

DEPARTAMENT DE NUTRICIÓ I BROMATOLOGÍA

UNIVERSITAT DE BARCELONA

Programa de Doctorado: Nutrición, Tecnología e Higiene de los Alimentos

**DESARROLLO Y APLICACIÓN DE UN SISTEMA DE
INFORMACIÓN PARA LA ELABORACIÓN DE
TABLAS DE COMPOSICIÓN DE ALIMENTOS**

ANDREU FARRAN CODINA

Tesis Doctoral

Barcelona 2004

Directora:
Dra. Magda Rafecas Martínez

Tutor:
Dr. Rafael Codony Salcedo

A la Gemma i la Clàudia

A la meva mare i al meu pare

AGRAÏMENTS

El meu més sincer agraïment a la Dra. Magda Rafecas per la seva direcció, estímul constant i dedicació fins i tot en moments certament complicats.

Al Dr. Josep Boatella, una de les primeres persones al nostre país que va expressar la conveniència d'aprofundir en l'adequació dels valors que oferien les taules de composició d'aliments i la metodologia d'elaboració de les taules. Ell em va introduir en aquest tema i em va facilitar la possibilitat de formar-me en la producció i maneig de dades de composició dels aliments.

Al Dr. Rafael Codony, per la seva ajuda i disponibilitat a resoldre pacientment els meus dubtes.

Al CESNID i, més concretament la seva directora Sra. Pilar Cervera, per haver posat a la meva disposició els mitjans per dur a terme la major part d'aquesta feina, pel seu suport constant i, sobretot, per la seva confiança.

A en David Cantós i en Raül Zamora amb qui he compartit llargues discussions d'anàlisi i un munt de maldecaps. En David ha aportat solucions informàtiques cabdals: ha refinat l'anàlisi de l'aplicació informàtica i ha fet les tasques de programació. Tampoc s'han escapat de compartir maldecaps en Raimon Milà i la Martina Miserachs. A tots ells, gràcies per ser tan bons companys de feina.

Als companys de Greixos I, que han esta molts: Pep Pascual, Miquel Illera, Anna Jordán, Francesc Guardiola, Xavi Parcerisa, Marta Roca, Anna Gibert, Stella Grimpa, Joan Ayté i tot el personal del departament de Nutrició i Bromatologia. Ep!... i la Lourdes Ribas i en Pau Gascón. I segur que em deixo algú... disculpeu-me.

A la resta de companys del CESNID: Imma Palma (m'has avançat!), Anna Puchal, Laura Padró, Magda Reixach, Rita Rigolfas i la resta de professors. Tots ells son excel·lents companys i sovint m'han aportat el seu punt de vista o correccions. I no em vull deixar la Montse, Mercè, Maria, Cristina, Pruvi i José per tot el bon humor que desprenen.

Al Dr. Carlos Alberto González, coordinador del grup EPIC-Espanya, per facilitar-me dades sobre consum d'aliments a la cohort d'EPIC.

A la Vicki Pons i Raül Bescós, per facilitar-me les dades comparatives de l'anàlisi de menús utilitzant diferents programes informàtics.

Haig d'expressar també el meu agraïment al Departament d'Agricultura Generalitat de Catalunya per l'ajut en el projecte de recopilació de dades de productes catalans, projecte que va aportar experiències importants sobre la compilació de dades de composició d'aliments.

ÍNDICE

1	Introducción	13
1.1	Desarrollo histórico de las tablas de composición de alimentos	15
1.1.1	¿Tablas o bases de datos?	18
1.2	La necesidad de disponer de datos de composición de alimentos	19
1.2.1	Limitaciones de las BDCA	21
1.3	La elaboración de tablas / bases de datos de composición de alimentos	23
1.3.1	Tipos de datos de composición	24
1.3.2	El proceso de datos en la elaboración de BDCA	25
1.3.3	Tipos básicos de BDCA	26
1.3.4	Validación de bases de datos de composición de alimentos	27
1.4	La compilación de datos de composición	38
1.5	Las fuentes de información de datos de composición	42
1.6	Identificación de alimentos	44
1.6.1	Sistemas de clasificación de alimentos	45
1.6.2	Sistemas descriptivos de alimentos	46
1.6.3	La utilización combinada de sistemas de descripción y clasificación	50
1.7	Identificación de componentes	52
1.7.1	Los códigos INFOODS y EUROFOODS para identificar componentes	53
1.7.2	El <i>Component Aspect Identifier</i> (CAId)	55
1.8	Métodos de obtención de los valores de composición	58
1.8.1	Métodos analíticos	58
1.8.2	Estimaciones y cálculos a partir de otros componentes	59
1.8.3	El “préstamo” de valores	64
1.9	Descripción de los valores de composición	66
1.10	Evaluación de la calidad de los datos de composición	68
1.10.1	La calidad de los datos analíticos	68
1.10.2	La calidad de los datos no analíticos	70
1.10.3	Chequeos sobre los valores durante el proceso de compilación	70
1.11	Agregación de los valores de composición	72
1.12	Aplicaciones de gestión de BDCA	75
1.12.1	Modelo conceptual de la información: el modelo entidad-relación	76
1.12.2	Modelo físico de la información: las bases de datos relacionales	79
1.12.3	Los sistemas de gestión de bases de datos	79
2	Justificación y objetivos	81
3	Materiales y métodos	87
3.1	Descripción general del sistema de información	89
3.1.1	Especificaciones generales del Sistema de Información	91
3.1.2	Descripción general de los datos	92
3.1.3	Descripción general de los procesos	94
3.1.4	Organización del Sistema de Información	95
3.2	Archivo	98
3.2.1	Formulario MUESTRA	99
3.2.2	Formulario COMPOSICIÓN	105
3.2.3	Formularios para importación de datos	107
3.2.4	Campos de relación y control	109
3.2.5	Proceso detallado de compleción y revisión de los formularios	110
3.3	Normalización de valores	114
3.4	Adjudicación de códigos de calidad	116
3.4.1	Sistema de evaluación de los datos	116
3.4.2	Correcciones del sistema de evaluación de datos	117
3.5	Generación de nuevos componentes	121

3.6	Detección de errores	124
3.7	Sistema Informático.....	127
3.7.1	Especificaciones del sistema informático.....	127
3.7.2	Descripción general del sistema informático.....	128
3.7.3	Base de datos	130
3.7.4	Sistema de gestión de la base de datos	136
3.8	Agregación de datos.....	158
3.9	Estimación de los valores desconocidos.....	165
3.9.1	Estimación utilizando recetas o fórmulas	165
3.9.2	Estimación por similitud.....	166
3.9.3	Estimación por valor de grupo.....	166
3.9.4	Extrapolación del valor	167
3.9.5	Ponderaciones o recetas sencillas.....	167
3.9.6	Estimación de un valor de composición a partir de un estándar o de una regulación legal	167
3.9.7	Cero directo.....	167
3.9.8	Valores obtenidos “por diferencia”	170
3.9.9	Estimaciones lógicas.....	171
3.9.10	Datos obtenidos de etiquetas	171
3.9.11	Estimaciones diversas.....	171
3.10	Extracción de datos hacia la base de datos del usuario, revisión final y generación de nuevos alimentos	172
3.11	Elaboración de la tabla de composición de alimentos impresa	174
3.11.1	Especificaciones respecto a los alimentos	174
3.11.2	Especificaciones respecto a los nutrientes y otros componentes	174
3.11.3	Especificaciones respecto a los valores	175
3.11.4	Especificaciones respecto a la información auxiliar para el usuario.....	176
4	Resultados	179
4.1	Recopilación de documentos	181
4.2	Base de datos de referencia.....	183
4.2.1	Origen de los datos	186
4.2.2	Documentación de los datos	189
4.2.3	Calidad de los datos.....	192
4.3	Base de datos del usuario	195
4.3.1	Origen de los datos	197
4.3.2	Tipo de valor.....	198
4.3.3	Tipo de método	199
4.3.4	Calidad de los datos.....	202
4.4	Tabla de composición de alimentos	204
5	Discusión.....	205
5.1	Discusión sobre el método	207
5.1.1	Aplicación de los códigos EUROFOODS utilizados para describir los valores y los métodos	207
5.1.2	Códigos utilizados para describir los componentes	211
5.1.3	Utilización de formularios para la recopilación y primer tratamiento de los datos	213
5.1.4	Sistema de evaluación de la calidad de los datos.....	214
5.2	Discusión sobre los resultados.....	216
5.2.1	Recopilación de documentos	216
5.2.2	Base de datos de referencia	219
5.2.3	Base de datos del usuario y tabla de composición	221
5.2.4	Resultados claves del estudio.....	228

5.2.5	Limitaciones del estudio	229
5.2.6	Implicaciones de este trabajo	230
5.2.7	Sugerencias para posteriores estudios	231
6	Conclusiones	235
7	Bibliografía	239
	Anexo I	257
	Anexo II	263
	Anexo III	291
	Anexo IV	297

1 INTRODUCCIÓN

1.1 DESARROLLO HISTÓRICO DE LAS TABLAS DE COMPOSICIÓN DE ALIMENTOS

Puede considerarse que el inicio del desarrollo de la Nutrición como disciplina científica se produce cuando se empieza a estudiar la composición química de los alimentos y se intentan describir en términos de entidades o grupos de entidades químicas con propiedades nutritivas únicas (1). De este modo, surge muy pronto la necesidad de compilar los datos analíticos generados en forma de tablas de composición de alimentos (TCA) con la finalidad de disponer cómodamente de esta información. La primera TCA conocida data de 1818, aunque fueron los trabajos de Von Voit (2, citado en 3) y König ((4), citado en (3)) en Alemania, Atwater (5) en Estados Unidos y Plimmer ((6), citado en (3)) en Gran Bretaña, durante el periodo de tiempo que va de 1860 a finales de la Primera Guerra Mundial, los que consolidaron los inicios del conocimiento sobre composición de alimentos (7). Las TCA británicas iniciadas por Plimmer fueron completadas entre 1920 y 1930 por los doctores McCance y Widdowson (8) a partir de análisis realizados *ex profeso* para las tablas e introduciendo datos procedentes de análisis para los azúcares y almidón (hasta la fecha las TCA incluían datos de carbohidratos calculados “por diferencia”). El Departamento de Agricultura de EEUU, responsable de las TCA norteamericanas, abandonó la realización de análisis de alimentos en la década 1940-50 y adoptó el método de recopilar datos ya existentes, procedentes de la industria alimentaria, la literatura científica y contratos puntuales para la realización de análisis (8,9). A mediados de la década de los 80 se estableció la realización de un programa de análisis de alimentos para incrementar la cantidad y calidad de los datos generados a través de los contratos establecidos externamente (10,11).

A lo largo del siglo XX se han ido desarrollando otras TCA, la mayoría de ellas de ámbito nacional. Actualmente existen más de 450 TCA en todo el mundo (12). Las tablas norteamericanas y británicas han continuado siendo un referente mundial y sus diferentes versiones y ediciones han constituido una fuente de datos muy a menudo utilizadas en la elaboración de otras tablas (13-17).

El desarrollo y mantenimiento de TCA no es fácil y sí muy costoso. Esto llevó a que en la década de 1980-90 diversos investigadores que trabajaban en la compilación de datos de composición iniciaran una creciente actividad de colaboración internacional con un claro objetivo principal: poder intercambiar datos de composición (18-21). Tal actividad condujo a la creación de un organismo internacional que funciona como una red de colaboración bajo los auspicios primero de la *United Nations University* (UNU) y diversas agencias y organismos norteamericanos (20) y, actualmente, de la *Food and Agriculture Organisation* y

la UNU (22). Este organismo se denomina *International Network of Food Data Systems* (INFOODS) y fue creado formalmente en un reunión mantenida en Italia el año 1983 (23).

Para poder alcanzar el objetivo principal, INFOODS se organizó en grupos de trabajo orientados a aportar soluciones a aquellos aspectos que dificultaban el intercambio de datos. Las tareas que cada grupo debía completar fueron las siguientes (24,25):

- Desarrollar unas guías internacionales que establecieran cómo deben recopilarse los datos de composición de alimentos.
- Desarrollar una terminología y nomenclatura estandarizada.
- Explorar como podían utilizarse los sistemas de información en el campo de la composición de alimentos.
- Producir un directorio internacional de bases de datos de composición de alimentos.
- Crear grupos regionales de colaboración dentro de la red INFOODS.
- Desarrollar una descripción detallada de las necesidades de los usuarios de BDCA.

Posteriormente se añadieron nuevas tareas a estas iniciales (26):

- Desarrollar un sistema para identificar los componentes (nutrientes o no nutrientes) de los alimentos.
- Desarrollar y mantener bases de datos de composición de alimentos regionales.

Los grupos de trabajo dieron lugar a una serie de guías y sistemas de uso internacional (9,15,27-29) o bien regional (30-33) cuya aplicación ha sido desigual pero que sin duda han aportado avances importantes en el conocimiento y la colaboración internacional sobre el tema (30,34,35). Otros organismos como la FAO han desarrollado sus propias iniciativas internacionales en el campo de la composición de alimentos, especialmente en países en vías de desarrollo, en colaboración con otras instituciones incluyendo INFOODS (36).

En España, los primeros trabajos dirigidos a establecer la composición de los alimentos consumidos de los que se tiene constancia son dos tesis doctorales publicadas el año 1932 (37,38). Posteriormente se han ido sucediendo la publicación de diferentes tablas de composición elaboradas por diferentes autores (Tabla 1-1) aunque no hubo ninguna iniciativa en el sentido de establecer una TCA oficial hasta la publicación en 1995 de las TCA del Ministerio de Sanidad (39,40).

Tabla 1-1 *Relación de TCA españolas publicadas hasta la fecha. No se han incluido las obras de carácter más divulgativo o las traducciones literales de obras extranjeras.*

- Comenge M. Principios bioquímicos de dietética normal y tablas de composición de los alimentos españoles. Madrid: A.G.I., 1946.
 - Alonso Samaniego JM. Factores dietéticos y tablas de composición de alimentos. Madrid: Alter - Departamento de Investigación, 1951.
 - Casares Lopez R, García Olmedo R, Valls Pallés C. Tratado de bromatología (5ª ed). Madrid, Universidad Complutense, 1978. *Contiene unas tablas de composición elaboradas a partir de datos de la FAO, la Escuela de Bromatología y las Fuerzas Armadas. La primera edición es de 1942.*
 - Andújar Arias M, Moreiras Varela O, Gil Extremera F. Tablas de composición de alimentos. Madrid, Instituto de Nutrición (CSIC). 1980. *Se han publicado sucesivas ediciones con algunos cambios en la lista de autores. La última edición ha sido de 2003.*
 - Marcos A, Fernández Salguero J, Esteban A, León F, Alcalá M, Beltrán de Heredia FH. Quesos españoles. Tablas de composición, valor nutritivo y estabilidad. Córdoba: Departamento de Tecnología de los Alimentos - Universidad de Córdoba, 1985.
 - Vivanco F, Palacios JM. Tabla de composición de alimentos españoles. Dirección General de Salud Pública - Ministerio de Sanidad y Consumo, 1985.
 - Casamitjana N. Taula de composició d'aliments per a ús clínic. Fundació Sardà Farriol. 1986.
 - Jiménez Cruz A, Cervera Ral P. Tabla de composición de alimentos. Barcelona: Wander SAE, 1988. *Se han continuado publicando en sucesivas ediciones bajo patrocinio de la empresa Wander, que pasó a llamarse Sandoz y posteriormente Novartis.*
 - Mataix Verdú J, Mañas Almendros M, Llopis González J, Martínez de Victoria E. Tabla de composición de alimentos españoles. Granada: Universidad de Granada, 1993. *Se han publicado 4 ediciones, la última en 2003.*
 - Martín Peña G. Tabla de composición de alimentos (versión 2.1). Madrid: Nutricia, 1997.
 - Ministerio de Sanidad y Consumo. Tablas de composición de alimentos españoles. Madrid: Ministerio de Sanidad y Consumo, 1995. *Se han publicado sucesivas ampliaciones, la última en 1999.*
 - Bello Gutiérrez J, Candela Delgado M, Astiasarán Anchía I. Tablas de composición para platos cocinados. Madrid: Díaz de Santos, 1998.
-

La evolución de las TCA se ha visto marcada por las necesidades que han ido surgiendo en el campo de la Nutrición. Según Dwyer (41), pueden establecerse tres etapas en la evolución del conocimiento sobre la composición de los alimentos. La primera se iniciaría con la descripción química y fisiológica de los compuestos que aportan energía. La segunda etapa está marcada por la necesidad de disponer de datos sobre composición química en minerales y vitaminas. La demanda de datos de composición de mejor calidad por parte de los investigadores que usan estos datos para poder establecer relaciones más claras entre

alimentación y enfermedad marca la tercera y actual etapa (42). La cuarta etapa estaría iniciándose en la actualidad, con la necesidad de disponer de datos de composición en compuestos químicos que no son nutrientes pero que pueden afectar la salud humana (43-46).

1.1.1 ¿Tablas o bases de datos?

Las TCA fueron durante muchos años la principal fuente de datos de composición de alimentos. Su elaboración era extremadamente laboriosa ya que todos los registros con los datos de composición e información complementaria (muestreo, tratamiento de la muestra, métodos analíticos, etc.) se almacenaban en formato impreso. La actualización de la TCA o el recálculo de algunos valores era una tarea tan importante que normalmente los datos anticuados permanecían más tiempo de lo recomendable dentro de la TCA (3).

La expansión de la informática vino a cambiar esta situación al introducir el formato electrónico para almacenar datos y los programas de gestión para manejarlos con mucha más rapidez. De este modo se pudo almacenar la información en un espacio mucho menor y gestionar grandes cantidades de información incorporando automatismos que hacen innecesaria la participación directa del compilador. Incluso el producto final del trabajo de compilación de datos ha dejado de ser una tabla de composición impresa y se trata ya de una base de datos que puede ser abierta y gestionada con aplicaciones habituales de ofimática o bien consultada en línea (47,48).

El uso de las herramientas informáticas se ha extendido tanto que hoy en día la literatura se refiere prácticamente toda a bases de datos de composición de alimentos (BDCA). Sin embargo, los principios sobre gestión de datos que se aplican a TCA y BDCA son exactamente los mismos ya que ambas son colecciones de datos ordenadas aunque sobre diferente soporte. En el presente texto se utilizará en todo momento el término “bases de datos” reservando el de “tablas” para el formato impreso de las bases de datos.

1.2 LA NECESIDAD DE DISPONER DE DATOS DE COMPOSICIÓN DE ALIMENTOS

La Nutrición tiene un objetivo básico: el de proporcionar un mejor estado de salud a la población. Para conseguirlo tiene que evaluar la alimentación de los individuos o poblaciones para establecer cuales son sus necesidades y, conociendo éstas, planificar su alimentación (9,49,50). Normalmente, una vez se ha elaborado esta planificación e intervenido sobre el individuo o población, es preciso volver a evaluar su alimentación (51).

