

EMCA
Gestión de la Calidad Asistencial

Evaluación y Mejora de la Calidad en Servicios de Salud

CONCEPTOS Y MÉTODOS

Pedro J. Saturno Hernández

Edita:
Consejería de Sanidad y Consumo
de la Región de Murcia

Imprime:
Imprenta Regional

I.S.B.N.: 84-95393-17-4

Depósito Legal: MU-2737-2000

Esta publicación no puede ser reproducida, transmitida total o parcialmente por cualquier medio electrónico o mecánico, ni por fotocopia, grabación u otro sistema de reproducción de información sin el permiso por escrito de los titulares del copyright

© Pedro J. Saturno Hernández

PRESENTACIÓN

Cuando en 1996 realizaba la presentación del Programa EMCA, manifestaba que la calidad no podía ser fruto del azar, sino del interés y la constancia en mejorar cotidianamente los servicios que, como instituciones públicas, prestamos a los ciudadanos de nuestra Región, para lo que se requiere el conocimiento y la utilización de técnicas específicas por parte de nuestros profesionales sanitarios.

La gran apuesta de esta Consejería de Sanidad y Consumo, en una productiva colaboración con la Universidad de Murcia, es un ambicioso programa de formación, asesoramiento e investigación sobre nuevos métodos y herramientas, que permita situar la Calidad como guía permanente de referencia del Sistema Sanitario, promoviendo y potenciando un cambio cultural donde el paciente se sitúa en el centro del sistema, articulando toda la estructura y funcionalidad en torno a él.

Transcurridos estos años, los más de 1000 alumnos formados, 100 cursos y seminarios celebrados, 60 proyectos de investigación financiados, y 2 Congresos Regionales celebrados pueden dar una idea de la magnitud y repercusión que la iniciativa de esta Consejería, en forma de programa institucional, ha tenido en el ámbito de la Región de Murcia.

Uno de los "productos estrella" del Programa EMCA son, sin duda, los cursos de Evaluación y Mejora de la Calidad Asistencial y Monitorización de la Calidad, verdaderas llaves de paso a los temas de Gestión de la Calidad, planificados y desarrollados con una metodología conscientemente elegida, diseñada y experimentada que, por su especial concepción, introducen y capacitan al profesional en el conocimiento teórico y práctico, mediante una docencia orientada a la solución de problemas, de forma secuencial y utilizando casos reales y cercanos.

Los materiales utilizados en dichos cursos han ido sufriendo a lo largo de estos años diferentes reestructuraciones y ampliaciones, todas con el objetivo de incrementar su adecuación e idoneidad hacia los profesionales a los que iban dirigidos, recogiendo sus sugerencias que han retroalimentado permanentemente su contenido. Así, se ha pasado de la recopilación y fotocopia de diferente bibliografía en una primera etapa, a la creación de un material de elaboración propia en este último año. Pero el deseo final, por encima de todo, era presentar en una monografía específica todo lo aprendido a lo largo de estos años.

Con la publicación de este libro que se presenta, culmina una etapa más, ese objetivo perseguido desde hace tiempo en el Programa EMCA: la de reunir en un texto único todo el conocimiento actual, rigurosamente seleccionado y tratado para, de una forma resumida y ágil, proporcionar a cualquier profesional sanitario el conocimiento y las herramientas necesarias que le permitan abordar grandes metas en la mejora de la Calidad de los servicios que presta.

El esfuerzo que la Consejería asume con su edición, constituye un reconocimiento al trabajo realizado por el Programa y a su autor, un agradecimiento a todos los profesionales de esta Región que han aceptado y asumido la política de esta Consejería con la mejora de la Calidad, y un nuevo compromiso público de esta Institución cuyo fin último es el de situar al más alto nivel los servicios recibidos por sus ciudadanos.

Francisco Marqués Fernández
Consejero de Sanidad y Consumo

INDICE

UNIDADES TEMÁTICAS

| | |
|---|-----|
| 1. CONCEPTO DE CALIDAD ASISTENCIAL | 7 |
| 2. EL CONTEXTO DE LOS SERVICIOS DE SALUD COMO MARCO DE REFERENCIA PARA LA MEJORA DE LA CALIDAD | 23 |
| 3. PROGRAMAS DE GESTIÓN DE LA CALIDAD: CONCEPTO, ACTIVIDADES Y UTILIDAD | 31 |
| 4. PROGRAMAS DE GESTIÓN DE LA CALIDAD: ENFOQUES Y RESPONSABILIDADES | 43 |
| 5. PROGRAMAS DE GESTIÓN DE LA CALIDAD: ESTRATEGIAS DE IMPLANTACIÓN | 57 |
| 6. IDENTIFICACIÓN Y PRIORIZACIÓN DE OPORTUNIDADES DE MEJORA | 71 |
| 7. MÉTODOS DE ANÁLISIS DE LOS PROBLEMAS DE CALIDAD | 93 |
| 8. LA CONSTRUCCIÓN DE CRITERIOS PARA EVALUAR LA CALIDAD | 111 |
| 9. DISEÑO DE ESTUDIOS DEL NIVEL DE CALIDAD. COMPONENTES DE UNA EVALUACIÓN | 133 |
| 10. DISEÑO DE ESTUDIOS DE INVESTIGACIÓN DE CAUSAS DE LA CALIDAD DEFICIENTE | 153 |
| 11. ANÁLISIS Y PRESENTACIÓN DE LOS DATOS DE UNA EVALUACIÓN | 171 |
| 12. DISEÑO DE INTERVENCIONES PARA MEJORAR | 197 |
| 13. REEVALUACIÓN, ANÁLISIS Y PRESENTACIÓN DE RESULTADOS COMPARATIVOS DE DOS EVALUACIONES | 213 |
| 14. ACTIVIDADES PARA LA MONITORIZACIÓN. CONSTRUCCIÓN Y ANÁLISIS DE INDICADORES. PLANES DE MONITORIZACIÓN | 239 |
| 15. EL MUESTREO PARA LA ACEPTACIÓN DE LOTES (LQAS) COMO MÉTODO DE MONITORIZACIÓN | 255 |
| 16. LOS GRÁFICOS DE CONTROL ESTADÍSTICO DE LA CALIDAD: GRÁFICOS DE DESARROLLO | 277 |
| 17. LOS GRÁFICOS DE CONTROL ESTADÍSTICO DE LA CALIDAD: GRÁFICOS DE CONTROL | 303 |
| 18. PROGRAMAS DE MONITORIZACIÓN EXTERNA DE INDICADORES. ANÁLISIS DE PERFILES. AJUSTE Y ESTANDARIZACIÓN DE INDICADORES COMPARATIVOS | 321 |
| 19. INTRODUCCIÓN A LA PLANIFICACIÓN O DISEÑO DE LA CALIDAD | 347 |
| FE DE ERRATAS | 361 |

UNIDAD TEMÁTICA 1

CONCEPTO DE CALIDAD ASISTENCIAL

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

En esta Unidad Temática se da un marco de referencia global para las actividades de Evaluación y Mejora de la Calidad. Se aborda la definición de calidad, necesaria para dar sentido (validez) a los aspectos que vamos a medir y mejorar.

ÍNDICE DE CONTENIDOS

1. Introducción: marco de referencia y contexto para la mejora de la calidad.
2. ¿Por qué preocuparse por la calidad?
3. ¿Por qué y cómo definir la calidad?
4. El concepto general de calidad.
5. El concepto de calidad de la atención sanitaria.
6. El concepto de calidad de un determinado servicio.
7. ¿Y la satisfacción del profesional? Concepto de "cliente interno".
8. El concepto de calidad total.

OBJETIVOS ESPECÍFICOS

1. Justificar el interés por la calidad de los proveedores, gestores y usuarios del sistema de salud.
2. Diferenciar definiciones genéricas de calidad, aplicables a cualquier tipo de producto o servicio, de las definiciones específicas de actividades y de productos.
3. Identificar o construir una definición de calidad de los servicios de salud, aplicable a su ámbito de trabajo (centro y sistema).
4. Describir las dimensiones o aspectos generales que entraña el concepto de calidad de los servicios de salud.
5. Distinguir los conceptos de cliente interno y cliente externo, aplicados a los servicios de salud.
6. Analizar críticamente terminologías y conceptos generales en relación a la calidad asistencial.
7. Justificar la importancia de la satisfacción de los clientes externos e internos en la conceptualización de la calidad.

1. INTRODUCCIÓN: MARCO DE REFERENCIA Y CONTEXTO PARA LA MEJORA DE LA CALIDAD

Vamos a empezar la casa por los cimientos. El objetivo central de la formación es adquirir los conocimientos y habilidades metodológicas para llevar a la práctica el compromiso con la mejora de la calidad que todos llevamos dentro; en esta tarea vamos a tener que ocuparnos de *medir* la calidad y entender las causas, tanto las modificables como las no fácilmente modificables, que influyen en ella. Podríamos empezar por ahí pero ello supondría que sabemos y estamos de acuerdo en las respuestas a las siguientes preguntas previas y básicas: ¿por qué preocuparnos por la calidad?; ¿qué entenderemos por calidad?; ¿es posible y necesario medirla?; ¿qué tiene que ver la medición de la calidad con la función de los centros sanitarios?; ¿y con su desarrollo futuro? En esta Unidad Temática (UT) vamos a responder a estas preguntas, lo cual significa cimentar, dar validez y sentido a nuestros esfuerzos para medir y mejorar la calidad.

Comenzaremos por justificar el interés existente por la calidad, inicialmente entendido como ente abstracto, para después definirla a tres niveles sucesivos de especificación, el último de los cuales conduce a su medición, y consideraremos a continuación los principales conceptos y precauciones generales a tener en cuenta al medirla. Finalmente, conectaremos el interés, la calidad, su medición y su mejora, con el quehacer actual y futuro de los centros, según es percibido por sus protagonistas principales: los profesionales que trabajan en ellos.

2. ¿POR QUÉ PREOCUPARSE POR LA CALIDAD?

Es evidente que existe un interés creciente por la calidad en todos los sectores y también en el sector salud. Un interés que afecta, además, a todos los implicados: pacientes o población receptora de los servicios, gestores, y profesionales de la salud. La población está cada vez mejor informada, exige más y reclama cada vez más sobre la calidad de los servicios; en algunos lugares de Estados Unidos se tiene acceso incluso al historial académico y profesional de los médicos. Los gestores, sobre la base de una preocupación por los costes en esta situación de crisis permanente en la que parece que está el sistema de salud, llega un momento en que se preguntan: ¿de qué?, es decir, ¿en qué se gasta el dinero?, ¿servicios necesarios o innecesarios?, ¿útiles o inútiles?, ¿con buena calidad o con mala calidad?; además muchos de ellos saben, o han oído decir, que hacer bien las cosas cuesta menos que hacerlas mal. Finalmente, los profesionales de la salud siempre han estado interesados en la calidad de forma al menos implícita, aunque sólo fuese por cuestiones de ética, que se relacionan con tres cuestiones fundamentales: (i) responder a la confianza que los pacientes depositan en ellos, (ii) para ser consecuentes con el compromiso científico que va inherente a la práctica de las ciencias de la salud; y (iii) por la responsabilidad de mantener la competencia profesional. Por todos estos motivos, los profesionales y los proveedores directos de los servicios en general son y deben ser los primeros en enarbolar y defender la bandera de la calidad, para que, entre otras cosas, no se vea afectada por recortes indiscriminados del gasto sanitario.

Cabe señalar que las primeras iniciativas formales de evaluación y mejora de la calidad tuvieron lugar en salud, y las protagonizaron profesionales preocupados por los resultados de la atención que recibían sus pacientes. Después ha sido

El concepto de calidad y el contexto de los centros sanitarios, vistos por sus profesionales, van a ser la base para las actividades de evaluación y mejora.

El interés por la calidad es compartido por usuarios, gestores y profesionales de la salud. En estos últimos la ética profesional supone un compromiso implícito con la calidad.

en la industria en donde ha habido un desarrollo metodológico más intenso y fecundo, motivado por la consecución de una mayor productividad, calidad y consiguientes beneficios para la empresa; pero los nombres de Florence Nightingale, enfermera, y Ernest Codman, cirujano, preceden a la aparición del interés formal por la calidad en la industria, y sus objetivos eran prioritariamente responder a los imperativos de la ética profesional y conseguir el mayor beneficio posible en la salud de los pacientes.

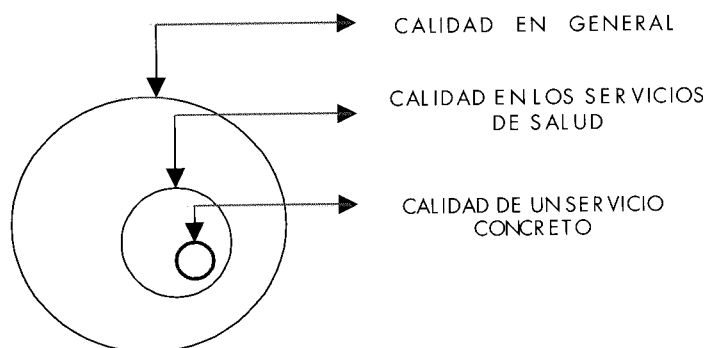
3. ¿POR QUÉ Y CÓMO DEFINIR LA CALIDAD?

Definir lo que se entiende por calidad de la atención no es una disquisición teórica irrelevante para la práctica, si lo que pretendemos es explicitar el marco de referencia para su evaluación y mejora. Las más elementales reglas de validez exigen tener claro qué queremos decir con "calidad de la atención", cuando va a ser ese concepto el que vamos a cuantificar. Asimismo, la gran cantidad de definiciones y matices existentes, convierte la conceptualización operativa de esta idea en una decisión consciente entre las diversas alternativas. Las variables, los factores, los componentes o las dimensiones, que van a ser medibles para caracterizar los niveles de calidad, y sobre las que enfocaremos nuestros esfuerzos de mejora son, asimismo, una consecuencia directa de la definición que escojamos. En resumen, vamos a necesitar definir para medir y medir para saber cómo podemos intervenir para mejorar aquello que previamente hemos definido.

Como punto de partida, la mejor definición de calidad de la atención que podemos ofrecer sería una hoja en blanco. Y no porque no existan definiciones, sino todo lo contrario, hay tantas y con una traducción práctica tan diferente que es conveniente subrayar que cada centro, institución o sistema va a tener que crear de manera indefectible, la suya propia o decantarse por la adopción y/o adaptación de una determinada. Una definición inicialmente en blanco sirve también para subrayar que se trata de un concepto en evolución, que ha ido perfilándose y enriqueciéndose, y que, casi con seguridad, va a continuar evolucionando.

¿Cómo proceder para hacer explícita nuestra definición? Para orientarnos en la concreción de lo que vamos a entender por calidad es conveniente distinguir tres niveles, de más general a más específico, tal como está representado en la Figura 1.1.

FIGURA 1.1. Los tres niveles del concepto de calidad



Necesitamos definir qué es calidad para, entre otras cosas, dar validez a su medición y mejora.

Un concepto general de calidad sería aplicable a cualquier producto o servicio de cualquier actividad o empresa; un segundo nivel de especificación sería una definición de calidad para un tipo determinado de productos o servicios, a nivel agrupado, genérico, por ejemplo calidad de la atención sanitaria o servicios de salud en general; un tercer nivel sería especificar qué es calidad en referencia a un producto o servicio concreto, como podrían ser cada uno de los muchos servicios diferentes que ofrecen los diferentes centros sanitarios (atención al infarto agudo de miocardio, al paciente diabético, niño sano, vacunaciones, urgencias, ...). Los dos primeros niveles más generales son marcos de referencia necesarios, que orientan pero no sirven para medir, mientras que el nivel más específico debe contener los elementos que nos pueden servir para medir hasta qué punto ese servicio concreto es o no de calidad.

Como es lógico, los tres niveles de conceptualización han de ser consistentes entre sí. Vamos a ver en qué puede consistir cada uno de estos tres niveles de conceptualización, desarrollando sobre todo el nivel más específico, que es el que vamos a medir.

4. EL CONCEPTO GENERAL DE CALIDAD

De una manera amplia, el concepto de *calidad* puede aplicarse a muy diversos aspectos de una determinada actividad industrial o de servicios, desde niveles filosóficos globales, a calidad de los procesos, la organización del personal, de los objetivos de la empresa, etc. Aquí nos vamos a referir exclusivamente a calidad *del producto o servicio* que se ofrece, objetivo o resultado último en el que confluyen las "calidades" de todos los elementos de la actividad industrial o de servicios que estemos considerando. La pregunta a responder sería: ¿cómo sabremos si un producto o servicio es de calidad? Veamos el ejemplo de las respuestas de algunos de los más influyentes autores sobre la calidad de la industria: Juran, Deming e Ishikawa.

Juran define la calidad como "idoneidad o aptitud para el uso". Un producto o servicio será de calidad si sirve para lo que está previsto que sirva. La aplicación de este concepto a un servicio o producto concreto exige una definición previa de qué necesidades y expectativas está previsto que sean satisfechas con el producto o servicio a evaluar; las cuales serían diferentes para según qué producto o servicio y determinarían las características que éste ha de tener para que sea de calidad.

Deming propone definir la calidad "en función del sujeto que ha de juzgar el producto o servicio". La idea básica es la misma: quien compra o demanda el servicio lo hace en función de unas necesidades y expectativas que son las que van a determinar si lo que recibe es o no de calidad.

Ishikawa utiliza la misma idea: la calidad del producto o servicio es la "satisfacción de los requisitos de los consumidores de ese producto o servicio". Hay que determinar, pues, cuáles son esos requisitos para poder saber si el producto o servicio es de calidad.

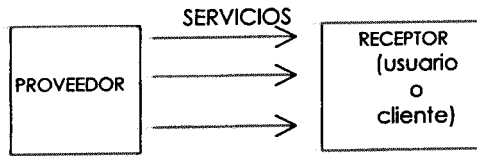
Como puede apreciarse hay un esquema básico subyacente a estas definiciones, que nos puede servir para describir y analizar la calidad de cualquier producto o servicio. Este esquema básico parte de tres componentes que hay que identificar y caracterizar: proveedor, servicio o producto que se ofrece y receptor

La calidad puede definirse a varios niveles. Vamos a identificar tres y a desarrollar el más específico, que es el que mediremos.

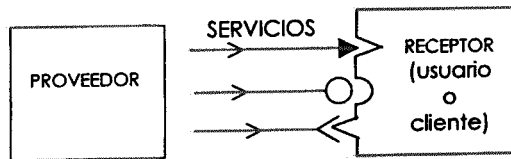
Nos centramos en definir calidad del producto o servicio, no de la organización u otros aspectos del sistema.

del servicio o producto. Para que haya calidad, el que ofrece el servicio ha de tener en cuenta las necesidades y expectativas de a quién va dirigido, y modelar el servicio o producto de manera que sintonicen perfectamente con estas últimas. La falta de sintonización será defecto de calidad. Este esquema básico está representado en la Figura 1.2.

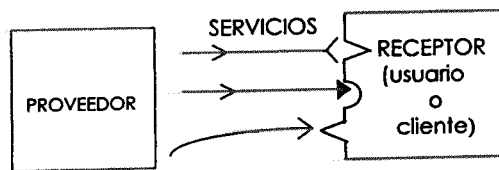
FIGURA 1.2. Esquema básico para definir y analizar la calidad



SERVICIOS DE ALTA CALIDAD: Responden a las necesidades y expectativas de los usuarios



SERVICIOS DE BAJA CALIDAD: No responden a las necesidades y expectativas de los usuarios



Una consecuencia de este esquema, aparte de necesitar de más concreción para que sirva para medir, es que la calidad, va a depender de lo correcta que sea la *interpretación* que hagamos de las necesidades y expectativas de los usuarios o clientes. Algo que explícita o implícitamente hacen todas las empresas públicas o privadas, aunque con diferencias a veces muy grandes entre ellas que pueden determinar el éxito o el fracaso de la empresa. Pero la primera gran diferencia es, naturalmente, el tipo de producto que se ofrece: no es lo mismo ofrecer automóviles que lavadoras, comida o servicios de salud. Es posible entonces, especificar un poco más la definición de calidad en el sentido de hacer explícito lo que entendemos que ha de ser un producto o servicio de calidad de nuestro tipo de "empresa" en concreto.

5. EL CONCEPTO DE CALIDAD DE LA ATENCIÓN SANITARIA

Hay múltiples definiciones de lo que se entiende por calidad de los servicios de salud. Todas responden a diversas interpretaciones sobre qué significa atender las necesidades de atención sanitaria de la población a la que se ofrecen los servicios cuya calidad queremos definir.

Revisando la variedad de definiciones existentes, se hace evidente una primera conclusión: no es posible elegir una y darla por universalmente válida, porque este tipo de definición no existe. Observamos, en cambio, que hay que tomar ciertas decisiones para perfilar el concepto de calidad que más se acomode a la misión y circunstancias del entorno en que va a ser aplicado. Vamos a ver cuáles son estas decisiones que dan carácter a la definición a la vez que analizamos la

En general, la definición de calidad de un producto o servicio ha de hacerse en función de las necesidades y expectativas de quien lo recibe.

que se adoptó para el Programa Ibérico, basándose en la adaptación que hizo R.H. Palmer de la definición propuesta por el Instituto de Medicina de Estados Unidos.

En la definición adoptada en el Programa Ibérico, calidad de la atención es “la provisión de servicios accesibles y equitativos, con un nivel profesional óptimo, que tiene en cuenta los recursos disponibles y logra la adhesión y satisfacción del usuario”. Esta definición incluye componentes incuestionables, puesto que la mayoría de las definiciones los contienen de forma explícita o implícita, como son el nivel profesional o *calidad científico-técnica*, y la adhesión y *satisfacción* (aceptabilidad como término más global) del usuario. La inclusión de la *accesibilidad*, la *equidad* y la *limitación impuesta por los recursos existentes* han de ser, sin embargo, argumentados y defendidos por quienes adopten esta definición.

Con relación a la *accesibilidad*, tan importante para las instituciones y sistemas con vocación de cobertura universal y enfocados a la atención de la comunidad, se cuestiona como componente del concepto de calidad de la atención en definiciones como las de Donabedian y la más reciente del Instituto de Medicina de EE.UU. (IOM). En éstos, como en otros casos, en los que no se considera la accesibilidad como parte del concepto de calidad, lo que parece que se obvia, de alguna manera, es gran parte de la dimensión social de la atención sanitaria, sesgando la conceptualización hacia la atención al individuo enfermo que “libremente” puede solicitar o no atención médica. En definiciones como la de Palmer, como ha sido interpretada para el Programa Ibérico y como pensamos que es aplicable en sistemas de atención con orientación comunitaria, se subraya la responsabilidad social del sistema y de sus profesionales. En nuestra opinión, con independencia de las decisiones individuales para solicitar o no asistencia, una atención de calidad óptima a la que sólo tiene acceso una parte de la población que la necesita, no puede considerarse de calidad desde un punto de vista social, comunitario. Incluso, en cuanto al usuario individual, el hecho de que los cuidados que precisa recibir sean más o menos accesibles, la medida en que se contribuya a que no existan obstáculos para restaurar, mantener o fomentar su salud, ha de formar parte de la calidad de la atención que puede ofrecerle el sistema y sus profesionales.

La inclusión del componente de *equidad* responde, asimismo, a un acercamiento comunitario en la definición, argumentable de forma parecida e íntimamente ligado a la accesibilidad. Se trata no sólo de que los cuidados científicamente óptimos sean accesibles, sino también de que el sistema y sus profesionales estén en disposición de ofrecer mayor atención a quién más la necesita, e igual atención a igual necesidad.

Sin embargo, en sistemas sin vocación de cobertura universal, o en entidades privadas sin responsabilidades de atención a grupos poblacionales definidos, puede prescindirse perfectamente de la accesibilidad y de la equidad como componentes de lo que es un servicio de salud de calidad, y ceñirse a los servicios que ofrecen y pueden obtener los clientes que pueden acceder a ellos.

Tener en cuenta *los recursos de que se dispone* al definir la calidad es aún más controvertido. De hecho, podríamos muy probablemente separar todas las definiciones existentes en dos grandes grupos, según contengan o no esta referencia. Su inclusión supone considerar la calidad como algo *relativo*, necesariamente peculiar de cada entorno; supone aceptar que no es comparable, en términos

No hay un único concepto de calidad de los servicios de salud. Cada institución debe elegir conscientemente el suyo. Aspectos como la accesibilidad, equidad y optimización de los recursos no son considerados en todas las definiciones.

absolutos, lo que puede exigirse a un hospital canadiense de lo que a uno en Centroamérica, a uno español o a otro en cualquier país africano: existe el límite que establecen los recursos que la sociedad y los individuos han decidido emplear en la provisión de servicios de salud, como apostillaba la definición adoptada por el IOM en 1974. Uno de los argumentos es que el estándar de calidad óptima está establecido por la sociedad y puede estar en conflicto con los estándares profesionales que "rara vez se definen en términos económicos" y "obedecen a los principios de beneficio y autonomía más que al principio de justicia distributiva".

Los principales argumentos parecen estar claros; corresponde a cada profesional, institución y sistema asumir unos u otros.

La definición que hemos analizado, como cualquier otra que decidamos adoptar como la más apropiada para nuestra "empresa" de provisión de servicios de salud, explicita una serie de dimensiones que nos ponen sobre la pista de hacia dónde mirar para valorar si hay o no calidad, utilidad del servicio en general. Es lógico que, al existir una gran variedad de definiciones, existan también un apreciable variedad de esquemas de dimensiones, factores, componentes o atributos (que de todas estas formas se les llama). La Tabla 1.1 contiene algunos de estos esquemas. El lector interesado puede consultar la bibliografía que se cita en esta UT; lo que vamos a hacer a continuación es revisar el significado de las dimensiones más habituales y terminar proponiendo un esquema de sólo tres, normalmente suficiente como marco de referencia, y que resulta acorde con la definición que hemos propuesto.

TABLA 1.1. Dimensiones de la calidad de la atención según diversos esquemas

| Palmer (1983) | Programa Ibérico (1990) | Vuori (componente) (1982) | Nutting et al (atributos) (1987) | Joint Comm. (componente) (1990) | Joint Comm. (factores) (1989) | Joint Comm. (características) (1990) |
|-----------------------------|----------------------------|----------------------------|--|---------------------------------|-------------------------------|--------------------------------------|
| Competencia profesional * | Calidad científico-técnica | Calidad científico-técnica | Competencia profesional* | — | — | — |
| Accesibilidad* | Accesibilidad | — | Distribución de la asistencia/Integralidad | Accesibilidad/Disponibilidad | Accesibilidad | Accesibilidad |
| Aceptabilidad/Satisfacción* | Satisfacción | — | — | Aceptabilidad | Participación | Accesibilidad |
| Efectividad | — | Efectividad | Efectividad | Efectividad | Efectividad | Efectividad |
| Eficiencia | — | Eficiencia | Eficiencia | Eficiencia | Eficiencia | Eficiencia |
| — | — | Adecuación | — | Adecuación | Adecuación | Adecuación |
| — | — | — | Continuidad/Coordinación | Continuidad | Continuidad | Continuidad |
| — | — | — | — | — | Eficacia | Eficacia |
| — | — | — | — | — | Oportunidad/temporal | — |
| — | — | — | — | — | Privacidad | — |
| — | — | — | — | — | Confidencialidad | — |
| — | — | — | — | — | Seguridad | — |
| — | — | — | — | — | Apoyo estructural | — |

5.1. LAS DIMENSIONES MEDIBLES

Para seleccionar aquellas que nos parezcan más convincentes, conviene repasar los conceptos a que se refieren los principales dimensiones de la Tabla 1.1: calidad científico técnica, efectividad, eficiencia, accesibilidad, satisfacción, adecuación y continuidad.

- La competencia profesional, **calidad científico-técnica** de la atención, es la dimensión central y menos discutible de la calidad, y hace referencia a la capacidad de los proveedores de utilizar el más avanzado nivel de conocimientos existente para abordar los problemas de salud, es decir, para producir salud y satisfacción en los usuarios. En la mayoría de las definiciones, incluye tanto los aspectos científico-técnicos, en sentido estricto, como los de trato interpersonal. Sin lugar a dudas, es la dimensión que mejor se entiende y más frecuentemente medida como representante de la calidad de los servicios de salud; significa atender de forma científica las necesidades sanitarias.
- **Efectividad** es la medida en que una determinada práctica o atención sanitaria mejora el estado de salud de la población concreta que la recibe. Durante algún tiempo ha venido utilizándose de forma confusa o intercambiable con el término *eficacia*, aunque en la actualidad, partiendo de las propuestas de White refrendadas con posterioridad por la OTA (*Office of Technology Assessment*) estadounidense, se entiende por eficacia “la probabilidad de beneficio de una determinada tecnología en condiciones ideales de uso”; mientras que efectividad sería “la probabilidad de beneficio de una determinada tecnología en condiciones de uso normales”. Lo más frecuente al definir calidad es la referencia a la efectividad de la atención.
- **Eficiencia** se define en términos de la relación entre coste y producto es decir, un máximo de efectividad o unidades de producto, dado un determinado coste, o un mínimo coste dadas unas determinadas exigencias de efectividad o unidades de producto. Por supuesto se asume que una mayor eficiencia proporciona una mayor calidad.
- La **accesibilidad** puede definirse como la facilidad con que la atención sanitaria puede obtenerse en relación con los aspectos (barreras) organizacionales, económicos, culturales y emocionales. Ya hemos mencionado que esta dimensión no se incluye en todas las definiciones de calidad. También es cierto que aquellas que la incluyen introducen matices, a veces, muy diferentes en su conceptualización operativa, medible. Palmer incluye en ella el concepto de equidad (accesibilidad es “la medida en que los cuidados apropiados son recibidos por la población de forma equitativa” o “la distribución equitativa y oportuna (a tiempo) de la atención sanitaria apropiada, para aquéllos con necesidades equivalentes”). En todo caso, lo que subyace es la necesidad de cuantificar si la atención sanitaria llega o no a quien la necesita y cuando la necesita. Uno de los criterios o indicadores de calidad más directo y con frecuencia medidos en relación con la accesibilidad es la cobertura; por ejemplo, en los programas de crónicos o en la evaluación de medidas preventivas (vacunación de hepatitis B, mamografías para detección precoz del cáncer de mama, ...), baja cobertura equivale a mala calidad por probable defecto en la accesibilidad.

Un matiz importante que podemos introducir es aplicar el concepto de accesibilidad no sólo a la obtención de servicios iniciales o contacto “principal”

Las dimensiones de la calidad nos indican qué valorar para saber si está presente.

con el sistema de salud, sino también a la existencia o no de barreras organizativas, culturales o de otro tipo una vez que se ha accedido al sistema de salud; este segundo aspecto lo podríamos llamar *accesibilidad interna* y englobaría toda la problemática de los tiempos de espera, falta de coordinación entre profesionales y entre niveles asistenciales, etc., que impiden la fluidez normal que debería tener un proceso asistencial visto en su totalidad.

- La **satisfacción** puede definirse como la medida en que la atención sanitaria y el estado de salud resultante cumplen con las expectativas del usuario. Según las distintas definiciones de calidad, esta dimensión se incluye o se confunde con el concepto más amplio de aceptabilidad. En todo caso, llámese satisfacción o aceptabilidad, de lo que se trata es de tener en cuenta la manera en que la atención es percibida por el usuario y sus reacciones como consecuencia de ello. Con algunas excepciones, la mayoría de las definiciones dan lugar prominente a esta dimensión al medir la calidad, a la vez que discuten sus variables, subcomponentes y la gran variedad de ajustes que hay que tener en cuenta para interpretar de forma correcta su medición. Este punto será revisado más en profundidad en la UT4 .
- La **adecuación** (*appropriateness*) es otra de las dimensiones mencionadas con frecuencia, pero de la que no existe una definición operativa uniforme. En la lista de la JCAHO (Joint Commission on Accreditation of Healthcare Organizations) estadounidense, la adecuación es la medida en que la atención se corresponde con las necesidades del paciente, es decir, adecuado como sinónimo de correcto, conveniente o necesario para el problema concreto que es atendido. Éste es un concepto que parece muy difícil de separar a la hora de ser medido, del concepto de calidad científico-técnica o competencia profesional.
- La **continuidad** es, con toda probabilidad, la dimensión de la calidad de la atención más compleja de conceptualizar y medir. De entrada, ya resulta inquietante que sea la dimensión sobre la que existen más versiones diferentes. La conceptualización más prevalente y atractiva, parece ser que es la que propuso Shortell: la medida en que la atención necesaria se provee de una forma ininterrumpida y coordinada, en cierto modo equivalente a lo que hemos llamado *accesibilidad interna*.

En la Tabla 1.1 se incluyen esquemas que contienen dimensiones, atributos o factores de la calidad diferentes a los comentados hasta aquí. Revisar cada uno de ellos sería muy prolijo, máxime cuando pensamos que lo más lógico es intentar simplificar.

5.2. ¿SE PUEDE SIMPLIFICAR? RAZONES Y ARGUMENTOS DE UN ESQUEMA MÁS SENCILLO: MEDIR PARA ACTUAR

El esquema de dimensiones adoptado se reduce a tres: calidad científico-técnica o competencia profesional, satisfacción y accesibilidad. ¿Por qué puede prescindirse de los conceptos de efectividad, eficiencia, adecuación y continuidad como dimensiones singulares de medición directa de la calidad? No es porque se considere que estas dimensiones no puedan estar directamente relacionadas con la calidad de la atención. Las razones fundamentales tienen que ver con la práctica de las evaluaciones sobre calidad de la atención, si se trata de evaluar la calidad para determinar qué hacer para mejorarla.

Proponemos un esquema de tres dimensiones: calidad científico-técnica, accesibilidad y satisfacción

Tomemos, por ejemplo, la efectividad y la eficiencia. Desde luego que más efectividad y más eficiencia son más calidad; pero, efectividad y eficiencia ¿de qué?, provocadas o decididas ¿por quién? La efectividad y la eficiencia son atributos de las acciones diagnósticas y terapéuticas, en sentido amplio, que los proveedores de los servicios han decidido utilizar o poner en marcha para la resolución de un determinado problema de salud. Depende de la competencia profesional del proveedor haber seleccionado y ejecutado las acciones más efectivas y eficientes que convenían al problema afrontado, con las limitaciones de los recursos disponibles. Es decir, *al medir efectividad y eficiencia medimos la calidad científico-técnica o la competencia profesional para actuar de forma adecuada.*

En relación a otras dimensiones finalmente no incluidas en nuestra propuesta de sólo tres ya hemos argumentado antes lo difícil que resulta considerar la *adecuación* de forma separada de la calidad científico-técnica. En relación con la *continuidad*, de indiscutible relevancia como concepto a la hora de medirla para mejorar la calidad, vamos a vernos obligados a pormenorizarla, pulverizarla casi, en sus diversos componentes o (sub)dimensiones. Componentes identificables en su mayoría como cuestiones que miden, en realidad, accesibilidad, satisfacción o competencia profesional, con mayor o menor importancia relativa, según el problema de salud que estemos evaluando y el aspecto de la continuidad que valoremos.

El entender y adoptar conscientemente una definición de calidad de los servicios de salud, de la que podemos deducir unas dimensiones que interpretan lo que pensamos que da satisfacción a las necesidades y expectativas de la población atendida, nos confiere un importante marco de referencia, pero aún insuficiente como instrumento para la medición directa de la calidad. Son muchos los servicios diferentes que ofrece cualquier centro sanitario y es necesario traducir estas dimensiones generales a requisitos *concretos* para cada servicio, problema de salud o tipo de paciente, de forma que podamos valorar si es o no de calidad. En cada tipo de paciente la importancia relativa de cada dimensión puede ser diferente, así como los requisitos de calidad que las caractericen; piénsese por ejemplo lo muy diferentes que han de ser los requisitos de calidad científico-técnica para las distintas enfermedades crónicas, que para procedimientos agudos, autolimitados, o para problemas de tipo preventivo. Por eso, vamos a necesitar de la definición de estos requisitos específicos de cada servicio o tipo de paciente para poder, finalmente, cuantificar la calidad.

Antes de pasar a considerar este tercer nivel, más específico y operativo, del concepto de calidad, conviene subrayar una serie de cuestiones controvertidas que es frecuente que nos surjan, y que deben quedar claras y resueltas al nivel de concepto más general de lo que entendemos por calidad en los servicios de salud.

5.3. LOS PUNTOS DE CONTROVERSIA

a) El conflicto entre calidad científico-técnica y satisfacción del usuario

Todos sabemos que los pacientes pueden estar muy satisfechos aunque la atención que reciban no sea del todo adecuada desde el punto de vista científico; también sabemos que pueden quedar descontentos aún haciendo lo que es cientí-

La definición general de calidad de los servicios de salud es aún insuficiente para medir. Tenemos que especificar cómo se traducen las dimensiones generales en requisitos de calidad de cada producto o servicio que se ofrece.

Calidad científico-técnica y satisfacción no son dimensiones contrapuestas sino simultáneas y complementarias al valorar la calidad.

La valoración de la calidad ha de hacerse sobre los datos objetivos de los requisitos de calidad de servicios específicos.

ficamente correcto, porque a veces ellos quieren o esperan algo que puede incluso ser perjudicial o como mínimo irrelevante o innecesario. ¿Qué es lo más importante?, ¿qué debemos elegir?, ¿en cuál de las dos situaciones hay más calidad: satisfacción con errores científico-técnicos o insatisfacción con una calidad científica impecable? Si lo planteamos de esta forma, está claro que no hemos entendido o que no hemos asumido el concepto de calidad propuesto. Aunque en determinadas situaciones la calidad científico-técnica debería prevalecer ("primero no dañar"), calidad no significa calidad científico-técnica o satisfacción, sino calidad científico-técnica **y** satisfacción; no es conveniente considerarlas de forma contrapuesta sino a la vez: si falla alguna de las dos, la calidad es mejorable; si para aumentar una de ellas provocamos una disminución de la otra, la estrategia empleada no es probablemente completa y, desde luego, no la mejor.

b) Calidad aparente/real; calidad global/específica

Dos lugares comunes y equivocados que se encuentran frecuentemente al hablar de calidad es valorar la calidad en base a las apariencias en vez de utilizando datos objetivos, y juzgar o intentar juzgar la calidad general, global, de un determinado centro o establecimiento sanitario, en vez de revisar cual es la calidad de los servicios concretos que ofrece. Un ejemplo de calidad "aparente" es la impresión que nos puede causar una mayor o menor sofisticación o cantidad del equipamiento con que cuenten los hospitales y hacer esta impresión equivalente a la calidad del centro, cuando lo importante es saber si en la provisión de servicios concretos este equipamiento se utiliza convenientemente, es decir, la calidad real y objetiva; puede ocurrir que en centros con peor aspecto o con menos medios, la calidad *real* sea mejor. De igual forma, el mito de la calidad "global" es a veces difícil de superar, por lo atractivo que resulta; sin embargo, es un hecho que cualquier centro ofrece una gran variedad de servicios y no hay ningún indicador que pueda resumir la calidad de todos; en cambio, lo correcto sería llevar la discusión al nivel de la calidad *específica* de cada uno de los servicios concretos que se ofrecen.

c) La calidad como algo subjetivo.

Una discusión clásica al definir calidad de la atención es la formulación diferente, sesgada, que puede producirse según cual sea el punto de vista –usuarios, proveedores directos, administración– que predomine. En general, si se llevan estas diferencias a sus rasgos más gruesos, caricaturescos, los profesionales aprecian, en primer lugar, un atención científicamente correcta, dentro de las limitaciones de la tecnología existente, y casi siempre sesgada a la producción de mejora del nivel de salud de cada paciente individualizado, en cuyo interés actúan; encuentran difíciles de incorporar las dimensiones sociales y poblacionales, y los criterios de eficiencia en las decisiones clínicas frente al paciente individual; por el contrario, estas últimas dimensiones serían las que habrían de valorar los gestores, la administración en general, como primordiales. Los usuarios, en cambio, tienden a coincidir con los profesionales en el rechazo a considerar las limitaciones de recursos en la atención al paciente individual, y aprecian como determinantes de la calidad el trato interpersonal, la accesibilidad y los resultados concretos de mejora en su salud y bienestar a corto plazo, más que consideraciones de índole científico-técnica en sentido estricto, sobre las cuales es obvio que no tienen, en general, elementos de juicio suficientes para valorarlas.

La calidad deja de ser subjetiva cuando la definimos.

Estos hechos evidentes, llevan muchas veces a concluir que la calidad es “subjetiva” y por lo tanto su medición va a estar siempre sesgada y su conceptualización es cuestionable. Como punto de partida puede no estar mal este cuestionamiento; sin embargo, una vez que especifiquemos en qué consiste la calidad, y definamos una serie de dimensiones y requisitos que son los que nos van a decir si está o no presente, la calidad deja de ser algo subjetivo y pasa a ser, como mínimo, *objetivable*. ¿Sesgadamente?, en cualquier caso una vez que explicitemos qué es calidad, es cuando estaremos en condiciones de añadir o quitar requisitos o dimensiones para aunar todos los puntos de vista posibles y necesarios. La calidad permanecerá como algo subjetivo e inmanejable *en tanto en cuanto no sea definida y medida*; y será tanto más objetiva y mejorable cuanto más descendamos en el nivel de especificación del concepto. La única calidad subjetiva es la que no se define.

6. EL CONCEPTO DE CALIDAD DE UN DETERMINADO SERVICIO

Consideremos ahora un tercer nivel de especificación, que es el que resulta realmente operativo, y por lo tanto el que vamos a manejar de manera preferente para su evaluación y mejora: la definición de calidad de un determinado producto o servicio dentro de los muchos que puede ofrecer nuestro centro. Esta definición de calidad es diferente y peculiar para cada servicio y se puede llegar a ella tratando de responder a las siguientes preguntas: para un servicio concreto, por ejemplo consulta de diabetes, la resolución de una solicitud de cita previa o disponer de la historia clínica en la consulta en el momento de la visita ¿qué quiere decir que sea accesible?, ¿cómo me aseguro de que no tiene problemas de accesibilidad?, ¿qué quiere decir que se ofrece con un nivel óptimo de calidad científico-técnica?, ¿qué características me indican este nivel?, ¿qué importancia tiene la aceptabilidad y satisfacción de los usuarios?, ¿cómo puedo saber si están presentes?

Una respuesta exhaustiva a todas estas preguntas nos producirá un listado de características o requisitos que nos definirían una atención de calidad para el problema que intentamos definir. Se trata como propone Ishikawa, de la descripción de las características de la calidad (aquellas que satisfacen las necesidades de los usuarios) de los servicios a evaluar, y que serían aquellas que conducen a la satisfacción de las necesidades y expectativas a las que se dirige el servicio a definir y evaluar. Este razonamiento está pensado, sobre todo, para los productos de la industria, pero la idea que subyace, también es relevante para los servicios de salud: definamos bien nuestro producto, incluidas las características que creamos que ha de tener para que sea de calidad y después intentemos medirlas de forma directa o indirecta. Tanto lo que es calidad como sus dimensiones medibles se convierten, de esta manera, en conceptos flexibles, adaptables a cada producto o servicio, y dependientes de la creatividad, responsabilidad y sentido de la función que debe desempeñar cada profesional, institución o sistema.

A partir de la UT6, iniciaremos este proceso de definición específica de la calidad, aplicado a un problema o aspecto concreto de los servicios que ofrecemos, de forma que podamos medirla, evaluarla y mejorarla.

La traducción de las dimensiones de la calidad en requisitos de calidad para un producto o servicio concreto es lo que nos va a permitir medirla.

La satisfacción del profesional forma parte de la calidad del sistema u organización, no del servicio que se ofrece, aunque pueda influir en él.

7. ¿Y LA SATISFACCIÓN DEL PROFESIONAL? CONCEPTO DE "CLIENTE INTERNO"

La satisfacción del profesional no ha sido mencionada en las definiciones de calidad revisadas, ni es frecuente encontrarla de forma explícita en ninguna definición. ¿Por qué? La razón principal es que, si recordamos el esquema básico (proveedor → servicio → receptor), (Figura 1.2), el producto o servicio que hemos intentado definir hasta ahora es aquél que recibe el usuario o cliente externo del proveedor, entendiendo por proveedor el individuo, la organización o el sistema que ofrece el servicio. Adicionalmente, hemos aclarado que íbamos a hablar de calidad del producto o servicio, y no de otras "calidades" como puede ser la de la organización, objetivos, personal, etc. La satisfacción del profesional ha de ser contemplada *dentro* de la "calidad" de la organización, pero no es un elemento de la calidad del servicio en sí, aunque puede *influir* en él, al igual que otros muchos factores del sistema o de la organización. Vamos a detenernos un poco más en detallar estos dos aspectos: (a) cómo influye la satisfacción del profesional en la calidad del servicio y (b) su relevancia como un elemento interno del sistema.

a) La satisfacción del profesional y la calidad del servicio prestado

Aunque la justificación fundamental implícita y subyacente para no incluir la satisfacción del profesional en la definición del servicio que se ofrece al usuario externo, es el que sea un elemento de la organización y no del servicio ofrecido, hay autores que dan justificaciones adicionales en relación con la influencia concreta que la satisfacción del profesional tiene sobre la calidad, concluyendo tal como hemos apuntado, que la satisfacción del profesional puede considerarse como un criterio de la calidad de la organización o institución, a modo de rasgo estructural que favorece el buen funcionamiento y la calidad del servicio.

La evidencia empírica sostiene de un forma imperfecta que, al igual que en la industria, la satisfacción del profesional puede influir en un mejor trabajo, así como el que una percepción del trabajo bien hecho produce satisfacción en los profesionales. Sin embargo, la relación satisfacción del profesional/calidad no es siempre la misma, ni lineal, ni consistentemente importante. Lo que parece más claro es una relación de la satisfacción con la productividad y la estabilidad o rotación en los puestos de trabajo.

b) La satisfacción del profesional como cliente interno

Los profesionales y los servicios que integran un determinado centro asistencial son, a su vez, proveedores y "clientes" unos de otros, aunque sus actividades deban estar interconectadas con el fin exclusivo de servir a los clientes externos. Este concepto de "clientes internos" que reciben servicios para la provisión de un determinado producto final, lleva a incluir la satisfacción de los profesionales como parte de la definición de calidad de los servicios internos o, al menos, como uno de sus determinantes de importancia y, desde luego, como un factor relevante para el funcionamiento del centro. Un "cliente interno" sería aquel que recibe servicios de otros elementos del sistema (radiografías, análisis, servicios administrativos, etc.) y su satisfacción con la calidad de ese *servicio interno* es tan importante como la definición de la calidad de los servicios externos.

Aparte de eso, aunque pueda argumentarse que la satisfacción de los profe-

La satisfacción del profesional como cliente interno que recibe servicios de otros elementos del sistema, debe formar parte de la definición de calidad de estos servicios internos.

sionales influya en la calidad de su trabajo para el cliente externo y viceversa, es muy arriesgado decir que al medir la satisfacción del profesional estamos midiendo calidad de la atención al cliente externo. Lo que sí estaremos midiendo en cambio es un aspecto de la calidad de la organización y de sus servicios internos.

8. EL CONCEPTO DE CALIDAD TOTAL

El adjetivo "total" cualificando al concepto "calidad" es algo de uso relativamente prevalente hoy en día en los medios sanitarios. Sin embargo, conviene aclarar que el uso de estas dos palabras juntas como concepto en castellano es fuente probablemente de una confusión, aunque, afortunadamente, no conduce, necesariamente a una confusión también en la práctica, como vamos a ver a continuación.

En la literatura anglosajona, en donde se ha originado el término, el adjetivo "total" unido a "calidad" se empezó a utilizar en los años cincuenta y se sigue utilizando para adjetivar no siempre a la palabra "calidad" sino también a una tercera palabra, que es en realidad la que se conceptualiza de una forma particular al añadirle los dos términos "total" y "calidad". Así encontramos "Total Quality Control", "Total Quality System" y "Total Quality Management", términos que pueden ser traducidos como "Control Total de la Calidad", "Sistema Total de Calidad" y "Gestión Total de la Calidad" respectivamente. ¿Qué se quiere significar al añadir la palabra "total"? Si vamos a las fuentes originales de los inventores del concepto queda claro que lo que introduce es un importantísimo matiz de ámbito de actuación al control de calidad, al sistema de calidad y a la gestión de calidad: que el ámbito en el que actúen sea "total" quiere decir que ha de incluir *todos los procesos y actividades* que puedan influir en la calidad del producto o servicio de que se trate; es decir se presta atención a la calidad de las materias primas, del personal, de la organización, servicio post venta, etc., y no solo los procesos directamente relacionados con la manufactura básica del producto o la prestación del servicio de que se trate. Como ejemplo más destacado de esta conceptualización cabe citar la explicación que da Feigenbaum, el primero que introdujo el concepto de "total" en 1951 en un libro clásico que tuvo recientemente una edición 40 aniversario. Feigenbaum, acuñó los términos "Total Quality Control" y "Total Quality System" para referirse al control de la calidad en *todos los aspectos del ciclo industrial* (marketing, diseño/ingeniería, compras, manufactura, inspección del producto, distribución, instalación y servicio post-venta) y a un sistema establecido *en toda la empresa* para guiar las acciones coordinadas de todos sus elementos en la consecución de la satisfacción del cliente de una forma eficiente. El mismo autor define "Total Quality Management" como aquella que abarca al ciclo completo del producto o servicio, desde su concepción hasta la atención al cliente que adquirió el producto o servicio, pasando por su producción.

No ser "total" significa un ámbito de actuación o atención explícita parcial (por ejemplo sólo a los procesos de producción) tanto si hablamos de control, como de gestión o de sistema.

Si trasladamos las explicaciones de Feigenbaum al concepto de calidad que vimos al principio de esta UT, cuando empleamos el adjetivo "total" debemos estar incluyendo una atención no sólo a la calidad del producto o servicio en sí, sino también a la calidad de la organización en un sentido amplio, incluyendo

El adjetivo "total" puede cualificar a la calidad o a la gestión de la calidad

Calidad Total implica una referencia a la calidad del producto o servicio y también a la calidad de la organización y sus relaciones con otras organizaciones

Tanto Gestión de la Calidad Total como Gestión Total de la Calidad deben significar que nuestra atención y actividades para mejorar se dirigen a todos los ámbitos y niveles, internos y externos, y a todos los servicios o productos, para clientes y de proveedores internos y externos, de nuestra institución.

relaciones con proveedores y otras organizaciones, así como las relaciones, clima organizacional y satisfacción interna del personal de la propia empresa.

¿Qué ocurre si utilizamos la palabra "total" como adjetivo de "calidad" en vez de cómo adjetivo de la función (control, gestión, etc.), como de hecho se hace en muchas ocasiones?: pues que al igual que si la aplicamos a "gestión", "control", o "sistema" nos debe quedar claro que lo que estamos indicando es calidad a todos los niveles y de todos los componentes, procesos y relaciones de nuestro "negocio", institución, centro o sistema de salud en el que trabajamos. Lo que debemos evitar es utilizar "total" para significar diferencias de métodos de medición, enfoque, o matices prioritarios, en vez de diferencias en el ámbito de actuación. Así por ejemplo a veces encontramos quien dice que las actividades son de "calidad total" porque se emplean métodos tomados de la industria, o porque se da mucha importancia a la satisfacción del cliente, o porque se tiene en cuenta el componente de costes, etc.; todos estas circunstancias pueden darse perfectamente dentro de un enfoque "parcial" de la calidad o de su gestión.

Lo lógico es que cuando hablemos tanto de Gestión Total de la Calidad como de Gestión de la Calidad Total, lo hagamos para significar que nuestra atención y actividades para mejorar se van a dirigir a todos los ámbitos y niveles, internos y externos, y a todos los servicios o productos, para clientes y de proveedores internos y externos, de nuestra institución.

BIBLIOGRAFÍA

- Saturno PJ, Quintana O, Varo J. ¿Qué es la calidad? En: Tratado de Calidad Asistencial en Atención Primaria. Madrid: Dupont Pharma; 1997. p. 17-45.
- Juran JM. La función de la calidad. En: Juran JM, Gryma FM. Manual de Control de la Calidad. Volumen I. Madrid: McGraw-Hill; 1993. p. 2.1-2.13.
- Ishikawa K. La esencia del control de calidad. En: Ishikawa K. ¿Qué es el Control Total de Calidad? Madrid: Parramón; 1994. p. 61-90.
- Deming WE. Calidad, productividad y competitividad. La salida de la crisis. Madrid: Díaz de Santos; 1989.
- Feigenbaum AJ. Total Quality Control. 3ª ed. (edición 40 aniversario). New York: McGraw-Hill; 1991.

**EL CONTEXTO DE LOS SERVICIOS DE
SALUD COMO MARCO DE REFERENCIA
PARA LA MEJORA DE LA CALIDAD**

EMCA
Gestión de la Calidad Asistencial

CONTENIDO GENERAL

En esta UT se define el contexto de la organización en la que realizar las actividades de mejora de la calidad. Los conceptos centrales son la misión, visión y las líneas estratégicas de la institución en la que se trabaja. La descripción y el conocimiento compartido y asumido del contexto de las organizaciones se ha revelado de una relevancia considerable en las estrategias de implantación de Gestión de la Calidad.

ÍNDICE DE CONTENIDOS

1. El contexto de los servicios de salud como marco de referencia.
2. El concepto de misión. ¿Qué somos?
3. El concepto de visión. ¿Hacia dónde queremos ir?
4. Establecimiento de líneas estratégicas.
5. Misión, visión y calidad. ¿Cómo se relacionan?

OBJETIVOS ESPECÍFICOS

1. Distinguir los conceptos de misión y visión de una institución.
2. Definir la misión de la institución en la que trabaja.
3. Definir la visión de la institución en la que trabaja.
4. Formular líneas estratégicas a corto plazo para cumplir con la misión asumida.
5. Establecer líneas estratégicas a medio y largo plazo para alcanzar la visión asumida.

1. EL CONTEXTO DE LOS SERVICIOS DE SALUD COMO MARCO DE REFERENCIA

Para la puesta en marcha de un programa de gestión de la calidad, suponemos ya asumido plenamente un interés por la calidad, aceptado conscientemente un concepto de lo que entendemos por calidad y vislumbrado cómo es posible medirla. Para todo ello hemos partido de un esquema básico (proveedor → servicio → usuario) en el que hemos desarrollado la idea de calidad del servicio en función de las necesidades y expectativas del usuario para el que está pensado. En todos los argumentos hemos considerado implícitamente que el proveedor está definido y sabe su función, y que los servicios que ofrece están explícitamente pensados para unos usuarios cuyas necesidades y expectativas se conocen. Si esto es así, podríamos pasar directamente a ver por dónde empezamos a medir para mejorar. Sin embargo, al igual que se hace necesario elegir entre diversos conceptos de lo que se entiende por calidad, no todas las organizaciones que proveen servicios de salud son iguales, ni los servicios que ofrecen y quieren ofrecer son los mismos. Se hace necesario, describir también, explicitar conscientemente, el *contexto* en que nos movemos como proveedores, para completar el marco de referencia de nuestras actividades de evaluación y mejora.

El *contexto* de una organización puede ser descrito en base a los componentes que conforman el medio en el que se realizan las actividades de dicha organización. Al igual que al hablar de calidad, podríamos considerar un concepto amplio de contexto, refiriéndonos al medio externo al centro asistencial concreto en que nos movemos. Podríamos hablar de los componentes de la organización (o "empresa") provisorora de servicios de salud a nivel de sistema, área o distrito. Sin embargo vamos a enfocar la descripción a un nivel más desagregado y cercano a la provisión directa de servicios: el nivel de centro asistencial.

Una posible descripción completa de un determinado contexto organizacional incluiría al menos cinco componentes: misión, visión, valores, principios, estrategias y objetivos. Simplificando, dado que muchos de estos conceptos están relacionados, vamos a detallar lo que sería la *misión*, la *visión* y las *líneas estratégicas*, como componentes básicos y necesarios para contextualizar adecuadamente las actividades del centro asistencial en que trabajamos, y con ellos las iniciativas de gestión de la calidad.

Lógicamente, por muy autónomos que sean, los centros asistenciales están inmersos en una organización más grande (comarca, área, distrito, sistema), cuyos componentes contextuales debían de ser consistentes con los del centro; de hecho estos niveles superiores deberían de *guiar* a los centros, al igual que las filiales deben seguir directrices de la casa central. Sin embargo, nuestro sistema de salud dista mucho de funcionar con esquemas organizativos empresariales, y es muy posible que la situación más común en los centros es que se desconozca, por falta de explicitación o de intentos por hacerlas compartir, cuáles son la misión, visión, valores, principios, estrategias y objetivos del sistema de salud en sus diversos niveles. Este hecho no debe ser obstáculo, para que intentemos aclarar estos conceptos a nuestro nivel. Tenemos mucho que ganar con ello, tal como veremos a continuación

El marco de referencia para la gestión de la calidad se completa describiendo el contexto organizacional de quien provee los servicios, tres de cuyas características son la misión, visión y las líneas estratégicas asumidas

2. EL CONCEPTO DE MISIÓN. ¿QUÉ SOMOS?

La misión de una determinada organización justifica su existencia (qué somos) y sus fines (para qué servimos)

La misión asumida por todo el centro es fuente de motivación y consolida valores más allá de los grupos de interés

La misión de una organización describe la razón o las razones por las cuales existe la organización. En su nivel de conceptualización más simple, la misión de una determinada organización es la función o tarea fundamental que la sociedad le asigna y espera de ella. Así por ejemplo, la misión de los tribunales de justicia sería la interpretación de las leyes y su aplicación, la misión de la policía sería la protección de la seguridad y bienestar de los ciudadanos, la misión de la universidad sería la enseñanza de las distintas profesiones universitarias y la investigación, etc, etc. Incluso a este nivel amplio y simple, la misión de un equipo de atención primaria puede distinguirse de la de otro centro asistencial. Por ejemplo, un centro de salud puede tener como misión mantener sana a la población a su cargo; mientras que un hospital de agudos puede tener como misión exclusiva la curación de enfermedades agudas. Sin embargo, la descripción de la misión de una determinada organización es conveniente que no sea tan simple, y que incluya un compendio resumido de los objetivos generales de la misma; se trata de reflejar de este modo las cualidades y características que justifican la existencia de la organización (quiénes somos) y sus fines (para qué existimos como organización). Se convierte así en un marco de referencia dentro del cual cada individuo dentro de la organización puede formular su misión particular, es decir, su contribución específica a la misión que se asume. Para comprender mejor este concepto vamos a ver a continuación tres ejemplos reales de misión elaborados por entes organizacionales diferentes: un centro de salud, un hospital universitario y una unidad de calidad de un complejo sanitario:

El centro de salud de Azogues, en Ecuador, en el marco de sus primeros pasos para evaluar y mejorar la calidad definió su misión como: "Brindar atención integral y de calidad a todos los usuarios, con criterio de riesgo (preferencia a niños, embarazadas, urbanomarginales y rurales), con metas de 100% de cobertura en programas preventivos, y propiciando la participación comunitaria y la autogestión".

El hospital universitario de Alberta (Canadá) definió su misión como: "Un compromiso con la provisión de educación y de cuidados de salud ejemplares, en una atmósfera de comprensión y cuestionamiento académico, respetando la dignidad y los derechos de los pacientes y sus familias...", detallando a continuación seis líneas principales de actividades y objetivos generales relacionados con esta misión.

El comité de calidad del complejo de centros sanitarios de la Universidad de Michigan definió como su misión: "Desarrollar políticas y procedimientos, diseñar estructuras de calidad, asignar recursos, e implementar estrategias para el avance del movimiento por la calidad en nuestra organización".

¿Qué utilidad tiene especificar la misión de nuestra organización e incluso de nosotros mismos como individuos? El conocimiento de la misión tiene utilidad tanto hacia fuera como hacia dentro de la organización. Hacia fuera, para comunicarle a nuestros usuarios y a otros centros asistenciales o niveles organizativos dentro de nuestro sistema lo que pueden esperar de nosotros; hacia dentro, como fuente de motivación y dirección. Al suministrar una unidad de propósitos que trasciende las necesidades individuales, locales y transitorias, fomenta un sentimiento de expectativas compartidas entre todos, consolidando valores más allá de los grupos de interés, y, en definitiva, afirmando el compromiso de la

organización con acciones que responden a las demandas esenciales de sus clientes o usuarios internos y externos.

En las empresas con misión explícita y conocida por todos los que trabajan en ella, lo normal es que haya sido elaborada por la alta dirección y comunicada de arriba a abajo; sin embargo, en los sistemas de salud, el margen de autonomía de los centros asistenciales puede ser muy amplio, dentro de unas líneas generales que de existir van poco más allá de una conceptualización del tipo "atender a la población y que esté satisfecha"; si existiese una misión explícita por parte de los niveles gerenciales, esta debe ser el punto de partida para la que elaboremos la de nuestro centro.

Conviene aclarar, no obstante, que la misión de una organización no debe formalizarse en términos de obtener utilidades o beneficios para la organización. Los manuales de administración enseñan que la obtención de "superávit" o "beneficio" no es la misión de las empresas, éstos pueden ser objetivos a lograr al ir en direcciones claramente definidas, alcanzando metas en la realización de una misión hacia la sociedad o el grupo de usuarios por el que la empresa existe. Llevado a nuestro contexto, ni el "ahorro" ni el "aumento de las recaudaciones" son elementos que deben formar parte de la misión de los centros asistenciales, que deben formularse en función de los servicios que prestan.

3. EL CONCEPTO DE VISIÓN. ¿HACIA DÓNDE QUEREMOS IR?

Si el tener clara la *misión* ayuda a comprender nuestra función en el contexto de nuestra organización y su utilidad para la sociedad, hay otro elemento del contexto organizacional que, si logramos establecerlo y compartirlo, nos motivará y guiará hacia el futuro: la *visión* que tenemos para nuestra organización.

La *visión* puede ser definida como una descripción de cómo queremos que sea nuestra organización en el futuro; una especie de sueño alcanzable a largo plazo; una situación a la vez posible y altamente deseable. La *visión* proporciona a la organización un sentido de dirección único y consistente hacia el futuro, que se ha de conseguir sobre la base de la *misión* asumida.

La *visión*, más aún que la *misión*, es elaborada, en las empresas que la tienen, por el nivel gestor, con una mayor o menor participación de los que trabajan en ellas; aunque no tiene ningún valor si no se consigue que sea compartida por todos. En los centros asistenciales puede ser útil tener una *visión* de futuro en la medida en que nos sintamos "dueños" de su funcionamiento y sea realmente posible un cierto grado de autonomía.

Una *visión* compartida por todos los miembros de la organización es un elemento motivador de primer orden, que fomenta el compromiso y da sentido a iniciativas estratégicas que superan el corto plazo. Tener claro el panorama de hacia dónde va la organización ayuda a alcanzarlo con éxito y confiere seguridad. La Tabla 2.1 contiene algunos ejemplos de *visión* de instituciones sanitarias.

La *visión* es una descripción de como queremos que sea nuestra organización en el futuro, algo posible, deseable y basado en la *misión* asumida

TABLA 2.1. Ejemplos de visión de algunas instituciones sanitarias

- **Hospital Universitario de la Universidad de Michigan:**
"A través de la mejora continua de la calidad nos convertiremos en el proveedor de primera elección para la población que servimos, ocuparemos un lugar prominente en la docencia y el avance de las ciencias médicas, y seremos el lugar de trabajo de primera elección".
- **Una HMO (Health maintenance Organization) de Chicago:**
"Nuestra organización será reconocida como el primer sistema de salud en nuestro entorno, a través de un esfuerzo continuo por superar los estándares de calidad profesionales, las actuaciones de otros colegas, satisfacer todas las expectativas de la población que servimos, y la promoción de un ambiente de trabajo excepcional".
- **Un centro asistencial religioso de Carolina del Sur:**
"Nos guiaremos por una filosofía de servicio cristiano y un compromiso por la excelencia en todo lo que hagamos y una visión holística de la salud para servir a los residentes de nuestra región. Respetaremos la individualidad de cada empleado y fomentaremos un lugar de trabajo en el que la creatividad y la productividad de los empleados será reconocida, valorada y recompensada. Nuestros éxitos se construirán en base a servicio, personas, ideas, respeto, innovación y trabajo en equipo".
- **Área 10 de Madrid**
"Pretendemos llegar a ser el Área sanitaria de referencia dentro del Sistema Nacional de Salud, por la calidad de los servicios sanitarios que presta a la población, el nivel de cualificación de sus profesionales, la eficiencia en la obtención de resultados y el grado de satisfacción".

Es absolutamente necesario que la visión sea positiva y contribuya a la visión personal, particular, de cada elemento de la organización. A pesar de no ser muy frecuente aún el que las empresas y las organizaciones sanitarias establezcan claramente una visión de futuro, ya empieza a haber ejemplos de visiones inefectivas, que no pasan de ser algo que se hizo en un determinado momento para un determinado plan de calidad y que jamás se ha intentado llevar a la práctica. Basados en estas experiencias negativas, o, mejor dicho, improductivas, hay autores que señalan lo que parecen ser los problemas más comunes en estos casos:

- La visión no es "inspiradora", y no se incorpora fácilmente a los valores personales, individuales de los que trabajan en la organización.
- La visión no es "auténtica". O no es realizable, o no responde a los valores y tradición de la organización.
- No se logró que la visión fuese compartida, por defecto de elaboración (de arriba a abajo exclusivamente), o de comunicación.
- La visión es demasiado abstracta, de interpretación ambigua y confusa, y por tanto no puede guiar las actividades de la organización.
- No se intentó traducirla en objetivos y líneas de acción en la organización.

Este último problema a evitar nos da pie para considerar el tercer componente del contexto organizacional que puede ser de interés para encuadrar adecuada y provechosamente las actividades de mejora de la calidad.

