

**ORIENTACIÓN SOBRE EL ANÁLISIS DE DATOS
PARA EL DESARROLLO DE NIVELES MÁXIMOS Y LA MEJORA DE LA RECOGIDA DE DATOS¹**

ÍNDICE	Párrafos
Preámbulo	1-7
Recogida y envío de datos	8-16
Rellenar la plantilla de SIMUVIMA/Alimentos	17-52
Extracción de datos	53-57
Selección/depuración de datos	
Consideraciones generales	58
Origen de los datos	59-60
Depuración de datos	61-65
Tratamiento de los datos	
Falta de información	66-73
Información incorrecta	74-77
Datos para los que la información sobre la porción analizada no está clara	78
Datos procedentes de muestras bajo sospecha de ser fraudulentas o de estar económicamente adulteradas	79
Datos de muestreo selectivo	80-81
Determinación de valores atípicos/extremos y su tratamiento	82-92
Consideraciones sobre el límite de cuantificación (LC) y el límite de detección (LD)	
Exclusión según el LC	93-96
Suma de componentes y LC	97
Análisis de datos: generación de un resumen de datos	
Resumen sobre países, número de puntos de datos y periodo de cobertura de datos	98
Cobertura geográfica de los datos de presencia facilitados	99-104
Plazos de la cobertura de periodos de los datos de presencia facilitados	105-110
Análisis estadísticos de los datos de presencia y tratamiento de los conjuntos de datos para el desarrollo de NM	
Consideraciones generales	111-113
Número mínimo de puntos de datos para estimar los valores del percentil alto	114
Tratamiento de conjuntos de datos	
Tratamiento de conjuntos de datos con un número reducido de puntos de datos	115-119
Tratamiento de conjuntos de datos en los que los datos disponibles sobre el/los producto/s o el/los alimento/s individuales son insuficientes, pero en los que los datos sobre el grupo de alimentos son suficientes	120-123

¹ El texto de color azul que aparece entre corchetes simples indica modificaciones realizadas de acuerdo con los cambios que podrían implementarse en una primera fase.
El texto de color morado que aparece entre corchetes dobles indica modificaciones que podrían realizarse en una fase más tardía.

Tratamiento de conjuntos de datos (incluido el uso de métodos de sustitución) con una gran proporción de puntos de datos censurados por la izquierda	124-126
Métodos de sustitución para conjuntos de datos con una elevada proporción de datos censurados por la izquierda	127-130

Decisión sobre el tratamiento de múltiples conjuntos de datos

Análisis de conjuntos de datos individuales y combinados, y toma de decisión sobre la necesidad de comparar los conjuntos de datos individuales antes de combinarlos, especialmente si los patrones de distribución son muy diferentes.	131-136
---	---------

Casos en los que se recomienda el análisis de los conjuntos de datos individuales	137-140
---	---------

Realización de análisis de datos por visualización	141
---	-----

Agregación de datos y cálculo de estadísticas descriptivas	142-144
---	---------

Cálculo de las tasas de rechazo con NM hipotéticos

Estimación de NM hipotéticos	145-148
------------------------------	---------

Cálculo de las tasas de rechazo con NM hipotéticos	149-150
--	---------

Evaluación del impacto de un NM en la tasa de rechazo	151-152
---	---------

Mejora del cálculo de las tasas de rechazo	153-155
--	---------

Cálculo preliminar de los efectos de NM sobre la reducción de la exposición dietética en el contaminante a partir del/de los producto/s meta a NM hipotéticos

Cálculo de la exposición dietética y la reducción a NM hipotéticos	156
--	-----

Mejora del cálculo de las tasas de reducción de la exposición	157
---	-----

Preparación de las recomendaciones definitivas al CCCF	158-164
---	---------

Presentación de datos en los informes del GTe para el CCCF

Presentación del análisis de datos: análisis estadísticos	165-166
---	---------

ANEXOS²

² Por desarrollar.

PREÁMBULO

1. Los pasos para el desarrollo de un nivel máximo (NM) pueden incluir:
 - Identificación de un nuevo problema comercial o de salud relevante para una combinación de contaminante - producto básico/alimento.
 - Elaboración de un documento de debate que explore los datos de presencia preliminares, los datos de exposición y la relevancia de la combinación de contaminante - producto básico/alimento a nivel global.
 - Acuerdo del Comité del Codex sobre Contaminantes de los Alimentos (CCCF) para iniciar un nuevo trabajo, incluido el debate sobre el mandato de un grupo de trabajo y el envío a la CAC de una propuesta de nuevos trabajos.
 - Elaboración de un documento que recomiende NM en el procedimiento de trámites del Codex y un análisis más exhaustivo de los datos de presencia, los datos de exposición, la relevancia de la combinación de contaminante - producto básico/alimento a nivel global y el impacto de los NM propuestos.
 - Recomendación de enviar los NM a la Comisión del Codex Alimentarius (CAC) para su adopción.
2. La fuente de datos primaria del CCCF es SIMUVIMA/Alimentos, una base de datos internacional gestionada por la Organización Mundial de la Salud que contiene datos sobre los niveles de contaminantes en distintos alimentos. Se anima a los países miembros a enviar datos procedentes de sus programas de seguimiento nacionales, ya sea de forma regular o como respuesta a peticiones de datos emitidas por el CCCF. Estos datos deben cumplir ciertos criterios para poder ser presentados, como incluir un límite de cuantificación (LC) o un límite de detección (LD) en el caso de datos no cuantificados. Se pueden referenciar datos externos, como los datos procedentes de artículos científicos, pero no suelen usarse para establecer NM.
3. Antes de empezar a trabajar en un documento de debate o un documento de NM, el CCCF establecerá un mandato para el grupo de trabajo, y podrá pedir a la Secretaría del Codex que emita una petición de datos. Tal y como se describe en el *Manual de procedimiento* de la CAC, en su 30.ª edición, el mandato deberá establecer claramente el objetivo u objetivos que se pretenden alcanzar con la creación del grupo de trabajo, el idioma o los idiomas que se van a usar, y el plazo de tiempo en que se espera completar el trabajo. La petición de datos suele identificar el contaminante y los alimentos/productos básicos de interés, así como la gama de datos de los datos solicitados. También ha habido casos de peticiones de datos que solicitaban información relativa a LC y LD del método analítico y nombres de muestras específicos, o indicaban en qué campos de la base de datos SIMUVIMA/Alimentos se debía introducir la información, e identificaban la base de resultados apropiada.
4. Establecer un mandato y planificar el ámbito de aplicación de una petición de datos es un paso importante del proceso de recogida de datos. Si se presta la debida atención al mandato y a la petición de datos, los datos recopilados serán de mejor calidad para establecer NM.
5. La gestión de datos desempeña un papel clave en la labor de elaboración de NM, por lo que redundaría en el interés común disponer de datos de buena calidad (como contar con datos que reflejen una imagen precisa de la contaminación alimentaria y que se presenten junto con un análisis estadístico). Idealmente, los datos de presencia deben obtenerse mediante el muestreo basado en estadísticas (ref. *Directrices generales sobre muestreo* (CXG 50-2004)), y el análisis debe llevarse a cabo mediante métodos validados con LC y LD adecuados en laboratorios que dispongan de sistemas de garantía de calidad. Idealmente, los datos presentados por los países miembros deberían ser representativos a nivel nacional.
6. El objetivo de este documento de orientación es facilitar los elementos que permiten disponer de datos de buena calidad y garantizar un uso y análisis armonizados de los datos de presencia disponibles por parte de los distintos grupos de trabajo electrónico (GTE) durante el desarrollo/establecimiento de NM del Codex.
7. Esta orientación es para uso interno en el CCCF, pero las autoridades nacionales/regionales pueden usar la información relevante contenida en este documento para el desarrollo/establecimiento de NM nacionales/regionales.

RECOGIDA Y ENVÍO DE DATOS

8. La página introductoria de la base de datos SIMUVIVA/Alimentos de la OMS es [Sistema Mundial de Vigilancia del Medio Ambiente \(SIMUVIMA\) / Programa de seguimiento y evaluación de la contaminación de los alimentos](#). La presentación de datos (carga) y el proceso de extracción de datos (descarga) empiezan en el sitio web: [Base de datos SIMUVIMA/Contaminantes de alimentos](#).

9. La página de la base de datos se abre a una página de bienvenida con dos pestañas, una pestaña con la página de inicio y una pestaña de búsqueda. Para garantizar la funcionalidad completa, los miembros deben registrarse e iniciar sesión en sus cuentas. Después de iniciar sesión, quien esté enviando los datos tendrá acceso a una pestaña de carga, además de a las pestañas de inicio y búsqueda. Asimismo, la persona podrá acceder a plantillas estándar y de carga masiva para cargar datos, a la herramienta de aprendizaje electrónico de SIMUVIMA/Alimentos, y a enlaces útiles, como el apartado de preguntas más frecuentes.
10. Antes de enviar los datos, quienes los envíen deberían revisar los materiales de la página web de SIMUVIMA/Alimentos ([Seguridad y Nutrición de los Alimentos \(who.int\)](#)) o de las páginas de SIMUVIMA/Alimentos vinculadas. El documento INSTRUCCIONES PARA EL ENVÍO ELECTRÓNICO DE DATOS SOBRE SUSTANCIAS QUÍMICAS EN LOS ALIMENTOS Y LA DIETA (de aquí en adelante, las INSTRUCCIONES) de la página web de la base de datos SIMUVIMA/Alimentos ofrece instrucciones detalladas. Este documento proporciona instrucciones sobre cómo abrir una cuenta, iniciar sesión en la base de datos SIMUVIMA/Alimentos, introducir datos en la plantilla Excel y cargarla. Se recomienda estar familiarizado con Excel.
11. A la base de datos SIMUVIMA/Alimentos se pueden enviar datos sobre cualquier alimento en todo momento, no solo como respuesta a una petición de datos sobre alimentos específicos o en los plazos de interés establecidos. Si los datos que se envían responden a una petición específica, recomendamos indicarlo en el campo de comentarios. También pueden enviarse datos que quedan fuera del plazo de tiempo establecido en la petición de datos, ya que pueden ser informativos para estudiar el nivel de contaminante a lo largo del tiempo.
12. Si las personas que envían los datos tienen alguna pregunta sobre los aspectos técnicos del envío de datos, deben ponerse en contacto con el administrador de la base de datos SIMUVIMA/Alimentos. Estas preguntas podrían estar relacionadas con mensajes de error en el momento de la carga, problemas con el registro, dudas sobre cómo nombrar las muestras, qué campos son obligatorios, la definición de los campos, problemas de correspondencia, etc.
13. Si se tiene alguna pregunta sobre si los datos se ajustan a una petición de datos específica, se debe consultar a la presidencia del GTE y, si procede, a las Secretarías del Codex o del JECFA. Por ejemplo, podría tratarse de preguntas sobre si las muestras corresponden a las definiciones incluidas en la petición de datos o al mandato del GTE.
14. Las personas que presentan los datos deberían desarrollar y guardar los metadatos relacionados con los envíos de datos. Estos metadatos ayudan a responder a las preguntas que pueda tener el GTE. Estos metadatos pueden incluir el año de producción, los rangos generales de LD y LC correspondientes a un conjunto de datos, información sobre etiquetas del producto, información sobre la situación de la recopilación (por ejemplo importación o venta minorista), nombres del personal que envió los datos y cuándo se enviaron, identificación del lote asociado al conjunto de datos enviado, etc.
15. La persona que envía los datos debería evaluar la calidad de los datos antes de cargarlos a SIMUVIMA/Alimentos. En caso de que surjan dudas cruciales sobre la calidad de los datos (falta de información, análisis sospechosos), no deben enviarse hasta que se hayan resuelto estas dudas.
16. Si la persona que envía los datos identifica un problema relativo a un conjunto de datos después de haberlo enviado, deberá consultar al administrador de la base de datos de SIMUVIMA/Alimentos sobre si retirarlo o corregirlo. Este conjunto de datos debería ser identificable mediante una identificación del lote.

Rellenar la plantilla de SIMUVIMA/Alimentos

17. La hoja de cálculo de la plantilla de los envíos estándar (no masivos)³ tiene cinco pestañas que incluyen 1) una lista de verificación sobre instituciones que envían los datos, 2) correspondencia de los alimentos de la muestra, 3) una plantilla para resultados de análisis individuales, 4) el sistema de clasificación de la OMS y FoodEx2, y 5) sustancias químicas que constan actualmente en la lista de opciones de envío en un menú desplegable.
18. El primer paso para enviar datos es rellenar la pestaña "1. Inicio", que contiene una lista de verificación para la institución que prepara el conjunto de datos para su envío, y que incluye la identificación de la sustancia química de interés. (Tenga en cuenta que las INSTRUCCIONES incluyen una opción para las sustancias químicas que no están disponibles en el menú desplegable.)
19. El segundo paso consiste en repasar los nombres del alimento/pienso/producto del conjunto de datos y establecer la correspondencia entre la clasificación alimentaria nacional y la clasificación de la OMS y FoodEx2. La pestaña "2. Correspondencia de los alimentos" contiene la herramienta para establecer correspondencias: el

³ Consulte las INSTRUCCIONES PARA EL ENVÍO ELECTRÓNICO DE DATOS SOBRE SUSTANCIAS QUÍMICAS EN LOS ALIMENTOS Y LA DIETA para saber cómo realizar envíos masivos de datos con la plantilla.