Muchas veces la alimentación de los individuos se tiene que evaluar en relación a una enfermedad determinada. Esto puede hacerse estimando la ingesta habitual de alimentos usando métodos fiables, describiendo la alimentación en términos de consumo de alimentos y explorando las posibles relaciones causales entre enfermedad y alimentación (44). El estudio de los patrones de consumo de alimentos de las poblaciones puede ofrecer información de interés. Sin embargo, también es necesario describir la alimentación en términos de ingesta de nutrientes: los nutrientes contenidos en los alimentos suelen ser muy importantes cuando estudiamos la causalidad biológica ya que pueden estar directamente relacionados con los fenómenos fisiopatológicos que se estén observando (44,52). Para calcular y evaluar la ingesta de nutrientes a partir de los alimentos consumidos es preciso disponer de datos de composición (51,53). Esta evaluación puede ser necesaria a nivel de individuo, de grupos de población, nacional o incluso internacional (3,15,54):

1. **Individual.** Suele ser necesaria para determinar si la alimentación del sujeto es equilibrada o cumple con alguna limitación impuesta por una prescripción dietética.
2. **Grupos de población.** El estudio comparativo de diferentes grupos de población puede ayudar a establecer la relación entre la alimentación y una determinada patología o algún índice de interés sanitario. Los grupos se pueden establecer según diferentes criterios: fisiológico, socio-económico, clínico, intervención, etc.
3. **Nacional.** Los datos estadísticos nacionales de producción agrícola, ajustados por las exportaciones, importaciones, usos no alimentarios y pérdidas se multiplican por la composición nutricional y se dividen por la población nacional para estimar la disponibilidad bruta de nutrientes per cápita. Los resultados permiten establecer la adecuación bruta del suministro de alimentos y son utilizados en el diseño de las políticas alimentarias nacionales.

4. **Internacional.** La comparación de datos procedentes de diferentes países puede ser útil para formular políticas nutricionales y alimentarias internacionales así como dilucidar mejor las relaciones entre dieta y enfermedad al disponer de datos correspondientes a patrones de consumo diferentes.

La planificación de la alimentación se puede realizar actuando a diferentes niveles pero siempre consiste en intentar proporcionar al individuo o población los nutrientes que precisa para un óptimo estado de salud a través del consumo de alimentos. Por lo tanto, es preciso decidir qué alimentos son los más adecuados para conseguir modular la ingesta de nutrientes y para ello se precisará conocer su composición. La planificación puede consistir en la prescripción de dietas terapéuticas, la planificación de dietas institucionales, la elaboración de políticas nutricionales y alimentarias, el establecimiento de regulaciones legales o la elaboración de programas de intervención.

1. **Prescripción de dietas terapéuticas.** Una dieta terapéutica ha de ser nutricionalmente equilibrada al tiempo que debe cumplir con la limitación impuesta por la prescripción en la ingesta de uno o varios componentes (nutrientes o no). El diseño de tales dietas requiere una preparación profesional adecuada y datos de composición (55).
2. **Planificación de dietas institucionales.** Muchos colectivos precisan de menús nutricionalmente equilibrados que, al mismo tiempo, cumplan con limitaciones de costos (56,57). Esta información también puede utilizarse para informar a los usuarios de los perfiles nutricionales de los menús (58).
3. **Elaboración de políticas nutricionales y alimentarias.** Estas políticas establecen a menudo objetivos para la ingesta de determinados nutrientes (59). Tales objetivos se traducen en objetivos de producción para el sector agrícola o en objetivos de consumo de alimentos para el mercado alimentario o el sector de salud pública (promoción de determinados alimentos).
4. **Regulaciones legales en el ámbito alimentario.** Los datos de composición de alimentos se utilizan para establecer estándares y máximos o mínimos legales (60), en la regulación legal del etiquetado nutricional o bien en su cálculo (61-64), y para determinar la equivalencia nutricional de un alimento que tiene que reemplazar a otro (65,66).
5. **Elaboración de programas de intervención.** Los programas de intervención orientados a la suplementación, suministro de alimentos u otras acciones (educación nutricional,

etc.) precisan que las necesidades nutricionales específicas se traduzcan en requerimientos alimentarios (41,67).

Además de los usos citados las tablas de composición son necesarias también en otros ámbitos:

1. **Laboratorio de análisis de alimentos.** Cuando se realizan análisis de alimentos es útil conocer los datos de composición tanto para ajustar las condiciones analíticas como para contrastar los resultados de los análisis.
2. **Diseño de nuevos productos alimentarios.** El diseño de nuevas fórmulas para productos alimentarios que sean equivalentes a los que han de sustituir o el diseño de nuevos productos que cumplan con ciertas especificaciones nutricionales precisan disponer de datos de composición de alimentos (60,68).
3. **Investigación en tecnología de los alimentos.** Las mejoras en las técnicas de producción y transformación agrícolas y ganaderas puede tener como objetivo introducir mejoras en el perfil nutricional de los alimentos. Naturalmente, esto debe partir de un buen conocimiento de su composición (60).
4. **Docencia.** La docencia en el ámbito de la Dietética, la Nutrición o la Tecnología de los Alimentos precisa de datos de composición de alimentos para realizar simulaciones de carácter educativo o para la simple consulta (7,67).

1.2.1 Limitaciones de las BDCA

La utilización de BDCA no está exenta de limitaciones, las cuales se entienden mejor cuando se conocen las dificultades de su elaboración (69). La primera limitación la impone el alto número de alimentos existentes en el mercado cuya cifra puede llegar a 100.000 en los países desarrollados (49). Es imposible incluir todos los alimentos que se consumen y, por lo tanto, es preciso dar prioridad a aquellos alimentos que se consumen más frecuentemente o que contribuyen en mayor medida a la ingesta de un componente (51,70) .

En segundo lugar, los alimentos tienen una composición que es variable. Esta variabilidad es debida a diferentes factores pero que pueden agruparse en tres grupos: factores genéticos, factores ambientales y factores relacionados con la transformación y conservación del alimento. Las BDCA reflejan una composición determinada en un momento

y lugar determinados, aunque el usuario de los datos actúa como si no fuera así (71). Normalmente los cambios en los alimentos del mercado son más rápidos de lo que puede reflejar una BDCA (72).

La tercera limitación es la cobertura de componentes. No siempre se dispone de recursos o de las técnicas analíticas adecuadas para analizar todos los posibles componentes de interés. De nuevo es necesario dar prioridad a aquellos componentes que sean de interés en salud pública y para los cuales existan métodos analíticos fiables y practicables (70,73).

Por último, las BDCA ofrecen datos representativos de los alimentos consumidos por una población pero estos datos nunca podrán predecir con exactitud la composición de una muestra concreta de alimento. Sin embargo, sí podrá predecir con aceptable exactitud la composición media de un alimento consumido por un grupo de individuos. Por esta razón, las BDCA encuentran su utilización óptima cuando se utilizan para evaluar la ingesta de poblaciones de individuos o bien cuando se evalúa la ingesta de un individuo durante un periodo largo de tiempo. En el resto de usos las BDCA se utilizan con un valor semi-cualitativo y el usuario ha de entender que los cálculos realizados con las tablas tienen un valor aproximado. Cuando se realizan determinados estudios clínicos que requieren una alta exactitud o estudios de balance metabólico, las BDCA no son apropiadas y es necesario recurrir a análisis químicos directos (3,71).

1.3 LA ELABORACIÓN DE TABLAS / BASES DE DATOS DE COMPOSICIÓN DE ALIMENTOS

Southgate (74) distingue dos métodos básicos para la elaboración de tablas / bases de datos de composición de alimentos:

Método directo. Consiste en la elaboración y ejecución de un plan de muestreo de alimentos y su análisis. El control sobre la selección de las muestras y los métodos analíticos es total permitiendo obtener datos de una alta fiabilidad. El inconveniente es que se trata de un método costoso tanto en término de recursos como de tiempo, convirtiéndose en un método difícil de asumir en muchos países.

Método indirecto. Consiste en la recopilación de datos ya existentes procedentes de fuentes publicadas y no publicadas. No se trata de un método fácil de llevar a cabo ya que requiere la evaluación crítica de los datos recopilados, la mayoría de los cuales se han obtenido con un objetivo diferente al de obtener una composición representativa de los alimentos consumidos por la población. Por lo tanto, es preciso determinar si aspectos como el método de cultivo o de cría y las variedades o razas son los habituales en el país, o la representatividad de las muestras analizadas. También es preciso evaluar el tratamiento de la muestra y los métodos analíticos utilizados, para comprobar que se adecuan a los considerados válidos. Una vez seleccionados qué valores pueden utilizarse es preciso determinar como obtener un valor representativo: no siempre la media de los valores es la mejor opción. Aunque el método indirecto no requiere el muestreo de alimentos y la infraestructura analítica del método directo, la cantidad de tiempo que comporta el escrutinio de los datos incrementa considerablemente los costes económicos.

Posteriormente, Greenfield y Southgate (9) añadieron un tercer método que en realidad se trata de la combinación de los dos anteriores:

Método combinado. Hoy en día es el mayoritariamente utilizado. Consiste en la obtención de datos de composición a través del método directo para los alimentos de mayor consumo por parte de la población o bien de aquellos que constituyen una fuente principal de algún nutriente, mientras que para los alimentos menos prioritarios los datos se obtienen a través del método indirecto.

1.3.1 Tipos de datos de composición

Aunque la mayor parte de los datos que se utilizan en las BDCA se generan a partir de la aplicación de métodos de análisis químico, normalmente no es posible obtener todos los datos de composición a partir de éstos. En tal caso, es posible utilizar otros métodos para generar valores que se pueden incluir en la TCA para evitar la presencia de valores desconocidos, especialmente en aquellos alimentos que son contribuyentes menores a la ingesta (75). Por lo tanto, es preciso distinguir entre diferentes valores de composición en función de cómo han sido obtenidos (3,53):

1. **Valores analíticos originales.** Son los datos obtenidos a partir de análisis químicos, ya sean realizados expresamente para la TCA o bien obtenidos de fuentes ya existentes. Es posible incluirlos en la TCA tal y como se han obtenido o pueden agregarse (cálculo de la media, la mediana, ponderaciones, valor central, etc.) para obtener valores más representativos. En este grupo se incluyen también aquellos valores que en realidad han sido calculados a partir de datos analíticos utilizando factores de conversión (por ejemplo, la proteína bruta a partir del nitrógeno total o los ácidos grasos a partir de los porcentajes de ácidos grasos en 100 gramos del total de ácidos grasos).
2. **Valores imputados.** Estos datos son estimaciones que se han obtenido a partir de datos analíticos aplicando un razonamiento que implica la asunción de unas premisas que pueden ser erróneas. La utilización de datos de un alimento similar, el cálculo de algunos valores a partir de una composición parcial (por ejemplo los carbohidratos por diferencia, el cloro calculado a partir del sodio) o el cálculo de una forma deshidratada de un alimento a partir de su forma hidratada son ejemplos de valores imputados.
3. **Valores calculados.** Se trata de valores calculados a partir de la receta y la composición de sus ingredientes, aplicando o no factores de corrección para tener en cuenta los cambios en la preparación y cocción de los alimentos. Existen diferentes tipos de algoritmos utilizados para calcular tales valores (76-78). Tales valores deben considerarse como aproximaciones ya que la lista de ingredientes, cantidades, preparación y condiciones de cocción pueden variar de manera importante.
4. **Valores prestados.** Son aquellos valores que se han tomado de otras TCA que no ofrecen información sobre el origen del dato. El escrutinio que se puede hacer sobre el dato es prácticamente nulo y la decisión se basa entonces en el prestigio de la TCA

utilizada como fuente. En algunos casos, los valores prestados tienen que ajustarse por el contenido en agua o lípidos.

5. **Valores asumidos.** Son aquellos valores que se han establecido únicamente a partir de una presunción. En esta categoría se incluyen los ceros asumidos (por ejemplo, adjudicar el valor cero al colesterol en los alimentos del grupo de las frutas frescas).

1.3.2 El proceso de datos en la elaboración de BDCA

La recopilación y tratamiento de datos que se realiza con el fin de elaborar una BDCA se realiza en varias etapas que se corresponden con cuatro niveles diferentes de tratamiento de la información (9,15,30,78).

Nivel 1: Fuentes de datos. Se trata de los documentos (en formato impreso o electrónico) que contienen datos de composición que se han generado mediante el método directo o bien los documentos que se han recopilado mediante el método indirecto. Se puede tratar de bases de datos, material publicado (artículos, libros, folletos), boletines de laboratorio, literatura gris, etiquetas de productos, etc.

Nivel 2: Archivo de datos. Los datos de composición se extraen de los documentos originales y se comprueba su consistencia para introducirlos en registros escritos (fichas o formularios) o bien informáticos. Estos registros han de contener toda la información necesaria para poder tomar las decisiones más adecuadas en el tratamiento posterior de los datos y también han de permitir que sea innecesario la consulta de las fuentes originales en caso de que se precise alguna aclaración. Por lo tanto, además de los valores de composición, deben informar también acerca de (78):

- unidades y modos de expresión
- cálculo del valor (por ejemplo, factores de conversión utilizados)
- método de muestreo
- número de muestras analizadas
- métodos analíticos utilizados
- información acerca de los procedimientos de aseguramiento de la calidad utilizados
- referencias bibliográficas de las fuentes
- el resultado del proceso de evaluación de los datos, si éste se realiza

Nivel 3: Base de datos de referencia. Está formada por la reunión de datos que ya han sido sometidos a escrutinio, convertidos a unidades estándar y expresados uniformemente, aunque los datos procedentes de diferentes fuentes se mantienen separados y sin agregar. Cada dato está ligado a la información referente al muestreo, métodos analíticos, laboratorio de origen, fecha de inserción, referencias bibliográficas y otra información relevante.

La base de datos de referencia suele formar parte de un sistema informático de gestión de base de datos, con programas informáticos o protocolos escritos para operar con la información almacenada. Los datos referentes a los factores, cálculos y recetas utilizadas también deben almacenarse en esta base de datos.

Nivel 4: Base de datos del usuario (tablas y bases de datos informáticas). A partir de la base de datos de referencia se genera la base de datos para los usuarios o bien la versión impresa (tablas de composición). Los datos de la base de datos de referencia se agregan para obtener valores representativos. Las bases de datos suelen contener más información y más detallada que las tablas de composición y presentar algún indicador de la calidad de los datos basado en los procedimientos de muestreo y analíticos.

Las bases de datos y tablas de composición del usuario han de incluir cuantos más alimentos y nutrientes pero dando prioridad a ofrecer datos completos (sin valores desconocidos). Para conseguirlo, a menudo es preciso recurrir a valores imputados, calculados, prestados o asumidos (13,14,51). Se debería incluir información sobre los procedimientos de muestreo, métodos analíticos, estimación o cálculo, y fuentes bibliográficas al nivel de nutrientes para que el usuario pueda realizar comparaciones con otras bases de datos o tablas de composición (3,15,79). Las bases de datos o tablas de composición pueden adaptarse a las necesidades de los usuarios realizando versiones simplificadas o bien adaptadas a propósitos especiales (14,15,80).

1.3.3 Tipos básicos de BDCA

Existen dos tipos básicos de BDCA (15):

1. **Las BDCA de referencia.** En este caso se trata de ofrecer una compilación lo más exhaustiva posible de datos de composición de alimentos consumidos por una determinada población (ejército, pacientes hospitalarios, habitantes de un país, etc.). En las BDCA de referencia predominan las fuentes originales de datos y se intentan incluir pocos datos “prestados” y estimados.

2. **Las BDCA destinadas a un propósito especial.** Se trata de bases de datos que se basan en una BDCA de referencia. Puede tratarse de bases de datos *ampliadas* o bien *reducidas*. Las *BDCA ampliadas* amplían la lista de alimentos y, por ello, precisan incorporar datos procedentes de fuentes diversas, incluyendo estimaciones. Suelen estar destinadas a tareas que requieren una larga lista de alimentos y un mínimo de valores desconocidos (13,16,53). Las *BDCA reducidas* intentan simplificar la lista de alimentos de la BDCA de referencia. Se trata de bases de datos que se han elaborado a partir de una BDCA de referencia de gran tamaño y que intentan ofrecer un nivel de exactitud similar a la base de datos “madre” pero con un número reducido de alimentos. Se trata de un caso muy concreto discutido ampliamente por Windham y colaboradores (14).

1.3.4 Validación de bases de datos de composición de alimentos

Una de las preguntas que el usuario puede realizarse es hasta que punto merecen credibilidad los datos que ofrece una BDCA (“¿los datos son buenos/correctos?”), o bien hasta que punto los valores obtenidos mediante el cálculo usando la BDCA se aproximan al valor verdadero. Aunque son dos preguntas que aparentemente tratan de lo mismo y cuyo planteamiento parece sencillo, es preciso analizar bien el contexto antes de establecer una posible respuesta a cada una de ellas.

Rand (50) señala que para definir adecuadamente qué debe entenderse por la “bondad de los datos” es preciso definir previamente cuales son las relaciones entre usuario y datos. La adecuación de los datos sólo puede evaluarse en función del fin al cual se destinan y esto depende de las actividades en las que se halle implicado el usuario. Por lo tanto, la bondad de un dato tiene que incluir el hecho de que sea adecuado para un propósito (una predicción o la toma de decisiones). El autor no define cuales son las otras propiedades que definen la bondad de los datos, aunque concluye que la variabilidad en la composición de los alimentos tiene que ser tenida en cuenta en la elaboración de BDCA y conocida por parte del usuario para asegurar que el dato es valorado y utilizado adecuadamente.

En la literatura sobre el tema se menciona a veces la “validación” de las BDCA (15,81,82), pero sin definir qué se pretende significar con el término utilizado. Según la Real Academia Española, validación es la “acción y efecto de hacer creíble una cosa” (83). En el ámbito de la química analítica se define como la “obtención de pruebas documentadas que demuestran que un método o proceso es lo suficientemente fiable como para producir el

resultado previsto dentro de intervalos definidos” (84). Si admitimos estas definiciones como aplicables a las BDCA quedaría determinar qué pruebas documentales es preciso obtener para considerar que una BDCA es fiable o creible. A este respecto ha habido diferentes aproximaciones que no son necesariamente excluyentes entre ellas.

En los apartados siguientes se revisan en primer lugar las diferentes aproximaciones a la obtención de pruebas documentales que avalen la fiabilidad de una BDCA (véase el apartado 1.3.4.1). En segundo lugar, el desarrollo de métodos para evaluar la calidad de los datos de composición (véase el apartado 1.3.4.2). En último lugar, se revisa el papel de los estudios que comparan los datos analíticos y los datos calculados con BDCA en la evaluación de la ingesta de individuos para comprobar la exactitud de la estimación efectuada con BDCA (véase el apartado 1.3.4.3).