La visión es un elemento motivador de primer orden, pero también puede ser inefectiva si no se comparte ni se traduce en objetivos y líneas estratégicas

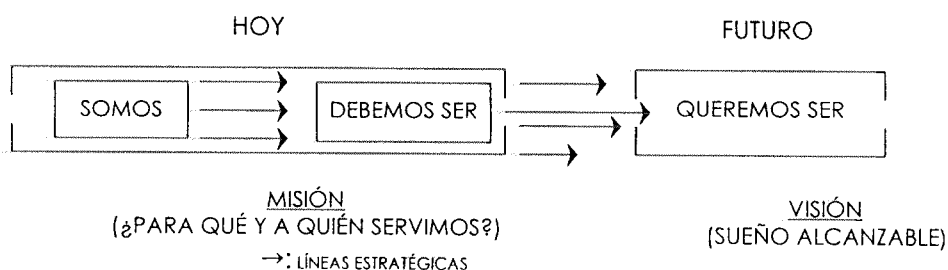
4. ESTABLECIMIENTO DE LÍNEAS ESTRATÉGICAS

Las *líneas estratégicas* establecen la dirección y priorización de las actividades de la organización durante un cierto periodo de tiempo; constituyen en conjunto un plan de acción para hacer realidad la misión y la visión de la organización. En buena lógica, han de surgir tras realizar un análisis de la situación y definir la problemática existente para cumplir la misión y conseguir la visión.

¿Por qué es útil establecer líneas estratégicas? En primer lugar, ayudan a combinar en la práctica el corto, medio y largo plazo; adicionalmente, sirven para optimizar iniciativas, recursos y actividades, y proveen un marco de prioridades para las actividades de la gestión de la calidad.

En la Figura 2.1 están representadas los tipos de líneas estratégicas que podemos trazar: unas a más corto plazo, para el cumplimiento de la misión, y otras a más largo plazo para progresar hacia la visión. Por ejemplo, si hemos previsto como misión el mantener sana a la población a nuestro cargo, una línea estratégica a corto plazo podría ser la mejora de las actividades de prevención y promoción de la salud; si nuestra visión incluye gestionar y coordinar la atención sanitaria de la población a nuestro cargo, una línea estratégica a medio y largo plazo sería conseguir una coordinación en el sentido deseado cada vez mayor, hasta ser total, con otros niveles del sistema de salud y servicios sociales.

FIGURA 2.1. Misión, Visión y Líneas Estratégicas



Las líneas estratégicas generales conforman la política de calidad de la organización. Por ejemplo, en el programa de gestión de calidad del Área 10 se han propuesto las siguientes líneas estratégicas: "El Área 10 de Atención Primaria del Insalud de Madrid se compromete a que la calidad constituya el eje central de todas sus actuaciones y a todos los niveles de la organización. Ello supone situar a los ciudadanos a los que prestan servicio y a la satisfacción de sus necesidades en el primer lugar de su orden de prioridades. Esto sólo será posible mediante la implicación y el compromiso de todos los profesionales del Área con la mejora continua de la calidad. La participación, formación y satisfacción de estos profesionales constituyen los elementos clave que permitirán el desarrollo de la Gestión de la Calidad Total."

Las líneas estratégicas constituyen un plan de acción para cumplir la misión y conseguir la visión

La misión, visión y líneas estratégicas de los centros completos, junto al concepto de calidad y el conocimiento de las expectativas y necesidades a los que se agrupan los servicios que se ofrecen, el marco de referencia para analizar y mejorar la calidad

5. MISIÓN, VISIÓN Y CALIDAD. ¿CÓMO SE RELACIONAN?

Es muy posible que todo lo relativo a la misión, visión y líneas estratégicas sea, no ya novedoso, sino hasta incluso chocante. Habrá incluso quien puede que no vea la necesidad de saber de estas cosas para medir y mejorar la calidad de los servicios que ofrecen los centros de salud. Sin embargo, la experiencia que va habiendo indica la conveniencia de tener claro el contexto interno del proveedor (misión, visión, líneas estratégicas) que se prolonga de forma fácil y consistente con los servicios que ofrece, que a su vez han de estar en sintonía con las necesidades y expectativas de la población a atender, que son a su vez las bases de la misión, visión y líneas estratégicas de los centros asistenciales, cerrándose así el círculo del esquema básico para analizar y mejorar la calidad.

Hablar de necesidades de salud y expectativas de los usuarios, y de los servicios que ofrecen los centros asistenciales, puede resultar más o menos complicado, pero es un terreno fácil y familiar. No es tan frecuente, en cambio, el lenguaje y significados unidos a la misión, visión y líneas estratégicas. Sin embargo, encontrarse con ellas y poder asumirlas con convicción es como encontrar un oasis en un desierto en el que podemos estar perdidos, un oasis donde, además tenemos la posibilidad de encontrar los caminos que nos lleven a donde queremos ir, haciendo posible salir del desierto para no volver si así lo queremos.

BIBLIOGRAFIA

- Shortell SM, Kaluzny AD. Health care management. Organizational design and behavior. 3^{er} ed. Albany: Delmar Pub; 1994.
- Andrews HA, Look LM, Davidson JM, Schurman DP, Taylos EW, Wersel RH. Organizational transformation in health care. San Francisco: Josey Bass; 1994.
- Melum MM, Simioris MK. Total Quality Management. The health care pioneers. Chicago: AHA; 1992.
- Robins SP. Organizational behavior. Concepts, controversies and applications. Practice Hall. Unglewood Cliffs, 1989. 4^{er} ed.

**PROGRAMAS DE GESTIÓN
DE LA CALIDAD: CONCEPTO,
ACTIVIDADES Y UTILIDAD**

EMCA
Gestión de la Calidad Asistencial

CONTENIDO GENERAL

Esta UT, partiendo de la necesidad de tener programas que tengan como fin explícito la gestión de la calidad, aclara, clasifica y analiza desde un punto de vista general y operativo (práctico) los diversos tipos de actividades que se pueden realizar en los programas de gestión, mejora o garantía de calidad. Las actividades se agrupan en función de los diversos puntos de partida y los objetivos directos que persiguen. En este esquema se encuadran las actividades de los Ciclos de Mejora, sin perder de vista su relación con otras actividades dentro de la globalidad de los programas de gestión de la calidad.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. ¿Qué es un programa de gestión de la calidad?
3. Elementos de un programa de gestión de la calidad.
4. Actividades de un programa de gestión de la calidad.
5. Otros modelos para representar las actividades de un programa de gestión de la calidad: Ciclo de garantía de calidad, Trilogía de Juran, Ciclo de Deming.
6. ¿Para que sirve un programa de gestión de la calidad?
7. Razones para no tener un programa de gestión de la calidad.

OBJETIVOS ESPECÍFICOS

1. Asumir un concepto operativo de lo que significa tener un programa de gestión de la calidad.
2. Enumerar las razones por las que son necesarios los programas de gestión de la calidad.
3. Distinguir los tres grandes grupos de actividades que componen los programas de gestión de la calidad.
4. Describir los puntos de partida (inicio) y los objetivos directos de cada grupo de actividades.
5. Establecer los puntos de interrelación entre los tres grupos de actividades.
6. Contrastar las diversas terminologías que describen estas actividades.

1. INTRODUCCIÓN

Con la UT 1 hemos asumido una definición de lo que entendemos por calidad, y posteriormente la habremos incorporado a la misión y visión particular que tenemos de nuestro entorno de trabajo; en esta UT nos vamos a ocupar de en qué puede consistir la materialización del compromiso con la calidad, traducido a una estructura organizativa y unas actividades que lo hagan real y operativo.

Una revisión rápida a la literatura sobre calidad nos informaría sin contradicciones sobre las ventajas de llevar a la práctica con iniciativas organizativas y actividades concretas el compromiso con la calidad; sin embargo, también veríamos que los acercamientos teóricos y prácticos a las actividades de evaluación y mejora de la calidad son frecuentemente parciales en cuanto a temas, protagonistas y/o niveles organizativos, y sin una conexión explícita con otras actividades y estructuras con las que, sin embargo, comparten el fin último de lograr un nivel de calidad cada vez mayor. Con esta realidad de fondo, esta UT ofrece una visión práctica e integradora de las diversas estructuras, protagonistas, objetivos y actividades tendentes al progreso en la calidad del servicio de salud. Vamos a llamar programa de gestión de la calidad a este conjunto de elementos estructurales y de actividades que tienen como fin específico la mejora continua de la calidad; es este un concepto equivalente en general al de sistema de gestión de la calidad utilizado en la industria.

El objetivo final de la mejora continua supone, cuando el programa funciona, avanzar en la efectividad, eficiencia y satisfacción de todos los protagonistas de la asistencia; un progreso consciente y cuantificable hacia la excelencia.

2. ¿QUÉ ES UN PROGRAMA DE GESTIÓN DE LA CALIDAD?

Vamos a llamar programa de gestión de la calidad (GC) al conjunto de elementos estructurales y de actividades que tienen como fin específico la mejora continua de la calidad. Este concepto puede encontrarse con diferentes nombres en la literatura sobre calidad, tanto en la industria como en los servicios de salud.

La confusión terminológica se acompaña con frecuencia de confusión conceptual. Hay tres componentes en el concepto global de programa de GC que no siempre son diferenciados, sino que podemos encontrarlos confusamente entrelazados o sustituidos unos por otros, y son: (i) la filosofía de mejora continua que guía las actividades de GC; (ii) la estructura y organización para la práctica de esta filosofía; y (iii) las actividades concretas conducentes a controlar y mejorar la calidad.

El término globalizador más frecuentemente empleado en la literatura de la industria es el de sistema de calidad. Sin embargo, la terminología empleada en el sector sanitario es diferente y particular. El término globalizador más frecuentemente empleado es el de "programa" y también, en nuestro caso, el de "plan". En planificación sanitaria, la acepción de "programa" probablemente con más predicamento es la propuesta por Pineault y Daveluy: "un conjunto organizado, coherente e integrado de actividades y de servicios, realizados simultánea o sucesivamente, con los recursos necesarios, y con la finalidad de alcanzar los objetivos determinados, en relación con problemas de salud precisos y ello para una población definida"; si sustituimos "problema de salud preciso" por "mejora

En esta UT se propone una visión práctica e integradora de lo que llamamos Programa de Gestión de la Calidad.

El concepto de programa de gestión de la calidad incluye la existencia de una filosofía o política explícita para mejorar la calidad, unos elementos estructurales y organizativos para llevarla a cabo, y unas actividades para hacerla efectiva

Ante la confusión terminológica existente optamos por el término "programa" por ser el de más tradición en el sector salud.

continua de la calidad" resulta una definición perfectamente compatible con la que proponemos. Estos mismos autores, aconsejan distinguir entre el plan del programa (plan de acción que contiene los elementos del programa) y la puesta en marcha operativa del mismo, como dos componentes del programa, a la vez que advierten que la distinción entre planificación y programación puede resultar artificial, otorgándole si acaso al término "programación" un sentido más específico y operacional.

¿Qué terminología adoptar? Aquí se ha optado por respetar la tradición sanitaria, llamándole programa para enfatizar su carácter finalista y operativo, pero incluyendo unos contenidos que incorporan ideas de la industria. Lo que importa en definitiva, se le llame como se le llame, es que podamos identificar una serie de elementos estructurales y de actividades que se justifican con la existencia previa de una filosofía o política explícita de mejorar continuamente la calidad.

3. ELEMENTOS DE UN PROGRAMA DE GESTIÓN DE LA CALIDAD

La lista de componentes o elementos que da carácter a un programa de GC puede ser muy larga, según el nivel de detalle a que queramos descender. Por otra parte, la forma concreta en que el compromiso con la calidad se traduce en una estructura organizativa y unas actividades concretas puede ser muy variada. Igual que el concepto de lo que se entiende por calidad, el programa que la gestione para procurar su mejora continua parece que ha de ser un "traje a medida" del centro, nivel asistencial o sistema de salud que acepte el reto de ponerlo en marcha. Todos los programas van a tener en común, desde luego, el objetivo general de la mejora de la calidad del servicio que se ofrece; sin embargo, la estructura concreta y las actividades priorizadas y rutinarias en que se basan los programas pueden ser muy diferentes de unos centros a otros. En sus rasgos organizativos más esenciales o elementales, a un programa de GC lo pueden caracterizar los siguientes tres componentes:

- **Una estructura identificable**, incluyendo una comisión o grupo visible que tiene como objetivo preocuparse por y promover las actividades de gestión de la calidad. Aunque esta comisión no ha de ser necesariamente quien se encargue de realizar las actividades de evaluación y mejora de la calidad, sí ha de promoverlas, apoyarlas y coordinarlas.
- **Normas de funcionamiento**, incluyendo reparto claro de responsabilidades, circuitos de información de los datos sobre calidad, y mecanismos de elaboración de planes de mejora de la calidad.
- **Las actividades que se realizan para mejorar la calidad**. Actividades identificables y clasificables no según su metodología sino en razón del fin u objetivo parcial al que sirven, dentro del objetivo común de la mejora de la calidad. Merece la pena detenernos en este punto y repasar una visión panorámica e integradora de las principales actividades que pueden realizarse en un programa de GC.

Los programas de GC son un "traje a medida" para cada centro, pero se pueden caracterizar describiendo sus componentes básicos: estructura, normas de funcionamiento y actividades que se realizan para la mejora continua.

4. ACTIVIDADES DE UN PROGRAMA DE GESTIÓN DE LA CALIDAD

Dentro del objetivo general de la mejora continua de la calidad, las diversas actividades a realizar son agrupables según sus tres principales puntos de partida en torno a los que puede funcionar el programa: (i) la identificación de oportunidades de mejora, (ii) la elaboración de indicadores y su monitorización, y (iii) el diseño de nuevos servicios o cambios substanciales en los existentes. Tres puntos de partida que se corresponden con tres áreas de actividades diferenciables dentro del programa: ciclos de mejora, monitorización y diseño de la calidad, que tienen a su vez objetivos inmediatos diferentes, como se refleja en la Tabla 3.1.

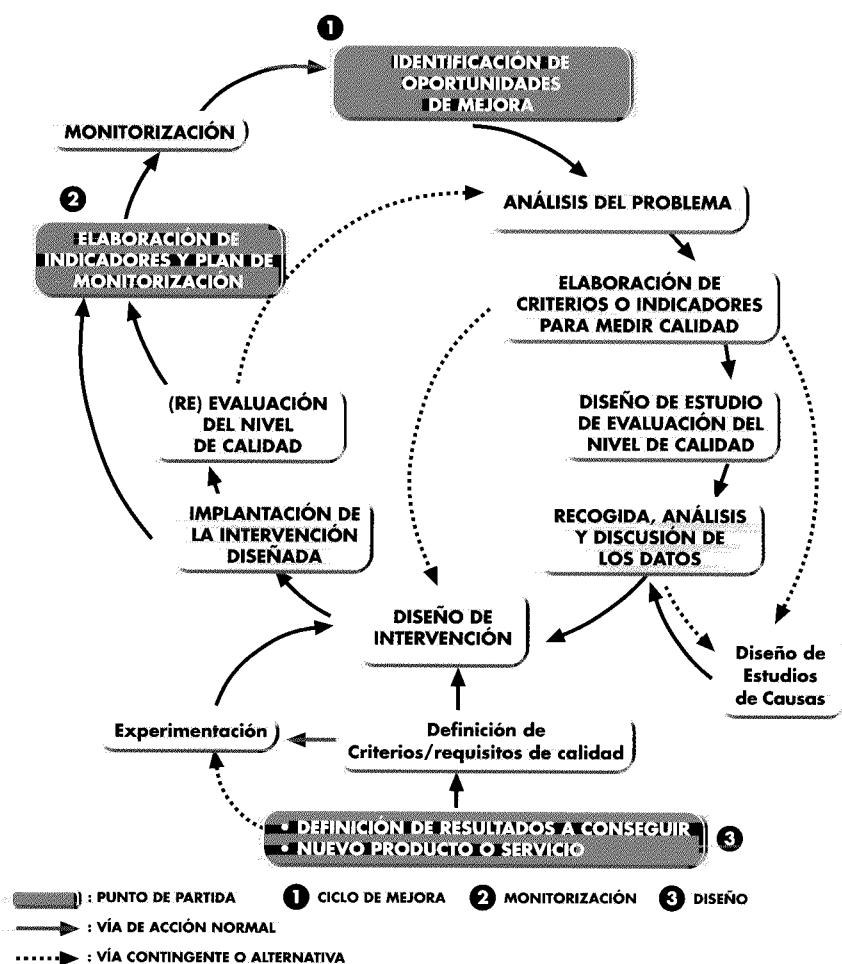
Las actividades de un programa de GC pueden agruparse según su punto de partida y objetivos inmediatos en ciclos de mejora de la calidad, diseño de la calidad y monitorización de la calidad

TABLA 3.1. Actividades de los Programas de Gestión de la Calidad

| ACTIVIDADES | PUNTO DE PARTIDA | OBJETIVO INMEDIATO |
|------------------|--|---|
| Ciclos de mejora | <ul style="list-style-type: none"> Identificación de un problema de calidad u oportunidad de mejora en algún aspecto de los servicios que se ofrecen. | <ul style="list-style-type: none"> Solucionar el problema. Aprovechar la oportunidad de mejora descubierta. |
| Monitorización | <ul style="list-style-type: none"> Identificación de aspectos relevantes de los servicios que se ofrecen y construcción de indicadores sobre su calidad. Selección de indicadores sobre problemas que hemos sometido a ciclos de mejora. | <ul style="list-style-type: none"> Identificación de problemas de calidad u oportunidades de mejora. |
| Diseño | <ul style="list-style-type: none"> Programación de un nuevo servicio a ofrecer. Identificación de necesidades y expectativas de los usuarios. Identificación de parámetros y resultados a conseguir. | <ul style="list-style-type: none"> Diseñar los procesos de atención para conseguir los resultados deseados predefinidos. |

Los principales pasos y actividades a realizar para la mejora de la calidad están reflejadas en la Figura 3.1, que nos va a servir de referencia para revisar las acciones incluidas en cada uno de los tres grandes grupos de actividades.

FIGURA 3.1. Relación entre las diversas actividades y cursos de acción en gestión de la calidad



4.1. CICLOS DE MEJORA

La evaluación, el ciclo de mejora, ciclo evaluativo o ciclo de garantía de calidad de la literatura tradicional sobre calidad en los servicios de salud, comienza con la identificación de una oportunidad de mejora en algún aspecto de los servicios que se están ofreciendo (el "problema" de calidad en la terminología tradicional). Este problema puede ser de tipo asistencial (calidad de los programas, diagnósticos, tratamientos, controles de salud, etc.), organizativo (falta de coordinación, tiempos de espera, circuitos anómalos, etc.) o de cualquier otro tipo (satisfacción, comunicación, información, etc.). El objetivo de las actividades que se generan a partir de que el tema a estudio ("oportunidad de mejora" o "problema" de calidad) se ha identificado y definido, es elevar el nivel de calidad del aspecto que se evalúa a los niveles máximos que permite el contexto en el que se prestan los servicios. Es decir: "aprovechar" la oportunidad de mejora o "solucionar" el problema identificado. Para ello, el curso normal de actividades es: (1) analizar el problema a la luz del conocimiento que se tenga sobre sus posibles causas; (2) explicitar los criterios, requisitos o especificaciones que nos sirvan para medir el nivel de calidad; (3) diseñar y realizar el estudio en sí, y, tras el análisis y discusión de los resultados; y (4) diseñar e implantar la intervención oportuna, cuyo efecto en la calidad ha de ser medido reevaluando de nuevo el tema en estudio.

Los ciclos de mejora aprovechan los problemas de calidad u oportunidades de mejora detectados para someterlos a análisis e intervención y conseguir solucionarlo.

A veces, con el solo análisis del problema y la consideración de los criterios o requisitos de calidad que queremos, resulta evidente lo que hay que hacer para mejorar y se puede diseñar directamente la intervención oportuna. Otras veces, sobre todo cuando el problema puede limitarse a un sólo requisito de calidad (tiempos de espera elevados, insatisfacción, fallos organizativos puntuales y repetidos, etc.) de difícil comprensión y sin relación sabida con criterios de calidad concretos, lo oportuno es hacer un estudio de causas o comprobación de hipótesis que expliquen el incumplimiento del criterio o requisito de calidad en estudio. Estas dos eventualidades están marcadas en el gráfico (Figura 3.1) como vías de acción contingentes o alternativas y completan las actividades relacionables directamente con la evaluación de la calidad, el clásico Ciclo Evaluativo o "Ciclo de garantía de calidad", en el que se evalúa la presencia o ausencia de criterios o requisitos de calidad que ya se sabe de antemano que son importantes (por ejemplo todos los requisitos de una actuación clínica adecuada a un tipo de paciente o problema de salud concreto). En las UT 10 y 11 se da una explicación más detallada de los diferentes tipos de estudio.

No es casual que el ciclo de mejora ocupe el centro de la Figura 3.1. Aparte de ser lo más conocido y fácil de entender es el grupo de actividades central y crucial de cualquier programa de gestión de la calidad. Es también el que más necesita de un enfoque interno, de autocompromiso, pero no es suficiente si queremos tener un programa completo, aunque podamos encontrar programas que se basen exclusivamente en ciclos de mejora.

4.2. MONITORIZACIÓN DE LA CALIDAD

La monitorización la entendemos como una medición periódica, planificada y rutinaria de una serie de aspectos-resumen o indicadores de la calidad. Su objetivo es asegurarnos que nos mantenemos a unos niveles previamente fijados como aceptables. En la práctica, para la monitorización se parte de una selección de aspectos o servicios para los que se elaboran los correspondientes indicadores, que han de ir acompañados de su correspondiente esquema de medición o plan de monitorización, en el que hay que especificar, entre otras muchas cosas, el método y la periodicidad de las mediciones.

Estos indicadores pueden derivarse de ciclos de mejora realizados o ser de entrada contruidos y seleccionados para monitorizar aspectos que consideramos relevantes en los servicios que ofrece el centro. El resultado o fin principal de la monitorización es identificar oportunidades de mejora, o problemas de calidad que merecen de una atención adicional más profunda y detallada.

La monitorización es igualmente útil en los programas o enfoques interno y externo, como después veremos, no como los ciclos de mejora que tienen su máxima aplicabilidad en los programas internos. A veces, junto con la acreditación que es otra forma de monitorizar si se hace de forma periódica, la monitorización es la única actividad desarrollada en los programas externos (de iniciativa externa al centro). Piénsese, por ejemplo, en los listados de indicadores o normas que se manejan para evaluar, desde los niveles centrales, la atención que se presta en los diversos centros sanitarios.

A veces, el análisis del problema sugiere la solución, pero lo normal es que tengamos que medir las posibles causas, sean estas de repercusión conocida o hipotética en el problema de calidad a solucionar

Los ciclos de mejora son una actividad imprescindible en los programas de gestión de la calidad.

La monitorización de la calidad es una medición periódica y planificada de una serie de indicadores, con el fin de controlar que están a los niveles deseables. Sirve para identificar problemas de calidad u oportunidades de mejora

A pesar de su falta de utilidad como actividad aislada, la monitorización resulta muy atractiva y es a veces la única actividad sobre calidad que se lleva a cabo

4.3. EL DISEÑO DE LA CALIDAD

Si los ciclos de mejora "curan" problemas que ya existen, y la monitorización tiene como fin detectar situaciones problemáticas o mejorables, diseñar la calidad es lo mismo que prevenir la aparición de problemas de calidad e implica, en la terminología de la industria, diseñar sistemas y procesos de manera que la calidad sea inevitable. Es decir: facilitar que las cosas tengan que hacerse siempre bien, de manera que el resultado sea el que precisamente hemos fijado como esperado.

El diseño de la calidad es el área de actividades que ha aparecido más tarde en los programas de gestión de la calidad, pero también a la que más importancia se le está dando y, desgraciadamente, la más compleja de dominar. Su práctica parte de definir qué queremos conseguir, en términos de expectativas y necesidades a satisfacer, resultados o parámetros del servicio a diseñar, y a partir de ello decidir qué hacer y cómo hacerlo para lograr el resultado esperado cuando tengamos ocasión de ofrecer el servicio que hemos diseñado.

Lo más común como actividades de diseño de la calidad son los diseños de intervención subsiguientes al análisis del problema o la discusión de los datos en un ciclo de mejora: el diseño de intervenciones es el punto en el que se unen evaluación y diseño de la calidad (Figura 3.1). Por otra parte, en los servicios de salud hay una cierta tradición en diseñar la calidad de la atención clínica: la construcción de protocolos o guías de práctica.

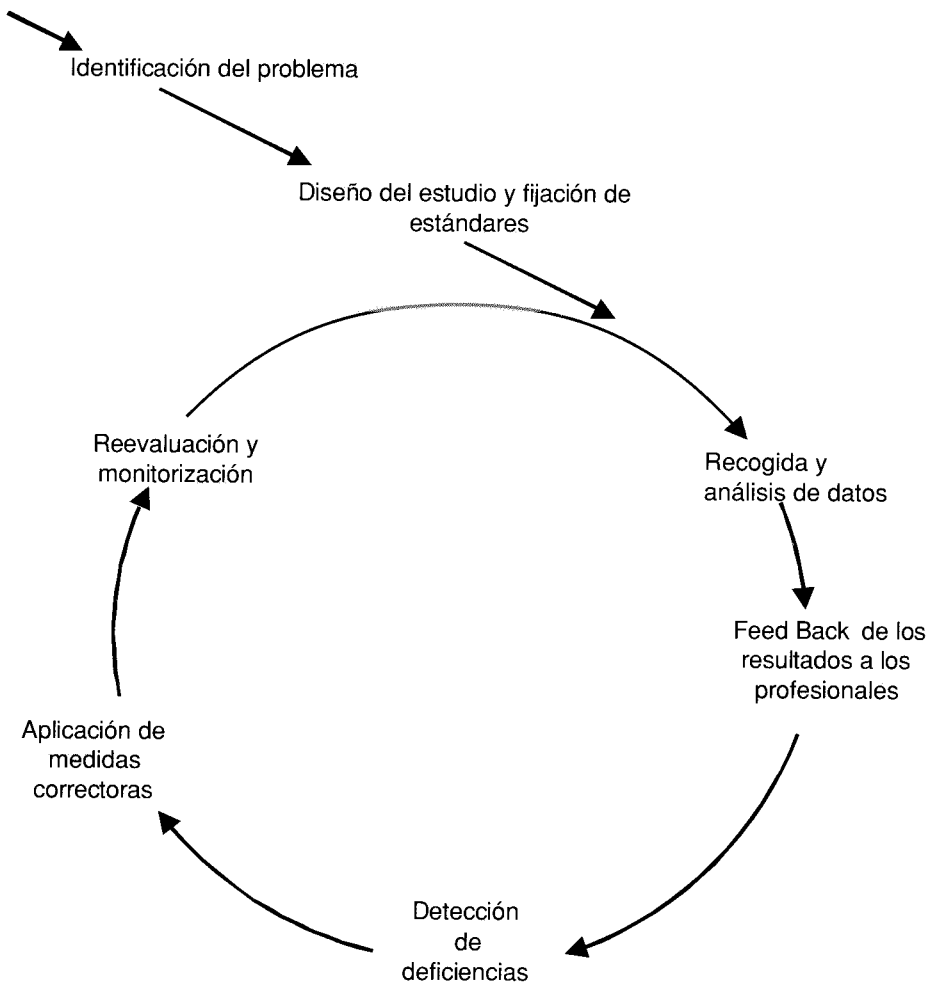
En la actualidad, el enfoque de diseño, con diversos nombres y acepciones (el más popular: reingeniería) es el que se propone como fundamental para los programas de GC, e incluso se llega a aconsejar empezar por ahí. Sin embargo, diseño solo tampoco es suficiente. Por muy perfecto que pensemos que sea, siempre hay que contrastarlo con la realidad, es decir, evaluar los criterios o indicadores que nos digan si funciona o no, y monitorizar aquellos que importan para detectar si se produce o no un deterioro con el tiempo.

5. OTROS MODELOS PARA REPRESENTAR LAS ACTIVIDADES DE UN PROGRAMA DE GC: CICLO DE GARANTÍA DE CALIDAD, TRILOGÍA DE JURAN, CICLO DE DEMING

El esquema de actividades que hemos propuesto y explicado tiene sus raíces en el ciclo de garantía de calidad (Figura 3.2), ampliamente asumido en el pasado en los ámbitos del sector salud, y que hemos ido ampliando a medida que la conceptualización y la práctica de los programas de gestión de calidad han ido necesitando de modelos más completos. El clásico ciclo de garantía de calidad o ciclo evaluativo, solo contempla explícitamente las actividades propias de lo que hemos llamado ciclos de mejora y, en ocasiones, con algunas dificultades de entendimiento en su aplicación, sobre todo cuando el problema identificado no es de tipo clínico, y por tanto no puede evaluarse la calidad directamente en base a criterios de buena práctica previamente conocidos y establecidos, cuya falta de cumplimiento sería la causa directa del problema de calidad.

Con el diseño de la calidad se intenta que las cosas se hagan siempre bien y a la primera según unos parámetros prefijados, previniendo la aparición de problemas de calidad

El Ciclo de Garantía de Calidad representa casi exclusivamente las actividades de los Ciclos de Mejora y resulta difícil de seguir si el problema de calidad no es de tipo clínico.

FIGURA 3.2. Ciclo de Garantía de Calidad


El esquema de actividades completo, compuesto por ciclos de mejora, monitorización y diseño de la calidad, aunque lo hemos ido desarrollando por la necesidad de modelar adecuadamente la realidad práctica y la visión de los programas que íbamos construyendo en centros sanitarios, se asemeja mucho a otros esquemas utilizados en el ámbito de la industria. Uno de los más parecidos y conocidos es la llamada trilogía de Juran. Los términos que representan, con alguna variación, lo que aquí hemos llamado diseño, monitorización y ciclos de mejora, son actividades de *planificación*, *control* y *mejora* para Juran, aunque este autor no ha descendido completamente al detalle de los pasos a incluir en cada grupo de actividades.

Otros esquemas y terminologías sobre las actividades que comprende un programa de gestión de la calidad (o cualquiera de los diversos nombres que pueda recibir tales como garantía de calidad, gestión de la calidad total, mejora continua de la calidad, etc.) se cuentan por decenas. Dentro de la nueva terminología o modelo conocido como "gestión de la calidad total" (Total Quality Management), han surgido multitud de esquemas. La mayoría de estos esquemas representan sólo uno de los tres grupos de actividades (normalmente los ciclos de mejora), confunden innecesariamente los tres, o son adaptaciones de otros es-

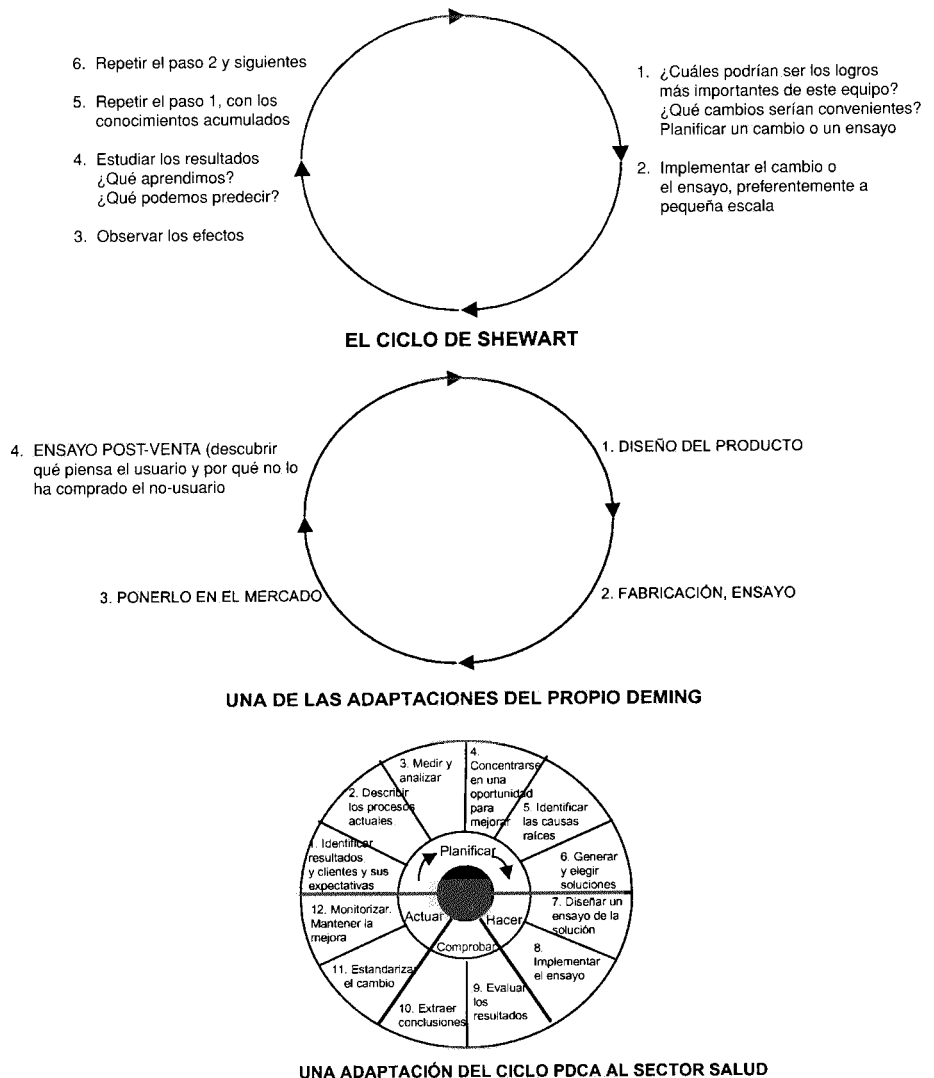
La trilogía de Juran (planificar, controlar y mejorar la calidad) se corresponde con lo que hemos llamado diseño, monitorización y ciclos de mejora

quemas con gran predicamento en la industria que se llegan incluso a desvirtuar. El ejemplo más paradigmático e ilustrativo de esto último son probablemente las múltiples adaptaciones del ciclo de Shewart, que Deming popularizó en Japón y que actualmente se conoce como PDCA (Plan-Do-Check-Act) o ciclo de Deming. El ciclo de Shewart fue mostrado por Deming como un esquema para la mejora de la calidad con base fundamentalmente en el diseño de innovaciones (Plan, planificar), que luego hay que ensayar a pequeña escala (Do, hacer) y medir sus efectos (Check, comprobar o estudiar), para después incorporar los rasgos positivos a la producción rutinaria de ese producto o servicio (Act, actuar); la inevitable interconexión entre los tres grupos de actividades, ha ido produciendo adaptaciones varias que explicitan más pasos de las otras dos (monitorización y solución de problemas), en detrimento de la idea inicial de ser un ciclo que se inicia como actividad de diseño (innovación).

Hay multitud de otros esquemas, de los cuales el más conocido es el PDCA o ciclo de Deming, profusamente adaptado e interpretado de maneras muy diversas

La Figura 3.3 contiene para su comparación el ciclo de Shewart tal como lo explica Deming, una primera adaptación de este autor en donde enfatiza el inicio en base a diseño, y una de las muchas adaptaciones al sector salud, en donde se pueden apreciar diferentes reinterpretaciones y añadidos para incluir las obviedades de los necesarios ciclos de mejora y la monitorización.

FIGURA 3.3. Diferentes adaptaciones del Ciclo de Shewart



6. ¿PARA QUÉ SIRVE UN PROGRAMA DE GESTIÓN DE LA CALIDAD?

La visión de lo que entendemos por un programa de GC no estaría completa sin una idea clara de lo que se puede conseguir en la práctica; es decir, cuáles son sus resultados y justificación.

Podemos afirmar sin temor a equivocarnos que el compromiso explícito, estructurado y funcionante con la calidad sirve para aumentar la efectividad, la eficiencia (a través sobre todo de una reducción drástica de los costes de la calidad deficiente), la satisfacción de proveedores y usuarios, y además nos pone en el camino hacia la excelencia. Es, si se juega bien, un juego sin perdedores: los gestores tienen ante sí una formidable herramienta para orientar su gestión; los profesionales de la salud trabajan mejor y, respondiendo de forma reconocida y responsable a sus necesidades éticas y científicas, se sienten más satisfechos; la población atendida solo puede esperar beneficios en su salud y satisfacción al mejorar la calidad de la atención sanitaria que reciben. La evidencia empírica es abundante tanto en la tradición de los programas de calidad en la industria como en los servicios de salud.

En el campo industrial, la representación más paradigmática de los beneficios que puede aportar la mejora de la calidad es la llamada "reacción en cadena" de Deming: mejor calidad → más productividad → disminución de costes → disminución de precios → aumento de cota de mercado → crecimiento de la empresa y de los beneficios. El énfasis se pone fundamentalmente, sin embargo, en los costes y los beneficios económicos, lo cual es comprensible porque es el lenguaje que mejor entienden los gestores y parece que esta literatura va dirigida sobre todo a este tipo de lectores. Pero hay más beneficios que los puramente económicos; por ejemplo, y continuando en el amplio campo de la industria, un resumen de los beneficios de la GC medidos en 22 compañías finalistas del Baldrige National Quality Award de Estados Unidos incluye, además de una mayor productividad y mayores beneficios de mercado, mejor clima laboral y mayor satisfacción de sus clientes.

En el sector salud la situación es la misma. Las publicaciones que describen los resultados de las iniciativas de GC documentan mejoras muy concretas, cuantificadas e importantes en la efectividad, eficiencia y satisfacción de personal y usuarios.

Todo lo mencionado hasta aquí son efectos y beneficios que tenemos al alcance de la mano y que podemos ver materializarse en el corto y medio plazo, pero hay también otros efectos más intangibles y que van cobrando sentido a más largo plazo: los programas de GC son el camino hacia la excelencia, un objetivo final, cambiante, e inalcanzable por los propios cambios a que está sujeta la visión de la excelencia, pero motor último del amor propio. Fernando Savater lo ha resumido de forma, muy afortunada: "El anhelo de excelencia y perfección, culminación del arte de vivir ético, es el producto más exquisito del amor propio adecuadamente considerado. Quien no desea ser excelente ni perfecto, quien crea que no se merece tanto o no se atreve a proponerse tanto, es que desde luego no se ama lo suficiente a sí mismo; o tiene una idea de la excelencia y perfección puramente ajena, pervertida, esclavizadora, incompatible con las urgencias inaplazables de su yo". Los programas de GC nos ayudan a ser excelentes, es decir, exquisitamente egoístas y a disfrutar con ello.

Los programas de GC aumentan la efectividad, la eficiencia y la satisfacción de proveedores y usuarios

La evidencia empírica justifica una actitud positiva hacia la implementación de programas de GC

Los programas de GC son el camino hacia la excelencia. Punto culminante de la autoestima

7. RAZONES PARA NO TENER UN PROGRAMA DE GESTIÓN DE LA CALIDAD

Si tenemos en cuenta el compromiso ético de los profesionales de la salud con los pacientes a los que atienden, el rechazo a los programas de GC sólo podría estar basado en que la alternativa de no tenerlos fuese más beneficiosa para la población atendida o, para los intereses de los profesionales. Igual argumentación podría utilizarse, tanto para los que proveen atención directa al paciente, como para el nivel gestor y otros proveedores de clientes internos.

También hemos visto que la evidencia empírica sobre los programas de GC es constante y consistentemente la misma: los beneficios son tangibles y manifiestos en comparación con su inexistencia. ¿Cómo se explica entonces que lo más frecuente sea el que no exista programa de GC? Los posibles argumentos para responder pueden resumirse en tres: ignorancia, inercia e incompetencia.

- **Ignorancia.** Está claro que no es exigible la existencia de programas de GC si no se sabe previamente en qué consisten y cómo ponerlos en marcha. La calidad no es aún un tema sobre el que tengan conocimiento y formación ni todos los profesionales de la salud ni todos los gestores. De ahí la importancia, y la necesidad, de realizar esfuerzos de formación como el presente curso.
- **Inercia.** Saber no es suficiente; hay también que querer, y poder superar los condicionantes adversos que se imponen reforzados con el tremendo poder de la inercia, fácilmente asociada a indecisión. Se puede argumentar que estos condicionantes adversos llegan a veces a hacer imposible el programa de GC. En realidad lo que puede ocurrir es que lo hagan más difícil, no imposible. Es en estas situaciones en las que es más evidente la importancia de la capacidad creativa y el ejercicio del liderazgo profesional para ir abriendo brecha. Se trata de dar el paso de tener un compromiso explícito y operativo con la calidad a partir del compromiso implícito que todo profesional tiene.
- **Incompetencia.** Cuando se sabe en qué consisten los programas de GC y existe la posibilidad cierta y evidente de ir venciendo la inercia, el no ponerlos en marcha es ejercer la profesión de forma incompleta. Un síntoma, sobre todo en el nivel gestor, de una cierta incompetencia.

BIBLIOGRAFIA

- Palmer RH. Definir la calidad asistencial. En: Palmer RH. Evaluación de la asistencia ambulatoria. Principios y práctica. Madrid: Ministerio de Sanidad y Consumo; 1990. p. 17-20.
- Palmer RH. Componentes de la evaluación de la asistencia sanitaria. En: Palmer RH. Evaluación de la asistencia ambulatoria. Principios y práctica. Madrid: Ministerio de Sanidad y Consumo; 1990. p. 23-42.
- Saturno PJ. Qué es y para qué sirve un programa de Gestión de la Calidad. En: Saturno PJ, Parra P, Gascón JJ. Tratado de Calidad Asistencial en Atención Primaria. Madrid: DuPont Pharma; 1997. p. 81-114.
- Deming WE. Calidad, productividad y competitividad. La salida de la crisis. Madrid: Diaz de Santos; 1989.
- Feigenbaum AJ. Total Quality Control, 3ª ed. (edición 40 aniversario). New York: McGraw-Hill; 1991.

La no existencia de programa de GC puede justificarse por ignorancia, inercia, incompetencia, o una mezcla de las tres cosas. Para ponerlos en marcha hay que saber, querer y poder

**PROGRAMAS DE GESTIÓN
DE LA CALIDAD.
ENFOQUES Y RESPONSABILIDADES**

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

En esta UT se aclaran los diversos enfoques (interno, externo) y responsabilidades (sistema, personal gestor, profesionales de atención directa, usuarios) para llevar actividades de un programa de gestión de la calidad a la práctica. Se hace un énfasis especial en la aportación del usuario y en el conocimiento de sus expectativas y satisfacción.

ÍNDICE DE CONTENIDOS

1. Responsabilidades y enfoques de los programas de gestión de la calidad. ¿Quién protagoniza los programas de gestión de la calidad?
2. El papel de los gestores.
3. El papel del personal de salud.
4. El papel de los usuarios.
5. Participación activa del usuario o cliente externo en la gestión de la calidad.
6. La participación que solicita el sistema: tipos y características.
7. Expectativas, necesidades y satisfacción.
8. Encuestas de satisfacción. Puntos clave para su utilidad e interpretación.

OBJETIVOS ESPECÍFICOS

1. Explicar las diferencias entre los enfoques interno y externo.
2. Situar los programas de acreditación y de valoración de indicadores propuestos por las gerencias dentro del esquema de grupos de actividades y enfoques.
3. Identificar el tipo de enfoque que corresponde a cada uno de los protagonistas (administración, personal de atención directa, pacientes/población) de los sistemas de salud.
4. Justificar la responsabilidad y el protagonismo que corresponde al personal que presta directamente los servicios, como proveedor y como cliente interno del sistema.
5. Distinguir las diversas formas en las que los pacientes/población pueden contribuir a los programas de Gestión de la Calidad.
6. Describir las diferencias entre encuestas de valoración subjetiva y objetiva de la Calidad.
7. Delimitar los conceptos de "expectativas", "necesidades" y "satisfacción" de los usuarios.
8. Distinguir diferentes tipos de expectativas.
9. Razonar la relación entre expectativas y satisfacción.
10. Analizar críticamente las encuestas de satisfacción.

1. RESPONSABILIDADES Y ENFOQUES DE LOS PROGRAMAS DE GESTIÓN DE LA CALIDAD. ¿QUIÉN PROTAGONIZA LOS PROGRAMAS DE GESTIÓN DE LA CALIDAD?

Como respuesta a esta pregunta, muy probablemente todos estaríamos de acuerdo en que la calidad en los servicios de salud tiene tres grupos de protagonistas directos (personal de gestión, personal que provee directamente los servicios y población atendida), y que los tres tienen algo que decir y hacer para mejorar la calidad. El problema no es, sin embargo identificar quienes son los protagonistas sino tener claro *cómo* ejercer este protagonismo. Para ello hay que distinguir en primer lugar dos tipos de enfoques y programas complementarios, externo e interno, y revisar después cuales de ellos están al alcance de cada uno, cuales son sus principales actividades, y cómo se justifica su realización para conseguir el objetivo final de la mejora continua de la calidad.

1.1. LOS ENFOQUES EXTERNO E INTERNO

Las iniciativas de evaluación y mejora de la calidad pueden tener dos enfoques -*interno* y *externo*- según que partan de los implicados directamente en aquello que se evalúa (autoevaluación, enfoque *interno*) o se pretenda evaluar las actuaciones ajenas (enfoque *externo*). Es esta una distinción que se ha subrayado de forma particular en el sector sanitario, mientras que en la literatura sobre calidad en la industria el énfasis casi exclusivo es en los programas internos. Una posible explicación sería, además de que como luego veremos la calidad solo puede ser mejorada consistentemente desde dentro, que las empresas son cada una de ellas un *sistema* en sí mismas, mientras que los sistemas de salud incluyen multitud de centros, instituciones o *microsistemas* con una gran autonomía de funcionamiento; en cada uno de estos "microsistemas" es posible definir un programa interno en el sentido de los que se describen en esta UT, mientras que a nadie escapa la conveniencia de tener otros programas que afecten a todo el sistema, o a segmentos importantes del mismo (atención primaria, hospitalaria, salud mental, etc.).

Algunos autores en el sector salud, distinguen una tercera categoría que podríamos llamar enfoque externo en programas internos, en referencia a las evaluaciones que se pueden hacer de unos profesionales sobre otros dentro de un centro de salud o en un hospital, siguiendo una línea jerárquica o de otro tipo (por ejemplo de tutores sobre sus residentes, de gerentes sobre los servicios, de coordinadores sobre coordinados, etc.) estas evaluaciones son internas en relación a la institución, pero no para el grupo que está siendo evaluado, por lo que se asemejan más a programas externos. Pero, ¿cuáles son las características principales de cada enfoque?, ¿cuáles sus ventajas e inconvenientes?, ¿cuál es más importante para lograr la mejora de la calidad? Vamos a intentar responder en base a la experiencia acumulada hasta ahora.

- **Enfoque interno.** En el enfoque *interno*, los que son evaluados *quieren* la evaluación, están implicados en el establecimiento de los criterios e indicadores que van a ser utilizados, y participan, de forma activa o delegada en la totalidad de actividades de los ciclos de mejora, monitorización y diseño. La utilidad de las iniciativas internas depende directamente del compromiso de los propios evaluados en ser rigurosos con su evaluación y con la mejora de la

La gestión de la calidad tiene tres protagonistas; gestores, personal de salud y población que recibe los servicios.

La iniciativa para evaluar marca dos enfoques diferentes: interno, si parte de quienes ofrecen el servicio que se evalúa; externo si parte de personas ajenas o protagonistas indirectos de la atención evaluada.

El éxito de los programas internos depende directamente del compromiso de los profesionales en ser rigurosos con su evaluación y con la mejora de la calidad. Hay que estar alerta a una serie de inconvenientes, el primero de los cuales es tener que buscar el necesario tiempo y esfuerzo para autoevaluarse y mejorar.

Los programas externos sirven a un propósito socialmente más amplio y permiten la comparación entre centros. Su principal inconveniente es que revistan características de fiscalización y utilicen indicadores poco válidos o poco fiables.

Los programas internos y externos se complementan. Pero los internos son imprescindibles para la mejora continua.

calidad. Una de sus principales ventajas es que reconoce el grado de responsabilidad y autonomía que indudablemente tienen en la prestación de los servicios los profesionales de la salud, además de que es internamente como mejor se entiende el contexto en que se da la asistencia y los medios disponibles para mejorarla; como consecuencia, los programas internos se adaptan a las condiciones locales, y al grado de sofisticación e interés de los profesionales; se enfatiza la profesionalidad y se canaliza el deseo de la mayoría de hacer las cosas de la mejor manera posible; además, el ejercicio interno del compromiso con la calidad es lógico que facilite la sinergia con las iniciativas externas para el mismo fin. Entre los inconvenientes, no es el menor al principio el tener que buscar el necesario tiempo y esfuerzo para autoevaluarse y mejorar; una vez superado éste pueden aparecer otros, como son determinados conflictos de autoridad, el desarrollo de métodos demasiado idiosincrásicos, con poca base científica, la autocomplacencia por la falta de una referencia externa para compararse, o la falta de intervenciones y seguimiento sobre determinados problemas internos que pueden aparecer.

- **Enfoque externo.** En el enfoque *externo*, aparte de la iniciativa en sí para evaluar, las normas de la evaluación o monitorización son establecidas por personas distintas de los profesionales que van a ser evaluados. Su principal ventaja es que, al menos en teoría, sirven para un propósito socialmente más amplio: el que corresponde a todo el sistema de salud en cualquiera de sus niveles; adicionalmente, los programas externos permiten las comparaciones entre centros semejantes, pueden llamar la atención sobre problemas que de otra forma pasarían inadvertidos, y hace mucho más difícil la autocomplacencia o la ocultación de problemas. Sin embargo, uno de sus principales inconvenientes es que revistan características de fiscalización o inspección, con el consiguiente rechazo de los inspeccionados, en vez de ayudar a la evaluación interna, a la identificación de problemas y, en definitiva al proceso de mejora de la calidad. Los programas externos precisan también de una considerable infraestructura organizativa y, sobre todo, precisan utilizar normas, criterios e indicadores suficientemente fiables y válidos para su propósito; los programas externos pueden utilizar estándares artificiales o irrelevantes, desconociendo o ignorando las circunstancias de la asistencia diaria y, lo que es peor, los medios con los que se puede mejorar la asistencia en un centro concreto.

¿Cómo proceder entonces? Parece obvio que ambos enfoques se complementan, ambos son necesarios para un acercamiento global, de sistema de salud, a la mejora de la calidad, y ninguno es por sí solo una respuesta suficiente desde el punto de vista de la población en su conjunto. Sin embargo, aún insistiendo en la complementariedad de los dos tipos de enfoques, lo que tenemos al alcance en cualquiera de los niveles que nos encontremos en el sistema de salud son las actividades de evaluación y autocorrección desde la perspectiva interna. Pero hay aún más razones para traer a primer término el enfoque interno. Ya hemos comentado cómo el entender y practicar la gestión de la calidad como algo inherente al funcionamiento diario de los respectivos centros y servicios, va a favorecer la superación de resistencias frente a propuestas de sistema que pueden ser percibidas como "fiscalización" del exterior, y no como una necesidad de optimizar profesional y socialmente los servicios de salud. Adicionalmente, los análisis realizados de los muchos años de experiencia con programas de todo tipo en el multisistema de Estados Unidos, concluyen sobre lo imprescindible

que resultan los programas internos para lograr la mejora de la calidad, poniendo en primer término el interés y la responsabilidad de los profesionales que la protagonizan. Como resume Palmer: desde fuera se puede medir, desde dentro se puede medir y mejorar.

2. EL PAPEL DE LOS GESTORES

Una de las principales características de los más recientes enfoques de los programas de GC es el reconocimiento de la máxima responsabilidad que tiene el personal gestor para que el programa exista y funcione. En los libros y manuales sobre implementación de programas de GC con los métodos tomados de la industria, se afirma de forma totalmente inequívoca que sin el compromiso activo de los directivos no hay programa posible. Sin embargo, el sector salud tiene ciertas peculiaridades, relativamente fáciles de reconocer. Se ha visto, por ejemplo, que la participación de los médicos es uno de los factores clave para el éxito del programa.

Los necesarios programas externos, tanto los de nivel de sistema como los de niveles intermedios, son una responsabilidad ineludible del nivel gestor. La forma en que estén diseñados e implementados y la medida en que contribuyan o no a la mejora de la calidad, va a depender ampliamente de la capacidad de gestión de sus impulsores. El nivel gestor es también responsable de impulsar y mantener los programas internos de los centros a su cargo; pero para ello, todos los autores subrayan la necesidad de tener en cuenta las peculiaridades de la actividad y el personal que ha de llevar adelante en la práctica las actividades de GC y, en definitiva, la mejora de la calidad.

3. EL PAPEL DEL PERSONAL DE SALUD

El personal encargado de proveer directamente la asistencia sanitaria es la base insustituible de los programas internos y de todo el sistema. Sin su compromiso y participación, la mejora de la calidad no es posible, aunque exista oficialmente un programa de GC. En uno de los primeros libros sobre las experiencias del enfoque industrial aplicado a la salud, editado entre otros por Donald Berwick, se reconoce taxativamente.

La participación del personal de salud en los programas internos de sus centros asistenciales, en su autoevaluación para asegurar y mejorar la calidad de la atención que ofrecen, puede justificarse por razones eminentemente prácticas: no quedarse al margen de algo que se va imponiendo hasta hacerse inevitable; sin embargo, su justificación más profunda es por razones de ética profesional; la evaluación y mejora de la calidad es un deber que tienen consigo mismo y con sus pacientes. El rechazo de los programas de GC sólo sería justificable si la alternativa de no participar se mostrase más beneficiosa para los intereses de los pacientes, de los que son ineludiblemente responsables, y del propio personal de salud; lo cual no parece ser el caso a la vista de los abundantes problemas de atención, muchos de ellos previamente inadvertidos, que afloran cada vez que se evalúa la calidad de los servicios de salud.

Por otra parte, el grado de autonomía en las decisiones a la hora de la pres-

Se ofrece a continuación una revisión más detallada del papel de los tres protagonistas, con un énfasis especial en lo relativo a la participación de los usuarios.

El nivel gestor es responsable de que los programas existan y funcionen, y de la actuación sinérgica de los programas externos e internos, que deben ser respetados y fomentados.. En los programas externos son además responsables de su diseño y efectividad.

Sin el compromiso y la participación del personal de salud no funcionan los programas internos, ni es posible la mejora de la calidad.

El rechazo a los programas de GC sólo se justificaría si el no tenerlos fuese mejor para los pacientes y los propios profesionales.

tación de servicios que tiene el personal de salud es probablemente de los más elevados que puedan encontrarse; en muchos de los aspectos un solo médico, una enfermera, o un mínimo equipo son en gran parte gestores de sí mismos, de su tiempo y de sus acciones, y las decisiones que tomen tienen una gran trascendencia para la calidad del servicio que ofrecen. Adquieren así a su nivel, las mismas responsabilidades que puedan ser atribuidas a los gestores de niveles superiores de la organización sanitaria.

Adicionalmente, en el sector salud la influencia de los líderes profesionales puede ser mucho mayor que la de los gestores a la hora de introducir actividades de GC; este hecho, comprobado empíricamente en más de una ocasión, confiere una responsabilidad adicional a los buenos profesionales en la práctica pionera de las actividades de GC.

4. EL PAPEL DE LOS USUARIOS

De una forma esquemática, podemos clasificar las principales formas en que el usuario puede contribuir a la GC en dos grandes grupos metodológicos, según que sea *activa* (a iniciativa propia) o *solicitada* por el sistema de salud. La distinción es importante metodológicamente pero, al igual que ocurre con los enfoques interno y externo, ambas iniciativas no deben considerarse como excluyentes. Ambas son útiles, ambas son teóricamente necesarias, y por tanto el grado de desarrollo de alguna de ellas no debería oscurecer a la otra, ni impedir que se avance en el conocimiento de métodos y mecanismos de optimización de cada una de las formas de participación posible. No puede decirse sin embargo que se haya demostrado un parecido nivel de utilidad en GC para todas ellas, ni que la metodología de su puesta en marcha y aprovechamiento efectivo para la GC esté igualmente desarrollada. En la actualidad, aún dentro de las muchas dificultades y puntos oscuros que todavía existen, parece que se sabe mucho más de los métodos y aprovechamiento rutinario de la evaluación *solicitada* que de las diversas formas de evaluación activa.

Pero, ¿cuáles son las formas concretas, activas o solicitadas, que pueden vehicular la contribución del usuario a la GC? Vamos a revisarlas a continuación.

5. PARTICIPACIÓN ACTIVA DEL USUARIO O CLIENTE EXTERNO EN LA GESTIÓN DE LA CALIDAD

Dejando al margen los cauces de participación política en general, el usuario puede *activamente* cambiar de proveedor (médico, equipo, centro u hospital), formular quejas y reclamaciones, actuar a través de asociaciones de usuarios y consumidores, y participar a través de los Consejos de Salud y órganos similares en la política de calidad de las instituciones concretas a las que se adscriban (Tabla 4.1).

El grado de autonomía en las decisiones que afectan a la calidad y la influencia del liderazgo profesional en la implementación de las actividades de GC, son dos razones adicionales para subrayar el importante papel de los profesionales.

El usuario puede participar por decisión propia o porque se le solicite. Ambas formas son útiles para la GC.

TABLA 4.1. Formas de participación activa del usuario en la evaluación y mejora de la calidad

- ELECCIÓN PROVEEDORES
 - QUEJAS Y RECLAMACIONES
 - ASOCIACIONES USUARIOS Y CONSUMIDORES
 - CONSEJOS SALUD Y OTROS ÓRGANOS DE PARTICIPACIÓN COMUNITARIA EN SALUD
- **El cambio de proveedor**, considerado como resultado final de una evaluación global de la atención que se recibe, está constantemente incorporado a las discusiones y reformas de los servicios de salud que intentan aumentar la eficiencia y la calidad a través de la competencia. El usuario iría a quién o quienes le ofrecieran mejor calidad o mejor relación calidad/coste, contribuyendo así a que las instituciones y profesionales se esforzaran en no perder "clientes" aumentando su calidad y su eficiencia. A pesar de su lógica aparente, no parece haber, sin embargo, estudios concluyentes sobre la efectividad real de esta manera de contribuir a la eficiencia y calidad de los servicios de salud. Es aún una discusión abierta. Una iniciativa interesante, aunque controvertida, para que el cliente potencial evalúe la calidad de los servicios y actúe en consecuencia es la de los informes de indicadores de calidad (*report cards*) que algunas instituciones estadounidenses ponen a disposición del público.
 - **Las "ventanillas", libros o puntos de recepción de quejas y reclamaciones** existen desde hace mucho tiempo en la práctica totalidad de instituciones sanitarias. Su utilidad más ampliamente reconocida para la GC es como mecanismo de identificación de problemas u oportunidades de mejora. Sin embargo su utilización real y la forma en que se estructura en general su presencia en la organización de las instituciones de salud, dista mucho de ser positiva y homogénea. No parece infrecuente el que no se utilicen en absoluto. Lo más normal es probablemente una franca infrautilización, acompañada de una consciente o inconsciente desincentivación al usuario para que las practique.
 - **Las asociaciones de consumidores**, bien desarrolladas en otros países y con un papel de exigencia creciente en cuanto a los datos a contrastar para poder valorar la calidad asistencial de las instituciones, apenas han intervenido en este sentido en nuestro país. En teoría es un campo potencialmente muy rico y que se debiera desarrollar más. Sin embargo, experiencias como la ocurrida en Estados Unidos de hacer públicas y juzgar la calidad de los hospitales por su tasa de mortalidad o de infecciones han sido tan justamente contestadas que han de alertar sobre la sobresimplificación y serios problemas de validez que tiene la idea de la calidad global y la consiguiente identificación y uso de indicadores supuestamente "representativos". Son demasiado numerosos los servicios diferentes que ofrecen las instituciones sanitarias y los factores de confusión a tener en cuenta, como para que sea simple y operativa la idea del indicador global. Lo cual no invalida, desde luego, el papel que pueden jugar las asociaciones de usuarios utilizando otro tipo de enfoques más válidos. De

Las diversas formas de participación activa tienen un desarrollo y aprovechamiento para la GC muy desigual, pero son un campo teóricamente prometedor.

la misma manera, podemos incluir en este grupo otros tipos de asociaciones de la población en general (por ejemplo, Asociaciones de Vecinos), de ex-pacientes, o de personal sanitario (Sociedades científicas, Colegios profesionales).

- **Los Consejos de Salud** y órganos similares de participación de la comunidad en las instituciones sanitarias, propugnados por la Ley General de Sanidad, son en teoría el ámbito ideal para la participación también en las actividades de GC, especialmente en su vertiente de enfoque interno, comprometido. Desde los Consejos de Salud se podrían establecer prioridades en relación a los problemas y oportunidades de mejora a abordar, ayudar en el diseño de los aspectos organizativos que afecten al trato y la accesibilidad del usuario, y ayudar también en la priorización, diseño y puesta en marcha de nuevos servicios o modificaciones de los existentes, optimizando los recursos disponibles. Completando el panorama de la utilidad potencial de la participación comunitaria a través de los Consejos de Salud, hay autores que la consideran imprescindible para poner en marcha y mejorar las actividades de promoción de la salud. Sin embargo, la realidad parece ser otra. Son aún escasas, aunque las hay, las experiencias abiertamente positivas. Lo más común entre los profesionales del sistema es una tendencia a considerar la articulación efectiva de este tipo de participación como un reto metodológico sin resolver o, directamente, prescindible o falta de interés.

En comparación, hay mucho más camino recorrido en relación a la participación *solicitada*, y es actualmente incuestionable su importancia para los programas de GC. Más aún: se tiende a considerarla como imprescindible. Por otra parte, la puesta en marcha y aprovechamiento rutinario de las formas de participación solicitadas al usuario para mejorar la calidad del servicio, cae directamente en el ámbito de responsabilidad del sistema. Aunque sólo sea por este conjunto de razones, merece la pena que nos detengamos un poco más en describir las características de este tipo de participación.

6. LA PARTICIPACIÓN QUE SOLICITA EL SISTEMA: TIPOS Y CARACTERÍSTICAS

El sistema de salud, a cualquiera de sus niveles, puede incorporar a sus actividades de GC la evaluación que hace el usuario de los servicios que se le ofrecen *preguntándole*. Esta es la idea simple y básica. Sin embargo su práctica correcta y eficiente precisa de una serie de consideraciones metodológicas, relacionadas con las características de lo que se pregunta y su utilización prevista para la GC (Tabla 4.2).

TABLA 4.2. Formas de participación de los usuarios solicitadas por el Sistema de Salud

- ENCUESTAS DE SATISFACCIÓN, OPINIÓN Y EXPECTATIVAS
- GRUPOS FOCALES, INFORMADORES CLAVE Y OTRAS TÉCNICAS DE INVESTIGACIÓN CUALITATIVA
- INFORMES DE LOS USUARIOS

Una primera distinción debe establecerse entre aquellas preguntas que exploran percepciones *subjetivas* y las que buscan el aporte de datos *objetivos* sobre el servicio recibido. El representante genuino prácticamente único de esta última modalidad son los llamados *informes de los usuarios (Patients' Reports)*. Entre las subjetivas se encuentran las encuestas de satisfacción, y expectativas del usuario, y también en general las comparativamente menos frecuentemente empleadas técnicas de investigación cualitativa, como los grupos focales, útiles tanto para detectar déficits sentidos por el usuario cómo para averiguar sus expectativas del servicio que recibe e intentar diseñar los servicios y las encuestas de satisfacción en torno al cumplimiento de estas expectativas.

La distinción entre indagar evaluación subjetiva u objetiva, tiene trascendencia tanto en el diseño del instrumento en sí, como en la utilidad y tipo de intervención que puede generar en GC. En el *informe* se va buscando información sobre si han ocurrido o no aspectos concretos de la atención. En las encuestas de opinión se busca la valoración sobre lo que se ha hecho, según la percepción del usuario. Un ejemplo de esta distinción puede verse en la Tabla 4.3

TABLA 4.3. Principales diferencias entre las preguntas de las encuestas de opinión y los informes

| | ENFOQUE | ASPECTOS TÉCNICOS DE LA ATENCIÓN | ASPECTOS INTERPERSONALES DE LA ATENCIÓN |
|---------|-----------|--|--|
| OPINION | Subjetivo | <i>Evaluación en una escala (ej: malo a excelente) de la atención profesional recibida</i> | <i>Evaluación en una escala (ej: mala a excelente) de la cortesía y respeto mostrado por los profesionales</i> |
| INFORME | Objetivo | <i>Indicación (si/no) de realización de acciones concretas (ej: toma de TA)</i> | <i>Indicación (si/no) de realización de acciones concretas (ej: si el médico le saludó al entrar)</i> |

7. EXPECTATIVAS, NECESIDADES Y SATISFACCIÓN

El estudio de las expectativas y necesidades de la población a "servir" va cobrando cada vez más fuerza como una actividad necesaria en las actividades de GC. Este énfasis es un trasplante más al sector de salud de lo que es un principio indiscutible en otras industrias de servicios. Para su introducción en los programas de GC, se dice incluso que el conocimiento de las necesidades y

expectativas de los usuarios es más importante que el de la satisfacción. Sin embargo, una de las aplicaciones en las que se ha utilizado este enfoque es no tanto o únicamente para *diseñar* los servicios a ofrecer según las necesidades y expectativas de la población a atender, sino para saber qué preguntar en las encuestas de satisfacción como posibles causas que ya sabemos son valoradas por los usuarios.

La diferencia entre necesidades y expectativas no siempre está clara. En otros sectores pueden coincidir, al considerarse ambas como algo subjetivo, sentido por el cliente. En el sector salud, en cambio, la necesidad de atención sanitaria definida desde el punto de vista de los profesionales no siempre es sentida y dimensionada por el usuario. En este caso, la definición "objetiva" de necesidad de salud quedaría como responsabilidad mayormente de los profesionales sanitarios, mientras que habría que contar con el usuario para averiguar las expectativas, y el componente de necesidad percibida, subjetiva, sobre problemas de salud. Las expectativas pueden explorarse con técnicas cualitativas (por ejemplo grupos focales) y con cuestionarios específicos o unidos a los de satisfacción.