Identificador de alimento local (columna A, texto libre) y dos niveles de clasificación en menús desplegados, a saber, Nivel 1: Grupo de alimentos de la OMS (columna B) y Nivel 2: Identificador de alimentos de la OMS (columna C). Una vez se hayan completado los campos del identificador del alimento local, el grupo de alimentos de la OMS y el identificador de alimentos de la OMS, se generan automáticamente el código de alimentos de la OMS, el código de FoodEx2 y el nombre de FoodEx2 en las columnas E, F y G de la pestaña 2.

20. Un motivo de confusión durante el envío de datos es con qué frecuencia hay que establecer correspondencias para cada alimento en la plantilla de correspondencia de alimentos. Por ejemplo: si la persona que envía los datos carga tres alimentos con los identificadores de alimentos locales de "Jengibre, cristalizado", "Jengibre en polvo, seco" y "Jengibre a tiras, seco", cada uno de ellos se debería introducir por separado en la plantilla de alimentos, pestaña 2, y se establecería la correspondencia en la OMS como "Hierbas, especias y condimentos" (columna B) y "Jengibre, raíz" (columna C). Sin embargo, si la persona que envía los datos está cargando 100 puntos de datos adicionales para "Jengibre, cristalizado", solo hay que establecer una sola vez la correspondencia para todas las muestras de "Jengibre, cristalizado".
21. Las INSTRUCCIONES también indican que, si la clasificación nacional es estable, la correspondencia solo debería establecerse una vez. Si bien es probable que algunos países o algunas regiones cuenten con sistemas de envío centralizado de datos, en otros países puede haber distintas instituciones o departamentos de instituciones con cuentas, por lo que pueden enviar sus datos por separado. En tal caso, las instituciones deberían intentar coordinar la manera de establecer correspondencias entre los alimentos de modo que los distintos envíos sean consistentes.
22. El tercer paso al rellenar la plantilla de SIMUVIMA/Alimentos es introducir los resultados de análisis individuales en la pestaña "3. Resultados de análisis individuales". Los campos incluyen el identificador de alimentos locales (cuya correspondencia con los códigos de la pestaña 2 se ha establecido anteriormente), la concentración química, las unidades de medición, LD, LC, etc. Puesto que se ha establecido una correspondencia de los identificadores de alimentos locales en la pestaña 2, las columnas B, C y D de la pestaña 3 se rellenarán automáticamente con la información de la correspondencia. La columna A mostrará automáticamente un error si algún campo de esta pestaña se ha completado de forma incorrecta. El resto de las columnas deben rellenarse siguiendo los pasos detallados de las INSTRUCCIONES.
23. Se debe tener en cuenta que las columnas con encabezados azules en la plantilla de SIMUVIMA/Alimentos son obligatorias. Las columnas con encabezados blancos son opcionales (se pueden dejar en blanco) en caso de que la información correspondiente no esté disponible.
24. Los campos actuales de los resultados de análisis individuales de la base de datos SIMUVIMA/Alimentos se enumeran en el Cuadro orientativo 1. El texto entre corchetes (tanto simples como dobles) indica cambios acordados pero que aún no se han implementado. El texto de color azul que aparece entre corchetes simples indica modificaciones realizadas de acuerdo con los cambios que podrían implementarse en una primera fase (en la 18.ª reunión del CCCF se informará de los plazos). El texto de color morado que aparece entre corchetes dobles indica modificaciones que podrían realizarse en una fase más tardía.
25. Desde el párrafo 26 hasta el párrafo 52, que pueden leerse después del Cuadro orientativo 1, se ofrece orientación adicional sobre los distintos campos a las personas que presentan los datos.

Cuadro orientativo 1: Campos actuales de la plantilla SIMUVIMA/Alimentos

Columna	Campo	Tipo de campo/Menú desplegable	Obligatorio u opcional	Marcador de texto (nuevos que se han propuesto o revisados)
E	Identificador de alimentos locales	Texto libre	Obligatorio	[Aportar un nombre breve y descriptivo del alimento, como "Reloj anaranjado" (en lugar de "Pescado") o "Arroz pulido/blanco (en lugar de "Arroz").]
F	Número de serie del informe	Texto libre	Obligatorio	[Se utiliza un número de serie para cada muestra. Los datos sobre distintos contaminantes en una misma muestra deberían tener el mismo número de serie. Las instituciones nacionales deberían coordinarse para seleccionar el número de serie y así garantizar que los números son informativos y no están duplicados.]
G	[Envío de] País/Región/Observador	Menú desplegable <ul style="list-style-type: none"> (Lista de países, regiones, observadores) Sin especificar 	[Obligatorio]	[Identifica el país, la región o el observador (región sin especificar) que envía los datos; no es el país productor. Si el observador no aparece en la lista del menú desplegable, se debe elegir "Sin especificar" y escribir en los comentarios el nombre del observador.]
H	Contaminante	Menú desplegable <ul style="list-style-type: none"> (Lista de contaminantes) 	Opcional	[Se debe elegir un contaminante de la lista. Es obligatorio incluir un contaminante, pero hacer una entrada manual en la "Columna H: Contaminante" es opcional si se ha indicado un contaminante en la Hoja de trabajo 1: Inicio. Si se ha seleccionado "Múltiples" en la Hoja de trabajo 1: Inicio, es obligatorio incluir manualmente los contaminantes en la Columna H.]
I	Origen de los alimentos	Menú desplegable <ul style="list-style-type: none"> Nacional Importación Origen mezclado Se desconoce 	Opcional	
J	Fecha de muestreo	Texto libre (AAAA)	Obligatorio	
K	Representatividad de la muestra [fiabilidad]	Menú desplegable <ul style="list-style-type: none"> Muestreo [[rutinario]] aleatorio Muestreo selectivo 	Obligatorio	[El muestreo selectivo se refiere a un seguimiento selectivo de resultados de contaminación específicos. El muestreo (rutinario) aleatorio se refiere a un muestreo que no es selectivo y que puede incluir una vigilancia rutinaria o un

Columna	Campo	Tipo de campo/Menú desplegable	Obligatorio u opcional	Marcador de texto (nuevos que se han propuesto o revisados)
		<ul style="list-style-type: none"> Se desconoce 		muestreo de tipos de alimentos específicos o de los países importadores.]
L	Identificación del laboratorio	Texto libre	Opcional	[El laboratorio que completó los análisis.]
M	Garantía de calidad de datos analíticos	Menú desplegable <ul style="list-style-type: none"> Solo garantía de calidad interna y normas de referencia. Participación correcta en pruebas de competencia relevantes/[comparaciones entre laboratorios] durante el período de muestreo y análisis. Acreditación oficial para los métodos relevantes durante el período de muestreo y análisis. Garantía de calidad del laboratorio desconocida. 	Opcional	
N	Unidades de medición para niveles de contaminantes	Menú desplegable [<ul style="list-style-type: none"> mg/kg µg/kg ng/kg pg/kg Bq/kg] 	Obligatorio	[Las unidades se deben verificar con cuidado. Se debe garantizar que las unidades elegidas del menú desplegable coinciden con los resultados de la muestra.]
O	LD	Texto libre	Obligatorio para resultados no cuantificados [(es decir, no detectados)] si no se proporciona LC. ([Opcional])	Se debe introducir un valor numérico mayor que 0 y menor que LC. Este campo contiene el límite de detección comunicado por el laboratorio. [LD ≠ o LC son obligatorios si no se aporta el resultado se introduce "no detectados" en Resultados (T).]

Columna	Campo	Tipo de campo/Menú desplegable	Obligatorio u opcional	Marcador de texto (nuevos que se han propuesto o revisados)
P	LC	Texto libre	Obligatorio para resultados no cuantificados si no se proporciona LD. [[Obligatorio]]	Se debe introducir un valor numérico mayor que 0 y que LD. Este campo contiene el límite de cuantificación comunicado por [el] laboratorio. [LD y/o LC son obligatorios si no se aporta el resultado se introduce "no detectados" en Resultados (T).]
Q	Resultados basados en	Menú desplegable <ul style="list-style-type: none"> • Contenido en grasa • Peso en seco • Tal cual (crudo, fresco [, tal como se vende]) • Tal como se consume 	Obligatorio	
R	Porción analizada	Menú desplegable <ul style="list-style-type: none"> • Solo comestible • [Alimento entero (comestible + no comestible)] 	Obligatorio	[Ejemplo: "fruto seco con cáscara" (comestible) en lugar de "fruto seco descascarado" (alimento entero)]
S	Estado del alimento analizado [[cocido/crudo]]	Menú desplegable <ul style="list-style-type: none"> • Cocido • Crudo • Se desconoce 	Opcional	[Ejemplo: "pescado crudo" en lugar de "pescado cocido"]
T	Resultados	Texto libre	Obligatorio	[Es obligatorio introducir un resultado: ya sea un cero, "no detectados" o un resultado numérico. Solo se puede introducir un cero o "no detectados" si se aporta LC o LD.]
U	Datos individuales versus agregados	Menú desplegable <ul style="list-style-type: none"> • Individuales • Agregados 	[[Obligatorio]]	
V	Confidencialidad de los datos	Menú desplegable <ul style="list-style-type: none"> • Sí • En blanco 	Opcional	[Todos los datos para los que se ha elegido "en blanco", o para los que no se ha seleccionado ninguna opción, se considerarán como no confidenciales a la hora de tratar y analizar los datos.]
W	Comentarios/Referencias	Texto libre	Opcional	

Columna	Campo	Tipo de campo/Menú desplegable	Obligatorio u opcional	Marcador de texto (nuevos que se han propuesto o revisados)
X	[[Año de producción/cosecha]]	[[Texto libre (AAAA)]]	[[Opcional]]	
Y	[[Información sobre la composición]]	[[Texto libre]]	[[Opcional]]	[[Información de las etiquetas o según lo que se determine analíticamente como ingredientes principales, contenido en grasa, contenido de agua o porcentaje total de sólidos de cacao.]]
Z-1	[[País/Región de producción del producto terminado]]	[[Menú <ul style="list-style-type: none"> • Se desconoce • Países/Regiones (A-Z)]]	[[Opcional]]	
Z-2	País/Región de origen de las materias primas	[[Menú <ul style="list-style-type: none"> • Se desconoce • Países/Regiones (A-Z)]]	[[Opcional]]	
AA	[[Tipo de producto]]	[[Menú: <ul style="list-style-type: none"> • Destinados a su posterior procesamiento • Listo para el consumo • No se aplica • Se desconoce]]	[[Opcional]]	[["Destinado a su posterior procesamiento" y "listo para el consumo" se definen para ciertos contaminantes y productos en la Norma general del Codex para los contaminantes y las toxinas presentes en los alimentos y piensos, CXS-193. Véase el párrafo 48.]]
BB	[[Situación del muestreo en la cadena de producción]]	[[Menú: <ul style="list-style-type: none"> • Se desconoce • Sitio de producción • Transporte en lotes a granel • Fronteras (importaciones/exportaciones) • Mercado/Comercio • Otros]]	[[Obligatorio]]	[[Aporta información sobre el punto de la cadena de producción en el que se obtuvo la muestra.]]
CC	[[Principio del método de análisis]]	[[Menú <ul style="list-style-type: none"> • Método A • Método B • Método Z]]	[[Opcional]]	

Columna	Campo	Tipo de campo/Menú desplegable	Obligatorio u opcional	Marcador de texto (nuevos que se han propuesto o revisados)
		<ul style="list-style-type: none">• Otros• Se desconoce]]		