1.3.4.1 Validación de BDCA

Rand y colaboradores (15) señalan que el concepto de validación aplicado a las BDCA es complejo y que debe incluir los siguientes requisitos:

- los números representan correctamente las concentraciones de nutrientes en los alimentos analizados
- los alimentos analizados son representativos de los alimentos listados
- los alimentos y nutrientes están identificados con precisión y cuidado
- la fuentes de los datos está bien identificada
- los datos no están representados con más precisión de la que tienen realmente; esto es, se incluyen los dígitos significativos que la variabilidad de los datos permite y no más (85)

Por otro lado, Schakel y colaboradores (81) utilizan el término “validación” aplicado tanto a datos como a métodos de estimación de datos de composición. En el primer caso se trata en realidad de comprobación de la consistencia de valores usando el término tal y como se utiliza en informática (86), y la documentación de la consistencia de los valores sólo es una parte de la validación de una BDCA (14). En el segundo caso, añade una consideración que

no tienen en cuenta Rand y colaboradores: los métodos de estimación de valores tienen que ser validados igual que los métodos analíticos (15,87).

Greenfield y Southgate (3) establecen 10 criterios que deben cumplir las BDCA para su aplicación en el ámbito de la Nutrición (Tabla 1-2), pero en ningún momento utilizan el término “validación” más que para referirse a la validación de los métodos analíticos. Tampoco se propone su utilización como criterios para demostrar la fiabilidad de una BDCA. Entre los 10 criterios no se incluye mención alguna a la validación de los métodos de estimación.

Tabla 1-2 *Criterios que deben cumplir las BDCA para su uso en Nutrición, según Greenfield y Southgate (3).*

1. Los datos tienen que ser representativos
2. Los datos tienen que ser de una calidad analítica válida
3. La lista de alimentos tiene que ser extensa
4. La lista de nutrientes tiene que ser extensa
5. Las descripciones de los alimentos tienen que ser claras
6. Los datos tienen que estar expresados de una manera consistente y sin ambigüedades
7. El origen de los datos tiene que proporcionarse a nivel de valores de composición
8. Las TCA y BDCA tienen que ser fáciles de usar
9. El contenido de diferentes BDCA tiene que ser compatible
10. Las BDCA tienen que tener pocos valores desconocidos

Puede considerarse que esta lista incluye, de una manera u otra, los criterios listados por Rand y colaboradores (15). El último criterio que establece que los datos no deben representarse con más precisión de la que realmente pueden ofrecer estaría incluida en el criterio número 6 de Greenfield y Southgate (3). Los criterios establecidos por Greenfield y Southgate se han utilizado como guía de los puntos a auditar en una auditoría realizada sobre la BDCA Neozelandesa (88).

Parpinel y colaboradores (82) proponen la validación de la BDCA italiana para uso en epidemiología llevando a cabo un muestreo representativo de los alimentos más consumidos en Italia (como mínimo consumidos una vez por semana) y realizando su análisis para todos los nutrientes incluidos en la BDCA. Los resultados obtenidos se compararían con los datos en la BDCA a través de un análisis de correlaciones. Según los autores, esto permitiría establecer si los métodos de compilación utilizados han producido una BDCA fiable.

Feinberg y colaboradores (89) establecen un proceso de validación en el cual cada dato agregado, con toda la documentación necesaria, es sometido al escrutinio por parte de un grupo de expertos para que determinen su validez desde un punto de vista analítico y nutricional (solo se mencionan estos dos). Se entiende que la base de datos surgida de la validación de cada dato será una BDCA validada. Aunque este esquema de trabajo se aplicó en dos de las tablas francesas (90,91), no se tiene constancia de que se haya seguido aplicando (17).

Murphy y colaboradores (92) realizan la validación de una BDCA, elaborada especialmente para la evaluación de la ingesta nutricional de una población rural de Kenya, a través de varias etapas. La primera consiste en la comprobación de la consistencia de los valores. La segunda consiste en el cálculo de 11 platos a partir de sus ingredientes y la comparación con los resultados obtenidos analíticamente. Finalmente, utilizan los valores aberrantes obtenidos del cálculo de las ingestas de nutrientes obtenidos con la BDCA para localizar posibles errores en la BDCA. También proponen completar la validación realizando la comparación de varias ingestas diarias determinadas analíticamente o bien calculadas con la BDCA, aunque no llegan a realizarla.

1.3.4.2 Códigos de calidad

La evaluación de la calidad de los valores partiendo de criterios pre-establecidos fue una estrategia propuesta a inicios de los años 80 en Estados Unidos (3,93) y que ha sido desarrollada y aplicada principalmente en este país (94-98), siendo pocas las experiencias en otros países (99-102). Consiste en establecer unos criterios frente a los cuales se evalúa la calidad de los valores de composición; por ejemplo el plan de muestreo, el número de muestras, la manipulación de la muestra, el método analítico y el control de calidad analítico (95). Dependiendo del grado de cumplimiento de estos criterios, los valores recibirán una puntuación menor o mayor. Las puntuaciones se traducen en unos códigos de calidad (o códigos de confianza) que distinguen aquellos valores en los que se puede depositar más confianza de los que son dudosos. El método depende en gran medida de que la fuente de los datos proporcione la documentación necesaria para confrontarlos a los criterios (un valor no documentado recibirá una puntuación baja independientemente de su calidad real), se trata de un método que no deja de ser arbitrario y los detalles de los criterios generales mencionados tienen que desarrollarse específicamente para cada nutriente. Además, su aplicación no es sencilla y puede estar sujeta a cierta subjetividad. Esto ha hecho necesario el desarrollo de sistemas expertos para la evaluación de los datos (63).

En principio, estas puntuaciones o códigos de calidad o confianza pueden estar a disposición del usuario y podrían utilizarse como indicador de la calidad relativa de un valor y de la confianza que el usuario puede depositar en él (95), siempre teniendo en cuenta que las puntuaciones son categorías y no números reales (71) y que la calidad depende también del uso que se pretenda hacer del dato (15,71). Greenfield y Southgate remarcan que la utilización de métodos analíticos válidos y la caracterización estadística de la composición de los alimentos son, en última instancia, los factores esenciales para determinar qué confianza puede adscribirse a un valor (3). En todo caso, es preciso señalar que no se ha planteado nunca la utilización de estos códigos como información única y suficiente para establecer la validez de una BDCA, y que las experiencias desarrolladas hasta la fecha con tales códigos de calidad los han revelado como una herramienta muy útil para el escrutinio sistemático de los datos y para sintetizar información diversa en un único dato.

Por otro lado, el concepto de “calidad de los datos” ha recibido un enfoque diferente, más generalista, en las recomendaciones elaboradas por Scholtke y colaboradores dentro de EUROFOODS (79,103). Según este enfoque, la calidad de los datos se define a través de un número de criterios de calidad los cuales pueden desglosarse en subcriterios más detallados. Los criterios principales señalan que los datos tienen que ser *accesibles*, *útiles*, *interpretables* y *creíbles*. Para ser *útiles*, los datos tienen que ser *oportunos* y *relevantes*. La *relevancia* la establece el usuario dependiendo de la aplicación a que destine los datos. Para que un dato sea *creíble* necesita ser *completo*, *consistente* y *exacto*. La *credibilidad* se conseguirá a través de controles de calidad y procedimientos de aseguramiento de la calidad (3). En cualquier caso, es importante documentar adecuadamente los datos para conseguir que los datos sean *interpretables* y que el usuario pueda decidir la *relevancia* del dato respecto a una aplicación concreta. Esta información descriptiva que ayuda a documentar los datos son los llamados “metadatos” (79).

1.3.4.3 Comparación con resultados analíticos de dietas o ingestas

Diferentes investigadores han realizado estudios para comparar los resultados obtenidos de la evaluación de la ingesta de nutrientes en grupos de población utilizando datos procedentes de TCA o BDCA y datos obtenidos directamente del análisis de los alimentos consumidos (53,104). Según Gibson (53) estos estudios permiten validar el cálculo de las ingestas de nutrientes con una determinada BDCA. Han sido utilizados extensamente para intentar comprobar la credibilidad de las estimaciones obtenidas con TCA o BDCA y merecen un tratamiento detallado pues ofrecen información interesante, pero al mismo

tiempo presentan ciertos puntos débiles debido a que son mucho más difíciles de llevar a cabo de lo que puede parecer a simple vista.

En la Tabla 1-3 se recogen los resultados obtenidos por algunos de los estudios más recientes (53,104-110). Como puede verse, éstos pueden variar de manera muy importante y no siempre los macronutrientes son los nutrientes que presentan diferencias menores, como cabría esperar de su menor variabilidad en los alimentos (51,53). Estas grandes diferencias entre resultados de diferentes estudios pueden quedar explicadas en parte por diferencias o problemas en el método:

1. **Los valores desconocidos se han substituido directamente por cero.** Idealmente la BDCA utilizada no debería tener valores desconocidos o tenerlos en proporciones muy bajas y para alimentos poco importantes. Sin embargo, muchas BDCA presentan bastantes valores desconocidos y dejan para el usuario la tarea de completarlos. Cuando se realizan cálculos con ellas es importante no substituir los valores por cero ya que puede subestimarse la ingesta de un componente (111). Algunos de los estudios señalan si han tenido este aspecto en cuenta, pero el resto no da ninguna indicación al respecto (105,107,108,110).
2. **Las correspondencias con los alimentos de la BDCA no son exactas.** Es importante informar sobre este aspecto ya que puede ocurrir que no encontremos correspondencia en la BDCA para un alimento incluido en la dieta o menú y que ha sido analizado. Evidentemente, es preciso encontrar un alimento similar (104) pero si esta sustitución se ha realizado en numerosas ocasiones puede introducirse un sesgo importante en los resultados obtenidos con las BDCA. Este sesgo provocaría que se extrajeran conclusiones erróneas (112).
3. **Errores en la valoración de los pesos de ingredientes o recetas que luego se utilizarán para los cálculos.** A menudo las cantidades de los alimentos consumidos se estiman a partir de medidas caseras o los ingredientes de las recetas se estiman a partir de recetas estándar. Éste es un problema que puede darse en algunos de los estudios aunque sólo adquiere importancia si los pesos sólo se utilizan para el cálculo con la BDCA y no para los cálculos con los datos analíticos (104,105) ya que entonces sí pueden introducir un sesgo.
4. **Problemas de representatividad de la muestra analizada.** Cuando se toman las muestras a analizar es necesario que sean representativas del conjunto y que luego sean homogeneizadas correctamente (104).

5. **Métodos analíticos que no son válidos.** La elección de los métodos analíticos que se utilizarán tiene que ser meticulosa. A menudo estos estudios utilizan muestras compuestas complejas resultado de la mezcla de diferentes ingestas. La extracción de los nutrientes de este tipo de matrices presenta dificultades adicionales (3,113). A modo de ejemplo, las diferencias observadas en los lípidos del estudio de Ribeiro y colaboradores (105) pueden ser debidas en gran parte a que han utilizado el método Soxhlet (114).
6. **La elección del laboratorio que realizará los análisis no ha sido adecuada.** Es importante comprobar que el laboratorio encargado de los análisis realizará correctamente los análisis con los métodos elegidos (104). Las diferencias en los resultados obtenidos entre laboratorios pueden ser importantes, incluso en el caso de los macronutrientes (115-117).

Si estas posibles fuentes de sesgo están minimizadas, entonces los resultados de los estudios comparativos con resultados analíticos pueden servir para validar la utilización de la BDCA en cierto tipo de estudios (104). En general, se ha comprobado en el caso de los macronutrientes que las diferencias entre el cálculo de ingestas de nutrientes con BDCA y el análisis químico directo de los alimentos consumidos en un grupo de individuos (o en un solo individuo durante varios días) es inferior al 10% (51,53,112). Esta diferencia crece hasta un 20% o más en el caso de nutrientes como los lípidos, hierro, sodio, potasio, vitamina C y vitamina A. Las principales razones de estas diferencias se encuentran en el contenido variable de grasa en la carne y los platos compuestos, las diferencias en el contenido de agua (que afectan la concentración del resto de nutrientes), la contaminación de metales procedente de cuchillos y utensilios de cocina, la alta variabilidad en el contenido de sal de los alimentos procesados o preparados en casa y la alta variabilidad en las pérdidas de minerales y vitaminas durante la preparación y cocción de los alimentos (3,15,53,106,118,119).

Las diferencias observadas entre valores calculados y analizados son importantes cuando se utilizan alimentos o recetas (58,109,120), pero tienden a disminuir cuando se utilizan comidas o ingestas diarias, o cuando se incrementa el número de días o individuos del estudio (109). Esto sería debido a una reducción en la variabilidad de las estimaciones debido al incremento del número de alimentos computados (75). En el primer caso se pone de manifiesto la existencia de una variabilidad intrínseca en la composición de los alimentos: recordemos que los valores de las BDCA son representativos de la población estadística de alimentos pero no de una muestra concreta. Cuando se incrementa el número de alimentos

y el número de días del estudio las diferencias entre valores calculados y analizados tienden a disminuir ya que se trata de diferencias aleatorias. Esta variabilidad en la composición de los alimentos limita el nivel de exactitud que puede conseguirse en los cálculos realizados con BDCA (11,121).

Tabla 1-3 Diferencias (%) entre resultados obtenidos en la composición de dietas/ingestas utilizando BDCA y análisis. Calculado como valor calculado menos valor analizado y posteriormente dividido por el valor analizado.

Referencia	(53)	(105)	(105)	(106)	(106)	(107)	(108)	(108)	(109)	(104)	(104)	(110)	
Tipo de estudio	1 día; porciones duplicadas	menús colectivos de 1400 kcal; 13 días; sin cocción	menús colectivos de 1400 kcal; 13 días; con cocción	dieta con grasa láctea predominante; 14 días; porciones duplicadas	dieta aceite de girasol; 25 días; porciones duplicadas	dieta aceite de colza; 25 días; porciones duplicadas	ingesta de 2 días; registro de la dieta; porciones duplicadas	dieta reducción de peso; 3 días, 15 muestras	dieta normal 10 individuos; 6 días; registro de dieta; porciones duplicadas	media de 60 recetas	dietas para 8 grupos de edad y sexo	Otros estudios citados	Ingesta 24 h; registro de la dieta
Energía	1,7	7,5	10			-13	27		5				
Agua		-2,1											
Proteína	8,1	12	19			-21	4		-7				
Carbohidratos	-5,6	-0,7	0,7				40		19				
Azúcares	-4,5												
Ceniza		2,5											
Fibra alimentaria							0		-6				
Lípidos	7,5	22	30	-0,1	-2,5	-31	12		-16	6 (59,5)			
AGS		4,3		4,3	-3,5		3		-21	8 (75,4)			
AGM		20		20	15		4		-28	5 (67,5)			
AGP		13		13	13		4		-19	94 (184)			
18:2(n-6)				7,4	13,0								
18:3(n-3)				13	11								
Colesterol		25		25	43								
Betastosterol		21		21	24								
Campesterol		42		42	19								
Estigmasterol		17		17	15								
Calcio	-46		-22				1	19		-8 (-9/-7)	-15 / 75	-26	
Magnesio							17	9		-7 (-8/-6)	-22 / 12		
Hierro	-11		-3,6				33	6		2,6 (-5,8/17)	-5,9 / 23,1		
Fósforo	-27		-21							-5 (-6/-4)	-26 / 17		
Sodio							7	29		63 (62,2)	-8 (-13/-6)	-149 / -4	
Potasio							9	5		-5 (-9/-2)	-42 / 16	-25	
Manganeso							20	-5		-0,6 (-3,4/-2,7)			
Cinc						1,9	29	3		-0,9 (-3,9/14,5)			
Cobre						-67	-26	-16		-11 (-16,1/-4,3)			
Selenio							-19	-22					
Molibdeno							65	42					

1.3.4.4 Conclusiones respecto a la validación de BDCA

La principal dificultad de definir un procedimiento de validación para las BDCA proviene del hecho de que son el resultado de múltiples procesos y para cada uno ha de establecerse el criterio de validación. Además, la validación de cada proceso puede ser multidimensional y requerir por lo tanto varios criterios de validación (15).

De la información extraída de la literatura consultada puede concluirse lo siguiente:

- 1. El muestreo y el método analítico son dos procesos fundamentales a validar (3,13,71).** Si se está realizando una recopilación de datos existentes y el muestreo no se puede controlar, entonces se tendrá que mejorar la representatividad de los datos a través de la agregación de datos procedentes de diferentes fuentes.
- 2. La lista de criterios (Tabla 1-2) elaborada por Greenfield y Southgate (3) tiene que considerarse como una referencia a cumplir.** Su carácter de guías obtenidas a partir del consenso de la comunidad de responsables de BDCA integrados en INFOODS hace conveniente tomar tal lista como referencia y, naturalmente, las guías completas. Sin embargo, las guías no los presentan como criterios de validación.
- 3. No es recomendable utilizar los códigos de calidad como único criterio de validación.** Estos códigos son únicamente indicadores de la calidad relativa de los datos, y las guías de Greenfield y Southgate señalan la caracterización estadística y la validación de los métodos analíticos como los aspectos últimos a tener en cuenta si se quieren ofrecer valores de confianza (3). Sin embargo, pueden ser útiles para facilitar la evaluación y documentación de los datos y su uso puede ayudar a determinar el estado de la BDCA.
- 4. La validación de la BDCA realizando análisis, aunque deseable, no es práctica.** Tanto la realización de análisis de los alimentos más consumidos y su comparación con los datos de la BDCA como la realización de estudios que comparan los datos de ingesta de nutrientes obtenidos del cálculo con una BDCA y del análisis de porciones duplicadas son iniciativas muy costosas desde el punto de vista económico. Además, en el primer caso es ciertamente paradójica ya que el intento de validar la BDCA genera nuevos datos que deben ser incluidos en la BDCA. En el segundo caso, la validación sería para una aplicación concreta de la BDCA (109) y esto no garantiza su validez en otras aplicaciones. Por otro lado, Southgate y Greenfield (121) remarcan que este

segundo tipo de estudios son más bien demostraciones y no pueden considerarse pruebas de la exactitud de una BDCA, aunque la confirmación repetida de la concordancia entre valores calculados y analíticos reduciría la probabilidad de que tales concordancias se hayan producido por azar.

5. La documentación exhaustiva de la BDCA se considera un punto fundamental.

Diferentes autores coinciden en este punto (3,15,79,89,122) y de hecho es la única manera de asegurar que algún criterio de validación puede aplicarse a la BDCA (89). La existencia de limitaciones en la validación de las bases de datos significa que la principal prueba de su fiabilidad reside en cómo ha estado elaborada (11,121). Esto requiere que se ofrezca toda la documentación necesaria y que sea posible conservar la trazabilidad de cada dato (89). El cumplimiento de los criterios mencionados por los diferentes autores o bien la obtención de pruebas documentadas que demuestran la validez de la BDCA son informaciones que, obviamente, tienen que estar a disposición del usuario.