Sin quitarle importancia alguna a las expectativas (de hecho se ha visto que pueden influir incluso en la sensación de mejora tras intervenciones quirúrgicas) hay que advertir no obstante sobre un peligro muy cierto de sobrevalorarlas, así como, por extensión, sobrevalorar también lo que se conoce como *calidad percibida* y sus posibles causas. El peligro consiste en olvidar o no prestar suficiente atención a otras expectativas, no siempre explicitadas por el usuario, y al componente de necesidad de salud no sentida. De los varios modelos que existen para explicar los distintos tipos de expectativas y su relación con la satisfacción, uno de los más útiles y simples es el de N. Kano, que distingue tres tipos de expectativas: de calidad *expresada*, de calidad *esperada* y de calidad *inesperada*.

Si le preguntamos sobre lo que es un servicio de calidad, el usuario va a expresar una serie de expectativas y características que responden a su visión más inmediata de la calidad. Es la calidad *expresada* o unidimensional del modelo de N. Kano, equivalente a las causas más frecuentemente encontradas en las encuestas de satisfacción (buen trato, poco tiempo de espera, etc.). La satisfacción será mayor o menor según que estas expectativas se hayan cumplido más o menos.

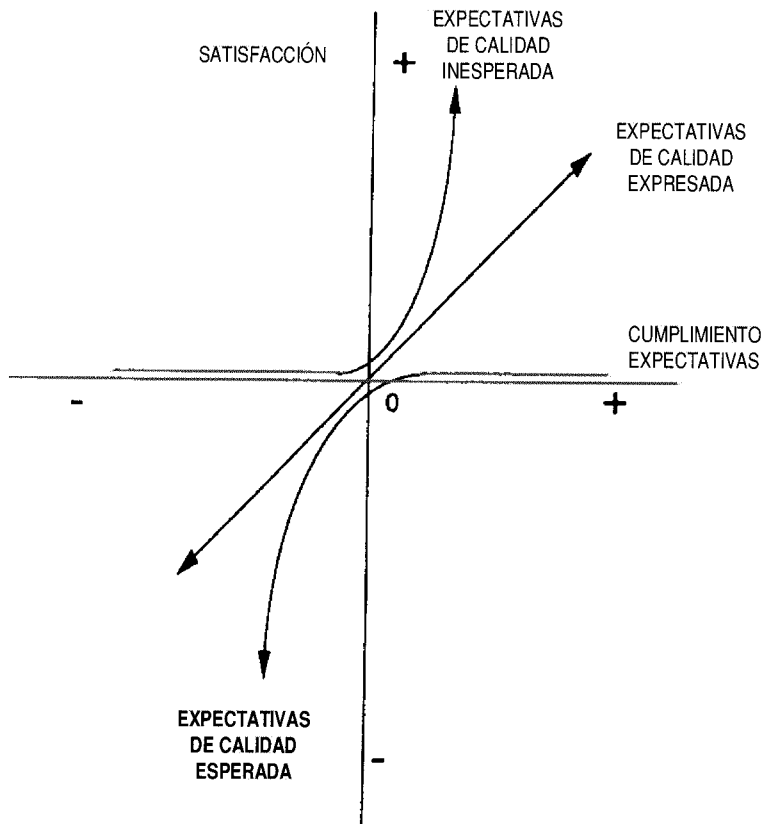
Las expectativas de calidad *esperada* son todo lo que se da por supuesto, y que en los servicios de salud se refiere casi exclusivamente a características científico-técnicas. Por ejemplo, el paciente espera que el médico sepa su oficio y que por lo tanto los diagnósticos y tratamientos sean correctos y sin complicaciones; igual que uno espera que el chófer de un autobús sepa conducir. La presencia de estas características se considera normal y su ausencia, si se percibe, produce insatisfacción en una medida exponencialmente negativa. Efectivamente, en el sector salud, la evidencia empírica señala que la calidad técnica de la atención sanitaria no es adecuadamente valorada por el usuario y que su variabilidad objetiva no se traduce en una variabilidad del mismo sentido en cuanto a satisfacción, mientras que los errores médicos percibidos producen una profunda insatisfacción. Sería en cambio irresponsable, cegados por su normalmente poca o nula influencia en la satisfacción, dada la escasa capacidad de discriminación de los pacientes en este sentido, no considerar la calidad científico-técnica al diseñar y evaluar la calidad de la atención.

El conocimiento de las necesidades y expectativas es útil para diseñar los procesos de atención y para construir encuestas de satisfacción contextualizadas

La calidad técnica de la atención no es adecuadamente valorada por el usuario y normalmente sólo influye en su (in)satisfacción si percibe que es mala.

Finalmente, las expectativas de calidad *inesperada*, se refieren a sobreatenciones, detalles de trato, consideración, rapidez, etc. que sorprenden al usuario porque no las considera normales (en la onda de la calidad expresada), ni jamás va a dar por supuesta su existencia (en la onda de la calidad esperada). La calidad *inesperada*, si aparece, maravilla al usuario, lo sobresatisface, en una medida exponencial. La Figura 4.1 representa en forma gráfica el modelo de Kano.

FIGURA 4.1. Relación entre los tipos de expectativas y la satisfacción



La conclusión es que los tres tipos de expectativas deben de tenerse en cuenta, y que es fácil sesgar las apreciaciones y las acciones hacia una de ellas (la *expresada*) en peligroso detrimento de la *esperada*, y que si lo que preocupa de forma prioritaria es la calidad aparente y la satisfacción del usuario, merecería la pena investigar o intentar incidir sobre la calidad *inesperada*, relacionada precisamente con expectativas inexistentes o raramente expresadas por el usuario de forma espontánea.

Sin embargo, una cuestión adicional a tener en cuenta es que las expectativas, y por lo tanto los motivos de satisfacción/insatisfacción, evolucionan de forma dinámica. Lo que serían expectativas de calidad inesperada pueden convertirse en calidad esperada y relacionarse de forma diferente con la satisfacción. Por ello se necesita una exploración periódica de las expectativas y satisfacción de la población que atendemos, una vez entramos de lleno en la dinámica de la mejora continua de la calidad.

Existen varios tipos o niveles de expectativas, en continua evolución, cuyo cumplimiento repercute de forma diferente en la satisfacción; de forma que se hace necesaria una exploración periódica de las expectativas y una relativización de las encuestas de satisfacción.

8. ENCUESTAS DE SATISFACCIÓN. PUNTOS CLAVE PARA SU UTILIDAD E INTERPRETACIÓN

La satisfacción es una de las dimensiones que forma parte de la mayoría de las definiciones de calidad asistencial, y además uno de los resultados a considerar como fruto del propio proceso asistencial. Las encuestas de satisfacción parecen un componente obvio de los programas de GC. Es difícil minimizar la importancia de poder medir la satisfacción, y hacerlo bien. Sin embargo, la medición debe de servir para diseñar intervenciones que mejoren, si ello es posible, la situación de partida. Estas dos condicionantes -la *validez* y la *utilidad* de la medición- no siempre se respetan en los diseños y explotación de las encuestas de satisfacción, que pueden convertirse con facilidad en ejercicios, inútiles y científicamente espermáticos.

Para ser *útil* en GC, una encuesta de satisfacción debe indagar además del grado de satisfacción o insatisfacción las *causas* con que se relaciona. De otro modo, no sabríamos qué hacer para mejorar si resultaran unos niveles de satisfacción bajos. Para ser *válida* y poder ser interpretada sobre todo en comparaciones con otras mediciones, hay que tener en cuenta en el diseño de la aplicación de la encuesta y/o en los propios apartados de la misma para incluirlos en el análisis, los *factores de confusión* que ya se sabe que influyen en los niveles de satisfacción y que en cambio no tienen nada que ver con que el servicio evaluado sea mejor o peor.

La Tabla 4.4 resume las principales causas y factores de confusión en torno a la satisfacción, según se desprende de la literatura publicada sobre el tema. Cada uno de los apartados merecería un comentario detallado. Destaquemos únicamente que entre las *causas* parece que la más constantemente y sobre la que hay más evidencia causal es la relativa al trato recibido, el componente interpersonal y la comunicación establecida entre el usuario y los profesionales. En relación a los *factores de confusión*, es notorio lo frecuente que viene siendo el no tenerlos en cuenta en comparaciones entre instituciones, provincias, Comunidades Autónomas e incluso países.

Para ser útiles las encuestas de satisfacción deben de realizarse a quienes hayan utilizado los servicios que se evalúan y deben de indagar las causas de insatisfacción, para así saber qué hacer para mejorar.

TABLA 4.4. Satisfacción del usuario. Principales causas y factores de confusión

CAUSAS

- RELACIÓN INTERPERSONAL (trato recibido)
- ACCESIBILIDAD: BARRERAS ORGANIZATIVAS (tiempos de espera, falta de coordinación entre niveles, etc.)
- RESOLUCIÓN DE PROBLEMAS SEGÚN EXPECTATIVAS
- PERCEPCIÓN DEL NIVEL TÉCNICO-CIENTÍFICO DE LOS PROFESIONALES
- CONDICIONES MEDIOAMBIENTALES DE LA ASISTENCIA

FACTORES DE CONFUSIÓN

- EDAD
- EXPERIENCIAS PREVIAS
- ESTADO DE SALUD
- NIVEL CULTURAL
- FORMA DE FINANCIACIÓN
- TIEMPO PASADO DESDE QUE SE RECIBIÓ LA ASISTENCIA HASTA LA CONTESTACIÓN DE LA ENCUESTA

Los resultados de las encuestas de satisfacción hay que analizarlos como mínimo estratificándolos por estos factores o variables que se sabe influyen por sí solos en la satisfacción. Por ejemplo, puesto que se sabe que los ancianos tienden a estar más satisfechos, es conveniente calcular la satisfacción por grupos de edad y no sólo la global; la satisfacción global, sin estratificar por edad, puede ocultar niveles preocupantes de insatisfacción en usuarios más jóvenes.

Las encuestas de satisfacción, razonablemente diseñadas, tienen su máxima aplicabilidad en su uso rutinario como parte de los programas de GC internos de los centros sanitarios, y como herramienta comparativa (si están bien diseñadas y analizadas) en los programas externos.

BIBLIOGRAFIA

- Saturno PJ. ¿Qué es y para qué sirve un programa de gestión de la calidad? Tratado de Calidad Asistencial en Atención Primaria. Madrid: Dupont Pharma; 1997.
- Juran JM, Gryna FM, Bingham RS. Manual de Control de la Calidad. 2ª ed. Barcelona: Reverté; 1990.
- Deming WE. Calidad, productividad y competitividad. La salida de la crisis. Madrid: Díaz de Santos; 1989.
- Lohr KN (ed.). Medicare: A Strategy for Quality Assurance. Washington DC: National Academy Press; 1990.

PROGRAMAS DE GESTIÓN DE LA CALIDAD: ESTRATEGIAS DE IMPLANTACIÓN

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

Una vez asumido el marco de referencia general (Concepto de Calidad, Misión y Visión del Centro, Líneas Estratégicas de actuación), e identificados los grandes grupos de actividades y responsabilidades en la Gestión de la Calidad, en esta UT se explicitan las posibles respuestas a la pregunta ¿cómo empezar? Se consideran los condicionantes y requisitos que implica empezar a practicar el compromiso con la mejora de la calidad, y se elige una de las formas de comienzo: la realización de un ciclo de mejora.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. Métodos para el diseño de estrategias. Análisis de campos de fuerza.
3. Análisis de campos de fuerza en nuestro contexto.
4. Introducción de innovaciones. Masa crítica y demografía del cambio.
5. Las primeras actividades para mejorar la calidad ¿diseñar, monitorizar o realizar ciclos de mejora?
6. Pasos de un Ciclo de Mejora.

OBJETIVOS ESPECÍFICOS

1. Describir los factores que favorecen y que obstaculizan el inicio de actividades de evaluación y mejora de la calidad en su lugar de trabajo.
2. Analizar estos factores a favor y en contra utilizando un diagrama de campos de fuerza.
3. Definir una estrategia de actuación consecuente con el análisis efectuado.
4. Aplicar los conceptos de la "demografía del cambio" en su centro de trabajo
5. Identificar aliados para la implantación de las actividades de evaluación y mejora de la calidad.
6. Razonar las ventajas e inconvenientes de las tres principales maneras de comenzar las actividades para mejorar la calidad.
7. Distinguir las diversas fases de un ciclo de mejora.

1. INTRODUCCIÓN

Llega el momento en que vamos a iniciar con experiencias propias actividades de evaluación y mejora de la calidad. Tenemos asumido después de revisar la UT1, un marco general que incluye un determinado concepto de calidad y un contexto como proveedores en el que hacerlo operativo; también hemos visto en las UT 3 y 4 qué es lo que significa tener un programa de gestión de la calidad en nuestro centro y cómo ejercer el protagonismo que nos corresponde en la mejora de calidad, según el lugar en que nos encontremos dentro del sistema de salud. Ahora queremos empezar. Va a ser más fácil si en nuestro entorno ya se sabe, y se entiende, aunque sea parcialmente, qué es eso de medir y mejorar la calidad. Sin embargo, en la mayoría de los centros, áreas, distritos y servicios de salud, es todavía una innovación en mayor o menor medida. Quiere esto decir que prestarle atención explícita a la calidad es un cambio, vamos a introducir una actividad que antes no se hacía o no se le prestaba suficiente atención, o quizá se hacía pero no se entendía bien por qué.

La introducción de innovaciones funciona mejor si se hace con una estrategia consciente y planificada, fruto del análisis de nuestra situación de partida concreta. En esta UT vamos a intentar que la práctica del compromiso con la calidad, y la realización misma de los trabajos prácticos que quedan en este curso, se encuadren y favorezcan un proyecto más amplio que tiene como final incorporar la gestión de la calidad a la rutina de nuestra actuación profesional. Vamos a revisar las condiciones generales de implantación de las actividades de gestión de calidad a nuestro sistema, para después describir algunos métodos sencillos de análisis de la situación existente en nuestro entorno más cercano, sobre la cual diseñar estrategias de implantación; vamos a ver también como podemos favorecer la implantación efectiva del compromiso con la calidad a través del análisis de los recursos humanos que nos rodean, y qué actividades concretas vamos a elegir para empezar, para lo cual propondremos la realización de un ciclo de mejora. En las siguientes UT, vamos a seguir paso a paso la realización de este ciclo de mejora.

2. MÉTODOS PARA EL DISEÑO DE ESTRATEGIAS. ANÁLISIS DE CAMPOS DE FUERZA

Una vez decididos a poner en marcha una innovación (en nuestro caso la evaluación y mejora de la calidad), la utilización de métodos estructurados puede facilitar el proceso de definir y acordar cómo tenemos que hacerlo. A veces, solamente discutiendo la información disponible puede ser suficiente, pero frecuentemente la visualización de las relaciones entre los diversos componentes del problema, a veces difíciles de ver hasta que no se intenta plasmarlos en un gráfico, acelera y da consistencia a las decisiones que se toman. No puede decirse, en cambio, que se pueda prescribir un método o métodos 100% infalibles; a medio plazo, lo correcto sería ir conociendo y aplicando diversos métodos para conocer de primera mano las ventajas e inconvenientes de cada uno. Algunos de estos métodos de utilidad para diseñar planes de acción (diagrama de afinidades, votación múltiple, diagrama de relaciones, matrices decisionales) van a ser descritos en otras UT como métodos para la identificación y priorización de oportunidades de mejora (UT 6) y para el diseño de intervenciones para mejorar un

En esta UT vamos a ver los métodos para diseñar una estrategia de implantación de actividades de GC propias de nuestro entorno, y propondremos la realización de un ciclo de mejora como primera actividad práctica.

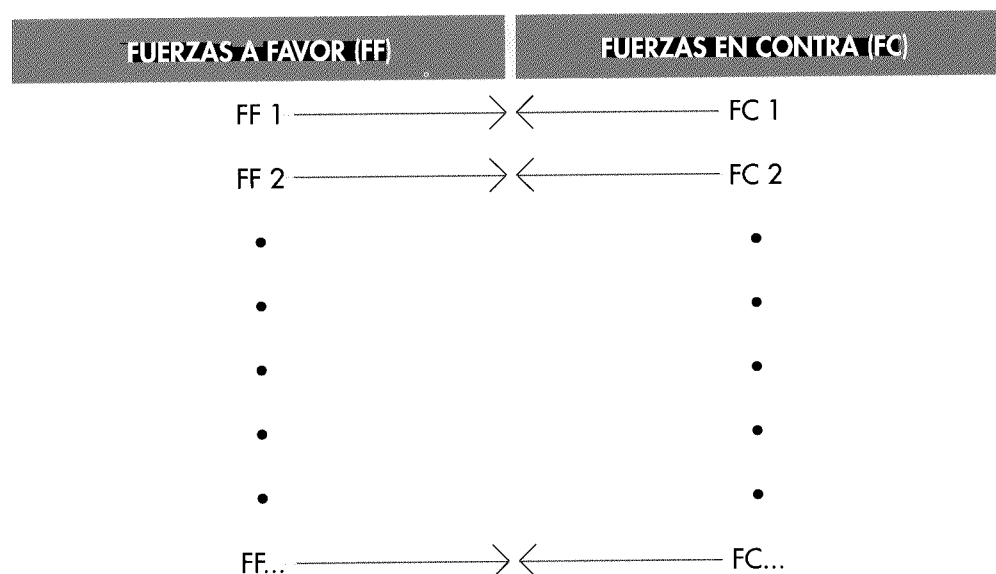
problema analizado (UT 12). En esta UT nos centraremos en la comprensión y utilización de uno de los métodos más sencillos y útiles: el análisis de campos de fuerza (ACF).

El ACF, tal como lo propuso Kurt Lewin, parte de asumir que para la introducción de un cambio o innovación (como puede ser iniciar actividades de evaluación y mejora de la calidad) es de gran ayuda empezar por examinar las circunstancias que contribuyen al *status quo*, de la situación de partida. Estas circunstancias se pueden expresar como fuerzas que se oponen al cambio que queremos introducir, y fuerzas que favorecen la implantación del cambio; cambio que ocurrirá cuando este último tipo de fuerzas supere a las que se oponen. Siguiendo este razonamiento, una estrategia de cambio puede ser enfocada reforzando las fuerzas a favor, reduciendo las fuerzas en contra, o combinando ambas cosas; sin embargo, según se desprende de la experiencia acumulada, la estrategia más efectiva suele ser la que actúa fundamentalmente para disminuir las fuerzas en contra.

¿Cómo se realiza un ACF? Podemos distinguir cinco pasos fundamentales:

1. Definición de la innovación a introducir, es decir, explicitar qué cambio queremos conseguir. Esta definición es un objetivo a alcanzar en el que debe concentrarse el grupo o, en el peor de los casos, el individuo que realiza el análisis.
2. Mediante lluvia de ideas, hacer un listado de factores a favor ("fuerzas favorables") del cambio a introducir, y de factores en contra ("fuerzas de resistencia"). Depurar los listados para evitar repeticiones.
3. Llevar el listado de factores a favor y en contra a un diagrama en el que las fuerzas a favor se colocan a la izquierda, y las contrarias a la derecha, unidas en ambos casos cada una de ellas a flechas dirigidas al centro del diagrama, que se divide en dos mitades con una línea central.

FIGURA 5.1. Esquema de diagrama para un análisis de campos de fuerza



De los métodos existentes para ayudar en el diseño de estrategias para implantar innovaciones, nos centraremos en la comprensión y utilización del análisis de campos de fuerza (ACF).

El ACF se basa en el conocimiento de los factores a favor y en contra de la innovación. La mejor estrategia es la que logra disminuir los factores en contra.

4. Seleccionar de forma realista unos pocos factores en contra sobre los que se pueda actuar para vencerlos o hacerlos disminuir; señalar también unos pocos factores a favor para apoyarse en ellos o potenciarlos si es posible.
5. Diseñar la estrategia de actuación correspondiente.

En algunos de los pasos del ACF pueden ser útiles otros métodos de trabajo grupal que veremos en otras UT, como por ejemplo el diagrama de afinidades (UT12) para la depuración de los listados y diseño del plan de acción (pasos 2 y 5), o cualquiera de los métodos de priorización (UT6) para seleccionar los factores sobre los que diseñar la estrategia (paso 4); pero también puede realizarse sin utilizar métodos adicionales.

Para facilitar su comprensión, vamos a ver dos ejemplos de ACF, uno sobre un problema clínico: ¿qué estrategia a seguir para inducir a la pérdida de peso? (Figura 5.2); y otro sobre un problema de gestión: ¿cómo llevar a cabo la concentración de las tareas de procesamiento de textos? (Figura 5.3); curiosamente el ACF clínico se explica en un manual de bolsillo sobre herramientas para la mejora de la calidad de origen industrial, mientras que el ejemplo de gestión es un caso real ocurrido en un complejo universitario de centros sanitarios, lo cual da una idea de la aplicación flexible y universal de esta herramienta.

No es preciso subrayar que el ACF es un método factible en todo tipo de organizaciones y situaciones, vamos a revisar el resultado de un caso concreto.

3. ANÁLISIS DE CAMPOS DE FUERZA EN NUESTRO CONTEXTO

La técnica de ACF la hemos utilizado de forma habitual en los seminarios introductorios sobre métodos de evaluación y mejora de la calidad. La innovación para la que proponíamos realizar el ACF era invariablemente implantar actividades de evaluación de la calidad en los lugares de trabajo de los participantes en los seminarios; en la mayoría de los casos era como dar respuesta a la pregunta ¿cómo empezar?

A lo largo de los años, el ACF resultante ha ido cambiando, a medida que el tema de la calidad se ha ido haciendo más familiar, al menos de oídas. Sin embargo, se mantienen una serie de características comunes. La primera, es que del lado de los factores a favor, siempre ha aparecido el interés de los profesionales por hacer bien su trabajo; la segunda es que del lado de los factores en contra figuran invariablemente la falta de formación o conocimientos sobre el tema, y toda una serie de circunstancias unidas a la inexistencia de experiencias que hayan disipado toda una serie de temores y prejuicios sobre lo que puede ser realizar actividades de evaluación y mejora de la calidad en los centros sanitarios.

Merece la pena resumir lo que sería una estrategia común y ajustada a nuestra realidad, acorde con lo que se evidencia en los ACF realizados. Esta estrategia, empleada aquí y en otros países, se basa en la *formación*, la *incentivación profesional* y la consecución de *experiencias reales* para ir venciendo la resistencia al cambio basada en el desconocimiento. Este último punto es de gran relevancia, como se desprende de los esquemas de manejo del cambio, y tal como ha sido comprobado en la práctica. Los esquemas conceptuales de la "demografía del cambio" nos han sido especialmente útiles para ir introduciendo con éxito en centros asistenciales la innovación que supone medir y mejorar la calidad.

Para la realización de algunos pasos del ACF pueden ser útiles otros métodos de trabajo grupal que revisaremos en otras UT.

FIGURA 5.2. Ejemplo de ACF para un problema de salud

OBJETIVO: diseñar una estrategia para perder peso

| FUERZAS A FAVOR | | FUERZAS EN CONTRA |
|-------------------------------------|-----|--|
| AMENAZA A LA SALUD | → ← | FALTA DE TIEMPO |
| OBSESIÓN CULTURAL DE SER DELGADO | → ← | CARACTERÍSTICAS GENÉTICAS |
| GRAN CANTIDAD DE MODELOS DELGADOS | → ← | FAMILIA Y AMIGOS POCO COMPRENSIVOS |
| VERGÜENZA | → ← | FALTA DE DINERO PARA HACER EJERCICIO FÍSICO |
| AUTOIMAGEN NEGATIVA | → ← | DESÁNIMO |
| ACTITUD POSITIVA HACIA EL EJERCICIO | → ← | MALOS CONSEJOS |
| LA ROPA QUE QUEDA CHICA | → ← | MUCHOS AÑOS DE MALOS HÁBITOS ALIMENTICIOS |
| | | ← CANTIDAD DE AZÚCAR Y GRASA EN LAS COMIDAS HABITUALES |

ESTRATEGIA:

- Ayudar a encontrar tiempo para hacer ejercicio.
- Reconducir hábitos alimenticios.
- Educar a la familia
- +
- Reforzar imagen positiva hacia el ejercicio.
- Reforzar autoestima.

Adaptado de: Manual de herramientas básicas para el análisis de datos. Guía de bolsillo con las herramientas para el mejoramiento continuo. GOAL/QPC, Mcthuen MA. 1990.

El ACF para diseñar estrategias de implantación de actividades de GC nos indica la importancia de la ética e incentivación profesional como factor a favor, y la formación y realización de experiencias reales como acciones para disminuir los factores en contra.

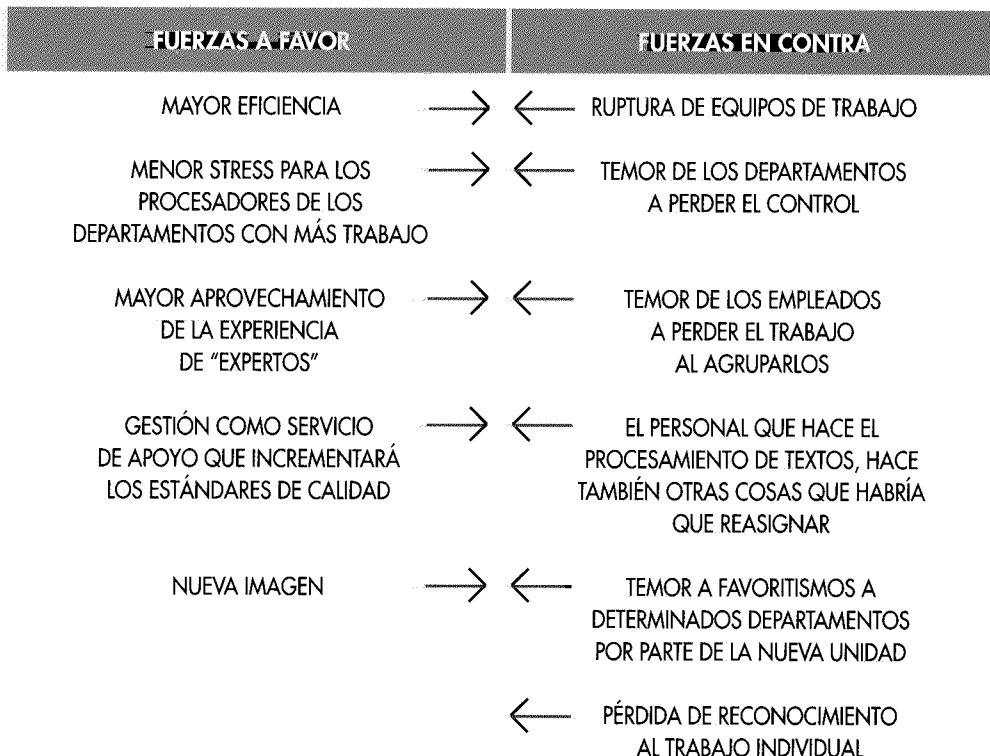
El análisis "demográfico" de la resistencia al cambio confiere un interés primordial al componente de los recursos humanos. Vamos a ver en qué consiste, a la vez que lo ilustramos con el caso de la introducción de actividades de evaluación de calidad en centros asistenciales.

4. INTRODUCCIÓN DE INNOVACIONES. MASA CRÍTICA Y DEMOGRAFÍA DEL CAMBIO

Cambiar es hacer las cosas de forma diferente; y para que ello ocurra según nuestros deseos, la literatura sobre gestión de recursos humanos y psicología de las organizaciones nos brinda múltiples modelos y consideraciones, cuyo denominador común es la necesidad de tener una estrategia determinada y planificada y prestar atención en todo el proceso al factor humano, que es en definitiva el que puede hacer progresar o fracasar la estrategia.

FIGURA 5.3. Ejemplo de ACF para un problema de gestión

OBJETIVO: Diseñar estrategia para implementar una unidad centralizada para el procedimiento de textos de todos los departamentos de un complejo asistencial universitario


ESTRATEGIA:

- Favorecer actividades para que se cree el sentimiento de equipo en la nueva unidad.
- Hacer explícita y pública la política de reubicación del trabajo, sin que nadie vaya a ser despedido.
- Construir un mecanismo para asegurarse que los trabajos se harán por orden y a tiempo.
- Asignar a cada persona de la nueva unidad un grupo determinado de departamentos como clientes preferentes.
- Considerar la necesidad de personal de apoyo para los departamentos que queden en precario.
- +
- Explicar a los departamentos cómo va a funcionar la nueva unidad, enfatizando:
 - Mejores estándares de calidad.
 - Circuito y distribución del trabajo para asegurar que se cumplen los plazos.
 - Plan de formación específico para aumentar las habilidades del personal de la unidad en el manejo de procesamiento de textos.

Adaptado de: Leebow W, Ersoz CJ: The health care manager's guide to continuous quality improvement, AHA, Chicago 1991.

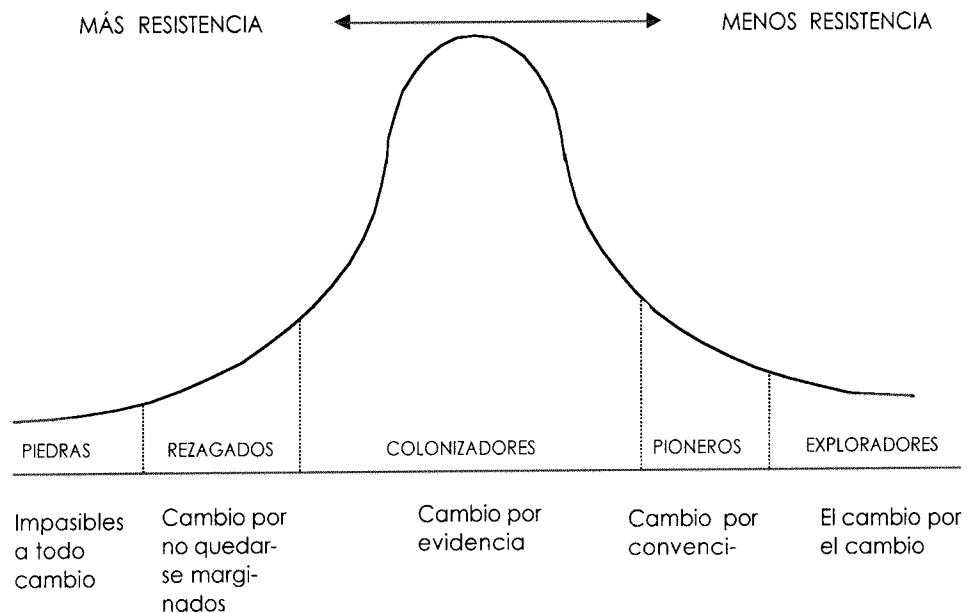
El análisis de la resistencia al cambio en el grupo en que queremos que se introduzca la innovación, es de gran utilidad para dirigir los esfuerzos de nuestra estrategia

Entender y gestionar *la resistencia al cambio* es el término más común que recoge esta atención preferente que hay que prestar a las personas.

De los muchos esquemas que ayudan a entender y gestionar la resistencia al cambio, la teoría de la "masa crítica" y entendimiento de la "demografía" del cambio pueden ser de especial utilidad. Tienen además la virtud de tener sentido común, ser lógicos y fáciles de entender y aplicar. Vamos a ver en qué consisten.

Se parte de un primer supuesto: frente a cualquier innovación, la resistencia a su implantación no se distribuye de forma homogénea entre todos los componentes del grupo que queremos que cambie; hay una cierta "tipología" de la resistencia al cambio que es conveniente conocer, porque son diferentes no sólo en cuanto a su grado de resistencia, sino también en relación a los argumentos que cada tipo entiende y los mecanismos que hay que emplear para vencer su resistencia. La Figura 5.4 contiene uno de los esquemas que representa esta teoría. En él se distinguen cinco grupos que -de menor a mayor resistencia- se denominan "exploradores", "pioneros", "colonizadores", "rezagados" y "pedras" (estos últimos son llamados "urbanitas" en el esquema original, por las razones que ahora veremos); esta tipología evoca la actitud de los americanos del este hacia la colonización del "salvaje oeste". Veamos por qué.

FIGURA 5.4. Tipología y distribución de frecuencias de la resistencia al cambio



El grupo de "exploradores" no opone ninguna resistencia, son aventureros que van a abrir caminos, gente que le gusta el cambio por el cambio, la novedad por la novedad; son importantes porque abren brecha, pero no fundamentales a medio plazo porque, entre otras cosas, la posibilidad de aventurarse con otro "cambio" puede hacerles perder el interés por el anterior. De estos en general hay pocos en cualquier organización, pero los hay.

El siguiente grupo son los "pioneros", su resistencia es algo mayor, pero la vence la razón, el convencimiento de que el cambio es prometedor, que cabe esperar cosas buenas de él. Son los que decidieron irse al lejano Oeste a buscar tierra en donde establecerse con sus familias, convencidos de que era posible y bueno aunque nadie lo había hecho antes. Este grupo, normalmente reducido, es de extraordinaria importancia para introducir el cambio; una vez convencidos van a "experimentar" el cambio de forma consciente y para quedarse, sin que necesiten de experiencias previas en su entorno inmediato, son innovadores conscientes y convencidos. ¿Por qué son tan importantes?, porque sin ellos sería muy difícil vencer la resistencia del siguiente grupo, los "colonizadores", que además suelen ser mayoría.

Los "colonizadores" necesitan *ver para creer*, su resistencia al cambio sólo se vence si se les muestra que es posible y bueno; necesitan de los ejemplos de los "pioneros" para decidir moverse a los nuevos asentamientos. Naturalmente, no son un grupo uniforme y es muy probable encontrar diversos grados de resistencia según se acerquen más al tipo "pionero" o al siguiente grupo. Pero lo fundamental es que este grupo necesita "hechos" no argumentos para vencer su resistencia, sólo así van a imitar, e incluso puede que superar con sus experiencias lo que ya saben que es positivo.

El extremo más resistente de los "colonizadores" son los "rezagados". Este grupo precisa no sólo de hechos sino también de garantías sobre la ausencia de riesgos que implica el cambio. Son los que no van a ir al oeste hasta que las ciudades estén funcionando y los indios en las reservas. Van cuando ya queda hasta feo que se nieguen a conocer los nuevos territorios.

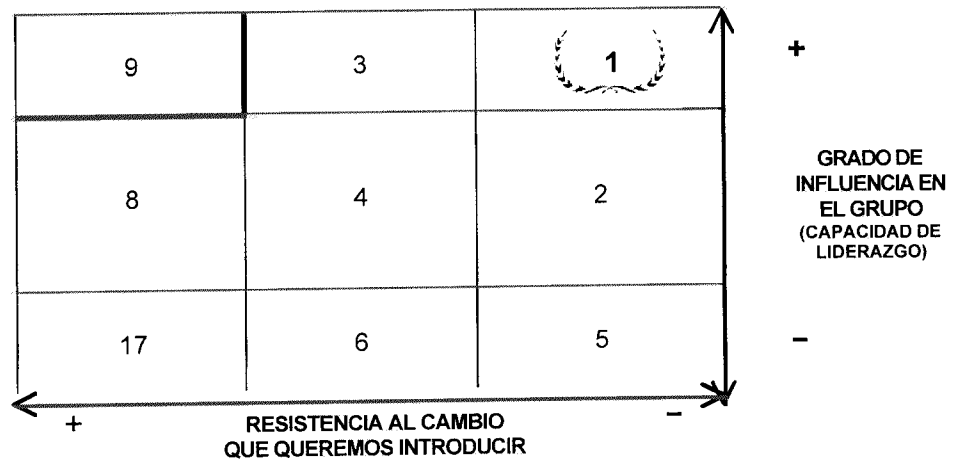
El último grupo, el más resistente al cambio, son los "urbanitas", aquellos para los cuales la civilización nunca va a existir fuera de las cómodas ciudades de la Costa Este. A éstos no hay quien los mueva: ni nada ni nadie. Por eso les viene bien el nombre de "piedras". En teoría son pocos: sin embargo representan un grave problema si ocupan puestos directivos o logran tener influencia de alguna manera sobre otros miembros de la organización.

Según este esquema, nuestros esfuerzos para convencer, a través de formación o por cualquier otra vía, deberían centrarse prioritariamente en los pioneros. Sin embargo, el conocer el grado y tipo de resistencia al cambio no es suficiente. Otra variable de interés para el diseño de nuestra estrategia es el grado de influencia que cada individuo tenga en el resto del grupo, es decir, su capacidad de liderazgo. Combinando resistencia al cambio y capacidad de liderazgo podemos identificar una serie de grupos; si por ejemplo, tal como se ha representado en la Figura 5.5, identificamos tres categorías en las dos variables (resistencia al cambio y capacidad de liderazgo), sería un total de nueve grupos, de los cuales los mejores aliados para la innovación serían los del grupo 1 (a favor del cambio y con un alto grado de influencia en los demás), y los peores el grupo 9 (máxima resistencia al cambio y alto grado de influencia en los demás). Si somos capaces de identificar cada grupo, nuestra estrategia debe incluir una búsqueda lógica de aliados (grupos 1 y 2) para la implantación de las primeras experiencias: lo cual no quiere decir que haya que cerrarse a nadie, sino que es conveniente saber y utilizar de forma priorizada nuestras energías.

La identificación de los pioneros es importante para que se realicen las primeras experiencias de GC.

El análisis de la "demografía" del cambio ayuda a buscar los mejores aliados para la implantación de las actividades de GC.

FIGURA 5.5. Resistencia al cambio y capacidad de liderazgo: análisis "demográfico" para buscar aliado y diseñar estrategias



- Los grupos están numerados por orden de preferencia, combinando importancia y facilidad, como aliados para introducir el cambio.
- Los grupos 1 y 2 son los mejores aliados para convencer a los 3 y 4.

Resumiendo lo expuesto hasta ahora, para el inicio de actividades de GC internas y voluntarias, nuestro ACF para diseñar la estrategia de implantación nos ha sugerido la ética profesional y su consiguiente incentivación como telón de fondo a favor, y la importancia de la formación y la consecución de experiencias prácticas para ir venciendo resistencias; el análisis "demográfico" de la resistencia al cambio, nos ha indicado en quienes nos deberíamos apoyar para acelerar la implantación de la GC. Pero no hemos decidido aún qué actividades acometer, cuál de los varios puntos de entrada a la mejora de la calidad es el que más nos conviene. Vamos a considerarlo a continuación.

5. LAS PRIMERAS ACTIVIDADES PARA MEJORAR LA CALIDAD ¿DISEÑAR, MONITORIZAR O REALIZAR CICLOS DE MEJORA?

Como vimos en la UT3 podemos identificar tres grandes grupos de actividades en los programas de GC, que responden a otros tantos puntos de partida: diseño de la calidad, monitorización y ciclos de mejora. Todos ellos terminan estando interrelacionados, pero en teoría es posible comenzar nuestras actividades por cualquiera de los tres. Sin embargo, ya hemos adelantado que lo que vamos a proponer es la realización de un ciclo de mejora. ¿Por qué? Hay tres razones: la primera es que, dentro de su posible complejidad, es la actividad más fácil de realizar y comprender; la segunda es que como actividad de aprendizaje es muy completa y básica, de modo que abona el terreno para comprender y practicar mejor las otras dos; y la tercera es que posibilita la obtención de resultados a un relativo corto plazo, de modo que motiva a continuar.

Elegimos comenzar con un ciclo de mejora por su facilidad de comprensión, por su valor didáctico para la GC y por su efecto motivador si se realiza con éxito.

Schein, uno de los grandes nombres del desarrollo organizacional, resume en siete puntos sus recomendaciones estratégicas para introducir cambios; los dos que se refieren a actividades concretas son: "comenzar como si fuese un experimento" y "empezar con pasos poco ambiciosos y adoptar una estrategia de

implantación progresiva". Aplicado a nuestro tema, comenzar por actividades de diseño, aunque es tremendamente atractivo e incluso hay una tendencia actual a proponerlo como primer paso, no sería una buena opción al revestir en general mayor complejidad que los otros dos. También es muy frecuente que se propugne comenzar por elaborar un listado de indicadores y monitorizar. Sin embargo, optar por la realización de un ciclo de mejora es un comienzo más simple, fácil de enfocar, y a la vez más enriquecedor y didáctico. Sin embargo, no es descartable ni excluyente comenzar *identificando problemas* a través de la medición de indicadores bien seleccionados. Todo depende del grado de madurez, compromiso y recursos iniciales para la GC: han de ser mayores para comenzar con monitorización. Ha de quedar claro además que monitorizar no es evaluar para mejorar. Monitorizar es asegurarse de que nos mantenemos a unos niveles de calidad aceptables y preestablecidos, o hacer despistaje de existencia de problemas, mientras que evaluar un ciclo de mejora es intentar buscar las causas y lograr la solución de esos problemas. Los ciclos de mejora son la esencia *imprescindible* de cualquier programa de GC, tienen sentido por sí mismos. La monitorización es necesariamente una actividad complementaria que aunque de gran relevancia en los programas estables, es perfectamente prescindible al inicio de los mismos.

6. PASOS DE UN CICLO DE MEJORA

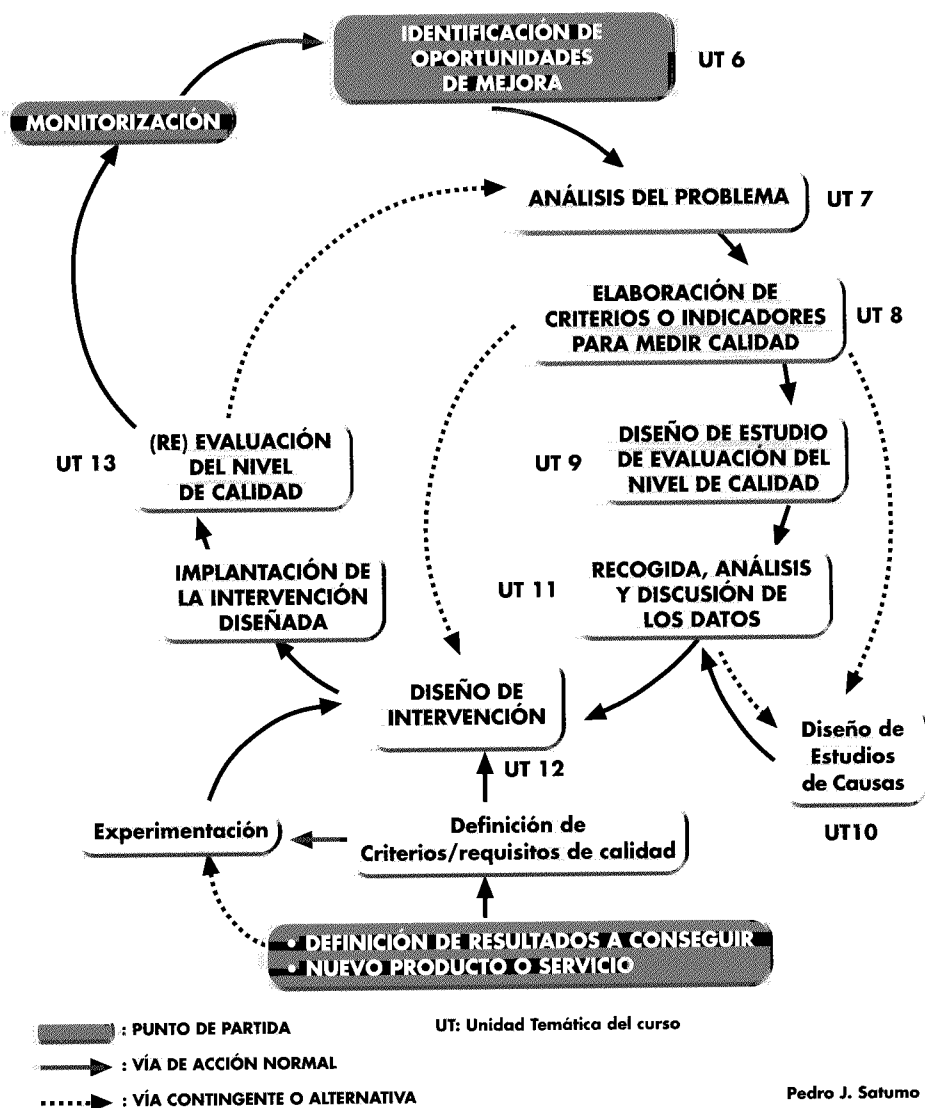
Vimos en la UT 3 que el ciclo de mejora o ciclo evaluativo comienza con la identificación de una oportunidad de mejora o problema de calidad, siendo el objetivo de las actividades que se generan a partir de ahí el solucionar el problema identificado o, lo que es lo mismo, aprovechar la oportunidad para mejorar. Para ello podemos establecer una serie de pasos o fases, para cuya realización nos vamos a ayudar de determinadas herramientas y métodos, que iremos detallando en cada UT a la vez que avanzamos en el curso, de acuerdo con el esquema representado en la Figura 5.6.

- **Paso 1:** *Identificación y priorización de la oportunidad de mejora (UT6)*. De los muchos métodos utilizables, proponemos unos pocos, de los que destacamos la lluvia de ideas, la técnica de grupo nominal, y las matrices decisionales, aparte de mencionar otros métodos muy atractivos, pero que precisan de datos y una determinada infraestructura para realizarlos. En este paso debemos elegir cuál es el servicio o el aspecto del mismo que queremos mejorar.
- **Paso 2:** *Análisis del problema de calidad (UT7)*. Crucial para decidir qué hacer a continuación. Veremos la utilidad de diversos métodos como el diagrama de causa y efecto, flujograma, histograma, estratificación y diagrama de Pareto, utilizables con y sin datos previos sobre el problema. De este paso se derivan tres posibles cursos de acción: diseño de intervención, diseño de estudio de nivel de calidad y diseño de estudio de causas hipotéticas. En estos dos últimos se necesita previamente tener claro cuáles son los criterios que nos miden la existencia o no del problema.
- **Paso 3:** *Construcción de criterios para evaluar calidad (UT8)*. En este paso deben identificarse criterios válidos y fiables para medir la calidad del servicio o aspecto del servicio en estudio. Junto a las indicaciones formales para valorar la validez, veremos un esquema de fácil realización para medir la fiabilidad.

La monitorización es un contenido más complejo y necesita además de ciclos de mejora para tener utilidad en la GC.

Los ciclos de mejora se realizan siguiendo una serie de pasos concretos, que iremos viendo en las siguientes UT, y que se benefician cada uno de ellos de métodos y herramientas concretos que serán objeto de revisión en la UT correspondiente.

FIGURA 5.6. Los pasos del ciclo de mejora y su correspondencia con las Unidades Temáticas



- **Paso 4:** *Diseño de estudio de nivel de calidad o de causas hipotéticas* (UT9 y 10). Vamos a proponer y detallar los componentes para un estudio de nivel de calidad, aunque queda abierta la posibilidad de hacer estudios de causas hipotéticas, incluso después de realizado un estudio de nivel de calidad. En este paso hay que utilizar unos conocimientos básicos de estadística para beneficiarse de las ventajas de poder evaluar con muestras.
- **Paso 5:** *Análisis y presentación de los datos de la evaluación* (UT11). Tras recoger los datos de acuerdo con el diseño realizado, veremos la forma más eficiente de presentarlos y analizarlos. Los datos estimados los daremos con intervalos de confianza si hemos evaluado una muestra, y veremos la forma de realizar un análisis gráfico con diagrama de Pareto y otras representaciones gráficas, de forma que nos resulte más evidente sobre qué incidir para mejorar.

- **Paso 6:** *Diseño de intervenciones para mejorar (UT12)*. La discusión de los datos de la evaluación puede sugerir por sí sola qué hacer para mejorar, sin embargo es conveniente conocer una serie de herramientas que pueden ser útiles para decisiones grupales sobre diseño de planes de acción para la mejora. Entre estas herramientas está el análisis de campos de fuerza, que hemos visto en esta UT, y el diagrama de afinidades y votación múltiple que serán revisados en la UT12 para ayudar en la realización de este paso.
- **Paso 7:** *Implementación de la intervención diseñada*. En esencia cumplimiento del plan de acción resultante del Paso 6.
- **Paso 8:** *Reevaluación y documentación de la mejora conseguida (UT13)*. Es el paso que cierra el ciclo de mejora si hemos tenido éxito con la intervención diseñada e implementada. Para discutir la mejora veremos como efectuar la comprobación estadística de que la diferencia es significativa, si hemos utilizado muestreo, y una forma sencilla y muy completa de visualizarlo como es un diagrama de Pareto antes-después.

Tras la reevaluación podemos decidir monitorizar, reanalizar el problema si no hemos conseguido mejorar, o emprender otros ciclos de mejora sobre un aspecto o servicio diferente.

BIBLIOGRAFÍA

- Saturno PJ. Estrategias para la puesta en marcha de actividades de evaluación y mejora de la calidad asistencial en Atención Primaria (I). Consideraciones generales. Centro de salud 1993; 1(9): 586-592.
- Leebow W, Ersoz CJ. The health care manager's guide to continuous quality improvement. Chicago: AHA; 1991.
- Robbins SP. Organizational behavior. Concepts, controversies and applications. Prentice Hall, Englewood Cliffs; 1989.
- Manual de herramientas básicas para el análisis de datos. Guía de bolsillo con las herramientas para el mejoramiento continuo. Methuen: GOAL/QPC; 1990.
- Scholtes PR, et al. The team handbook. Madison: Join et Associates; 1988.

IDENTIFICACIÓN Y PRIORIZACIÓN DE OPORTUNIDADES DE MEJORA

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

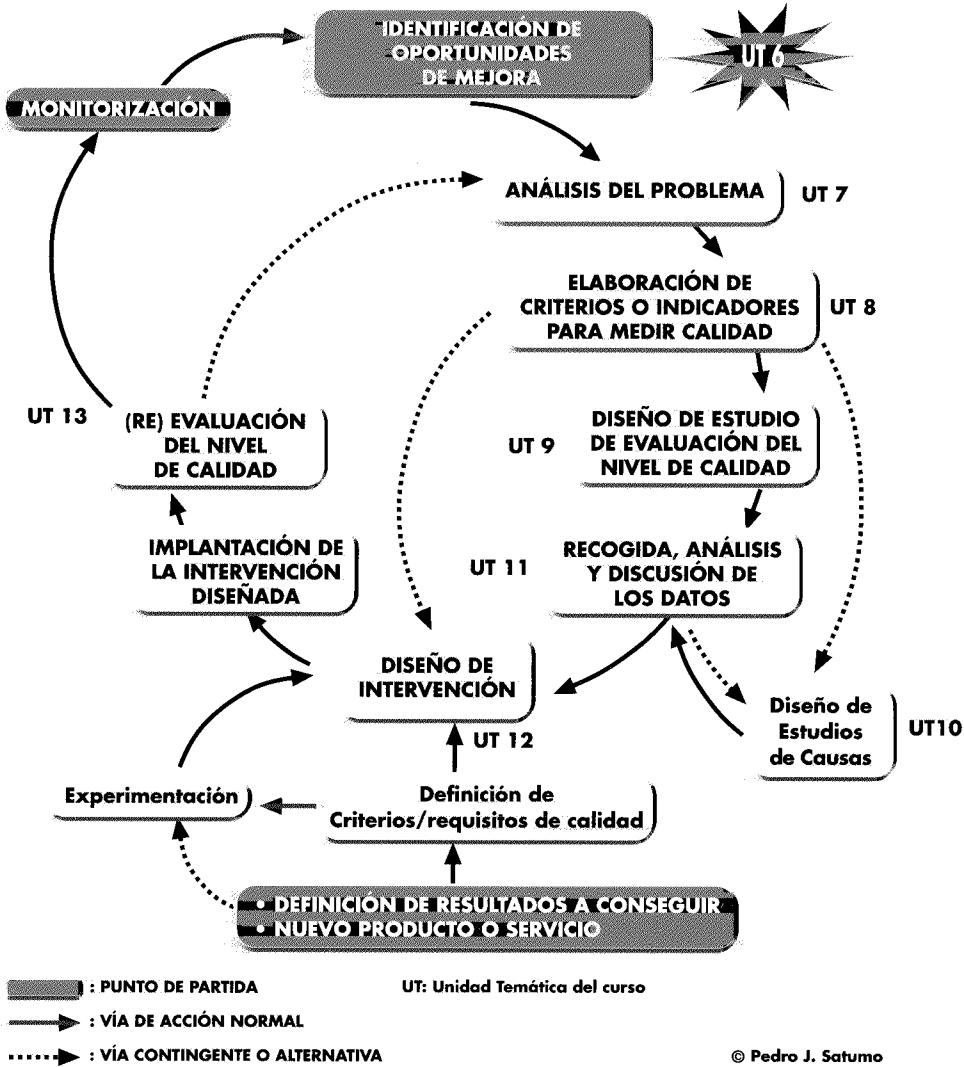
Esta UT versa sobre el primer paso de los ciclos de mejora: la selección y descripción operativa del problema u oportunidad de mejora que vamos a someter a evaluación. Se hace una revisión de diversos métodos utilizables, ilustrándolos con ejemplos, y se insta al participante que seleccione su objetivo de estudio. Para facilitar la realización y seguimiento del ciclo de mejora que se inicia con esta UT se aconseja utilizar un cartel narrativo.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. Concepto de problema de calidad u oportunidad de mejora.
3. Identificación de oportunidades de mejora.
4. Normas generales para la selección de una oportunidad de mejora
5. Métodos de identificación de oportunidades de mejora. Consideraciones generales.
6. Métodos para la identificación de oportunidades de mejora que no precisan datos
7. Métodos para la identificación de oportunidades de mejora utilizando datos
8. El cartel narrativo: un apoyo para la realización y el seguimiento del ciclo de mejora.

OBJETIVOS ESPECÍFICOS

1. Describir diferentes métodos aplicables para la identificación de problemas de calidad u oportunidades de mejora.
2. Escoger la metodología de identificación de problemas más adecuada a su entorno de trabajo.
3. Resumir los resultados de la lluvia de ideas de un grupo de trabajo.
4. Dirigir la realización de un Grupo Nominal.
5. Distinguir diferentes métodos de priorización de problemas.
6. Aplicar matrices de decisión a la priorización de problemas.
7. Establecer prioridades basándose en datos.
8. Identificar en qué formas puede tenerse en cuenta al usuario para la identificación de oportunidades de mejora.
9. Definir qué es un problema de calidad u oportunidad de mejora.
10. Describir un problema de calidad en términos operativos, concretos.
11. Diferenciar problemas genéricos y específicos.
12. Diferenciar problemas estructurales de problemas relativos a las actividades y resultados.
13. Delimitar de forma estructurada y secuencial diversas fases en el proceso de atención.
14. Escoger la oportunidad de mejora que va a ser sometida a evaluación.



"Un defecto puede ser un tesoro. El descubrimiento de la imperfección nos brinda la posibilidad de reducir el espacio que nos separa de la excelencia".

(citado por Berwick DM)

1. INTRODUCCIÓN

Cualquier actividad de evaluación de la calidad debe comenzar por plantearse ¿Qué es lo que se quiere mejorar? De ésta forma iniciamos el ciclo de evaluación con la identificación de la oportunidad de mejora. Probablemente ésta sea la fase más importante. Una elección acertada será crucial para que tengamos éxito en la evaluación. De hecho, la decisión que tomemos en este momento condicionará el desarrollo metodológico de las siguientes etapas del ciclo (análisis del problema, evaluación e intervención). Si la elección no es acertada las siguientes fases no se podrán desarrollar con la agilidad requerida y sin tropiezos. Peor aún, puede ocurrir, en un caso extremo, que el desarrollo del ciclo de mejora entre en una vía sin salida.

En esta unidad temática se inicia el ciclo de evaluación, mediante la identificación de la oportunidad de mejora, y se revisan diversos métodos para llevarla a cabo.

Por tanto es conveniente, *sobre todo en las primeras experiencias*, que para identificar la oportunidad de mejora reflexionemos sobre las implicaciones de su elección, priorizando las que sean más sencillas en su desarrollo a lo largo del ciclo de mejora y, sobre todo, en las que vislumbremos la factibilidad de su solución.

Vimos en la anterior UT que, de forma refleja, es posible que la idea de mejorar en conjunto la calidad del centro pueda parecernos atractiva. Sin embargo, una evaluación global puede presentar numerosas dificultades prácticas, especialmente cuando se aplica a una institución en la que se llevan a cabo numerosos procesos asistenciales y administrativos para resolver condiciones clínicas con requisitos de calidad muy diferenciados, como obviamente ocurre con la mayoría de los centros sanitarios.

Dado que la revisión simultánea de todos los servicios prestados es generalmente impracticable como primer paso e incluso es desaconsejable afrontar varias oportunidades de mejora a la vez, es preferible que seleccionemos solamente una. Algo que queremos hacer especialmente bien o que es inadmisibile que se haga mal y en la que a la vez sospechamos que hay una diferencia importante entre la atención proporcionada y la que sería razonable prestar. Las ventajas de esta manera de proceder son numerosas. Focalizamos la evaluación de la calidad en un aspecto concreto y tangible, facilitando el compromiso de los profesionales con la mejora de la calidad. Además, en ocasiones, la delimitación de una oportunidad de mejora será suficiente para que se promuevan los cambios que conducen a la resolución del problema, sobre todo cuando sus causas son evidentes.

En esta UT revisaremos una serie de herramientas conducentes a delimitar y priorizar oportunidades de mejora. Veremos que para la identificación de las áreas problema tenemos diversas técnicas. La decisión de utilizar un método determinado debe condicionarse a las características de cada centro y principalmente a la sencillez y factibilidad de su empleo. En este sentido, vamos a subrayar el empleo de la lluvia de ideas, la técnica del grupo nominal y la matriz decisional. Esta elección se debe a que estos métodos no requieren un contexto especial de aplicación (no necesitan fuentes de datos adicionales), son fáciles de comprender e involucran a los que quieren evaluar y mejorar la calidad. Son métodos versátiles que pueden adaptarse a la mayoría de las situaciones.

2. CONCEPTO DE PROBLEMA DE CALIDAD U OPORTUNIDAD DE MEJORA

Podemos proponer una oportunidad de mejora si percibimos o intuimos que alguna de las características del servicio que ofrecemos no está como creemos que debe de estar. Desde el punto de vista cuantitativo la podemos definir como una diferencia, que potencialmente se puede reducir, entre lo que es razonable que exista (estructura), se haga (proceso) u ocurra (resultado), con lo que existe, se hace o se proporciona en la práctica habitual. Por ejemplo podemos identificar una oportunidad de mejora cuando percibimos que la captación de hipertensos es muy baja en relación a lo que podríamos esperar, sospechamos que el seguimiento de las dislipemias se realiza de forma esporádica, intuimos que se utilizan antibióticos en la mayoría de los pacientes diagnosticados de infección

respiratoria alta, vemos que es demasiado frecuente el que no esté disponible la historia clínica a la hora de la consulta, etc. Una oportunidad de mejora es una *divergencia entre la situación actual y la que queremos que exista*. ¿Cómo proceder para identificarlas en nuestro entorno? Hay una serie de procedimientos que sirven para hacerlas explícitas, y que veremos a continuación.

3. IDENTIFICACIÓN DE OPORTUNIDADES DE MEJORA

Es evidente que para determinar una posible divergencia entre la atención proporcionada y la que se debería proporcionar, tenemos que haber establecido previamente un marco de referencia, en el que se refleje de forma explícita y consciente las características que a nuestro juicio y en función de las necesidades del usuario debería reunir la atención sanitaria para cumplir con la finalidad para la cual ha sido concebida y con el menor coste posible. Este marco de referencia corresponde al concepto de calidad de los servicios sanitarios que ya habremos establecido previamente adoptando o adaptando las diversas definiciones posibles tal como vimos en la UT 1.

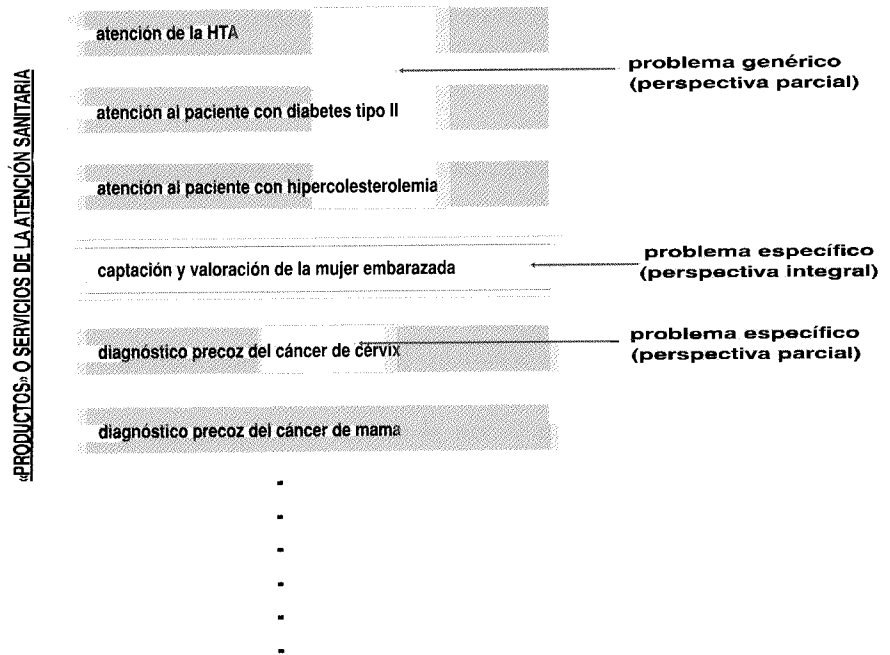
Sin embargo, este marco de referencia, puede resultar aún excesivamente ambiguo y poco operativo para concretar una oportunidad de mejora. Es preferible que se reflexione acerca de la gran variedad de "productos" o servicios prestados al usuario y definir a continuación sus respectivos requisitos de calidad. Por ejemplo, la "atención de pacientes crónicos con hipercolesterolemia" es un producto asistencial con unos requisitos de calidad que le confieren una identidad propia y que le diferencian de otros "productos asistenciales" como puede ser el caso de la "atención al paciente con hipertensión arterial" o "la gestión de las citas".

Una vez llegamos a este nivel de especificación debe plantearse para uno o varios de los servicios que nos preocupen especialmente ¿Qué tipo de atención es razonable que se proporcione a nuestros pacientes? ¿Qué tipo de atención estamos proporcionando? y ¿Qué diferencia suponemos o percibimos que hay entre ambas? Ésta aproximación a los elementos básicos de la atención sanitaria es probablemente más adecuada para apreciar las posibles divergencias entre lo que es y lo que debería ser y permitirá delimitar con más facilidad diversas oportunidades de mejora de la calidad asistencial que pueden coincidir completamente -perspectiva integral- o en parte -perspectiva parcial- con uno o con varios servicios de la actividad asistencial. Si atañe a un solo servicio le vamos a llamar problema *específico*, mientras que si afecta a más de un servicio le llamaremos problema *genérico* (Figura 6.1). Desde esta perspectiva es más patente si, por ejemplo, el control del paciente con hipertensión arterial, el control del paciente con hipercolesterolemia (ambos enfoques específicos) o la atención de pacientes en la consulta programada (enfoque genérico que engloba parcialmente, entre otros, a los dos servicios anteriores) pueden ser áreas potenciales de mejora.

La oportunidad de mejora es cualquier situación en la que se observan diferencias, que se pueden reducir, entre lo óptimo y la realidad.

Las oportunidades de mejora se identifican más fácilmente si pensamos en servicios concretos y sus requisitos de calidad.

FIGURA 6.1. Determinación de oportunidades de mejora a partir del listado de los "productos" o servicios de la atención sanitaria*



**(en sombreado: áreas potenciales de mejora identificadas)*

Con estas reflexiones, espontáneamente o con métodos estructurados, es posible identificar un listado normalmente extenso de oportunidades de mejora; sin embargo para empezar un ciclo de mejora vamos a seleccionar una, y es conveniente que prestemos atención a una serie de normas generales que nos pueden ayudar a que la selección sea adecuada.

4. NORMAS GENERALES PARA LA SELECCIÓN DE UNA OPORTUNIDAD DE MEJORA

Cualquier oportunidad de mejora es válida desde el punto de vista operativo siempre que su desarrollo metodológico sea posible (puede analizarse y medirse sin dificultad) y conduzca a su resolución. Para valorar esta validez debemos considerar las características, tanto extrínsecas (factibilidad de su resolución) como intrínsecas (características del problema), que pueden determinar la mayor o menor garantía de éxito en la mejora de la calidad, y en base a ello priorizar y seleccionar una.

4.1. FACTIBILIDAD DE LA RESOLUCIÓN DEL PROBLEMA DE CALIDAD

La identificación y selección de una oportunidad de mejora no debe consistir exclusivamente en revelar o denunciar una atención incorrecta sino tener como objetivo corregirla. Por tanto, es básico que se valore, al mismo tiempo que se identifica la oportunidad de mejora, las posibilidades de éxito de esa evaluación, en el sentido de que sea factible la mejora.

Cualquier problema de calidad es válido para iniciar un ciclo de mejora siempre que se pueda cuantificar, analizar y resolver.

En primer lugar tendríamos que averiguar si existe una relación evidente entre las posibles soluciones y nuestra capacidad de intervención. En la mayoría de las ocasiones la capacidad de resolver los problemas relacionados con aspectos estructurales de tipo cuantitativo (p.e. falta de personal, recursos insuficientes, horario inadecuado, etc.) dependen en gran medida de organismos o de niveles jerárquicos distintos a los que han identificado el problema. Por tanto, en estos casos el margen de maniobra de resolución del problema será escaso y la identificación del problema se limitará con mucha probabilidad a una simple denuncia de la situación.