26. **E. Identificador del alimento local.** Campo obligatorio. Si es posible, la persona que envía los datos debería proporcionar los nombres en inglés. Añadir detalles al nombre puede ayudar al analista de datos a clasificar las muestras (por ejemplo "zumos de piña-naranja" en lugar de "zumos"). Por otro lado, un nombre de muestra demasiado largo (por ejemplo si se enumeran todos los ingredientes de un alimento con múltiples ingredientes) puede complicar el trabajo de los analistas. La información suplementaria del nombre también se puede añadir a la columna de comentarios.
27. **F. Número de serie.** Campo obligatorio. Debería utilizarse un número de serie (ID de muestra) para cada muestra. Si se envía información sobre múltiples contaminantes en una misma muestra, debería utilizarse el mismo número de serie. Las instituciones nacionales tienen la responsabilidad de coordinar el uso del mismo número de serie para todos los envíos de la misma muestra. [Las instituciones nacionales deberían coordinarse para seleccionar el número de serie y así garantizar que los números son informativos y no están duplicados.](#) Las convenciones de numeración deberían aportar una información que pueda conducir a la persona que envió los datos, a la fecha del envío y al programa de seguimiento. Se deben evitar numeraciones con números de serie que no aportan información, como 1, 2, 3, etc. En caso de que los envíos de datos se hagan como respuesta a dos peticiones de datos distintas, la persona que envía el segundo conjunto de datos debería garantizar que los datos nuevos no reemplazan a los datos anteriores; en caso necesario, poniéndose en contacto con el administrador de la base de datos SIMUVIMA/Alimentos.
28. **G. [Envío de] País/Región/Observador.** [Campo obligatorio. Identifica al país, la región o el observador (región sin especificar) que envía los datos; no es el país productor. Si el observador no aparece en la lista del menú desplegable, se debe elegir "Sin especificar" y escribir en los comentarios el nombre del observador.]
29. **H. Contaminante.** [Campo opcional. Se puede añadir un contaminante a "Hoja de trabajo 1: Inicio" o mediante una entrada manual en "Columna H: Contaminante". Es obligatorio identificar un contaminante, pero es opcional incluir el nombre del contaminante en la Columna H si ya se ha indicado un contaminante en la Hoja de trabajo 1. Sin embargo, si se ha seleccionado "múltiples" en la Hoja de trabajo 1: Inicio, se requiere la entrada manual de los contaminantes en la Columna H.]
30. **I. Origen de los alimentos.** [Campo opcional. Identifica si las muestras son nacionales, importadas, de origen mezclado o si se desconoce.]
31. **J. Fecha de muestreo.** [Campo obligatorio. Identifica la fecha de muestreo.]
32. **K. Representatividad de la muestra.** [Campo obligatorio. El término "muestreo (rutinario) aleatorio" se refiere a un muestreo que no es selectivo y que puede incluir una vigilancia rutinaria o un muestreo de tipos de alimentos específicos o de los países importadores específicos. Por ejemplo, analizar una amplia variedad de muestras importadas de una determinada categoría de alimentos para detectar la presencia de un determinado contaminante es "aleatorio". El término "muestreo selectivo" debe elegirse para el muestreo de seguimiento después de dar con hallazgos específicos de contaminación. Por ejemplo, si un país observa que una muestra de un fabricante en concreto presenta niveles elevados de un contaminante, tomar muestras adicionales del mismo lote o lotes producidos al mismo tiempo por el mismo fabricante sería "selectivo".]
33. **L. [Identificación del laboratorio.** Campo opcional. Identifica el laboratorio que completó los análisis.]
34. **M. [Garantía de calidad de datos analíticos.** Campo opcional. Aporta información sobre la garantía de calidad de datos analíticos del laboratorio encargado de los análisis.]
35. **N. Unidades de medición para niveles de contaminantes.** [Campo obligatorio. Aporta unidades para niveles de contaminantes. Las unidades se deben comprobar con cuidado. Se debe garantizar que las unidades elegidas del menú desplegable coinciden con los resultados de la muestra. Se debe garantizar que la unidad indicada es la misma en los resultados, el LD y el LC. Idealmente, quien envía los datos debería incluir tanto LC como LD, aunque ahora estos campos solo son obligatorios para resultados no cuantificados.]
36. **O. Límite de detección (LD).** [El campo de LD contiene el límite de detección comunicado por el laboratorio. Se debe introducir un valor numérico mayor que 0 y menor que LC. Se debe tener en cuenta que LD \leq LC son obligatorios si se introduce "no detectados" en Resultados (T).] [\[\[Aunque informar del LD sea opcional, es recomendable.\]\]](#) (Observación: Los campos O. LD y P. LC se pueden cambiar de orden, de forma que el campo obligatorio de P. LC aparezca en primer lugar).
37. **P. Límite de cuantificación (LC).** [El campo de LC contiene el límite de cuantificación comunicado por el laboratorio. Se debe introducir un valor numérico mayor que 0 y que el LD. Se debe tener en cuenta que LD \leq LC son obligatorios si se introduce "no detectados" en Resultados (T).] [\[\[LC es obligatorio para las muestras enviadas después del XX/XX/XXXX. Antes del XX/XX/XXXX, LC solo era obligatorio para resultados no cuantificados si no se proporciona LD.\]\]](#) (Observación: Los campos O. LD y P. LC se pueden cambiar de orden, de forma que el campo obligatorio de P. LC aparezca en primer lugar).

38. **Q. Resultados basados en.** Campo obligatorio. Aporta información sobre si los resultados se basan en un análisis "tal cual (crudo, fresco o tal como se vende)", "tal como se consume" o según el contenido en grasa o el peso en seco. Entre los ejemplos se incluyen 1) "contenido en grasa": resultados para contaminantes liposolubles (en grasas) en el tejido de la carne animal, basados en el contenido en grasa; 2) "peso en seco": resultados para contaminantes en el tejido animal en los casos en los que la muestra se seca en el laboratorio hasta que se elimina la humedad antes del análisis; 3) "tal cual (crudo, fresco o tal como se vende)": hortalizas y frutas secas vendidas en comercios minoristas, así como los resultados para el preparado en polvo para lactantes, no reconstituido con agua y 4) "tal como se consume": resultados para el preparado en polvo diluido para lactantes.
39. **R. Porción analizada.** Campo obligatorio. Aporta información sobre si se analiza el alimento entero o solo la porción comestible (por ejemplo, frutos secos con cáscara (comestible) en lugar de frutos secos descascarados (alimento entero)).
40. **S. Estado del alimento analizado.** [Campo opcional. Aporta información sobre si la muestra del alimento está cocida o cruda (por ejemplo, pescado crudo en lugar de pescado cocido).]
41. **T. Resultados.** [Es obligatorio introducir un resultado: ya sea un cero, "no detectados" o un resultado numérico. Solo se puede introducir un cero o "no detectados" si se aporta LC o LD.]
42. **U. Muestra agregada.** Opcional [[Obligatorio.]] Los datos agregados se refieren a resultados basados en muestras asociadas, como muestras de Estudios de Dieta Total. A menudo los datos agregados quedan excluidos de los análisis de la tasa de infracción que se llevan a cabo para establecer NM apropiados, que se basan en observar la distribución de los datos y en que los porcentajes superiores excedan los NM propuestos. Sin embargo, los datos agregados pueden incluirse en la base de datos SIMUVIMA/Alimentos, y en el pasado se han incluido datos limitados en los análisis del CCCF. Antes de incluir datos agregados, habría que consultar al administrador de la base de datos de SIMUVIMA/Alimentos o a un estadístico. Si se incluyen datos agregados en un análisis de NM, debe indicarse en el informe del GTE.
43. **V. Datos confidenciales.** [Campo opcional. Se debe tener en cuenta que todos los datos para los que se ha elegido "en blanco", o para los que no se ha seleccionado ninguna opción, se considerarán como no confidenciales a la hora de tratar y analizar los datos.] Los países pueden enviar datos y marcarlos como "confidenciales" si quieren limitar su acceso a la FAO, la OMS y otros organismos técnicos vinculados, como el Codex. El administrador de la base de datos de SIMUVIMA/Alimentos puede proporcionar registros clasificados como "confidenciales" a las presidencias de los GTE; por lo tanto, las presidencias de los GTE siempre deberían consultar al administrador de la base de datos de SIMUVIMA/Alimentos en relación con la extracción de datos antes de descargarlos. Si algún país envía datos como "confidenciales" en respuesta a una petición de datos, este país debería informar también de ello a la presidencia del GTE durante la fase de extracción y análisis de datos.
44. **W. Comentarios/referencias.** [Campo opcional que puede utilizarse para añadir información relevante para la que no hay un campo en específico. Algunos ejemplos de información que se ha incluido en esta columna son la información en las etiquetas de los productos (como ingredientes o nombres detallados de productos) o la información detallada sobre el método de análisis. Otro dato que se puede introducir en esta columna es la referencia a una petición de datos específica.]
45. **[[X. Año de producción/cosecha.** Campo opcional. Aporta información sobre el año de producción, por ejemplo, para los alimentos de origen animal, o sobre el año de la cosecha, por ejemplo, para los alimentos de origen vegetal, en caso de que esta información esté disponible.]]
46. **[[Y. Información sobre la composición.** Opcional. Información de las etiquetas, como ingredientes principales, contenido en grasa, contenido de agua o porcentaje total de sólidos de cacao para el chocolate.]]
47. **[[Z. País/Región de origen/producción.** Opcional. Aporta información sobre el país o la región de origen o producción del [alimento, la materia prima, el lote muestreado o la partida]. En caso de que se desconozca esta información, se puede seleccionar "se desconoce".]]
48. **[[AA. Tipo de producto.** Opcional. Aporta información sobre si la muestra está "destinada a su posterior procesamiento" o "lista para el consumo", o si esta información "no se aplica" o "se desconoce". "Destinados a su posterior procesamiento" y "Listo para el consumo" se definen en distintas partes de la [Norma general para los contaminantes y las toxinas presentes en los alimentos y piensos](#) (CXS-193-1995). Para las aflatoxinas en el maíz, el sorgo, las nueces producidas por árboles, el maní y los higos secos, así como para el deoxinivalenol en los cereales en grano, "destinado a su posterior procesamiento" significa que tienen como fin someterse a un tratamiento/una elaboración adicional que haya demostrado reducir los niveles de aflatoxinas o de deoxinivalenol antes de ser utilizados como ingredientes en productos alimenticios, procesados de otro modo u ofrecidos para el consumo humano. Para las aflatoxinas en las nueces producidas por árboles y los higos secos,

"listos para el consumo" significa que no se van a someter a un tratamiento/una elaboración adicional que haya demostrado reducir los niveles de aflatoxinas antes de ser utilizados como ingredientes en productos alimenticios, procesadas de otro modo u ofrecidas para el consumo humano. Para los productos que se acaban de mencionar, se puede seleccionar "destinado a su posterior procesamiento», "listo para el consumo" o "se desconoce"; para otros productos se debería seleccionar "no se aplica".

49. La norma CXS-193-1995 también se refiere en varios lugares a comidas para lactantes "listas para el consumo" y a leche líquida "para su posterior procesamiento". El campo AA no se aplica a estos usos de "listo para el consumo" o de "para su posterior procesamiento".]]
50. **[[BB. Situación del muestreo en la cadena de producción.** Obligatorio. Aporta información sobre el punto de la cadena de producción en el que se obtuvo la muestra, por ejemplo, el lugar de producción (empresas, plantas de producción de alimentos, etc.), el transporte en lotes a granel, las fronteras (por ejemplo, si los análisis se hicieron en la importación o la exportación), el mercado/comercio u otro tipo de información. En caso de que se desconozca esta información, se puede seleccionar "se desconoce". Si se selecciona otro tipo de información, esta información puede añadirse en las observaciones.]]
51. **[CC. Principio del método de análisis.** Opcional. Se puede seleccionar el método de análisis, como ICPMS o GC-MS, del menú desplegable.]]
52. **Errores.** Antes de cargar los datos, la persona que los envía debería revisar el fichero cuidadosamente para asegurarse de que no contiene errores. Durante la carga, se escaneará el fichero de datos para detectar problemas antes de incorporar los datos a la base de datos. Quien envía los datos es responsable de corregir los errores y volver a enviar la plantilla. Los conjuntos de datos se pueden rechazar por varios motivos, algunos de los cuales se enumeran a continuación. Se puede contactar al administrador de la base de datos de SIMUVIMA/Alimentos en caso de necesitar ayuda.
 - a. Resultado introducido < LD, ausencia de LC o LD cuando el resultado es **no se ha detectado** **[[ausencia de LC]],** LD introducido > LC.
 - b. Fechas introducidas en un formato equivocado.
 - c. No se han completado los campos obligatorios.
 - d. Entradas duplicadas en la hoja de trabajo actual o la base de datos.

EXTRACCIÓN DE DATOS

53. El proceso de extracción de datos empieza en el sitio web de la base de datos: [Base de datos SIMUVIMA/Contaminantes de los alimentos](#). Como se ha indicado anteriormente, los analistas deben registrarse e iniciar sesión en sus cuentas para acceder a todas las funcionalidades. Una vez iniciada la sesión, los analistas verán una página de bienvenida con dos pestañas, una pestaña de página de inicio y una pestaña de búsqueda. La página de inicio contiene un número limitado de conjuntos de datos extraídos y preparados por región y contaminante. Para hacer búsquedas específicas, el analista debe seleccionar la pestaña de búsqueda. La función de búsqueda permite filtrar los datos por región de la OMS, contaminante, categoría de alimentos, nombre del producto y período de muestreo. Estos filtros permitirán al analista identificar datos que coincidan con una petición de datos o mandato concretos.
54. Para identificar el conjunto de datos más adecuado para extraer datos para desarrollar propuestas de NM se recomienda consultar con el administrador de la base de datos de SIMUVIMA/Alimentos. Las personas que presentan los datos pueden tomar decisiones durante el envío de los datos que provoquen la pérdida de información durante la extracción. Por ejemplo, los datos introducidos como "alimentos para lactantes y niños" podrían quedar fuera de una búsqueda limitada a "zumos de fruta y hortalizas". Otro ejemplo: el zumo puede categorizarse erróneamente como "fruta y productos de fruta" aunque el identificador de alimento local o el campo de comentarios identifiquen las muestras claramente como zumo. Consultar con el administrador de la base de datos de SIMUVIMA/Alimentos antes de extraer los datos puede ayudar al GTE a garantizar que se extraen todos los datos relevantes para el análisis de NM en SIMUVIMA/Alimentos.
55. La confidencialidad de los datos es otro motivo por el cual las presidencias de los GTE siempre deberían consultar con el administrador de la base de datos de SIMUVIMA/Alimentos en relación con la extracción de datos antes de descargarlos. El administrador de la base de datos de SIMUVIMA/Alimentos puede proporcionar registros calificados de "confidenciales" a las presidencias de los GTE. Estos registros no aparecen en una búsqueda estándar como las que se han descrito más arriba. Los miembros de un GTE que estén interesados en un análisis más detallado de datos confidenciales pueden consultar con la presidencia del GTE.
56. Es importante registrar todos los filtros y las palabras clave para el informe del GTE.