1.4 LA COMPILACIÓN DE DATOS DE COMPOSICIÓN

El proceso de compilación de datos de composición es complejo y comprende un número de sub-procesos que requieren una planificación e integración clara y cuidadosa. Para conseguir este grado de planificación es preciso definir previamente el propósito con el cual se compilará la BDCA ya que su uso último determinará tanto el contenido (los datos) como la forma (organización y medio) (15,89).

La elaboración y utilización de una BDCA se puede asimilar a un flujo de datos desde las fuentes de datos hasta los usuarios (Figura 1-1). Los datos son recopilados de las fuentes y son sometidos a una serie de procesos (véase el apartado 1.3.2) que se ordenan en una secuencia determinada y cuyo objetivo es conseguir datos de calidad (3,79) para el usuario final. Los procesos que se efectúen sobre los datos dependerán básicamente de las fuentes de datos que se utilicen y de las especificaciones que marquen los usuarios finales. El diseño del proceso general puede ser aparentemente variable (Figura 1-1) y los procesos particulares pueden variar, pero existen 4 procesos que se mantienen comunes (3,15,89,100) y que podríamos calificar de básicos:

1. **Recopilación de los datos necesarios.** La reunión de los datos procedentes de fuentes diversas es un proceso por si mismo. Por datos hemos de entender no solo los valores numéricos sino también cualquier tipo de información que sea necesaria en posteriores procesos (formas de expresión de los valores, métodos de muestreo, métodos analíticos, etc.). Algunos autores se refieren a estos datos sobre los valores como “metadatos” (79).
2. **Homogeneización de los datos.** Es preciso homogeneizar los datos procedentes de diferentes fuentes para poder operar con ellos conjuntamente (por ejemplo, realizar comparaciones entre ellos). Esto no solo se refiere a las formas de expresión sino también a información como los factores de cálculo o conversión utilizados con los datos, la codificación estándar del método analítico o la descripción del alimento con una nomenclatura también estándar.
3. **Validación de los datos / control de la calidad de los datos.** Tal como indican Greenfield y Southgate (3) es preciso realizar un control de calidad sobre los datos. Algunos autores optan por utilizar códigos de calidad (95,99-102) mientras que otros (81,89,92) utilizan sistemas de validación. En todo caso, es preciso controlar que los datos sean adecuados para los usos que se les pretende dar.

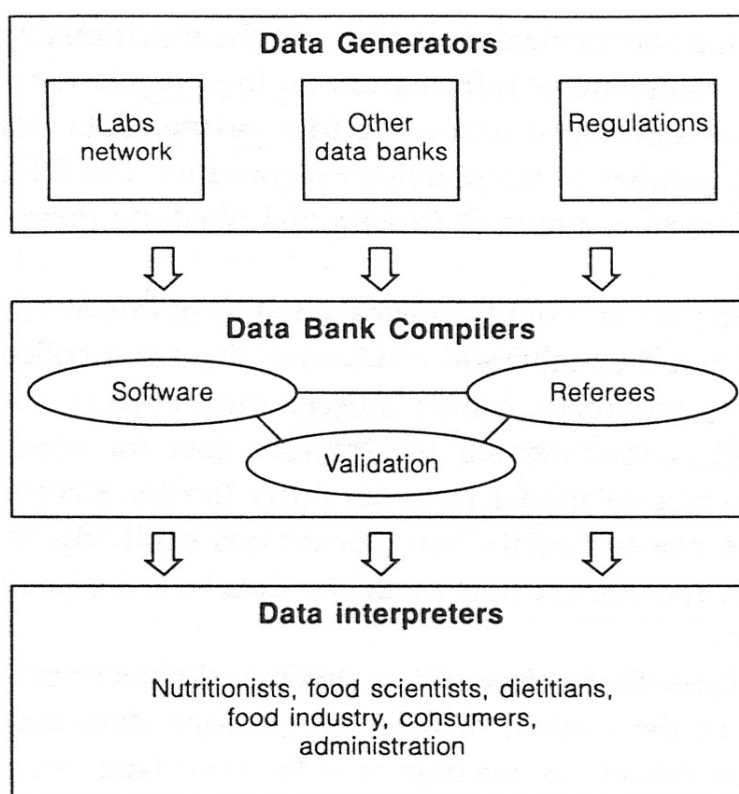
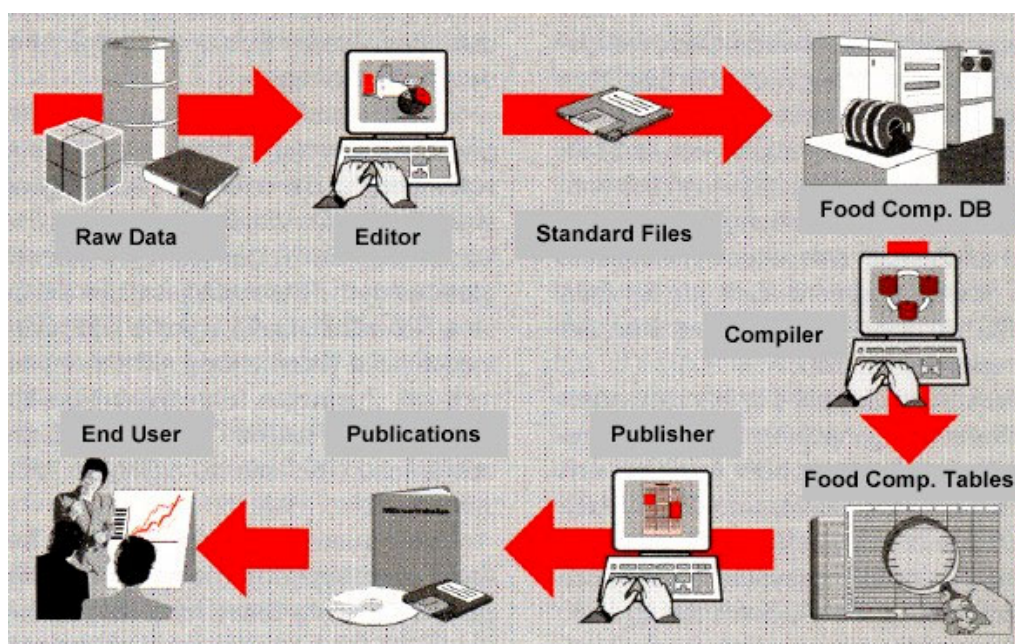


Figura 1-1 Ejemplos de cómo se estructuran diferentes procesos de compilación de datos: BDCA suiza (100) y BDCA francesa (89,123).

- 4. Agregación de los datos.** Los datos homogeneizados pueden agregarse de diferentes maneras con la finalidad de obtener la BDCA con el máximo de valores completos y de la máxima representatividad posible. La agregación puede implicar la simple yuxtaposición de datos procedentes de diferentes fuentes o el cálculo (media, mediana, ponderación, etc.) de valores representativos a partir de valores procedentes de diferentes fuentes. En cualquier caso, se trata de tomar una decisión para la cual necesitamos que los datos estén bien documentados ya que precisamos saber si los métodos analíticos utilizados por las diferentes fuentes proporcionan valores comparables, si la descripción de los alimentos es suficiente para establecer que los alimentos analizados por las diferentes fuentes pueden considerarse el mismo alimentos, si la muestra se considera representativa del alimento consumido, etc.

Es evidente que sería posible realizar una BDCA que implicara una simple recopilación de datos, sin efectuar ningún tratamiento sobre ellos, pero su utilidad sería muy limitada ya que al usuario le sería bastante difícil usarla en las aplicaciones más habituales. El tratamiento de los datos a través de procesos de datos es necesario para obtener una BDCA útil. Los procesos que se definan son importantes cuando se quiere concretar qué información es preciso recopilar para elaborar una BDCA. Por ejemplo, si se considera necesario homogeneizar los valores de proteína para recalcularlos todos con el factor 6,25 será preciso recopilar el factor original utilizado. Si se desea realizar un control de calidad se deberá recopilar la información que ayude a establecerlo (método de muestreo, método analítico, etc.). Si se quiere agregar varios valores de composición calculando la media será preciso saber el número de muestras han sido utilizadas para calcular cada valor.

Schlotke y colaboradores (89) destacan además que la gestión e intercambio de datos de composición están íntimamente relacionadas (Figura 1-2). La posibilidad de que se puedan producir diferentes tipos de intercambios de datos que den lugar a circularidades no deseadas (señaladas con una advertencia de peligro en la Figura 1-2) hace necesario que los datos estén bien documentados a todos los niveles de trabajo y que el intercambio de datos incluya todos los metadatos necesarios para permitir al compilador evitar tales situaciones.

En definitiva, en el diseño del método que permita elaborar una BDCA es preciso definir en primer lugar las necesidades del usuario final. A continuación se deben establecer los procesos que se aplicaran a los datos para conseguir obtener la BDCA que cumpla con las

necesidades establecidas anteriormente. Si se desea facilitar el intercambio de datos con otras BDCA se tendrán que tener en cuenta los requisitos adicionales que se establezcan.

Sea cual sea el método de elaboración de la BDCA y los procesos establecidos para generarla, existen ciertos puntos que son críticos. La literatura especializada los desarrolla especialmente y han sido objeto de especial consideración por parte de diversas iniciativas internacionales (3,15,20,27,28,79,124,125). Estos puntos se tratan en los apartados que siguen.

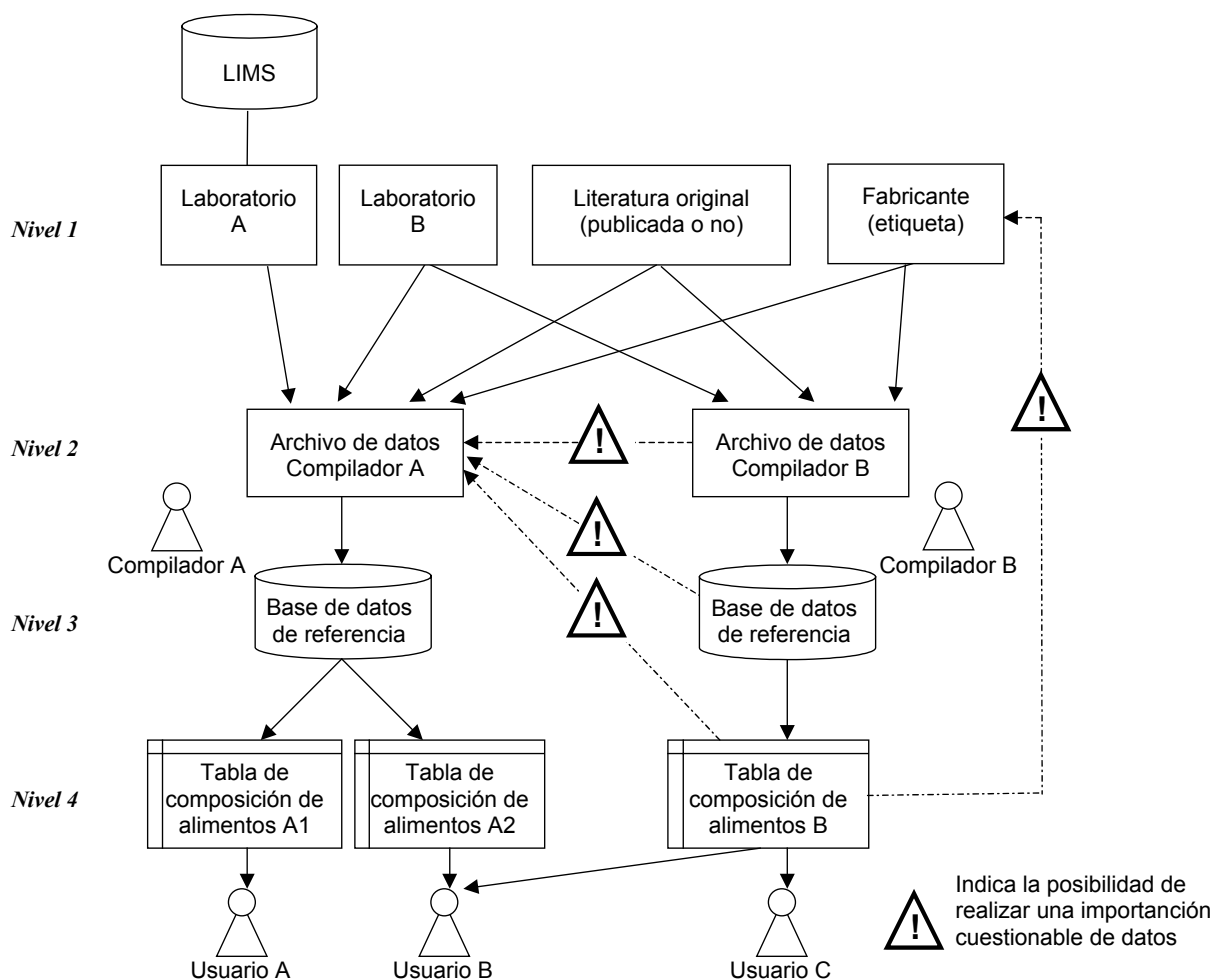


Figura 1-2 Gestión e intercambio de datos a diferentes niveles, según Schlotke y colaboradores (89). Los cuatro niveles especificados se refieren a los comentados en el apartado 1.3.2.

1.5 LAS FUENTES DE INFORMACIÓN DE DATOS DE COMPOSICIÓN

La decisión sobre qué fuentes de información se utilizarán en la compilación de una BDCA depende de diversos aspectos, siendo el principal las necesidades que se pretenden cubrir con la BDCA (véase el apartado 1.3.3).

Basándonos en la clasificación de los datos contenidos en las BDCA elaborada por Greenfield y Southgate (véase el apartado 1.3.1) podríamos distinguir entre fuentes analíticas documentadas, el propio sistema de compilación de datos (ya que pueden generarse datos a partir de datos preexistentes o mediante la realización de estimaciones) y fuentes no documentadas (por ejemplo, valores procedentes de etiquetas). Las fuentes analíticas serán siempre preferibles, pero dentro del sistema pueden generarse datos de composición perfectamente válidos (15). Las fuentes no documentadas son, naturalmente, más difíciles de evaluar.

Teniendo en cuenta sólo las fuentes externas al sistema de compilación de datos y fijándose especialmente en aquellas características de la fuente más relacionadas con el aseguramiento de la calidad de los datos de composición, Greenfield y Southgate distinguen cuatro grandes tipos de fuentes de datos de composición (3):

1. **Informes analíticos.** Cabe distinguir entre los elaborados específicamente para la BDCA y los elaborados con otros propósitos (por ejemplo, el control de los alimentos comercializados). En el primer caso los datos se han obtenido a través de un plan de muestreo y un programa analítico especialmente diseñado, mientras que en el segundo caso es muy posible que el plan de muestreo o el método analítico no se ajusten a las especificaciones marcadas por los compiladores de BDCA.
2. **Publicaciones primarias.** Se trata de artículos publicados en revistas científicas que contienen información sobre composición de los alimentos. El espectro de revistas científicas que pueden ser de interés para el compilador de BDCA no se limita a las relacionadas con las ciencias de los alimentos o la nutrición sino que también incluye revistas relacionadas con el desarrollo de métodos químicos, agronomía, fisiología vegetal, edafología, nutrición animal, ganadería, acuicultura y otros (3,13). Aunque muchas de estas revistas disponen de revisores es importante tener en cuenta que la revisión se ha realizado en función de los objetivos de la revista y del artículo. A menudo la sección experimental de los artículos proporcionan detalles insuficientes en relación a

las necesidades del compilador de BDCA e, incluso, para una mínima revisión crítica del manuscrito (126).

3. **Publicaciones secundarias.** Se trata de revisiones, libros o capítulos de libros, artículos no revisados y compilaciones de datos de composición (incluyendo TCA y BDCA). Los valores ofrecidos por estas publicaciones suelen ser difíciles de evaluar debido a la insuficiencia en la información ofrecida y a que, a menudo, no se detallada la fuente original del valor. Esto es especialmente cierto en el caso de TCA y BDCA. Puede ser necesario ponerse en contacto con los autores para obtener más información sobre los datos, aunque ello comporta una importante dedicación en cuanto a tiempo por parte del compilador (95) y no siempre se obtiene respuesta (97).
4. **Informes no publicados.** En este apartado se incluyen los documentos que incluyen datos de composición y que tienen una circulación limitada, normalmente dentro de una organización, institución, agencia o empresa. La documentación de los valores puede ser muy desigual y su evaluación, por lo tanto, puede requerir contactos con los autores. En este apartado podrían incluirse también las tesis doctorales y trabajos de final de carrera no publicados.

Tanto los datos analíticos procedentes de laboratorios como los informes no publicados suelen ser fuentes de difícil acceso (15). Los datos procedentes del etiquetado nutricional deben utilizarse con prudencia. Es conveniente contactar con la empresa productora para determinar cómo se han obtenido los datos. Si no se dispone de otros datos, los valores de las etiquetas pueden utilizarse como “último recurso” y ser reemplazados tan pronto como se disponga de mejores valores. No obstante, la información de las etiquetas puede ser útil en muchos otros aspectos: identificación y descripción de los alimentos, identificación de formulaciones, estimación de valores de composición, etc. (13,15,81).

1.6 IDENTIFICACIÓN DE ALIMENTOS

La nomenclatura y descripción de alimentos es uno de los aspectos más importantes a considerar tanto cuando se trata de la elaboración de BDCA (3,15,74,79,127) como de su utilización en diferentes aplicaciones (44,51,53,119,128-132). La preparación de BDCA fiables requiere una identificación precisa de los alimentos. Esta identificación tiene que realizarse superando la enorme complejidad cultural y comercial de los alimentos humanos y permitir que los alimentos sean descritos a un nivel que haga posible que otros usuarios de los datos de composición, especialmente aquellos de otros países, puedan reconocer claramente un alimento y distinguirlo de otros alimentos (29).

No obstante, disponer de esta precisión puede resultar difícil cuando se utilizan datos existentes (método indirecto) ya que normalmente el objetivo para el cual se han generado los datos difiere bastante del objetivo del proyecto de generar una BDCA (3) y puede ser que los responsables de los datos no hayan incluido aspectos en la descripción de los alimentos que pueden ser importantes para el compilador de BDCA. De este modo, aunque los datos de composición sean de buena calidad pueden ser una fuente de error si el alimento al que corresponden no está claramente descrito. Polacchi (127) y Granado y colaboradores (130) ilustran esta situación con diversos ejemplos obtenidos de la literatura científica y de TCA.