Uno de los errores más comunes en los primeros pasos en la mejora de la calidad consiste en plantear la oportunidad de mejora en términos de "falta de..."; sin embargo la ausencia de algo no es evaluable; puede ser una *causa* de problemas o una *solución* a un problema, pero nunca el problema a evaluar. Es recomendable que descartemos como foco de evaluación los problemas de estructura en términos de cantidad de recursos humanos o de equipamiento.

¿Quiere esto decir que sólo debemos someter a ciclos de mejora a los problemas para los que se intuye una fácil solución? A medio y largo plazo probablemente no, pero a corto plazo, para las primeras experiencias, rotundamente sí. Es una cuestión de estrategia de implantación de la innovación que supone prestarle una atención explícita y cuantificada a la mejora de la calidad.

Los problemas de calidad no deben formularse en términos de "falta de..."

4.2. CARACTERÍSTICAS INTRÍNSECAS DE LA OPORTUNIDAD DE MEJORA E IMPLICACIONES METODOLÓGICAS

Desde un punto de vista operativo, además de la visión integral o parcial que ya hemos visto, es útil que clasifiquemos las oportunidades de mejora en función de dos características intrínsecas (Figura 6.2): (i) en función de que coincidan con un solo "producto" o servicio asistencial -problema específico- o se refieran a varios -problema genérico-, tal como hemos apuntado más arriba; y (ii) según afecten a un componente determinado en la prestación de los servicios -problemas de estructura *versus* problemas de proceso/resultado-.

FIGURA 6.2. Clasificación de los problemas de calidad

| | PROCESO/RESULTADO | ESTRUCTURA (ORGANIZACIONAL) |
|------------|---|--|
| ESPECÍFICO | <p>Problema tipo I (ejemplo: atención de pacientes con hipercolesterolemia)</p> | <p>Problema tipo II (ejemplo: organización del equipo/servicio en la atención de pacientes con hipercolesterolemia)</p> |
| GLOBAL | <p>Problema tipo III (ejemplo: control de pacientes con enfermedades crónicas)</p> | <p>Problema tipo IV (ejemplo: organización del equipo/servicio en la atención de pacientes de la consulta programada)</p> |

La elección de uno de estos tipos de problema debe hacerse de forma consciente valorando las consecuencias de la elección en relación a su medición y capacidad de resolución. En general la posibilidad de desarrollar metodológicamente el problema y las posibilidades de éxito de la evaluación son mayores con los problemas de tipo I (específicos y de proceso/resultado) que con los de tipo IV (genéricos y estructurales) pasando por los de tipo II (específicos y estructurales) y tipo III (genéricos y de proceso/resultado).

- **Problemas específicos mejor que genéricos.** Los problemas genéricos presentan un gran atractivo por su tendencia globalizadora de resumir diversos servicios asistenciales en una sola evaluación. Sin embargo, precisamente por su tendencia globalizadora, su evaluación puede reflejar espejismos de buena calidad. Cualquier problema se complica y se difumina cuando se trata de forma genérica. Por ejemplo la organización del trabajo en equipo para los atendidos en el centro de salud o en un servicio hospitalario es una oportunidad de mejora de tipo genérico en la que no se puede determinar con precisión sobre qué aspecto hay que incidir para mejorar ¿Sobre la cita previa? ¿El tiempo dedicado a cada consulta? ¿La coordinación con enfermería para la atención a un tipo de paciente concreto? Sin embargo, la organización del trabajo para atender a un tipo determinado de paciente (problema específico) es más fácil de comprender y de abordar que la organización del trabajo en general (problema genérico). Del mismo modo, los requisitos del registro de los datos relacionados con el control de una patología (problema específico) son más evidentes y fáciles de admitir y concretar que la calidad del registro en la historia clínica en general (problema genérico).
- **Problemas relacionados con el proceso y/o resultado mejor que los problemas relacionados con la estructura.** En la mayoría de los casos, la capacidad de resolución de los problemas estructurales por parte del personal de los centros sanitarios se limita exclusivamente a aquellos que están relacionados con aspectos organizativos. Sin embargo, los aspectos estructurales, aunque son necesarios, no son suficientes para que la calidad aumente. Por ejemplo, la inexistencia de esfigmomanómetro o la falta de fiabilidad del mismo (problema estructural) imposibilita u ocasiona un déficit en la calidad del diagnóstico y seguimiento de pacientes hipertensos. Sin embargo, si se plantea como un problema de calidad y éste se resuelve, se *facilitará* una atención adecuada pero, no hay ninguna garantía de que esto realmente se haga. En el mismo sentido, un equipo perfectamente organizado no es sinónimo de que los problemas de salud se atiendan adecuadamente (sin embargo, una mala organización puede ser la causa de defectos en la atención sanitaria). No obstante, un problema de organización que imposibilita que un determinado proceso se pueda llevar a la práctica, es equiparable en importancia a los problemas del proceso y/o resultado. Por ejemplo, en un programa de búsqueda de casos de riesgo cardiovascular elevado, un sistema de recepción de pruebas complementarias en el que son frecuentes los extravíos dificultará la adecuada valoración diagnóstica de la población incluida en el programa de prevención, y es importante, aunque este problema estructural se hace también evidente si elegimos como oportunidad de mejora la valoración del riesgo cardiovascular *per se*. En resumen, un problema estructural si es relevante, es candidato a ser detectado como causa de alteraciones en el proceso o resultado, mientras que la ausencia de problemas estructurales no significa

La factibilidad y facilidad de análisis de los problemas específicos y de proceso/resultado es en general mayor que la de los problemas genéricos y de estructura.

que el proceso y resultado consiguientes no sean defectuosos. A veces los problemas estructurales específicos son buenos candidatos por su fácil abordaje y solución. Pero normalmente no son los mejores. El análisis de problemas de proceso o resultado es más exhaustivo y permite determinar diferentes causas, entre las que se encuentran los problemas estructurales subyacentes.

5. MÉTODOS DE IDENTIFICACIÓN DE OPORTUNIDADES DE MEJORA. CONSIDERACIONES GENERALES

Ya tenemos claro que vamos a seleccionar un problema u oportunidad de mejora que, por cuestión de factibilidad del ciclo de mejora completo, es conveniente que tenga una serie de características. Pero no hemos revisado aún la metodología para identificar oportunidades de mejora para elegir. Esto es lo que vamos a ver a continuación.

Hay muchos métodos y, en principio, cualquier método puede ser válido; a la hora de utilizar uno u otro es conveniente ser flexibles y adaptarse a las posibilidades del contexto asistencial. Tampoco hay por que utilizar sólo uno, no son excluyentes. Uno de los principales factores determinantes de nuestra elección del método o métodos será el que dispongamos de datos sobre los servicios que prestamos o no.

- **Si no disponemos de datos**, la fuente de información que nos permite apreciar si hay problemas de calidad será exclusivamente la propia experiencia, el conocimiento que tengamos de la actividad de nuestros centros. Como se puede suponer la principal limitación de estos métodos, entre los que destacan la "lluvia de ideas" y la "técnica del grupo nominal", radica en que su efectividad es directamente proporcional al grado en el que conocemos de forma *directa* los servicios que se quieren mejorar. Lo que implica la necesidad de que seamos proveedores de esos servicios o que tengamos un contacto directo con los mismos, para que estos métodos se puedan utilizar con provecho. Por otra parte cuando los utilizamos, podemos sospechar la existencia de un problema de calidad pero, aunque es muy probable que éste exista, mientras no se cuantifique sólo será una oportunidad de mejora o un problema de calidad potencial, hipotético, de la que no tendremos constancia de su magnitud y por tanto si constituye en realidad un problema. Sin embargo, las ventajas de los métodos que no precisan datos sino el conocimiento del contexto en que se dan, superan a sus posibles limitaciones. Si los aplicamos adecuadamente, la oportunidad de mejora identificada se corresponderá con las preferencias de los profesionales evaluados, lo que facilitará el compromiso de estos con la resolución del problema.
- **Los métodos que requieren datos**, entre los que se encuentran el micromuestreo, la monitorización de indicadores, el análisis de perfiles, y la participación del usuario, no siempre estarán al alcance de cualquier centro ya que precisan sistemas de información específicos en algunos casos. Obviamente serán más precisos en la identificación de oportunidades de mejora ya que cuantifican la dimensión del problema.

Vamos a ver con más detalle ambos tipos de métodos, aunque enfatizando los métodos sin datos ya que pueden tener una validez similar a la de los que utilizan datos, pero además son más versátiles, sencillos y están al alcance de todo el

Hay diversos métodos para identificación de oportunidades de mejora, y todos pueden ser útiles. Es conveniente ser flexibles y realistas a la hora de seleccionar el o los que vamos a utilizar.

La identificación de problemas mediante métodos cuantitativos pueden tener una mayor precisión, sin embargo los métodos cualitativos no precisan de ninguna infraestructura previa y favorecen el compromiso de los profesionales con la mejora. Ambos tipos de métodos no son excluyentes si no complementarios.

mundo. La Tabla 6.1 contiene una relación de los métodos que vamos a revisar, agrupados según precisen o no datos.

TABLA 6.1. Métodos de indentificación y priorización de problemas

1. MÉTODOS QUE NO REQUIEREN DATOS

- 1.1. LLUVIA DE IDEAS (BRAINSTORMING).
- 1.2. TÉCNICA DE GRUPO NOMINAL.
- 1.3. MATRIZ DECISIONAL.
- 1.4. COMPARACIÓN POR PARES.
- 1.5. COMUNICACIÓN DE INCIDENCIAS.

2. MÉTODOS QUE REQUIEREN DATOS

- 2.1. MICROMUESTREO.
- 2.2. MONITORIZACIÓN DE INDICADORES.
- 2.3. ANÁLISIS DE PERFILES ESTADÍSTICOS.
- 2.4. LA PARTICIPACIÓN DEL USUARIO.

6. MÉTODOS PARA LA IDENTIFICACIÓN DE OPORTUNIDADES DE MEJORA QUE NO PRECISAN DATOS

Dentro de estos métodos vamos a exponer la *lluvia de ideas*, la *técnica del grupo nominal*, la *matriz decisional*, la *comparación por pares* y la *notificación de incidentes*. La opción por una u otra técnica dependen principalmente de la experiencia previa en la toma de decisiones en equipo, y del grado de conclusión al que queramos llegar durante el proceso de identificación (necesidad de completarla posteriormente con una técnica de priorización o no). Todos ellos se complementan y pueden utilizarse aislada o conjuntamente.

6.1. LA LLUVIA DE IDEAS (BRAINSTORMING)

Vamos a emplear esta técnica principalmente cuando no tenemos mucha experiencia en la toma de decisiones en grupo y cuando el objetivo es sacar a la luz una gran cantidad de oportunidades de mejora. Sin embargo, hay que considerar que el grado de conclusión de los resultados será menor en relación a la técnica del grupo nominal y que habrá que completar el proceso de identificación de problemas con alguna técnica de priorización como por ejemplo la matriz decisional o la técnica del grupo nominal. Además, los resultados pueden polarizarse, reflejando sobre todo los problemas aportados por aquellos miembros del grupo con mayor tendencia a la participación y a la exteriorización de sus impresiones.

El principio básico para llevar a cabo esta técnica consiste en evitar cualquier tipo de valoración o crítica cuando los participantes exponen sus ideas. De esta manera, no se coaccionará la libertad de los participantes ni se dificultará un proceso creativo que puede ser incompatible con un juicio crítico anticipado.

Podemos identificar cuatro normas de interés en la ejecución de este método (Tabla 6.2). Tras la introducción podemos empezar con la propuesta de una oportunidad de mejora, que en principio, pueda parecer banal, con el fin de crear un ambiente distendido que promueva a la participación y favorezca el planteamiento del mayor número de ideas. A partir de este momento, el encadenamiento y la asociación de ideas es la base que facilita una mayor creatividad. Cada participante planteará un solo problema en cada intervención, cuando considere oportuno y sin que se establezca un límite de intervenciones. Durante los episodios de silencio el moderador puede aportar nuevas ideas o variaciones sobre el tema.

La lluvia de ideas favorece la creatividad y la identificación de una gran cantidad de oportunidades de mejora, pero es difícil que sirva por sí sola para seleccionar una.

TABLA 6.2. Normas para el desarrollo de la lluvia de ideas

NORMAS

1. EXPOSICIÓN DEL OBJETIVO DE LA REUNIÓN POR PARTE DEL MODERADOR.
2. RECORDAR LAS REGLAS DEL MÉTODO: NO CRITICAR LAS IDEAS PROPUESTAS.
3. PLANTEAMIENTO, POR PARTE DEL ANIMADOR, DE UN EJEMPLO QUE PUEDE PARECER BANAL, PERO QUE FAVORECE UN AMBIENTE DISTENDIDO Y LA GENERACIÓN DE IDEAS.
4. EXPOSICIÓN, POR CADA PARTICIPANTE, DE UNA SOLA IDEA A LA VEZ CUANDO CONSIDERE OPORTUNO Y SIN LIMITACIÓN EN SU NÚMERO.

En ocasiones tendremos que completar este método con otras técnicas de consenso, especialmente cuando se necesiten clarificar los problemas planteados y redefinir aquellos que presenten puntos en común. Con la lluvia de ideas solamente es difícil seleccionar *una* oportunidad de mejora.

6.2. TÉCNICA DE GRUPO NOMINAL

Podemos aplicarla cuando los participantes tienen experiencia sobre el tema planteado, el grupo no es muy heterogéneo y tiene un tamaño inferior a 10 personas. Con esta técnica conseguiremos normalmente una gran validez en las decisiones realizadas a nivel grupal, a la vez que se favorece la resolución constructiva de los conflictos despersonalizando la confrontación entre los desacuerdos.

Este método se basa en utilizar la reflexión individual para generar ideas, y para su priorización. A lo largo de su desarrollo se procura la participación paritaria de todos los componentes del grupo para asegurar que se expresan también las posturas e ideas minoritarias, que quedarían ocultas o no se llegarían a expresar en una discusión abierta. Su desarrollo se realiza de una forma muy estructurada y requiere que se sigan las siguientes etapas:

- **Generación de ideas de forma individual.** Después de plantear en forma de pregunta escrita, concreta y sin ambigüedades, la cuestión sobre la que hay que reflexionar (en nuestro caso, las oportunidades de mejora existentes) los participantes escriben sus propuestas, de forma individual, sin que intercambien información y sin que se realicen aclaraciones sobre el tema. Para mayor

operatividad es conveniente limitar el número de propuestas por cada individuo, por ejemplo a tres o cinco.

- **Registro del listado individual de ideas.** Para ello solicitaremos a cada participante que de forma rotatoria expongan su listado de problemas (oportunidades de mejora) por turno de intervención para registrarlos en un listado común.
- **Clarificación de ideas.** Los problemas planteados en el listado común se someten a una discusión que tendrá como principal objetivo aclarar su significado, evitando duplicaciones y malentendidos. Posteriormente se puede volver a elaborar el listado eliminando los conceptos que se encuentren repetidos y concretando o matizando aquellos que sean ambiguos o que no respondan completamente a la cuestión planteada.
- **Priorización individual de las ideas (votación preliminar).** Durante esta fase se valoran de forma individual e independiente (sin discusión de grupo) las oportunidades de mejora del listado definitivo, otorgándoles una puntuación según se consideren más o menos relevantes. De esta manera se evitarán las presiones de los compañeros, y se garantizará que la toma de decisiones refleje lo que opinan todos los participantes. Obtendremos la decisión del grupo calculando el promedio de las priorizaciones individuales. La puntuación es el método de priorización más simple y el que se emplea habitualmente en el grupo nominal (por ejemplo de 0 a 5 puntos para cada problema, repartiendo un número de votos fijo entre todos ellos, o de otra cualquier otra forma que acuerde el grupo). Sin embargo, en este paso pueden utilizarse también otros métodos complementarios como la matriz decisional que veremos más adelante.
- **Discusión del voto preliminar.** Cada participante expresa su votación y se suman todas ellas para obtener la priorización del grupo. La decisión del grupo puede someterse a una nueva valoración en la que se intenta discutir las votaciones obtenidas. Se consideran especialmente los apartados con patrones de votación heterogéneos, o puntuaciones extremas (muy elevadas o muy bajas). En estos casos es importante que se descarte la existencia de factores artificiales, como por ejemplo la información insuficiente o parcial, así como interpretaciones incorrectas sobre la cuestión planteada. Si se ve oportuno tras esta discusión puede procederse a la votación final.
- **Voto final.** Con esta etapa se finaliza el proceso de toma de decisiones procediéndose de forma análoga a la 4ª etapa. El problema u oportunidad de mejora más votado será el elegido por el grupo para su evaluación y mejora.

6.3. OTROS MÉTODOS DE SELECCIÓN DE OPORTUNIDADES DE MEJORA: NOTIFICACIÓN DE INCIDENTES, COMPARACIÓN POR PARES Y MATRIZ DECISIONAL O PARRILLA DE ANÁLISIS DE CRITERIOS MÚLTIPLES

Hay otros varios métodos, enfocados sobre todo a la priorización, que establecen para valorar la relevancia del problema unos criterios de valoración. Para definir estos criterios pueden ser orientativas las indicaciones de la *Joint Commission on Accreditation of Health Care Organization*, en el sentido de seleccionar procesos que sean conjunta o separadamente muy *frecuentes*, de *alto*

La técnica del grupo nominal favorece la participación paritaria de todo el grupo en la identificación de la oportunidad de mejora

La técnica de grupo nominal es muy estructurada y se realiza siguiendo una serie de pasos determinados.

riesgo para los usuarios o con *tendencia reconocida a ser problemáticos*. Además se pueden añadir la dependencia interna de la solución y el coste de la misma. Cada centro determinará la adopción, adaptación o incorporación de estos u otros criterios de priorización, así como su grado de relevancia.

Los criterios que decidamos para valorar la oportunidad de mejora los podemos aplicar de forma implícita o bien explícita, como por ejemplo en la *comparación por pares* y en la *matriz decisional* respectivamente. Esta última normaliza las valoraciones realizadas por los diversos participantes en el proceso de priorización, a diferencia de lo que ocurre con la comparación por pares que al ser implícita es más subjetiva. Otro método sencillo y de utilidad, sobre todo cuando ya tenemos una estructura organizativa para las actividades de gestión de la calidad, es el establecimiento de un sistema de *recogida de incidencias o problemas (problem report)* identificados por los profesionales del centro. Estos problemas son considerados como posibles oportunidades de mejora.

Revisamos a continuación en qué consiste la comparación por pares y la matriz decisional.

a) La comparación por pares.

Con este método el proceso de valoración se realiza concentrándose en dos problemas como máximo de forma simultánea, comparando un problema con el otro y eligiendo aquel que se considere más importante. No es imprescindible, por tanto, que se apliquen criterios explícitos de valoración. Podemos distinguir en su desarrollo las siguientes etapas: (i) elaboración de una tabla en la que se detallan como cabeza de las filas y las columnas todos los problemas a considerar; (ii) confrontación de todas las parejas posibles, combinando los problemas de dos en dos, al cruzar filas y columnas; (iii) selección, dentro de cada pareja, del problema más importante, ese problema se anota en la celda correspondiente al cruce donde se encuentra la pareja comparada. Cada pareja se compara sólo una vez; (iv) cómputo final, contabilizando el número de veces en los que ha sido elegido cada problema durante la fase anterior; y (v) priorización de los problemas clasificándolos en orden decreciente a la puntuación obtenida. La Tabla 6.3 esquematiza este proceder aplicándolo a la priorización de cuatro problemas.

La comparación por pares prioriza con criterios implícitos, subjetivos.

TABLA 6.3. Priorización de problemas de calidad mediante el método de comparación por pares

| PROBLEMA | A | B | C | D |
|----------|---|---|---|---|
| D | D | D | C | 2 |
| C | C | C | 3 | |
| B | A | 0 | | |
| A | 1 | | | |

(Resultado C=3; D=2; A=1; B=0)

- A = utilización de antibióticos.
- B = calidad del registro de la historia clínica.
- C = control del paciente hiperlipémico.
- D = tratamiento de la infección respiratoria de vías altas.

b) La matriz decisional de criterios múltiples o parrilla de análisis

Esta herramienta tiene dos modalidades de aplicación en función de que los criterios empleados se consideren igual de importantes o bien se jerarquicen.

En el primer caso la parrilla de análisis adopta una forma matricial en la que se incluyen en la primera columna los distintos problemas identificados y en la primera fila los criterios de valoración (Tabla 6.4). Los problemas de calidad se valoran de forma separada concediéndoles una puntuación en cada uno de los criterios, para contabilizar posteriormente el total de puntos obtenidos. El nivel de priorización se establece ordenando los problemas de calidad en orden decreciente según la puntuación obtenida.

TABLA 6.4. Parrilla de decisión mediante criterios univalentes

- A= uso no racional de antibióticos;
- B= calidad del registro de la historia clínica;
- C= control del paciente hiperlipémico;
- D= tratamiento de la infección respiratoria de vías altas

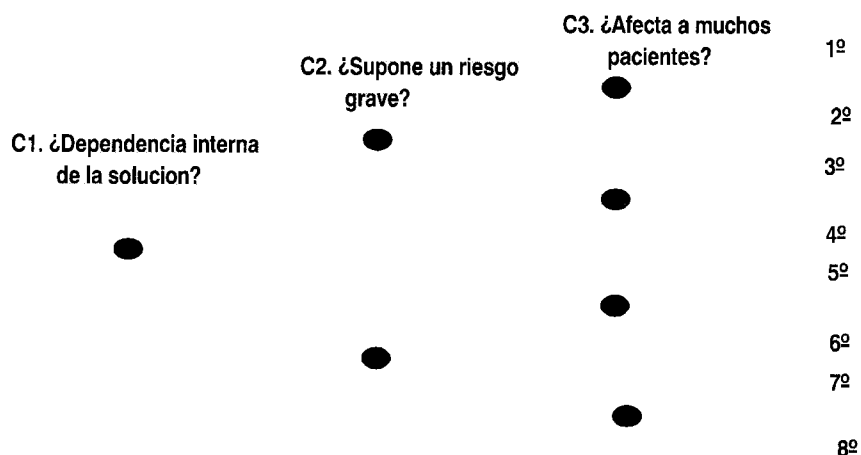
CRITERIOS DE DECISIÓN (máximo=5; mínimo=1)

| PROBLEMA | ¿AFECTA A MUCHOS PACIENTES? | ¿SUPONE UN RIESGO GRAVE PARA LA SALUD? | ¿DEPENDENCIA INTERNA DE LA POSIBLE SOLUCIÓN? | ¿ES UNA SOLUCIÓN BARATA? | TOTAL |
|----------|-----------------------------|--|--|--------------------------|-------|
| A | ++ | ++ | +++ | ++++ | 11 |
| B | ++ | + | +++ | ++++ | 10 |
| C | ++ | +++++ | +++ | ++++ | 14 |
| D | +++++ | + | +++ | ++++ | 13 |

En el segundo caso, previa jerarquización o priorización de los criterios decisionales, se diseña un gráfico de forma arborescente en el que los criterios de valoración se representan en forma de nódulos sucesivos (empezando por el más importante) y su valoración positiva o negativa se representa como un segmento superior o inferior respectivamente, que se inicia en ese nódulo y conecta con el siguiente (Figura 6.3). El criterio más importante se sitúa en el primer nódulo, continuándose con los demás criterios en función del orden decreciente de importancia. Cuando se valora un problema se sigue el recorrido determinado por la valoración positiva o negativa de los criterios y una vez que se llega al último se señala su nivel de prioridad.

La matriz decisional prioriza en base a criterios explícitos.

FIGURA 6.3. Parrilla de análisis aplicando criterios jerarquizados en orden de importancia (C1>C2>C3) (la línea continua corresponde a la valoración positiva del criterio)



La matriz decisional puede utilizarse como método de priorización individual en el paso 4 de la técnica de grupo nominal. La combinación de *lluvia de ideas*, *grupo nominal* y *matriz decisional* sin jerarquización de criterios es probablemente una de las más frecuentes y mejores bases para realizar la identificación y priorización de oportunidades de mejora, exclusivamente con el conocimiento que se tiene del entorno de trabajo en el que se prestan los servicios a evaluar.

7. MÉTODOS PARA LA IDENTIFICACIÓN DE OPORTUNIDADES DE MEJORA UTILIZANDO DATOS

El uso de este tipo de métodos está condicionada por la disponibilidad de fuentes de información que suministren los datos necesarios para aplicarlos. La posibilidad de identificar oportunidades de mejora puede ser más amplia con el micromuestreo y los métodos de participación del usuario mientras que la monitorización y el análisis de perfiles se suelen circunscribir a aspectos asistenciales concretos y previamente seleccionados.

7.1. MICROMUESTREO DE REGISTROS

Este método requiere que se registren los datos de la actividad asistencial. Cuando disponemos de estos registros (historias clínicas, documentos de derivación del paciente a otros dispositivos, recetas médicas...) podemos hacer una revisión rápida de una pequeña parte de los mismos (por ejemplo 15 o 20) para detectar áreas potencialmente problemáticas. Para aumentar la validez de este método se suele revisar cada registro por dos profesionales con experiencia en la materia. Éstos valoran si a su juicio existe alguna actuación incorrecta en la atención proporcionada al paciente comparando y resumiendo posteriormente sus anotaciones, a partir de las cuales se puede elaborar un listado de problemas. Pueden buscarse bien problemas genéricos (de antecedentes, prevención, utilización de analítica, etc.) o bien específicos si hacemos el mismo muestreo sobre un tipo concreto de pacientes.

7.2. MONITORIZACIÓN DE INDICADORES

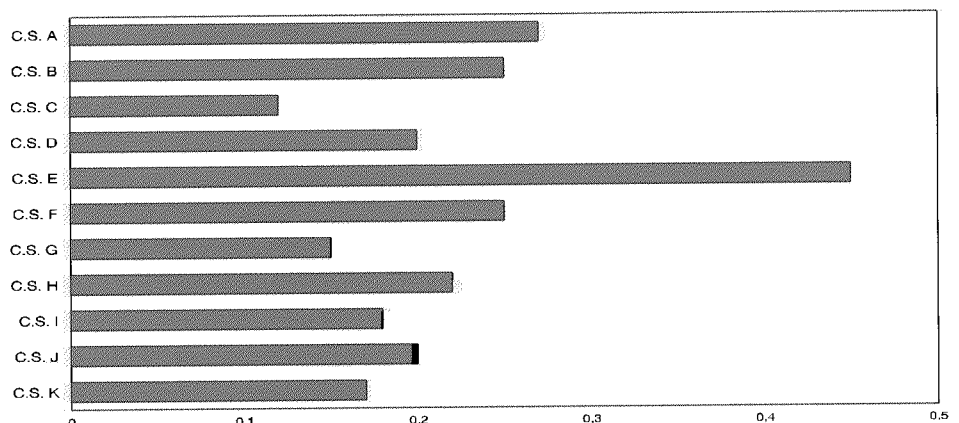
Los sistemas de monitorización de la calidad tienen como objetivo principal identificar problemas de calidad. Por tanto, si tenemos un sistema de medición periódica de indicadores sobre un aspecto esencial concreto (p.e. tiempos de espera, control del paciente hipertenso, del paciente diabético, etc.) podemos identificar problemas relacionados con esa área asistencial. Cuánto más amplio sea el listado de indicadores utilizado, mayor será también el espectro de problemas y servicios sobre los que será posible identificar la existencia de problemas, pero siempre va a ser más restrictivo que los métodos que no precisan datos.

7.3. ANÁLISIS DE PERFILES

Cuando la medición de un indicador se realiza entre distintos proveedores (facultativos, centros sanitarios, etc.), los resultados se pueden representar mediante un gráfico de barras. Cada barra refleja el nivel alcanzado por cada uno de los proveedores en la medición del indicador. Una amplia variabilidad en la distribución de los perfiles sugiere la existencia de un problema de calidad en alguna o varias de las áreas asistenciales que se encuentran relacionadas con el indicador valorado, especialmente en aquellos proveedores que se encuentran alejados en más de una o dos desviaciones estándar en relación a la media. Este método se ha empleado con frecuencia para identificar posibles oportunidades de mejora en la prescripción de medicamentos. Entre otros se han utilizado indicadores como la proporción de medicamentos de utilidad terapéutica baja, la proporción de antibióticos recetados en atención primaria pertenecientes al grupo de las quinolonas o las cefalosporinas, etc. Obviamente, precisa de una priorización previa de los aspectos a medir. La Figura 6.4 es un ejemplo de análisis de perfiles.

FIGURA 6.4. Análisis de perfiles

Proporción de quinolonas y cefalosporinas en relación al total de antibióticos recetados en 11 centros de salud. El área sombreada incluye a la media \pm una desviación estándar.



7.4. LA PARTICIPACIÓN DEL USUARIO EN LA IDENTIFICACIÓN DE PROBLEMAS

El paciente es el destinatario final del “producto” de la atención sanitaria. Su papel en la identificación de problemas de calidad es esencial en aquellos aspectos que están relacionados con el incumplimiento de sus expectativas, a la vez que constituye una fuente de información adicional para identificar oportunidades de mejora de procesos asistenciales que no se registran habitualmente. Su implicación en la identificación de oportunidades de mejora por parte de los centros asistenciales se puede articular mediante el análisis de las quejas y reclamaciones, las encuestas de satisfacción y los informes del usuario. Todas ellas precisan de una infraestructura previa para su realización y aprovechamiento. Resumimos aquí lo fundamental de estas herramientas como métodos de identificación de problemas.

7.5. ANÁLISIS DE QUEJAS Y RECLAMACIONES

Las ventanillas o los libros de recepción de quejas y reclamaciones son un medio clásico de identificación de problemas, aunque con frecuencia se haya desvirtuado su naturaleza considerándolo principalmente como un sistema punitivo. El número de reclamaciones recibidas en un centro sanitario puede ser un reflejo de su calidad y de las facilidades prestadas al usuario para participar en la identificación de problemas de calidad. En ocasiones tendremos que valorarlas con precaución ya que pueden responder a una interpretación errónea y no señalar en realidad un problema de calidad, pero en general nos van a ser de utilidad para identificar problemas relacionados con los aspectos que más valora el usuario, y pueden ser la punta del iceberg de problemas que debemos solucionar.

7.6. ENCUESTAS DE SATISFACCIÓN

Con las encuestas de satisfacción podemos identificar problemas similares a los señalados en el análisis de quejas y reclamaciones (no cumplimiento de las expectativas del usuario). Pero las ventajas sobre este último radican en que: (i) si la encuesta es metodológicamente correcta, identifican una mayor cantidad de problemas (relacionándolas directamente con la satisfacción) y (ii) reflejan los puntos de vista de una muestra representativa de pacientes, de manera que se evita el sesgo de incluir sólo la opinión de los pacientes que expresan su insatisfacción a iniciativa propia. En definitiva, las encuestas de satisfacción son en general una fuente de información más adecuada para sacar a la luz problemas relacionados con el trato recibido, el contexto en el que se desarrolla la asistencia y también problemas de tipo organizacional.

7.7. EL INFORME DEL USUARIO

De forma complementaria o como un sustituto de los registros habituales en los que se basa el micromuestreo, se puede preguntar al propio usuario sobre aspectos de la asistencia recibida que él conoce (p.e. ¿le informaron cuantas veces al día tenía que tomar esa medicación?, ¿cuándo le tomaron la tensión

arterial por última vez?). La principal ventaja de este método consiste en revelar áreas/problema de aspectos asistenciales que habitualmente no se registran y ser en ocasiones un método más eficiente de recopilación de datos en la identificación de problemas que la revisión de criterios.

8. EL CARTEL NARRATIVO: UN APOYO PARA LA REALIZACIÓN Y EL SEGUIMIENTO DEL CICLO DE MEJORA

El *story board* o *cartel narrativo* del proyecto de mejora es una representación paso a paso del ciclo de mejora. Aunque puede empezar a utilizarse a cualquier altura dentro del ciclo de mejora, lo más lógico es establecerlo una vez decidido el problema a evaluar y mejorar o, como mínimo, una vez decidida la intervención para mejorar. Los elementos del cartel completo se pueden ver en la Figura 6.5. Al principio estarán los espacios en blanco y se irán ilustrando y llenando a medida que vaya avanzando el proyecto. Para cada paso del ciclo de mejora se va informando de las actividades clave y sus herramientas utilizadas. El que esté todo ello expuesto en un lugar visible fomenta la responsabilidad y compromiso del grupo de mejora encargado del proyecto, a la vez que hace posible y fácil recibir comentarios y sugerencias de otros profesionales del centro.

Esta herramienta de ayuda a la implementación de los proyectos se originó de forma casi casual en la industria (concretamente la popularizó Florida Power and Light, la primera compañía americana que ganó el Premio Deming), pero se ha extendido a todos los sectores, incluido el sanitario, por su sencillez y utilidad, fomentando la transparencia de las actividades de mejora de la calidad y, como ya hemos apuntado, el compromiso de los grupos de mejora implicados. Adicionalmente, es una herramienta total y absolutamente compatible con todas las demás; de hecho, cualquier método que se utilice en cualquiera de los pasos del ciclo de mejora puede y debe verse reflejado en el cartel, que tiene así, además, una función educativa para todo el personal del centro.

Se incluye como anexo las instrucciones para su elaboración, adaptado para uso en países latinoamericanos por el proyecto de Garantía de Calidad (QA Project) que se realizó con fondos de la AID (Agency for International Development) norteamericana, y gestionado por la URC (University Research Corporation), en el cual hemos tenido la fortuna de colaborar.

Los carteles narrativos fomentan la transparencia de las actividades de mejora de la calidad, el compromiso de los grupos de mejora implicados y tiene además una función educativa para todo el personal

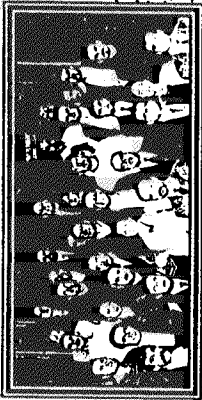
FIGURA 6.5. Ejemplo de cartel narrativo

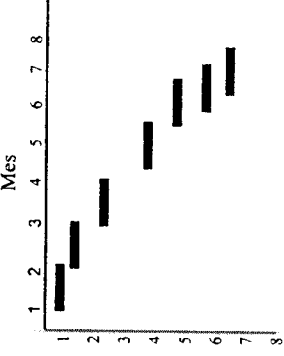
DEFINICIÓN DEL PROBLEMA

EQUIPO

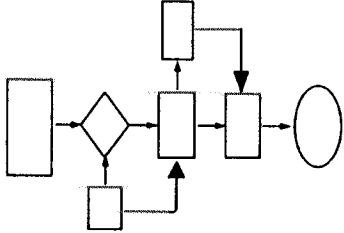
FOTO

CRONOGRAMA

Nombres: 

Roles: 

ANÁLISIS DEL PROBLEMA



Flujograma

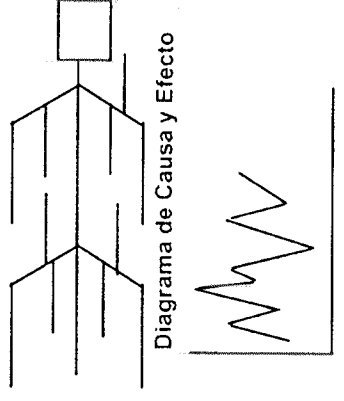


Diagrama de Causa y Efecto

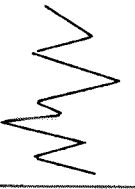


Gráfico de Tendencia

| CRITERIO | Nº OCURRENCIAS (Presentación actual) | Nº OCURRENCIAS (objetivo) | Frecuencia (objetivos) |
|----------------------------|--------------------------------------|---------------------------|------------------------|
| 1. Diagnóstico correcto | 31 | 15.8 | 15.8 |
| 2. Antecedentes personales | 45 | 12.2 | 12.2 |
| 3. Antecedentes personales | 42 | 12.3 | 12.3 |
| 4. Consumo de alcohol | 20 | 10.5 | 10.5 |
| 5. Consumo de alcohol | 30 | 10.5 | 10.5 |
| 6. Peso y talla y/o IMC | 17 | 7.9 | 7.9 |
| 7. Auscultación cardiaca | 15 | 4.4 | 4.4 |
| 8. Exploración abdominal | 15 | 4.4 | 4.4 |
| 9. Pulsos periféricos | 12 | 5.5 | 5.5 |
| 10. Soplos cardíacos | 4 | 2.8 | 2.8 |
| 11. Eteimas | 6 | 2.8 | 2.8 |
| 12. Fondo de ojo | 242 | 100 | 100.0 |

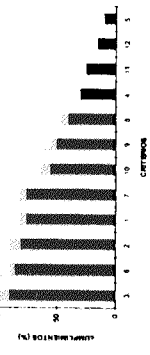
Lista de Criterios

- Diagnóstico correcto
- Antecedentes personales
- Antecedentes personales
- Consumo de alcohol
- Consumo de alcohol
- Peso y talla y/o IMC
- Auscultación cardiaca
- Exploración abdominal
- Pulsos periféricos
- Soplos cardíacos
- Eteimas
- Fondo de ojo

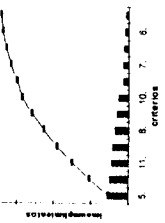
Lista de chequeo

| Cuestionario | si/no |
|-------------------|--------|
| 1. XXXXXXX? A B C | XXXXXX |
| 2. XXXXXXX? A B C | XXXXXX |
| 3. XXXXXXX? A B C | XXXXXX |
| 4. XXXXXXX? A B C | XXXXXX |
| 5. XXXXXXX? A B C | XXXXXX |

Instrumentos



Criterios



Presentación y Análisis de datos

CAUSAS PRINCIPALES

- _____
- _____

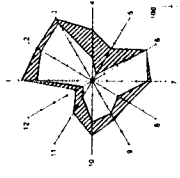
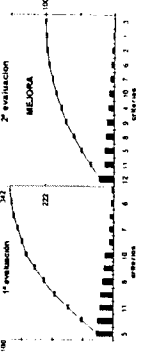
SOLUCIÓN

- _____
- _____

RESULTADOS

| Antes | Después | Manteniendo lo alcanzado |
|-------------|-------------|--------------------------|
| % _____ | % _____ | % _____ |
| Fecha _____ | Fecha _____ | Fecha _____ |

Análisis gráfico

ANEXO: INSTRUCCIONES PARA LA ELABORACIÓN DEL CARTEL NARRATIVO**CARTEL NARRATIVO DE EVALUACIÓN Y MEJORAMIENTO DE LA CALIDAD⁽¹⁾**

Enunciado del problema.

Describir la oportunidad que existe de mejorar en términos de sus consecuencias. El enunciado del problema debe definir con claridad la diferencia entre la situación deseada y la existente. No debe hacer ninguna referencia a la causa del problema; no debe culpar a otros, ni debe expresarse como una solución. Deberá ser evidente dónde comienza y termina el problema, de manera que se marque la diferencia respecto de otros problemas del sistema en general. Si se desea, también puede describirse el método o los métodos utilizados para seleccionar un problema o campo para mejorar.

El equipo de EMC (Evaluación y Mejora de la Calidad).

Identificar a todos los integrantes del equipo, sus funciones en las organizaciones y las funciones que cumplirán en el esfuerzo de mejora de la calidad. Si es posible se puede incluir una fotografía del equipo. También se puede incluir un cronograma general que enumere todas las actividades principales y cuándo se espera iniciarlas y terminarlas. No es necesario seguir exactamente lo que indique el cronograma general, dado que sólo proporciona una guía general para los miembros del equipo.

Análisis del problema.

El análisis del problema debe documentar los esfuerzos realizados por el equipo para determinar cuáles son las causas de fondo del problema. No es necesario incluir todos los elementos que se enumeran a continuación, pero deben mencionarse la causa o las causas de fondo del problema, las que constituirán el punto central de las actividades de mejora.

Diagrama de flujo. Sirve para mostrar el proceso que existe actualmente e indicar cuáles son los límites del esfuerzo para mejorar la calidad. Indicar los impedimentos o los puntos problemáticos con una flecha. Trazar un círculo alrededor de los campos en los que será necesario recopilar datos.

Diagrama de causa y efecto. Mostrar el análisis de las causas y efectos utilizando un diagrama de causa y efecto. Marcar aquellas causas sobre las que deben recogerse datos. Recordar que se trata de causas hipotéticas hasta que estén fundamentadas con datos.

Indicadores o criterios principales. Criterios o indicadores que cuantifican la diferencia entre el desempeño real y el deseado, así como los que determinen la frecuencia de las causas posibles que se marcaron antes, en caso de que sea un estudio de causas hipotéticas.

Instrumentos para la recolección de datos. Incluir los instrumentos para la recolección de datos que se desarrollaron para medir los indicadores anteriores.

Exhibición de los datos. Presentar los resultados en cuadros adecuados, incluyendo la representación gráfica (histogramas, diagramas de barras o diagramas "Pareto").

Identificar causas de fondo. Basándose en el análisis, el equipo debe indicar qué causa o causas quisieran abordar a fin de solucionar el problema y mejorar la calidad. Es importante acordarse de concentrarse en las "pocas causas fundamentales" del problema que los datos hayan revelado.

Formular una solución. Formular una descripción clara y sencilla de las mejoras propuestas para el proceso. Resumir los pasos clave del plan para poner en ejecución la solución. Describir los métodos utilizados para diseñar y seleccionar una solución (opcional).

Resultados. Mostrar los datos que demuestran los adelantos logrados en términos de mejoras en el proceso y los resultados deseados. Se pueden utilizar los mismos indicadores que se utilizaron para analizar el problema, o un subconjunto de los mismos. Los resultados deben demostrar que la discrepancia demostrada en el enunciado del problema se ha reducido, logrando que el desempeño real se parezca más al desempeño deseado.

Es conveniente que los resultados se midan nuevamente después de 6 meses o un año para asegurar que las mejoras se mantienen. Si se ha producido el problema nuevamente, el equipo debe estudiar la situación para comprender el motivo y se deben tomar otras medidas correctoras.

⁽¹⁾ Batalden P, Gillen T: "Hospitalwide Quality Improvement Storytelling", Quality Resource Group, Hospital Corporation of America, Nashville, Noviembre de 1989.

BIBLIOGRAFÍA

- Palmer RH. Identificación de problemas en la asistencia al paciente. En: Palmer RH. Evaluación de la asistencia ambulatoria: Principios y práctica. Madrid: Ministerio de Sanidad y Consumo; 1990. p. 63-77
- Saturno PJ. Monitorización del uso de medicamentos para mejorar la calidad de la prescripción: problemas, métodos e indicadores. Aten Primaria 1996; 18(6):331-338.
- Saturno PJ. Los métodos de participación del usuario en la evaluación y mejora de la calidad de los servicios sanitarios. Rev Esp Salud Pública 1995; 69:163-75.
- Delbecq AL, Van de Ven AH, Gustafson DH. Técnicas grupales para la planeación. Mejico DF: Trillas; 1984.
- Pineault R, Daveluy C. La determinación de prioridades. En: La planificación sanitaria: Conceptos, métodos, estrategias. Madrid: Masson; 1987. p. 43-211.
- Sannazaro PJ, Mills DH. A critique of the use of generic screening in quality assessment. JAMA 1991; 265:1977-81.
- Joint Commission on Accreditation of Health Care. Monitoring and Evaluation Process (cap. 4). En: Quality Assurance in Ambulatory Care. 2º ed. Chicago: JCHO, 1990; 39-82.

7

**MÉTODOS DE ANÁLISIS
DE LOS
PROBLEMAS DE CALIDAD**

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

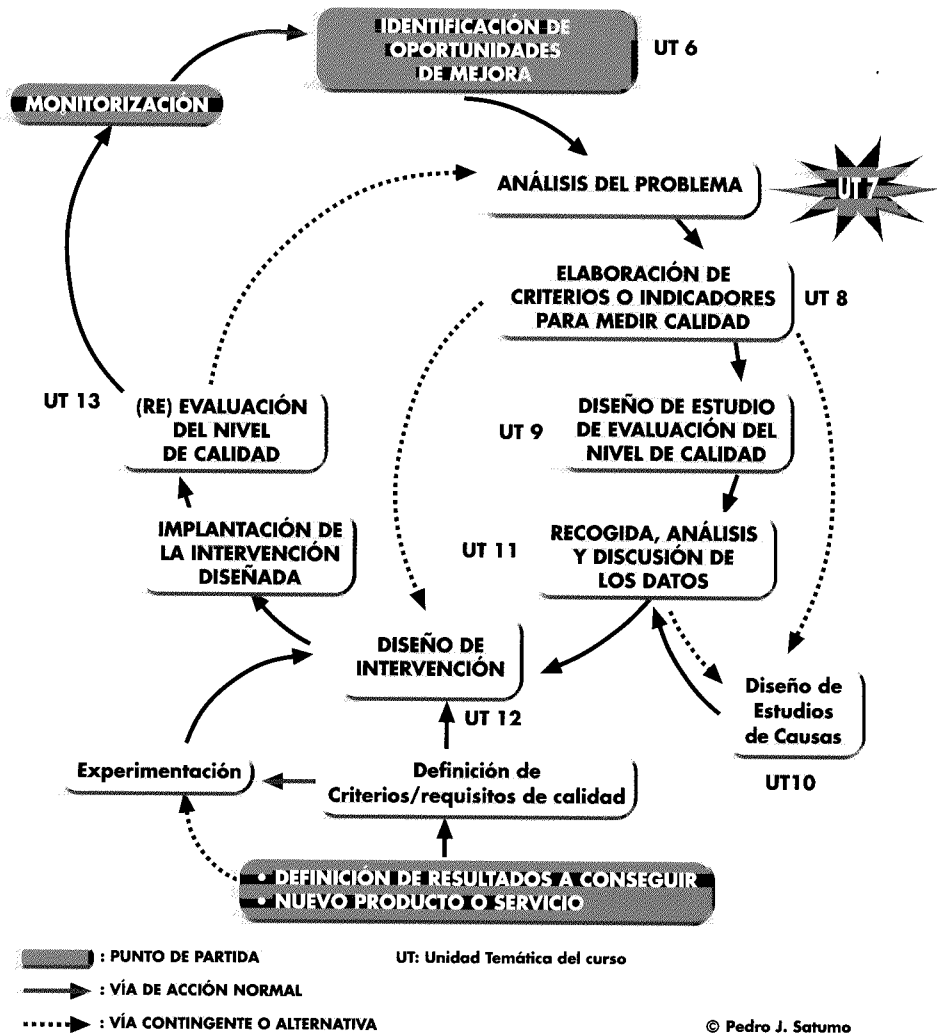
Una vez delimitado en la UT6 el problema de calidad u oportunidad de mejora a evaluar, esta UT desarrolla la metodología para analizarlo con y sin datos, y sienta las bases para las acciones o pasos siguientes. Se revisan diversos métodos útiles para analizar problemas de calidad, explicándose en detalle y con ejemplos prácticos la utilización para estos fines del diagrama de Pareto, histograma, flujograma y diagrama de Causa y Efecto.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. ¿Por qué se realiza el análisis del problema de calidad u oportunidad de mejora?
3. Métodos de análisis de problemas.
4. Análisis del problema con métodos que no requieren datos.
5. Herramientas de análisis que requieren datos sobre el problema.

OBJETIVOS ESPECÍFICOS

1. Describir los objetivos que se persiguen al analizar los problemas de calidad con cualquiera de los métodos de análisis que se utilice.
2. Enumerar al menos dos métodos utilizables sin disponer de datos y dos métodos que precisen de la existencia previa de datos.
3. Identificar los principales tipos de problemas o situaciones para los que está indicado cada uno de los métodos.
4. Especificar las condiciones de aplicación de los métodos de análisis sin datos.
5. Especificar las limitaciones en la interpretación de los resultados de los análisis de problemas sin datos.
6. Definir el significado y distinguir los símbolos utilizados en un flujograma.
7. Analizar estructuralmente un flujograma.
8. Realizar un diagrama de Causa y Efecto para analizar problemas de calidad.
9. Analizar los resultados de un diagrama de Causa y Efecto.
10. Distinguir entre los tres principales tipos de causas que pueden ser identificadas al analizar los problemas de calidad.
11. Formular diversos cursos de acción sobre la base del análisis efectuado.
12. Decidir el curso o cursos de acción consiguientes, según los resultados del análisis efectuado.
13. Tener en cuenta las implicaciones metodológicas y los riesgos de los diversos cursos de acción.



"No podemos considerar nada verdad si no reconocemos claramente que lo es. Para conseguir esto, puede que sea necesario dividir un problema complejo en cuantas partes parciales sea posible. Entonces se puede empezar por las ideas más sencillas"

(Descartes)

1. INTRODUCCIÓN

En la unidad temática anterior habremos identificado una oportunidad de mejora. Para avanzar en el ciclo vamos a ver que, además de que hay varias metodologías de utilidad para analizar las oportunidades de mejora, incluso sin tener datos, un análisis correcto nos aclara el camino a seguir, y que éste, puede ser múltiple. Para decidir las alternativas de acción que conduzcan, al final, a la mejora del probable problema de calidad identificado, debemos reflexionar acerca del grado de conocimiento que tenemos sobre el mismo (¿Cómo es de importante? ¿Cómo se puede corregir?), y realizar un *análisis estructurado* para identificar los factores implicados en él. Sobre ésta base vamos a decidir qué hacer.

En esta UT se sientan las bases para el camino a seguir en la continuación del ciclo de mejora

2. ¿POR QUÉ SE REALIZA EL ANÁLISIS DEL PROBLEMA DE CALIDAD U OPORTUNIDAD DE MEJORA?

Una vez que hemos seleccionado la oportunidad de mejora podemos preguntarnos de qué forma puede abordarse para conseguir su resolución. Sin embargo, es probable, de hecho es lo más frecuente, que en este segundo paso del ciclo de mejora tengamos un conocimiento insuficiente del problema y no estemos preparados para pasar al planteamiento de soluciones. Para ello requerimos, indagar sobre sus causas e incluso cuantificar hasta qué punto existe o no el problema. No es fácil cuantificar el problema ni identificar las medidas de intervención cuando éste se aprecia como un todo, en conjunto y sin un conocimiento explícito, visualizado, de sus implicaciones. Normalmente conviene que lo examinemos y lo dividamos en aspectos (subproblemas o causas) que sean lo suficientemente manejables para determinar, en cada una de ellas, el tipo de contribución al problema analizado y el curso de acción consiguiente que conviene emprender. En consecuencia, el análisis estructurado de la oportunidad de mejora lo vamos a realizar para conseguir los siguientes objetivos: (i) dimensionar el problema; (ii) identificar las causas; (iii) clasificarlas y priorizarlas; y (iv) decidir cómo continuar en el camino hacia la mejora.

- **Dimensionar el problema.** ¿Constituye la oportunidad de mejora un problema de calidad real? ¿Qué magnitud tiene el problema? La respuesta es más fácil si los métodos empleados en la identificación de la oportunidad de mejora han sido de tipo cuantitativo tales como la monitorización, el análisis de perfiles o el informe del usuario, o bien disponemos ya de datos relacionados con ese tema. Sin embargo, con el resto de los métodos intuimos una divergencia entre lo que se observa y lo que es razonable que exista (se ha sospechado un problema de calidad) pero no se ha objetivado todavía. En estos casos la reflexión sobre sus posibles causas y la frecuencia con la que se dan éstas puede confirmarnos de nuevo intuitivamente la importancia del problema. Sin embargo, el dimensionado real no se conocerá (a menos de que tengamos ya cuantificadas las causas), hasta que no se mida.
- **Identificar las causas.** Consiste en descomponer el problema en unidades elementales que sean susceptibles de medición y que a la vez se puedan traducir en medidas operativas que faciliten la resolución del problema. Por ejemplo la existencia de hojas sin numerar puede ser una de las causas de que no se pueda revisar la historia clínica de los pacientes atendidos (que a su vez puede contribuir como causa al problema de calidad que estamos analizando). Además de ser medible se puede convertir en una medida correctora, anulando, contrarrestando o modificando la propia causa, que en este caso sería simplemente la numeración de las hojas de la historia clínica.
- **Clasificación y priorización de las causas.** Con el análisis estructurado es posible identificar las principales *subáreas-problema* (componentes principales del problema) y/o las *causas* del problema (componentes medibles y operativos para la intervención) a la vez que permite priorizar el orden de intervención para su corrección o estudio, según su importancia y el grado de conocimiento que tengamos sobre su frecuencia.
- **Decidir cómo continuar.** Éste es el principal objetivo del análisis del problema u oportunidad de mejora. Como veremos más adelante lo más posible es que nos encontremos con diversos tipos de causas frente a las cuales la acción

El análisis del problema de calidad u oportunidad de mejora se realiza para dimensionar el problema, reflexionar sobre sus causas y decidir qué hacer a continuación en el camino hacia la mejora.

más oportuna es también diferente. Resumidamente, habrá tres tipos de causas: (i) causas sobre las que no podremos actuar fácilmente, típicamente causas estructurales en relación a la adquisición de recursos, (ii) causas sobre las que será posible la intervención directa, porque sabemos que influyen en el problema de calidad analizado, y sabemos su grado de presencia en nuestro centro: también suelen ser causas estructurales y organizacionales pero a nuestro alcance: por ejemplo, la inexistencia de protocolización, la falta de numeración de las páginas en la historia, circuitos innecesariamente complicados o redundantes, etc. y (iii) causas que tendrán que ser investigadas y cuantificadas. Si resumimos aún más, los dos cursos de acción a los que nos va a guiar el análisis del problema son la intervención directa y/o el diseño de estudios para un conocimiento cuantificado de las causas, sobre el que basar la intervención.

Todos estos objetivos generales pueden lograrse utilizando métodos muy diversos, cuya aplicabilidad depende del tipo de problema y de la existencia o no de datos sobre el mismo.

Vamos a revisar a continuación los métodos que son en general de más utilidad, clasificados según que precisen o no de la existencia previa de datos.

3. MÉTODOS DE ANÁLISIS DE PROBLEMAS

No hay reglas fijas para elegir un método de análisis determinado. Lo importante es que el método o métodos elegidos se adapten al problema y permitan alcanzar los objetivos de éste análisis. La única limitación radica en la disponibilidad o no de datos que cuantifiquen las subáreas y/o las causas del problema. En la Tabla 7.1 se resumen los métodos que vamos a revisar.

TABLA 7.1. Herramientas de análisis de la oportunidad de mejora

1. SIN DATOS

- 1.1. DIAGRAMA DE FLUJO.
- 1.2. DIAGRAMA DE CAUSA Y EFECTO.
- 1.3. ANÁLISIS DE CAMPOS DE FUERZA.

2. CON DATOS

- 2.1. HISTOGRAMA.
- 2.2. ESTRATIFICACIÓN.
- 2.3. DIAGRAMA DE PARETO

- **Cuando no tenemos datos**, podemos analizar el problema básicamente con tres tipos de herramientas: el flujograma, el diagrama de causa-efecto y el análisis de campos de fuerza (ver algoritmo representado en la Figura 7.1). El flujograma nos será de gran utilidad en problemas relacionados con procesos que tienen numerosas actividades secuenciadas en el tiempo, como en el caso de los problemas de tipo organizacional, circuitos, coordinación de profesionales y niveles, etc., (por ejemplo, la cita programada, la organización de médicos con enfermeros en el seguimiento de pacientes, el circuito de obten-

El principal objetivo del análisis de los problemas de calidad es decidir el curso de acción a seguir; que puede ser la intervención directa y/o el diseño de estudios para cuantificar las causas y fundamentar así la intervención

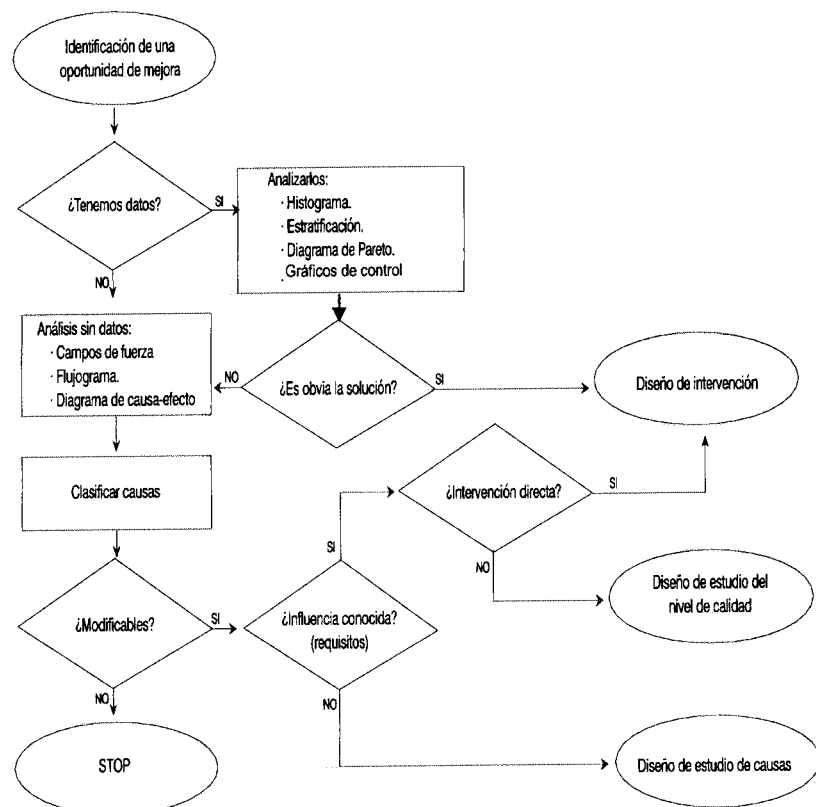
ción de pruebas complementarias desde que se solicitan hasta que se obtienen, etc.), mientras que el diagrama de causa-efecto tiene utilidad para todo tipo de problemas. El análisis de campos de fuerza, que ya vimos como método de diseño de estrategias, puede ser también utilizado para identificar causas sobre las que intervenir o cuantificar. Por último, aunque cada herramienta tiene utilidad por sí sola, pueden también utilizarse combinadas: por ejemplo un flujograma puede identificar dónde están los problemas de un circuito o proceso determinado, y se puede aplicar un diagrama de causa y efecto para analizar las causas.

Aunque no precisan de datos, la principal condición para el empleo de estos métodos es que se tenga experiencia sobre el problema analizado con el fin de aportar la información necesaria para el análisis. Su empleo será adecuado, por tanto, en la exploración de oportunidades de mejora que nos sean conocidas de forma directa. Por otra parte, aunque en principio los podemos elaborar de forma individual, la información obtenida es mayor, y en consecuencia el análisis es más exhaustivo, si se hace en grupo; más aún si este grupo, conocedor del tema, es el que se va a encargar de solucionar el problema, es decir, si es lo que se llama un *grupo de mejora*.

- **Si tenemos datos del problema** el análisis se puede realizar utilizando el histograma, la estratificación, el diagrama de Pareto, y, si tenemos una serie temporal de mediciones, los gráficos de control (Figura 7.1). Con estos métodos estaremos más seguros sobre qué hacer a continuación; qué intervenciones son oportunas y/o hacia dónde dirigir estudios más en profundidad. Es obvio que su gran limitación se encuentra en la disponibilidad de los datos necesarios.

Los problemas de calidad pueden analizarse con y sin datos. Los métodos sin datos son muy útiles pero precisan de un conocimiento directo del problema analizado y sus hallazgos son hipotéticos. Los análisis con datos son más seguros para tomar decisiones pero es preciso disponer de los datos necesarios.

FIGURA 7.1. Algoritmo decisional del análisis de la oportunidad de mejora



Una vez enunciados, vamos a ver en qué consisten todos estos métodos, a excepción del análisis de campos de fuerza que ya vimos en la UT5.


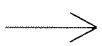

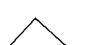
4. ANÁLISIS DEL PROBLEMA CON MÉTODOS QUE NO REQUIEREN DATOS

4.1. FLUJOGRAMA

Con el diagrama de flujo o flujograma vamos a describir, de forma estructurada, la secuencia de actividades de un proceso determinado.

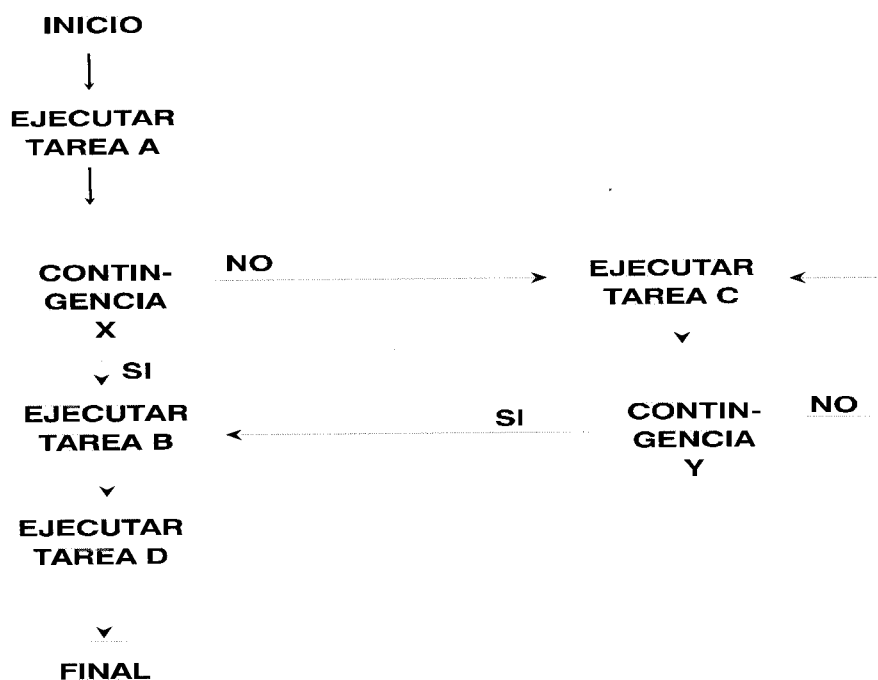
Es una de las herramientas más útiles para analizar problemas organizacionales o de procesos que incluyen numerosas tareas secuenciadas a lo largo del tiempo (por ejemplo, los circuitos de petición de pruebas complementarias). Con él simplificamos un proceso complejo dividiéndolo en actividades definidas en las que sea más asequible apreciar en cuál de ellas o en qué paso entre ellas puede haber problema. De hecho, la percepción directa de procesos complejos sería análoga a captar en una sola instantánea un objeto en movimiento, con la consiguiente pérdida de información. Mediante el flujograma conseguiremos apreciar a cámara lenta este movimiento para identificar cada tarea del mismo y apreciar las interconexiones y relaciones existentes.

Entre los principales componentes a plasmar en un flujograma podemos distinguir: (i) una entrada o inicio; (ii) actividades concretas; (iii) situaciones de contingencia, donde se comprueba si ocurre o no determinada circunstancia decisiva y (iv) una salida. Estos elementos se representan mediante símbolos preestablecidos y su secuencia representa el proceso analizado, tal como se puede ver en la Figura 7.2. Estos símbolos son:

-  *Elipse*: representa el inicio (lo que se requiere para comenzar el proceso) y el final del proceso analizado.
-  *Flecha*: señala la dirección en la que fluye el proceso dentro de la secuencia de tareas.
-  *Rectángulo*: representa una tarea o una acción determinada dentro del proceso. Normalmente se sigue de un solo flujo de actividad, de lo contrario requerirá que se intercale un bloque de contingencia o decisión.
-  *Rombo*: señala una situación de contingencia o comprobación que dividirá la dirección del flujo en dos o más opciones, y en la que la conveniencia de elegir una dependerá de que se dé una circunstancia determinada. Por tanto, se sigue de tantos flujos como posibilidades tiene esta contingencia, aunque, habitualmente, suele ser dicotómica: "sí" o "no".

El flujograma es una herramienta muy útil para analizar circuitos, problemas organizacionales y, en general procesos que incluyen numerosas tareas secuenciales

FIGURA 7.2. Representación gráfica de un proceso mediante el diagrama de flujo



Para elaborar el flujograma se pueden seguir los siguientes pasos:

1. Identificación de las actividades o tareas (por ejemplo mediante una lluvia de ideas o por cualquier otro medio de reflexión grupal) y determinar quienes son sus responsables respectivos.
2. Clasificación de las tareas en orden cronológico y determinar los flujos del proceso (¿qué dirección sigue?).
3. Identificar situaciones en las que una determinada tarea no se pueda empezar porque la anterior no se ha concluido adecuadamente o no ha ocurrido algo que se supone debía haber ocurrido y determinar qué actividad o contingencia es necesaria para que se resuelva esta situación y en qué orden del cronograma se colocaría.
4. Representar el flujograma y sustituir las situaciones en las que surge más de un flujo de una sola tarea por un nódulo de contingencia (con un rombo). Posteriormente habrá que señalar de qué tipo de contingencia se trata (dentro del rombo) y cuáles son sus posibilidades al lado de sus flujos respectivos. En este punto se puede dar por terminado el flujograma como representación gráfica del proceso, aunque generalmente es posible continuar seguir identificando actividades que probablemente no se hayan identificado en un principio, concentrando nuestro foco de atención entre dos actividades consecutivas y cuestionándonos si se requiere alguna actividad o alguna contingencia entre ellas.

Una vez que hemos elaborado el flujograma podemos plantear las siguientes cuestiones generales para analizar el problema de calidad:

Para representar un proceso mediante un flujograma se utilizan una serie de símbolos preestablecidos con significados concretos

- *Análisis del flujo general:* ¿Es lógico el flujo? ¿Hay diferencias entre el flujo ideal y el realizado de forma habitual? ¿Cuáles son las causas de estas diferencias?
- *Análisis de cada etapa del proceso:* ¿De qué forma esta etapa no puede cumplir su función? ¿Cuáles son los problemas potenciales de calidad en esta etapa? ¿Cuáles son las posibles causas de estos problemas?
- *Análisis de los bucles o pasos hacia atrás:* ¿Cuáles son las causas que motivan un retraso en el desarrollo del proceso?