57. [[Los cambios se llevaron a cabo en la base de datos de SIMUVIMA/Alimentos en 2025/6 para hacer los LC obligatorios y para introducir nuevos campos, como X, Y, Z1, Z2, AA, BB y CC. Los datos heredados enviados antes de estos cambios deberían considerarse válidos y utilizarse para el desarrollo de NM.]]

SELECCIÓN/DEPURACIÓN DE DATOS

Consideraciones generales

58. En este apartado se ofrece orientación para la selección y la depuración de datos para su envío y análisis en el contexto del trabajo para el CCCF. Tiene como objetivo apoyar el tratamiento de datos coherente y transparente en todas las distintas fuentes, al mismo tiempo que acepta que las decisiones clave (por ejemplo, si es necesario desarrollar NM) son parte de la competencia del CCCF. Los procedimientos que aquí se describen pretenden facilitar el análisis de datos de alta calidad, además de promocionar la armonización entre los contribuidores.

Origen de los datos

59. Para el desarrollo de los NM en función de los datos de presencia, solo se deberían utilizar los datos de la base de datos de SIMUVIMA/Alimentos. Los datos que no procedan de SIMUVIMA/Alimentos solo pueden utilizarse para apoyar análisis complementarios con el objetivo de informar sobre el entendimiento de los datos, por ejemplo, cuando únicamente se disponga de datos limitados en la base de datos de SIMUVIMA/Alimentos para determinados periodos de tiempo o regiones, en particular los datos limitados procedentes de los países productores primarios.
60. Los datos que no procedan de SIMUVIMA/Alimentos, como los datos enviados directamente al GTE por el/los país/es o el/los observador/es de los miembros del Codex, o los obtenidos a través de la búsqueda bibliográfica, también se someterán a los procedimientos de depuración, según sea necesario.

Depuración de datos

61. El objetivo de la depuración de datos es eliminar del conjunto de datos las muestras que no son pertinentes o adecuadas antes de analizar el conjunto de datos para recomendar NM.
62. Entre los ejemplos de las muestras que se deberían eliminar de un conjunto de datos se incluyen:
- Las muestras que quedan claramente fuera del mandato del trabajo, por ejemplo, las muestras de ketchup para los trabajos sobre las salsas de tomate.
 - Las muestras que quedan fuera del intervalo de fechas del mandato del trabajo, por ejemplo, las muestras que tienen una antigüedad de 20 años si el mandato se refiere a los datos de los 10 años anteriores. Este punto es especialmente importante en aquellos casos en los que se haya adoptado un código de prácticas (CdP) para la prevención y la reducción de la contaminación por contaminantes relevantes en los alimentos desde la generación de datos más antiguos.
 - Las muestras a las que les falta información (véase los párrafos 66-73).
 - Las muestras con LC inaceptablemente elevados (véase los párrafos 93-96).
63. La depuración de datos se refiere únicamente al conjunto de datos extraído. Los datos originales de la base de datos de SIMUVIMA/Alimentos no se modificarán ni se verán afectados por los pasos que se indican a continuación, a no ser que la persona que envíe los datos pida correcciones u otro tipo de cambios al administrador de la base de datos de SIMUVIMA/Alimentos.
64. Para la depuración de datos, se recomienda solicitar la ayuda de un experto en el contaminante específico que pueda saber qué patrones de datos son irregulares y cuáles no.
65. Se deben registrar y describir todas las medidas tomadas en la depuración de datos en el documento de trabajo final del Codex presentado por el GTE en la reunión plenaria del CCCF; por ejemplo, la información detallada sobre el motivo de la exclusión de datos (por ejemplo, el LD/LC especificado es superior a los NM hipotéticos que se están considerando, valores atípicos discutibles, etc.), el número de puntos de datos excluidos durante el proceso de depuración, incluyendo también, si es posible, un desglose que explique cuántos se excluyeron en cada paso, etc.

Tratamiento de los datos

Falta de información

66. Una vez se hayan eliminado todos los datos irrelevantes o inadecuados del conjunto de datos, no se deberían excluir datos si todos los campos están cumplimentados (véase la sección de recogida y envío de datos) y los datos cumplen con los criterios para su carga en la base de datos de SIMUVIMA/Alimentos. Se debe tener en cuenta que, aunque todos los campos estén cumplimentados y los datos cumplan con los criterios, es posible que los datos sigan estando incompletos para determinar los NM.
67. En aquellos casos en los que se necesite la información de la muestra de la base de datos de SIMUVIMA/Alimentos, pero esta esté incompleta (por ejemplo, falta la información o no es clara), el primer paso será dirigirse al punto de contacto del país, la organización o el observador que envió los datos para conseguir la información que falta. Este paso es especialmente importante si la información que falta es indispensable para una petición de datos. También se puede pedir al administrador de la base de datos de SIMUVIMA/Alimentos que lleve a cabo esta tarea.
68. En caso de que la información que falta esté disponible, el país, la organización o el observador que envió los datos deberá entregar un fichero de datos corregidos al GTE y al administrador de la base de datos de SIMUVIMA/Alimentos. Es posible que puedan continuar los análisis con las muestras correctas. Se deberá tomar nota de la petición y las correcciones en el documento de trabajo del GTE.
69. En el caso de que no se pueda obtener la información que falta, y si el GTE llega a la conclusión de que los datos se deben excluir del análisis debido a esta falta de información, el GTE deberá tomar nota de la exclusión de los datos y de los posibles impactos de dicha exclusión en su documento de trabajo.
70. Estos son algunos ejemplos de falta de información por la que probablemente los datos se tengan que excluir del análisis posterior de los datos:
 - Falta el estado del alimento analizado (por ejemplo, cocido en lugar de crudo si el análisis solo se debe llevar a cabo con alimentos crudos).
 - La descripción del producto en el campo del identificador del alimento local no es adecuada (por ejemplo, el análisis se realiza con la "caballa", pero el producto se describe como "pescado", y el análisis depende de la identificación de las especies de pescado).
 - Etcétera.
71. Estos son algunos ejemplos de falta de información que no impediría el análisis posterior de los datos (según la revisión caso a caso):
 - Información sobre el muestreo: tipo de muestreo, situación del muestreo en la cadena de producción.
 - Estado del producto; por ejemplo, no se indica si está cocido o crudo en el campo S, pero la información se puede deducir de otra información aportada, como porque la muestra se describe como pescado cocido.
 - Principio del método de análisis utilizado.
 - Cuando el NM se ha basado en una suma de componentes y los datos no se notifican para todos los componentes, sino para aquel que contribuye significativamente a la suma, o en caso de que la presencia se notifique como una suma.
72. La presidencia del GTE no debería excluir todos los datos a los que les falta un campo opcional, pero sí debería considerar qué detalles se necesitan para el desarrollo y la elaboración de NM. Por ejemplo, es posible que algunos productos necesiten información adicional para desarrollar NM, como en el caso de los granos crudos: puede que la información sobre la fase de la elaboración sea importante para proponer NM; pero esta información no sería necesaria para los productos alimentarios terminados.
73. Tal y como se ha mencionado anteriormente, los datos que no procedan de SIMUVIMA/Alimentos y que estén considerados como parte de un análisis complementario pueden carecer de información importante. La presidencia del GTE debería aplicar el mismo criterio y las mismas preguntas (por ejemplo, si es de vital importancia incluir la información que falta) a los datos que no procedan de SIMUVIMA/Alimentos.

Información incorrecta

74. En casos en los que haya claros indicios de que la unidad de datos o la base sobre la que se asientan los datos son incorrectas, el primer paso sería dirigirse al punto de contacto del país, la organización o el observador que envió los datos y pedir que revisen las entradas por si necesitan correcciones. También se puede pedir al administrador de la base de datos de SIMUVIMA/Alimentos que lleve a cabo esta tarea.

75. En caso de que el país, la organización o el observador que envió los datos esté de acuerdo con que se necesita una corrección, deberán entregar un fichero de datos corregidos al GTE y al administrador de la base de datos de SIMUVIMA/Alimentos. Es posible que puedan continuar los análisis con estas muestras. Se deberá tomar nota de la petición y las correcciones en el documento de trabajo del GTE.
76. Si no se puede confirmar la exactitud de los datos y no se pueden hacer correcciones, estos datos deberían excluirse de los análisis posteriores.
77. Algunos ejemplos de errores aparentes que deberían conducir a contactar con quienes hayan enviado los datos para una posible corrección y reenvío son:
 - Todos los datos de un conjunto de datos con 200 muestras se expresan en $\mu\text{g}/\text{kg}$, excepto 5 puntos de datos cuantificados expresados en mg/kg . Al trazar estos datos en una curva de distribución de frecuencias, tras su conversión a la misma unidad, los cinco puntos de datos en mg/kg se identificarían como posibles valores atípicos (véase los párrafos 82-92 y el Anexo I).
 - 195 resultados de un conjunto de datos con 200 muestras de alimentos con un contenido en grasa típico del 5 % se sitúan en el rango de 0-20 mg/kg ; sin embargo, 5 puntos de datos se sitúan en el rango de 100-400 mg/kg , lo que sugiere que se notificaron en base a la grasa y no en base al peso total designado. Al trazar estos datos en una curva de distribución de frecuencias, se identificarían como posibles valores atípicos (véase los párrafos 82-92 y el Anexo I).

Datos para los que la información sobre la parte analizada no está clara.

78. Para determinados alimentos (por ejemplo, frutas o arroz), si la porción analizada no está clara (por ejemplo, fruta pelada en lugar de entera, o arroz descascarillado en lugar de arroz pulido), se podrá contactar con el punto de contacto del país, la organización o el observador que envió los datos para que lo clarifique. En el caso de que no se proporcione ningún tipo de clarificación, se debería reflejar si la información que no está clara es importante para el contaminante en cuestión y la concentración final hallada en el producto. Además, para algunos alimentos puede suponerse que la porción se analizó en el estado en el que suele venderse/consumirse; por ejemplo, los cítricos suelen ser frescos a no ser que estén claramente identificados como en conserva. Estas suposiciones deberán registrarse e incluirse en el documento final que presente el GTE a la reunión plenaria del CCCF. En el caso de que no pueda hacerse una suposición razonable, estos datos deberían excluirse del análisis posterior, a menos que se obtenga la información necesaria.

Datos procedentes de muestras bajo sospecha de ser fraudulentas o de estar económicamente adulteradas

79. A la hora de evaluar si las muestras contaminadas son el resultado de una adulteración fraudulenta o económica, primero se debe tener en cuenta la naturaleza del contaminante (por ejemplo, plomo frente a micotoxinas). Las grandes diferencias de año a año pueden deberse a una variabilidad natural (por ejemplo, un nivel elevado de micotoxinas debido a las condiciones climáticas específicas de una determinada región o de un año de producción). En otros casos, las grandes diferencias pueden deberse a malas prácticas (por ejemplo, con el plomo). Posibles indicios de muestras fraudulentas o económicamente adulteradas son:
 - Hay determinadas muestras de órdenes de magnitud superiores a las de otras, por ejemplo, de 0,1 en lugar de 100 $\mu\text{g}/\text{kg}$.
 - La variabilidad temporal de los datos (según el contaminante); por ejemplo, los datos son mucho más altos en un año determinado del conjunto de datos.

Los datos que estén claramente relacionados con muestras fraudulentas o económicamente adulteradas deberían excluirse del análisis, y esta exclusión debe documentarse.

Datos procedentes del muestreo selectivo y sesgos

80. El muestreo selectivo es distinto del muestreo aleatorio en cuanto que el muestreo selectivo se refiere a un seguimiento selectivo de resultados de contaminación específicos. En principio, estos datos no deberían utilizarse en la derivación de los NM, puesto que no son representativos de la población general, y puede que no reflejen los niveles alcanzables en situaciones habituales.
81. También cabe señalar que puede producirse algún sesgo en el muestreo aleatorio, ya que puede haber motivos para realizar un muestreo más selectivo en regiones o tipos de productos específicos. Dichos datos podrían incluir niveles superiores o inferiores al rango normal, y no deberían excluirse sin más consideración, pues reflejan la variación natural de los datos de presencia.