El problema de la descripción de alimentos se extiende a las BDCA: debido a que las diferentes bases de datos utilizan métodos diferentes para identificar los alimentos, el intercambio de datos resulta difícil entre países, entre organizaciones del mismo país o incluso dentro de una misma institución (35). Para superar estos inconvenientes se han llevado a cabo diferentes iniciativas que han intentado introducir cierto grado de sistematización y estandarización en este campo.

Pueden distinguirse dos tipos de aproximaciones al problema de la identificación de alimentos. La primera consiste en intentar clasificar los alimentos en categorías “universales”, de manera que los alimentos que se clasifican del mismo modo corresponden al mismo alimento; estos sistemas de clasificación o agrupamiento pueden también considerarse como sistemas de codificación de los alimentos (29) y por lo tanto sirven para identificarlos. La segunda aproximación trata de establecer una descripción detallada de los alimentos incluidos en la BDCA a través de un sistema estándar y pre-establecido que en algunos casos intenta estandarizar incluso los términos descriptivos. Ambos sistemas tienen diferentes objetivos: los sistemas de clasificación tienden a “agregar” los alimentos con

características similares mientras que los sistemas descriptivos pretenden dar una descripción de los alimentos lo más precisa posible sin necesidad de agregarlos. Los sistemas de clasificación son más convenientes para el usuario final de los datos, mientras que los sistemas descriptivos se ajustan mejor a las necesidades de los compiladores de datos. Los principales ejemplos de cada caso se tratarán a continuación, pero si se desea obtener más información sobre este tema puede consultarse el trabajo de Ireland y Møller (35).

1.6.1 Sistemas de clasificación de alimentos

Existen diferentes sistemas de clasificación o categorización que se utilizan en diversos ámbitos: el establecido por la *Confederation de Industries Agroalimentaries* (CIAA), la clasificación Codex de los alimentos y piensos de la FAO/WHO, el sistema de codificación de la *World Trade Organization* (WTO), el sistema de codificación Eurocode-2 y otros de menor importancia. Todos estos sistemas han sido diseñados para un uso determinado (facilitar el establecimiento de listas positivas de aditivos, facilitar el comercio internacional, clasificar los alimentos en los estudios de consumo de alimentos, etc.) y resulta difícil encontrar la versatilidad necesaria en la compilación de BDCA. Además, la clasificación que realizan de los alimentos no acostumbra a ser suficientemente detallada para su uso en aplicaciones usuales de las BDCA como, por ejemplo, el cálculo de aportes nutricionales (35).

La inclusión de estos sistemas de clasificación de los alimentos se ha tenido en cuenta en el desarrollo de BDCA (79,100) pero no consta en la literatura que hayan tenido un amplio uso. De todos ellos, Eurocode es seguramente el que más se ha acercado a un sistema de clasificación de los alimentos de uso internacional en el ámbito de las BDCA (129,133). Sin embargo, presenta una serie de limitaciones que han frenado su utilización: resulta difícil definir reglas lógicas para la asignación de los alimentos a los grupos correspondientes (especialmente para los alimentos complejos como las recetas) y además presenta una serie de definiciones de ciertos grupos de alimentos que no se ajustan a los estándares establecidos por el *Codex Alimentarius* o las directivas de la Comisión Europea (35). Recientemente se ha llevado a cabo una revisión de este sistema con la finalidad de facilitar su uso en los estudios de evaluación del consumo de alimentos y nutrientes y para la clasificación y agregación de alimentos en las BDCA (33).

1.6.2 Sistemas descriptivos de alimentos

El primer sistema creado específicamente para la identificación de alimentos, concretamente piensos animales, fue el sistema INFIC/ENFIC (23,35). El sistema se basa en un tesoro multilingüe de términos que sirven para describir diferentes facetas de los alimentos destinados a animales. El vocabulario se controla a través de la estandarización de los términos y su definición estricta, así como con el establecimiento de relaciones jerárquicas entre ellos: para cada faceta existe una jerarquía de términos que va de los más generales a los más específicos. Sin embargo, no resulta aplicable a las necesidades de la alimentación humana (23). En la Tabla 1-4 se muestran dos ejemplos de alimentos descritos mediante el sistema INFIC/ENFIC.

Tabla 1-4. Ejemplos de alimentos descritos usando el sistema INFIC/ENFIC.

	Alfafa seca	Patata hervida
1. Origen	<i>Medicago sativa</i> , alfalfa	tubérculos sin piel
2. Parte	parte aérea	hervido
3. Proceso	deshidratada y triturada	
4. Estadio de crecimiento	vegetativo tardío	
5. Recolección/corte	corte 1	
6. Grado	19, 1-21% de proteína	
Número de identificación	22-703	13-228

Basándose en esta estructura por facetas del sistema INFIC/ENFIC se han desarrollado dos sistemas descriptivos específicamente orientados a la descripción de los alimentos propios de la alimentación humana: LanguaL y el sistema INFOODS.

1.6.2.1 LanguaL

El CFSAN (*Center for Food Safety and Applied Nutrition*) de la FDA (*Food and Drug Administration*) desarrolló el *Food Factored Vocabulary* (FFV) a finales de la década de los 80 (23,134). El FFV estaba pensado para su utilización en el ámbito estadounidense y se ajustaba a las necesidades legales de este país. Sin embargo, pronto se decidió internacionalizar su uso, intentando adaptarlo a la utilización en diferentes países (123,135,136), cambiando su nombre por el de LanguaL (acrónimo de *langua alimentaria*) y creando un comité técnico internacional que se encargaba del desarrollo, mejora y difusión

del sistema. LanguaL se define como “un método para describir, capturar y recuperar datos acerca de los alimentos” (31,137).

LanguaL se estructura por facetas (como el sistema INFIC/ENFIC) y utiliza un tesoro normalizado multilingüe de términos en el cual el idioma dominante es el inglés. Existen 13 facetas que hacen referencia al grupo de alimento en el cual debe encuadrarse, origen del alimento, atributos físicos, procesado a que ha sido sometido, envase, usos alimentarios, origen geográfico y otras características (Tabla 1-5).

Tabla 1-5. Facetas del sistema LanguaL.

CARACTERÍSTICA	FACETAS
GRUPO DE ALIMENTO	A. Tipo de producto
ORIGEN DEL ALIMENTO	B. Fuente del alimento C. Parte del animal o planta
ATRIBUTOS FÍSICOS	E. Estado físico, tamaño o forma
PROCESADO	F. Extensión del tratamiento térmico aplicado G. Método de cocción H. Tratamiento aplicado J. Método de conservación
ENVASADO	K. Medio de envasado M. Contenedor o envoltorio N. Superficie en contacto
USOS ALIMENTARIOS	P. Grupo de consumidores / Uso dietético
ORIGEN GEOGRÁFICO	R. Zonas y regiones geográficas
CARACTERÍSTICAS ADICIONALES	Z. Características adicionales

Para cada faceta pueden utilizarse una serie de términos tomados del tesoro. Estos términos tienen relaciones establecidas entre ellos que son de tipo jerárquico. Las relaciones jerárquicas pueden ser sencillas o polijerárquicas, en las cuales un término puede estar en una situación jerárquica inferior respecto a varios términos a la vez (Figura 1-3). Cada término tiene asociado un código formado por la letra de la faceta a la cual pertenece y un número de cuatro dígitos. Así, por ejemplo, el término *BROWN SUGAR ADDED* tiene el código H0361 (faceta H: Tratamiento aplicado) y el término *BREAKFAST CEREAL* tiene el código A0258 (faceta A: Tipo de producto). La información que debe intercambiarse entre usuarios del sistema LanguaL no son los términos descriptivos sino estos códigos alfanuméricos (31,32). En la Tabla 1-6 se ofrece un ejemplo de descripción de un alimento usando el sistema LanguaL.

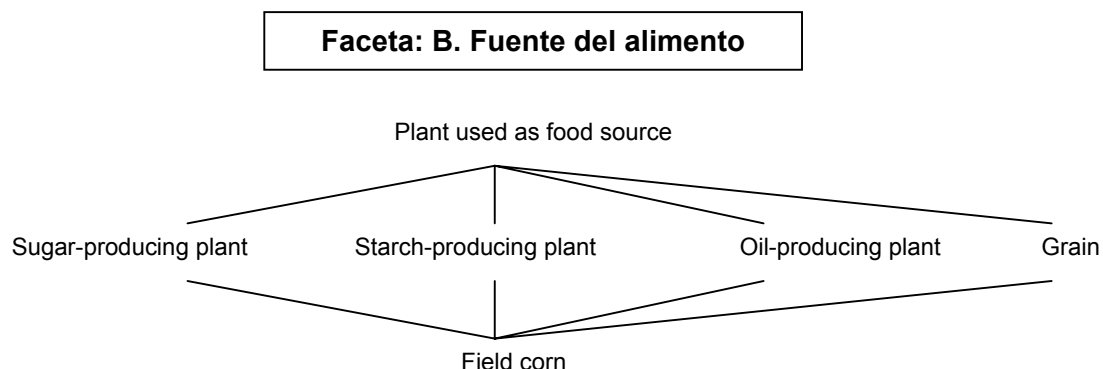


Figura 1-3. Ejemplo de relación polijerárquica entre términos del tesouro. El término FIELD CORN se halla simultáneamente bajo cuatro términos de nivel superior en la jerarquía (31).

Tabla 1-6. Ejemplo de descripción de un alimento usando las facetas y términos de facetas propios de LanguaL, y asignación de códigos para el tratamiento informático de la información (31,32). En una BDCA junto al nombre del alimento se almacenaría su descripción según el sistema LanguaL, pero usando únicamente los códigos alfanuméricos. En este ejemplo sería: A0258 · B1379 · C0208 · E0153 · F0014 · G0003 · H0158 · H0100 · H0367 · H0163 · H0181 · H0274 · H0138 · J0116 · K0003 · M0148 · N0015 · N0039 · P0024. Los literales que corresponden a cada código se almacenarían aparte para recuperarse cuando fuera necesaria una presentación en pantalla.

NOMBRE COMERCIAL: Kellogg's Corn Flakes

DESCRIPCIÓN: Copos de maíz tostados

LISTA DE INGREDIENTES: Maíz, azúcar, extracto de malta, sal, niacina, hierro, vitamina B6, riboflavina (B2), tiamina (B1), ácido fólico, vitamina D y vitamina B12.



BREAKFAST CEREAL	A0258
FIELD CORN	B1379
SEED OR KERNEL, SKIN REMOVED, GERM REMOVED (ENDOSPERM)	C0208
WHOLE, SHAPE ACHIEVED BY FORMING, THICKNESS <0,3 CM	E0153
FULLY HEAT TREATED	F0014
COOKING METHOD NOT APPLICABLE	G0003
SUCROSE ADDED	H0158
FLAVORING OR SPICE EXTRACT OR CONCENTRATE ADDED	H0100
SALT ADDED	H0367
VITAMIN ADDED	H0163
IRON ADDED	H0181
FLAKED	H0274
WATER REMOVED	H0138
DEHYDRATED OR DRIED	J0116
NO PACKING MEDIUM USED	K0003
PAPERBOARD CONTAINER WITH PAPER LINER	M0148
WAX	N0015
PAPERBOARD OR PAPER	N0039
HUMAN FOOD, NO AGE HUMAN FOOD, NO AGE SPECIFICATION	P0024

El sistema LanguaL presenta diversas ventajas. Permite una descripción muy detallada de los alimentos utilizando un vocabulario controlado que permite superar barreras culturales e idiomáticas. De este modo permite un intercambio fiable de datos entre BDCA, aunque sean de países diferentes (15,136). La estructura por facetas facilita la elaboración de la descripción de los alimentos al proporcionar un recordatorio sobre cuales son los aspectos que deben describirse (128). Además, las búsquedas por ordenador de alimentos en una base de datos utilizando los términos descriptivos de LanguaL pueden ser muy flexibles. Es posible realizar búsquedas utilizando los términos más generales en la jerarquía o bajar hasta los términos más específicos, aplicar múltiples estrategias de búsqueda combinando los términos de una misma faceta o de facetas diferentes y utilizar diferentes conectores lógicos (AND, OR, XOR, etc.) entre términos descriptivos. El hecho de que los términos se codifiquen permite que el vocabulario pueda ser actualizado fácilmente sin que esto afecte la base de datos donde se almacenan las descripciones de los alimentos (31,128,134).

Las principales desventajas del sistema es que los usuarios suelen encontrarlo demasiado complejo y que la concordancia entre las descripciones realizadas por diferentes personas para un mismo alimento no siempre es satisfactoria, aunque ésta se incrementa cuando la experiencia del codificador en el uso de LanguaL es mayor (138). La complejidad en la codificación de los alimentos utilizando LanguaL puede atenuarse si se utilizan programas informáticos para esta tarea (139). Algunas de las facetas de LanguaL precisan una mayor clarificación y la lista de términos es aún insuficiente. Además, el mantenimiento del vocabulario precisa de herramientas informáticas de las que aún no se dispone (35).

Actualmente se tiene constancia de que LanguaL está siendo utilizado activamente en el proceso de compilación de datos de composición en Estados Unidos, Francia y Dinamarca (35). La implantación de LanguaL no es aún, ni mucho menos, universal (31).

1.6.2.2 Sistema INFOODS

Las Guías INFOODS para la Descripción de los Alimentos fueron preparadas por el Comité de Nomenclatura de los Alimentos y Terminología de INFOODS en 1987 (29). El propósito de este sistema de nomenclatura es proporcionar un marco para el intercambio de datos entre las fuentes de datos y los compiladores de BDCA. Se trata de un amplio sistema organizado en facetas (como LanguaL) pero cuyos descriptores no se restringen a un vocabulario organizado en un tesoro sino que puede utilizarse cualquier término. Está diseñado para capturar el máximo posible de información descriptiva sobre los alimentos

(Tabla 1-7). Aunque las descripciones se realizan con texto libre, estaba previsto desarrollar un vocabulario de términos, pero no se acabó de concretar (35).

El sistema INFOODS para la descripción de los alimentos está siendo utilizado, más o menos adaptado a las necesidades particulares, en Nueva Zelanda (140), el Pacífico Sur, algunos países de la ASEAN, dos países africanos y 10 países de Latinoamérica. Está siendo incorporado también en los sistemas de trabajo en Oriente Próximo y el Sur Asiático (35). Otros países utilizan sistemas descriptivos organizados por facetas similares al sistema INFOODS o LanguaL (141).

1.6.3 La utilización combinada de sistemas de descripción y clasificación

Últimamente ha cobrado fuerza una tendencia a la utilización combinada de diferentes sistemas de identificación de los alimentos (descripción y clasificación) para aprovechar las ventajas que cada uno de ellos ofrece. Así, se han desarrollado sistemas de correspondencia (*mapping systems*) entre Eurocode 2 y LanguaL (35), interfaces (*International Interface Standard for Food Databases*, desarrollado por la FDA) que permiten realizar consultas sobre bases de datos usando diferentes sistemas y facilitan la recuperación de información y estandarizan la descripción de los alimentos (136) y, por último, recomendaciones que definen los atributos necesarios para describir los alimentos especificando más de 50 propiedades que pueden afectar la composición de los alimentos (35,79). En realidad, estas últimas recomendaciones reúnen diferentes sistemas (LanguaL, INFOODS, CIAA, etc.) en una única lista de atributos.

Tabla 1-7. Resumen del sistema INFOODS para describir los alimentos. Los idiomas permitidos son el inglés, francés o español excepto para B1 y B2 (29).

A. Fuente del nombre del alimento y términos descriptivos.

B. Nombre e identificación del alimento

1. Nombre del alimento en uno de los idiomas nacionales del país.
 2. Nombre local del alimento.
 3. Término equivalente más próximo para este alimento en inglés, francés o español.
 4. País o área en la cual se obtuvo la muestra del alimento.
 5. Grupo de alimentos y código para este alimento en la BDCA del país (proporciónese referencia bibliográfica).
 6. Grupo de alimentos y código para este alimento en la BDCA de la región (proporciónese referencia bibliográfica).
 7. Grupo de alimentos según Codex Alimentarius o sistema INFOODS
-

C. Descripción de alimentos “sencillos”

1. (a) Fuente del alimento (en inglés, francés o español)
(b) Nombre científico del alimento (Latin)
(c) variedad, raza o línea
 2. Parte del animal o planta
 3. País o área de origen
 4. Nombre del fabricante y dirección. Número de lote o remesa.
 5. Otros ingredientes (incluyendo aditivos).
 6. Procesado y/o preparación del alimento.
 7. Método de conservación
 8. Grado de cocción
 9. Condiciones de producción agrícola.
 10. Madurez.
 11. Condiciones de almacenamiento.
 12. Grado.
 13. Contenedor y superficie de contacto con el alimento.
 14. Estado físico, tamaño o forma.
 15. Color.
 16. Otros descriptores no contemplados.
 17. Disponibilidad y localización de una fotografía o dibujo del alimento.
-

D. Descripción de alimentos “mixtos” (multi-ingrediente).

1. Ingredientes y cantidades; fuente de la información sobre los ingredientes.
 2. Procedimiento de la receta.
 3. Lugar donde se elaboró el alimento.
 4. Disponibilidad y localización de una fotografía o dibujo del alimento.
 5. Nombre y dirección del fabricante.
 6. Contenedor y superficie de contacto con el alimento.
 7. Método de conservación.
 8. Condiciones de almacenamiento.
 9. Preparación final de este alimento multi-ingrediente.
-

E. Usos

1. Peso de la porción típica y medida casera correspondiente.
 2. Disponibilidad; frecuencia y estación de consumo.
 3. Lugar usual del alimento dentro de la alimentación (momento del día, lugar en la colación, etc.)
 4. Consumidores del alimento.
 5. Propósitos especiales del alimento; alegaciones especiales.
-

F. Muestreo y manipulación de la muestra en el laboratorio.

1. fecha de recolección.
 2. Peso de las muestras.
 3. Porcentaje de la porción comestible; naturaleza de la porción comestible.
 4. Porcentaje de desperdicio; naturaleza del desperdicio.
 5. Lugar de recolección: suministradores; tipo de comercio.
 6. Manipulación entre el suministrador y el laboratorio.
 7. Manipulación en el laboratorio.
 8. Almacenamiento en el laboratorio y manipulación subsiguiente.
 9. Estrategia para los análisis.
 10. Razón para realizar los análisis.
-

1.7 IDENTIFICACIÓN DE COMPONENTES

Cuando se comparan TCA o BDCA procedentes de diferentes países es posible apreciar diferencias considerables, y a veces realmente sorprendentes, entre las definiciones que se dan de los nutrientes y otros componentes (26). A menudo el problema surge porque los nutrientes o componentes no se definen desde un punto de vista únicamente químico sino que se tiene en cuenta su efecto en la fisiología del organismo humano (142,143). Por ejemplo, en algunas tablas de composición (144) el valor de vitamina E incluye sólo uno de sus vitámeros, el α -tocoferol, mientras que otras incluyen también otros vitámeros (17,145). Los valores de vitamina E en ambas tablas pueden presentar diferencias debido a que la definición de lo que se entiende por vitamina E no es la misma. Otros ejemplos de diferencias habituales en las definiciones de los nutrientes pueden encontrarse en el caso de la energía, los carbohidratos, la fibra alimentaria, la grasa, los ácidos grasos, la proteína y algunas vitaminas (3,142,146-152).