Esta herramienta se puede utilizar de forma combinada con el diagrama de causa-efecto cuando queremos indagar acerca de las causas de los puntos críticos del proceso.

La Figura 7.3 contiene un ejemplo de aplicación de flujograma, que puede servir para ver más claramente en qué consiste.

4.2. EL DIAGRAMA DE CAUSA-EFECTO

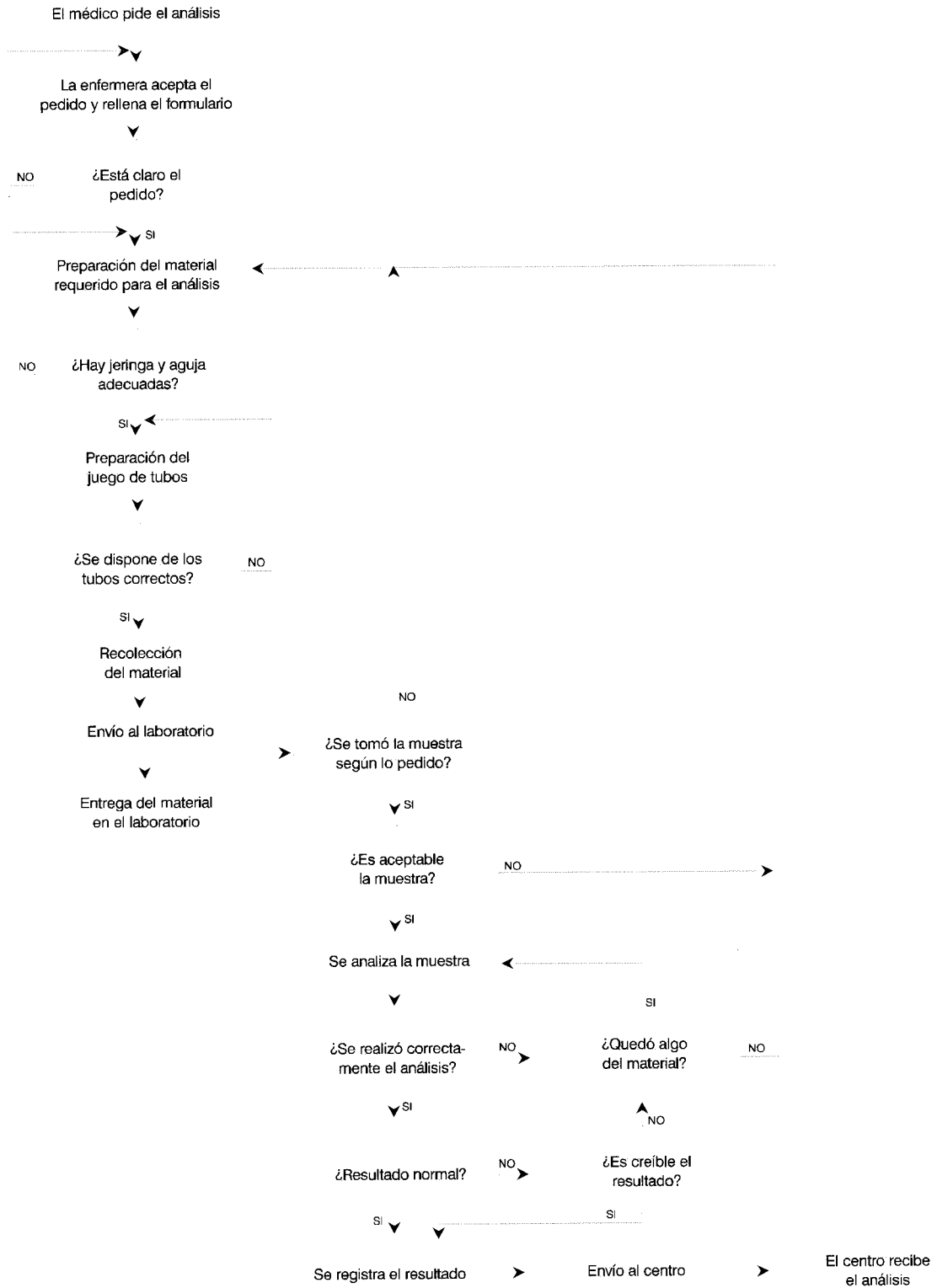
El análisis una oportunidad de mejora mediante el *diagrama de causa-efecto* es probablemente lo más frecuente e incluso aconsejable cuando no disponemos de datos del problema. En la mayoría de las situaciones en las que se identifica un problema se pueden apreciar de forma más o menos inmediata un reducido número de causas. Pero, si no se examina con más profundidad pueden pasar desapercibidos otros factores que aunque no son tan evidentes sí que son esenciales para comprender el origen del problema.

Con esta herramienta podemos descomponer una oportunidad de mejora en elementos identificables medibles y operativos para la intervención (causas potenciales del problema). Una vez que se han identificado las causas, trasladamos la reflexión sobre el grado de conocimiento del problema en conjunto, a la reflexión sobre el grado de conocimiento de sus causas potenciales. En este nivel de acercamiento estamos en mejores condiciones para conocer la dimensión (frecuencia) o la cuantificación de parte de sus causas. Cuando se tiene este conocimiento, el curso de acción más oportuno (diseño de intervención y/o estudio) aparece con mucha mayor claridad.

El diagrama de causa-efecto, se conoce también como diagrama de Ishikawa (su inventor) o "diagrama de espina de pescado" (por el aspecto que tiene) es probablemente una de las herramientas más sencillas y, sin embargo, más eficientes para determinar las causas de un problema de calidad o detectar subáreas problemáticas. Representa la relación entre un problema de calidad determinado (efecto) y sus causas potenciales según se esquematiza en la Figura 7.4. Para ello se exploran de forma sistematizada y en sucesivas fases de profundización en el análisis, las distintas categorías o clases de causas, los factores que contribuyen al problema, identificándolos y relacionándolos unos con otros.

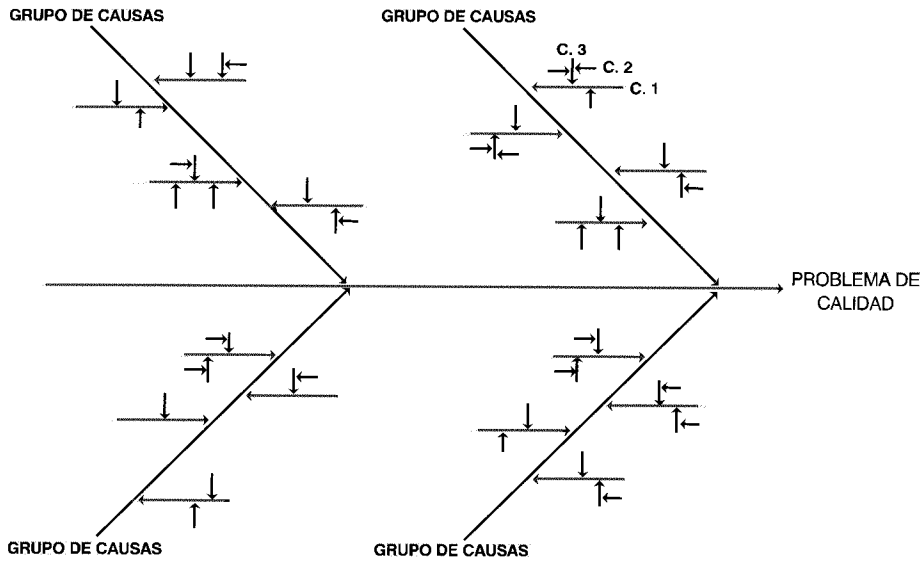
El análisis del flujograma se realiza considerando una serie de aspectos concretos. Adicionalmente puede utilizarse combinado con otros métodos de análisis para indagar las causas de los puntos señalados como críticos o problemáticos

FIGURA 7.3. Flujograma del circuito de petición y realización de análisis



Conclusión: identificadas las posibles causas del retraso (bucles hacia atrás) diseñamos un estudio para medir cómo es de frecuente cada una y si se asocian o no al problema analizado (no tener a tiempo el análisis).

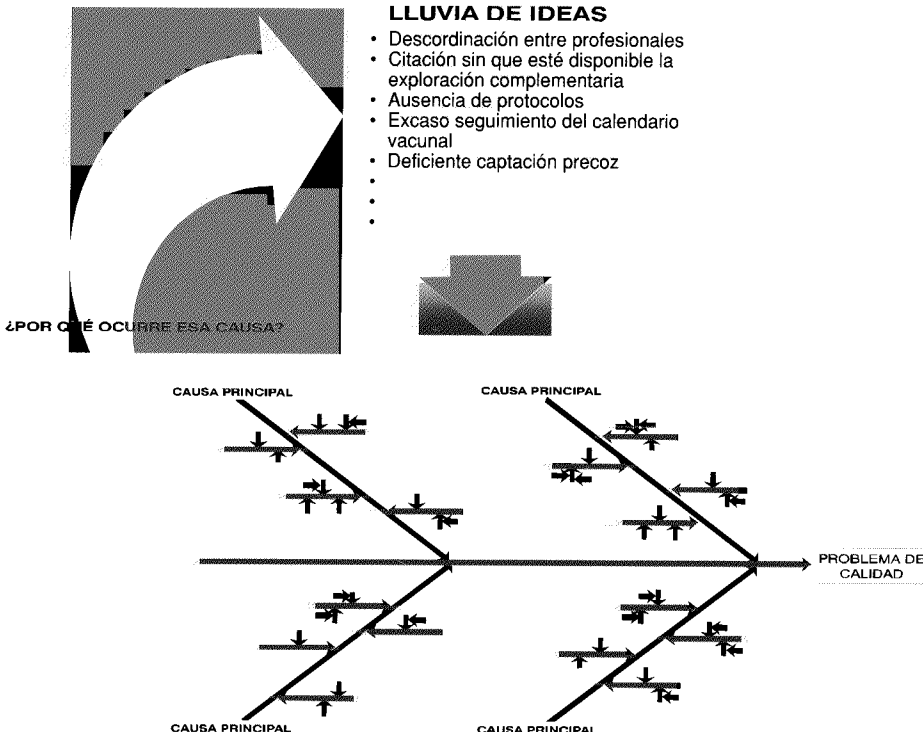
FIGURA 7.4. Diagrama de causa-efecto: causas de primer orden (C.1) de segundo orden (C.2) y de tercer orden (C.3)



El diagrama de causa-efecto explora de forma sistematizada y en sucesivas fases de profundización las distintas categorías o clases de causas que influyen en el problema analizado. Para realizarlo hay que seguir una serie de pasos establecidos, partiendo de la definición escueta y clara del problema a analizar.

Para elaborarlo se pueden seguir una serie de pasos en los que siempre hay que preguntarse ¿Por que ocurre la situación anterior? (problema, grupos de causas, causas de primer orden, causas de segundo orden...). Para ello se suelen utilizar técnicas que facilitan la creatividad como es el caso de la lluvia de ideas (Figura 7.5). Los pasos concretos para su elaboración son los siguientes:

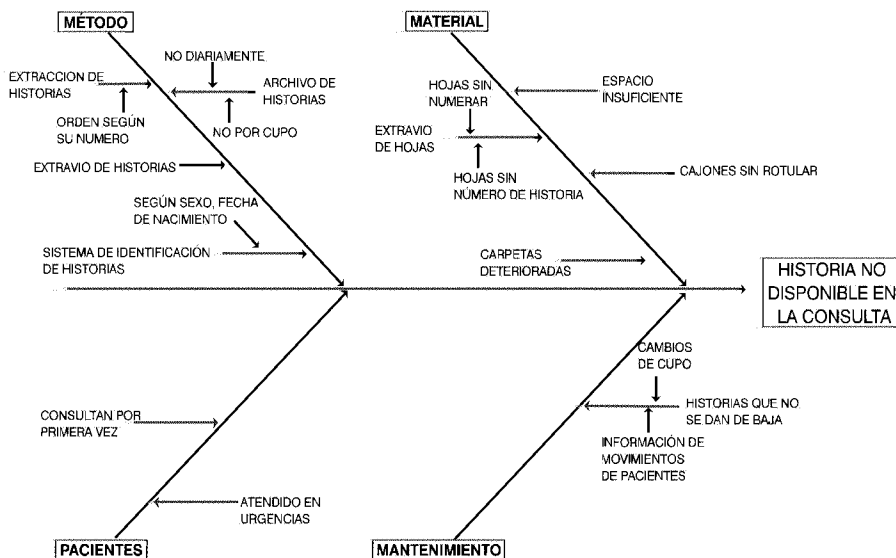
FIGURA 7.5. Identificación secuencial de las causas de un problema



1. **Anotación del problema:** se traza una flecha, centrada entre los márgenes superior e inferior, y se coloca el problema de calidad o efecto en su parte derecha. El problema es la "cabeza" de la espina. Debe estar definido de forma escueta y clara (por ejemplo: "seguimiento irregular del paciente hipertenso"; "captación deficiente"; "historia clínica no disponible en el momento de la consulta"; "exceso de pruebas complementarias", etc.); puede incluso ser cuantificado (por ejemplo: "el 30% de los pacientes no vienen a las consultas de seguimiento", "el número de hipertensos en programa es menos del 50% de los esperados", "el 10% de los pacientes tienen que ser vistos sin su historia", "el 60% de los hipertensos no tienen controlada su cifra tensional tras seis meses de tratamiento", etc.)
2. **Análisis de grupos de causas:** se determinan las causas principales del problema agrupadas en categorías genéricas o en número reducido de grupos de causas (de 4 a 6), como por ejemplo los métodos de trabajo, el personal, los recursos físicos (infraestructura), los pacientes, etc.. Cada una de ellas se representa como una flecha oblicua a la flecha central. La determinación de este grupo reducido de grandes apartados puede hacerse con discusión simple o con lluvia de ideas. Incluso puede postponerse y realizarse después de hacer el listado de causas de primer orden.
3. **Análisis de causas de primer orden:** mediante una lluvia de ideas se determinan las causas de cada una de las categorías genéricas; consiste en cuestionarse, para cada una de las categorías genéricas ¿Por qué sucede ésta?; por ejemplo ¿por qué los métodos de trabajo son causa del problema? Se representan como una flecha paralela a la flecha central y dirigida en uno de sus extremos hacia la flecha correspondiente a la causa genérica. Si se comienza la indagación de causas a este nivel, después de tener el listado hay que agruparlas en 4 a 6 grupos genéricos como vimos en el paso anterior.
4. **Análisis de causas de segundo, tercer y hasta "n" orden.** Se procede de forma análoga a la anterior. Para cada causa hay que plantearse ¿Por qué ocurre?, y esta a su vez ¿por qué?, hasta llegar a la más elemental. Las respuestas se representan como ramificaciones de esa causa.

Un ejemplo simple de diagrama de causa y efecto, aplicado al problema "historia clínica no disponible en el momento de la consulta" puede verse en la Figura 7.6.

FIGURA 7.6. Diagrama de causa-efecto aplicado en el análisis de los "pacientes atendidos sin historia clínica"



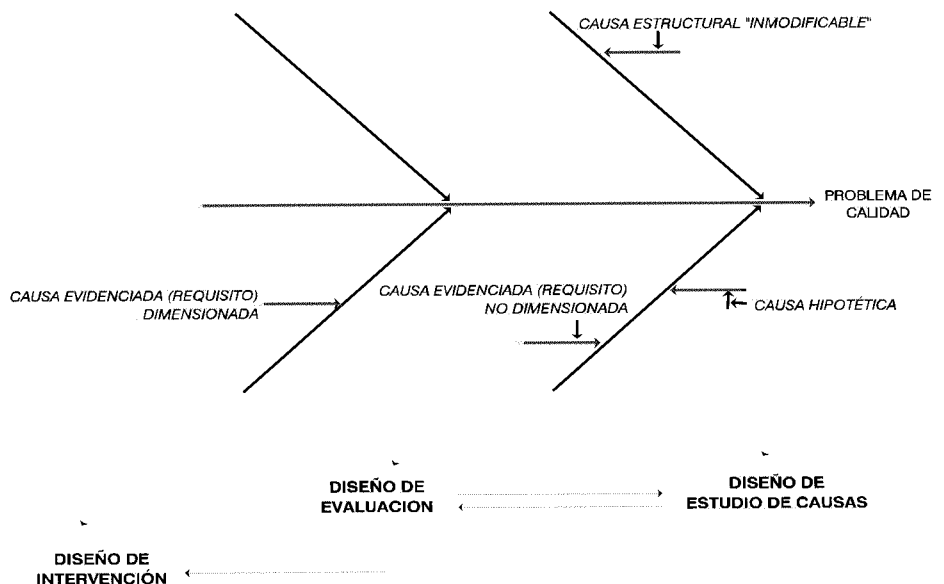
Un punto siempre controvertido a la hora de hacer un diagrama de causa y efecto es hasta qué nivel de causas llegar, es decir, cuando parar de interrogarse; otro problema que suele surgir es encontrarse causas que encajan en más de un grupo, ¿dónde colocarlas? No hay una norma inflexible ni, por tanto, respuesta única a estas cuestiones. En relación al *nivel* al que hay que llegar en la explicitación de causas, en teoría debe ser hasta que uno se quede sin respuestas; pero en la práctica se detiene el proceso de indagación cuando se ve que esta causa ya es medible y se entiende bien como puede influir en el problema analizado. En relación a la *ubicación* de causas que pueden estar en más de un grupo, tampoco hay regla inflexible; lo importante es que estén, que se hayan identificado: si se desea se pueden poner incluso en más de un lugar, aunque lo lógico sería elegir aquel grupo al que contribuya más, o sea más afín; por ejemplo, una causa como "no se indagan adecuadamente los hábitos alimenticios del paciente", que puede contribuir a fracasos en el control del riesgo cardiovascular, puede ser incluida en el grupo de "métodos" o en el de "personal", pero parece más claro que sea en el de "métodos"; otra cosa sería si la causa fuese "el personal desconoce la importancia de los factores dietéticos en el control del riesgo cardiovascular".

Una vez que se han terminado de identificar las causas potenciales del problema de calidad se pueden clasificar éstas, desde un punto de vista operativo, en las siguientes categorías (esquematisadas en la Figura 7.7):

- 1. Causas estructurales inmodificables.** Son factores de la producción de servicios: (recursos físicos y recursos humanos) sobre los que no podemos intervenir y que por tanto, en principio, limitan y condicionan la calidad. Es bueno, en todo caso, conocerlos y tenerlos presentes.
- 2. Causas con repercusión conocida en el problema analizado (requisitos de calidad).** Son factores en los que hay evidencia previa de la existencia de un nexo causal con el problema de calidad. A su vez se pueden distinguir dos tipos:

- 2.1. Frecuencia conocida (dimensionadas).** Son requisitos de calidad normalmente estructurales y de organización sobre los que disponemos de datos sobre la frecuencia con la que ocurren, que suele ser siempre o nunca, y sobre los que podemos intervenir directamente y mejorar la calidad. Por ejemplo la ausencia de normas, protocolos, registros, existencia de circuitos redundantes, etc..
- 2.2. Frecuencia desconocida (no dimensionadas).** Son requisitos de la calidad que no están cuantificados y deberán medirse para determinar el nivel de calidad. Lo más frecuente es que sean actividades de proceso que no sabemos hasta qué punto se realizan. Se siguen, en consecuencia con el diseño de evaluación, para saber su nivel de cumplimiento.
- 3. Causas de repercusión hipotética en el problema analizado.** Son factores en los que intuimos que existe una relación con el problema de calidad, pero sin que exista evidencia previa sobre esta relación. Su curso de acción consiguiente es el diseño de investigación de causas, para saber no sólo hasta qué punto están presentes o no, sino *también* hasta qué punto influyen o no en el problema de calidad analizado. Supongamos por ejemplo que nuestro problema es el seguimiento irregular del paciente hipertenso; para explicarlo podemos pensar en el cumplimiento de los criterios establecidos en cuanto a periodicidad de controles y solicitud de exploraciones complementarias, y éstos serían requisitos que iríamos a medir para saber su nivel de cumplimiento: serían causas con repercusión previamente conocida pero frecuencia desconocida (la situación 2.2 explicada más arriba); pero también podríamos pensar como hipótesis que los controles no se cumplen porque las analíticas se pierden, o el paciente cambia las citas con demasiada frecuencia, etc; todo esto son hipótesis plausibles pero que hay que medir y ver si en realidad ocurren y repercuten en los seguimientos irregulares (por ejemplo, un análisis que se pierda puede repetirse y no perjudicar en definitiva el seguimiento del paciente; los controles de citas igual, etc.).

Figura 7.7. Clasificación de las causas de un problema de calidad y vías de acción consiguientes



Los métodos de análisis de los problema de calidad que no precisan datos van a ser los que se van a utilizar con más frecuencia, sin embargo son complementarios con los métodos que utilizan datos, en los casos en que dispongamos de esta información.

5. HERRAMIENTAS DE ANÁLISIS QUE REQUIEREN DATOS SOBRE EL PROBLEMA

5.1. HISTOGRAMA

Cuando el problema a analizar esta expresando en términos de un solo indicador o criterio de tipo cuantitativo (es decir no mide presencia o ausencia de una cualidad sino cantidad de una determinada variable, como puede ser tiempo de espera), su resumen estadístico simple suele no ser suficiente y es conveniente realizar un análisis gráfico para saber qué hacer con el problema. Este análisis gráfico es el histograma tal como se estudia en los textos básicos de estadística, sólo que interpretándolo de una forma un tanto peculiar. El histograma, tal como se aprende en las primeras lecciones de cualquier curso de estadística, es un gráfico de barras en el que se muestra la distribución de un determinada variable cuantitativa a partir de los datos de una tabla de frecuencias de los intervalos o valores sucesivos en que se ha dividido la variable que se ha medido; por ejemplo, distribuimos el tiempo de espera para la consulta en intervalos de tiempo de 5 minutos, y vemos cuantos han esperado de 1 a 5 minutos, de 6 a 10 minutos, etc. Lo que se consigue al representarlo gráficamente es básicamente ver el *tipo de variabilidad* que presenta, lo cual aclara la situación y nos ayuda a decidir como proceder a continuación. Por ejemplo, si tenemos medido el tiempo de espera para la consulta, o los días que se tarda en recibir una determinada prueba analítica que se solicita, sólo con saber el tiempo medio o la media de días no es suficiente, porque desconocemos si estos tiempos medios representa a una distribución homogénea o no. Tomemos el caso del tiempo de espera (Figura 7.8), al hacer el histograma podemos encontrar que hay poca variación y que la media representa bien lo que la gente espera; en este caso, si el tiempo de espera es elevado, el problema es generalizado y hay que diseñar un estudio de causas hipotéticas sin distinción de grupos de pacientes. Sin embargo, si nos encontramos que es bimodal (con dos picos) o francamente sesgado a un lado, nos conviene identificar las características de los que esperan más (2º pico del histograma o cola extrema), porque son los que representan realmente el problema; o enfocarlo nuestro estudio a distinguir cuáles son las características diferenciales de los que esperan poco y los que esperan mucho (puesto que hemos visto que son como dos grupos diferentes), y en segundo lugar, si es necesario, cuáles son los factores adicionales que pueden asociarse a que se espere más.

Un histograma es la representación gráfica de un problema medido con un indicador cuantitativo, para averiguar la cantidad y tipo de variabilidad que presenta. Según estas características va a ser aconsejable enfocar el estudio de causas consiguiendo de una u otra manera.

FIGURA 7.8. Posibles histogramas del tiempo de espera antes de la consulta para decidir los estudios subsiguientes



Conclusión: problema general; estudio de causas sin distinción de grupos.



Conclusión: problema no general; estudiar causas de esperas extremas.



Conclusión: problema no general; estudiar causas de variabilidad; ver sobre todo lo que puede ocurrir a las esperas por encima de la media.



Conclusión: problema no general; estudiar características diferenciadas de los dos grupos.

Al igual que para las demás herramientas de análisis estructurado de los problemas de calidad, la construcción de un histograma se realiza siguiendo una serie de indicaciones y pasos determinados

Para elaborar un histograma podemos seguir los siguientes pasos:

1. Recopilar los datos del problema o causa potencial del problema en estudio (por ejemplo, tiempo de espera antes de la consulta).
2. Calcular el rango de la variable (restar al valor máximo el valor mínimo), y decidir sobre la amplitud del intervalo en que vamos a dividirlo. Supongamos que el rango resulta de 40', y que 5' nos parece un intervalo razonable.
3. Establecer el número de barras del histograma (debe resultar un mínimo de 6 y un máximo de 20) dividiendo el rango por el intervalo establecido. En nuestro ejemplo tendríamos un histograma con 8 barras (40'/5').
4. Determinar la frecuencia para cada intervalo. Es decir, ver cuántos pacientes han esperado un tiempo incluido en cada uno de los intervalos.
5. Representar en un gráfico con dos ejes, un eje vertical "y" (frecuencia de cada intervalo) y uno horizontal "x" (los intervalos de la variable) las barras correspondientes a cada intervalo.

Este gráfico reflejará el patrón de distribución de la variable analizada, así como la variabilidad de esta distribución, y nos indicará el mejor camino a seguir.

5.2. ESTRATIFICACIÓN

Una modalidad de analizar un problema de calidad consiste en presentar los datos de forma estratificada. En este caso, cuando tenemos los datos en conjunto de la oportunidad de mejora analizada (por ejemplo, tiempo transcurrido entre

la recepción y la petición de pruebas de laboratorio) podemos saber el grado de contribución de una determinada causa potencial (por ejemplo, dos laboratorios distintos) recogiendo los datos de forma desglosada para cada uno de ellos y elaborando los dos histogramas. La contribución de esta causa potencial (laboratorio al que se solicitan las pruebas) puede apreciarse por el grado en que se diferencian ambos histogramas. Naturalmente que ante dos histogramas muy diferentes decidiremos estudiar y/o intervenir sobre el laboratorio más problemático.

La estratificación puede realizarse también cuando el problema sobre el que tenemos datos no es cuantitativo sino cualitativo; supongamos, por ejemplo, que tuviésemos un servicio de radiología en nuestro centro y que el problema es que está colapsado, con listas de espera inaceptables; un posible acercamiento, si tenemos datos de los motivos de prescripción de radiografías, sería estratificar por estos motivos de prescripción para ver cuál es más frecuente; con este análisis por delante lo lógico sería ver si están claras las indicaciones de radiografías para los motivos más frecuentes, e intervenir directamente (para acordar suprimirlos) si aparecen motivos de solicitud de radiografía inadecuados; sin la estratificación podríamos haber estado tentados de estudiar la adecuación de la prescripción radiográfica en general, en vez de centrarnos en las causas de prescripción más frecuentes e intervenir directamente sobre las inadecuadas. Situaciones como ésta, en la que la estratificación es de gran ayuda, se producen con frecuencia si el problema a analizar es de utilización de recursos (por ejemplo, derivación a especialista → estratificar por especialista o incluso por tipo de enfermedad; utilización de analítica → estratificación por tipo de análisis; prescripción de antibióticos → estratificación por patologías y/o tipo de antibiótico, etc.) como veremos más adelante, la estratificación es uno de los métodos de gran utilidad en el análisis de los datos de un estudio de causas hipótéticas.

5.3. EL DIAGRAMA DE PARETO

Este gráfico se estudiará con detalle en la UT 11 como forma de análisis y presentación gráfica de los resultados de una evaluación. En general lo podemos aplicar para determinar en qué grado contribuyen las distintas causas y/o estratos de un problema en el mismo; en ambos casos lo que tenemos es un cierto número de categorías o de grupos y la frecuencia de su presencia o ausencia (variable cualitativa). Una vez medida la frecuencia con que se produce cada categoría, podemos ordenarlas de mayor a menor frecuencia y focalizar nuestra atención en las categorías más frecuentes. El diagrama de Pareto consiste básicamente en un diagrama de barras, en el que cada barra es una de las categorías, criterios, o estratos considerados, y en el que *las barras se ordenan de mayor a menor frecuencia*. De esta forma se hace gráficamente evidente qué es prioritario analizar. El ejemplo de la estratificación de las radiografías por tipo de patología o cualquiera otra estratificación por variables categóricas (variables con varias categorías o posibilidades excluyentes) pueden ser presentadas gráficamente con diagramas de Pareto.

Hay que advertir, no obstante, que el diagrama de Pareto se construye de una forma diferente y particular, y tiene utilidad más para decidir sobre diseño de intervención, que sobre estudios sucesivos a realizar.

La estratificación ayuda a entender la influencia de un determinado factor en la magnitud del problema y ayuda a priorizar los estudios de causas a realizar

Un diagrama de Pareto sirve para priorizar las categorías o causas principales sobre las que dirigir nuestra atención.

BIBLIOGRAFIA

- Banks NJ. Constructing Algorithm Flowcharts for Clinical Performance Measurement. *International Journal for Quality in Health Care* 1996; 8(4):395-400.
- Ishikawa K. Introducción al control de calidad [cap. 4]. Madrid: Diaz de Santos; 1994. p. 221-310.
- Manual de Heramientas básicas para el análisis de datos. Methuen: GOAL/QPC; 1990. p 9-13; 17-23; 25-29; 36-43.
- Leebov W, Ersoz CJ. *The Health Care Manager's Guide to Continuous Quality Improvement* Chicago: AHA; 1991.

0

**LA CONSTRUCCIÓN DE
CRITERIOS PARA
EVALUAR LA CALIDAD**

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

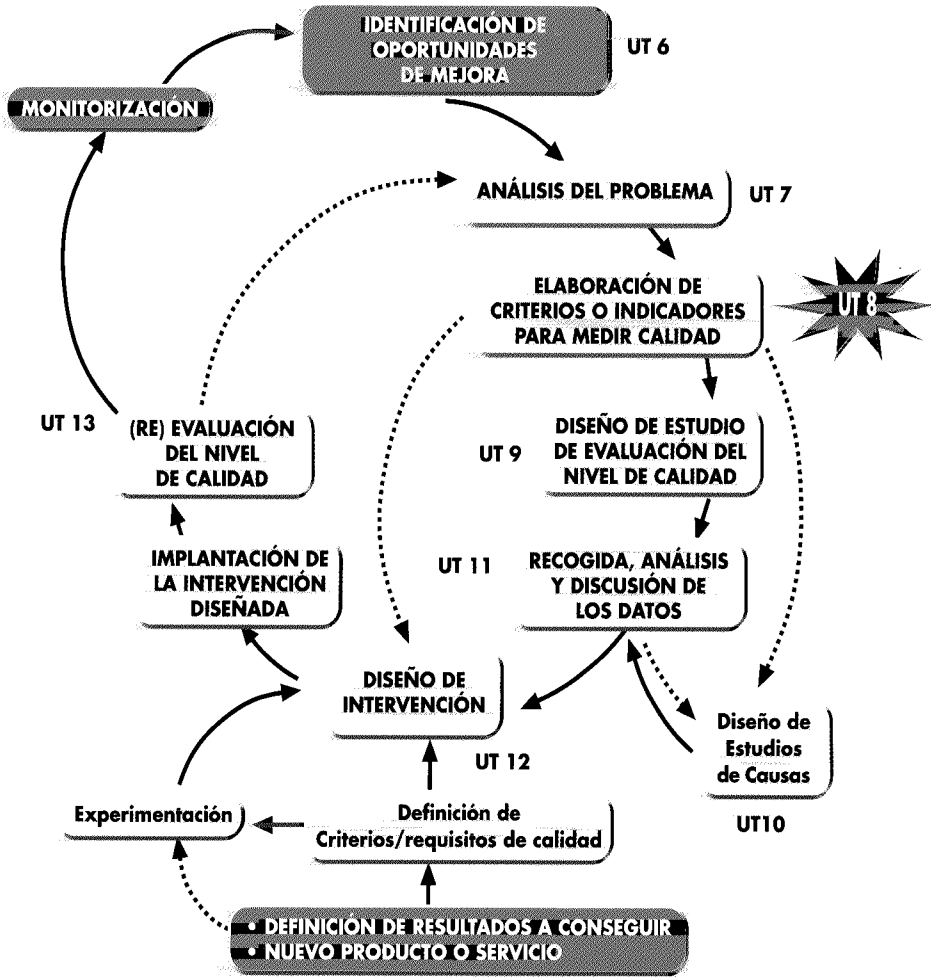
El siguiente paso en el Ciclo de Mejora, sea cuál sea el curso de acción que decidamos tras analizar el problema, es la construcción de los criterios, indicadores, herramientas o referentes con los que vamos a medir la calidad en la situación de partida, y que luego emplearemos para documentar la mejora conseguida. Aparte de otras consideraciones, los criterios o indicadores a emplear deben ser herramientas válidas y fiables, para lo cual es preciso abordar su construcción con un rigor metodológico que se detalla en esta UT. Junto al marco teórico-práctico general a tener en cuenta en la construcción de criterios e indicadores, esta UT ofrece un esquema sencillo y práctico para asegurarnos su fiabilidad.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. ¿Qué es un criterio?
3. Tipos de datos: estructura, proceso y resultados.
4. Tipos de criterios.
5. ¿Cómo se construye un criterio?
6. Excepciones y aclaraciones a los criterios.
7. Validez.
8. Fiabilidad.

OBJETIVOS ESPECÍFICOS

1. Desarrollar criterios o requisitos para medir la calidad de un determinado producto o servicio, o de un aspecto del mismo.
2. Distinguir los criterios aislados, transversales, de los secuenciales.
3. Distinguir los criterios empíricos de los normativos.
4. Discutir las ventajas y las desventajas de la ponderación de los criterios.
5. Establecer mecanismos para asegurarse la validez de los criterios a utilizar.
6. Describir los enunciados de los criterios acompañados de excepciones y aclaraciones de los términos empleados.
7. Explicar el significado de las excepciones al criterio a la hora de medir la calidad.
8. Determinar si las excepciones han de ser tratadas aparte o conjuntamente con los casos normales.
9. Estar alerta a la posibilidad de que las excepciones sean un factor de confusión si se tratan conjuntamente con los casos normales.
10. Diseñar un pilotaje para comprobar la fiabilidad de los criterios que se definan para evaluar la calidad.
11. Analizar la fiabilidad de los criterios.
12. Calcular e interpretar la concordancia o grado de acuerdo global entre evaluadores.
13. Calcular e interpretar el grado de acuerdo específico entre evaluadores.
14. Calcular e interpretar el índice Kappa para valorar el acuerdo entre parejas de evaluadores.
15. Analizar críticamente los valores del índice Kappa en función del número de evaluadores y la prevalencia del criterio que se mide.



■ : PUNTO DE PARTIDA
 → : VÍA DE ACCIÓN NORMAL
→ : VÍA CONTINGENTE O ALTERNATIVA

UT: Unidad Temática del curso

© Pedro J. Saturno

"Todo tiene que medirse, y lo que no se puede medir, hay que hacerlo medible"

(Galileo)

1. INTRODUCCION

Paso a paso hemos ido avanzando por el ciclo de mejora. Después de la identificación del problema de calidad (UT 6) y el análisis del problema seleccionado (UT 7), en esta unidad temática llegamos a una parte crucial, la construcción de criterios válidos y fiables con los que medir la calidad.

Vamos a ver que hay diferentes tipos de criterios con implicaciones metodológicas peculiares para su medición. Propondremos realizar las evaluaciones preferentemente con criterios explícitos y normativos y estableceremos las precauciones a tomar para asegurarnos que tengan validez y fiabilidad.

En esta UT se aborda la construcción de criterios válidos y fiables para evaluar la calidad

2. ¿QUÉ ES UN CRITERIO?

Un criterio es un referente o herramienta para medir la presencia, ausencia o diversos niveles de calidad.

Los criterios son las herramientas o referentes con las cuales se mide la calidad. Aspectos esenciales y relevantes de la atención sanitaria en torno al problema o área que queremos evaluar. Ese aspecto relevante de la atención sanitaria, lo mediremos para conocer su grado de cumplimiento, de forma que su ausencia o presencia o grado de cumplimiento podemos decir que nos está indicando a su vez, presencia o diferentes niveles de actuación correcta, es decir, diferentes niveles de calidad.

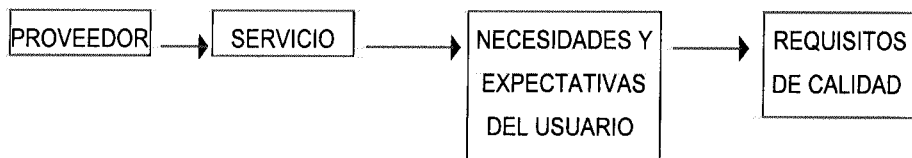
Un criterio se deriva de la reflexión sobre cómo han de ser los requisitos de los distintos servicios para que sean acordes con las necesidades y expectativas de quienes los van a recibir. En este sentido es a la vez un objetivo de calidad que queremos que se cumpla. De ahí la tremenda importancia que tiene el que sean relevantes, realistas y que puedan ser aceptados sin controversia por todos los profesionales implicados en su consecución. Una correcta elaboración y selección de criterios es probablemente, junto a una priorización adecuada del tema a evaluar, el principal factor determinante del éxito o fracaso de los ciclos de mejora.

Por ejemplo, supongamos que nos preocupe el nivel de calidad alcanzado en el programa de atención al cáncer en la mujer. No habría que discutir mucho para acordar que uno de los aspectos más relevantes para una atención correcta, de calidad, es la detección precoz. Así de una manera general la existencia o no de un programa de *screening* para el cáncer femenino es una medida de calidad.

Sin embargo, un criterio estructural como el que hemos definido rara vez es suficiente. Lo habitual es que haya que especificar además como se ha de llevar a cabo el proceso, cuál es la traducción operativa de la satisfacción de las necesidades de salud que queremos cubrir. Siguiendo con nuestro ejemplo, la detección precoz ha de consistir en la práctica de citologías vaginales con una pauta determinada, y realización también explícitamente pautada de exploraciones mamarias. El criterio o herramienta de medida sería en este caso el cumplimiento o incumplimiento de las pautas de actuación acordadas como concretas. De igual manera podríamos ir reflexionando sobre cuáles serían los requisitos para que no hubiera problemas de accesibilidad, satisfacción, etc. La Tabla 8.1 contiene el resumen de este esquema básico de razonamiento para definir criterios de calidad.

La definición de criterios de calidad parte de las necesidades y expectativas de los usuarios del servicio a analizar. Son requisitos que conducen a la satisfacción de estas necesidades y expectativas

TABLA 8.1. Ejemplo de aplicación del esquema de razonamiento básico para definir criterios para evaluar la calidad



| Centro de Salud | Servicio | Calidad científico-técnica | Requisitos de Calidad |
|--|-------------------------------|---|---|
| Hospital | Prevención de cáncer femenino | <ul style="list-style-type: none"> • <u>Diagnóstico precoz</u> • Prevención primaria. • Tratamiento adecuado. • Seguimiento adecuado. • . • . <p><u>Accesibilidad.</u></p> <ul style="list-style-type: none"> • Inexistencia de barreras culturales. • No hay lista de espera • horario conveniente. • . • . • . <p><u>Satisfacción</u></p> <ul style="list-style-type: none"> • Trato considerado. • Exploraciones suficientes y comprensibles. • . • . • . | <ul style="list-style-type: none"> - Existencia de programa de screening de cáncer. - Citologías realizadas según pautas determinadas. - Mamografías realizadas según pautas determinadas. - Calidad citológica adecuada. - Calidad adecuada de la exploración mamográfica. - . - . - . |
| <p>FORMULAR CON FORMATO DE CRITERIOS MEDIBLES, VÁLIDOS Y FIABLES</p> | | | |
| <p>BASE DE LOS REQUISITOS DE CALIDAD</p> | | | |

CADA NECESIDAD Y EXPECTATIVA PUEDE RELACIONARSE CON UNOS DETERMINADOS REQUISITOS DE CALIDAD

Los términos "criterio", "requisito", "norma", "estándar", no son siempre utilizados de la misma manera en la literatura sobre calidad asistencial, más bien son utilizados de forma intercambiable y a veces confusa. La distinción terminológica tiene una traducción práctica generalmente irrelevante. Lo importante para diseñar una evaluación no es tanto el nombre que le demos a la herramienta de medida, sino la secuencia consciente del razonamiento que nos lleva a utilizarla, y tener claro el concepto de lo que vamos a medir.

Con cualquiera de los servicios que consideremos como problema de calidad y queramos evaluar su nivel, el listado de criterios que podríamos construir es normalmente grande. Es conveniente, sin embargo, priorizar y prestar atención a un cierto número de características esenciales que distinguen los criterios buenos de los no tan buenos. Entre estas características que deben cumplir los crite-

Los criterios han de ser relevantes, realistas, aceptables, medibles, pocos, válidos y fiables.

rios (Tabla 8.2), ya hemos hablado de la *relevancia*, una característica inherente al concepto mismo de criterio que hace referencia a su importancia y a que sea un aspecto esencial del problema a evaluar. También hemos hablado de la conveniencia de que sean *realistas* y *aceptables*, en cuanto a que sea algo que el centro y sus pacientes se pueden permitir (realistas) y que los profesionales y la bibliografía los sustente como científicamente demostrados de calidad (aceptables y válidos). Hay que añadir a los requisitos que debe cumplir un criterio el que sea *medible*, se debe poder cuantificar el nivel de calidad. También es aconsejable que sean *pocos* criterios, los menos posibles, algunos autores dan menos de 10 como referencia útil. Otra característica fundamental de los criterios, que veremos más adelante, es que sean *fiables*. Fiabilidad y validez van a ser dos características de los criterios a los que vamos a prestar un interés especial.

TABLA 8.2. Características de los criterios para evaluar la calidad

- RELEVANTES
- REALISTAS
- ACEPTABLES
- VÁLIDOS
- MEDIBLES
- FIABLES
- POCOS

Si existen datos de estudios previos sobre un determinado problema objeto de evaluación, ello nos permitirá establecer unos estándares o niveles óptimos de aplicación del criterio, los cuales se fijaran antes de iniciar la evaluación, de acuerdo con la situación asistencial concreta del ámbito donde se realice. Normalmente, se expresan en forma de porcentaje. El "estándar" lo definimos como el nivel de cumplimiento patrón, que separa lo aceptable de lo no aceptable. Es un límite que establecemos como nivel mínimo exigible de cumplimiento del criterio. Aunque es aconsejable establecerlo antes de evaluar tal como hemos señalado, tampoco es incorrecto establecerlo después de una primera evaluación.

3. TIPOS DE DATOS: ESTRUCTURA, PROCESO Y RESULTADO

Cuando medimos la calidad, los criterios a utilizar van a estar constituidos por datos de *estructura*, *proceso*, *resultado*, o una combinación de ellos. Esta distinción, propuesta por Donabedian y profusamente utilizada en el sector salud, no forma parte de la terminología utilizada normalmente en sectores de la industria pero reviste un gran interés por las diferentes precauciones que hay que tomar en cuanto a su *validez*, y para tener argumentos en las discusiones que periódicamente aparecen dando preferencia a uno u otro tipo de datos. Vamos a aclarar en primer lugar qué se entiende por estructura, proceso y resultado de los servi-

El estándar (nivel de cumplimiento patrón) de los criterios puede fijarse antes o después de evaluar.

Los datos pueden ser de estructura, proceso y resultado.

cios de salud; y después reflexionaremos sobre las ventajas, inconvenientes y precauciones en cuanto a validez que hay que tomar según el tipo de datos que utilizemos.

- **Estructura** se refiere a todo lo que tenemos para la provisión de servicios - recursos físicos, equipamiento, financiación, personal (número y cualificación); y la forma en que están organizados. En resumen, estructura es lo que hay.
- **Proceso** es todo lo que el proveedor de los servicios de salud hace para el usuario, y también lo que el usuario hace como consecuencia de las actividades de los proveedores. Los procesos utilizan la estructura de una determinada manera, es lo que se hace.
- **Resultado** es lo que se consigue con los procesos, es decir, la incidencia en la función física, satisfacción y nivel de salud en sentido amplio, de los servicios que se ofrecen. Resultado es lo que acontece al ejecutar los procesos.

Para evaluar la calidad de cualquier servicio, va a ser posible identificar una serie de criterios o requisitos que, además de responder a una o más de las dimensiones que vimos en el nivel de concepto general, van a estar constituidas por datos de estructura, proceso y resultado, y es conveniente que lo consideremos explícitamente. ¿Por qué?, para reflexionar a continuación sobre si su medición constituye o no una medición *válida* de la calidad; es decir si su presencia es equivalente de presencia de calidad en el servicio evaluado, y sus diferentes valores o niveles son sinónimo también de diferentes niveles de calidad. Los posibles problemas de validez a considerar son diferentes según que los datos sean de estructura, proceso o resultado, y vamos a resumirlos a continuación.

- **Los criterios de estructura** son condición necesaria pero no suficiente para afirmar que el servicio es de calidad. Su presencia nos dice que la calidad es *posible*, pero no podemos concluir que *realmente* ocurra. Si están bien elegidos y encontramos que no se cumplen al evaluar su presencia, podemos decir que la calidad no es posible o va a estar probablemente dañada; pero no podemos concluir que el servicio es de calidad si encontramos que se cumplen. Por ejemplo, un criterio estructural para el programa de hipertensión es contar con esfignomanómetros calibrados; otro puede ser el que estén bien definidos los circuitos y la organización del equipo de médicos y enfermeras para las tareas de detección de hipertensos; así podríamos ir identificando una serie de criterios estructurales de relevancia específica para el servicio de atención al hipertenso; si los medimos ¿podemos concluir que el servicio es de calidad? Obviamente no; lo que podemos decir es que existen las condiciones para que el servicio sea de calidad, pero no podemos garantizar que los esfignomanómetros se utilicen adecuadamente, o que cada miembro del equipo ejecute correctamente y a tiempo las tareas que le correspondan.
- **Los criterios de proceso** pueden ser medidas válidas de la calidad solo si estamos seguros que su realización o no repercute en un resultado diferente y relacionado con aspectos que importan a las necesidades y expectativas de quien recibe el servicio. En relación a la calidad científico-técnica, solo serían criterios *válidos*, aquellos sobre los que hubiese evidencia científica de su contribución a la solución o control del problema de salud a que se refiere el servicio a evaluar. A veces se utilizan como válidos criterios sobre los que hay un amplio consenso sobre su importancia, o simplemente aquellos que el equipo de profesionales acuerda que son los correctos; en estos casos, puede haber

La identificación de los datos como datos de estructura, proceso o resultado es importante para reflexionar sobre su validez para medir calidad.

Los datos de estructura no miden la calidad, solo indican si es posible o no.

Los datos de proceso son válidos si se ha probado su relación con el resultado.

problemas de validez. Es conveniente que exista algún tipo de comprobación o validación previa (externa o interna) para aceptar un criterio de proceso como válido. Ejemplos de criterios de proceso en la atención al hipertenso serían los relacionados con la forma en que realizamos la captación de casos, realizamos el diagnóstico, el estudio inicial, seguimiento, etc. Los criterios de proceso son profusamente utilizados en las mediciones de la calidad; son generalmente accesibles para su medición y muy útiles para saber qué hacer para mejorar, pero hay que estar muy atentos para no incluir criterios irrelevantes, de validez dudosa (efecto no sabido o comprobado previamente).

- **Los criterios de resultado** son probablemente los más atractivos. ¡Qué mejor criterio de calidad que verificar que se consigue mejorar la salud y la satisfacción de los que reciben nuestros servicios! Sin embargo, a la hora de definirlos como referentes o herramientas de medición hay que tener dos precauciones en relación a su validez: (i) asegurarnos que son aspectos sobre los que influye la manera en que se hagan las cosas, es decir el *proceso*; y (ii) asegurarnos que tenemos en cuenta para su medición e interpretación los *factores de confusión*, cuestiones generalmente características de los usuarios de los servicios, que influyen *por sí solas* en el resultado, con independencia de que hagamos mejor o peor las cosas. Un ejemplo de la primera precaución y de un criterio de resultado no válido por no estar normalmente influido por el proceso sería el de curación de determinadas enfermedades virales, por ejemplo el resfriado común o la gripe; ejemplos de factores de confusión que pueden explicar variaciones en el resultado no atribuibles a la calidad del servicio serían la edad, la comorbilidad, el estado de salud en general de los usuarios, etc., de igual forma el que un hipertenso se controle o no puede depender, además de que lo hagamos bien, de múltiples factores (edad, tipo de hipertensión, enfermedades concomitantes, etc) que habrá que tener en cuenta al medir la calidad en base a este criterio de resultado.

El resumen es que, según el tipo de datos que se utilicen los problemas de validez pueden ser de diferente origen, tal como se resume en la Tabla 8.3. Lo cual no quiere decir que haya necesariamente que primar un tipo de datos sobre otro; los tres pueden ser útiles, según el servicio o el aspecto del servicio cuya calidad estamos interesados en medir. La Tabla 8.3 incluye también una valoración sobre la facilidad con que normalmente se puede obtener cada tipo de datos, y la utilidad inmediata de su conocimiento para saber qué hacer para mejorar.

TABLA 8.3. Validez, facilidad de obtención y utilidad inmediata de los diferentes tipos de datos

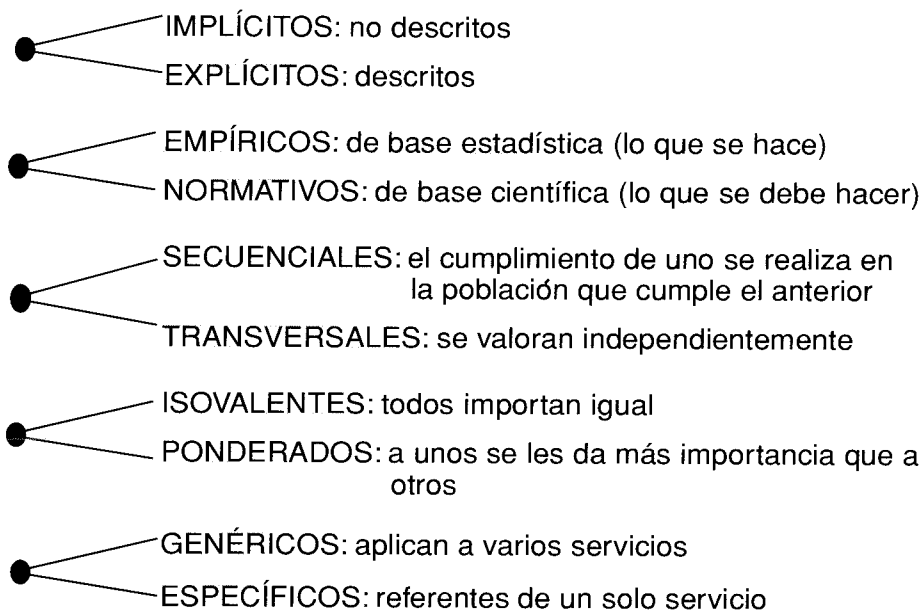
| TIPO | VALIDEZ | FACILIDAD DE OBTENCIÓN | UTILIDAD INMEDIATA |
|------------|---|------------------------|--------------------|
| ESTRUCTURA | ES CONDICIÓN NECESARIA PERO NO SUFICIENTE | +++ | +++ |
| PROCESO | DEPENDE DE LA RELACIÓN CON EL RESULTADO | ++ | +++ |
| RESULTADO | -DEPENDE DE LA RELACIÓN CON EL PROCESO -ES PRECISO AJUSTAR PARA ELIMINAR FACTORES DE CONFUSIÓN | + | + |

Los datos de resultado son válidos si se ha probado su relación con el proceso. Pueden además estar sujetos a factores de confusión que hay que medir.

4. TIPOS DE CRITERIOS

Al hablar de criterios de calidad podemos encontrar diferentes tipos (Tabla 8.4):

TABLA 8.4. Tipos de criterios



4.1. SEGÚN ESTÉN O NO ESPECIFICADOS

- **Criterios implícitos.** Son aquellos en los que no se especifica cuales son los criterios de calidad, van en función de los conocimientos del sujeto que lo realiza. No es recomendable utilizar este tipo de criterio de forma exclusiva. Pueden emplearse más como mecanismo de detección de problemas o de segunda revisión de los casos clasificados de calidad deficiente.

Ejemplo de criterio implícito: Evaluar si el diagnóstico es correcto sin que se especifique cuál es el concepto de buen diagnóstico que emplea el evaluador.

- **Criterios explícitos.** Son aquellos en los que se describe en qué consisten. Se aconseja utilizar este tipo de criterio. Permite tener claro cuáles han sido los referentes para la medición, que puede ser criticados o justificados, si es preciso. Ejemplo de criterio explícito: la toma de TA se realizará con esfigmomanómetro de mercurio, con el paciente en reposo y sin que haya fumado ni tomado café en 30 minutos previos a la medición.

4.2. SEGÚN EL ORIGEN

- **Criterios empíricos.** No dicen como tienen que ser las cosas para ser de calidad. Describen como es la realidad y se acepta como válido lo más frecuente. Por ejemplo, dar como referente de calidad de la prescripción la media de

prescripciones de utilidad baja de una determinada área o Comunidad Autónoma, o dar como referente de estancia media adecuada para una patología la media de los hospitales de una región.

- **Criterios normativos.** Dicen explícitamente como tienen que ser las cosas para que sean de calidad aceptable. Son generalmente los más adecuados.

4.3. SEGÚN SU DEPENDENCIA O NO DE OTROS CRITERIOS

- **Longitudinales, secuenciales o encadenados.** El criterio secuencial es aquel que precisa para su evaluación de los datos de cumplimiento de otro criterio anterior. Se evalúa después del precedente. Su distinción es importante por las implicaciones que tienen para su medición; más concretamente para los esquemas de muestreo, al tener que asegurarnos un número mínimo de casos para poder evaluar todos los criterios, tal como veremos en la UT 9.

Por ejemplo, para evaluar el *screening* de cáncer de cérvix podemos tener lo siguiente: 1^{er} criterio. Toda mujer entre 35 y 65 años deberá tener realizada la citología y registrado el resultado (normal o no) en la historia clínica. 2^o criterio. Si la citología es patológica, la mujer deberá ser derivada al ginecólogo de zona.

Para que se pueda cumplir el 2^o criterio previamente se ha de cumplir el 1^o, por tanto, el 2^o criterio es un criterio *secuencial* o *encadenado* al 1^o.

- **Transversales.** Se evalúan todos de forma independiente el uno del otro.

4.4. SEGÚN SU PONDERACIÓN

- **Isovalentes.** A todos los criterios se les da el mismo valor como requisitos de calidad.
- **Ponderados.** Al resumir la evaluación se le da a algunos criterios más valor que a otros.

Por ejemplo: Criterio 1: A todo paciente hipertenso se le realizará un ECG en el año del diagnóstico. Criterio 2: A todo paciente hipertenso se le realizará anamnesis sobre antecedentes familiares de HTA, cardiopatía, y dislipemia. Criterio 3: Todo paciente hipertenso se considera bien controlado si su TA \leq 140/90 en los tres últimos controles.

Si nosotros a la hora de evaluar decimos que el criterio 3, vale el doble que el 1 y el 2, estamos ante un criterio ponderado; si por el contrario todos nos parecen igual de importantes, serían criterios isovalentes. La forma más frecuente de ponderación es, al aplicar un listado cerrado de criterios, hacer que todos sumen 100 ó la unidad y ponderar cada uno en función del % o proporción que se supone importa sobre la valoración global que se haga.

La ponderación suele ser arbitraria y discutible, por lo que en general es preferible evitarla.

4.5. SEGÚN EL NÚMERO DE SERVICIOS A LOS QUE INCLUYEN

- **Genéricos.** Incluyen a varios servicios diferentes.
- **Específicos.** Sólo afectan a un servicio.

Esta distinción es semejante a la que vimos a la hora de hablar de problemas de calidad en la UT 6. De hecho, si el problema es genérico, va a tener que ser evaluado con toda probabilidad con criterios genéricos. Algunos ejemplos: "No deben existir reacciones adversas a medicamentos" (o con un cierto estándar), sería criterio genérico de calidad de prescripción; si lo referimos a un tipo de enfermedad concreta sería específico: "no debe haber reacciones alérgicas a la penicilina en el tratamiento de la amigdalitis".

5. ¿COMO SE CONSTRUYE UN CRITERIO?

El proceso de elaboración y selección de criterios es algo que atañe a todos los implicados en su cumplimiento. Todo el personal sanitario debe participar en la selección de criterios, en la medida en que están involucrados en el servicio a evaluar, y en particular en el cumplimiento de los criterios que se decida utilizar en la evaluación.

En instituciones grandes no es operativo que todo el personal participe en la selección de criterios y puede delegarse al grupo de mejora su selección inicial. Para ello pueden utilizarse técnicas de toma de decisiones en grupo. Se ahorra tiempo si se elabora un borrador previo sobre los criterios que hay que discutir y luego se consensuan con el resto de profesionales.

A continuación, se sugieren algunas recomendaciones para la elaboración de criterios:

- Manejar criterios simples.
- Limitar el número de criterios (el menor número posible. Máximo 10 criterios).
- Incluir solo los mínimos parámetros esenciales.
- Elaborar criterios que se adapten a los recursos de los profesionales y el entorno.
- Asegurarse el consenso y la aceptabilidad por parte de quienes los tienen que cumplir.
- Asegurarse de que el contenido esté actualizado y sea válido.
- Prestar atención cuidadosa a las excepciones y aclaraciones de los términos empleados.
- Asegurarse la fiabilidad mediante pilotaje, sobre todo si va a haber más de un evaluador.

Previamente a la elaboración de criterios hay que revisar y actualizar la bibliografía existente, consensuar con el grupo de trabajo, ver si son factibles y realizables en nuestro medio y con nuestras características peculiares. Copiar criterios ya elaborados sin tener en cuenta estas reflexiones es erróneo, y puede resultar contraproducente para el ciclo de mejora en el que se quieran emplear.

Los criterios pueden clasificarse en función a diversas características. De todas ellas destacamos como más adecuados los criterios explícitos, normativos, isovalentes y específicos.

La elaboración o adaptación de criterios es conveniente que sea participativa y asegurándose su idoneidad por el equipo que los va a utilizar, así como su validez y fiabilidad.

Si hemos de hacer referencia a las clasificaciones de criterios que vimos en el apartado 4, en la elaboración de criterios debemos tender a que sean *explícitos, normativos, isovalentes y específicos*, que son los que resultan en general más útiles y entendibles para iniciar actividades de mejora de la calidad.

6. EXCEPCIONES Y ACLARACIONES A LOS CRITERIOS

Una primera y elemental precaución para que los criterios sean válidos y fiables (es decir que sean interpretados de igual manera por todos los evaluadores) es definir lo más exhaustivamente posible las excepciones y aclaraciones que los acompañan.

Hay circunstancias en las cuales no es exigible el cumplimiento del criterio, a estas circunstancias se las llama *excepciones* al criterio. Por ejemplo, si el criterio es "los pacientes con otitis media aguda deberán ser tratados con amoxicilina clavulánico en 1ª instancia", una clara excepción sería pacientes alérgicos a la penicilina. Está claro que aunque la amoxicilina es el fármaco de elección, no debe darse a un paciente alérgico. Otro ejemplo frecuente son las excepciones a los criterios de diagnóstico, en caso de que el diagnóstico sea antiguo o haya sido realizado en otro nivel asistencial.

Las excepciones al criterio, no deben ser muchas ni frecuentes y deben de tratarse con cuidado a la hora de evaluar la calidad, ya que pueden ser un factor de confusión para las comparaciones; sobre todo si las consideraciones como cumplimiento del criterio, como veremos más adelante.

6.1. FRECUENCIA DE LAS EXCEPCIONES

Si a la hora de evaluar un criterio, vemos que nos aparecen muchas excepciones debemos eliminar ese criterio, ya que corremos el riesgo de que no tenga validez, como criterio de calidad o que dificulte la evaluación exigiendo diseños muestrales complicados. Por ejemplo, un criterio puede ser, "Debe constar en la H.C. de toda gestante que se da consejo sobre tabaco en el embarazo", excepción, "paciente en cuya historia consta que no fuma".

A la hora de evaluar, el número de veces en que se da consejo a una gestante sobre tabaco, si son muchos los casos en que aparecen excepciones, no sabremos realmente si se da el consejo en un nº suficiente de casos (porque tenemos una muestra menor, solamente aquellas que realmente fuman). Habría probablemente que rediseñar el criterio dirigiéndolo solo a las que fuman o encadenándolo a la exploración sobre hábito tabáquico.

Si tenemos muchas excepciones corremos el riesgo de terminar no midiendo aquello para lo que diseñamos el criterio.

6.2. LAS EXCEPCIONES ¿SON CUMPLIMIENTOS O INCUMPLIMIENTOS DE LOS CRITERIOS?

Otro problema que plantean las excepciones es *la forma de valorarlas en la evaluación*. Si a la hora de la evaluación, las excepciones que aparecen las

Los criterios deben ir acompañados de excepciones y aclaraciones para contribuir a la validez y fiabilidad de las mediciones.

Un criterio con muchas excepciones cuando se mide no es un buen criterio.

consideramos como cumplimiento del criterio, estamos perdiendo la información "real" de lo que se está haciendo (no estamos midiendo aquello que queríamos medir). Volviendo a nuestro ejemplo del consejo antitabáquico en la embarazada, si evaluamos una HC de una paciente gestante que no fuma (excepción), y lo consideramos como cumplimiento del criterio, la evaluamos igual que si fuera una gestante fumadora a la que se dio consejo antitabáquico. En el primer caso no estamos valorando el cumplimiento del criterio al dar consejo antitabáquico a la gestante (no estamos midiendo aquello que queríamos medir). Por otra parte si no consideramos a las excepciones como cumplimientos y les damos el mismo tratamiento que si no cumple el criterio, tampoco refleja la realidad de los hechos, ni nos deja ninguna opción de mejora. En nuestro ejemplo si en esas mujeres que no fuman decimos que no cumplen el criterio tampoco estamos reflejando fielmente la realidad, porque seguramente no se les dio consejo por no ser del grupo de riesgo.

¿Qué decisión adoptar? Lo que parece más lógico en general es considerarlas como cumplimiento de los criterios; sin embargo, lo más recomendable es considerarlas *aparte*. De no hacerlo así, pueden ser un importante *factor de confusión* para las comparaciones entre evaluaciones y entre centros; piénsese que un cumplimiento alto de un criterio podría ser debido a una proporción alta de excepciones, más que a la realización concreta del criterio en sí; podría incluso darse la paradoja que en los pocos casos *reales* del criterio a medir no se hiciera correctamente, mientras que en otro centro con menos excepciones el cumplimiento *real* fuese mayor, pero aparecer disminuido por la baja proporción de excepciones. Un ejemplo claro, frecuente y real es comparar la calidad de los diagnósticos de hipertensión, considerando como cumplimiento de los criterios si el diagnóstico es antiguo o se ha realizado en otro nivel; la diversa proporción de estas excepciones entre los diferentes centros a comparar, confunde claramente el criterio de calidad que queremos medir, si las consideramos como cumplimiento.

Las excepciones es mejor, en general, considerarlas como casos aparte, no como cumplimientos del criterio. De lo contrario pueden ser un importante factor de confusión al comparar mediciones.

6.3. IMPORTANCIA DE LAS ACLARACIONES

Se entiende por *aclaración al criterio* la descripción inequívoca de cualquier término que aparezca en el criterio o sus excepciones, que pueda estar sujeto a interpretaciones. Para garantizar el empleo exacto y repetible de los términos técnicos, hay que definirlos explícitamente. Hay que dar una definición lo suficientemente detallada para que la persona que evalúa no tenga que hacer ningún juicio personal sobre si un caso cumple o no un criterio determinado. Por ejemplo, un criterio sería: "los pacientes con urocultivo positivo deberán ser tratados en los tres días posteriores al cultivo. Excepción: "paciente tratado en la visita inicial". *Aclaración*: "urocultivo positivo significa 100.000 o más colonias/cc registrado en el informe del cultivo".

Veamos otros ejemplos de criterios utilizados en la realidad para el uso de la hipertensión:

- **Criterio de diagnóstico.** "Haber sido diagnosticado mediante tres tomas de tensión arterial (TA), separadas entre sí un período máximo de tres meses, cuyo promedio sea igual o mayor a 140/90 mm de Hg". Excepción: paciente con diagnóstico previo y documentado o paciente menor de 18 años. Si el

Las aclaraciones correctas y exhaustivas evitan que el profesional que evalúa tenga que "interpretar" el criterio. De este modo, las aclaraciones contribuyen a que éste sea fiable

criterio se utiliza tal cual y no hay aclaraciones, ¿cuál es el período mínimo entre tomas de TA? Si dos TA son tomadas el mismo día en diferentes posiciones o tras reposo ¿se considera una sola toma de tensión o dos?; ¿tienen que haber sido tomadas con el mismo aparato?; ¿sirven las TA que aporta el enfermo de su domicilio o de la farmacia?; ¿qué se entiende por diagnóstico previo y documentado?; ¿es suficiente con que el enfermo lo refiera?; ¿es suficiente con que tome antihipertensivos?; ¿se tienen en cuenta fármacos, situaciones estresantes o dietas temporales que pudieran haber influido en el diagnóstico? Este sería, pues, un criterio insuficientemente aclarado, y que se presta por tanto a múltiples interpretaciones.

- **Criterio de exploración física.** "Tener realizada una auscultación cardiopulmonar y pulsos periféricos". Al no haber aclaraciones, este criterio también puede ser problemático. Si sólo tiene realizada la auscultación y no los pulsos ¿cumple el criterio?; ¿quién puede realizar esa auscultación?; ¿sirve si está en algún informe de especializada?; ¿o tiene que ser realizada por médicos de primaria?; ¿es suficiente poner normal o un símbolo alusivo o hay que detallarla?; ¿es suficiente con una localización?
- **Criterio de seguimiento.** "Tener realizado e informado un ECG en los dos últimos años. Excepción: pacientes inmovilizados". Sin aclaraciones, no está lo suficientemente claro: ¿Que se entiende por ECG informado?; ¿es suficiente poner normal o patológico?; ¿sirve la descripción de un ECG realizado en otro nivel asistencial?