Determinación y tratamiento de valores atípicos/extremos

82. La presidencia del GTE debería revisar los conjuntos de datos para determinar si hay valores atípicos/extremos que se deberían eliminar. Este apartado ofrece orientación sobre los diferentes enfoques para identificar valores atípicos/extremos y los protocolos para eliminarlos en caso necesario.
83. Por regla general, es importante no descartar los valores atípicos/extremos, a no ser que haya un motivo válido para hacerlo.
84. Los valores extremos pueden ser valores válidos, por ejemplo, debido a la heterogeneidad de la distribución del contaminante (como en los focos de micotoxinas) o a la variación natural de los contaminantes medidos (por ejemplo, como resultado de las condiciones atmosféricas, las condiciones del suelo, etc.). Otros valores extremos pueden ser erróneos, por ejemplo, en el caso de los errores a la hora de medir y procesar los datos, así como los cálculos incorrectos o el uso de las unidades de medida incorrectas, o como resultado de un comportamiento fraudulento (adulteración económica).
85. Los valores erróneos o los valores que surgen como resultado de una adulteración económica siempre se deberían eliminar del conjunto de datos definitivo antes de determinar NM. Los valores extremos válidos se deberán revisar caso por caso, primero mediante la inspección visual de los datos (véase el Anexo III para ver ejemplos) y, después, con métodos estadísticos (Anexo I).
86. La presencia de valores atípicos en los conjuntos de datos tiene un impacto significativo en la media aritmética y en los valores extremos, pero no en la mediana. Ahora bien, se debe tener en cuenta el porcentaje de cualquier valor atípico posible en todo el conjunto de datos. Como son los valores de percentil alto, y no los valores máximos, los que se usan como base para el desarrollo de NM en función de las tasas de rechazo, en general el impacto de los valores atípicos en los NM determinados será escaso. Sin embargo, en aquellos casos en los que se excluye un porcentaje notable de puntos de datos (por ejemplo, entre el 2 y el 5 %) del conjunto de datos por considerarse un valor atípico, podría afectar a los valores de percentil alto (véase también el párrafo 92). En estos casos, lo apropiado es incorporar un comentario sobre el efecto de la exclusión de los valores atípicos en la viabilidad de los NM (es decir, en la tasa de rechazo) que se están considerando.
87. El Anexo I a este documento orientativo ofrece más detalles sobre los enfoques estadísticos para identificar valores atípicos.
88. Los "valores atípicos" definidos en la *Directrices sobre terminología analítica* (CXG 72-2009) parten de una distribución normal en un conjunto de datos compuesto de resultados de análisis repetidos de la misma muestra. Se establece que "los valores atípicos estadísticos se descartan, a menos que el estadístico decida conservarlos por una razón justificada". En cambio, los conjuntos de datos abordados en este documento orientativo son resultados analíticos de una variedad de muestras y de diferentes métodos analíticos. Al desconocerse la distribución de estos datos (en muchos casos, no seguirá una distribución normal), y dado que es posible que se combinen a partir de múltiples fuentes, resulta difícil predecir el rango de variación dentro de un conjunto de datos. Por consiguiente, esta orientación recomienda no descartar los valores atípicos estadísticos, a menos que se identifique y se explique científicamente una razón justificada para su exclusión.
89. Los valores extremos que estén identificados como errores deberían abordarse tal y como se describe en los párrafos 74-77. Sin embargo, los valores extremos que carezcan de una explicación clara deberían guardarse y evaluarse para determinar si deberían ser tratados como valores atípicos en el conjunto de datos definitivo. Se puede llevar a cabo un análisis de sensibilidad para entender qué impacto puede tener la inclusión de los valores atípicos en toda la evaluación general.
90. Como puede haber muchas causas para los valores extremos, y algunos de estos valores pueden no considerarse extremos si se combinan con datos de otras fuentes (países/regiones, años diferentes, etc.), se debería determinar la posible exclusión de un valor extremo como valor atípico según el conjunto de datos combinado y depurado. En el caso de que se analicen por separado los conjuntos de datos individuales, se debería considerar con más detenimiento excluir los valores extremos como valores atípicos.
91. Puede haber casos en los que los valores extremos sean científicamente válidos en función de la producción, las condiciones atmosféricas u otros posibles factores, como erupciones volcánicas, etc. Teniendo en cuenta las características de la distribución del contaminante de los datos de presencia en los alimentos, no se recomienda excluir sin más los valores extremos en función de los resultados de las pruebas estadísticas de valores atípicos u otros métodos, como la inspección visual. Dado que el rango de la distribución de la concentración que puede asumirse empírica o teóricamente varía significativamente en función del tipo de contaminante (metales pesados, micotoxinas, etc.), el tratamiento de los valores extremos debe determinarse caso por caso. Por ejemplo, debe prestarse especial atención a las micotoxinas, cuyas concentraciones pueden variar significativamente en función de los métodos de muestreo utilizados debido a la distribución heterogénea en un lote, así como a drásticas variaciones interanuales.

92. En el caso de que se excluyan los valores atípicos, se recomienda informar con claridad del motivo de la exclusión en el documento final que presente el GTE a la reunión plenaria del CCCF. Se pueden llevar a cabo análisis de sensibilidad para mostrar cómo la exclusión o no exclusión de los valores atípicos podría afectar o no al cálculo de los valores de percentil alto. Cabe reiterar que el hecho de que queden unos pocos valores extremos en el conjunto de datos apenas tendrá efecto en el cálculo de los valores de percentil alto, siempre que el número total de puntos de datos del conjunto de datos supere suficientemente el número mínimo de puntos de datos necesario para calcular los valores de percentil alto.

Límite de cuantificación (LC) y límite de detección (LD)

Exclusión según el LC

93. Los distintos métodos de análisis arrojan LD y LC diferentes. Un LC elevado no significa automáticamente que los datos deban excluirse.
94. Orientaciones para la inclusión/exclusión de datos en distintos escenarios de LC/LD
- En caso de que no se proporcione ningún LC/LD para un conjunto de datos específico
 - El primer paso será dirigirse al país, la organización o el observador que envió los datos para que facilite dicha información (es decir, el LD y/o el LC).
 - En caso de que el conjunto de datos solo contenga (prácticamente) resultados cuantificados: el conjunto de datos podría utilizarse⁴.
 - En caso de que el conjunto de datos contenga una cantidad significativa de datos censurados por la izquierda (datos individuales sin valores cuantificados (finitos), generalmente considerados como datos por debajo de los LC/LD notificados): el conjunto de datos no debería utilizarse.
 - En caso de que se proporcione un LC:
 - Se debe identificar un nivel de corte para el LC en el análisis en función de los NM que se consideren (por ejemplo: $LC < NM$ en debate, $LC < 1/2 NM$ en debate).
 - En caso de que el conjunto de datos contenga una cantidad significativa de datos censurados por la izquierda: se debe considerar más información contextual (por ejemplo, si los datos provienen de un país importador o un país productor).
95. Si casi todos los datos del conjunto de datos están por debajo de LC o si se notifican como no detectados (ND, < LC), no es posible estimar valores de percentil alto para establecer los NM propuestos. Cuando únicamente haya un número bajo de valores cuantitativos, el conjunto de datos debería tratarse caso por caso siguiendo las orientaciones recogidas en el apartado "Tratamiento de conjuntos de datos con una elevada proporción de datos censurados por la izquierda". En este caso, a la hora de proponer NM no resulta conveniente calcular valores de percentil alto utilizando únicamente los valores cuantitativos, pues podría resultar en NM innecesariamente elevados.
96. El documento de trabajo del GTE debería esbozar con claridad los criterios por los que se excluyeran determinados datos del conjunto de datos debido a LD/LC elevados (por ejemplo, el LD/LC notificado para algunas muestras es mayor que el NM propuesto, o el LD/LC notificado para algunas muestras es de "x" órdenes de magnitud mayor que el LD/LC más bajo de la amplia mayoría de otras muestras en el conjunto de datos, lo que sugeriría que hay un error), o si todo el conjunto de datos debería excluirse del análisis, ya que la eliminación de datos individuales podría introducir sesgos.

Suma de componentes y LC

97. En caso de que los niveles de contaminantes sean una suma de componentes (por ejemplo, las aflatoxinas totales), se debería considerar lo siguiente:
- Por lo general, los niveles de contaminantes que son una suma de componentes se notifican como límite inferior, es decir, que para los niveles de componentes no cuantificados se ponen los valores en 0.

⁴ En la plantilla actual de la base de datos de SIMUVIMA/Alimentos, no es obligatorio notificar el LD o el LC en caso de resultados cuantificados. Por lo tanto, estos datos, comunicados con esta plantilla, podrían utilizarse, en particular en caso de no disponer de suficientes datos, con la indicación clara en el documento presentado por el GTE al pleno del CCCF de que no se ha proporcionado ningún LC/LD para estos datos, de conformidad con las normas vigentes en el momento en que se presentaron los datos.

- Así, es necesario que los LD o LC para los componentes individuales se proporcionen para los componentes no cuantificados por debajo de LD o LC.
- Si solo se notifican datos de componentes individuales, los datos individuales pueden sumarse en un resultado total.
- En casos específicos, puede resultar apropiado notificar niveles de contaminantes que son una suma de componentes utilizando un enfoque de límite medio (es decir, los niveles no cuantificados de los componentes se ponen en $\frac{1}{2}$ de LC, o de LD en caso de que no se proporcione el LC) o un enfoque de límite superior (es decir, los resultados no cuantificados se ponen como iguales al LC, o al LD en caso de que no se proporcione el LC); sin embargo, estos casos deberían identificarse claramente por adelantado en las instrucciones de la petición de datos antes de enviar los datos.

ANÁLISIS DE LOS DATOS: GENERAR UN RESUMEN DE DATOS

Resumen de países, número de puntos de datos y periodo de cobertura de datos

98. Una vez depurado el conjunto de datos, se considera que los datos restantes son de calidad suficiente para el análisis. Deberá facilitarse en un cuadro un resumen de estos datos restantes con detalles (por ejemplo, el país de producción del producto terminado, el año de producción/cosecha, la cantidad de datos incluidos y excluidos), así como describirse en el documento de trabajo. Junto con el resumen deberán facilitarse todos los pasos seguidos en la depuración de datos, y en la justificación y las hipótesis formuladas. Además, podría ser útil proporcionar información (por ejemplo, de la FAO) sobre cuáles son las principales regiones productoras del producto objeto de debate. En función de este resumen, el GTE también podrá presentar un análisis más centrado de las áreas geográficas específicas y de los periodos de tiempo.

Cobertura geográfica de los datos de presencia facilitados

99. A la hora de evaluar un NM para un producto en particular, el conjunto de datos debería incluir una representación de las regiones productoras importantes para el comercio internacional. Por lo tanto, es útil (y puede que sea necesario en el futuro) que se notifique en la base de datos de SIMUVIMA/Alimentos el país/la región que produce los productos terminados (véase la sección de recogida y envío de datos). Además, es probable que las regiones más importantes de producción de determinados productos sean geográficamente limitadas; si este es el caso, la presidencia del GTE puede notificar esta información en el documento de trabajo. En este contexto, deberían considerarse los datos de las regiones productoras, puesto que los datos de los países importadores de los alimentos podrían incluir sesgos si los alimentos deben cumplir los requisitos del país importador, como podría ser un NM que ya está establecido en ese país. Si una región tiene NM muy restrictivos que provocan que los niveles de contaminación de los alimentos importados presenten una asimetría hacia la izquierda, podría tratarse de una representación poco precisa de la variabilidad en la contaminación de las regiones productoras. Sin embargo, los datos de los países importadores también reflejan los alimentos (ingredientes) comercializados internacionalmente y tal y como se consumen, por lo que deberían considerarse en cierta medida. Efectivamente, podría producirse contaminación adicional durante el transporte desde el país productor (por ejemplo, producción de micotoxinas).
100. En algunos casos, podría ser conveniente dar prioridad a los conjuntos de datos de los países productores con respecto a los de los países importadores. Sin embargo, en tal caso, los conjuntos de datos de los países productores deben reflejar la implementación de las buenas prácticas agrícolas y de producción (BPA, BPF) tal y como se recogen en los códigos de prácticas del Codex, si los hay, y ser representativos de los productos que se comercializan a nivel internacional.
101. Otra opción sería ofrecer análisis separados para los conjuntos de datos de los países productores e importadores. De ser posible, se debería realizar un análisis de sensibilidad sobre el uso de datos de los países productores en comparación con el de los países importadores con el fin de orientar la selección de datos utilizados para establecer NM.
102. Solo si hay suficientes datos que muestren un indicio de grandes diferencias en los niveles notificados entre las regiones o entre los países de una región se podría realizar un análisis por región o por país. Cabe señalar que, en el caso de un enfoque por países, se debería hacer para los principales países productores de la región, además de que se tendría que disponer de suficientes datos. (Véase la sección Análisis estadísticos de los datos de presencia/Tratamiento de conjuntos de datos para el desarrollo de NM).
103. En resumen, a pesar de haber distintos enfoques posibles para la observación de datos de los países importadores en comparación con los de los países productores o de datos regionales, es importante tener en cuenta que los NM del Codex son normas a nivel mundial, por lo que el enfoque por defecto para el procesamiento de datos debería ser analizar los datos a nivel mundial.

104. Orientación para conjuntos de datos sin cobertura geográfica:

- Si la región o las regiones sobre las que faltan datos son regiones productoras importantes, y si las regiones se comprometen expresamente a aportar datos adicionales, se suelen conceder varios años adicionales (por ejemplo, entre 2 y 3 años) para la recogida de datos antes de proseguir el debate sobre las propuestas de NM, siempre y cuando no haya ninguna urgencia y el CCCF esté de acuerdo. Una vez transcurridos los años adicionales concedidos, se proseguiría el debate sobre los NM a partir de los datos disponibles, independientemente de que se haya alcanzado o no la cobertura geográfica.
- Si la región o regiones productoras importantes no se comprometen a proporcionar los datos adicionales, o si no es factible recoger los datos adicionales dentro del plazo acordado (por ejemplo, entre 2 y 3 años), se continuará la consideración de los NM a partir de los datos disponibles, o se interrumpirá.
- Si la región o las regiones sobre las que faltan datos no son regiones productoras importantes, se proseguirá la consideración de los NM a partir de los datos disponibles.
- En caso de que exista una necesidad urgente de establecer un NM para garantizar la protección de la salud de los consumidores, se tendrá que alcanzar un compromiso en el CCCF para establecer un NM a partir de los datos disponibles. En tales casos, se podrá revisar el NM a los entre 3 y 5 años para evaluar si es necesario efectuar ajustes en cuanto haya más datos disponibles.