A menudo las diferencias en las definiciones de los nutrientes y componentes se corresponden con diferencias en los métodos de obtención de los valores, ya sean analíticos o estimaciones (153-155). Por ejemplo, la fibra alimentaria puede analizarse mediante el método de la AOAC o el método Englyst (142) pero ambos métodos no son equivalentes: parten de definiciones diferentes de lo que se entiende por "fibra alimentaria" y no determinan los mismos compuestos químicos, con lo cual los resultados obtenidos difieren en mayor o menor medida dependiendo del alimento (156). Por otro lado, la energía metabolizable de los alimentos suele estimarse multiplicando los nutrientes que aportan energía por un factor de cómputo calórico (3). Estos nutrientes suelen ser la grasa, la proteína y los carbohidratos en todas las tablas, pero algunas BDCA añaden otros componentes que pueden aportar energía como los ácidos orgánicos o los polialcoholes (148). Las últimas recomendaciones de la FAO proponen contabilizar también la energía procedente de la fibra alimentaria (157), con lo cual es de esperar que algunas BDCA la incluyan también. A esto hay que sumar también que las definiciones que se den de grasa, proteína y carbohidratos pueden afectar la estimación de la energía. Por último, las definiciones de los nutrientes pueden variar en función de los avances y la aparición de nuevas necesidades que se produzcan en diferentes campos científicos (158).

Las formas de expresión utilizadas por diferentes fuentes pueden no coincidir y esto también está relacionado con diferencias en la definición del nutriente. Podemos encontrar ejemplos de estos casos en las vitaminas. Para calcular la actividad total de vitamina E en un alimento

es preciso multiplicar la cantidad presente de cada vitámero por un factor que expresa la actividad del vitámero en relación a la actividad del α -tocoferol (143). A continuación se suman las cifras obtenidas para cada vitámero en un valor total. Se obtiene así una estimación de la *actividad total de vitamina E* en el alimento expresada como *equivalentes de α -tocoferol*. Sin embargo, también podría darse la situación que no se aplicara ningún factor de actividad del vitámero o que sólo se incluyera uno de ellos (como se ha comentado más arriba). Así pues, en este caso la definición del nutriente que se proporciona esta relacionada con la forma de expresión que se utilice.

Estas diferencias en las definiciones de los nutrientes o componentes pueden llevar a los usuarios o a otros compiladores de BDCA a cometer errores importantes cuando utilicen los datos de composición (3,15,130,146,148,154). Por tal razón, el desarrollo de sistemas para asegurar un identificación inequívoca de los componentes se ha considerado un aspecto crítico en el intercambio de datos de composición entre bases de datos (27,79).

1.7.1 Los códigos INFOODS y EUROFOODS para identificar componentes

Aunque ya existían algunas guías para la nomenclatura de diferentes componentes presentes en los alimentos, las primeras orientadas específicamente a la elaboración de BDCA e intercambio de datos de composición fueron las publicadas por INFOODS (27). Actualmente pueden consultarse, debidamente actualizadas, por internet (159).

Estas guías establecen, partiendo de una serie de consideraciones previas, una lista de definiciones de componentes de los alimentos. Para cada uno de ellos las guías proporcionan la siguiente información:

1. Una **abreviatura** (*tagname*) sencilla y única que se utiliza en el intercambio de datos. Esta abreviatura tiene que ser, en la medida de lo posible, fácilmente interpretable por parte de los usuarios.
2. Un **nombre** o definición descriptiva.
3. Una **unidad de medida** común o por defecto a utilizar en la expresión de la cantidad por 100 gramos de porción comestible de alimento.
4. **Sinónimos** por los cuales también es conocido el componente en cuestión.
5. **Comentarios** para aquellos componentes que precisan aclaraciones adicionales.
6. Una lista de **tablas seleccionadas** en las cuales aparece el componente en cuestión.
7. Información adicional: **notas, palabras claves y ejemplos**.

La principal consideración que se establece es que el uso de las abreviaturas es determinar si los valores asociados pueden ser comparados o combinados. Todas las otras consideraciones deben contribuir a que la principal pueda cumplirse. En la Tabla 1-8 pueden encontrarse ejemplos de las abreviaturas más habituales para algunos nutrientes. Es importante observar como el código INFOODS puede incluir información sobre la identidad del alimento, el método de obtención del valor (ya sea analítico o mediante una estimación) y la forma de expresión usada. En el Anexo I se ha incluido, también a modo de ejemplo, la descripción detallada de las abreviaturas para el componente “proteína”.

Tabla 1-8. Ejemplos de abreviaturas más habituales para los nutrientes proteína, grasa total y carbohidratos totales según la guía INFOODS para la identificación de componentes (27).

Abreviatura	Unidad	Descripción
PROCNA	g	Proteína total; calculada a partir del nitrógeno aminoacídico
PROCNP	g	Proteína total; calculada a partir del nitrógeno proteico
PROCNT	g	Proteína total, calculada a partir del nitrógeno total
PRO-	g	Proteína total; método de determinación desconocido
FAT	g	grasa total
FATCE	g	grasa total; obtenida de análisis utilizando el método de extracción continua
FATNLEA	g	grasa total; según la definición de la NLEA (equivalentes de triacilglicéridos de ácidos grasos)
CHO-	g	carbohidratos totales; método de determinación desconocido
CHOAVL	g	carbohidratos disponibles
CHOAVLM	g	carbohidratos disponibles; expresados en equivalentes de monosacáridos
CHOCDF	g	carbohidratos totales; calculados por diferencia
CHOCSM	g	carbohidratos totales; calculados por sumación

Los códigos INFOODS para los componentes han sido modificados y ampliados en sucesivas revisiones (146,159) pero han conservado su estructura original. Dentro de las guías no se realiza ninguna propuesta sobre como deben agruparse los diferentes componentes.

Posteriormente, las recomendaciones elaboradas por EUROFOODS (30,79) realizaron una propuesta de códigos para los componentes presentes en los alimentos que parte de los abreviaturas elaboradas por INFOODS pero que modifican algunos aspectos para conseguir que la utilización de los códigos se adapten a las recomendaciones o para facilitar la interpretación de las abreviaturas. Los cambios que se introducen son los siguientes:

1. **La abreviatura EUROFOODS no contiene referencia alguna a la forma de expresión de los valores.** Las recomendaciones EUROFOODS ya prevén la existencia de un campo en la BDCA destinado a almacenar esta información.
2. **La abreviatura EUROFOODS no contiene referencias sobre el método de obtención del valor excepto en un caso: los carbohidratos obtenidos por diferencias.** Las recomendaciones EUROFOODS ya prevén la existencia de una serie de campos en la BDCA que informan acerca de cómo se ha obtenido el valor de composición.
3. **Algunas abreviaturas han sido modificadas** con la intención de facilitar su uso. Éste es el caso de algunos componentes (carbohidratos, ácidos grasos, etc.) de los factores para cálculo de algunos componentes (factor de cómputo proteico, factor de conversión para los ácidos grasos) y algunas propiedades (cómputo químico, desperdicio). En otros casos se han introducido algunas abreviaturas que no existían en las guías INFOODS.
4. **Los componentes se han agrupado** para facilitar la localización y la realización de consultas con medios informáticos.

En definitiva, es necesaria una adecuada identificación de los componentes tanto en la etapa de compilación de datos como durante la utilización de las BDCA por parte de los usuarios. Durante la compilación, esta identificación puede realizarse usando una nomenclatura propia que tenga en cuenta los aspectos importantes para identificar adecuadamente un componente. Pero para el intercambio de datos es preciso disponer de un sistema común de identificación (27,79) y para el usuario de BDCA puede resultar útil disponer de un sistema internacional que le facilite la comparación de datos procedentes de diferentes BDCA. Esta última razón ha llevado a que algunas TCA o BDCA incluyan los códigos INFOODS en sus publicaciones destinadas al usuario final de los datos (17,48).

1.7.2 El Component Aspect Identifier (CAId)

La necesidad de identificar inequívocamente los componentes y la estrecha relación que tiene esta identificación con aspectos tales como el método analítico y la forma de expresión, ha llevado a la búsqueda de soluciones específicas para el compilador de datos de composición que precisa reunir información compleja en un espacio limitado para poder comparar listas de valores de composición.

El CAId es una solución que surge de las necesidades específicas del compilador de datos de composición con el propósito de recopilar toda la información relacionada con un componente y poder distinguir entre valores que corresponden a componentes diferentes (160). Consiste en una cadena de códigos que aportan información sobre las siguientes categorías:

- **Identidad del componente.** Utiliza el código INFOODS para identificar los componentes.
- **Modo de expresión.** Utiliza los sufijos correspondientes del código INFOODS.
- **Origen del valor.** Utiliza tres códigos: el primero indica el método de obtención (analítico, calculado, estimado, etc.), el segundo especifica cual ha sido el método utilizado para los valores analíticos y el tercero proporciona un vínculo con la referencia bibliográfica que describe el método analítico.
- **Fuente del valor.** Utiliza dos códigos: el primero describe a grandes rasgos la fuente que ha proporcionado el valor (TCA, BDCA, artículo científico, empresa, etc.) y el segundo proporciona un vínculo a la referencia bibliográfica de la fuente.
- **Calidad del valor.** Código que proporciona una indicación de la calidad del datos. Para los valores estimados este campo puede utilizarse para indicar alguna observación respecto a la estimación realizada.

Estos códigos se reúnen en una secuencia que acompaña a cada valor. Por ejemplo, la secuencia:

STARCH · M · a · Polarimetry · J0002 · f · B0007 · A

indica que se trata de un valor de almidón (*STARCH*) expresado en equivalentes de monosacárido (*M*) por 100 gramos de alimento, obtenido analíticamente (*a*) por polarimetría según el método analítico publicado en la referencia *J0002*, procede de una TCA (*f*) cuya referencia bibliográfica tiene el código *B0007* y la calidad del valor es alta (*A*).

El CAId no pretende substituir a los códigos de identificación mencionados en el apartado anterior, antes al contrario, hace uso de ellos. El CAId trata simplemente de categorizar y codificar la información asociada al componente, de igual modo como lo hace LanguaL.

Aunque sólo se tiene constancia que se haya usado en la compilación de la BDCA sueca (160), muchos aspectos del CAId han influido notablemente en las recomendaciones de EUROFOODS aunque no haya sido incorporado como tal a estas recomendaciones (79). Un problema que parece presentar el CAId es que pone de manifiesto una superposición entre lo que sería la identificación del componente y la identificación del método con el cual se ha obtenido, aunque esto ya sucedía con el código INFOODS (79,160).

1.8 MÉTODOS DE OBTENCIÓN DE LOS VALORES DE COMPOSICIÓN

Una vez identificado adecuadamente el alimento y los componentes, el siguiente paso en la compilación de datos de composición es evaluar cómo se han obtenido éstos. Como es fácil deducir de lo expuesto en el apartado 1.3.1, existen diferentes métodos para obtener datos de composición de los alimentos. Aunque no se ha encontrado una clasificación recomendada que establezca los diferentes tipos de métodos de obtención, de la consulta de los principales textos es posible establecer la existencia de cuatro tipos de métodos: analíticos, estimaciones, cálculo a partir de otros componentes y “préstamo” de datos (3,15,79,124).

Cuando se juzga la calidad de un dato de composición es necesario tener en cuenta el método con el cual se ha obtenido (3). La idoneidad de cada método de obtención depende de diferentes factores, que pueden variar según el caso, de manera que no siempre es fácil determinar qué método es más efectivo para obtener datos de composición fiables. Es importante valorar bien todos estos factores y no caer en simplificaciones como la de considerar siempre que los mejores datos de composición se obtienen realizando análisis de alimentos. No siempre los datos analíticos son los más exactos (15) y, además, el coste de su obtención es alto.

1.8.1 Métodos analíticos

Los métodos analíticos son aquellos que se basan en procedimientos analíticos realizados en el laboratorio que determinan la cantidad del componente de interés presente en el alimento, ayudándose de patrones o estándares. En una situación de recursos ilimitados, se trata del mejor método para obtener datos de composición. Sin embargo, es un método costoso en términos de tiempo y dinero por lo que a menudo es preciso recurrir a otros métodos (véanse los apartados siguientes).

Para evaluar si un método analítico puede considerarse adecuado para generar datos que se utilizarán en una BDCA es preciso tener en cuenta qué uso se hará de la BDCA. Normalmente, el uso en aplicaciones nutricionales es el perfil de base que se utiliza para valorar los métodos analíticos, pero algunos usos particulares (por ejemplo, monitorización de la ingesta de tóxicos) pueden obligar a ser más exigentes en cuanto al método analítico (49). Conociendo el uso que se hará de los datos podemos establecer nuestras

especificaciones respecto a los métodos analíticos. No obstante, ¿qué criterios deben utilizarse para evaluar la calidad de un método analítico?.

La respuesta a esta pregunta no es sencilla. De hecho, ha ido cambiando a lo largo del tiempo y es previsible que siga cambiando en el futuro: mientras que antes se juzgaba la calidad de un método analítico a partir de su descripción minuciosa, la respuesta actual está más enfocada hacia los resultados analíticos (161). Los indicadores del grado de rendimiento de un método analítico que se pueden utilizar como criterios de validación se obtienen mediante la realización de pruebas de validación y analizando estadísticamente los resultados (71,84,161-163). Estos datos son muy útiles para juzgar la adecuación de un método analítico, pero no deja de ser preciso disponer de una buena descripción del método, así como de detalles sobre el tratamiento al cual se ha sometido la muestra analítica, información sobre los protocolos de aseguramiento de la calidad seguidos y la comparación con datos procedentes de la literatura científica. Desgraciadamente, muy a menudo no se ofrece toda esta información tan necesaria (126).

Existe también algún estudio que intenta obtener datos representativos de la composición de los alimentos al mismo tiempo que proporciona medidas de las incertidumbres asociadas a éstos, analizando las diferentes fuentes de variabilidad (164). Sin embargo, su complejidad y coste económico no hace posible la aplicación sistemática de este enfoque a la compilación de BDCA.

Sale fuera del marco de esta introducción revisar en detalle, componente por componente, los métodos analíticos existentes y su validez para la compilación de datos BDCA para uso en Nutrición. Existen textos que tratan este tema y que son fuentes de referencia en el ámbito de las tablas y bases de datos de composición de alimentos (3,113,114,165), así como artículos de revisión o estudios de comparación entre métodos o laboratorios que ofrecen información más detallada (115,116,166,167).

1.8.2 Estimaciones y cálculos a partir de otros componentes

Los métodos de estimación son aquellos que se basan en un razonamiento o algoritmo (convenientemente documentado) para estimar qué cantidad de un componente de interés está presente en el alimento. Los métodos de estimación son necesarios porque a menudo no es posible cubrir todos los alimentos y componentes con datos obtenidos con métodos analíticos o porque realizar un muestreo y análisis para ese alimento y/o componente no es rentable en términos de coste/beneficio o bien, sencillamente, no es necesario (15). En

definitiva, los métodos de estimación permiten reducir o eliminar por completo el número de valores desconocidos presentes en una BDCA. Los valores desconocidos pueden conducir a una infraestimación de las ingestas calculadas usando BDCA, siendo preferible la realización de estimaciones antes que adjudicar el valor cero (3,14,111).

Existen diferentes situaciones en las cuales los métodos analíticos no son necesarios para obtener datos de composición con la fiabilidad requerida. Cuando se trata de determinar el contenido de un componente que se sabe, partiendo de conocimientos bioquímicos bien establecidos, que no está presente en el alimento o bien lo está en cantidades ínfimas es posible “asumir” que el dato de composición es el valor cero (81). Por ejemplo, el colesterol o la vitamina B12 en alimentos de origen vegetal.

Otro ejemplo al respecto podemos encontrarlo cuando se desea determinar la composición de una receta casera. Normalmente, la variabilidad de la composición de estos alimentos es alta debido a los diferentes métodos de preparación seguidos por la población, diferentes gustos o incluso a los utensilios de cocina utilizados. Torelm (164) realizó un extenso e interesante estudio analizando las fuentes de variabilidad en los datos de composición de alimentos y recetas caseras. En este estudio queda patente que la variabilidad introducida por las diferentes preparaciones de las recetas es alta (Tabla 1-9). El componente con una mayor variabilidad es el hierro, debido según los autores a la utilización de diferentes recipientes de cocina que pueden liberar hierro (por ejemplo, el hierro fundido). En el caso de las recetas suele ser más conveniente calcular su composición a partir de los ingredientes en lugar de realizar un muestreo (¡que debería realizarse en los puntos donde se consume, esto es, hogares y hostelería!) y los análisis correspondientes.

Si se quisiera determinar la composición de una receta mediante el método directo sería necesario recoger un alto número de muestras en los hogares donde se preparan, tal y como se realiza en el estudio antes citado (164). Este tipo de estudios son extremadamente costosos. Por otro lado, la preparación y análisis en el laboratorio tampoco garantizaría una exactitud alta en los resultados. El cálculo de recetas mediante algoritmos, especialmente si tienen en cuenta los cambios nutricionales, es económico y se puede conseguir un grado de precisión aceptable para la mayor parte de aplicaciones en el ámbito de la Nutrición (168,169).

Tabla 1-9 Variaciones (% de la desviación estándar relativa) debidas a la operación de preparación de recetas en dos situaciones diferentes. Según Torelm (164), modificado.

Componente	Preparación estándar por investigadores ¹	Preparación realizada por familias ²
Humedad	1,0	3,4
Cenizas	3,9	15
Nitrógeno total	5,9	10
Grasa	9,8	27
Calcio	6,4	22
Fósforo	4,6	12
Hierro	6,3	55
Sodio	6,3	26
Potasio	9,0	24

(1) Menú de 10 platos; (2) Menú de 3 platos

Existen algunos componentes que se pueden calcular a partir de otros componentes. A veces su valor puede obtenerse a través de métodos analíticos y otras no, pero siempre pueden calcularse agregando los valores de los componentes que lo constituyen si se dispone de tales datos. Encontramos ejemplos en las agregaciones usuales de ácidos grasos (saturados, poliinsaturados, monoinsaturados, etc.), carbohidratos (carbohidratos totales, azúcares, monosacáridos, etc.), vitaminas (la vitamina E puede calcularse como la agregación de sus vitámeros, con o sin corrección por potencia vitamínica), etc. Parece discutible incluir los valores calculados como estimaciones cuando, por definición, sólo pueden obtenerse mediante cálculo. Algunos autores consideran que los valores calculados a partir de otros componentes son estimaciones (81) aunque otros no (79), existiendo además ciertas diferencias de matiz.