La existencia o no de aclaraciones, cuando éstas son necesarias, y la suficiencia de las mismas repercute directamente en la *fiabilidad* de los criterios. La fiabilidad podemos definirla como el grado de reproducibilidad de los resultados cuando el criterio es utilizado por observadores diferentes. Criterios sujetos a interpretación diversa por distintos evaluadores no van a poder ser utilizados de forma fiable para comparar evaluaciones ni dentro de los centros ni entre centros.

6.4. INCLUSIÓN DE LAS EXCEPCIONES Y ACLARACIONES EN EL DISEÑO DE LOS CRITERIOS

Hemos visto que las excepciones pueden influir en la fiabilidad y en la validez de los criterios como factores de confusión y que las aclaraciones influyen en la fiabilidad. Validez y fiabilidad son dos atributos importantísimos para los criterios a medir. En consecuencia, aparte de que consideremos adicionalmente tanto validez como fiabilidad, es muy aconsejable utilizar un formato para la descripción de criterios que incluyan explícitamente excepciones y aclaraciones. Caso de no ser necesarias se consignará así, pero es conveniente asegurarse de que, al menos, reflexionamos sobre la necesidad de constatarlas. Este formato de descripción de criterios que incluye excepciones y aclaraciones es el que se describe en la Tabla 8.5. Si se utiliza, es más probable que construyamos criterios válidos y fiables, pero es conveniente considerar aparte y explícitamente estas dos características antes de dar por aceptado el criterio.

Es conveniente presentar los criterios con un formato que incluya excepciones y aclaraciones.

TABLA 8.5. Formato para la descripción de criterios para evaluar

| CRITERIO | EXCEPCIONES | ACLARACIONES |
|---|--|--|
| Descripción del referente a utilizar como medición (Contenido del criterio) | Especificación de las situaciones en las que es aceptable que el criterio no se cumpla | Definición detallada de cualquier término del criterio o las excepciones que pueda ser interpretado de más de una forma o a la hora de evaluar |

7. VALIDEZ

Un criterio válido es aquel que mide aquello que está pensado que mida; es decir, aquel que mide realmente la calidad del servicio evaluado, de forma que su presencia es equivalente a presencia de calidad, y sus diferentes valores o niveles son sinónimos también de diferentes niveles de calidad. En general, para ser válido, debe estar clara su influencia en el resultado de la atención (cumplimiento de expectativas y necesidades de los usuarios del servicio) y deben ser tenidos en cuenta los posibles factores de confusión (que hacen variar por sí solos el cumplimiento del criterio, sin que responda esta variación a actuaciones realmente diferentes).

La validez es, probablemente, el principal atributo exigible a un criterio. Para evaluar la calidad, no se trata de tener muchas medidas de validez y fiabilidad dudosa, sino de disponer de medidas clave válidas y fiables. La validez es un concepto ligado al error sistemático: si el criterio no es válido vamos a estar sistemáticamente midiendo una cosa diferente a la prevista.

En los textos sobre psicometría y métodos cuantitativos para las investigaciones sociales en general, aparece el concepto de validez seguido de diversos apellidos que lo descomponen en varios aspectos: validez facial (aparente o lógica), validez de contenido, validez de construcción, validez de consenso y validez de criterio. Para la construcción y análisis de criterios para evaluar calidad, nos interesa sobre todo conocer y prestar atención a la validez facial, de contenido y de criterio; fundamentalmente esta última. Vamos a ver en qué consiste este análisis de la validez y como podemos asegurarnos su presencia al construir o adoptar criterios para evaluar la calidad.

La validez es el principal atributo exigible a los criterios. El análisis de su presencia en un criterio puede realizarse considerando diversos tipos de validez de entre los que vamos a destacar la validez facial, de contenido y de criterio.

7.1. VALIDEZ FACIAL, DE CONTENIDO Y DE CRITERIO

- **La validez facial, lógica o aparente**, hace referencia a que el criterio parezca importante para el problema a evaluar sin que sea preciso dar muchas explicaciones; es su grado de relevancia obvia. No es un aspecto de la validez que determine la validez real, pero si es una buena guía para priorizar y excluir criterios obviamente no válidos. Los criterios tienen que *tener que ver* con el problema de calidad a evaluar. Por ejemplo, el estado vacunal o los antecedentes de cáncer en la familia no parece que sean relevantes como criterios de calidad para la anamnesis en el hipertenso.
- **La validez de contenido**, hace referencia a que el criterio mida el concepto que se quiere medir; en nuestro caso, sería asegurarnos que está en relación

La validez facial o aparente hace referencia a la relevancia del criterio a primera vista. Debe ser obvia su importancia sin demasiadas explicaciones

con la calidad; es decir, que se refiere a algunas de las dimensiones con las que hemos definido calidad, y se relaciona con las necesidades y expectativas de los usuarios. La valoración de este tipo de validez no dejará de ser, sin embargo, de tipo fundamentalmente subjetivo, igual que la de la validez facial, pero también nos puede servir para descartar criterios no claramente relacionados con el concepto o aspecto de calidad que queramos medir. Por ejemplo, si no hemos incluido la satisfacción del profesional como una de las dimensiones del servicio que se ofrece al cliente externo, un criterio que la mida no tiene validez de contenido para evaluar la calidad del servicio que recibe el usuario externo. De igual manera, si queremos medir la calidad científico-técnica, no serán válidos criterios que haga referencia a aspectos estructurales que tengan relación básicamente con la accesibilidad.

- **La validez de criterio**, es el grado en que la variable escogida se correlaciona con un criterio de referencia objetivo, fiable y que esté ampliamente aceptado como una buena medida del fenómeno de interés. En nuestro caso sería la relación probada con un mejor resultado, si es un criterio de proceso, con diferencias en el proceso si es un criterio de resultado y con factibilidad del proceso, si es un criterio de estructura. En ausencia de una medida de referencia, valorar una adecuada validez es un tema complejo. Muchas veces resulta imposible usar este abordaje por falta de estudios que lo avalen. Sin embargo, es de una gran importancia y debemos procurar siempre reflexionar sobre ella. La forma de realizarlo es *evaluando la evidencia científica* que justifica los criterios. Lo ideal es que pueda estar especificado el método empleado para revisar las evidencias científicas cuantitativas en las que se fundamenta; y que consten las fuentes de información utilizadas. Para evaluar el nivel de evidencia científica se ha desarrollado el siguiente sistema de graduación:

1. Evidencia obtenida de al menos un ensayo diseñado apropiadamente, randomizado y controlado.
2. Evidencia a partir de estudios no randomizados o no controlados.
 - 2.1. Evidencia obtenida de ensayos controlados bien diseñados, sin randomización.
 - 2.2. Evidencia obtenida de estudios analíticos bien diseñados de cohortes o caso-control preferiblemente de más de un centro o grupo de investigación.
 - 2.3. Evidencia obtenida de series múltiples de tiempo con o sin la intervención. Los resultados muy llamativos en experimentos no controlados pueden ser considerados también como evidencia de este tipo.
3. Opiniones de expertos basadas en experiencia clínica, estudios descriptivos o informes de comités de expertos. Este es el grado más bajo de validez.

Obviamente, es muy poco práctico analizar de forma exhaustiva la validez de todos los criterios a utilizar, sin embargo, es necesario al menos reflexionar sobre ello y cuestionarse aquellos que se hacen exclusivamente por consenso. Por otra parte, muchas de las decisiones a tomar en los procesos asistenciales no están, no ya probadas, sino ni siquiera estudiadas; no es incorrecto entonces basarse en el consenso para elaborar criterios, pero sabiendo que serían en cierto modo criterios "provisionales".

La evidencia científica tiene varios niveles, el más bajo de los cuales es el consenso. Los criterios basados en el consenso son criterios "provisionales".

Cuando lo que vamos a explorar no son criterios ya probados, sino lo que hemos llamado en UT anteriores *causas hipotéticas*, el propio diseño del estudio nos va a informar (al medirlas concurrentemente con el criterio que representa el problema a evaluar) si tienen al menos validez interna, es decir, para el entorno en que se realiza el ciclo de mejora.

Una vez hemos reflexionado sobre la validez, hasta donde nos sea posible, lo que si hemos de asegurarnos es que los criterios a emplear sean *fiabiles*. El criterio que elijamos para representar el problema a evaluar tenemos que considerar que es efectivamente válido, *antes* de la medición.

8. FIABILIDAD

Si la validez asegura que la herramienta mida lo que queremos medir, la fiabilidad se define como el grado de reproducibilidad de los resultados cuando la herramienta es utilizada por observadores diferentes. Es decir, que cada uno de los evaluadores obtenga el mismo resultado que el otro o los otros evaluadores al evaluar la misma cosa con el mismo criterio. Sin fiabilidad no hay validez, porque estamos asignando valores a la medición de una forma inconsistente. La realización de un pilotaje anterior a la generalización de las mediciones, es considerado un paso inexcusable para asegurar la fiabilidad. El pilotaje permite identificar, discutir y corregir, si las hubiera, diferencias en la interpretación de los criterios o indicadores. El proceso de pilotaje debe repetirse tantas veces como sea necesario hasta alcanzar cotas de fiabilidad aceptables.

Para medir la fiabilidad se utilizan diversos índices de acuerdo, generalmente fáciles de calcular e interpretar: índice kappa, índice de concordancia general, índice de concordancia específica. Si el criterio es fiable, estos índices deben ser elevados. El más extendido es probablemente el estadístico kappa de Cohen que se usa para corregir el acuerdo debido al azar y permite estimar la significación estadística. Vamos a ver como se calculan e interpretan cada uno de ellos.

8.1. ÍNDICE KAPPA

El índice kappa (k) mide la concordancia total que existe si excluimos la debida al azar, o el acuerdo real más allá del azar. Su cálculo responde a la fórmula general:

$$k = \frac{P_o - P_e}{1 - P_e}$$

donde P_o es la concordancia observada, y P_e es la concordancia esperada debida exclusivamente al azar. El valor de k va de 0 a 1, aunque pueden resultar también valores negativos, siendo 0 el valor que representa la concordancia por azar, y 1 el valor de concordancia máximo, que respondería a un fiabilidad perfecta.

Una vez calculados P_o y P_e , el cálculo de k puede realizarse según la fórmula indicada por Fleiss tanto para parejas de evaluadores como para el caso de evaluadores múltiples. Existen otras alternativas, al parecer estadísticamente más correctas que la fórmulas de Fleiss, para el caso de que sean más de dos evaluadores, pero considerablemente más complejas.

La validez de criterio equivale a comprobar la evidencia científica.

La fiabilidad se asegura pilotando los criterios y analizándolos para introducir las modificaciones oportunas, antes de generalizar el uso del criterio.

La fiabilidad se mide con diversos índices como el kappa, la concordancia general y la concordancia específica.

Veamos un ejemplo de cálculo del índice kappa para una pareja de evaluadores:

Para obtener los datos necesarios el criterio cuya fiabilidad queremos conocer es utilizado de forma independiente por dos evaluadores para evaluar las mismas historias. Vamos a suponer que seleccionamos una muestra de 30; con los resultados de los dos evaluadores podemos construir una tabla en la que se reflejan las veces que han coincidido en que el criterio se cumple o que no se cumple. La tabla, ya con los datos obtenidos en nuestro ejemplo simulado, sería la siguiente:

| Evaluador 1 | Evaluador 2 | | |
|-------------|-------------|-----------|----|
| | Cumple | No cumple | |
| Cumple | 9 | 1 | 10 |
| No cumple | 6 | 14 | 20 |
| | 15 | 15 | 30 |

Quiere esto decir que según el evaluador 1, 10 historias cumplen el criterio y 20 no; mientras que para el evaluador 2, 15 historias cumplen el criterio y 15 no. Las celdas de la tabla indican la distribución de las que cumplen y no cumplen según cada evaluador, en función de lo que dijo el otro. Así por ejemplo, de los 20 casos en los que el evaluador 1 dijo que el criterio no se cumplía, en 14 coincidió con el evaluador 2, y en 6 no coincidió porque el evaluador 2 dijo que sí cumplía el criterio.

La concordancia observada (P_o) es el total de coincidencias, dividido por el total de casos evaluados. Según los datos de la tabla $P_o = 9+14 / 30 = 0,77$ es decir en 77%. ¿Y cuál sería P_e , la concordancia esperada por puro azar? P_e se calcula sumando el producto de las proporciones de cumplimiento que han dicho los dos evaluadores, al producto de las proporciones de incumplimiento.

En nuestro ejemplo, según el evaluador 1, la proporción de cumplimiento es $10/30 = 0,33$

Según el evaluador 2, la proporción de cumplimiento es $15/30 = 0,5$

El producto de ambas = $0,33 \times 0,50 = 0,165$

Según el evaluador 1, la proporción de incumplimiento es $20/30 = 0,67$

Según el evaluador 2, la proporción de incumplimiento es $15/30 = 0,5$

El producto de ambas = $0,67 \times 0,50 = 0,335$

P_e es igual a la suma de ambos productos = $0,165+0,335= 0,5$

Con ella podemos calcular $\frac{P_o - P_e}{1 - P_e} = \frac{0,77 - 0,5}{1 - 0,5} = 0,54$

¿Qué podemos concluir ante valores como este?, ¿es el criterio fiable o no? En esencia, ¿qué valores de kappa son adecuados? Como respuesta, existen diversas escalas de valoración en la literatura pero en general, la mayoría de los autores coinciden en dar por *aceptables* kappas mayores de 0,4. Para ser buenos deberían de ser mayores de 0,6. Por ejemplo, según criterio de Fleiss se considera un kappa que traduce buena fiabilidad si es mayor o igual a 0,6. Según Landis y Kock, kappa negativo o cero indica acuerdo bajo o inexistente,

hasta 0,2 acuerdo ligero, por encima de 0,4 acuerdo moderado, a partir de 0,6 acuerdo notable, a partir de 0,8 acuerdo alto y 1 acuerdo absoluto. Según todos los autores kappas por debajo de 0,4 no son aceptables. El kappa de nuestro ejemplo sería aceptable.

Sin embargo, y a pesar de su extendido uso, el índice kappa no está exento de inconvenientes. El valor y la interpretación del índice kappa debe expresarse con algunas cautelas, en relación con diversos factores que pueden hacerlo variar con independencia de la fiabilidad real del instrumento valorado. Estos factores son la *prevalencia* de la cualidad que el instrumento mide, la *sensibilidad* y *especificidad* de los diversos evaluadores, y la propia *adecuación de la fórmula* aplicada cuando se analiza la fiabilidad de más de dos evaluadores simultáneamente.

La influencia de la *prevalencia* (o proporción verdadera de presencia de la característica evaluada) en el valor de kappa es que para prevalencias (expresadas en nuestro caso como proporción de cumplimientos de los criterios) cercanas a 1 o a 0, kappa tiende a ser *cero* con independencia del grado de acuerdo entre los evaluadores, a menos de que éste sea total, en cuyo caso el valor de kappa para prevalencia de 1 o 0 es infinito. A la vista de los cálculos simulados realizados por diversos autores para diversas prevalencias, manteniendo constante en cada caso la sensibilidad y especificidad de los evaluadores, parece evidente que, como regla práctica general, el valor de kappa no debe ser interpretado de la forma habitual para prevalencias menores de 0,15 o mayores de 0,85.

En nuestro caso, como puede ser el caso en la práctica mayoría de las veces en que se pilote y evalúe la fiabilidad de los criterios o indicadores para medir calidad, la prevalencia es inicialmente desconocida; lo que proponemos hacer, es estimarla en base a calcular la media de las prevalencias (proporciones de cumplimiento) que propone cada evaluador. Está claro que lo ideal es conocer previamente la prevalencia (proporción de cumplimientos real); sin embargo, aunque puede haber considerables desviaciones en relación al valor real, las situaciones extremas van a ser identificadas. En resumen y como norma, una correcta interpretación de un índice kappa bajo necesita del conocimiento, o al menos de un cálculo estimado, de la prevalencia de lo que valora el instrumento analizado; el valor de kappa solo deberá ser concluyente por sí solo si es alto: en kappas bajos debemos asegurarnos de que no estamos ante prevalencias extremas.

La *sensibilidad* y *especificidad* de los evaluadores en relación a lo evaluado, en general no es posible conocerla ni estimarla de antemano, pero es, sin embargo un factor también de gran influencia en los valores de kappa. De hecho según cálculos simulados publicados, si la sensibilidad (capacidad de ver la existencia de la cualidad o criterio si realmente está) fuese inferior al 70%, el valor de kappa va a resultar necesariamente inferior a 0,4, porque la baja sensibilidad conduce por sí misma a discordancias. La manera más común de normalizar (y elevar) la capacidad discriminatoria de los evaluadores es entrenándolos. En ausencia de entrenamiento, la mejor medida es construir el instrumento de valoración con unas especificaciones para su manejo (excepciones, aclaraciones, mecánica de evaluación) lo suficientemente detalladas para que sea *necesariamente* fiable en las condiciones normales de aplicación.

El cálculo del valor de *kappa para múltiples evaluadores* merece también un

El índice kappa es relativamente sencillo de calcular, en base al conocimiento de la concordancia observada y la esperada por azar. Para ser aceptable ha de tener un valor mayor de 0,4; para ser bueno ha de ser mayor de 0,6.

Para interpretar adecuadamente el índice kappa, hay que tener en cuenta la prevalencia de lo que mide el criterio que se evalúa.

Prevalencias extremas hacen que el índice kappa tienda a ser cero y su valor no se puede interpretar en la forma habitual.

comentario. El hecho de que la valoración del cumplimiento la efectúen más de dos observadores aumenta la complejidad del cálculo, y se pueden utilizar modelos jerárquicos, o bien la fórmula de Fleiss para más de dos observadores. La fórmula propuesta por Fleiss, es cuestionada por algunos autores, pero la complejidad del cálculo de las fórmulas alternativas para varios evaluadores, junto al hecho de que el kappa para dos evaluadores sigue sin cuestionar, aconseja basar nuestros cálculos rutinarios en los valores de kappa por parejas. Si el pilotaje se realiza con más de dos evaluadores, puede calcularse el kappa para las diversas combinaciones de parejas.

Finalmente, la correcta interpretación del índice kappa precisa del cálculo de su intervalo de confianza, de forma que daremos por aceptables solo los kappas cuyo límite inferior del correspondiente intervalo de confianza sea mayor de 0,4. La fórmula para el cálculo del intervalo de confianza para el índice kappa puede encontrarla el lector interesado en la bibliografía que se cita (Fleiss) al final de esta UT.

Después de todas estas aclaraciones y consideraciones sobre las particularidades del índice kappa vamos a ver los otros índices para valorar la fiabilidad.

8.2. ÍNDICE DE CONCORDANCIA GENERAL

Indica la concordancia de los evaluadores al clasificar el criterio. Se define como la proporción de casos en que el juicio de ambos evaluadores coincide. Se calcula dividiendo el número de acuerdos de ambos observadores por el total de casos evaluados. Es equivalente a P_o o proporción de acuerdo observado, en la fórmula del cálculo de kappa.

Para su interpretación aislada para valorar la fiabilidad hacen falta valores muy altos para que sea considerada buena. Según proponen Palmer y Banks concordancias $\geq 95\%$ van a ser siempre sinónimos de herramientas fiables. Estos niveles de concordancia son exigibles, así como kappas como mínimo buenos ($>0,6$) cuando el criterio o indicador se va a generalizar y utilizar para comparar diversos centros en programas externos.

8.3. ÍNDICE DE ACUERDO ESPECÍFICO

Es el acuerdo al clasificar como negativo un criterio (concordancia al decir que no se cumple). Se calcula dividiendo el número de casos en que ha habido acuerdo al decir que el criterio no se cumple, por el número de casos a los que al menos un observador ha dicho que no se cumple. Tiene interés cuando nos interesa especialmente que no se clasifique como incumplidores, a quienes puedan estar haciéndolo bien. Su interpretación es semejante a la de la concordancia general.

8.4. PROCEDIMIENTO PRÁCTICO PARA EVALUAR LA FIABILIDAD

Una forma simple, intuitiva y pragmática de obviar toda la complejidad del índice kappa en cuanto a cálculo e interpretación es basar inicialmente la valora-

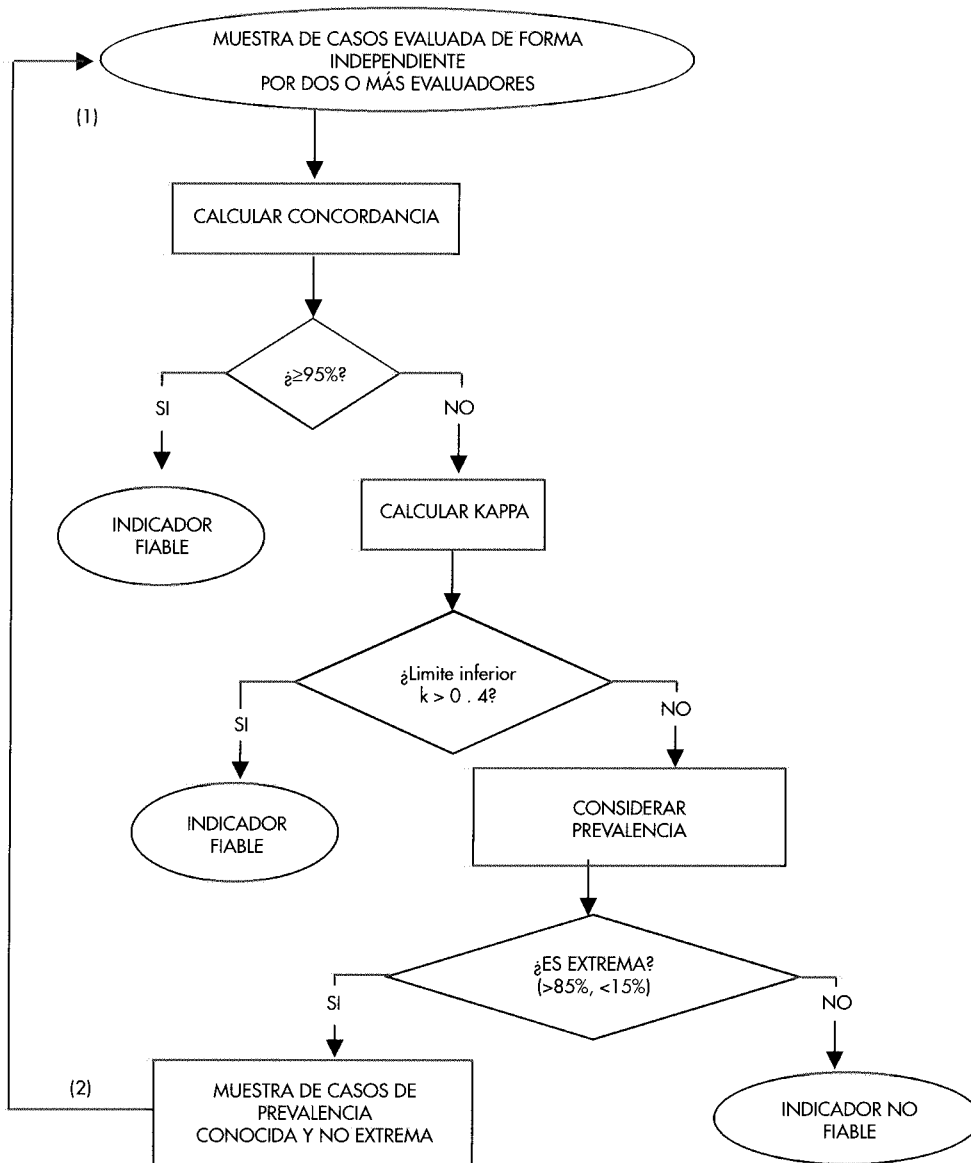
El índice de concordancia general es la concordancia observada. Son exigibles valores muy elevados (mayor o igual a 95%) para utilizarlo con seguridad como medida aislada de fiabilidad.

ción de la fiabilidad en el cálculo de la concordancia simple, y solo si ésta no es alta ($\geq 95\%$) calcular el índice kappa, ya que en ocasiones concordancias solo moderadamente altas van unidas a kappa aceptables. Además, el índice kappa debería acompañarse, como hemos visto, de un conocimiento previo o una estimación de la prevalencia de la cualidad que evalúa el instrumento cuya fiabilidad queremos analizar.

La Figura 8.1 contiene un esquema, en forma de algoritmo, que proponemos para valorar la fiabilidad de los indicadores, en base a la discusión que acabamos de exponer.

La valoración de la fiabilidad puede exigir la consideración de la concordancia, el kappa y la prevalencia, en una secuencia práctica que explicita la Figura 8.1.

FIGURA 8.1. Procedimiento de análisis de la fiabilidad de los indicadores para medir calidad



(1): punto de inicio cuando la prevalencia es desconocida.

(2): punto de inicio si conocemos la prevalencia.

REFERENCIAS BIBLIOGRÁFICAS

- Palmer RH. Planificación de una evaluación. En: Palmer RH. Evaluación de la asistencia ambulatoria. Principios y prácticas. Madrid: Ministerio de Sanidad y Consumo; 1990. p. 39-80.
- Saturmo PJ. Como diseñar una evaluación sobre la calidad asistencial en atención primaria. IV Congreso de la Sociedad Andaluza de Medicina Familiar y Comunitaria. Sevilla: Consejería de Salud; 1992.
- Hernández Aguado I, Porta Serra M, Miralles M, García Benavides F, Bolúmar F. La cuantificación de la variabilidad en las observaciones clínicas. *Med Clin (Barc)* 1990; 95: 424-429.
- Fleiss JL. *Statistical Methods for rates and proporciones*. New York: John Willey & Sons; 1981.
- Thompson WD, Walter SD. A reappraisal of the kappa coefficient. *J Clin Epidemiol*, 1988; 41(10): 949-958.

DISEÑO DE ESTUDIOS DEL NIVEL DE CALIDAD. COMPONENTES DE UNA EVALUACIÓN

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

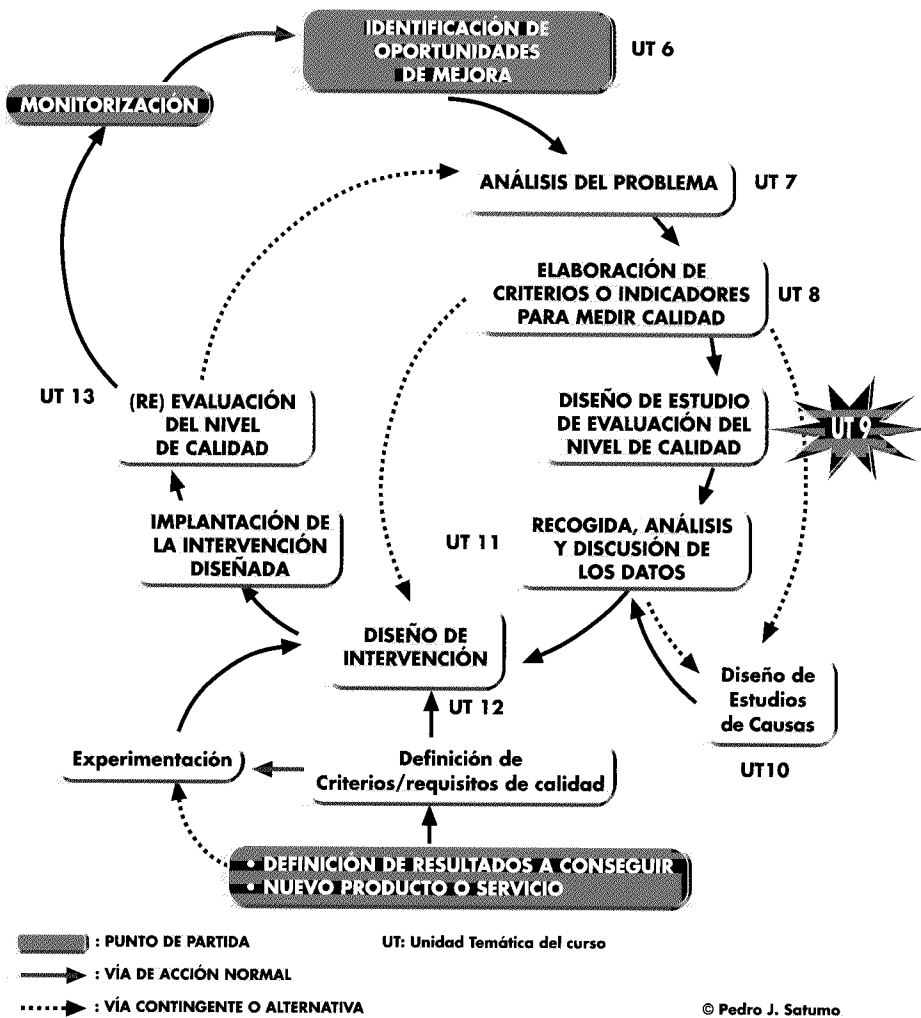
Una vez establecidos los criterios o requisitos de calidad de los aspectos a evaluar, hemos de diseñar cuidadosamente su medición. Veremos en esta UT que los estudios de evaluación de la calidad pueden tener diversos niveles de complejidad, pero que podemos agruparlos en dos que denominamos, a falta de una terminología mejor, estudios del nivel de calidad y estudios de causas de calidad deficiente. Se explica más en detalle el primero de ellos, consistente en medir el nivel de presencia de los criterios o requisitos de calidad preestablecidos, y se ofrece un esquema detallado de los componentes del estudio que deben de quedar bien definidos, sea cual sea el servicio o producto objeto de evaluación.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. Estudios del nivel de calidad y estudios de investigación de causas de la calidad deficiente.
3. Diseño de estudios de nivel de calidad: componentes a especificar.

OBJETIVOS ESPECÍFICOS

1. Diferenciar los diversos tipos de estudio que pueden diseñarse para evaluar la calidad, en especial aquellos que miden requisitos de calidad de aquellos que buscan causas de calidad deficiente.
2. Definir las características de las unidades de estudio (casos) en las que se va a realizar la medición.
3. Distinguir entre marco temporal para el muestreo de los casos y periodo de tiempo en el que los criterios deben cumplirse.
4. Determinar si todos los criterios se miden en idénticas unidades de estudio o no.
5. Definir cuáles van a ser las fuentes de los datos que se necesitan en el estudio.
6. Determinar el número de casos a evaluar.
7. Valorar las ventajas e inconvenientes de utilizar diversos tamaños de muestra.
8. Explicar en qué consiste una muestra representativa.
9. Distinguir muestreos aleatorios de muestreos de conveniencia.
10. Escoger el método de muestreo que más se ajuste al contexto de su centro.
11. Desarrollar paso a paso la metodología a seguir para extraer la muestra.
12. Definir quién debe encargarse de recoger los datos.
13. Resumir en el diseño del estudio las decisiones tomadas en relación a la dimensión de la calidad que se va a evaluar y los tipos de datos a utilizar.



"Desde que tenemos el programa de gestión de calidad nada cuenta hasta que no se cuantifica. Sólo nos creemos lo que medimos"

(Gerente de hospital distinguido con el Premio a la mejor Gestión. Colombia 1996)

1. INTRODUCCIÓN

En Unidades Temáticas previas hemos tomado decisiones importantes y se han resumido las implicaciones teóricas y prácticas en relación a los tres primeros pasos de los ciclos de mejora: la identificación del problema de calidad u oportunidad de mejora (UT6), el análisis del problema seleccionado (UT7) y la elaboración de criterios válidos y fiables para medir la calidad (UT8). Esta UT aborda el paso siguiente: diseñar adecuadamente la recogida de datos para estudiar y dimensionar el problema de calidad.

Para ello vamos a distinguir en primer lugar entre dos tipos de estudios diferentes que podrían plantearse: estudios de nivel de calidad y estudios de causas de calidad deficiente. A continuación detallamos los componentes del diseño del estudio en referencia al primero de ellos (nivel de calidad), que es el que vamos a practicar la mayoría de las veces en que estemos interesados en evaluar la calidad de los procesos clínicos.

En esta UT se avanza un paso más en el Ciclo de Mejora: el diseño de estudio para medir calidad.

2. ESTUDIOS DEL NIVEL DE CALIDAD Y ESTUDIOS DE INVESTIGACIÓN DE CAUSAS DE LA CALIDAD DEFICIENTE

Hay dos grandes grupos de estudios cuyos resultados nos van a servir para decidir qué hacer para mejorar: los estudios del nivel de calidad y los estudios de investigación de causas de la calidad deficiente. En los primeros, lo que vamos a medir es el grado de cumplimiento de los criterios o requisitos de calidad previamente definidos y que sabemos que su grado de presencia o ausencia es equivalente al grado (o nivel) de calidad de lo que vamos a evaluar. Éste es el caso normalmente de las evaluaciones de calidad de problemas de atención clínica, en donde vamos a comprobar si estamos haciendo las cosas de forma correcta. En estos estudios medimos directamente lo que sabemos que son requisitos de calidad.

Una situación diferente es cuando el problema no es de tipo clínico sino organizacional (tiempo de espera, falta de coordinación entre los profesionales, etc.), podemos tenerlo incluso ya medido, pero para su mejora tenemos que averiguar las causas que lo provocan; una situación que puede darse también después de los estudios de nivel de calidad, si no sabemos qué hacer para mejorar un determinado criterio clínico que no se cumple. En ambos casos, lo que tenemos va a ser un serie de causas hipotéticas que investigar, pero no la certeza de que su presencia o ausencia va unida a presencia o ausencia de calidad, como ocurre en los estudios que hemos llamado del nivel de calidad. Por ejemplo, si determinamos que el tiempo de espera para la consulta es criterio de calidad y queremos medir la proporción de pacientes que esperan más de un cierto tiempo, diseñaríamos un estudio de nivel de calidad; si por el contrario ya sabemos que el tiempo de espera es excesivo y no tenemos idea de por qué es así, diseñaríamos un estudio para averiguar de todas las causas hipotéticas (dependientes del paciente, de la organización del centro, horarios, normas, etc.) cuáles se asocian a tiempos de espera prolongados en nuestro caso concreto. En los estudios de investigación de causas hipotéticas es necesario medir concurrentemente con las posibles causas, el criterio o los criterios que nos representan la calidad del aspecto en estudio. Por ejemplo, no podemos averiguar las causas de los tiempos de espera prolongados si no medimos a la vez las causas y los tiempos de espera a los que se asocian, por eso el estudio de causas debe ir precedido de la definición de los criterios o indicadores de calidad del problema en estudio. Estos mismos criterios o indicadores son los que tendremos que medir tras la intervención para asegurarnos que hemos mejorado.

3. DISEÑO DE ESTUDIOS DEL NIVEL DE CALIDAD: COMPONENTES A ESPECIFICAR

Los estudios sobre calidad han de ser una actividad rigurosa, planificada y factible. Para asegurarnos que es así, resulta útil seguir una cierta sistemática, que nos ayude a considerar todos los aspectos que van a ser relevantes en su realización; adicionalmente, la especificación homogénea de los componentes del estudio sirve también para poder replicarlo, y comparar evaluaciones posteriores de un mismo centro o realizadas en centros distintos.

El esquema que proponemos tiene sus raíces en el Programa Ibérico, modificado en base a experiencias posteriores. Consta de una serie de componentes, la

En los estudios del nivel de calidad medimos criterios que ya se sabe que están relacionados con la calidad.

Se propone un esquema de componentes que conviene definir separada y exhaustivamente.

mayoría de los cuales están interrelacionados, aunque conviene considerarlos aisladamente y procurar su definición exhaustiva. Una vez definidos todos, queda definido asimismo el diseño de la evaluación.

Estos componentes son:

- Criterios para evaluar la calidad.
- Dimensión estudiada.
- Tipos de datos.
- Unidades de estudio.
- Fuente de datos.
- Identificación y muestreo de las unidades de estudio.
 - Marco muestral.
 - Número de casos.
 - Método de muestreo.
- Tipo de evaluación.
 - Según relación temporal.
 - Según quién tomó la iniciativa.
 - Según quién obtiene los datos.

Varios de estos componentes han sido vistos y decididos en UT anteriores y haremos una breve reseña de los mismos; los demás vamos a revisarlos más en detalle.

3.1. CRITERIOS PARA EVALUAR LA CALIDAD

Los criterios van a ser la herramienta con la cual mediremos la calidad. Su aplicación nos va a servir para discernir buena de mala calidad y los diferentes niveles en que se encuentre. Hay que recordar que, en la práctica, aceptar unos criterios determinados, equivale a establecer y aceptar asimismo unos determinados objetivos de calidad en la asistencia que prestamos. Un criterio es a la vez un requisito de calidad que queremos que se cumpla. De ahí la tremenda importancia que tiene el que sean relevantes, realistas y que puedan ser aceptados sin controversia por todos los profesionales implicados. Una correcta elaboración y selección de criterios es, junto a una priorización adecuada del tema a evaluar, el principal factor determinante del éxito o fracaso del ciclo de mejora (este punto se puede revisar en la UT 8).

3.2. DIMENSIONES ESTUDIADAS

A la vista del aspecto a evaluar y los criterios que vamos a medir, hemos de describir cuál de las varias dimensiones de la calidad vamos a valorar.

Como vimos en la UT1, proponemos la utilización de un listado reducido de tres dimensiones: calidad científico-técnica, accesibilidad y satisfacción.

La mayoría de las evaluaciones que realizan los profesionales incluyen la medición, muchas veces como única dimensión, de algún aspecto de la calidad científico-técnica. Esto es así no sólo porque es la dimensión más amplia y con

El primer componente es la descripción de los criterios a utilizar, especificando excepciones y las aclaraciones necesarias.

El segundo componente es aclarar qué dimensión o dimensiones (calidad científico-técnica, accesibilidad, satisfacción) va a ser evaluada.

más subcomponentes; también influye el que sea la que se ve más cercana a la actuación de los profesionales, y aquella sobre la que más directa y fácilmente podemos actuar. Es también la dimensión más comúnmente aceptada como indiscutible para caracterizar la calidad de la atención prestada. Sin embargo, no debe olvidarse la importancia de garantizar asimismo unos niveles adecuados de accesibilidad y satisfacción. Accesibilidad tanto interna (en el sentido de la existencia o no de barreras a la continuidad de cuidados) como externa, si pensamos que actuamos sobre poblaciones y no exclusivamente sobre pacientes que nos visitan; y la satisfacción, como un resultado a conseguir, y para garantizar la adecuada y necesaria colaboración de los pacientes y población en los programas y tratamientos.

En el diseño del estudio hay que hacer explícita cual es la dimensión o dimensiones que hemos decidido medir. Con ello aclaramos asimismo cual es la dimensión concreta que intentamos mejorar, y avisamos implícitamente sobre las dimensiones que van a escapar a nuestra medición, sobre las que no podemos afirmar si están o no presentes, o si conseguimos o no mejorar. Naturalmente se asume que, aunque no haya cambios en las otras dimensiones, la mejora de una de ellas supone un salto cualitativo en la asistencia que ofrecemos.

3.3. TIPOS DE DATOS

Los criterios a utilizar van a estar referidos a datos de estructura, proceso o resultado. El explicitarlo como uno de los componentes del diseño es para recordar la necesaria reflexión sobre los problemas de validez que puedan plantear cada uno de ellos (este punto puede revisarse en la UT8). También es conveniente constatarlo para dejar claro que sabemos de la existencia de otros datos potencialmente relevantes que no van a ser tenidos en cuenta.

3.4. DEFINICIÓN DE LAS UNIDADES DE ESTUDIO

La descripción de las unidades de estudio para la evaluación, es el equivalente a la definición de "caso" a evaluar. Se trata, en esencia, de determinar cuáles son las características de la unidad básica, más desagregada, sobre la que vamos a efectuar la medición de los criterios de calidad asistencial.

Habitualmente tendremos que especificar tres aspectos:

- **Características del paciente o población que entra en la evaluación.** Debe ser coincidente con el grupo de población objeto o receptor de la atención que evaluamos. Este grupo puede ser definido en función de su edad, sexo, pertenencia a determinado grupo de riesgo, diagnosticados de determinada enfermedad, sometidos a determinado tratamiento, intervención, etc. Por ejemplo, en el caso de la evaluación de la atención preventiva del cáncer femenino, podemos decidir que será objeto de estudio la atención recibida en este sentido por toda mujer mayor de 18 años residente en nuestra área de cobertura, o bien que haya acudido al menos una vez a nuestro centro. Para la detección de hipertensos, todos los adscritos a nuestro centro y que pertenezcan a unos ciertos grupos de edad, etc.
- **Características de los profesionales y nivel asistencial evaluado.** Por ejemplo, un médico de un determinado centro sanitario, todos los médicos de un

El tercer componente es la constatación explícita del tipo de datos (estructura, proceso, resultados) que vamos a utilizar.

Las unidades de estudio se delimitan definiendo las características de la población que recibe la atención evaluada, los proveedores de la atención y el tiempo en que la atención tiene que haber sido prestada, a cada paciente.

centro sanitario, el personal de enfermería, todo el equipo de un servicio, todos los centros de un área, etc. Como en el caso anterior, esta característica es excluyente: sólo vamos a considerar aquellos casos poblacionales de características determinadas que además haya dependido en relación al problema a evaluar de una determinada instancia o nivel asistencial, que es el que evaluamos.

- **Período de tiempo evaluado.** Coordinada que no hace referencia necesariamente al tiempo físico convencional (un año determinado, un mes, etc.), sino al segmento del proceso asistencial, tiempo de contacto o relación entre cada caso que entra en el estudio y el nivel asistencial evaluado, en el cual queremos que se vea cumplido el criterio de calidad que vamos a medir. Por ejemplo, podemos definir que en nuestro programa de detección precoz de cáncer de mama, la exploración de mama debe realizarse en la primera visita de toda mujer incluida el grupo de edad que hemos definido. De este modo, el periodo de tiempo que define nuestra unidad de estudio es una visita, y más concretamente la *primera visita* a nuestro centro, de manera que al obtener los datos para nuestra evaluación no vamos a fijarnos en las actuaciones realizadas en visitas sucesivas. En otros casos, especialmente en el seguimiento de enfermedades crónicas, podemos plantearnos evaluar la atención prestada en un año, en las visitas siguientes al diagnóstico, etc. Otros criterios *pueden no tener referencia temporal*, como es el caso de muchos de los criterios de estructura.

En una evaluación con varios criterios, es perfectamente posible que cada uno de ellos se refiera a segmentos del proceso diferentes. La experiencia nos ha mostrado que este aspecto de la definición de la unidad de estudio es la más difícil de entender y especificar, especialmente cuando se quiere evaluar la atención en caso de pacientes crónicos.

3.5. FUENTES DE DATOS

En nuestro diseño tenemos que hacer explícito de dónde va a ser posible extraer los datos que necesitamos para la evaluación. En evaluaciones retrospectivas, vamos a utilizar muy probablemente las historias clínicas y/o cualquier otro sistema o soporte de registro de datos en relación al tema evaluado. Otra fuente de datos frecuente (a veces obligada por el tema o por la inexistencia de registros sobre el dato de interés, o porque lo que nos interesa es lo que puede decirnos el usuario) son las encuestas *ad hoc*. En otras ocasiones (habilidades manuales, técnicas de entrevista clínica, etc.), la evaluación sólo es posible mediante observación directa.

En cualquier caso, una regla a seguir siempre es procurar comenzar diseñando criterios que puedan ser evaluados con los datos o sistemas de registro ya existentes. En aras de la factibilidad (y su correspondiente repercusión en el éxito de la implantación de las actividades de GC), hay que procurar no empezar con evaluaciones que precisen el montaje de nuevos registros o la realización de encuestas complicadas.

Desafortunadamente el relativo déficit de buenas fuentes de datos limita con frecuencia las posibilidades y la extensión de las evaluaciones de la calidad asistencial. En relación a ello, es frecuente encontrar como la primera o una de

A continuación hay que describir de qué fuente vamos a identificar los casos (listados, libros de registro, censos, etc.) y de dónde extraer los datos para la evaluación. Evitar en lo posible tener que crear fuentes de datos.

las primeras actividades de GC en los centros, la evaluación de la "calidad" de los registros e historias clínicas. Sin quitarle mérito a esa actividad, pensamos que, una vez garantizada la existencia de registros, es más efectivo el que los déficits que tengan aparezcan evidentes en función de su repercusión en la buena o mala asistencia que se presta a un problema de salud determinado (por ejemplo registro de edad y sexo para saber si esta persona es susceptible de entrar en un programa de detección precoz de cáncer femenino, registro de peso y talla y resultados de la analítica en pacientes cuyo diagnóstico, tratamiento o seguimiento adecuados dependen precisamente del conocimiento explícito de esos valores, etc.); de hecho los registros han de ser contemplados como un medio más para lograr una buena asistencia y no como un fin en sí mismos.

3.6. IDENTIFICACIÓN Y MUESTREO DE LAS UNIDADES DE ESTUDIO

a) Marco Muestral

Una vez definida la unidad de estudio, sobre todo en lo que respecta a las características de la población atendida cuya asistencia queremos evaluar, nos hemos de plantear los mecanismos para identificar los casos de pacientes o población que responden a la definición de caso, y el marco muestral o universo de donde extraeremos una muestra representativa de los mismos para su evaluación. Los mecanismos de identificación de casos y la facilidad relativa para realizar un muestreo varían según el tipo de atención a evaluar y las características de cada centro.

Frecuentemente, sobre todo en los casos en que la atención a evaluar se encuadra en un determinado programa o grupo bien definido (hipertensión, diabetes, encamados en domicilio, niño sano, etc.), existen ficheros específicos que facilitan el muestreo e identificación de los casos a evaluar. Otras veces, es posible utilizar a tal efecto los registros de morbilidad, o cualquier otro registro rutinario en el que conste explícitamente la característica que define la unidad de estudio. Otras veces, el universo para muestreo e identificación de casos han de ser directamente las historias clínicas, o un muestreo aleatorio de la población cubierta o atendida. Todo depende del tema concreto a evaluar, la definición que hagamos de la unidad de estudio y las características estructurales de cada centro.

Es aconsejable tener presentes muy en primer término las posibilidades de identificación de casos y obtención de los datos pertinentes, al escoger el tema y diseñar los criterios para evaluar: en primera instancia es preferible abordar evaluaciones con criterios aplicables a casos de identificación posible y fácil, aunque resulten menos atractivas que otras para las que el proceso de identificación de casos pueda resultar costoso o complicado.

Si el total de casos (universo o marco muestral) que responde a la definición de unidad de estudio es pequeño, podemos plantearnos evaluarlos todos; sin embargo, lo más frecuente es que sea más conveniente extraer una muestra.

El muestreo, aparte de ser el proceder de elección cuando el universo de unidades de estudio es relativamente grande, y obligatorio cuando es muy grande, tiene la ventaja de que es menos costoso (por lo menos en tiempo) que evaluar la totalidad de casos, y permite obtener resultados útiles a corto plazo. Los

Es aconsejable tener muy claro en el diseño la delimitación del universo o marco muestral y la factibilidad de los mecanismos para la obtención de los casos a evaluar y los datos para la evaluación.

Evaluar con muestreo es una manera eficiente y útil de tener datos, que se convierte en obligatoria si el universo de casos o marco muestral es muy grande.

inconvenientes, si es que pueden llamarse así, es que hay que hacerlo bien, y para ello hay que manejar con cierta soltura unos conceptos estadísticos básicos; y además puede haber escépticos que no terminen de dar crédito a los resultados por el simple hecho de ser una muestra. Es muy aconsejable prepararse para afrontar estos dos inconvenientes, porque las ventajas superan con mucho los esfuerzos que puede suponer el poder evaluar con muestras.

b) Número de casos necesarios

La definición del número de casos o unidades de estudio necesarios para la evaluación, es uno de los puntos que más angustia suele producir al diseñar una muestra; es también un punto en el que el enfoque que se aconseja para los estudios de evaluación de la calidad puede diferir substancialmente de los trabajos de investigación de otro tipo. Podemos adelantar que un tamaño de muestra manejable y habitualmente suficiente es de 50-60 casos, pero conviene entender por qué y aprender a extraer la muestra de forma que sea representativa. Para ello, tras decidir cual es el nivel o unidad organizativa para la que queremos extraer conclusiones, hay que tener claros los conceptos de confianza, precisión y representatividad de la muestra.

- **Unidad organizativa para la que queremos extraer conclusiones.** Al definir la unidad de estudio en lo relativo al proveedor del servicio que se evalúa, hemos decidido si nos interesa que sea un profesional individual, un equipo, todo el centro, etc.; con ello hemos contribuido a definir también el marco muestral. Sin embargo, esta decisión es además de una gran trascendencia para determinar el tamaño de la muestra: el número de casos que decidamos extraer nos va a permitir extraer conclusiones sobre la calidad al nivel organizativo que previamente hayamos decidido, pero no a un nivel inferior. Es decir, una muestra pensada para evaluar la calidad de atención al hipertenso a nivel de servicio, no va a servir para evaluar cada médico por separado. Si queremos evaluar cómo lo hace cada médico, vamos a necesitar el mismo tamaño de muestra pensado para el servicio, *pero para cada uno de los médicos*. Es factible la agrupación de muestras pensadas para niveles organizativos más bajos para evaluar niveles superiores (por ejemplo, agrupar los datos de cada médico para evaluar el servicio), pero no al revés. Esto hay que tenerlo muy claro y muy en cuenta al diseñar el estudio.

- **Representatividad, precisión y confianza de la muestra.** Estos tres conceptos, con estos y otros nombres, suelen considerarse relacionados con el tamaño de la muestra y se utilizan a veces de forma confusa e intercambiable, aunque son diferentes. Vamos a aclarar su significado.

Supongamos que estamos interesados en saber la altura media de los pacientes que están en un determinado día en las salas de espera de un centro de salud grande. Si queremos saberlo con toda exactitud, mediremos la talla de todos ellos y calcularíamos la media; en este caso, como no es una muestra no hay problema de representatividad, precisión ni confianza del resultado. Pero ¿y si son cientos de personas? ; podemos considerar entonces la posibilidad de hacer un *estimación* midiendo la talla sólo en una muestra de estas personas. Al ser una muestra, el resultado que obtengamos está sujeto a errores; pero esos errores pueden conocerse y controlarse de manera que los tengamos en cuenta a la hora de dar los resultados de nuestro muestreo. Estos errores son de tres tipos y se relacionan con la representatividad, la precisión y la confianza de los resultados o estimaciones de la muestra.

En las evaluaciones de calidad solemos utilizar inicialmente muestras manejables, relativamente pequeñas (50-60 casos) pero no exentas de rigor científico.

El tamaño de muestra decidido para evaluar niveles organizativos superiores (por ejemplo centro de salud), no sirve para evaluar niveles más desagregados (por ejemplo cada médico del centro).

La representatividad, la precisión y la confianza representan tres tipos de errores diferentes que podemos tener al medir con muestreo

La representatividad hace referencia al error de sesgo y no depende del tamaño de la muestra sino del método de obtención de la misma

El primer error que podemos tener es hacer una selección *sesgada* de las personas que incluimos en la muestra. Por ejemplo, supongamos que todos los cogemos de las consultas de pediatría: obviamente esta muestra no es *representativa* de los pacientes que esperan en las consultas del centro de salud, y ello *con independencia de cuántos niños hayamos incluido en la muestra*. El problema, el error, es que en nuestra muestra están sobrerrepresentados los niños, hayamos medido muchos o pocos. Para ser *representativa* tendríamos que haber hecho la selección de casos para la muestra de manera que todos los pacientes de todas las consultas hubieran tenido la misma oportunidad de entrar en la muestra, y no únicamente los de pediatría. El control del error de sesgo para que la muestra sea *representativa* depende del *método de muestreo*, no del tamaño de la muestra. La muestra va a ser *representativa* (no sesgada) si se ha realizado de forma aleatoria, lo cual equivale a decir que se le da a todas las unidades del universo muestral la misma oportunidad de entrar en la muestra. Otra cosa es, una vez que nos aseguramos que la muestra no es sesgada, hasta qué punto las estimaciones que hagamos con nuestra muestra responden o no al valor real que tiene, en el universo muestral, lo que estamos midiendo con una muestra. Es decir, siguiendo con nuestro ejemplo, hasta qué punto podemos inferir la altura media real de todos los que están esperando en las consultas a partir de una muestra *representativa* (no sesgada) de los mismos. Para saberlo, tenemos que tener en cuenta los otros dos errores que, éstos sí, varían con el tamaño de la muestra.

Una de las cosas que nos va a preocupar es la *precisión* de la estimación. Sabemos que el valor que encontramos en la muestra puede no ser el real, pero ¿y si tenemos un medio de, a partir de la estimación de la talla media que hemos obtenido en la muestra y teniendo en cuenta la variación (dispersión) de valores diferentes, y el tamaño de la muestra, poder afirmar que el valor real se encuentra entre unos determinados límites?; así podríamos conocer la magnitud de este error. Pues bien, el medio de calcularlo existe, y la forma de reconocerlo y controlarlo es dar el resultado de la muestra con un *intervalo* de valores, dentro del cual sabemos (por la teoría de la probabilidad) que está el valor real de lo que estamos midiendo; aunque no sepamos a ciencia cierta cuál de los valores incluido en el intervalo es el real (¡para eso tendríamos que medir a todos, no a una muestra!). Cuanto más pequeño sea el intervalo que nos resulta, más *precisa* es nuestra estimación, es decir, más pequeño será este error de muestreo. La estimación va a ser lógicamente más precisa, con intervalo de posibles valores reales más pequeño (menos error) cuanto más grande sea la muestra. En nuestro ejemplo, a cuanta más gente midamos la talla, más se acercará nuestra estimación al valor real. Los cálculos matemáticos oportunos nos mostrarían que podemos dar nuestra estimación con intervalos de valores posibles cada vez más pequeños.

Pero aún nos queda otra posibilidad de error: ¿hasta qué punto estamos seguros de que el valor real está realmente comprendido en el intervalo de valores que hemos calculado? Está claro que sólo tendremos un 100% de certeza si medimos a todos los pacientes que esperan. Como hemos obtenido una muestra, tenemos una cierta probabilidad de equivocarnos si afirmamos que los resultados obtenidos en la muestra serían los que obtendríamos si midiéramos a todos los pacientes. Esta probabilidad de estar o no equivocado al afirmar que el valor real es uno de los valores del intervalo que hemos calculado para saber la precisión, es lo que se conoce como significación o

La precisión de los datos medidos en una muestra se refiere a un error de muestreo que podemos calcular y reconocer dando el resultado de la evaluación con un intervalo de valores, dentro del cual va a estar el valor real. La precisión aumenta (es decir, el intervalo de valores disminuye) a medida que aumenta el tamaño de la muestra.

grado de confianza de la estimación que hemos obtenido con el muestreo. La confianza en el resultado del muestreo aumentará, naturalmente, si aumentamos el tamaño de la muestra, pero también si decidimos dar el resultado con menos precisión (con un intervalo de valores más amplio). Pensemos en nuestro ejemplo: si decimos que según los resultados de la muestra representativa (es decir, sin sesgo, extraída de forma aleatoria), la altura media estimada es de 160 cm y que la media real está en un intervalo de ± 20 cm en torno a esa media estimada (es decir, entre 140 y 180 cm), tendremos menos posibilidades de equivocarnos (más confianza), para un mismo tamaño de muestra, que si afirmamos que el valor real está entre 160 cm ± 5 cm (es decir, entre 155 y 165).

Dado que ambos errores (representados por la precisión y la confianza), aunque diferentes, están indisolublemente unidos y ambos dependen del tamaño de la muestra, cuando se dan los resultados de mediciones en base a una muestra se suelen (en realidad se *deben*) acompañar de la medida de estos dos errores, que es lo que se conoce conjuntamente como *intervalo de confianza* del resultado obtenido; el *intervalo* hace referencia a la precisión y se expresa como \pm una cierta cantidad, que hay que sumar y restar al valor estimado que hemos obtenido en la muestra, y la *confianza* hace referencia al grado de certeza del resultado que hemos obtenido (intervalo incluido) y se suele expresar en %; este % expresa la probabilidad de que el valor real sea de verdad uno de los que está incluido en el intervalo; por ejemplo un intervalo de confianza del 95%, quiere decir que estamos un 95% seguros de que el valor real está dentro del intervalo, o, lo que es lo mismo, que la probabilidad de que el valor real sea un valor de fuera del intervalo es sólo del 5%. En nuestro ejemplo, si el resultado de la medición de la talla en la muestra representativa es de 160 cm ± 10 cm, para una confianza de 95% (o lo que es lo mismo, el intervalo de confianza de 95% es 160 ± 10 cm), quiere decir que hay un 95% de probabilidades de que sea cierto que la talla media real está entre 150 cm y 170 cm (o un 5% de probabilidades de que el valor real, el que resultaría de medir a todos, sea inferior a 150 cm o superior a 170 cm).

Una vez entendido el significado de los tres errores que hay que tener en cuenta al medir con muestra, y cómo uno de ellos (el sesgo, traducido como representatividad) no depende del número de casos de la muestra sino del método de muestreo, podemos volver sobre la cuestión ¿cuántos casos hacen falta para que la evaluación sea útil?, es decir suficientemente precisa y confiable.

El proceder habitual en los trabajos de investigación es calcular mediante fórmulas el tamaño de la muestra fijando, previamente la precisión y la confianza que queremos tener, de forma que los resultados tengan el nivel de interpretabilidad deseada.

Las fórmulas que se emplean con más frecuencia son las siguientes:

$$n = \frac{z^2 \cdot p \cdot (1 - p)}{i^2}$$

Si la variable a medir es cualitativa, es decir si en cada unidad de estudio se mide la presencia o ausencia de algo (una cualidad determinada), y el resultado de toda la muestra se resume en proporción o porcentaje de individuos que tienen esa cualidad. Por ejemplo cumplimiento o no de un determinado criterio.

La confianza es una medida de otro posible error del muestreo, y representa la probabilidad de que el resultado que hemos obtenido con la muestra (precisión incluida) sea cierto. Esta confianza o probabilidad de certeza se expresa en % y se suele establecer como valor aceptable un 95%.

En los trabajos de investigación suelen emplearse fórmulas para calcular el número de casos necesarios que garantice una precisión y confianza previamente definidas.

Si la variable es cuantitativa, es decir si lo que medimos en cada individuo de la muestra es la magnitud o cantidad en que la variable en estudio está presente (por ejemplo, peso, talla, TA, tiempo de espera, etc.), y el resultado de toda la muestra se resume como cantidad media (peso medio, talla media, etc.), se utiliza la siguiente fórmula:

$$n = \frac{z^2 \cdot \sigma^2}{i^2}$$

En estas fórmulas,

n = tamaño de la muestra que necesitamos

z = un número que depende de la confianza que queremos tener en el resultado; este número representa valores de la distribución normal que se busca en las tablas correspondientes, y que es **1,96** para una confianza de 95%, que es la más habitual en estudios muestrales.

i = el intervalo o precisión, expresado en proporción (tanto por uno) para las variables cualitativas y en unidades de medida de la variable en estudio para las cuantitativas. Representa la precisión deseada. Por ejemplo 0,1, si queremos que el valor real esté en el intervalo comprendido entre lo que resulte en la muestra $\pm 0,1$ (equivale a $\pm 10\%$).

p = la proporción esperada de lo que se va a medir. Como habitualmente no se sabe se pone 0,5 (equivalente a 50%), que es el valor de p que da un número mayor de n , para garantizar que la muestra tendrá en cualquier caso la precisión y la confianza prefijadas.

σ = la desviación estándar que se espera va a tener la variable a medir, expresada en unidades de medida de esa variable.

Si el universo o marco muestral (que se representa como N) es suficientemente pequeño (como regla, si la n que hemos calculado es mayor del 10% de N) se calcula un tamaño de muestra más reducido (n'), pero que mantiene la misma precisión y confianza, con la fórmula

$$n' = \frac{n}{1 + \frac{n-1}{N}}$$

Dado que los criterios a utilizar para evaluar el nivel de calidad son habitualmente variables cualitativas, se adjunta una tabla con el número de casos necesarios para asegurarnos una confianza de 95% y diferentes intervalos (precisión), para diversas proporciones de cumplimiento esperadas (p). Se ha destacado la columna que corresponde a la proporción esperada del 50% y los intervalos de $\pm 0,05$ y $\pm 0,1$. (es decir $\pm 5\%$ y $\pm 10\%$) (Tabla 9.1).

Para las evaluaciones de calidad se aconseja utilizar un tamaño de muestra manejable (50-60 casos), que garantice la factibilidad del estudio.

El utilizar estas reglas prácticas para determinar el tamaño de la muestra puede parecer restar "cientifismo" a las evaluaciones de la calidad asistencial. Sin embargo, una vez realizado el muestreo aleatorio de este número relativamente pequeño de casos, nada impide enmendar la falta de "cientifismo" y calcular a posteriori el intervalo de confianza que corresponde a los valores en-

La muestra de 50-60 casos se extrae de forma aleatoria para que sea representativa, y después calculamos el intervalo de confianza del resultado obtenido

TABLA 9.1. Número de casos necesarios para asegurarnos una confianza de 95% y diferentes intervalos (precisión), para diversas proporciones de cumplimiento esperadas (p).

| (p) | 0,05 | 0,10 | 0,15 | 0,20 | 0,25 | 0,30 | 0,35 | 0,40 | 0,45 | 0,50 | 0,55 | 0,60 | 0,65 | 0,70 | 0,75 | 0,80 | 0,85 | 0,90 | 0,95 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0,01 | 1825 | 3457 | 4898 | 6147 | 7203 | 8067 | 8740 | 9220 | 9508 | 9604 | 9508 | 9220 | 8740 | 8057 | 7203 | 6147 | 4898 | 3457 | 1825 |
| 0,02 | 456 | 864 | 1225 | 1537 | 1801 | 2017 | 2185 | 2305 | 2377 | 2401 | 2377 | 2305 | 2185 | 2017 | 1801 | 1537 | 1225 | 864 | 458 |
| 0,03 | 203 | 384 | 544 | 683 | 800 | 896 | 971 | 1024 | 1056 | 1067 | 1056 | 1024 | 971 | 896 | 800 | 683 | 544 | 384 | 203 |
| 0,04 | 114 | 216 | 306 | 384 | 450 | 504 | 546 | 576 | 594 | 600 | 594 | 576 | 546 | 504 | 450 | 384 | 306 | 216 | 114 |
| 0,05 | 73 | 138 | 198 | 246 | 288 | 323 | 350 | 369 | 380 | 384 | 380 | 369 | 350 | 323 | 288 | 246 | 198 | 138 | 73 |
| 0,06 | 51 | 96 | 136 | 171 | 200 | 224 | 243 | 256 | 264 | 267 | 264 | 256 | 243 | 224 | 200 | 171 | 136 | 96 | 51 |
| 0,07 | 37 | 71 | 100 | 125 | 147 | 165 | 178 | 188 | 194 | 196 | 194 | 188 | 178 | 165 | 147 | 125 | 100 | 71 | 37 |
| 0,08 | 29 | 54 | 77 | 96 | 113 | 126 | 137 | 144 | 149 | 150 | 149 | 144 | 137 | 126 | 113 | 96 | 77 | 54 | 29 |
| 0,09 | 23 | 43 | 60 | 76 | 89 | 100 | 108 | 114 | 117 | 119 | 117 | 114 | 108 | 100 | 89 | 76 | 60 | 43 | 23 |
| 0,10 | 18 | 35 | 49 | 61 | 72 | 81 | 87 | 92 | 95 | 96 | 95 | 92 | 87 | 81 | 72 | 61 | 49 | 35 | 18 |
| 0,11 | 15 | 29 | 40 | 51 | 60 | 67 | 72 | 76 | 78 | 79 | 78 | 76 | 72 | 67 | 60 | 51 | 40 | 29 | 15 |
| 0,12 | 13 | 24 | 34 | 43 | 50 | 56 | 61 | 64 | 66 | 67 | 66 | 64 | 61 | 56 | 50 | 43 | 34 | 24 | 13 |
| 0,13 | 11 | 20 | 29 | 36 | 43 | 48 | 52 | 55 | 56 | 57 | 56 | 55 | 52 | 48 | 43 | 36 | 29 | 20 | 11 |
| 0,14 | 9 | 18 | 25 | 31 | 37 | 41 | 45 | 47 | 49 | 49 | 49 | 47 | 45 | 41 | 37 | 31 | 25 | 18 | 9 |
| 0,15 | 8 | 15 | 22 | 27 | 32 | 36 | 39 | 41 | 42 | 43 | 42 | 41 | 39 | 36 | 32 | 27 | 22 | 15 | 8 |
| 0,20 | 5 | 9 | 12 | 15 | 18 | 20 | 22 | 23 | 24 | 24 | 24 | 23 | 22 | 20 | 18 | 15 | 12 | 9 | 5 |
| 0,25 | 6 | 8 | 8 | 10 | 12 | 13 | 14 | 15 | 15 | 15 | 15 | 14 | 14 | 13 | 12 | 10 | 8 | 6 | 6 |

Se ha destacado la columna que corresponde a la proporción esperada del 50% y los intervalos de $\pm 0,05$ y $\pm 0,1$. (Es decir $\pm 5\%$ y $\pm 10\%$)

contrados (lo cual realizaremos en la UT11). Muy probablemente, estaremos suficientemente satisfechos con la información obtenida por las siguientes razones:

- **Necesidad relativa de estimaciones muy precisas.** En GC no es de primordial interés en general el tener una gran precisión sobre el grado de cumplimiento de los criterios, sino el tener una idea aproximada que ilustre sobre todo si nos encontramos muy por debajo o notablemente por encima. Intervalos de hasta $\pm 0,15$ (o a veces más) no nos van a modificar las conclusiones sobre el tema en relación a las necesidades de introducir medidas correctoras. Piénsese por ejemplo en que quisiéramos un cumplimiento de determinado criterio en al menos un 75% de los casos, y nos sale en la muestra un 30% ± 20 : tendríamos igualmente que actuar si la proporción real de cumplimiento fuese del 10% que si fuese del 50%, que son los dos extremos del intervalo de valores posibles.
- **Importancia de la factibilidad del proyecto.** En GC, tanto al inicio como para el mantenimiento de las actividades, es de gran trascendencia mantener unos niveles lo más altos posible de factibilidad de los proyectos. Factibilidad que no hay que sacrificar a la precisión del muestreo, aunque sí combinar ambas características sabiendo que la primera va a ser determinante del éxito de la implantación rutinaria de las actividades. Con muestras relativamente pequeñas, las estimaciones pueden resultar poco precisas, pero, si el muestreo está bien hecho, respetando la aleatoriedad, también sabemos que la estimación nos aporta información válida.
- **Importancia de evitar muestreos ineficientes.** En las fórmulas habituales para calcular el número necesario de casos en la muestra, por ejemplo para una variable cualitativa que es lo más frecuente en GC, hay que introducir no sólo la precisión que buscamos y el nivel de confianza que queremos obtener, sino también el valor estimado a priori que tiene la variable en estudio en la población de donde se extrae la muestra. Esta estimación habitualmente no la tenemos al estudiar por primera vez un tema concreto en GC. Como consecuencia, es habitual calcular la fórmula situándonos "en el peor de los casos", es decir aquel valor esperado (p) de la variable en estudio que precisaría de un número más elevado de casos para ser estudiada con resultados que tengan intervalos de confianza no más grandes que el previamente fijado. En variables cualitativas "el peor de los casos" es que la variable tenga un valor estimado de $p=0,5$ (50%). La aplicación de esta regla, sumamente extendida en los trabajos de investigación que no pueden permitirse realizar muestreos que produzcan resultados de una precisión relativamente baja, conduce en el campo de la GC a la realización de un trabajo extra, al muestrear un número de casos mayor que el necesario para decidir, y que no se traduce en unos mejores resultados para el Programa.

