chem., vol. 73, no. 8, p. 1381–1386, 2001.

[43] ISO, In vitro diagnostic medical devices – Information supplied by the manufacturer (labelling) – Part 1: Terms, definitions and general requirements (ISO 18113-1), Geneva: ISO, 2009.

[44] European Network of Forensic Science Institutes, ENFSI Guideline for evaluative reporting in forensic science, ENFSI, 2015.

[45] V. Morgado, C. Palma and R. J. N. Bettencourt da Silva, “Microplastics identification by Infrared spectroscopy – Evaluation of identification criteria and uncertainty by the Bootstrap method,” Talanta, vol. 224, p. 121814, 2021.

[46] A. J. Nunes, P. Paixao, J. Proenca and R. J. N. Bettencourt da Silva, “Early warning of suspected doping from Biological Passport based on multivariate trends,” Int. J. Sports Med., vol. 41, pp. 44-53, 2020.

[47] N. Pinto, M. Magalhaes, E. Conde-Sousa, C. Gomes, R. Pereira, C. Alves, L. Gusmao and A. Amorim, “Assessing paternities with inconclusive STR results: The suitability of bi-allelic markers,” Forensic Sci. Int. Genet., vol. 7, pp. 16-21, 2013.

[48] B. Meijer, J. Thijs, J. Kleibeuker, A. van Zwet and R. Berrelkamp, “Evaluation of eight enzyme immunoassays for detection of immunoglobin G against Helicobacter pylori,” J. Clin. Microbiol, vol. 35, no. 1, pp. 292-294, 1997.

[49] B. G. Armitage P, Statistical methods in medical research, 3rd ed., Cambridge: Blackwell Science, 1994.

[50] D. Zwillinger and S. Kokoska, Standard probability and statistics tables and formulae, Boca Raton (FL): Chapman & Hall/CRC, 2000.

[51] C. B. Agresti A, “Approximate is better than "exact" for interval estimation of binomial proportions,” Am. Stat., vol. 52, no. 2, pp. 119-126, 1998.

[52] D. Altman, D. Machin, T. Bryant and M. Gardner, Statistics with confidence, 2nd ed., M. D. B. T. G. M. Altman DA, Ed., London: BMJ Books, 2000.

[53] R. Newcombe, “Two-sided confidence intervals for the single proportion: comparison of seven methods,” Stat. Med., vol. 17, no. 8, pp. 857-872, 1998.

[54] E. Wilson, “Probable inference, the law of succession, and statistical inference,” JASA, vol. 22, no. 158, pp. 209-212, 1927.

[55] J. A. Sphon, “Use of Mass Spectrometry for Confirmation of Animal Drug Residues,” J. Assoc. Off. Anal. Chem., vol. 61, pp. 1247-1252, 1978.

[56] R. Baldwin, R. Bethem, R. Boyd, W. Budde, T. Cairns, R. Gibbons, J. Henion, M. Kaiser, D. Lewis, J. Matusik, J. Sphon, R. Stephany and R. Trubey, “1996 ASMS FALL WORKSHOP: Limits to Confirmation, Quantitation, and Detection,” J. Am. Soc. Mass Spectrom., vol. 8, pp. 1180-1190, 1997.

[57] K. S. Webb and D. Carter, "GC Report number LGC/VAM/1998/010," LGC Limited, London, 1998.

[58] W. G. De Ruig, R. W. Stephany and G. Dijkstra, "Criteria for the detection of analytes in test samples," J. Assoc. Off. Anal. Chem., no. 72, pp. 487-490, 1989.

[59] Clinical Laboratory and Standards Institute, MM17 - Validation and Verification of Multiplex Nucleic Acid Assays (2nd ed.), Wayne (PA): CLSI, 2018.

[60] J. Gaddum, “Medical Research Council, Special Report Series no. 183,” Br Med J, 1933.

[61] C. I. Bliss, “The method of probits,” Science, vol. 79, no. 2037, pp. 38-39, 1934.

[62] D. J. Finney, Probit analysis, Cambridge: Cambridge University Press, 1947.

[63] Clinical Laboratory and Standards Institute, EP-17A2 Evaluation of detection capability for clinical laboratory measurement procedures, 2nd ed., Wayne (PA): CLSI, 2020.

[64] Clinical Laboratory and Standards Institute, MM53-A - Criteria for Laboratory Testing and Diagnosis of Human Immunodeficiency Virus Infection, Wayne (PA): CLSI, 2011.

[65] S. L. R. Ellison and A. Williams, (Eds); Eurachem/CITAC guide: Quantifying Uncertainty in Analytical Measurement, Third edition, 2012.

[66] JCGM, Evaluation of measurement data – The role of measurement uncertainty in conformity assessment (JCGM 106:2012), Sevres: BIPM, 2012.

[67] I. Kuselman, F. Pennecchi, R. J. N. Bettencourt da Silva and D. B. Hibbert, "IUPAC/CITAC Guide: Evaluation of risks of false decisions in conformity assessment of a multicomponent material or object due to measurement uncertainty (IUPAC Technical Report)," Pure Appl. Chem., vol. 93, no. 1, pp. 113-154, 2021.

[68] Association of Forensic Science Providers, "Standards for the formulation of evaluative forensic science expert opinion," Science and Justice, vol. 2009, pp. 161-164, 2009.

[69] L. A. Currie, "Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995)," Pure Appl. Chem., vol. 67, no. 10, pp. 1699-1723, 1995.

[70] European Food Safety Authority, The 2016 European Union report on pesticide residues in food, EFSA, 2018.

[71] I. Kuselman, F. Pennecchi, R. J. N. Bettencourt da Silva and D. B. Hibbert, "Risk of false decision on conformity of a multicomponent material when test results of the components' content are correlated," Talanta, no. 174, pp. 789-796, 2017.

[72] R. B. Silva and A. Williams, (Eds.), Eurachem/CITAC Guide: Setting and Using Target Uncertainty in Chemical Measurement, Eurachem, 2015.

[73] EU, Commission Regulation 2017/626 of 31 March 2017, EU, 2017.

[74] D. R. Cox, "The regression analysis of binary sequences (with discussion)," J R Stat Soc B, vol. 20, no. 2, pp. 215-242, 1958.

[75] D. J. Finney, Probit analysis, 2nd ed., Cambridge: Cambridge University Press, 1952.

[76] S. L. R. Ellison and A. Williams, (Eds.), Traceability in Chemical Measurement, 2nd ed., UK: Eurachem, 2019.