Llegados a este punto es importante hacer notar una precisión terminológica. Los términos “calculado” e “imputado” son usados a menudo para denominar aquellos datos que no son analíticos, asignándose a “calculado” más fiabilidad que a “imputado”, aunque no existe consenso al respecto: algunos autores prefieren hablar de una manera general de valores “estimados” (15) mientras que otros utilizan “imputado” como término general (13) o bien no distinguen entre “imputado” y “estimado” (79). En el presente texto se ha preferido utilizar exclusivamente el adjetivo “estimado” y especificar a qué tipo de estimación se refiere.

Existen pocos trabajos que hayan intentado sistematizar los diferentes métodos de estimación y cálculo utilizados. El más detallado y extenso es el de Schakel y colaboradores (81), el cual distingue siete procedimientos diferentes para estimar la composición de los alimentos.

1. Utilización de valores procedentes de un alimento diferente pero parecido.

Consiste en utilizar los datos procedentes de un alimento lo más parecido al alimento problema, realizando si es necesario un ajuste por materia seca. En este apartado los autores incluyen también lo que llaman “perfiles nutricionales”: la composición general de un grupo de alimentos establecida a partir de datos analíticos existentes y usada para la estimación de valores en alimentos de este mismo grupo. Es el mismo tipo de estimación que apunta Southgate (49) como usual para adjudicar un número a los valores desconocidos. El grado de similitud entre dos alimentos debe juzgarse en función de diversos aspectos: relación biológica (pertenencia a la misma especie, género, etc.), condiciones de cultivo o cría, madurez, morfología, procesado, similitud de los ingredientes, etc. (15).

2. Cálculo de valores a partir de una forma diferente del mismo alimento. La aplicación de los factores de retención real (RR) de nutrientes (170) para calcular el contenido de un componente en un alimento procesado (cocinado, desecado, etc) partiendo de su forma en crudo entraría dentro de esta categoría. Es posible que deban realizarse ajustes adicionales por materia seca si los cambios de contenido hídrico son importantes. Otro método incluido en esta categoría sería el cálculo de la composición de un alimento a partir de la composición de sus partes.

3. Cálculo de valores a partir de otros componentes del mismo alimento. En esta categoría se incluyen los cálculos para obtener componentes como la energía, la actividad total de las vitaminas y provitaminas, los carbohidratos por diferencia o los perfiles de ácidos grasos o aminoácidos calculados a partir de otro alimento.

4. Cálculo de valores a partir de recetas caseras o formulaciones industriales.

Consiste en utilizar la composición de los ingredientes y, mediante los algoritmos y factores necesarios, calcular la composición de la receta o fórmula. Existen muchos algoritmos diferentes para realizar tales cálculos (15,77,145,168) y la principal limitación es disponer de los factores necesarios (169), normalmente retenciones reales o porcentaje de pérdida (170), para realizar las correcciones por cambios en los componentes inducidos por el procesado. Otras veces se puede recurrir a la utilización

de las listas de ingredientes y algunos valores de composición de las etiquetas de los alimentos para llegar, a través de métodos de optimización, a ajustar la composición completa del producto (87).

5. **Conversión de valores procedentes de etiquetas nutricionales de productos comerciales.** Se incluyen en este apartado todos aquellos ajustes y conversiones que sea preciso realizar para utilizar los datos de composición procedentes de etiquetas nutricionales.
6. **Calculo de valores a partir de un producto estándar.** Los estándares, ya sean legales o no, pueden utilizarse en algunos casos para establecer el contenido más probable de un componente.
7. **Ceros asumidos.** Cuando existen suficientes evidencias de que un componente está ausente en un alimento, puede estimarse directamente como cero.

Estos siete métodos resumen de manera general las diferentes estrategias seguidas por los compiladores para estimar aquellos valores para los cuales se carece de datos, pero cada autor puede utilizar diferentes métodos o variaciones (15,49,171,172). Por ejemplo, también es habitual la utilización de relaciones matemáticas entre concentraciones de componentes (lineales o de otro tipo) en un grupo de alimentos para realizar estimaciones (141,156,172), aunque estas no están incluidas explícitamente en la clasificación de Schakel y colaboradores (81).

Las recomendaciones EUROFOODS incluyen una lista de posibles métodos para estimar valores de composición, aunque no realizan un estudio y definición con detalle al respecto (79). La clasificación mantiene algunas similitudes con la clasificación establecida por Schakel y colaboradores (81) y conserva la distinción entre tres grandes grupos de métodos de obtención de valores: analíticos, calculados y estimados. Dentro de los valores calculados se distingue entre:

1. **Calculado genérico.** Valor calculado, sin más especificación.
2. **Calculado como agregación de alimentos.** Cuando la composición se obtiene a partir de la suma de la composición de sus ingredientes o de sus partes.
3. **Calculado como receta.** Cuando la composición se obtiene a partir de la composición de los ingredientes, aplicando además factores por cambios debidos a la cocción.

4. **Calculado a partir del perfil de un componente.** Se utiliza el perfil del mismo componente procedente de otro alimento (por ejemplo, utilizando el perfil de ácidos grasos de la leche entera para calcular el perfil de ácidos grasos del yougurt entero natural).
5. **Sumación de componentes constituyentes.** Por ejemplo, la suma de azúcares y polisacáridos para calcular los carbohidratos totales. Se incluye también la sustracción (por ejemplo, el cálculos de carbohidratos por diferencia).
6. **Cálculos utilizando factores de conversión.** En este apartado se incluiría el cálculo de la energía, la proteína bruta o las actividades vitamínicas.
7. **Calculado de un alimento relacionado.** Por ejemplo, cuando se calcula la composición del pan tostado a partir del pan teniendo en cuenta las pérdidas de agua.

Dentro de los valores estimados se distingue entre:

1. **Estimado genérico.** Valores estimados, sin más especificación.
2. **Estimado de un alimento relacionado.** Se utiliza la composición de un alimento similar o relacionado.
3. **Estimado de otro alimento y otro componente.** Se utiliza la composición de una alimento similar y de otro componente que no es exactamente el mismo.
4. **Estimado de acuerdo con regulaciones legales.**

Aunque se trata de una clasificación que recoge las experiencias propias (79) y de otros autores (3,15,81,173), los autores de las Recomendaciones reconocen que se trata de códigos que deben ser refinados a través de su aplicación.

1.8.3 El “préstamo” de valores

A veces es preciso recurrir a valores de los cuales no se tienen adecuadas referencias. Sería el caso de un valor extraído de una BDCA que no indican cómo se han obtenido los valores y cuya documentación puede ser deficiente (no indica el método de elaboración de

las tablas, las fuentes, etc.). Tal vez el término podría extenderse a otras fuentes de datos como etiquetas, folletos, etc. pero Greenfield y Southgate (3) establecen el término “valor prestado” específicamente para el caso de las BDCA.

El problema de los valores prestados es que no pueden ser evaluados adecuadamente: no se sabe como se han obtenido y a veces se carece de otra información fundamental o bien no está suficientemente clara (por ejemplo, la forma de expresión o la definición del nutriente). La evaluación de estos valores reside básicamente en su consistencia y en la reputación del organismo que ha elaborado la BDCA (15).

1.9 DESCRIPCIÓN DE LOS VALORES DE COMPOSICIÓN

Como se ha visto en el apartado anterior, los valores de composición necesitan estar adecuadamente documentados por lo que al método de obtención se refiere. Adicionalmente, cada valor de composición tiene que estar adecuadamente descrito en lo que se refiere a su significado. Los aspectos que deben describirse son diversos:

1. **Las unidades y modos de expresión.** Los valores de composición, sean cuales sean, se refieren a concentraciones. Se trata de cantidades del componente expresadas en determinadas unidades (*equivalentes de monosacáridos, mL, mg, etc.*) que se hallan presentes en una determinada cantidad de alimento. La cantidad de alimento (*100, 1000, 1, etc.*) puede expresarse en diferentes unidades (*g, mL, kg, etc.*) y de diferentes modos (*porción bruta, porción comestible, nitrógeno, etc.*) de manera que tenemos diferentes posibilidades para expresar las concentraciones. Es importante que las BDCA incluyan las unidades utilizando el Sistema Internacional de unidades y evitando expresiones ambiguas (3,27,89). Durante el proceso de compilación de la BDCA es preciso homogeneizar las diferentes formas de expresión de los valores procedentes de fuentes diferentes antes de agregarlos para obtener valores representativos.
2. **El significado de los valores.** Como ya se ha visto, los valores incluidos en las BDCA pueden ser numéricos, trazas, cero o desconocidos. Los tres primeros llevan asociadas unidades y modos de expresión pero, además, a menudo es preciso describir el significado del valor incluido en la BDCA. Por ejemplo, un valor traza puede simplemente incluirse como “*tr.*” en la base de datos o bien como el límite de cuantificación del método analítico (79). En este último caso, el valor numérico no será una estimación de la concentración de componente al alimento, sino el límite por debajo del cual se supone que se halla el valor verdadero. Si la BDCA incluye valores como los límites de cuantificación, los límites de detección, valores mínimos, etc. es preciso que queden identificados adecuadamente. Los valores desconocidos también han de quedar perfectamente identificados como tales (3).
3. **Los dígitos significativos.** Se trata de un aspecto al cual normalmente no se le presta la suficiente atención. Cuando se analiza la composición química de los alimentos, se sabe que los resultados llevan asociada una incertidumbre. Esta incertidumbre es debidas a las varianzas en las técnicas analíticas y a las varianzas en las muestras. Por consiguiente, es preciso expresar los resultados de modo que el usuario de los datos no tenga duda de qué dígitos se conocen con certeza y cuales no (15,85,174). La

convención actual es que los resultados deben expresarse con todos los dígitos que se conocen con certeza y solo el primer dígito incierto. El número de dígitos significativos no tiene nada que ver con el número de decimales. Por ejemplo, el número “0,00125” tiene tres dígitos significativos y el número “125” también.

Para determinar el número de dígitos significativos que debe atribuirse a un valor es preciso conocer la desviación estándar asociada. Existe aproximadamente una probabilidad igual a 0,95 de que un resultado se sitúe en el rango:

$$\bar{x} \pm 2 \cdot s$$

Por lo tanto, la incertidumbre de la media del resultado analítico es aproximadamente dos veces la desviación estándar. A partir de este dato y siguiendo un procedimiento sencillo, puede establecerse el número de dígitos significativos. Si la desviación estándar no se conoce entonces los datos deben expresarse con los dígitos significativos apropiados para los tipos de análisis utilizados (85) o bien los sugeridos por las guías elaboradas al respecto (132). La utilización de estas convenciones permite al usuario de los datos discernir si dos valores pueden considerarse diferentes o no. Cuando se utilizan datos procedentes de fuentes diversas (publicaciones, libros, etc.), en la mayoría de casos no se siguen estas convenciones (85).

La descripción de los valores presenta algunas lagunas. Algunos autores señalan que la descripción de un valor debe incluir también los datos estadísticos asociados al valor, aunque pocas BDCA incluyen información al respecto (79). Por otro lado, las definiciones de los diferentes tipos de valores (traza, desconocido, etc.) no concuerdan siempre (28,111). Finalmente, existe cierta confusión en lo que se refiere a qué información debe ser considerada como parte de la descripción de un valor debido a que a veces se superpone con la información que se refiere a la descripción del componente (160).

1.10 EVALUACIÓN DE LA CALIDAD DE LOS DATOS DE COMPOSICIÓN

La calidad puede entenderse como el atributo de adecuación al propósito de utilización. Cuando se trata de datos de composición podría decirse que los datos de calidad alta son aquellos que cubren las necesidades de la mayoría de los usuarios mientras que los datos de calidad baja no las cubren. No obstante, los datos de baja calidad podrían ser adecuados para usuarios menos exigentes (41,71).

1.10.1 La calidad de los datos analíticos

La elaboración de una BDCA implica tener en cuenta las necesidades de los usuarios. Estas pueden ser muy diversas aunque presentan ciertos requerimientos comunes (21,71):

1. Los datos tienen que ser **representativos** de los alimentos del ámbito al cual va destinada la BDCA.
2. Los datos tienen que haber sido obtenidos a través de **métodos analíticos apropiados** y que se hayan **aplicado correctamente**.
3. Los datos tienen que reflejar con **exactitud** la composición de los alimentos.

Rand y colaboradores (15) reducen estas tres propiedades de los datos a una de sola: la **exactitud**. Consideran estos autores que la exactitud de un dato es función de la representatividad de la muestra estadística respecto a la población estadística y de la exactitud y precisión de la medida analítica. A su vez, la precisión de la medida analítica es una función del número de muestras y del cuidado en la aplicación de la técnica analítica. Sea cual sea el enfoque, los principales autores coinciden en que existen tres puntos críticos cuando se evalúan datos destinados a una BDCA: el **muestreo**, el **método analítico** y la **aplicación del método analítico**. Evidentemente, los principios que se aplican para diseñar un plan de muestreo correcto y un plan de análisis adecuado se aplican también al escrutinio de los datos analíticos procedentes de fuentes publicadas o no publicadas (71), aunque demasiado a menudo las fuentes de datos no proporcionan suficiente información al respecto (126). Se han desarrollado también sistemas para codificar o describir la calidad de los datos en las BDCA que se basan en la evaluación de los tres puntos críticos antes citados. Estos sistemas ya se han comentado en el apartado 1.3.4.2.

1.10.1.1 Muestreo

Los datos que se compilan para generar una BDCA tienen que ser representativos de los alimentos consumidos por la población a la cual va dirigida la BDCA. Cuando se utiliza el método directo esto significa que el plan de muestreo tiene que diseñarse intentando obtener muestras representativas de los alimentos habitualmente consumidos por la población (70), teniendo en cuenta que cada componente es un caso particular pues su variabilidad puede diferir de manera importante (71). Más aún, el número de muestras necesario para conseguir un grado de error en el muestreo aceptable depende de la variabilidad del componente en el alimento (70) y muy a menudo esta variabilidad no es conocida, con lo cual el número de muestras se asigna intuitivamente basándose en información parcial sobre la distribución del componente en el alimento (71,175).

Los protocolos de muestreo se basan en el conocimiento sobre los alimentos bajo estudio (producción, transformación, distribución, etc.) y su consumo por parte de la población a la cual va destinada la base de datos. Es especialmente importante que la muestra se proteja de posibles alteraciones que modifiquen su composición. En la guías elaboradas por Greenfield y Southgate se trata en extensión este tema (3) así como en numerosos artículos (10,70,71,176-178), libros (15,20) y otras publicaciones científicas (179).

Es importante señalar que el diseño de protocolos de muestreo detallados que tienen en consideración todas las fuentes posibles de variación conocida es laborioso y llevarlos a término es muy costoso. Normalmente se reservan para aquellos alimentos con un peso muy importante en la alimentación de la población. Para los alimentos con un peso menor se elaboran protocolos menos ambiciosos, lo cual disminuye la calidad de los datos obtenidos (71,174,177). Para disminuir el número de análisis y disminuir costos sin perder representatividad puede recurrirse a las muestras compuestas, pero ello implica la pérdida de información sobre la distribución y variabilidad del componente (70,180).

1.10.1.2 Método analítico

La elección del método analítico es tal vez el punto más importante (4,15). Los requisitos que debe cumplir un método analítico para ser considerado adecuado en la elaboración de BDCA están explicadas con detalle en la literatura especializada (15,162,174), especialmente la obra de Greenfield y Southgate (3). Es importante tener en cuenta que los conocimientos sobre análisis de alimentos van incrementándose con el paso del tiempo y métodos aceptados en el pasado pueden considerarse no adecuados actualmente. Por otro

lado, el hecho de que un método sea aceptado como válido o como oficial en determinados campos no significa que lo sea también en el desarrollo de BDCA de aplicación en Nutrición (166). También se han de tener en cuenta los aspectos prácticos y económicos cuando se elige el método analítico (181).

1.10.1.3 Aplicación del método analítico

Aunque un método analítico sea adecuado y su rendimiento óptimo, si no se aplica convenientemente puede generar valores inexactos o imprecisos. Para evitarlo es necesario establecer un control de calidad y un sistema de aseguramiento de la calidad (3,174). El control de la calidad de los datos está orientado a que los resultados estén bajo control estadístico, esto es, que en la aplicación rutinaria del método analítico la exactitud y precisión de las medidas se mantengan en el tiempo (182). El aseguramiento de la calidad incluye todas las actividades orientadas a procurar que la información generada por un laboratorio sea correcta (3). Las Buenas Prácticas de Laboratorio (BPL) o los sistemas certificables (ISO 9000, EN 45001, etc.) son ejemplos de sistemas de aseguramiento de la calidad (183).

1.10.2 La calidad de los datos no analíticos

No se ha establecido hasta la fecha ningún sistema que permita evaluar la calidad de los datos no analíticos. Se han realizado algunos estudios para valorar la efectividad de algoritmos para el cálculo de la composición de un alimento a partir de sus ingredientes (87,168,170). Normalmente, se asigna una mayor confianza a los valores obtenidos con tales algoritmos que a los valores obtenidos con otros tipos de estimaciones (15).

1.10.3 Chequeos sobre los valores durante el proceso de compilación

Los valores pueden someterse a dos tipos de chequeos para detectar posibles valores erróneos. Estos chequeos son parte del sistema de control de la calidad de los datos compilados.

Los **chequeos internos** consisten en comprobar la consistencia de los valores. Esto puede realizarse con unas sencillas operaciones matemáticas y la comparación del resultado obtenido con el esperado. Por ejemplo, la suma de los componentes principales de un alimento expresados en gramos por 100 gramos de porción comestible tiene que situarse

entre 97 y 103, o la suma de los ácidos grasos expresados en gramos por 100 gramos tiene que ser inferior a la grasa total. Estos chequeos los puede establecer el propio compilador, aunque en la literatura pueden encontrarse algunas sugerencias al respecto (3).

Los **chequeos externos** consisten simplemente en la comparación con otros valores procedentes de diferentes muestras y/o fuentes de datos. Esto puede realizarse visualmente o utilizando métodos estadísticos para detectar valores extremos y valores aberrantes (89,182,184). Este chequeo se realiza durante la agregación de los valores para obtener valores centrales representativos.