Plazos de la cobertura de periodos de los datos de presencia facilitados

105. Es conveniente que los datos de presencia proporcionados se refieran a varios años de producción para el desarrollo de NM. Los requisitos pueden ser diferentes para distintos tipos de contaminantes (por ejemplo, micotoxinas, toxinas vegetales, biotoxinas marinas, contaminantes de procesado y ambientales), y dependen de la variación interanual supuesta o de la evolución de la contaminación en el tiempo.
106. En el caso de contaminantes como las micotoxinas, de las que se sabe que presentan una variación interanual, los datos de los últimos 10 años pueden proporcionar una muy buena representación de la variación interanual; sin embargo, puede haber casos en los que deban considerarse más de 10 años de datos (por ejemplo, si se ha reducido el esfuerzo de muestreo en los últimos años, o si se dispone de menos conjuntos de datos de mayor calidad). Para otros contaminantes, la variación interanual es menos relevante y posiblemente se puedan seleccionar datos más recientes (o un intervalo temporal más corto). En cualquier caso, debe debatirse si para el análisis son relevantes los datos con una antigüedad de más de 10 años.
107. Además, podría ser relevante investigar/incluir datos más antiguos para saber si determinadas especies/subgrupos de un grupo/categoría de alimentos tienden a presentar niveles más elevados.
108. En determinados casos podría ser conveniente realizar un análisis de las tendencias temporales. En estos casos, deben tenerse en cuenta los datos de más de 10 años de antigüedad para determinar si las concentraciones han cambiado/están cambiando con el tiempo, y si eso se podría tener en cuenta para determinar si debe utilizarse un determinado número de años de datos para el desarrollo de NM a fin de representar las concentraciones actuales.
109. En caso de que se haya establecido e implementado un CdP del Codex, los datos considerados deberían ser de los años posteriores a la implementación de dicho CdP para reflejar las buenas prácticas agrícolas y de producción.
110. Si el GTE excluye datos por haber sido recogidos antes de la aplicación del CdP y sin que el país haya indicado que ya se habían aplicado buenas prácticas a lo largo de la cadena de producción antes del establecimiento del CdP, las exclusiones y la justificación deben documentarse claramente en el documento final presentado por el GTE al pleno del CCCF.

ANÁLISIS ESTADÍSTICOS DE LOS DATOS DE PRESENCIA/TRATAMIENTO DE CONJUNTOS DE DATOS PARA EL DESARROLLO DE NM**Consideraciones generales**

111. En los siguientes apartados se explican las consideraciones que se deben tener en cuenta antes de realizar un análisis estadístico de los datos extraídos/depurados y cómo deberían presentarse los resultados del análisis estadístico en los GTE para desarrollar NM aplicables a nivel global.
112. El análisis estadístico debe realizarse a partir de los datos extraídos/depurados. A la hora de evaluar los datos, el primer paso sería analizar la distribución del conjunto de datos. En general, la distribución de los datos de los contaminantes en los alimentos tiende a presentar una asimetría hacia la derecha, por ejemplo, una distribución log-normal (véase la Figura 1). Para estas distribuciones no normales, no es apropiado el uso de métodos

estadísticos paramétricos, que se basan en la distribución normal.

113. La norma CXS193 indica en su Anexo I que "los NM deberían establecerse a un nivel que sea (ligeramente) superior al rango de variación normal de los niveles en alimentos y piensos". Esto significa que para desarrollar un NM es necesario estimar valores de percentil alto (generalmente valores de percentil 95) con un nivel de confianza alto. En el ámbito de la inocuidad alimentaria, se suele utilizar un nivel de confianza del 95 %. La siguiente figura (Figura 1) explica, a partir de una distribución modelizada, la relación entre un valor de percentil alto, un NM hipotético (por lo general un valor redondeado del valor del percentil) y el porcentaje de puntos de datos que superan el NM hipotético cuando el NM es el mismo que el valor del percentil 98.

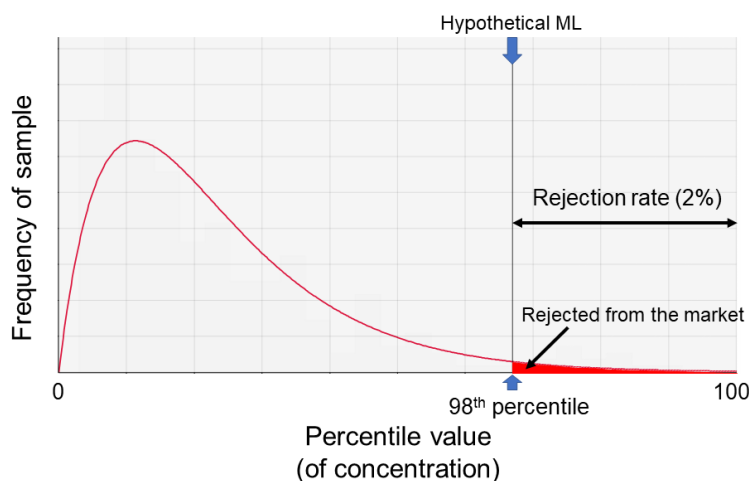


Figura 1. Representación simplificada de la relación entre un valor de percentil alto, un NM hipotético, la tasa de rechazo y el porcentaje de muestras que superan el NM hipotético.

Nota: En la figura de arriba se supone que el NM hipotético es el mismo que el valor del percentil 98. La elección del percentil 98 para el NM hipotético en esta figura es simplemente para ofrecer un ejemplo.

Número mínimo de puntos de datos para estimar los valores del percentil alto

114. Para desarrollar un NM, es necesario estimar los valores de percentil alto (generalmente, los valores de percentil 95) de un conjunto de datos. Se requiere un mínimo de 59 puntos de datos para una estimación del percentil 95 con un 95 % de confianza (véase la Opción 1 del Anexo II). El Anexo II aporta detalles adicionales sobre las opciones alternativas para calcular el número mínimo de puntos de datos necesarios para estimar los valores del percentil alto.

Tratamiento de conjuntos de datos

Tratamiento de conjuntos de datos con un número reducido de puntos de datos

115. Si la evaluación del Comité Mixto FAO/OMS de Expertos en Aditivos Alimentarios (JECFA) sugiere que hay un riesgo significativo para la salud debido a la exposición a un contaminante, se consideraría adecuado contar con un número menor de puntos de datos del número mínimo de puntos de datos (es decir, 59 puntos de datos, véase el párrafo 114 y el Anexo II) para desarrollar un NM, siempre y cuando el nivel de confianza de los valores estimados del percentil alto solo sean ligeramente inferiores al nivel de confianza alto estimado, por ejemplo, un 95 %. (Véase el Anexo II para un ejemplo de cómo calcular el nivel de confianza.) Por ejemplo, si se necesita urgentemente un NM para proteger la salud del consumidor, el GTE encargado de recomendar los NM debería considerar recomendar al CCCF el desarrollo de NM, incluso si solo hay disponible un número escaso de puntos de datos. Si en el futuro se llegara a disponer de datos suficientes, se podría considerar la revisión del NM previamente establecido.
116. Si no hay un riesgo inmediato para la salud y el número de puntos de datos no es suficiente para desarrollar un NM, se podrían emitir peticiones de datos adicionales. Sin embargo, si tras repetidas peticiones de datos el número de puntos de datos disponibles sigue siendo muy inferior al número mínimo de puntos de datos requeridos, se deberá recomendar al CCCF y caso por caso si se desarrolla un NM utilizando el conjunto de datos limitado o si se interrumpe el trabajo. Otra opción podría ser extender el NM a un grupo de alimentos más amplio si se justifica un NM para el grupo más extenso.

117. Para productos que no suelen consumirse de forma rutinaria o que no se comercializan a nivel internacional, la disponibilidad de los datos de presencia podría ser insuficiente. En tales casos, el GTE encargado de recomendar los NM debería considerar indicar al CCCF que el NM solicitado para la combinación producto/contaminante podría no cumplir los criterios descritos en el documento CXS193 y el *Manual de procedimiento* (Sección 4.5, "Política del Comité del Codex sobre Contaminantes de los Alimentos para la evaluación de la exposición a contaminantes y toxinas presentes en grupos de alimentos o piensos") para el desarrollo de NM.
118. Si el número de puntos de datos es significativamente inferior al número mínimo requerido de puntos de datos (es decir, 59; véase el párrafo 114 y el Anexo II), y no hay ninguna razón de peso para desarrollar un NM de inmediato, no se necesita realizar más análisis estadísticos. En caso de que se necesiten datos adicionales para establecer un NM estadísticamente sólido, puede que sea necesario emitir más peticiones de datos.
119. Al revisar los NM existentes, incluso si solo se dispone de un número escaso de puntos de datos de regiones limitadas o no se van a generar nuevos datos, el GTE encargado de recomendar NM no debería recomendar automáticamente revocar el NM debido al número reducido de puntos de datos, a menos que el NM no se ajuste a las buenas prácticas agrícolas o de producción, o a los datos toxicológicos actuales. En el caso de que el consumo del producto conllevara un riesgo potencialmente significativo, una opción sería recomendar al CCCF mantener el NM existente, o si ya no existe un riesgo significativo para la salud o no hay barreras comerciales, otra opción sería recomendar al CCCF revocar el NM. En algunos casos, a la hora de revisar un NM, si solo se dispone de un número escaso de puntos de datos para un producto en particular, podría ser posible considerar incorporar el producto en el grupo de alimentos del que originalmente se excluyó (por ejemplo, eliminar una exclusión para *Brassica* en conserva de un NM para hortalizas en conserva [véase REP18/CF12, párr. 32]).

Tratamiento de conjuntos de datos en los que los datos disponibles sobre el/los producto/s o el/los alimento/s individuales son insuficientes, pero en los que los datos para el grupo de alimentos sí son suficientes

120. Incluso cuando los puntos de datos son suficientes para todo un grupo de alimentos, si los datos se separan según los alimentos individuales de ese grupo de alimentos podría ocurrir que los puntos de datos sean demasiado escasos para el desarrollo de NM para alimentos individuales. En general, el análisis de los datos actualmente disponibles debería empezar con el desarrollo del documento de debate. Según los datos disponibles, el GTE debería recomendar un enfoque preliminar para el establecimiento de NM para alimentos individuales en lugar de para grupos de alimentos, así como recomendar los textos para una nueva petición de datos. Si tras la petición y la recogida de datos se comprueba que se dispone de menos puntos de datos de los previstos inicialmente, podría ser necesario cambiar los alimentos seleccionados para el NM a una variedad más amplia de alimentos, por ejemplo, de alimentos individuales a subgrupos de alimentos, o de subgrupos de alimentos a grupos de alimentos.
121. La idoneidad de desarrollar un NM para un grupo de alimentos depende de si la distribución de los valores de concentración del contaminante en los alimentos individuales dentro del grupo es similar. Pueden utilizarse pruebas estadísticas no paramétricas, como la prueba U de Mann-Whitney (para 2 conjuntos de datos) o la prueba H de Kruskal-Wallis (para 2 o más conjuntos de datos), para determinar si la distribución de esos alimentos en el grupo puede considerarse similar, incluso cuando el número de puntos de datos es relativamente reducido. Sin embargo, como los niveles de contaminación de los productos de un grupo de alimentos pueden variar significativamente, es importante ser flexible a la hora de elegir los métodos e interpretar las pruebas estadísticas. Si el número de puntos de datos es relativamente reducido, también resulta útil la comparación de conjuntos de datos mediante diagramas de caja y bigotes, siempre que el porcentaje de datos censurados por la izquierda sea inferior al 25 % del conjunto de datos correspondiente.
122. Si un alimento individual muestra una distribución diferente de las concentraciones del contaminante a la de otros alimentos del grupo de alimentos, podría ser necesario establecer dos NM diferentes: uno para el grupo de alimentos que excluye el alimento individual, y otro para el alimento individual. Pueden adoptarse enfoques similares para el/los subgrupo/s del grupo de alimentos. Si no se dispone de datos suficientes para que los alimentos individuales puedan alcanzar el número mínimo de puntos de datos requeridos, podrán emitirse peticiones de datos adicionales para aquellos alimentos para los que se considere necesario establecer NM. Si el consumo de un alimento individual que muestra un patrón de distribución diferente al del grupo de alimentos no contribuye de forma significativa a la exposición total al contaminante que es motivo de preocupación, podría considerarse insignificante desde el punto de vista de la protección de la salud del consumidor. En tales casos, no será necesario emitir ninguna petición de datos adicional, y se podría considerar excluirlo del desarrollo del NM para el grupo de alimentos (por ejemplo, en un NM para el plomo en la sal, un grado alimentario que excluya la sal de marismas [véase REP18/CF, párrafos 39-41]). En cuanto a los grupos de alimentos y sus subgrupos, se puede hacer referencia a los productos incluidos en las normas pertinentes del Codex sobre productos, la *Clasificación de alimentos y piensos* (CXA 4-1989) desarrollada por el Comité del Codex sobre Residuos de

Plaguicidas (CCPR) y el sistema de clasificación de alimentos utilizado por el Comité del Codex sobre Aditivos Alimentarios (CCFA) para los alimentos procesados.