Nada impide si se tienen los recursos suficientes (ganas y tiempo) emplear las fórmulas habituales para calcular el tamaño de la muestra; con ello nos garantizamos una precisión determinada; sin embargo, para una evaluación de calidad asistencial que inicia un ciclo de mejora, es preferible realizar de entrada una muestra pequeña (hasta un máximo de 100 casos), calcular a posteriori el intervalo de confianza, y, en caso de que la información obtenida no resulte satisfactoria para los objetivos de GC ampliar el número de casos de la muestra. Para lo cual ya dispondremos como mínimo de una estimación del valor de p , que va a ser mejor, en cuanto al tamaño de la muestra necesario, que situarnos "en el peor de los casos".

En las evaluaciones de la calidad, estimaciones no muy precisas dan información suficiente para decidir

Los proyectos de mejora de calidad deben de ser factibles.

Muestras grandes pueden impedir que el proyecto se lleve a cabo por el esfuerzo inicial que supone

Utilizar en principio muestras pequeñas evita muestreos ineficientes. Respetando la aleatoriedad, siempre tenemos la opción de ampliar el tamaño de la muestra si es necesario

El principal inconveniente que puede tener el evaluar la calidad en base a muestras pequeñas es que al comparar los resultados iniciales con evaluaciones posteriores, sólo vamos a encontrar significativas las diferencias importantes; las variaciones pequeñas no van a ser detectadas o, si aparecen en los resultados, probablemente carecerán de significación estadística. En base a lo alejados que se encuentren de nuestro estándar los primeros resultados, podremos decidir ampliar la muestra para la segunda evaluación o para ambas.

Una circunstancia frecuente e importante que puede complicar el muestreo es el hecho de que los diferentes criterios que vayamos a utilizar en la evaluación no sean aplicables a la misma unidad de estudio sino a "sub" unidades. Por ejemplo, en el caso de utilizar criterios secuenciales, encadenados, en los cuales el primero de ellos ha de cumplirse en todo un grupo de población (por ejemplo detección de cáncer en mujeres de edad determinada), mientras que los siguientes aplican sólo a sectores incluidos en el primer grupo (por ejemplo criterios de seguimiento caso de encontrar citología o exploración mamaria positivas). En estos casos el universo de referencia para cada criterio es diferente: en nuestro ejemplo, una muestra realizada para evaluar la realización o no de las actividades de screening, difícilmente va a servirnos también para evaluar adecuadamente el cumplimiento de los criterios de seguimiento. En consecuencia, puede ser conveniente realizar muestreos diferentes para el estudio de cada criterio (con referencia a su universo específico de aplicación), o realizar un muestreo amplio hasta cubrir un mínimo de 30 casos en los que pueda evaluarse el criterio aplicable al subgrupo de población más pequeño. Siempre pensando en la factibilidad de los proyectos, es aconsejable evitar en las primeras evaluaciones criterios que sólo sean aplicables a subgrupos de pacientes o población de número reducido y que no puedan ser fácilmente identificados para su muestreo.

c) Métodos de Muestreo

Ya hemos señalado que la *representatividad* de la muestra depende del método de muestreo, y más concretamente de que éste no sea sesgado y esté hecho de manera que todos los elementos del universo o marco muestral tengan la misma oportunidad de ser incluidos en la muestra, es decir, que el muestreo sea *aleatorio*.

Existen diversos métodos de muestreo que respetan la aleatoriedad. Los más comunes y útiles para los estudios de calidad son el muestreo aleatorio simple, el sistemático, el estratificado y combinaciones de ellos (por ejemplo estratificado y sistemático).

Los métodos que respetan la aleatoriedad nos aseguran la representatividad estadística, pero hay otros métodos de muestreo, encuadrables en lo que se conoce como muestreos de conveniencia, relativamente frecuentes y útiles pero sobre cuyas limitaciones debemos estar avisados.

Vamos a revisar las características y mecánica de todos ellos. En el diseño del estudio hay que dejar bien claro cuál es el método de que vamos a utilizar, a menos que midamos todos los casos, lo cual debe también ser consignado.

- **Muestreo aleatorio simple.** En el muestreo aleatorio simple, cada caso de la muestra se extrae al azar y de forma independiente. Para ello se utilizan tablas de números aleatorios, u otra forma de generarlas (calculadoras, listas de lotería, etc.).

Si queremos detectar diferencias pequeñas en la reevaluación, conviene utilizar muestras más grandes.

En las evaluaciones con criterios secuenciales (encadenadas) considerar un muestreo que nos garantice un mínimo de 30 casos para el criterio que se evalúa en el último eslabón de la cadena.

Para que la muestra sea representativa el muestreo debe ser aleatorio

En el diseño del estudio se debe especificar el método de muestreo a emplear

Para el muestreo aleatorio simple todos los elementos del universo han de estar numerados y localizables.

El muestreo sistemático es muy cómodo de realizar, sólo exige saber el total del universo y el número de casos que queremos. Con ello se calcula el intervalo de muestreo. Hay que estar alerta, en cambio, a los factores que intervienen en la forma en que está ordenado el universo para evitar sesgos.

Para realizarlo es preciso que todos los elementos del universo o marco muestral estén numerados y cada número sea fácilmente localizable. La forma en que funciona, una vez que se determina el tamaño de la muestra (n), es como sigue:

- 1º Se dispone de un universo o marco muestral (N) previamente definido y en el que todos sus elementos están numerados (de 1 a N) y localizables.
- 2º Con una tabla de números aleatorios, u otra manera de generar números aleatorios, se van seleccionando n números que han de estar comprendidos entre 1 y N .
- 3º Estos números corresponden a los casos que hay que extraer del universo de casos, y que van a componer la muestra. Por ejemplo, supongamos que tenemos un universo de 500 historias de hipertensos y queremos sacar una muestra aleatoria simple de 50 historias. Para ello generamos 50 números aleatorios comprendidos entre 1 y 500 (o entre los 500 números que le hayamos asignado a las historias), y extraeremos para evaluar las 50 historias que tengan asignadas los números aleatorios obtenidos.

El principal inconveniente del muestreo aleatorio simple es la necesidad de que el universo esté numerado y cada número fácilmente localizable (piénsese que los números aleatorios pueden saltar de una parte a otra del universo).

- **Muestreo Sistemático.** El muestreo sistemático se realiza extrayendo un caso para la muestra cada cierto intervalo fijo de casos. Este intervalo, conocido como intervalo de muestreo (K) se calcula dividiendo el número total de casos del marco muestral (N) por el tamaño de la muestra (n). Para que sea aleatorio, *el primer caso tiene que extraerse de manera aleatoria*, de entre todo el universo o, lo que es más cómodo, de entre los primeros K casos.

En el ejemplo anterior $N= 500$; $n= 50$

$$k = \frac{500}{50} = 10 \rightarrow$$

elegimos un número al azar del 1 al 10, y ésta será la primera historia clínica para la muestra; a continuación escogemos *sistemáticamente* una historia de cada 10. Es decir, contando nueve historias entre caso y caso para nuestra muestra.

Este proceder es mucho más cómodo que el muestreo aleatorio simple, pero, dado que los casos se extraen de forma cíclica, tenemos que estar seguros que el universo no está ordenado de forma también cíclica en base a un variable o factor que influya en el criterio que queremos medir. Por ejemplo, supongamos que queremos medir la proporción de visitas a las que se le toma la tensión y que utilizamos como marco muestral las hojas de morbilidad o un registro similar de los pacientes atendidos cada día; supongamos además que algunos de los médicos del centro tienen por costumbre tomar tensiones los miércoles. Si realizamos un muestreo sistemático sobre las hojas de consulta de, digamos un mes, puestas en orden cronológico día tras día corremos el riesgo de sobrestimar o subestimar la variable que queremos medir porque en el universo va a estar sistemáticamente agrupada en determinados días que pueden saltarse sistemáticamente o ser incluidos sistemáticamente en la muestra. En este caso el muestreo sistemático no sería aconsejable.

- **Muestreo Estratificado.** El muestreo estratificado es aquel en el que la muestra (n) se subdivide en estratos o submuestras extraídas cada una de los diferentes estratos en que a su vez está dividido el universo.

Los estratos del universo o marco muestral se definen en función de las categorías o posibilidades de la variable o factor por el que nos interesa estratificar, y que habitualmente es algo que sabemos influye en lo que queremos medir, y que por lo tanto no interesa controlar su representatividad en la muestra.

En el ejemplo anterior, dado que sabemos que el día de la semana influye en la proporción de pacientes a los que se le toma la tensión, nos interesaría hacer un muestreo estratificado por días de la semana; de forma que dividiríamos la muestra en cinco submuestras (exceptuamos sábado y domingo) y cada una se extraería de entre los pacientes atendidos en cada uno de los cinco días de la semana en los que hay consulta normal.

Otra situación muy frecuente y que aconseja la estratificación es cuando hacemos un muestreo que queremos que sea representativo de todo el servicio, pero sabemos o sospechamos que puede haber grandes diferencias en el cumplimiento de los criterios entre los diferentes médicos. En este caso conviene también asegurarnos la representatividad de la muestra, considerando a cada médico como un estrato y realizando un muestreo estratificado.

La estratificación de la muestra puede realizarse de forma proporcional o no a la contribución de cada estrato al total del marco muestral.

Para la estratificación *proporcional* se ha de calcular primero en qué proporción contribuye cada estrato al total del universo o marco muestral. Por ejemplo supongamos que queremos estratificar por médico y que en el servicio trabajan 4 médicos; y que el total de casos del universo es 500, pero que a estos 500, el médico 1 contribuye con 100, el médico 2 con 50, el médico 3 con 200 y el médico 4 con 150; supongamos además que queremos extraer una muestra de 60 casos y la queremos estratificar, de forma proporcional. Calcularemos entonces en qué proporción contribuye cada médico al universo:

$$\text{Médico 1: } = \frac{100}{500} = 0,2 \rightarrow 20\%$$

$$\text{Médico 2: } = \frac{50}{500} = 0,1 \rightarrow 10\%$$

$$\text{Médico 3: } = \frac{200}{500} = 0,4 \rightarrow 40\%$$

$$\text{Médico 4: } = \frac{150}{500} = 0,3 \rightarrow 30\%$$

La muestra se divide entonces en estratos proporcionales a estos mismos valores. Así:

De los pacientes del médico 1 se extraen un 20% de 60 \rightarrow 12 casos

De los pacientes del médico 2 se extraen un 10% de 60 \rightarrow 6 casos

De los pacientes del médico 3 se extraen un 40% de 60 \rightarrow 24 casos

De los pacientes del médico 4 se extraen un 30% de 60 \rightarrow 18 casos.

En el muestreo estratificado la muestra se subdivide en submuestras que se extraen de los correspondientes "estratos" en que se ha dividido el universo. Los estratos son agrupaciones de los casos en función de la variable (por ejemplo médicos) que queremos controlar con la estratificación.

El muestreo estratificado puede ser proporcional o no proporcional al tamaño de los estratos. En cada estrato el muestreo ha de ser aleatorio, simple o sistemático.

En el diseño de estudio hay que considerar el mecanismo de sustitución de los casos que no se encuentren o no puedan ser evaluados.

En el muestreo estratificado no proporcional se extrae un número fijo o un porcentaje de la muestra fijo de cada estrato, con independencia del tamaño del estrato. El que se haga proporcional o no, va a influir en los cálculos necesarios para valorar el resultado, como veremos en la UT 11.

Cada submuestra (la muestra de cada estrato) hay que extraerla a su vez de forma aleatoria, sea de forma aleatoria simple o sistemática. El muestreo estratificado y sistemático es muy frecuente en los estudios del nivel de calidad.

- **Muestreo de conveniencia.** El muestreo de conveniencia se caracteriza por ser de fácil obtención pero no tiene garantizada la representatividad estadística. Son muestreos en los que se eligen los casos que están más a mano, para tener una idea de si hay o no problema. Por ejemplo, decidimos evaluar los pacientes que vengan hoy, o los que vinieron ayer, o la semana pasada, o las primeras veinte historias, etc. Normalmente se establece de antemano un número de casos también cómodo (muestreo tipo cuota), e incluso se aplican metodologías semejantes a la estratificación (muestreos a propósito) eligiendo cuotas de casos extraídos a conveniencia de grupos diferentes de pacientes (de diferentes médicos, por ejemplo) sobre los que se quiere saber si existen o no diferencias.

No es un muestreo aconsejable para los estudios de nivel de calidad, pero no es infrecuente el que se realice este tipo de muestreos (por ejemplo evaluar los casos atendidos en el último mes, o los últimos 6 meses) sin reflexionar hasta qué punto es una muestra representativa o no. A veces, puede parecer que un muestreo de conveniencia sí es representativo, y al revés: muestreos que se suponen representativos son en realidad de conveniencia, depende de si el marco muestral escogido para el muestreo de conveniencia representa razonablemente la población o marco muestral más amplio al que queremos generalizar los resultados de la medición. Por ejemplo, si los pacientes de la última semana representan los que podían verse en cualquier semana del año, los resultados pueden ser generalizados a todo el año. La reflexión sobre la representatividad de la muestra (control del error de sesgo) debe de hacerse *siempre* en el momento del diseño del estudio.

Una cuestión común a todos los tipos de muestreo y que hay que tener previsto también en el diseño del estudio es el *mecanismo de sustitución de los casos* cuando éstos no se encuentran (por ejemplo historias que no están). El mecanismo de sustitución debe explicitarse de antemano y ser siempre el mismo en todo el muestreo. Una forma común es elegir el número o la historia que hay en el listado inmediatamente después de la que ha fallado.

3.7. TIPO DE EVALUACIÓN

El tipo de evaluación se define en función de la relación temporal con la acción evaluada, quién toma la iniciativa de evaluar y quién obtiene los datos.

a) Tipo de Evaluación, según relación temporal

Si lo que se evalúa ha ocurrido ya, es una evaluación *retrospectiva*; si está ocurriendo a la vez que se evalúa, es *concurrente*. La mayoría de las evaluaciones del nivel de calidad son retrospectivas: se evalúa algo que ha ocurrido y por lo tanto si ha habido problema o defecto de calidad es ya inevitable; la utilidad

Los muestreos de conveniencia, aunque puedan aportar información de interés, generalmente no son estadísticamente representativos.

de estas evaluaciones estriba en aprovechar esta información para diseñar intervenciones que eviten que el problema vuelva a ocurrir.

En las evaluaciones concurrentes, puesto que se realizan antes de que se dé por terminada la acción que se evalúa, es posible introducir las medidas correctoras antes de que se consume un acto asistencial de mala calidad. Por ejemplo, si evaluamos la calidad de la información que recibe el paciente sobre su tratamiento, en base a un diseño de encuesta realizada a los usuarios antes de que abandonen el centro sometido a evaluación, y en la evaluación encontramos que el paciente no tiene claros algunos criterios empleados en la evaluación (dosis, pauta, contraindicaciones, tiempo de duración del tratamiento), se le aclara y nos aseguramos su comprensión antes de que abandone la institución. En la asistencia hospitalaria hay más ejemplos de evaluación concurrente, algunos de ellos realizados de forma rutinaria aunque no siempre con criterios de evaluación explícitos. Citemos como ejemplos relativamente rutinarios la solicitud de segundas opiniones antes de proceder a determinada intervención o exploración, la práctica de aprendizaje supervisada por un docente que interviene en caso necesario modificando un curso de acción determinado que le parece erróneo, o la realización de biopsias intraoperatorias antes de proceder a la realización de determinadas intervenciones quirúrgicas. Las evaluaciones concurrentes, si se encuentra la ocasión y la oportunidad de ser realizadas, pueden convertirse con facilidad en parte integrante del proceso asistencial en sí.

La evaluación prospectiva (antes de que ocurra lo que se evalúa), es un término equívoco que preferimos no utilizar al hablar de estudio de nivel de calidad, puesto que responde más bien a actividades de diseño de la calidad, que tratan de prevenir la aparición de problemas.

b) Tipo de evaluación, según quién toma la iniciativa

Existen dos tipos complementarios de organización de las actividades de GC: programas internos y programas externos. El esquema metodológico es en esencia el mismo, pero se diferencian en quienes toman la iniciativa de las evaluaciones y demás actividades del Programa: en los programas internos la responsabilidad de las actividades y la iniciativa de las mismas recae en los propios profesionales cuya asistencia se evalúa; en los programas externos, la responsabilidad de las actividades y la iniciativa de las mismas recae en organismos externos a los centros o profesionales evaluados; estas instancias externas pueden ser niveles superiores de la organización administrativa del sistema de salud, u organizaciones profesionales de carácter independiente.

Merece la pena aclarar que, el tipo de evaluación sea interna, no quiere decir necesariamente que sean los propios profesionales del centro los que lleven a cabo la recogida física de datos: en determinadas situaciones estructurales favorables es perfectamente posible, y a veces deseable, que la recogida de datos la realice personal ajeno al centro. Una revisión interna se caracteriza por responder a un diseño y llevar a una explotación de los resultados para mejorar la asistencia, que responden a la iniciativa consciente y responsable de los profesionales de la institución.

c) Tipo de evaluación, según quién obtiene los datos

Ya hemos visto que la iniciativa de la evaluación puede ser interna o externa,

También hay que explicitar si la evaluación es interna, decidida por los profesionales evaluados, o externa.

La evaluación puede realizarse de forma cruzada, por personal ajeno o evaluándose uno mismo sus propios datos. El proceder concreto debe consignarse en el diseño del estudio.

y que debe ser consignado este punto en el diseño del estudio. Adicionalmente hay que considerar y explicar quién se encarga de obtener los datos y cómo.

En las evaluaciones internas, es importante considerar si la evaluación será cruzada (un profesional evalúa la actuación de otro) o autoevaluación (cada uno recoge los datos de su propia actuación). Esta segunda modalidad no es totalmente desdeñable en los estudios de calidad, porque se supone que nadie está interesado en engañarse a sí mismo, sino en saber en qué debe mejorar; sin embargo, al menos en los primeros proyectos, quizá sea más aconsejable la evaluación cruzada o mixta. También puede ser realizada por personal ajeno al proceso evaluado.

BIBLIOGRAFÍA

- Palmer RH. Evaluación de la asistencia ambulatoria. Principios y práctica. Madrid: Ministerio de Sanidad y Consumo; 1990.
- Saturno PJ. Cómo diseñar una evaluación sobre la calidad asistencial en Atención Primaria. IV Congreso de la Sociedad Andaluza de Medicina Familiar y Comunitaria. Sevilla: Consejería de Salud; 1992.
- Briones G. Métodos y técnicas de investigación para las ciencias sociales. México: Trillas; 1990.
- Lemeshow S, Hosmer DW, Klar J, Lwanga SK. Adequacy of sample size in health studies. Sussex: WMO/John Wiley & Son; 1992.

10

**DISEÑOS DE ESTUDIOS DE
INVESTIGACIÓN DE
CAUSAS DE LA CALIDAD
DEFICIENTE**

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

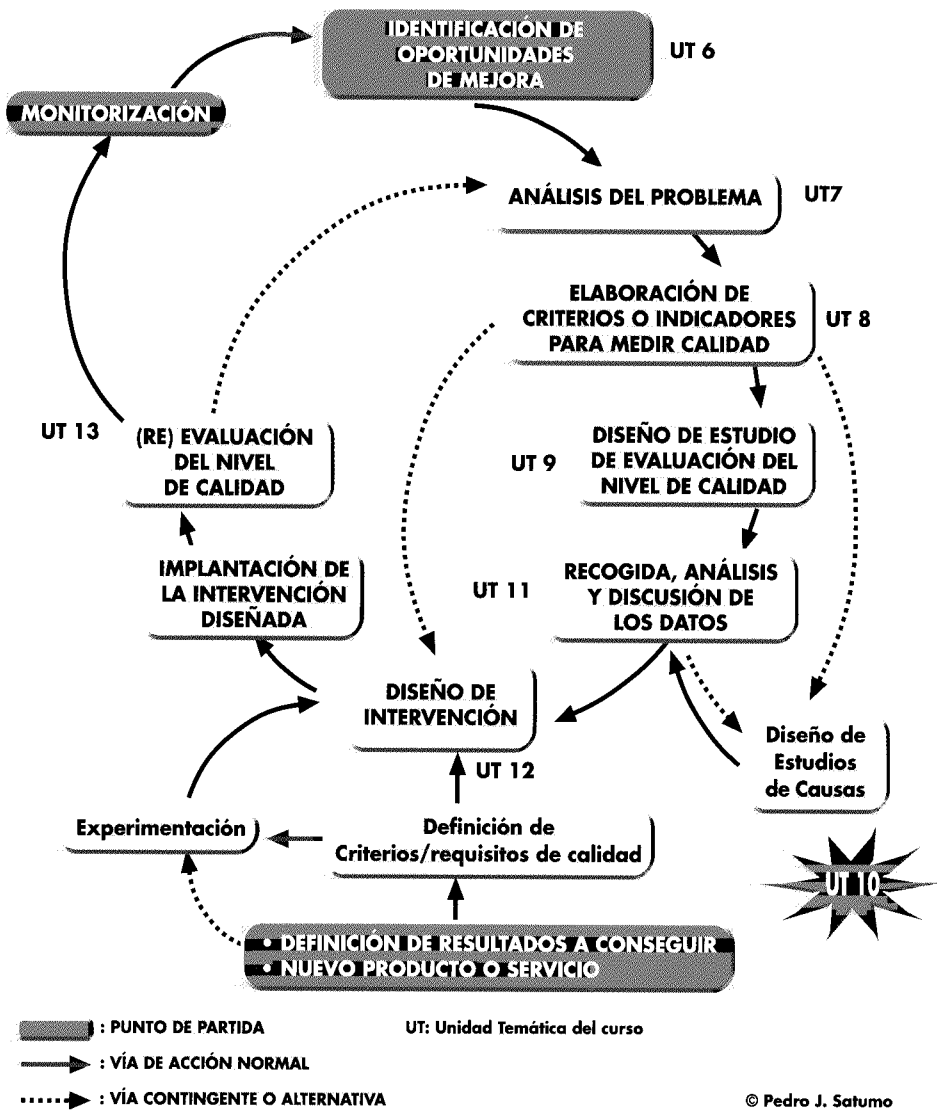
Esta UT complementa a la anterior, detallando las características y metodología de los estudios de causas hipotéticas. Se revisan métodos con y sin datos para identificar las causas sobre las que intervenir para mejorar. Para el estudio de causas con datos se indican los métodos, test estadísticos y herramientas gráficas que más conviene utilizar en función del tipo de variables (cuantitativa o cualitativa) con que se miden las causas hipotéticas y los criterios o indicadores de calidad.

ÍNDICE DE CONTENIDOS

1. Introducción.
2. Métodos generales para los estudios de causas hipotéticas.
3. Búsqueda de causas hipotéticas sin datos.
4. Búsqueda de causas con datos: comprobación de hipótesis.

OBJETIVOS ESPECÍFICOS

1. Identificar el tipo de estudio que conviene para avanzar en el ciclo de mejora, según las causas a indagar.
2. Distinguir métodos de estudio de causas hipotéticas con y sin datos.
3. Diseñar y analizar una encuesta de causas.
4. Diseñar y analizar una votación múltiple sobre causas hipotéticas.
5. Identificar el tipo de variable que corresponde a cada tipo de variables.
6. Seleccionar el test estadístico que corresponde a cada tipo de variables.
7. Describir el fundamento de los análisis de estratificación.
8. Describir el fundamento de los análisis de correlación y regresión.
9. Utilizar e interpretar adecuadamente un diagrama de dispersión.
10. Realizar la prueba de los signos en un diagrama de dispersión.



1. INTRODUCCIÓN

Tras seleccionar la oportunidad de mejora o problema de calidad (UT6) podemos preguntarnos de qué forma puede abordarse para conseguir su resolución. Para ello requerimos indagar sobre sus causas. En consecuencia realizamos un análisis estructurado de la oportunidad de mejora (UT 7). Al realizar este análisis, lo más probable es que nos encontremos con diversos tipos de causas frente a las cuales la acción más oportuna es también diferente. Los dos cursos de acción a los que nos va a guiar el análisis del problema son la intervención directa y/o el diseño de estudios para un conocimiento cuantificado de las causas, sobre el que basar la intervención.

Los estudios para un conocimiento cuantificado de las causas pueden definirse en dos grandes grupos, según que las causas a indagar sean criterios o requisitos de calidad (estudios de nivel de calidad, UT9) o bien factores hipotéticos que pueden o no relacionarse con el problema de calidad analizado (estudios de

En esta UT nos vamos a ocupar de los estudios de causas hipotéticas de la mala calidad: aquellas que no son requisitos de calidad (criterios), pero que pensamos que pueden influir en que esos requisitos se den o no.

causas hipotéticas). Este último tipo de estudios es el contenido fundamental de esta unidad.

2. MÉTODOS GENERALES PARA LOS ESTUDIOS DE CAUSAS HIPOTÉTICAS

Los estudios de causas hipotéticas son los que realizamos en general cuando el problema a analizar y mejorar no es de tipo clínico sino organizacional (tiempo de espera, falta de coordinación entre los profesionales, etc.), que podemos tenerlo incluso ya medido, pero que para su mejora tenemos que averiguar las causas que lo provocan; o bien después de los estudios de nivel de calidad, si no sabemos qué hacer para mejorar un determinado criterio clínico que no se cumple. En ambos casos, lo que tenemos va a ser una serie de causas hipotéticas que investigar y es necesario medir concurrentemente con las posibles causas del criterio o los criterios que nos representan la calidad del aspecto en estudio.

En ambos casos también, es necesario tener definido el criterio o criterios que nos miden el nivel de calidad del problema a analizar y mejorar.

En los estudios de causas hipotéticas, lo más conveniente sería diseñar estudios que las cuantifiquen. Sin embargo, en la práctica de la evaluación y mejora de la calidad en los servicios de salud se han aplicado también otros métodos que aprovechan de forma estructurada el conocimiento que pueden tener sobre el problema sus propios protagonistas. Son métodos como la *encuesta de causas* o la *votación múltiple* que parten o se complementan con los de análisis del problema, y en los cuales se "cuantifican" las causas según la impresión subjetiva de los proveedores.

Cuando queremos que la cuantificación sea objetiva, podemos recurrir a un diseño de investigación que no se diferencia de los habituales estudios de evaluación de servicios de salud. Pero a veces procedimientos más simples, respetuosos sin embargo con el rigor científico, pueden ser suficientes para el grado de conocimiento que necesitamos. De todos estos métodos de estudio, tanto sin datos como con datos, más propios del campo de la evaluación y mejora de la calidad, tratan los siguientes apartados de esta UT.

3. BÚSQUEDA DE CAUSAS HIPOTÉTICAS SIN DATOS

Tras la identificación de posibles causas de un determinado problema de calidad por medio de métodos de consenso u otras como el diagrama de causa y efecto realizados sin datos, lo lógico sería diseñar un estudio que cuantificase su importancia relativa. Sin embargo, en algunos casos, por lo complicado del diseño, por la urgencia o la seguridad que puede tener el grupo en el conocimiento que tienen del problema, se puede proceder a señalar cuales son las principales causas a remediar por medio de métodos que tratan de cuantificar ese conocimiento subjetivo. Para ello se realiza un listado de las causas hipotéticas según el resultado del análisis realizado, y se someten bien a encuesta, bien a votación múltiple.

Las causas hipotéticas no son requisitos de calidad sino factores que pueden influir en ellos. Para estudiar si esta influencia es cierta hay que diseñar estudios que miden a la vez los requisitos de calidad (traducidos a criterios) y las causas hipotéticas.

La valoración de la importancia de las causas hipotéticas puede hacerse de forma subjetiva (sin datos empíricos) o cuantificándolos objetivamente

3.1. ENCUESTA SOBRE CAUSAS

En este enfoque, el listado de causas hipotéticas se convierte en un listado de preguntas, semejante a lo que se realiza en las encuestas de satisfacción, pidiendo al personal implicado que valore la importancia de cada una de ellas, con escalas de varias opciones (por ejemplo "no importante", "poco importante", "importante", o con más opciones).

Tanto el formato de las preguntas como la escala de valoración pueden ser de varios tipos, pero finalmente lo que persigue es tener una cuantificación/priorización de las diversas causas en relación a su importancia. El resultado de la encuesta puede presentarse y analizarse con un diagrama de Pareto (ver UT11), en donde se subrayen y prioricen las causas sobre las que diseñar la intervención en función de la valoración recibida.

Este método ha sido empleado, por ejemplo, en algunos de los proyectos pioneros de aplicación a sistemas de salud de los métodos de evaluación y mejora de la calidad propios de la industria.

La Figura 10.1 contiene un ejemplo real de resultado de encuesta de causas del retraso intrahospitalario en la aplicación de trombolíticos en el infarto de miocardio. En este caso la cuantificación, aunque subjetiva, fue acertada y la intervención diseñada sobre esta base fue efectiva. La Figura 10.2 contiene, sin embargo, un ejemplo de cuantificación mediante encuesta, también real, de las causas de no tener la historia clínica en la consulta que no resultó acertada, a tenor del escaso efecto de la intervención diseñada con estos datos.

3.2. VOTACIÓN MÚLTIPLE

La votación múltiple es también un método para seleccionar las causas más importantes de la lista que se haya elaborado. Consiste en una serie de votaciones entre los implicados en el proceso a mejorar, cada una de las cuales reduce la lista, hasta que quedan seleccionadas las causas sobre las que intervenir.

Los pasos a seguir son los siguientes:

1. Preparar el listado de causas.
2. Tras examinarlo, cada persona implicada entrega una lista en la que ha seleccionado un tercio de las causas de la lista (por ejemplo, si el listado tiene 30 causas, selecciona 10; si tiene 15, selecciona 5, etc.).
3. Se cuentan los "votos", viendo el número de menciones para cada causa de la lista, y se eliminan aquellas con menos menciones, reduciendo la lista. El número mínimo de "votos" para mantener la causa en la lista depende del tamaño del grupo que vota; la regla general es eliminar aquellas que tienen un número de votos inferior a $1/3$ del tamaño del grupo.
4. Se repiten los pasos 1 a 3, hasta que el listado es lo suficiente reducido e importante como para ser objeto de diseño de intervención para contrarrestar las causas que contiene.

La votación múltiple no deja de ser una técnica de consenso, que tiene el riesgo, como todas las técnicas que no manejan datos, de equivocarse si las

La encuesta de causas y la votación múltiple son los dos métodos más comunes para cuantificar la valoración subjetiva de las causas hipotéticas

La encuesta sobre causas pide al personal implicado en el problema a evaluar que señale la importancia relativa de cada una de las posibles causas, sobre la base de su conocimiento directo del problema

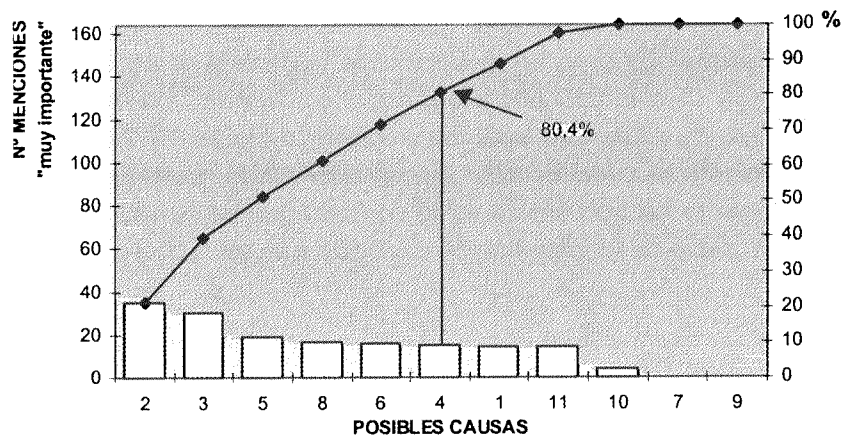
La votación múltiple busca el consenso de los implicados en el problema evaluado sobre cuales son las causas principales, a través de un proceso estructurado de votación y reducción consecuente del listado de causas inicial.

FIGURA 10.1. Encuestas de causas de retraso en la aplicación de trombolíticos

La encuesta se pasó a personal de Urgencias (n=65), pidiendo su valoración en una escala de tres opciones.

LISTADO DE POSIBLES CAUSAS DE DEMORA (en el orden en que figuran en la encuesta)

1. Pocos recursos humanos para el traslado del paciente hasta la UCI.
2. Tiempo de preparación del paciente para el traslado (bala de oxígeno, desfibrilador, etc.).
3. Inadecuada estructura física de comunicación Urgencias-UCI.
4. Interpretación del ECG por residentes.
5. Necesidad de interconsulta residentes/staff
6. Demora por evaluación de pruebas complementarias (CPK, etc.)
7. Inadecuado sistema de telefonía interna para comunicación Urgencias-UCI.
8. Demora por realización de diversas técnicas o exploraciones (Rx, vía venosa, etc.)
9. Demora por no priorización.
10. Demora en el ingreso en UCI por ocupación.
11. Otras causas.

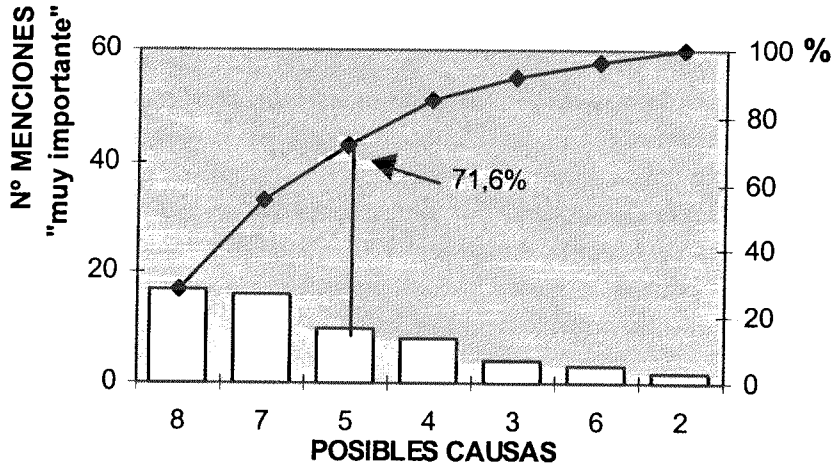


| Posibles causas de demora | Nº menciones como "muy importante" | % | % ACUMULADO |
|---------------------------|------------------------------------|--------------|--------------|
| 2 | 35 | 21,3 | 21,3 |
| 3 | 30 | 18,3 | 39,6 |
| 5 | 19 | 11,6 | 51,2 |
| 8 | 17 | 10,4 | 61,5 |
| 6 | 16 | 9,8 | 71,3 |
| 4 | 15 | 9,1 | 80,4 |
| 1 | 14 | 8,5 | 89,0 |
| 11 | 14 | 8,5 | 97,5 |
| 10 | 4 | 2,4 | 100,0 |
| 7 | 0 | 0,0 | 100,0 |
| 9 | 0 | 0,0 | 100,0 |
| total | 164 | 100,0 | 100,0 |

FIGURA 10.2. Encuesta de causas de no tener la historia clínica en la consulta

LISTADO DE POSIBLES CAUSAS DE NO DISPONER DE LA HISTORIA CLÍNICA EN LA CONSULTA

1. No tiene historia clínica
2. No se identifica bien de qué cupo son las historias
3. Retraso en la comunicación del cambio de médico al personal administrativo
4. La historia no esta archivada
5. La historia esta siendo utilizada por otro profesional en otra consulta
6. La historia esta siendo utilizada para trabajos de investigación
7. La historia no se encuentra por estar incorrectamente archivada
8. El nombre de la cita y el de la historia no coinciden



| Posibles causas de demora | Nº menciones como "muy importante" | % | % ACUMULADO |
|---------------------------|------------------------------------|--------------|-------------|
| 8 | 17 | 28,3 | 28,3 |
| 7 | 16 | 26,7 | 55,0 |
| 5 | 10 | 16,7 | 71,6 |
| 4 | 8 | 13,3 | 85,0 |
| 3 | 4 | 6,7 | 91,6 |
| 6 | 3 | 5,0 | 96,6 |
| 2 | 2 | 3,3 | 100,0 |
| total | 60 | 100,0 | |

percepciones del grupo no son las correctas. La reevaluación de la calidad en base al criterio/indicador que representa el problema, y su comparación con la situación (medición) de partida va a ser la única forma de saber si el grupo estuvo o no acertado.

Todos estos métodos, al igual que otras técnicas de consenso como el grupo nominal (UT6) o el diagrama de afinidades (UT12), pueden utilizarse tanto para priorizar causas como para diseñar la estrategia o intervención para mejorar, corriendo eso sí, el mismo riesgo de equivocarse.

4. BÚSQUEDA DE CAUSAS CON DATOS: COMPROBACIÓN DE HIPÓTESIS

Los estudios para cuantificar las posibles causas de un problema de calidad y averiguar su asociación con el criterio o indicador con el que lo medimos no son diferenciables en cuanto a su diseño de los que pueden ponerse en marcha para evaluar el nivel de calidad (UT9). La diferencia es que, en la muestra de casos en las que evaluamos la calidad (normalmente mediante un solo criterio o indicador que define el problema de calidad, por ejemplo retraso en la fibrinólisis o porcentaje de consultas sin historia, en los ejemplos de las Figuras 10.1 y 10.2) se miden también las causas, características o factores que sospechamos son de interés. Con estos datos lo normal sería proceder a valorar la asociación entre el indicador de calidad y sus posibles causas; para ello, según el esquema de la Figura 10.3, disponemos de una serie de test estadísticos (correlación de Pearson, correlación de Spearman, ANOVA, χ^2) empleables según la naturaleza estadística de las dos variables (indicador de calidad y causa) cuya asociación queremos ver. Si quisiéramos considerar simultáneamente más de una posible causa, utilizaríamos entonces técnicas de análisis multivariable, sin embargo, en la práctica normal de los ciclos de mejora es frecuente no llegar a utilizar estos análisis estadísticos.

En la práctica, los estudios de causas hipotéticas pueden resumirse en tres tipos de situaciones, y los análisis en dos enfoques principales, estratificación y correlación, adecuados cada uno de ellos según el tipo de variables de que se trate. Tanto la estratificación como la correlación pueden acompañarse o no de análisis de significación estadística, utilizando entonces el test que resulte más apropiado según el algoritmo de la Figura 10.3.

Las tres posibles situaciones que podemos tener, se definen en función de la caracterización de las causas a investigar. Son las siguientes:

A) Las causas hipotéticas son una serie o conjunto de características, circunstancias o factores de carácter cualitativo, dicotómico, de forma que mediremos su presencia o ausencia, de manera similar a lo que haríamos si fuese un listado de criterios de calidad. Por ejemplo, supongamos que por medio de un flujograma como el de la Figura 10.4 hemos identificado una serie de puntos problemáticos en el proceso de petición, realización y remisión de analítica que pueden provocar vueltas a atrás al flujo normal de actividades y por tanto causar repeticiones y retrasos; en nuestro estudio de causas mediremos la frecuencia de cada una de estas posibles vueltas a atrás (presencia/ausencia) cuando se ha producido un retraso problemático (previamente definido como problema de calidad); para estar más seguros podemos medir si estas vueltas

El diseño de los estudios para cuantificar las causas hipotéticas es idéntico al de los estudios de nivel de calidad, si bien en cada caso hay que medir tanto el criterio o indicador de calidad como las causas en consideración. La mayor diferencia entre los dos tipos de estudios radica en el análisis estadístico de los datos

a atrás se producen también en casos que no se atrasan y comparar las frecuencias en ambos grupos, pero si el flujograma es correcto y todos los puntos problemáticos son indeseables, su presencia puede ser considerada como defectos de calidad y analizados con diagrama de Pareto (ver UT11) para priorizar las intervenciones para mejorar. Un caso semejante sería el análisis a efectuar sobre las causas de insatisfacción a través de encuesta: cada posible causa se traduce en una pregunta que es un factor o variable que puede considerarse para el análisis separadamente como satisfactoria o insatisfactoria y realizar un análisis con diagrama de Pareto para evidenciar cuales son los factores/causas más frecuentemente valorados como insatisfactorios.

B) La causa o factor a indagar es una variable cualitativa, categórica, con dos o más valores. Los casos más típicos son tipo de centro, médicos diferentes dentro de un equipo, turnos, tipo de consulta, tipo de especialista, días de la semana, etc. El análisis oportuno en estos casos es la *estratificación* o análisis estratificado del indicador de calidad según los diversos valores del factor o causa a indagar. Este método se explica más adelante en esta UT.

Cuando el criterio de calidad que representa el problema es dicotómico (cumple/no cumple) y las causas hipotéticas son asimismo un listado de características dicotómicas (presencia/ausencia) su análisis puede realizarse con un diagrama de Pareto de las veces que cada causa ha estado presente en casos con defecto de calidad.

FIGURA 10.3. Test estadísticos para asociación entre dos variables (indicador de calidad y una probable causa)

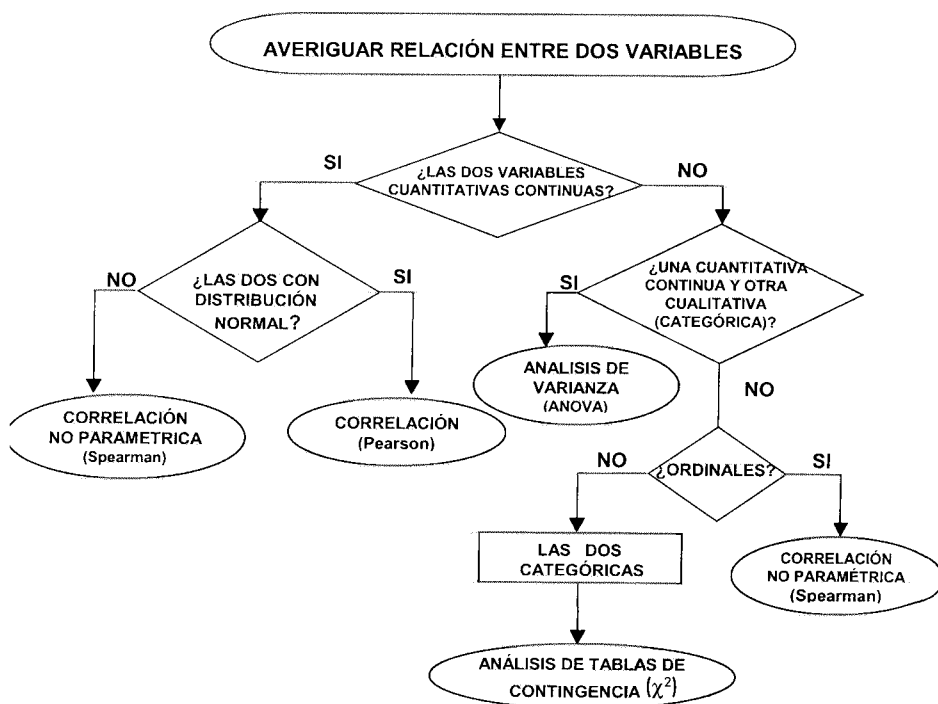
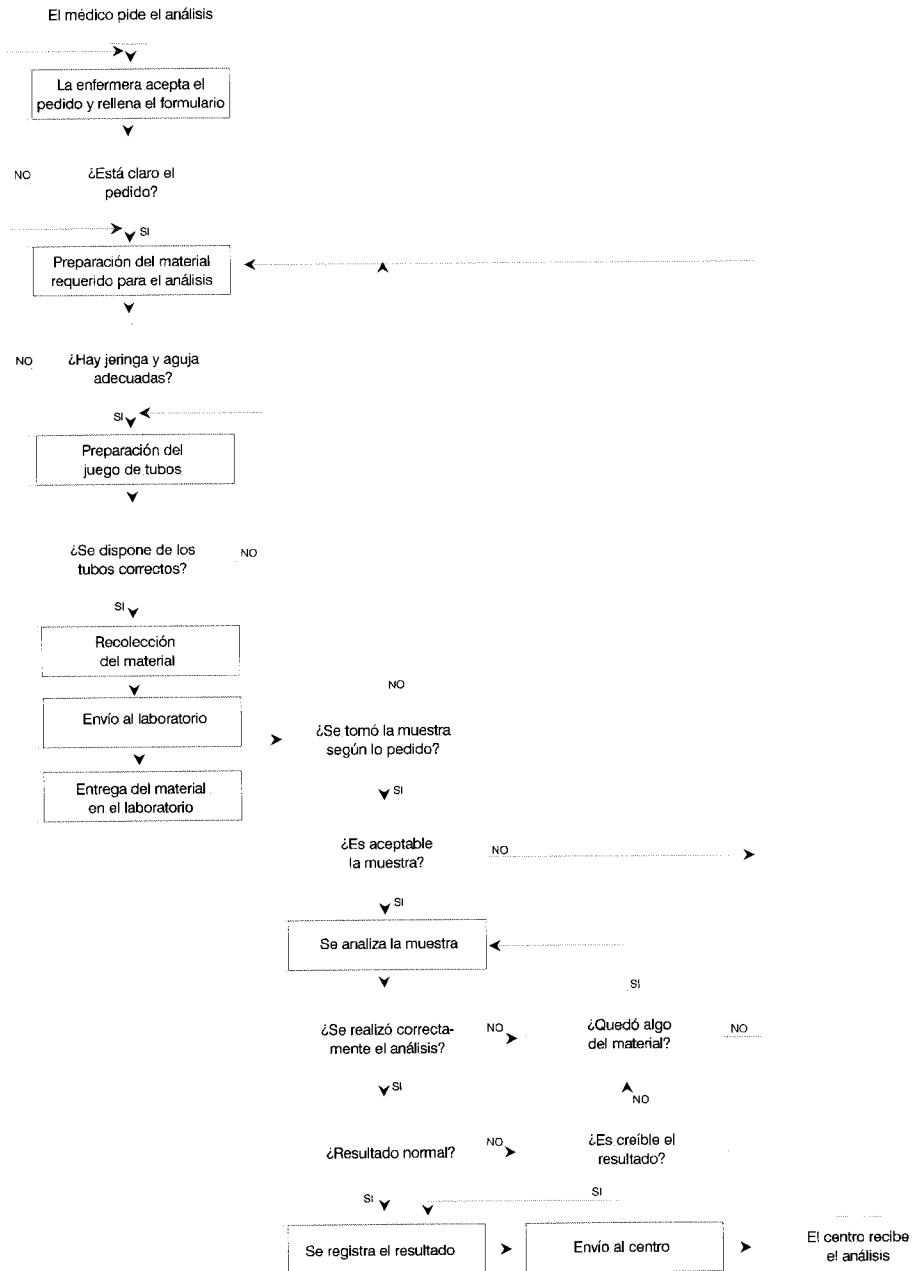


FIGURA 10.4. Flujograma del circuito de petición y realización de análisis, para identificar causas de retraso



Conclusión: identificadas las posibles causas del retraso (bucles hacia atrás) diseñamos un estudio para medir cómo es de frecuente cada una y si se asocian o no al problema analizado (no tener a tiempo el análisis).

C) La causa o factor a indagar es una variable cuantitativa, porque se mide así en cada individuo de la muestra (por ejemplo tiempo de espera, nº de visitas, etc.) o porque las unidades de estudio son agrupaciones de individuos y lo que se mide es la proporción de una determinada cualidad (por ejemplo proporción de consultas a demanda en un día, siendo el día la unidad de estudio tanto para el indicador de calidad como para el posible factor o causa asociada). En estos casos se analizará la correlación con el indicador de

calidad si éste es de tipo cuantitativo, o se realizará análisis estratificado de la probable causa según las dos posibilidades (cumple/no cumple) del indicador de calidad, si éste es de tipo cualitativo. La correlación puede analizarse gráficamente por medio de un *diagrama de dispersión (Scatter Diagram)*.

Es conveniente señalar también aquí que todos estos tipos de análisis pueden combinarse, de hecho a veces combinarlos es lo más adecuado, a medida que queremos profundizar en el análisis de las posibles causas. Por ejemplo un análisis de correlación puede ir seguido o acompañado de análisis de correlación estratificado, si la reflexión sobre las causas así lo aconseja porque la considerada en primer lugar resume una probable variabilidad interna que encubre la verdadera causa (por ejemplo una media diaria puede encubrir diferencias importantes por horas, turnos, médicos, etc.).

Vamos a ver con más detalle en qué consiste la estratificación (útil para la situación B descrita más arriba) y el análisis con diagrama de dispersión (útil para la situación C).

4.1. ANÁLISIS ESTRATIFICADO

La estratificación puede ser definida como un método para identificar la fuente o causa de variación en los criterios o indicadores de calidad *clasificando las mediciones según las categorías o posibilidades (estratos) del factor a indagar*; un factor que pensamos *puede* causar diferencias en el indicador de calidad. De esta forma se describe de forma sistemática el cumplimiento o los valores del criterio o indicador en cada uno de los *estratos*, con lo que obtendremos una comprensión más detallada de la estructura y funcionamiento del proceso analizado. Un *estrato* puede ser definido como un segmento o categoría que se diferenciará, probablemente, del resto de categorías con respecto a la característica de calidad que se está analizando.

Supongamos por ejemplo que nos interesa mantener un balance entre las consultas a demanda y las programadas y elegimos como indicador para medir este balance el porcentaje de consultas a demanda, sobre el total de consultas o visitas del día. Tenemos la sensación de que son pocas las consultas programadas y que actuamos sólo a demanda. Para valorar el nivel de este indicador hemos recogido los datos diarios de las últimas cinco semanas, y nos ha resultado una media de 80,5% y los valores diarios que figuran en la Tabla 10.1.

TABLA 10.1. Porcentaje de consultas a demanda sobre total consultas. Resultados diarios en las últimas cinco semanas

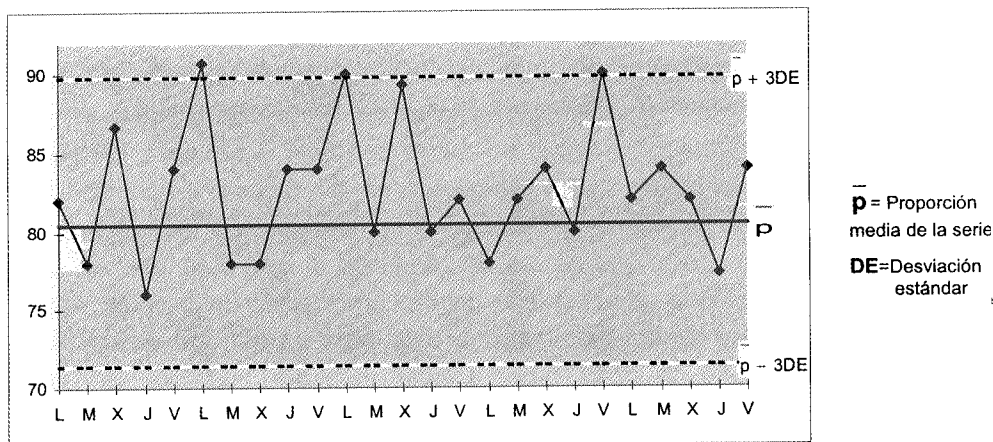
| | | | | | | | | |
|------|-----------|------|-----------|------|-----------|------|-----------|------|
| 82,0 | Lunes | 90,7 | Lunes | 90,0 | Lunes | 78,0 | Lunes | 82,0 |
| 78,0 | Martes | 78,0 | Martes | 80,0 | Martes | 82,0 | Martes | 84,0 |
| 86,7 | Miércoles | 78,0 | Miércoles | 89,3 | Miércoles | 84,0 | Miércoles | 82,0 |
| 76,0 | Jueves | 84,0 | Jueves | 80,0 | Jueves | 80,0 | Jueves | 77,3 |
| 84,0 | Viernes | 84,0 | Viernes | 82,0 | Viernes | 90,0 | Viernes | 84,0 |

$$\bar{p} = 80,5$$

Cuando la causa hipotética es cualitativa el análisis se asocia con el criterio o indicador puede realizarse mediante estratificación; cuando tanto el criterio de calidad como la causa hipotética son cuantitativas se utiliza la correlación y el análisis gráfico mediante diagrama de dispersión.

Una primera forma de analizar los datos sería ver si la media es estable, o existe una gran dispersión, lo que traducirá o bien un patrón rígido, insensible a las diferencias o altibajos que puede tener la demanda día a día o bien la inexistencia de un patrón de conducta establecido. Este análisis puede hacerse con un gráfico de control (UT17) que en nuestro ejemplo nos ha resultado ser como el que se muestra en la Figura 10.5. En él puede verse que la proporción de consulta a demanda es muy variable, con grandes altibajos y diferencias con respecto del valor medio de la serie. Para intentar explicar esta variabilidad podemos pensar en varios *factores o causas hipotéticas* como son el día de la semana, porque sabemos que la demanda varía por días, el propio número diario de consultas a demanda (la hipótesis sería que mantenemos un número adecuado diario de consultas programadas, pero las diferencias se dan por las fluctuaciones de la demanda, que no conocemos bien y no tenemos en cuenta), etc.

FIGURA 10.5. Porcentaje de consultas a demanda sobre total consultas. Evolución de los resultados diarios en las últimas cinco semanas. Gráfico de control



Si nos fijamos primero en el primer factor mencionado, influencia del día de la semana, haríamos una *estratificación* de los datos según día de la semana (con el resultado que puede verse en la Tabla 10.2), y procederemos a comprobar si hay o no diferencia significativa entre ellos (por ejemplo, con una c^2). También podríamos hacer un análisis gráfico, aprovechando que ya tenemos nuestro gráfico de control, si bien su comprensión precisa del conocimiento y manejo de este tipo de gráficos.

TABLA 10.2. Porcentaje de consultas a demanda sobre total consultas. análisis estratificado según días de la semana

| | % de consultas a demanda |
|-----------|--------------------------|
| Lunes | 85,1 |
| Martes | 80,2 |
| Miércoles | 83,0 |
| Jueves | 79,1 |
| Viernes | 84,5 |

$\bar{p} = 80,5$

Igual procedimiento podemos seguir con cualquier otro factor (médicos, tipo de centro, etc.) que nos parezca de interés, siempre que éste sea una variable categórica (de varias opciones). Si el factor o causa de interés es cuantitativa, como sería en nuestro ejemplo en vez del día de la semana, la influencia de la fluctuación en el nº de consultas a demanda, el análisis lo haríamos por medio de correlación, y, a nivel gráfico, con un diagrama de dispersión, tal como veremos más adelante.

Cuál de los factores por los que queremos estratificar va a mostrarse relevante es algo que no se sabe hasta que se realiza el análisis. En los manuales sobre control de calidad se encuentran ejemplos a veces tan llamativos como el análisis provocado por un incremento inusual de los accidentes de tráfico en Japón en 1970; como en los impresos sobre los accidentes había un apartado llamado "causa" se realizó un primer análisis viendo cuál era la más frecuente; ésta resultó ser "conducción imprudente"; sin embargo estratificando cada "causa" por las características de la carretera y la localización específica de los accidentes se evidenciaron frecuencias significativamente más elevadas según determinadas localizaciones y condiciones de las carreteras, las verdaderas causas sobre las que el correspondiente Departamento de Carreteras decidió actuar, mostrándose después una drástica reducción del número de accidentes de tráfico.

La estratificación del nivel de calidad según los diversos factores o causas hipotéticas nos puede indicar cual de ellas se asocia a diferencias significativas en el nivel de calidad. Sin embargo la utilidad de análisis va a depender en primer lugar de una correcta selección o identificación de las causas a analizar

4.2. CORRELACIÓN Y DIAGRAMA DE DISPERSIÓN

La correlación es un análisis estadístico para comprobar la asociación entre dos variables de tipo cuantitativo. El resultado se expresa por medio de un coeficiente (r) llamado *coeficiente de correlación*, que puede ir en valores absolutos de 0 a 1, expresando 1 el grado máximo de correlación, que si es positiva (cuando una variable aumenta, la otra también) se traduce en valores también positivos para el coeficiente de correlación, y si es negativa (cuando una variable aumenta, la otra disminuye) dará valores negativos para el coeficiente de correlación. La *regresión* es un análisis semejante, pero en el que estamos interesados no sólo en la existencia o no de asociación sino también en cuantificar en qué medida la variación de una variable (que llamamos independiente) provoca variación en la otra (que se llama dependiente), para lo cual se construye un modelo o ecuación que represente esa relación de dependencia y que puede llevarse a un gráfico por medio de la línea que resulte al aplicar a la ecuación las parejas de valores que hemos obtenido para las dos variables en estudio; este acercamiento tiene una intención predictiva, (encontrar la fórmula con la cual pueden predecirse los valores de la variable dependiente en función de los que tenga la variable independiente), mientras que en la correlación sólo estamos interesados por la existencia de asociación, en realidad como si ambas variables fuesen variables "dependientes" entre sí.

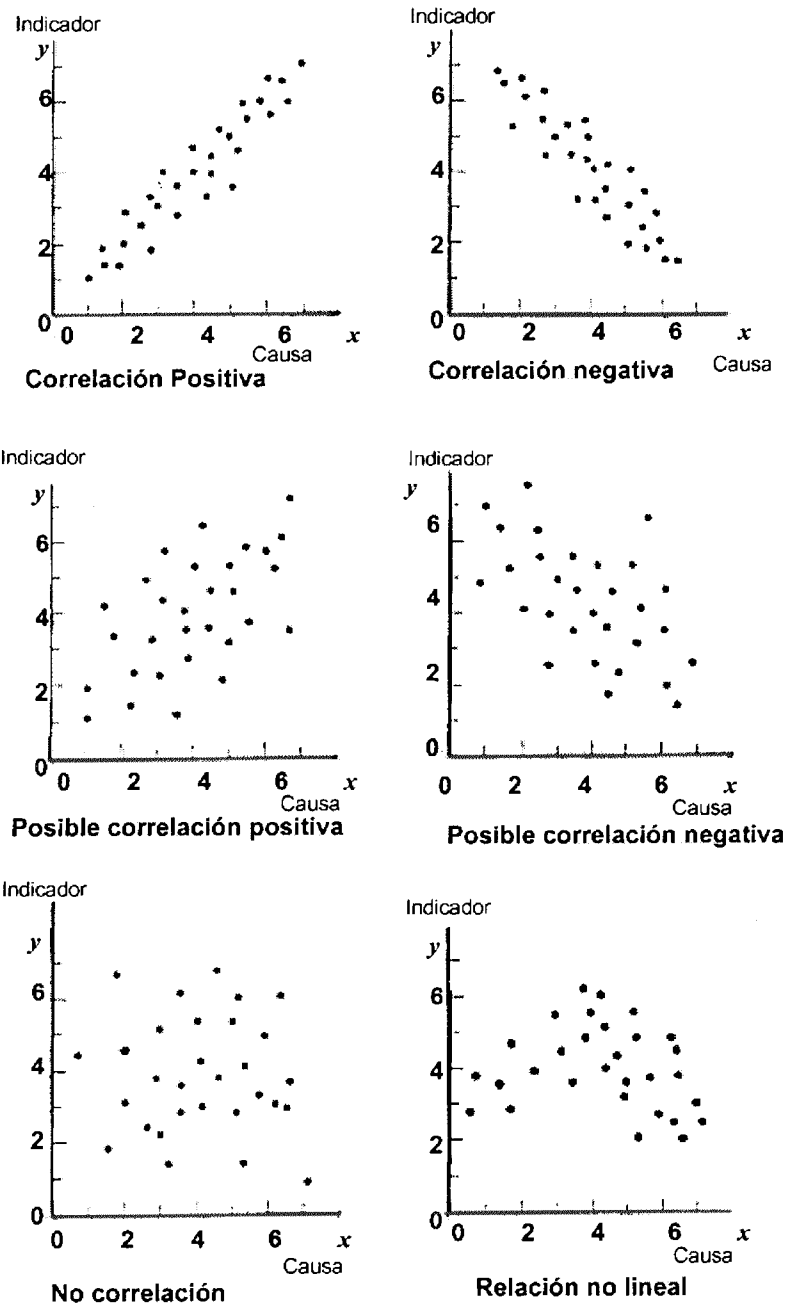
En los análisis que se realizan en los ciclos de mejora, es más frecuente el enfoque de correlación que el de regresión pero en ambos casos (correlación y regresión), el análisis puede plasmarse, e incluso realizarse, de forma gráfica. De hecho es un consejo muy extendido y aceptado que antes de aplicar ningún cálculo estadístico se debe observar la distribución gráfica de las mediciones.

Al plasmar nuestras mediciones en un gráfico (llamado gráfico de correlación o gráfico de dispersión), utilizamos un eje de coordenadas, siendo lo habitual el

que en el eje de ordenadas (y) se represente el criterio o indicador de calidad que estamos analizando como problema, y en el de abscisas (x) la causa o factor influyente potencial; cada pareja de datos (valor del indicador y valor de la causa en estudio) se representa como un punto en el gráfico. Como resultado podemos encontrarnos, en general, con los patrones que se reproducen en la Figura 10.6. Si queremos ser más detallados en el análisis puede calcularse el coeficiente de correlación y su correspondiente significación estadística, o bien realizar una prueba no paramétrica sencilla basada en el simple conteo de los puntos en referencia a su posición en el gráfico (prueba de los signos).

La correlación o asociación entre un criterio de calidad cuantitativo (por ejemplo tiempo de espera) y una causa hipotética también cuantitativa (por ejemplo número de pacientes de consulta a demanda) se puede valorar mediante análisis estadístico (coeficiente de correlación) y también mediante un diagrama de dispersión, que puede adicionalmente analizarse calculando el coeficiente de correlación y su significación estadística, o mediante la prueba de los signos.