1.11 AGREGACIÓN DE LOS VALORES DE COMPOSICIÓN

Como se ha visto en el apartado anterior, los datos que se compilan para generar una BDCA tienen que ser representativos de los alimentos consumidos por la población a la cual va dirigida la BDCA. Si se dispone de un número alto de valores analíticos obtenidos con el método directo, los valores obtenidos a partir del análisis de las muestras serán representativos de la población y, normalmente, podrán agregarse realizando un simple media aritmética o bien una mediana (3). Si se dispone de pocos valores y estos exhiben una importante variabilidad entonces la operación es más complicada y requiere un detallado estudio de los metadatos por parte del compilador a partir del cual deberá decidir qué tipo de agregación es la más adecuada (véase más abajo). Naturalmente, si la muestra es compuesta no será necesaria la agregación ya que sólo se dispondrá de un valor (70).

Cuando la BDCA se elabora con el método indirecto entonces la representatividad de los datos no puede asegurarse en el muestreo. El diseño del plan de muestreo lo establece la fuente de datos en función de sus propios objetivos lo cuales, naturalmente, no tienen porque coincidir con los nuestros (70,122,174). En tal caso, la única posibilidad de influir en la representatividad de los datos de composición es durante la agregación de los valores procedentes de diferentes fuentes. En esta etapa es posible seleccionar los valores y agregarlos de modo que sea posible acercarse al máximo a un valor representativo de la población de alimentos que nos interesa. Para ello es preciso disponer de los metadatos relevantes y de la posibilidad de realizar diferentes tipos de agregaciones (3,15).

Los metadatos relevantes son aquellos que permiten seleccionar las muestras en función de los factores que pueden afectar su composición y que se tienen en cuenta cuando se diseña un plan de muestreo. Pueden ser la marca comercial, la variedad o raza, el tipo de cultivo, la zona de cultivo, la zona de venta, el punto de venta (comercio, grandes superficies, hostelería, etc.), ingredientes, etc (3,49,79). En general, aquellos metadatos contemplados en los sistemas descriptivos (véase el apartado 1.6). Sin embargo, muy habitualmente las fuentes de información utilizadas en la compilación de BDCA no proporcionan la suficiente información al respecto y el compilador tienen que conformarse con información parcial (176, 102).

Existen diferentes tipos de agregaciones que se pueden efectuar: la media, la mediana, la media geométrica, la media ponderada según el número de muestras, la ponderación y el cálculo del valor central son las más habituales (15,79,89). La decisión sobre cual de las

agregaciones es la más adecuada depende de la información proporcionada por los metadatos y de la distribución de los valores, aunque normalmente no se dispone de suficientes valores para apreciar correctamente su distribución. Debido a la importancia de conocer la distribución estadística de los valores, se recomienda introducir los valores analíticos procedentes de diferentes muestras y fuentes sin agregar aunque, de nuevo, muy a menudo esto no es posible porque las fuentes de información los proporcionan sólo agregadas (15). Si se dispone de una cantidad suficiente de datos para estudiar su distribución estadística y se observa que ésta es simétrica entonces la mejor elección es la media de los valores individuales, o la media ponderada por el número de muestras de cada valor si se trata de valores ya agregados y se desea obtener la máxima representatividad. Si se aprecia un sesgo puede utilizarse la mediana o, si se aprecia una distribución log-normal, la media geométrica. Si se distinguen diferentes poblaciones, será preciso investigar en profundidad los datos y analizar la significación nutricional para decidir si deben separarse o bien mantenerse juntas. Cuando se pretende obtener un valor representativo de un alimento con un consumo o composición diferente a lo largo del año (o en función de otra variable) puede recurrirse a las ponderaciones. Holden y colaboradores (94), por ejemplo, ponderan según la puntuación obtenida por el plan de muestreo en su sistema de evaluación de la calidad con la finalidad de dar más peso a aquellas muestras más representativas. Evidentemente, además de calcular estos valores centrales será preciso calcular, a ser posible, también los parámetros estadísticos que informen de la variabilidad (desviación estándar, rango, error estándar, intervalo de confianza del 95%, etc.) y el número de muestras que intervienen en el cálculo. Si se dispone de pocos valores y, además, éstos presentan una alta variabilidad entonces la situación es más compleja. Normalmente estos casos tienen que resolverse a favor de aquellos valores que tengan un nivel de confianza mayor, excluyendo el resto. Todas estas operaciones están condicionadas también por la conservación de la consistencia de los valores (3,185).

En cualquier caso, si el compilador no puede calcular un valor agregado representativo, entonces debe seleccionarse aquel valor que merezca mayor confianza. Siempre que sea posible, es conveniente que la BDCA destinada al usuario incluya algún parámetro estadístico que informe de la variabilidad así como una descripción de las muestras utilizadas en la agregación (15).

La agregación de datos es la etapa final en el proceso de compilación. El valor generado será el que se utilizará en la BDCA destinada a los usuarios. Como se ha visto, normalmente la agregación puede requerir una revisión de los metadatos. Esto implica que es necesario guardar la documentación asociada a cada dato y conservar la trazabilidad

para poder revisar las operaciones efectuadas. Para poder cumplir con estos dos requisitos es imprescindible disponer de una aplicación informática para la gestión de la BDCA (4).

1.12 APLICACIONES DE GESTIÓN DE BDCA

La compilación de datos de composición de alimentos con el objetivo de generar una BDCA comporta la recopilación, escrutinio, selección y tratamiento de una cantidad considerable de información muy diversa. Se trata de una tarea que puede realizarse manualmente sobre papel y contando con un buen sistema de archivo para coleccionar datos y organizarlos adecuadamente. De hecho, así se realizaba no hace muchos años. No obstante, actualmente todas estas tareas se realizan utilizando las potentes herramientas informáticas disponibles y substituyendo el soporte papel por las *bases de datos informáticas*. Una base de datos informática la constituyen los ficheros de datos y las interrelaciones entre los ficheros.

Sea cual sea el soporte y las herramientas que se utilicen (papel y archivos o bien ordenador), para compilar datos de composición es preciso construir un *sistema de información*. Un sistema de información (SI) se define como “una entidad identificable del entorno donde se encuentra, dotada de una estructura definida con un conjunto de elementos constituyentes, que recibe y transmite informaciones y, eventualmente, trata y gestiona dichas informaciones, justificando así su razón de ser, y que es susceptible de evolucionar en el tiempo”. Una *aplicación* dentro de un sistema de información es “un subsistema de información que trata de resolver una problemática, con una misma finalidad y una fuerte cohesión entre sus componentes” (86). Un ejemplo de sistema de información podría ser el sistema de gestión de una empresa y una aplicación dentro de este sistema sería un programa de contabilidad general. A menudo sistema y aplicación coinciden de tal modo que se hace imposible diferenciarlos.

La elaboración de un SI (incluyendo las aplicaciones) contempla una serie de etapas cuyo conjunto forma el ciclo de vida del sistema. Los nombres que pueden recibir estas etapas pueden variar según el autor, pero en general se reconocen las siguientes (86):

1. **Etapas de especificación**, donde se estudia y define el nuevo sistema a crear.
2. **Etapas de diseño**, donde se detallan y modelan los componentes del nuevo sistema.
3. **Etapas de implantación**, donde se prueba y se pone en marcha el nuevo sistema.

En el diseño de una aplicación, y de cualquier SI en general, es imprescindible definir adecuadamente los procesos que se desean llevar a cabo, y los datos a tratar mediante tales procesos. Para identificar las diferentes entidades de datos del sistema, sus

interrelaciones y sus atributos es necesario elaborar un *modelo conceptual* de los datos. Posteriormente será necesario trasladar este modelo a un nivel más físico en el cual ha de tenerse en cuenta el sistema de gestión de bases de datos (SGBD) que se utilizará en el diseño de la aplicación; es decir, será preciso elaborar un *modelo físico* de los datos. Un SGBD es sencillamente un programa orientado a facilitar el uso de las bases de datos, el diseño, programación, mantenimiento, etc. (86). Actualmente, la práctica más extendida en el desarrollo de aplicaciones es la utilización del *modelo de entidad-relación* como modelo conceptual durante el análisis y las primeras etapas de diseño de los datos, y la utilización del *modelo relacional* como modelo físico para acabar el diseño y construir la base de datos con un SGBD (186). Ambos modelos se explican en los siguientes apartados, así como las características de los SGBD.

Existe poca información sobre desarrollo de aplicaciones informáticas destinadas a la recopilación de datos de composición de alimentos y se restringen principalmente a los modelos conceptuales o físicos de la información (79,89,187). Del mismo modo, se han publicado pocos trabajos sobre aplicaciones informáticas para la compilación de BDCA y la descripción de sus funcionalidades acostumbran a ser muy someras. Se trata de aplicaciones desarrolladas para la gestión de las BDCA de Nueva Zelanda (140,188), Suiza (100) y Suecia (189). Éste último sistema es tal vez el que mejor descrito está, existiendo una descripción detallada que se puede consultar por Internet (190). Por otro lado, Klensin (28) desarrolló un formato para el intercambio de datos por encargo de INFOODS que llevaba implícito un cierto modelo para los datos. A pesar de la elegancia del sistema, no tuvo éxito debido a la dificultad de desarrollar las aplicaciones informáticas que soportaran el formato que proponía (79). A principios de los noventa se desarrolló en España una aplicación para gestionar una BDCA que, sin embargo, no tuvo continuidad (191). También se han desarrollado aplicaciones para la consulta de BDCA a través de internet (99,192) que presentan conceptos e ideas de interés para el desarrollo de aplicaciones de gestión de BDCA.

1.12.1 Modelo conceptual de la información: el modelo entidad-relación

Un modelo muy extendido para realizar una análisis lógico de la información es el modelo entidad-relación (MER), basado en el álgebra relacional (89,100,140,141) Este modelo considera que aquello que nos rodea puede ser percibido como una colección de objetos, que denominaremos “entidades”, y relaciones entre objetos.

Según el MER, una entidad es un objeto que existe y que es distinguible. Un grupo de entidades son entidades del mismo tipo. Para distinguir las diferentes entidades es preciso conocer sus atributos. Por ejemplo, podemos hablar de las entidades “manzana”, “pan” o “zanahoria” que pertenecen al grupo de entidades [alimentos] cuyos atributos son el “nombre en castellano”, “nombre en catalán” y “grupo de alimentos”.

Pueden crearse relaciones o asociaciones entre diferentes grupos de entidades. Por ejemplo, una relación bastante clara es la que existe entre [alimentos], [componentes] y [valores]. Para visualizar mejor el modelo puede representarse cada grupo de entidades como tablas de formato clásico. En estas tablas, cada fila o registro correspondería a una entidad y los atributos corresponderían a las columnas o campos. Las relaciones entre tablas permiten generar nuevas tablas (Figura 1-4).

Las relaciones entre grupos de entidades no siempre son del mismo tipo, depende del número de entidades con las cuales una entidad concreta pueda asociarse a través de una relación. A este concepto se le denomina “cardinalidad de la relación”. Si suponemos que A y B son dos grupos de entidades distintos, y siendo N y M dos números cualesquiera, las cardinalidades existentes pueden ser:

(1,1) Una entidad en A está asociada con una y solo una entidad en B .

(1, N) Una entidad en A está asociada con N entidades en B , pero una entidad en B solo puede asociarse con una y solo una entidad en A .

(M , N) Una entidad en A está asociada con N entidades en B , y una entidad en B está asociada con M entidades en A .

La aplicación práctica de este modelo requiere en primer lugar de un análisis lógico de la información para definir adecuadamente los grupos de entidades y sus atributos.

El MER proporciona una importante flexibilidad cuando se trata de organizar la información y facilita la incorporación de nueva información al sistema. Sin embargo, el modelo no contempla las relaciones de tipo *isa* (contracción del inglés “...is a...”) que también se dan a menudo. Por ejemplo, el grupo de entidades [Publicaciones] mantiene una relación *isa* con los grupos de entidades [Libros], [Artículos], [Informes] (un libro es *una* publicación, un artículo es *una* publicación, etc.).

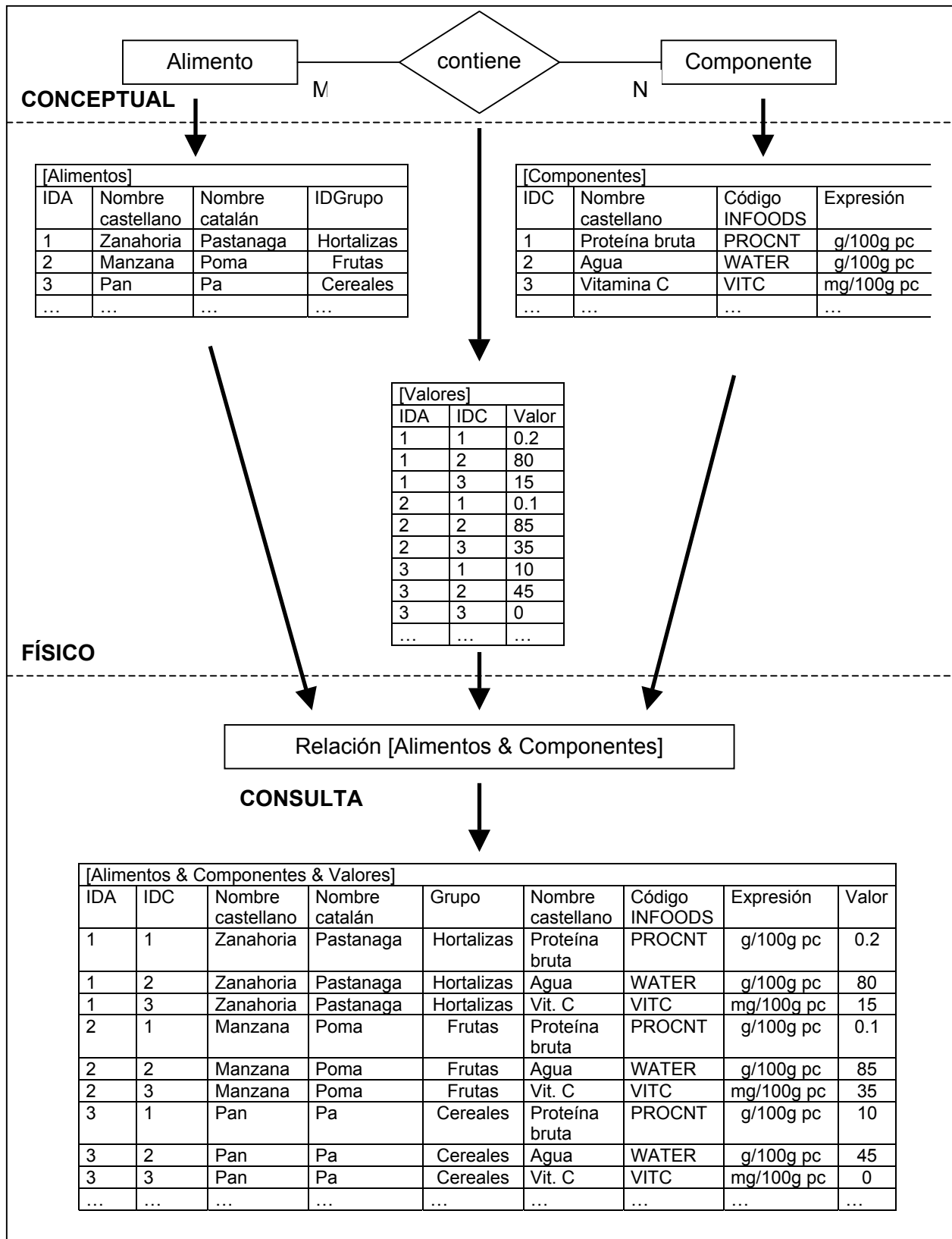


Figura 1-4 Ejemplo de modelización conceptual de la información usando un diagrama entidad-relación y cómo puede traducirse a un formato de tablas (modelo relacional). La tabla ilustra también el paso del modelo conceptual al modelo físico de datos durante el diseño de la BDCA, y la elaboración de consultas a partir de las tablas. IDA: clave o identificador del alimento; IDC: clave o identificador del componente; pc: porción comestible.

1.12.2 Modelo físico de la información: las bases de datos relacionales

Los diagramas de entidad-relación suponen un primer nivel de análisis de los datos de un sistema. Sin embargo, posteriormente el modelo conceptual se tiene que trasladar un entorno físico que está condicionado por el tipo de SGBD que utilizemos. Los modelos físicos para la información pueden ser diversos: jerárquico, en red, relacional, etc. No obstante, el modelo relacional ha tenido una fuerte expansión a partir de los años ochenta y la mayoría de SGBD existentes hoy en día son de tipo relacional (186).

Una base de datos que sigue el modelo relacional (base de datos relacional) consiste en un conjunto de tablas con diferentes campos y registros que se relacionan entre ellas. Las tablas corresponderían a las entidades del modelo conceptual y los atributos serían los campos de cada tabla. En cada tabla, además de definir los campos que contendrán la información de cada atributo, es preciso crear las claves o identificadores. Estas claves sirven para identificar de manera inequívoca y sencilla las diferentes entidades o registros de un grupo de entidades o tablas. Sin estas claves, para mencionar una entidad sería preciso listar todos sus atributos. Las claves acostumbran a ser series numéricas cuyos elementos nunca se repiten. Cuando las relaciones entre grupos de entidades se establecen a través de estas claves se habla de “claves externas” (Figura 1-4). Además, también es preciso definir las reglas de relación entre tablas a partir de las cardinalidades. Estas reglas servirán para evitar incoherencias dentro de la base de datos (89).

1.12.3 Los sistemas de gestión de bases de datos

Como ya se comentó, los SGBD son programas destinados a facilitar el uso de las Bases de datos. El desarrollo de aplicaciones informáticas que gestionan bases de datos es mucho más fácil gracias a estos sistemas pues proporcionan una serie de ventajas (186):

1. **Permiten realizar consultas no predefinidas y complejas sobre los datos.** No es preciso elaborar un programa informático cada vez que se desee realizar una consulta.
2. **Proporcionan flexibilidad e independencia.** Los usuarios y programadores pueden realizar sus tareas sin preocuparse de las características físicas de la base de datos.

3. **Eliminan la redundancia en los datos.** Permiten que el usuario pueda establecer datos redundantes, pero se encargan también de eliminar automáticamente aquellos datos repetidos que no interesa conservar.
4. **Conservan la integridad de los datos.** Evitan que se transgredan las reglas de relación entre datos establecidas al definir la base de datos.
5. **Permiten la concurrencia de usuarios.** De este modo, varios usuarios pueden consultar o modificar la base de datos simultáneamente.
6. **Seguridad.** Permiten establecer autorizaciones de acceso a diferentes niveles y sistemas de encriptación de la información.

Cualquier aplicación desarrollada en un lenguaje determinado que esté destinado a unos propósitos específicos para los cuales precise gestionar la información de una base de datos y utilice un SGBD para acceder a ella constituirá también por sí mismo un nuevo sistema de gestión de base de datos. Así, por ejemplo, el programa MS Access es un SGBD sobre el cual podemos desarrollar aplicaciones (utilizando lenguajes como Access Basic o Visual Basic) que constituirán a su vez nuevos SGBD, con usos mucho más específicos.