123. A la hora de desarrollar un NM para un grupo de alimentos más amplio, y debido a la disponibilidad limitada de datos para alimentos o subgrupos individuales, es posible que algunos alimentos (o subgrupos) muestren patrones de distribución diferentes a los de otros del mismo grupo de alimentos. En caso de que no haya datos suficientes para desarrollar un NM aparte para estos alimentos, el GTE encargado de recomendar NM podría recomendar al CCCF si se debería excluir estos alimentos del desarrollo de NM hasta que haya datos suficientes disponibles.

Tratamiento de conjuntos de datos (incluido el uso de métodos de sustitución) con una gran proporción de puntos de datos censurados por la izquierda

124. El "conjunto de datos" de esta sección se refiere a un conjunto o conjuntos de datos que se encuentran entre los conjuntos de datos seleccionados para utilizarse para el desarrollo de NM. Esta sección es especialmente pertinente si los conjuntos de datos utilizados para el desarrollo del NM contienen una proporción elevada de puntos de datos no cuantificados (por ejemplo, debido a la baja sensibilidad de los métodos analíticos disponibles para la concentración en las muestras, a una frecuencia de presencia extremadamente baja, etc.) después de la depuración de datos.
125. Si bien no se ha encontrado ninguna definición oficial del término "censurado por la izquierda" en ninguno de los documentos del Codex, en estadística, los datos individuales sin valores (finitos) cuantificados se denominan "datos censurados por la izquierda", que por lo general se refieren a los datos inferiores a los LC/LD notificados.
126. Para el análisis estadístico de los conjuntos de datos que contienen datos censurados por la izquierda, se deberían considerar los métodos de sustitución convencionales, en especial a la hora de calcular estadísticas como el percentil 95, o cuando se estiman las tasas de rechazo de las muestras y las reducciones en la exposición de los productos meta en función de los NM hipotéticos. Si el conjunto de datos contiene una elevada proporción de datos censurados por la izquierda, no se recomienda analizarlo estadísticamente utilizando únicamente valores cuantificados, puesto que esta práctica introduce un sesgo en los resultados del análisis estadístico.

Métodos de sustitución para conjuntos de datos con una elevada proporción de datos censurados por la izquierda

127. El enfoque convencional para tratar los datos censurados por la izquierda para el análisis estadístico es el uso de una o más de las siguientes hipótesis de sustitución:
- Hipótesis de límite inferior (LB): los resultados por debajo del LC se sustituyen por cero o por el LD, si se conoce el LD (los resultados $< LD$ se sustituyen por cero).
 - Hipótesis de límite superior (UB): los resultados por debajo del LC se sustituyen por el valor del LC notificado.
 - Hipótesis de límite medio (MB): Una estimación puntual entre las dos hipótesis extremas (LB y UB), la asignación de un valor de $LC/2$, la raíz cuadrada del LC o $(LD + LC)/2$ si se conoce el LD para resultados analíticos por debajo del LC notificado. Por lo general, $LC/2$ es el que más se usa.

Para cada una de estas hipótesis, si no se notifica el LC y solo se notifica el LD, debe utilizarse el LD como alternativa.

128. Por lo general, en función de la distribución de los datos, estos métodos de sustitución pueden utilizarse para calcular medidas de tendencia central, como la media aritmética al calcular la exposición dietética (véase EHC 240⁵). Si el GTE encargado de recomendar NM debe llevar a cabo cálculos preliminares para la reducción de la exposición, la elección de las hipótesis de LB, MB o UB puede afectar a la media aritmética calculada y a la exposición estimada de los productos meta basada en la media aritmética. Sin embargo, para el desarrollo de un NM, el efecto de los valores censurados por la izquierda en la estimación empírica de los valores de percentil alto podría ser insignificante, y el impacto en el NM podría ser ínfimo, independientemente de la hipótesis que se elija, a no ser que una gran mayoría de puntos de datos esté censurada por la izquierda (es decir, $< LD$).
129. Los conjuntos de datos con una gran proporción de datos censurados por la izquierda deberían tratarse caso por caso en función del resultado de la evaluación por parte del JECFA del riesgo del contaminante y del consumo del alimento en cuestión. Lo ideal sería calcular y presentar las hipótesis de LB, MB y UB. Es muy importante conocer la distribución de los valores cuantificados en caso de que se dé un elevado porcentaje de datos

⁵ ENVIRONMENTAL HEALTH CRITERIA 240, *Principles and methods for the risk assessment of chemicals in food* (OMS, 2009)

censurados por la izquierda a la hora de estimar valores de percentil alto utilizando una distribución modelada para el establecimiento de NM.

130. Si la dispersión de valores cuantificados se sitúa dentro de un rango estrecho (valores próximos entre sí) y cercano al LC notificado, el desarrollo de un NM podría ser innecesario, a menos que el contaminante sea altamente tóxico. El GTE podrá comunicar su recomendación al CCCF con respecto a la idoneidad de un NM en esta situación.

Tratamiento de múltiples conjuntos de datos

Análisis de conjuntos de datos individuales y combinados y toma de decisiones sobre la necesidad de comparar los conjuntos de datos individuales antes de combinarlos, especialmente si los patrones de distribución son significativamente diferentes.

131. Dado que los NM del Codex son de aplicación global, lo ideal sería que se basaran en conjuntos de datos globales. Si bien el enfoque por defecto para el desarrollo de NM es el uso de los conjuntos de datos globales combinados, se proporcionan conjuntos de datos individuales por año o por región para su consideración adicional. Dependiendo del caso, y según los análisis estadísticos descritos en este apartado, será el CCCF quien decida si un NM debe basarse en un conjunto de datos globales o en un conjunto de datos de una región/año concretos.
132. Hay métodos estadísticos paramétricos disponibles para comparar patrones de distribución de conjuntos de datos individuales por región/país o por año. La hipótesis nula es que se supone que todos los conjuntos de datos siguen la misma distribución. Estas pruebas incluyen la prueba t (para 2 conjuntos de datos) o ANOVA (para 2 o más conjuntos de datos).
133. En internet se pueden encontrar muchas plantillas de métodos estadísticos no paramétricos. Por ejemplo, en el sitio web de la JMPR de la FAO se pueden descargar plantillas de MS Excel para realizar las pruebas U de Mann-Whitney (para 2 conjuntos de datos) y las pruebas H de Kruskal-Wallis (para 2 o más conjuntos de datos)⁶.
134. Además, se aconseja elaborar diagramas de caja y bigotes o histogramas de cada conjunto de datos para comprobar si existen diferencias visuales en las distribuciones antes de combinar los conjuntos de datos. Solo resulta preferible elaborar un histograma cuando el conjunto de datos contiene un número suficiente de puntos de datos (véase el párrafo 114 y el Anexo II). Para un conjunto de datos con un número menor de puntos de datos, resulta difícil conocer la forma de la distribución mediante un histograma, por lo que resulta más útil elaborar un diagrama de caja y bigotes (véase el Anexo III).
135. Normalmente, en los GTE se proponen NM utilizando conjuntos de datos combinados (conjuntos de datos globales). En los casos en los que no se observa una diferencia significativa entre las distribuciones de múltiples conjuntos de datos procedentes de distintas fuentes, se considera poco importante llevar a cabo análisis de datos adicionales y la comparación de cada conjunto de datos individual (aunque lo ideal sería hacerlo si los recursos y el tiempo lo permiten).
136. Cuando el número de puntos de datos es significativamente diferente entre conjuntos de datos individuales de diferentes regiones/países, el conjunto de datos combinado resultante refleja principalmente las condiciones de un país/una región con un mayor número de puntos de datos, en lugar de reflejar las de todos los países/las regiones que presentan los datos. Para resolver este problema, si bien implicaría un proceso complejo, sería teóricamente viable equilibrar los conjuntos de datos ponderándolos en función de la proporción del volumen de producción o comercio, o cualquier otro factor razonable. Sin embargo, en el CCCF no se había considerado en el pasado la metodología ni la justificación para el uso de la ponderación de datos y, debido a su complejidad y a la carga de trabajo asociado, no se ha previsto llevarla a cabo hasta que se elabore en el futuro una orientación factible para ello. En caso de que preocupe la distorsión en la distribución del conjunto de datos combinado debido a la presencia de un conjunto de datos muy extenso, se debería considerar caso por caso. Por el momento, también se podría recomendar llevar a cabo análisis para cada conjunto de datos individual, tal y como se expone en los siguientes apartados, o solicitar el consejo del JECFA sobre el análisis de datos.

Casos en los que se recomienda el análisis de los conjuntos de datos individuales

137. Si una prueba estadística indica una diferencia significativa entre la distribución de conjuntos de datos múltiples, y si la diferencia es considerable, se recomienda analizar los conjuntos de datos individuales junto con los conjuntos de datos combinados para el desarrollo de NM. No obstante, esto debe decidirse caso por caso, puesto que el alcance de las diferencias en la distribución suele depender de la combinación del producto y de los contaminantes que se están examinando. Se deberá aportar la justificación del análisis de conjuntos de datos

⁶ "Appendix XIV Electronic Attachments (2020_Nov)" y "XIV 12 Spreadsheet for Kruskal_Wallis 20 group.xls" para realizar la prueba U de Mann-Whitney y la prueba H de Kruskal-Wallis.
<https://www.fao.org/agriculture/crops/thematic-sitemap/theme/pests/jmpr/jmpr-docs/en/>

concretos y por separado junto con el conjunto de datos combinado y, en caso de que no se pueda encontrar una justificación, se debería utilizar en su defecto el conjunto de datos combinado.

138. Al considerar el uso de conjuntos de datos individuales, se recomienda comparar los resultados estadísticos, como los valores de percentil alto de los conjuntos de datos separados con los conjuntos de datos combinados. Cabe señalar que no se pueden obtener valores sólidos de percentil alto para conjuntos de datos individuales o combinados cuyas muestras tengan un tamaño inferior al número mínimo de puntos de datos requerido (véanse el párrafo 114 y el Cuadro 1 del Anexo II).
139. Cabe destacar que cuando se consideran varios conjuntos de datos individualmente, es posible que se identifiquen varios NM posibles. Queda fuera del ámbito de aplicación de esta orientación aportar orientaciones sobre qué NM posibles se deben elegir, ya que el CCCF es quien decide sobre la elección del NM adecuado, siempre teniendo en cuenta elementos que pueden diferir según el caso.
140. Si los conjuntos de datos de regiones/países diferentes se analizaran por separado mediante los métodos estadísticos recomendados en esta orientación, el GTE encargado de recomendar NM podría proporcionar elementos para la elección del conjunto de datos sobre el que se basará el NM para su consideración y decisión por parte del CCCF. Por ejemplo, si se sabe con seguridad que los conjuntos de datos con concentraciones elevadas corresponden a productos producidos con arreglo a las buenas prácticas (CdP del Codex o BPA, BPF, etc.), habrá que centrarse en los conjuntos de datos con concentraciones elevadas para considerar los NM aplicables a escala global.

Realización de análisis de datos por visualización

141. Existen distintos métodos para ilustrar los cuadros/gráficos y diagramas para mostrar la distribución de los datos de presencia y los valores estadísticos para evaluar la idoneidad del anteproyecto del NM propuesto. El Anexo III incluye ejemplos de los análisis exploratorios de datos y las estadísticas. En función del contaminante, el número de puntos de datos y la distribución de los datos de presencia (paramétricos frente a no paramétricos) se podrán realizar análisis exploratorios de datos por visualización y según el caso.

Agregación de datos y cálculo de estadísticas descriptivas

142. La siguiente información y las estadísticas resumidas pueden presentarse para los conjuntos de datos de presencia, ya que cumplen con los requisitos mínimos:
 - Número de puntos de datos totales.
 - Número de puntos de datos inferiores a los LC notificados (o LD) y/o proporción del número de datos < LC (o < LD) entre el número total de puntos de datos.
 - Rango de LC (LD) entre los datos excluidos o incluidos en el conjunto de datos definitivo utilizado para recomendar NM.
 - Media (media aritmética); en caso de que el conjunto de datos contenga puntos de datos inferiores al LC (LD), podrían prepararse tres medias aritméticas basadas en tres hipótesis de sustitución de LB, MB y UB (si la distribución es o se aproxima a la normal y simétrica).
 - Si la distribución es muy asimétrica, una opción podría ser obtener una media geométrica utilizando el mismo enfoque que el descrito arriba, pero es una opción que todavía no se ha utilizado en el CCCF.
 - Mediana (valores del percentil 50), pero si más del 50 % de los puntos de datos son inferiores al LC, la mediana podría notificarse como "< LC" (o < LC).
 - Valores de percentil muy alto (por ejemplo, valores del percentil 95, 97 y 98, según sea necesario y dependiendo de los debates en el GTE sobre la tasa o las tasas de rechazo apropiadas); si más del 95 %, 97 %, etc., de los puntos de datos están por debajo del LC, los percentiles asociados podrían notificarse como "< LC" (o < LC).
 - Mínimo.
 - Máximo: en los casos en los que se identificó el máximo como un valor atípico potencial y el valor máximo aún no se había excluido del conjunto de datos, podría resultar útil notificar el segundo valor más alto, el tercer valor más alto, etc., para ampliar el contexto.
 - Rango de datos cuantificados.
 - Desviación estándar, que es una medida de la cantidad de variación de una distribución paramétrica.
 - Valores intercuartílicos (véase el Anexo III), que son una medida de la cantidad de variación de una distribución no paramétrica.

143. Muchas de estas estadísticas pueden obtenerse fácilmente utilizando las funciones de Excel, mediante el menú de estadísticas descriptivas en las herramientas de análisis de datos de MS Excel o desde cualquier otra aplicación estadística. Las distintas aplicaciones estadísticas utilizan protocolos de cálculo diferentes y, por lo tanto, devuelven valores de percentil diferentes para el mismo conjunto de datos. Por consiguiente, cuando se calculen valores de percentil utilizando aplicaciones informáticas, los valores obtenidos deben comprobarse cuidadosamente teniendo en cuenta las funciones utilizadas; además, se deben indicar el nombre y, si es necesario, la información relevante de la aplicación utilizada para el cálculo.
144. Si la mayor parte del conjunto de datos está formada por datos censurados por la izquierda, puede que no sea posible calcular valores de percentil alto. En estos casos, se recomienda utilizar el método de sustitución con las hipótesis de LB, MB o UB. Aunque se deba hacer caso por caso y en función del número de valores cuantificados disponibles y de la distribución del contaminante, pueden utilizarse métodos como la estimación de valores de percentil alto a partir de funciones de densidad de probabilidad mediante la modelización de la distribución de los datos de presencia.

Cálculo de las tasas de rechazo con NM hipotéticos

Estimación de NM hipotéticos

145. A partir de un valor de percentil alto (por lo general, ligeramente superior al percentil 95) del conjunto de datos meta, se identifica un valor candidato para un NM teniendo en cuenta la precisión del método analítico actual y las cifras importantes de los resultados analíticos. Los valores numéricos de los NM deberán ser preferentemente cifras a intervalos regulares en una escala geométrica (0,01; 0,02; 0,05; 0,1; 0,2; 0,5; 1; 2; 5; etc.), a menos que esto pueda plantear problemas en cuanto a la aceptabilidad de los NM (véase el Anexo I a CXS 193-1995).
146. Una vez se haya determinado el valor numérico candidato de un NM, los siguientes valores superiores o inferiores también se propondrán como NM hipotéticos. (Por ejemplo, para un NM candidato de 0,5 mg/kg, los NM hipotéticos adicionales podrían ser de 0,4 y 0,6 mg/kg). En el caso de revisión de NM existentes, el NM existente también debería añadirse como uno de los NM hipotéticos. Además, los valores obtenidos de los valores de percentil alto (por ejemplo, los valores de percentil 95, 97 y 98) también pueden utilizarse directamente como NM hipotéticos.
147. Cuando se decida analizar por separado múltiples conjuntos de datos con distintos patrones de distribución, los NM hipotéticos se determinarán a partir de los valores de percentil alto de cada conjunto de datos. Si los patrones de distribución son significativamente diferentes, los NM hipotéticos de los conjuntos de datos individuales podrían ser significativamente diferentes (véase el párrafo 139).
148. No existe ninguna regla sobre el número de NM hipotéticos que deben proponerse, pero es preferible tener más de 2 NM hipotéticos (que se deben valorar caso por caso) para considerar sus efectos en la exposición dietética en productos meta y el impacto económico a consecuencia de las tasas de rechazo; todo esto deberá debatirlo más a fondo el GTE responsable de recomendar NM al CCCF.
149. En el documento que se prepare para el CCCF se debe explicar con suma claridad el razonamiento de la elección de un NM. El razonamiento debe basarse claramente en el análisis y no en un valor elegido de antemano.

Cálculo de las tasas de rechazo con NM hipotéticos

150. La tasa de rechazo se define mediante la ecuación que figura a continuación. Puede obtenerse con facilidad utilizando directamente las funciones de MS Excel (como la función COUNTIF) o mediante aplicaciones estadísticas o de modelización/simulación después de modelizar cada conjunto de datos. Si se utiliza un método diferente para calcular la tasa de rechazo, deberá indicarse claramente en el documento de trabajo.

$$\text{Tasa de rechazo (\%)} = (\text{número de puntos de datos} > \text{NM hipotético}) / (\text{número total de puntos de datos}) \times 100$$

151. Cabe destacar que la tasa de rechazo puede ser diferente de la prevista a partir del percentil alto por el redondeo. Cuanto menor sea el número de puntos de datos utilizado para calcular las tasas de rechazo, mayor será la incertidumbre en la estimación de la tasa de rechazo. A la hora de calcular la tasa de rechazo, se supone que las muestras que superan el NM hipotético quedan excluidas del mercado con una probabilidad del 100 % mediante la aplicación del NM.

Evaluación del impacto de un NM en la tasa de rechazo

152. Para evaluar el impacto en el comercio internacional del producto, debe utilizarse el conjunto de datos combinado global y, si es necesario, los conjuntos de datos de cada región. Aunque no se recomienda al GTE encargado de proponer NM calcular las tasas de rechazo país por país, es posible que algún miembro del Codex pida al GTE/CCCF que considere el impacto económico que ciertos NM hipotéticos puedan tener en su país en caso de que esta información no se refleje lo suficiente en la evaluación del impacto del NM hipotético en la tasa

de rechazo para su región/país.

153. Para los contaminantes conocidos por presentar una gran variación anual en las concentraciones, la tasa de rechazo debería calcularse para el conjunto de datos de cada año, si es posible, para la comparación interanual de las tasas de rechazo.

Mejora del cálculo de las tasas de rechazo

154. Con diferentes NM hipotéticos, es posible que las tasas de rechazo calculadas no cambien significativamente, dependiendo de la distribución del contaminante. Dada la naturaleza asimétrica de los datos del contaminante, el número de puntos de datos en el rango del percentil alto suele ser mucho más bajo que en el rango de percentil bajo, lo que afecta a la estimación de NM hipotéticos y a las tasas de rechazo (véase la Figura 1 para conocer la forma de la distribución).
155. Si el patrón de distribución del conjunto de datos combinado (potencialmente globales) muestra un único pico, se pueden utilizar aplicaciones de modelización/simulación (como @Risk, Crystal Ball, R, etc.) para modelizar la distribución a fin de estimar continuamente la distribución cerca de los valores del percentil alto a partir de la función de distribución, por lo que será posible obtener mejores estimaciones de la tasa de rechazo.
156. En caso de que se necesite una evaluación de impacto más detallada o mejorada en relación con la tasa de rechazo, se puede solicitar una evaluación al JECFA.

Cálculo preliminar de los efectos de NM sobre la reducción de la exposición dietética en el contaminante a partir del producto meta a NM hipotéticos

Cálculo de la exposición dietética y la reducción a partir del producto meta a NM hipotéticos

157. Para garantizar que el NM propuesto proteja la salud de los consumidores, puede que resulte apropiado evaluar cuantitativamente el efecto de un NM hipotético en la reducción de la exposición dietética a partir del producto meta comparando la exposición sin un NM y NM hipotéticos bajo consideración. En caso de revisión de un NM existente, la exposición dietética a partir del producto meta con el NM ya establecido se compara con la exposición con el nuevo NM hipotético (NM revisado).

Se puede efectuar la evaluación preliminar de la exposición a partir del producto meta como información de referencia en los GTE encargados de recomendar NM en caso de que haya recursos disponibles.

En caso de que se necesite una evaluación detallada o más compleja (por ejemplo, una evaluación general de la exposición), el CCCF puede pedir al JECFA que lleve a cabo una evaluación de los efectos de los NM hipotéticos en la reducción del riesgo derivado de la exposición dietética, pues el cálculo de la exposición dietética es una función de evaluación de riesgos que debería efectuar el JECFA. En el Anexo IV se incluyen ejemplos sobre cómo calcular la tasa de rechazo de la exposición dietética a partir del producto meta.]

Mejora del cálculo de la reducción de las tasas de reducción de la exposición dietética

158. En caso de que se necesite una evaluación de impacto más detallada o mejorada en relación con la exposición dietética, se puede solicitar una evaluación al JECFA.

Preparación de las recomendaciones definitivas al CCCF

159. A la hora de finalizar un análisis, es posible que el GTE encargado de recomendar NM se enfrente a cuestiones como si se necesita un NM en función de una evaluación preliminar de la exposición, si se necesita proponer numerosos anteproyectos de NM, qué conjuntos de datos se deberían utilizar o qué tasa de rechazo es la apropiada. Es posible que surjan perspectivas distintas a las previstas cuando se elaboró el mandato para el trabajo. El GTE no debería tomar una decisión final, sino que debería preparar preguntas y recomendaciones para presentarlas al CCCF en su conjunto.
160. Si la exposición dietética estimada para UB está muy por debajo del valor de referencia basado en efectos sobre la salud (HBGV), incluso sin un NM, y el NM propuesto se sitúa en o cerca del valor del LC, el NM tendría poco impacto en la reducción de la exposición dietética, por lo que no sería necesario. Para aquellos contaminantes para los que no se ha establecido un HBGV, si todos los datos están censurados por la izquierda, se podría recomendar al CCCF establecer temporalmente el NM en el valor de LC en caso de que haya posibles motivos de preocupación para la salud. Sin embargo, si la mayoría de los datos están censurados por la izquierda y no hay o apenas hay motivos de preocupación para la salud, se podría indicar al CCCF que no es necesario establecer NM. Por ejemplo, un conjunto de datos combinado de plomo en huevos de gallina frescos contenía, después de la depuración de datos, un 99 % de puntos de datos censurados por la izquierda comprendidos entre 0,001 y 0,257 mg/kg. La reducción calculada de la exposición con un NM hipotético fue baja, y el NM propuesto se situaba en el rango de los LC notificados. Por lo tanto, se interrumpió el desarrollo de NM para el plomo en huevos de gallina (véase CX/CF 22/15/7).

161. Si la exposición dietética estimada para UB está cerca o supera el HBGV, o si el margen de exposición es bajo, se podría recomendar al CCCF desarrollar un NM aunque un NM propuesto esté cerca de los LC notificados, siempre y cuando exista un método o métodos analíticos validados con un LC adecuado. Si es necesario, se puede recomendar emitir peticiones de datos adicionales utilizando métodos analíticos más sensibles (LC más bajos).
162. Si los valores cuantificados muestran una gran variación y alcanzan valores significativamente altos, se podría recomendar al CCCF desarrollar un NM para eliminar los alimentos altamente contaminados del mercado internacional. Si el contaminante es altamente tóxico o genotóxico/cancerígeno y se halla en alimentos que se consumen en grandes cantidades, se podría recomendar al CCCF establecer un NM para proteger la salud de los consumidores, incluso si la tasa de rechazo es baja. Por ejemplo, el conjunto de datos combinado de aflatoxinas totales en granos de sorgo contenía, después de la depuración de datos, un 94 % de datos censurados por la izquierda, y el rango superior de concentraciones cuantificadas en este conjunto de datos superaba los 200 µg/kg. Esto indica que un NM basado en una tasa de rechazo baja seguiría teniendo un gran impacto en la reducción de la exposición dietética a las aflatoxinas de los granos de sorgo. Se recomendó al CCCF un anteproyecto de NM para las aflatoxinas en sorgo con una tasa de rechazo del 0,9 %, así como una reducción de la ingesta del 46,5 % (véase CX/CF 22/15/9).
163. En general, si se evalúa la exposición dietética, los GTE encargados de proponer NM al CCCF deberán evaluar el equilibrio entre la tasa de rechazo y la reducción de la exposición dietética a partir del producto meta en cada NM hipotético y determinar el nivel más bajo que pueda alcanzarse razonablemente u ofrecer opciones al CCCF para fundamentar la decisión del Comité.
164. Si bien queda fuera del ámbito de aplicación de esta orientación determinar qué tasa de rechazo es la más apropiada, el GTE encargado de recomendar NM al CCCF también debería considerar los patrones de consumo regionales e internacionales para establecer un anteproyecto de NM propuesto entre los NM hipotéticos con respecto a la protección de la salud del consumidor y la garantía de la inocuidad de los alimentos y el comercio justo.
165. Los miembros del Codex tienen la responsabilidad de comprobar el impacto del anteproyecto de NM (o NM hipotético(s)) comparándolo con sus propios datos nacionales/regionales, así como de aportar observaciones sobre los resultados de su análisis estadístico enviado al GTE o a la reunión plenaria.

PRESENTACIÓN DE DATOS EN LOS INFORMES DEL GTE AL CCCF

Presentación del análisis de datos: análisis estadísticos

166. Es importante que los datos se presenten en el informe del GTE al CCCF como un documento de trabajo que permita un debate informado sobre los NM apropiados para su deliberación a través del procedimiento de trámites del Codex. Esto significa que los datos se notifican incluyendo todos los supuestos, por ejemplo, cuántos datos se excluyeron y las razones para ello, cómo se gestionaron los datos censurados por la izquierda, si se tuvieron en cuenta los datos no incluidos en la base de datos de SIMUVIMA/Alimentos, etc.
167. El Anexo V proporciona los elementos y los ejemplos de plantillas que puede utilizar el GTE para presentar los resultados de los análisis estadísticos de los datos de presencia para el desarrollo de NM. El GTE puede utilizar las plantillas o modificarlas caso por caso, puesto que los detalles notificados dependerán de la cantidad de datos disponibles y de la naturaleza del contaminante.