FIGURA 10.6. Posibles patrones de los gráficos de dispersión



a) Análisis de la correlación: prueba de los signos

Supongamos que nuestro indicador de calidad es la proporción de pacientes que esperan más de 45' para ser atendidos, y que pensamos que a pesar de nuestros esfuerzos por mantenerlo bajo, la propia variabilidad de la demanda influye en que nuestro indicador varíe, de forma que si hay más demanda, la proporción de pacientes que tienen que esperar más se eleva. Si esto es así, debemos intentar controlar mejor la demanda o establecer mecanismos organizativos que tengan en cuenta sus fluctuaciones para que no influyan en el tiempo de espera de los pacientes. Para estudiar esta relación entre tiempo de espera y porcentaje de consultas a demanda, recogemos los datos de ambas variables cada día durante un cierto tiempo (queremos tener al menos 30 mediciones) con el resultado que se recoge en la Tabla 10.3

TABLA 10.3. Porcentaje de pacientes que esperan más de 45' (y), y porcentaje de consultas a demanda (x)

| | FECHA | y | x |
|-------|-------|------|------|
| MAYO | 1 | 8,6 | 88,9 |
| | 2 | 8,9 | 88,4 |
| | 3 | 8,8 | 87,4 |
| | 4 | 8,8 | 89,1 |
| | 5 | 8,4 | 87,4 |
| | 8 | 8,7 | 88,6 |
| | 9 | 9,2 | 91,1 |
| | 10 | 8,6 | 91,2 |
| | 11 | 9,2 | 89,5 |
| | 12 | 8,7 | 89,6 |
| | 15 | 8,4 | 89,4 |
| | 16 | 8,2 | 86,4 |
| | 17 | 9,2 | 92,2 |
| | 18 | 8,7 | 90,9 |
| | 19 | 9,4 | 90,5 |
| | 22 | 8,7 | 89,2 |
| | 23 | 8,5 | 87,7 |
| | 24 | 9,2 | 88,5 |
| | 25 | 8,5 | 86,6 |
| | 26 | 8,3 | 89,6 |
| 29 | 8,7 | 89,6 | |
| 30 | 9,3 | 92,8 | |
| 31 | 8,9 | 88,6 | |
| JUNIO | 1 | 8,9 | 90,8 |
| | 2 | 8,3 | 88,1 |
| | 5 | 8,7 | 88,2 |
| | 6 | 8,9 | 90,4 |
| | 7 | 8,7 | 91,2 |
| | 8 | 9,1 | 92,5 |
| | 9 | 8,7 | 87,2 |

El correspondiente diagrama de dispersión es la Figura 10.7, en el cual parece que, efectivamente, hay una relación positiva. Pero, ¿cómo podemos saber si es o no significativa? Por una parte podemos calcular el coeficiente de correlación y ver su significación estadística, lo cual nos hará el ordenador con cualquier paquete estadístico u hoja de cálculo de que dispongamos, pero también podemos realizar este análisis visualmente si no disponemos de esos recursos o si, como ocurre con frecuencia, tenemos un número pequeño de mediciones. Para ello seguimos los siguientes pasos:

1. Trazamos una línea vertical a nivel de la mediana de los valores de la causa potencial (representada en el eje de abscisas, x), de forma que deja a cada lado la mitad de los valores para esta variable en nuestra serie. Para nuestro ejemplo la mediana de esta causa potencial (el porcentaje de consultas a demanda) es 89,1.
2. Trazamos una línea horizontal a nivel de la mediana de valores del indicador analizado. En nuestro ejemplo el indicador de calidad es la proporción de pacientes que esperan más de 45', la mediana es 8,7, valor que deja la mitad de valores por encima y la mitad por debajo.
3. Al trazar estas líneas tendremos el gráfico dividido en cuatro cuadrantes (n_1, n_2, n_3, n_4). Procedemos entonces a contar el número de puntos en cada cuadrante, y sumamos los que hay en los cuadrantes contrapuestos (n_1+n_3 y n_2+n_4 , ver Figura 10.7). En estos cálculos no se debe incluir ningún punto que esté justo en cualquiera de las dos medianas.
4. La menor de estas dos sumas se compara con los valores de la tabla que se reproduce como Tabla 10.4, y si este valor es *igual o menor* que el que aparece en la tabla para el correspondiente número total de puntos del gráfico, concluiremos que la correlación es significativa. Adicionalmente, si n_1+n_3 es mayor que n_2+n_4 la correlación es positiva, mientras que si n_1+n_3 es menor que n_2+n_4 la correlación es negativa, aunque esto es visible a simple vista en el gráfico.

En nuestro ejemplo:

$$\begin{array}{cccc}
 n_1=6 & n_2=4 & n_3=8 & n_4=3 \\
 & n_1+n_3=14 & n_2+n_4=7 &
 \end{array}$$

el más pequeño es 7 y buscando en la tabla el valor que corresponde para $n=30$, como es igual que el correspondiente al nivel de significación 0.05 (e incluso 0,01) concluiremos que ambas variables están significativamente correlacionadas.

FIGURA 10.7. Relación entre la proporción de pacientes que esperan más de 45' (y) y las de consultas a demanda (x) (según los datos de la Tabla 10.3).

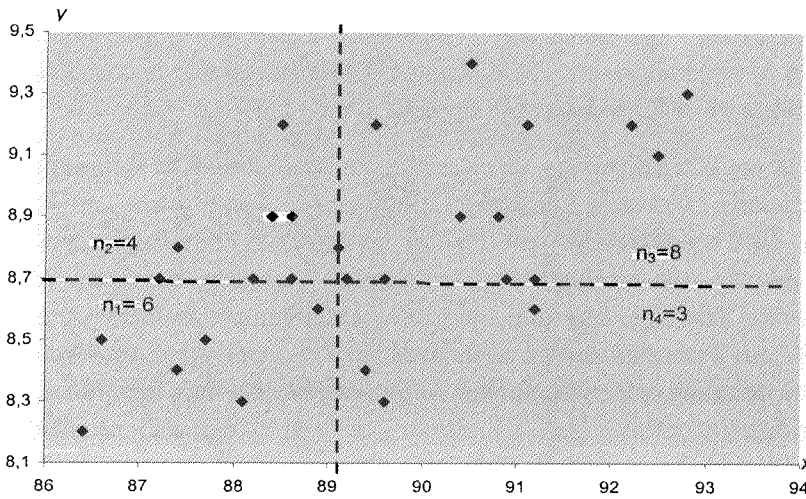


TABLA 10.4. Tabla de la prueba de los signos.

| n | 0,01 | 0,05 | n | 0,01 | 0,05 | n | 0,01 | 0,05 |
|----|------|------|----|------|------|-----|------|------|
| 20 | 3 | 5 | | | | | | |
| 21 | 4 | 5 | 46 | 13 | 15 | 71 | 24 | 26 |
| 22 | 4 | 5 | 47 | 14 | 16 | 72 | 24 | 27 |
| 23 | 4 | 6 | 48 | 14 | 16 | 73 | 25 | 27 |
| 24 | 5 | 6 | 49 | 15 | 17 | 74 | 25 | 28 |
| 25 | 5 | 7 | 50 | 15 | 17 | 75 | 25 | 28 |
| 26 | 6 | 7 | 51 | 15 | 18 | 76 | 26 | 28 |
| 27 | 6 | 7 | 52 | 16 | 18 | 77 | 26 | 29 |
| 28 | 6 | 8 | 53 | 16 | 18 | 78 | 27 | 29 |
| 29 | 7 | 8 | 54 | 17 | 19 | 79 | 27 | 30 |
| 30 | 7 | 9 | 55 | 17 | 19 | 80 | 28 | 30 |
| 31 | 7 | 9 | 56 | 17 | 20 | 81 | 28 | 31 |
| 32 | 8 | 9 | 57 | 18 | 20 | 82 | 28 | 31 |
| 33 | 8 | 10 | 58 | 18 | 21 | 83 | 29 | 32 |
| 34 | 9 | 10 | 59 | 19 | 21 | 84 | 29 | 32 |
| 35 | 9 | 11 | 60 | 19 | 21 | 85 | 30 | 32 |
| 36 | 9 | 11 | 61 | 20 | 22 | 86 | 30 | 33 |
| 37 | 10 | 12 | 62 | 20 | 22 | 87 | 31 | 33 |
| 38 | 10 | 12 | 63 | 20 | 23 | 88 | 31 | 34 |
| 39 | 11 | 12 | 64 | 21 | 23 | 89 | 31 | 34 |
| 40 | 11 | 13 | 65 | 21 | 24 | 90 | 32 | 35 |
| 41 | 11 | 13 | 66 | 22 | 24 | 100 | 36 | 39 |
| 42 | 12 | 14 | 67 | 22 | 25 | | | |
| 43 | 12 | 14 | 68 | 22 | 25 | | | |
| 44 | 13 | 15 | 69 | 23 | 25 | | | |
| 45 | 13 | 15 | 70 | 23 | 26 | | | |

La estratificación y el diagrama de dispersión forman parte de las siete herramientas consideradas por algunos autores como las básicas para el control de la calidad, y de conocimiento indispensable por todo el personal de una organización con programa de gestión de la calidad

Es preciso recordar, sin embargo, que encontrar correlación entre dos variables no quiere decir que la relación sea *causal*. Puede que la relación se establezca a través de terceros factores y, en cualquier caso, la relación de causalidad va a precisar de otro tipo de estudios, entre otras cosas, el resultado que obtengamos en la reevaluación una vez hayamos actuado sobre la causa potencial, nos confirmará o no si la relación era causal.

En la búsqueda de causas de los problemas de calidad, el "viaje diagnóstico" según algunos autores o "recorrido diagnóstico" según otros, no hay recetas mágicas. Es frecuente tener que utilizar sabiamente combinadas varias herramientas de una manera secuencial, razonada y con sentido común. De todas las herramientas posibles, hay una cierta coincidencia en destacar siete como las más útiles, las llamadas por los japoneses Siete Herramientas, en memoria de las (allá) famosas siete armas de Benkei, un guerrero-sacerdote japonés, con las que venció todas sus batallas (en USA algunos autores las han rebautizado como los "7 magníficos"). Estas siete herramientas son las que figuran en la Tabla 10.5; entre ellas están, junto al listado de criterios, tanto la estratificación como los diagramas de dispersión, lo cual da una idea de su relevancia potencial. Estas herramientas son consideradas por Ishikawa como correspondientes al nivel elemental, dentro de los tres niveles de métodos estadísticos para el control de la calidad, definidos en función de su dificultad. Son, además, según Ishikawa, indispensables para los programas de calidad y deben de poder ser utilizadas por todo el personal de la organización.

TABLA 10.5. Las siete herramientas elementales para el control de la calidad

1. DIAGRAMA DE PARETO
2. DIAGRAMA DE CAUSA-EFECTO
3. ESTRATIFICACIÓN
4. HOJA DE COMPROBACIÓN DE PRESENCIA DE EFECTOS (Listado de criterios para su cuantificación)
5. HISTOGRAMA
6. DIAGRAMA DE DISPERSIÓN
7. GRÁFICOS DE CONTROL

Fte.: Adaptado de Ishikawa K.

BIBLIOGRAFÍA

1. Goldfield N, Pine M, Pine J. Techniques for analyzing processes and making team decisions. En: Measuring and managing health care quality. Vol I. Gaithersburg: Aspen Pub; 1995.
2. Ishikawa K. EL enfoque estadístico y algunas herramientas estadísticas sencillas. En Ishikawa: Introducción al control de calidad. Madrid: Díaz de Santos; 1994.
3. Scholtes PR, Joiner BL, Streibel BJ. Getting started: Learning the tools. En: The team handbook. 2ª ed. Madison: Joiner Associates; 1996.
4. Rosner B. Fundamentals of Bioestistics. 4ª ed. Belmont: Duxbury Press; 1995.

ANÁLISIS Y PRESENTACIÓN DE LOS DATOS DE UNA EVALUACIÓN

EMCA

Gestión de la Calidad Asistencial

CONTENIDO GENERAL

Esta UT detalla la forma en que los datos obtenidos en el estudio del nivel de calidad deben ser presentados y analizados desde un punto de vista estadístico, y para que sirvan como base para diseñar la intervención que haga mejorar la calidad. Se subraya la necesidad de calcular los intervalos de confianza de los valores encontrados, si se ha utilizado una muestra, y se especifica paso a paso la construcción de un diagrama de Pareto para el análisis de las deficiencias de calidad a subsanar. Se muestra asimismo la construcción de otros tipos de gráficos (estrella, Box and whiskers) que pueden ser utilizados para la presentación global de los datos.

ÍNDICE DE CONTENIDOS

1. Introducción
2. Los datos de la evaluación: estimación del nivel de cumplimiento de los criterios e intervalos de confianza.
3. Influencia del método de muestreo en la estimación de intervalos de confianza.
4. Presentación gráfica de los datos de una evaluación.
5. Análisis gráfico de los datos para priorizar intervenciones: Diagrama de Pareto
6. Análisis de los datos de estudios de causas hipotéticas.

OBJETIVOS ESPECÍFICOS

1. Resumir de manera informativa los datos de una evaluación.
2. Calcular los intervalos de confianza de las estimaciones de cumplimiento encontradas.
3. Distinguir los conceptos de "intervalo" y "confianza" e interpretar sus valores.
4. Ordenar los datos de forma adecuada para construir un diagrama de Pareto sobre las deficiencias de calidad encontradas.
5. Hacer una tabla con los porcentajes de defectos y sus porcentajes acumulados según criterios, para su utilización en la construcción del diagrama de Pareto.
6. Construir un diagrama de Pareto completo, con los resultados de la evaluación.
7. Explicar los resultados de la evaluación utilizando el diagrama de Pareto.
8. Presentar gráficamente los resultados de la evaluación utilizando un gráfico tipo estrella o radar.
9. Presentar gráficamente los resultados de la evaluación utilizando un gráfico "box and whiskers".
10. Describir correctamente las presentaciones tabular y gráfica de los resultados de una evaluación.

2. LOS DATOS DE LA EVALUACIÓN: ESTIMACIÓN DEL NIVEL DE CUMPLIMIENTO DE LOS CRITERIOS E INTERVALOS DE CONFIANZA

La mayoría de las evaluaciones del nivel de calidad se realizan utilizando muestras. En estas condiciones, lo que realizamos no son mediciones directas y sin error posible sino *estimaciones* de los valores reales que tiene la variable (criterio de calidad) que queremos evaluar; estos valores reales, sean proporciones o medias, se conocen como *parámetros* en la jerga estadística. Dicho de una forma más precisa lo que hacemos (o debemos hacer) cuando medimos el nivel de calidad con una muestra es realizar una *inferencia estadística* sobre las características (parámetros) de poblaciones o grupos amplios de casos, sobre la base de información aportada por la muestra; siempre que, recordémoslo, la muestra sea aleatoria (representativa), no sesgada, tal como veíamos en la UT 9 al hablar del diseño del estudio. Las estimaciones a inferir sobre el parámetro que nos interese pueden ser de tipo puntual (un solo valor) o de intervalos de valores.

2.1. ESTIMACIÓN PUNTUAL DEL NIVEL DE CUMPLIMIENTO DE UN CRITERIO

La primera *inferencia* que podemos hacer con los datos de la muestra es la propia estimación del nivel de cumplimiento que resulta en la medición realizada para cada uno de los criterios. Por ejemplo, si hemos evaluado la proporción de hipertensos correctamente diagnosticados en una muestra aleatoria de 60 casos y hemos encontrado que 45 de estos 60 cumplen el criterio de diagnóstico correcto, podemos *inferir* o *estimar* que una proporción de $45/60 = 0,75$ o, en porcentaje, un 75% de los hipertensos están correctamente diagnosticados. Esta estimación se conoce como *estimación puntual* del parámetro (valor poblacional real) que queremos medir, en este caso la proporción de hipertensos correctamente diagnosticados.

Sin embargo, todos sabemos que al medir algo por medio de una muestra el resultado está sujeto a errores propios del muestreo, y que, por tanto, la estimación puntual puede no ser exactamente igual al valor real que encontraríamos si midiésemos no una muestra sino *todos* los casos. Es más, no sólo somos conscientes de que la estimación puntual puede no ser igual al valor real, sino que también está claro que no sabemos, dando sólo la estimación puntual, lo cerca o lo lejos que puede encontrarse del valor real. Para solucionar este problema hemos de completar nuestro proceder de inferencia estadística estimando no sólo un valor (estimación puntual) sino un *intervalo de valores*, dentro de los cuales podamos tener una cierta seguridad de que se encuentra el valor real, poblacional, de lo que hemos medido a través de una muestra.

2.2. ESTIMACIÓN DEL INTERVALO DE CONFIANZA DEL CUMPLIMIENTO DE UN CRITERIO

Esta segunda estimación es lo que se conoce como *intervalo de confianza* de la medición realizada. El *intervalo de confianza* expresa la *precisión* o exactitud probable (¡todo en estadística es probabilidad!) de la medición realizada, en

Los datos de una muestra representativa nos sirven para estimar por medio de inferencia estadística el parámetro que medimos.

La primera inferencia a realizar es la estimación puntual del parámetro (por ejemplo, proporción de cumplimiento de un criterio) que hemos medido. Para ello hacemos los cálculos con los datos de la muestra como si se tratase de la totalidad del universo de donde se ha extraído la muestra.

relación con el valor real; obviamente, cuanto más amplio sea el intervalo, menos precisa es nuestra estimación. En la amplitud del intervalo intervienen, como luego veremos, el tamaño de la muestra, la variabilidad de los datos de la muestra, y la confianza (certeza o probabilidad) que queramos tener de que sea cierto que este intervalo contiene el valor real.

No debe de extrañarnos la afirmación de que las estimaciones de cumplimiento de criterios en base a muestras deberían incluir *siempre* la estimación del intervalo de confianza. Pero, ¿cómo podemos estimar este intervalo de valores?; es fácil: sólo necesitamos los datos de nuestro muestreo y decidir sobre la *confianza* que queremos tener en nuestra estimación. De hecho, conviene entender que, aunque relacionados, el concepto *intervalo de confianza* entraña dos aspectos diferentes: el "intervalo" y la "confianza". El "intervalo" es la gama de valores dentro de la cual estimamos que está el valor real (poblacional) del parámetro evaluado (en nuestro caso este parámetro sería la proporción de cumplimientos de un criterio), mientras que la "confianza" es la probabilidad de que sea *cierto* que el intervalo contiene el valor real del parámetro evaluado. Sin embargo, dado que ambos aspectos van necesariamente unidos en la inferencia estadística (para una misma muestra si aumentamos la confianza de la estimación, aumenta la amplitud del intervalo), es costumbre considerarlos como un concepto único. Vamos a ver como se calcula.

a) Cálculo del intervalo de confianza del nivel de cumplimiento de un criterio

El razonamiento y fórmula general para estimar el intervalo de confianza del parámetro medido con muestra es: *estimación puntual* $\pm i$, siendo

$$i = z \cdot \text{Desviación estándar del parámetro estimado}$$

donde "z" es un valor de la distribución normal (la distribución de probabilidades más conocida en estadística) relacionado con la *confianza* que queramos tener en la estimación (probabilidad de que el intervalo contenga el valor real del parámetro estimado); y la "desviación estándar del parámetro estimado" es lo que se conoce también como "error del muestreo" o "error estándar", que se estima en base a la desviación estándar de los datos de la muestra (una expresión de su dispersión o variabilidad) y el propio tamaño de la muestra. Así pues, el cálculo del intervalo de confianza entraña los siguientes dos pasos previos: (1) Determinar el valor de z; y (2) Calcular el error estándar.

- **Determinar el valor de "z".** Para ello hay que decidir sobre la "confianza" que queremos tener para la estimación; es decir lo seguros que queramos estar de que el valor real esté contenido en el intervalo que vamos a estimar. Está claro que una "confianza" del 100% exigiría medir todos los casos, no una muestra, o bien dar como intervalo toda la gama posible de valores, lo cual no tendría ninguna utilidad como estimación. En consecuencia, hay que "arriesgarse" algo; habrá que aceptar una cierta probabilidad de equivocarse; es decir estimar un intervalo de valores con una confianza inferior al 100%. ¿Cuál debe ser esta confianza? En general, se considera una "confianza" aceptable si ésta es del 95%. Es decir, si tenemos un 95% de probabilidades de que el intervalo contenga el valor real, y sólo un 5% de probabilidades de que el valor real sea un valor externo al intervalo (probabilidad de equivocarnos en nuestra estimación).

Para una confianza de 95%, z es 1,96. Este valor es el que aparece en las

La estimación puntual puede y debe completarse con la estimación del intervalo de confianza; un intervalo de valores dentro del cual tengamos un cierto nivel de seguridad (confianza) de que se encuentre el valor real.

El intervalo de confianza implica dos aspectos que, aunque en la práctica han de ir necesariamente juntos, conviene entender y considerar por separado; el intervalo (conjunto de valores entre los cuales sea cierto que estimamos que está el valor real) y la confianza (probabilidad de que sea cierto que el intervalo contenga al valor real).

El intervalo de confianza (estimación puntual $\pm i$) se calcula fácilmente a partir de los datos de la muestra, una vez fijada la confianza y estimando el error llamado error estándar ($i = z \cdot \text{error estándar}$, siendo z un valor dependiente de la confianza).

tablas de la distribución normal, para esta confianza. La Tabla 11.1 contiene los valores de z , correspondientes a los diversos y más frecuentemente empleados niveles de confianza.

TABLA 11.1. Valores de "z" para diversos niveles de confianza de los intervalos estimados

| NIVEL DE CONFIANZA | Z |
|--------------------|------|
| 90% | 1,64 |
| 95% | 1,96 |
| 99% | 2,58 |

- **Calcular el error estándar o desviación estándar del parámetro estimado.** Anotado habitualmente como s_p (cuando el parámetro estimado es una proporción, como ocurre en la mayoría de los criterios de calidad), o como $S_{\bar{x}}$ (cuando el parámetro estimado es una media). En el caso de la estimación del nivel de calidad como proporción de cumplimiento de un criterio la fórmula de cálculo es:

$$S_p = \sqrt{\frac{p(1-p)}{n}}$$

dónde p es la *estimación puntual* de la proporción de cumplimiento, y n es el tamaño de la muestra utilizada.

En el ejemplo empleado de la estimación de la proporción de hipertensos correctamente diagnosticados, teníamos una estimación puntual de 75% de hipertensos correctamente diagnosticados, en una muestra de 60; la estimación del *intervalo de confianza* para el cumplimiento de este criterio, sería, con una confianza de 95%:

$$z = 1,96 \text{ (como corresponde a una confianza de 95\%)}$$

$$p = 0,75 \qquad 1-p = 0,25 \qquad n = 60$$

$$S_p = \sqrt{\frac{(0,75)(0,25)}{60}}$$

$$\text{con lo cual } i = z \cdot s_p = 1,96 \sqrt{\frac{(0,75)(0,25)}{60}} = 0,11$$

y el intervalo de confianza: *estimación puntual* $\pm i = 0,75 \pm 0,11$

¿Qué quiere decir esto?, pues que estamos un 95% seguros de que el valor real de cumplimiento del criterio es un valor que está entre 0,64 (límite inferior del intervalo: $0,75-0,11$) y 0,86 (límite superior del intervalo: $0,75+0,11$); en porcentaje: entre 64% y 86%.

Esta información es mucho más útil y correcta que dar solamente la estimación puntual. Merece la pena subrayar, en cambio, la sencillez de su cálculo: sólo se precisa una calculadora de mano.

Por otra parte, sólo con ver la fórmula para su cálculo (n está en el denominador de s_p) es evidente que cuanto mayor sea el tamaño de la muestra, más pe-

Es habitual estimar el intervalo de confianza con una confianza de 95%, con lo cual en la fórmula para su cálculo $z=1,96$.

Una vez calculada la desviación estándar del parámetro estimado (error estándar), se puede estimar el intervalo de confianza. Este intervalo nos indica los valores mínimo y máximo entre los cuales se encuentra el valor real del parámetro estimado, con una confianza (seguridad de que sea cierto) prefijada que habitualmente decidimos que sea del 95%.

queño va a resultar el intervalo; es decir: más precisa puede ser nuestra estimación. De esta forma, si al estimar el nivel de cumplimiento de un determinado criterio el intervalo de confianza nos resulta demasiado amplio para decidir sobre su mejora, podemos obtener una mayor precisión ampliando la muestra; siempre que los casos adicionales los obtengamos también de forma aleatoria y del mismo marco muestral. Sin embargo, la influencia del tamaño del universo o marco muestral es, la mayoría de las veces, irrelevante, tal como veremos a continuación.

2.3. INFLUENCIA DEL TAMAÑO DEL UNIVERSO O MARCO MUESTRAL

¿Qué ocurre si extraemos la muestra de marcos muestrales (total de casos, universo o población) de tamaños diferentes?; ¿influye el tamaño del marco muestral en las estimaciones que podamos inferir? La respuesta es *no siempre*.

En principio el tamaño del universo muestral no influye a efectos prácticos a menos de que la muestra sea de un tamaño superior al 10% del total del universo o marco muestral de donde se ha extraído. En este caso, si bien no hay influencia en cuanto a la estimación puntual, sí que conviene recalcular la estimación del intervalo de confianza porque, manteniendo la misma confianza, el intervalo va a resultar más pequeño: es decir la estimación va a ser más precisa.

El cálculo del intervalo ajustado por el tamaño del universo se realiza multiplicando por la fracción:

$$\sqrt{\frac{N-n}{N-1}}$$

donde N es el tamaño del universo y n es el tamaño de la muestra.

La fórmula es entonces: $i' = i \cdot \sqrt{\frac{N-n}{N-1}}$

donde i es el cálculo realizado sin tener en cuenta el tamaño del universo, tal como hemos visto antes. En nuestro ejemplo de estimación de la proporción de hipertensos correctamente diagnosticados, realizada en una muestra de 60 casos, realizaremos este ajuste para la estimación del intervalo de confianza siempre que el tamaño del universo (total de hipertensos) sea inferior a 600, puesto que en esos casos el tamaño de la muestra, 60, va a ser mayor que el 10% del tamaño del universo.

Supongamos, por ejemplo, que el total de hipertensos diagnosticados en nuestro centro (N , marco muestral o universo para nuestro estudio) es de 300; el intervalo de confianza, ajustado por el hecho de que $n > 10\% N$, sería

$$i' = i \cdot \sqrt{\frac{300-60}{299}} = 0,10,$$

lo cual quiere decir que podemos afirmar, con una confianza del 95%, que la proporción real de cumplimiento es un valor comprendido entre 65% y 85% ($0,75 \pm 0,10$), en lugar de 64% y 86% como teníamos antes.

Obsérvese, que para $n=N$ (todos los casos), la fracción $\sqrt{\frac{N-n}{N-1}}$

es cero, lo cual es lógico porque no hablaríamos entonces de inferencia esta-

Para una misma confianza, si aumentamos el tamaño de la muestra, más precisión obtenemos al estimar el intervalo; es decir más pequeña es la gama de valores entre los cuales estimamos que está el valor real.

El tamaño del universo (N) de donde se extrae la muestra (n) no influye en la estimación del intervalo de confianza a menos que $n > 10\% N$.

dística sino de medición directa del parámetro; de igual manera, a medida que la diferencia entre N y n se agranda, la fracción de ajuste tiende a ser 1, lo cual indica que el tamaño del universo no va a influir en la estimación del intervalo de confianza. De hecho, como regla, no merece la pena hacer el cálculo ajustado para cualquier $n < 10\% N$.

2.4. INFERENCIA ESTADÍSTICA SOBRE MEDIAS

La mayoría de los criterios para medir nivel de calidad son cualitativos, categóricos: se mide en cada caso la cualidad de si lo cumple o no, clasificamos los casos en estas dos categorías (las que cumplen y las que no), y estimamos la proporción de una de ellas, la que más nos interese. Sin embargo, en ocasiones el nivel de calidad se mide con criterios que son variables cuantitativas cuyo parámetro característico es una media (por ejemplo media de minutos por consulta, media de minutos que tiene que esperar el paciente antes de ser atendido, etc.). En estos casos, aunque el razonamiento estadístico es igual, las fórmulas para la inferencia de las estimaciones tanto puntuales como de intervalo son diferentes. La Tabla 11.2 contiene las fórmulas necesarias, tanto para el caso de proporciones (como en nuestro ejemplo de hipertensos bien diagnosticados), como de medias.

Si el criterio mide una cantidad cuyo parámetro resumen es una media, la fórmula para la estimación puntual y del intervalo de confianza son peculiares, aunque el razonamiento estadístico es idéntico al visto para proporciones.

TABLA 11.2. Inferencia estadística en base a muestras no sesgadas: estimaciones puntuales e intervalos de confianza de proporciones y medias

| PARÁMETRO A ESTIMAR | TAMAÑO DE MUESTRA (n) | ESTIMACIÓN PUNTUAL | INTERVALO DE CONFIANZA (95%) (estimación puntual $\pm \hat{\delta}$ *) |
|---|--|---|---|
| PROPORCIÓN (p) ejemplo: proporción de cumplimientos de un criterio de calidad. | Menor que el 10% del tamaño del marco muestral (N) | $p = \frac{\text{n}^\circ \text{c u m p l i m i e n t o s}}{n}$ | $p \pm 1,96 \sqrt{\frac{p(1-p)}{n}}$ |
| | Mayor que el 10% del tamaño del marco muestral (N) | $p = \frac{\text{n}^\circ \text{c u m p l i m i e n t o s}}{n}$ | $p \pm \left(1,96 \sqrt{\frac{p(1-p)}{n}} \cdot \sqrt{\frac{N-n}{N-1}} \right)$ |
| MEDIA (\bar{x}) (ejemplo: media de minutos por consulta) | Menor que el 10% del tamaño del marco muestral (N) | $\bar{x} = \frac{\sum x_i}{n}$ | $\bar{x} \pm 1,96 \frac{s}{\sqrt{n}}$ |
| | Mayor que el 10% del tamaño del marco muestral (N) | $\bar{x} = \frac{\sum x_i}{n}$ | $\bar{x} \pm \left(1,96 \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} \right)$ |

x_i : los diversos valores individuales de cada caso incluido en la muestra

s : desviación estándar de los datos de la muestra = $\sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$

*: la fórmula general para calcular los intervalos de confianza es $i=z$. Desviación estándar del parámetro estimado, donde z es el valor en la distribución normal que corresponde a los diversos niveles de confianza (ver Tabla 11.1), y la desviación estándar del parámetro corresponde al llamado *error estándar*. Si el tamaño de la muestra no es suficientemente grande ($n > 30$), la distribución normal no puede asumirse y se debe emplear el valor de t (de la distribución de la t de student) en lugar de z ; t varía con el tamaño de la muestra.

3. INFLUENCIA DEL MÉTODO DE MUESTREO EN LA ESTIMACIÓN DE INTERVALOS DE CONFIANZA

En la UT 9 vimos que el muestreo puede realizarse con diversos métodos, de los que destacamos por su mayor aplicabilidad práctica en general el aleatorio simple, el sistemático y el estratificado (proporcional y no proporcional). El emplear uno u otro método tiene influencia en la variabilidad de los datos de la muestra, y por tanto en las estimaciones que hagamos con ellos, sobre todo en la estimación del intervalo de confianza. Esta influencia puede ser particularmente importante en el muestreo estratificado, y se basa en el hecho de que al estratificar controlamos uno de los factores (aquel por el que estratificamos) con influencia potencial en la variabilidad de los datos (y por tanto en la desviación estándar); la variabilidad *dentro* de cada estrato es de esperar que sea menor (es más homogéneo) que la que exista entre los estratos y por tanto en la muestra que podríamos realizar sin estratificar; por otra parte, las diferencias de tamaño de cada estrato también han de ser tenidas en cuenta, sobre todo en el muestreo estratificado no proporcional. El resultado es generalmente que en las muestras estratificadas la estimación es más precisa si aplicamos las fórmulas adecuadas que si no tenemos en cuenta la estratificación.

La Tabla 11.3 contiene las fórmulas para las estimaciones puntuales e intervalos de confianza según los diversos métodos de muestreo. La primera impresión al ver las fórmulas para las estimaciones en muestreos estratificados es que son de gran complejidad; en realidad no es así. Son fácilmente programables en hoja de cálculo, e incluso manualmente si preparamos una tabla debidamente ordenada para el cálculo de cada elemento necesario por separado, con una secuencia lógica. Veamos por ejemplo, como se procedería para el caso de un muestreo estratificado proporcional, probablemente uno de los más frecuentemente empleados al evaluar calidad.

Supongamos que hemos estratificado una muestra de 60 casos entre tres profesionales participantes en un estudio del nivel de calidad. Entre los tres tienen en programa un total (N) de 450 hipertensos.

Uno de ellos tiene 90 hipertensos diagnosticados, otro 160 y el tercero 200. Para estratificar de forma proporcional, asignamos a cada profesional un número de casos para la muestra en proporción a su número total de hipertensos. Así tendremos:

$N = 450$, de los cuales

- 90 (20%) son del profesional 1 (estrato 1)
- 160 (36%) son del profesional 2 (estrato 2)
- 200 (44%) son del profesional 3 (estrato 3)

$n = 60$, de los cuales extraemos

- el 20% de entre los hipertensos del profesional 1 ($n_1=12$)
- el 36% de entre los hipertensos del profesional 2 ($n_2=22$)
- el 44% de entre los hipertensos del profesional 3 ($n_3=26$)

Al evaluar las muestras de cada estrato, hemos encontrado 4 hipertensos bien diagnosticados en la muestra 1, 17 en la 2, y 24 en la 3; un total de 45 sobre el total de 60 casos evaluados. Con ello estimaríamos en un 75% ($45/60 \times 100$) el porcentaje de hipertensos bien diagnosticados. Al ser la estratificación propor-

El muestreo estratificado puede proporcionar estimaciones más precisas. Para ello hay que emplear las fórmulas de cálculo apropiadas.

cional, la estimación puntual se hace de igual forma que si fuera una muestra aleatoria o sistemática simple.

TABLA 11.3. Estimaciones puntuales de proporciones (p) y medias (\bar{x}) intervalos de confianza según tipo de muestreo utilizado

| TIPO DE MUESTREO | ESTIMACIÓN PUNTUAL | INTERVALO DE CONFIANZA (95%) |
|--------------------|---|---|
| - Aleatorio simple | Ver Tabla 11.2 | Ver tabla 11.2 |
| - Sistemático | Ver Tabla 11.2 | Ver tabla 11.2* |
| - Estratificado | | |
| - Proporcional | Ver Tabla 11.2 | $p \pm 1,96 \sqrt{\frac{1-f}{n} \sum W_h \frac{n_h}{n_h - 1} P_h (1 - P_h)}$ $\bar{x} \pm 1,96 \sqrt{\frac{1-f}{n^2} \cdot \sum n_h \cdot s_h^2}$ |
| - No proporcional | $p = \frac{1}{N} \sum N_h \cdot P_h$ $\bar{x} = \frac{1}{N} \sum N_h \cdot \bar{x}_h$ | $p \pm 1,96 \sqrt{\sum W_h^2 (1 - f_h) \frac{P_h (1 - P_h)}{n_h - 1}}$ $\bar{x} \pm 1,96 \sqrt{\sum W_h^2 (1 - f_h) \frac{s_h^2}{n_h}}$ |

f: fracción de muestreo = $\frac{n}{N}$; n = tamaño total de muestra; N = tamaño del universo o marco muestral.

h: número de estratos; n_h = tamaño de la muestra en cada estrato; N_h = tamaño de cada estrato;

\sum : sumatorio de todos los estratos.

P_h : proporción de cumplimientos de cada estrato; f_h = fracción de muestreo en cada estrato.

W_h : proporción de casos de cada estrato (N_h) en el total del universo muestral = $\frac{N_h}{N}$.

s_h : desviación estándar en cada estrato; \bar{x}_h = media en cada estrato.

*: Algunos autores proponen y razonan otras fórmulas, semejantes a las del muestreo estratificado.

Para estimar el intervalo de confianza, podemos proceder sin tener en cuenta que la muestra ha sido estratificada y encontraríamos, tal como vimos más arriba, un intervalo, para una confianza de 95%, de $0,75 \pm 0,11$, ó $0,75 \pm 0,10$ si ajustamos por el tamaño del universo dado que $60 > 10\% 450$:

$$0,11 \cdot \sqrt{\frac{450-60}{449}}$$

Sin embargo, si tenemos en cuenta la estratificación haríamos los cálculos según la fórmula de la Tabla 11.3 y en el orden que se muestra a continuación.

| ESTRATO | Proporción del estrato en el total (W_h) | Muestra en el estrato (n_h) | Proporción de cumplimiento en el estrato (P_h) | $1 - P_h$ | $W_h = \frac{n_h}{n} \cdot P_h (1 - P_h)$ |
|---------|--|---------------------------------|--|-----------|---|
| 1 | 0,20 | 12 | 4/12=0,33 | 0,63 | 0,045 |
| 2 | 0,36 | 22 | 17/22=0,77 | 0,23 | 0,067 |
| 3 | 0,44 | 26 | 24/26=0,92 | 0,08 | 0,034 |
| | | 60 | | | 0,146 |

$$f = \frac{n}{N} = \frac{60}{450} = 0,13$$

Una vez hecho esto, podemos aplicar la fórmula indicada en la Tabla 11.3 para las muestras estratificadas proporcionales:

$$S_{p,estr.} = \sqrt{\frac{1-f}{n} \sum W_h \frac{n_h}{1-n_h} P_h (1-P_h)} = \sqrt{\frac{1-0,13}{60} \cdot 0,146} = 0,046$$

Y con ello el intervalo, para una confianza del 95% sería: $0,75 \pm (1,96 \cdot 0,046) = 0,75 \pm 0,09$.

Una estimación más precisa que la de $0,75 \pm 0,11$ ó $0,75 \pm 0,10$ que habríamos estimado sin tener en cuenta la estratificación.

En general, el incremento de precisión que comporta la estratificación es tanto mayor cuanto mayor sea la desigualdad entre los estratos, de forma que si las estimaciones puntuales por cada estrato son parecidas, no merece la pena realizar los cálculos del intervalo de confianza teniendo en cuenta la estratificación. Sin embargo, si la estratificación que hemos realizado no es proporcional a los tamaños de los estratos, y encontramos proporciones de cumplimiento diferentes en los diferentes estratos, no es sólo conveniente sino *obligatorio* tenerla en cuenta al menos para la estimación puntual global. Veámoslo con un ejemplo: supongamos que la muestra de 60 casos para los mismos 3 profesionales del ejemplo anterior decidimos estratificarla extrayendo aleatoriamente 20 casos de entre los hipertensos de cada uno (estratificación no proporcional). En los 20 casos del profesional 1, encontramos 12 correctamente diagnosticados (proporción de 0,6), 15 en los del profesional 2 (proporción de 0,75), y 18 en los del profesional 3 (proporción de 0,9); un total de 45 cumplimientos pero con proporciones diferentes en los estratos. Para hacer nuestra estimación del nivel de cumplimiento de los 3 profesionales conjuntamente, no sería correcto decir 45/60 sino que utilizaríamos la fórmula indicada en la Tabla 11.3 para muestras estratificadas no proporcionales $p = 1/N \sum N_h \cdot P_h$, cuyos valores en nuestro ejemplo serán los siguientes:

| Tamaño de cada estrato (N_h) | Proporción de cumplimiento en cada estrato (P_h) | $N_h \cdot P_h$ |
|----------------------------------|--|-----------------|
| Estrato 1: 90 | 12/20=0,6 | 54 |
| Estrato 2: 160 | 15/20=0,75 | 120 |
| Estrato 3: 200 | 18/20=0,90 | 180 |
| $N=450$ | | 354 |

El cálculo del intervalo de confianza para muestras estratificadas no es más difícil que para los otros tipos de muestreo, aunque pueda llevar más tiempo se realiza con calculadora de mano. No obstante puede obviarse si las estimaciones puntuales en cada estrato son semejantes.

Si se ha realizado un muestreo estratificado no proporcional y las estimaciones puntuales (proporciones de cumplimiento) en cada estrato son diferentes, la estimación puntual global debe realizarse con la fórmula apropiada, teniendo en cuenta los datos de la estratificación.

De modo que: $P = \frac{1}{450} \cdot 354 = 0,787$; es decir, 78,7%.

Recuérdese que sin tener en cuenta que la muestra ha sido estratificada no proporcional, la estimación de cumplimiento hubiera sido $45/60 = 0,75$, ó 75%.

Hay otros tipos de muestreo (por ejemplo el muestreo en conglomerados y polietápico) que conllevan sus propias fórmulas para una correcta inferencia estadística. En la práctica de los programas de calidad lo normal es que utilicemos en la mayoría de los casos uno u otro de los métodos de muestreo que hemos revisado, pero debe quedar clara la idea de que, en cualquier caso, es conveniente reflexionar para la interpretación de los datos si el tipo concreto de muestreo que hemos utilizado puede o no influir en la forma de realizar los cálculos para la inferencia estadística.

4. PRESENTACIÓN GRÁFICA DE LOS DATOS DE UNA EVALUACIÓN

Una correcta estimación del nivel de cumplimiento de los criterios es fundamental para dar base real al conocimiento sobre el nivel de calidad de los aspectos evaluados y evidenciar en qué criterios estamos necesitados de mejora. Sin embargo, en calidad existe siempre una representación gráfica que visualiza y hace más comprensible el resultado de las mediciones.

De todas las representaciones gráficas posibles vamos a seleccionar las más útiles y prácticas, de entre las cuales elegiremos la que más nos interese, en función de lo que queramos enfatizar en relación a la evaluación realizada. La selección de uno u otro gráfico está representada en la Figura 11.1 en forma de algoritmo decisional.

Como puede verse una vez construida la tabla de presentación de los datos, las decisiones que afectan a la elección de gráfico son: (i) el tipo de variable con que medimos el criterio (categórica, o cuantitativa); (ii) si nos interesa representar sólo estimaciones puntuales o también intervalos; y (iii) si queremos representar cumplimientos o analizar incumplimientos (defectos de calidad). En relación a esta última decisión las opciones que tenemos para representar cumplimientos son el clásico gráfico de barras o el gráfico tipo estrella o radar (si queremos representar sólo las estimaciones puntuales), y el gráfico "box and whiskers" (si queremos representar los intervalos de confianza); para analizar los incumplimientos, objeto de la intervención para mejorar, aunque es posible utilizar estos mismos gráficos, es más ilustrativo y adecuado para evidenciar las prioridades para intervenir el diagrama de Pareto.

Vamos a ver en primer lugar los gráficos para representar cumplimientos y nos detendremos después en revisar paso a paso la construcción e interpretación de un gráfico de Pareto sobre defectos o incumplimientos. En todos los casos, utilizaremos los datos que figuran en la Tabla 11.4, resultado de una evaluación de los criterios de diagnóstico, anamnesis y exploración física inicial del paciente hipertenso.

El análisis de los datos de una evaluación se completa con su representación gráfica. Vamos a elegir el tipo de gráfico más adecuado en función del tipo de variable (cumple/no cumple o cantidad de algo) con la que medimos el criterio, el interés que tengamos en representar las estimaciones de los intervalos de confianza, y si queremos representar niveles de cumplimiento o analizar los incumplimientos.

FIGURA 11.1. Algoritmo decisonal para el análisis y presentación de datos en un estudio del nivel de calidad.

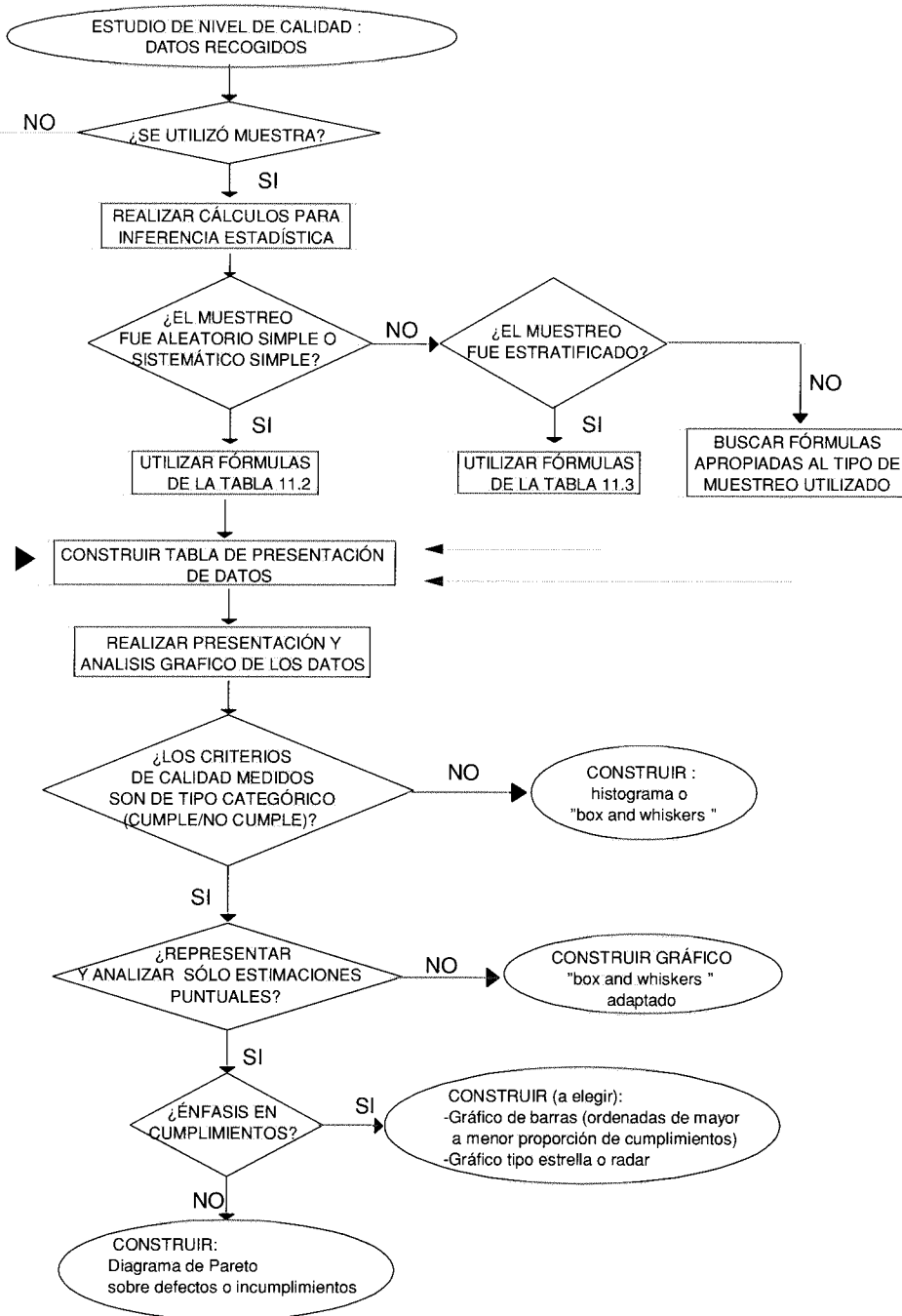


TABLA 11.4. Evaluación de la calidad del diagnóstico, anamnesis y exploración física inicial del paciente hipertenso. niveles de cumplimiento de los criterios en una muestra aleatoria de 60 casos

| CRITERIO* | PORCENTAJE DE CUMPLIMIENTO (Intervalo de confianza del 95%) | |
|----------------------------|---|---------|
| 1. Diagnóstico correcto | 75,0 | (±11,0) |
| 2. Antecedentes familiares | 80,0 | (±10,1) |
| 3. Antecedentes personales | 90,0 | (±7,6) |
| 4. Consumo de tabaco | 30,0 | (±11,6) |
| 5. Consumo de alcohol | 10,0 | (±7,6) |
| 6. Peso y talla y/o IMC | 85,0 | (±9,0) |
| 7. Auscultación cardiaca | 75,0 | (±11,0) |
| 8. Exploración abdominal | 40,0 | (±12,4) |
| 9. Pulsos periféricos | 50,0 | (±12,7) |
| 10. Soplos carotídeos | 55,0 | (±12,6) |
| 11. Edemas | 25,0 | (±11,0) |
| 12. Fondo de ojo | 15,0 | (±9,0) |

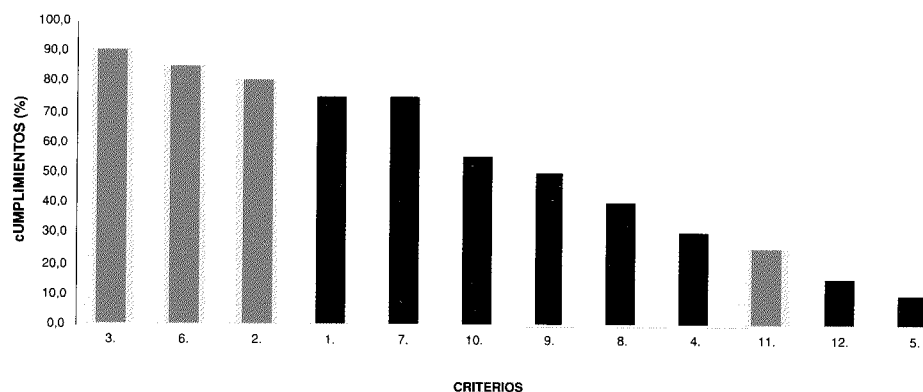
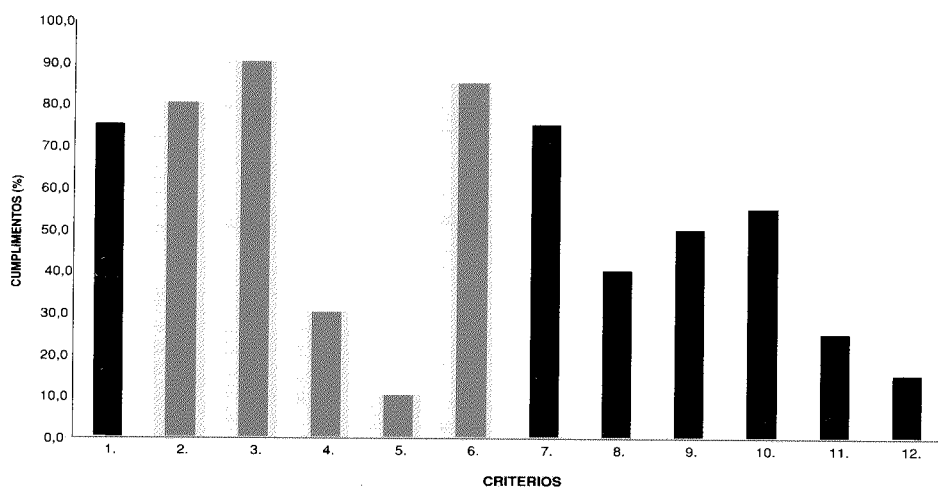
*: Enunciado de modo resumido para la Tabla.

4.1. GRÁFICO DE BARRAS DE LOS CUMPLIMIENTOS

Un gráfico de barras consiste en la representación alineada en el eje de abscisas de las barras que representan los diversos criterios evaluados, con una altura para cada una de ellas que corresponde al nivel de cumplimiento del criterio que representa, según una escala que se coloca como eje de ordenadas. De esta forma podemos comparar visualmente la diferencia en los niveles de cumplimiento de los distintos criterios. La escala de medida puede ser de números absolutos de cumplimientos (tendrá entonces como valor máximo el total de casos de la muestra) o, lo que es más común, porcentajes, yendo entonces de 0 a 100.

La representación de los datos de una evaluación por medio de un gráfico de barras es probablemente la representación gráfica más sencilla y común. Sin embargo, para que sea realmente informativa conviene que ordenemos los criterios y su correspondiente representación de mayor a menor cumplimiento. En caso contrario, si las representamos por el orden del listado de criterios, va a ser de poca ayuda para comparar más allá de distinguir las que se destaquen como más extremas. La Figura 11.2 contiene los gráficos de barras de los datos de la Tabla 11.4, ordenados y sin ordenar para que pueda apreciarse la diferencia.

El gráfico de barras es útil para representar las estimaciones puntuales de los niveles de cumplimientos, sobre todo si ordenamos los criterios de mayor a menor cumplimiento.

FIGURA 11.2. Gráfico de barras (datos de la Tabla 11.4)


CRITERIOS

- | | |
|----------------------------|--------------------------|
| 1. Diagnóstico correcto | 7. Auscultación cardiaca |
| 2. Antecedentes familiares | 8. Exploración abdominal |
| 3. Antecedentes personales | 9. Pulsos periféricos |
| 4. Consumo de tabaco | 10. Soplos carotídeos |
| 5. Consumo de alcohol | 11. Edemas |
| 6. Peso y talla y/o IMC | 12. Fondo de ojo |

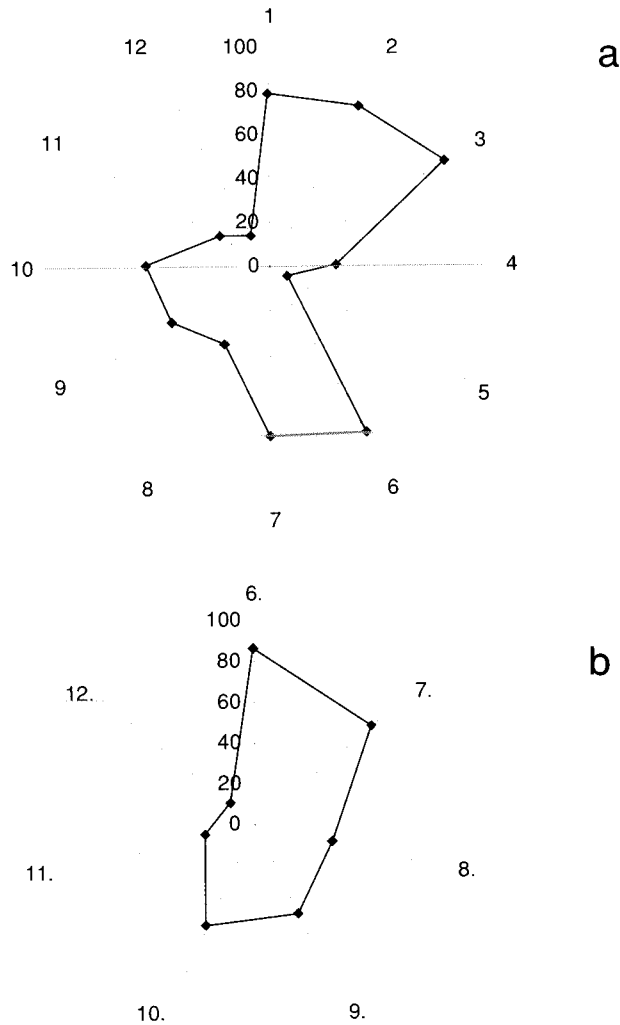
4.2. GRÁFICO TIPO ESTRELLA O RADAR

En este gráfico cada criterio se representa en un eje cuyo origen es el centro de una estrella o asterisco, de forma que cada eje (cada criterio) es como un brazo del asterisco. La escala de los ejes, todos iguales, puede ser sobre total de casos de la muestra o como porcentaje (de 0 a 100), que es como hemos representado los datos de nuestro ejemplo. Al representar cada criterio en un eje, puede verse fácilmente la distancia que le queda a cada uno para llegar al 100% de cumplimiento.

Este tipo de gráfico es interesante para comparar evaluaciones de varios centros o profesionales, cada uno con su "estrella" de la forma y tamaño que corres-

ponda a sus resultados, y también para visualizar la mejora sobre un mismo gráfico al comparar dos evaluaciones de un mismo centro, tal como veremos en la UT 13. Es en cambio una opción menos atractiva cuanto mayor es el número de criterios representados; véase por ejemplo la diferencia entre las Figuras 11.3a y 11.3b. En la primera se han representado todos los criterios de la Tabla 11.4, y en la segunda sólo los correspondientes a la exploración física inicial.

FIGURA 11.3. Gráfico tipo estrella o radar (Datos de la Tabla 11.4)



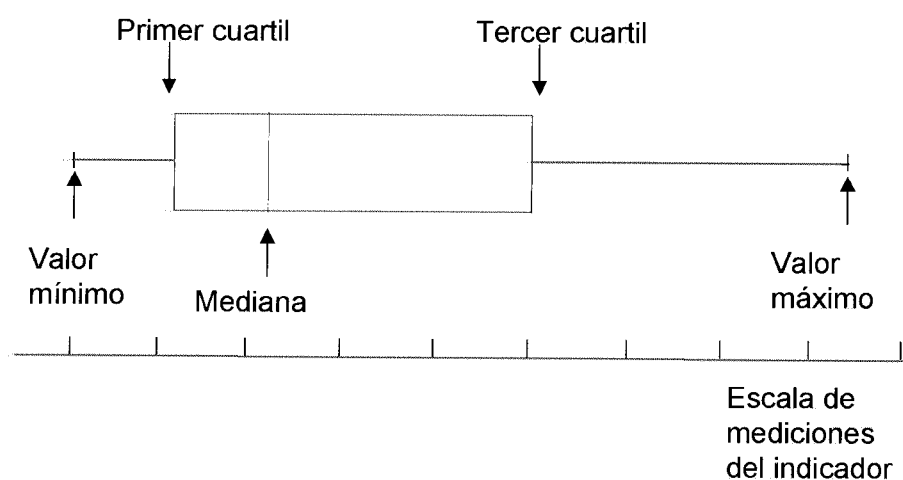
El gráfico tipo estrella es útil sobre todo para representar estimaciones puntuales de niveles de cumplimiento y realizar comparaciones. Es menos práctico cuanto mayor es el número de criterios representados.

4.3. GRÁFICO "BOX AND WHISKERS"

La traducción literal sería gráfico de "caja y bigotes", o, simplificando, gráfico de "cajas" (*box plot*). Es un tipo de gráfico que se propone a finales de los setenta para visualizar los resultados de una muestra de una variable cuantitativa, como una forma de representación de los llamados por Tukey "cinco números resumen" de los datos recogidos, que son el valor mínimo y el máximo, el primer y tercer cuartil (es decir los valores que corresponden al 25% y 75% de los casos),

y la mediana (el valor central de la serie). Los cuartiles delimitan los bordes de la "caja", en la cual está contenida y representada la mediana, mientras que los valores mínimo y máximo determinan la longitud de los "bigotes". Con ello se tiene una idea visual de la variabilidad de los datos del criterio o indicador medido, y permite una comparabilidad más clara que si representamos los datos de cada criterio con histogramas. Este esquema básico es el de la Figura 11.4.

FIGURA 11.4. Esquema básico del gráfico "box and whiskers"



Fte: Adaptado de NR Farnum

El gráfico "box and whiskers" es útil para representar la variabilidad de los datos de la muestra, además de la estimación puntual del parámetro (proporción o media) de interés. Es el gráfico de elección si queremos representar los intervalos de confianza y resulta visualmente más informativo si los criterios se ordenan de mayor a menor nivel de cumplimiento

La idea de representar no solamente las estimaciones puntuales sino también las características que describen la variabilidad de los datos de una muestra ha sido adaptada para múltiples aplicaciones y ha dado lugar a diversas modificaciones de interés. Algunas de ellas son muy sofisticadas, como su utilización para representar las series temporales de las mediciones incorporando en cada medición los resultados de las anteriores, pero otras son sencillas y de una utilidad práctica más inmediata; de estas últimas las tres más relevantes son probablemente la inclusión de la media en el gráfico, modificar los "bigotes" para evidenciar y representar la existencia de "outliers", y, lo que es más importante, para representar las estimaciones de los intervalos de confianza.

Las dos primeras modificaciones son útiles para la representación de criterios cuantitativos (resumidos como media, no como proporción o porcentaje), y consisten en lo siguiente:

- **Incluir la media.** Es simplemente añadir un punto más al gráfico; además de señalar la mediana dentro de la caja, se representa con un punto el valor que corresponde a la media, tal como se ve en la Figura 11.5.
- **Modificar los "bigotes" para señalar "outliers".** Si la muestra tiene casos con valores muy extremos, claramente alejados de la mayoría (lo que se conoce como "outliers") van a producir un gráfico con "bigotes" muy largos en el lado de los "outliers" y va a proporcionar una información distorsionada de la realidad, al no verse en el gráfico la distribución de valores a lo largo del "bigote". Para evitarlo, se limita la longitud de los "bigotes" a los valores que comprenden a $1,5 \times$ amplitud de la caja (recuérdese que la amplitud de la

caja es la distancia entre el primer y tercer cuartil). Los valores que quedan fuera de los "bigotes" modificados se representan como asteriscos y constituyen los "outliers" de la serie de datos representada. El tipo de gráfico que resulta es el de la Figura 11.6.

FIGURA 11.5. Esquema de gráfico "box and whiskers" con mediana y media

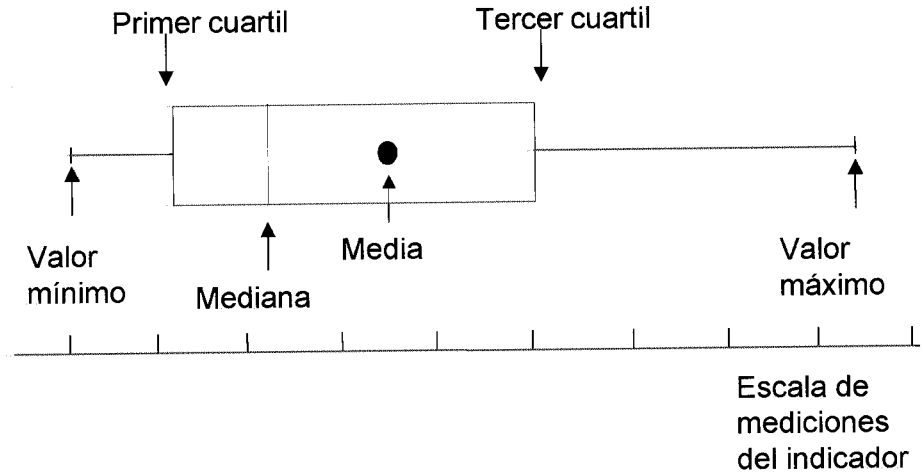
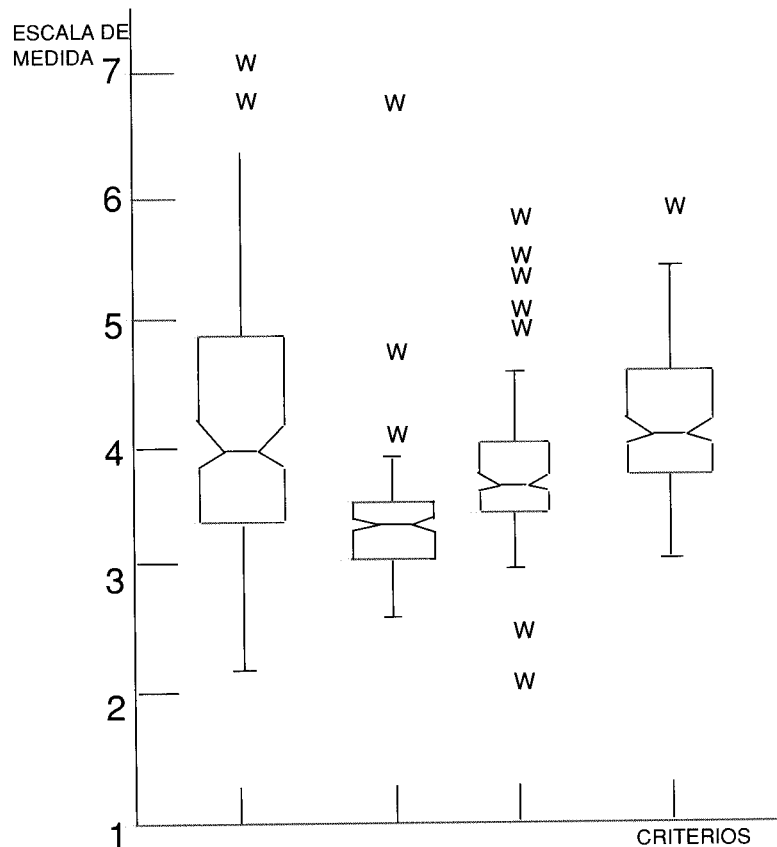


FIGURA 11.6. Gráfico "box and whiskers" modificado, con indicación de "outliers"

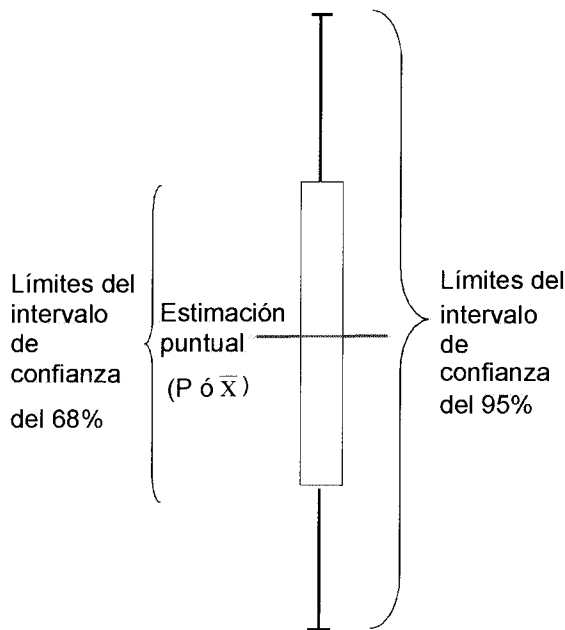


* OUTLIERS.

Fte.: Adaptado de HM. Wadsworth KS. Stephens y AB. Godfrey.

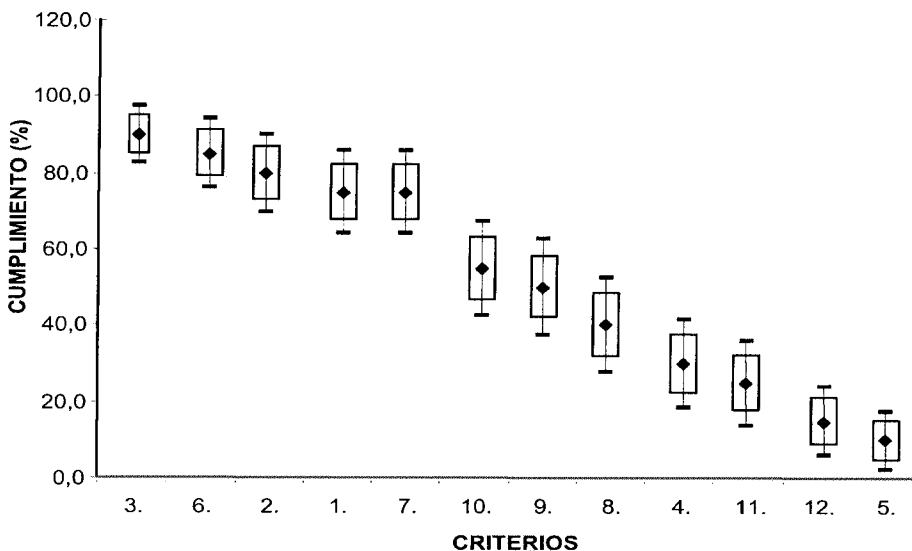
- Representación de los intervalos de confianza.** Ésta es la modificación que más nos interesa, puesto que es aplicable a la representación del nivel de calidad para cualquier tipo de criterio, sea este medido como media o como proporción de cumplimientos. En esta representación gráfica la "caja" delimita el intervalo de confianza del 68% (aproximadamente la estimación puntual ± 1 desviación estándar del parámetro estimado) y los límites de los "bigotes" son los valores correspondientes a los límites del intervalo de confianza del 95%. Este esquema es el que se representa en la Figura 11.7, y es el que hemos utilizado para representar en la Figura 11.8 los datos de la Tabla 11.4.

FIGURA 11.7. Esquema básico para el gráfico "box and whiskers" representando intervalos de confianza



Fte.: Adaptado de HM. Wadsworth KS, Stephens y AB. Godfrey.

FIGURA 11.8. Gráfico "box and whiskers" (Datos de la Tabla 11.4)



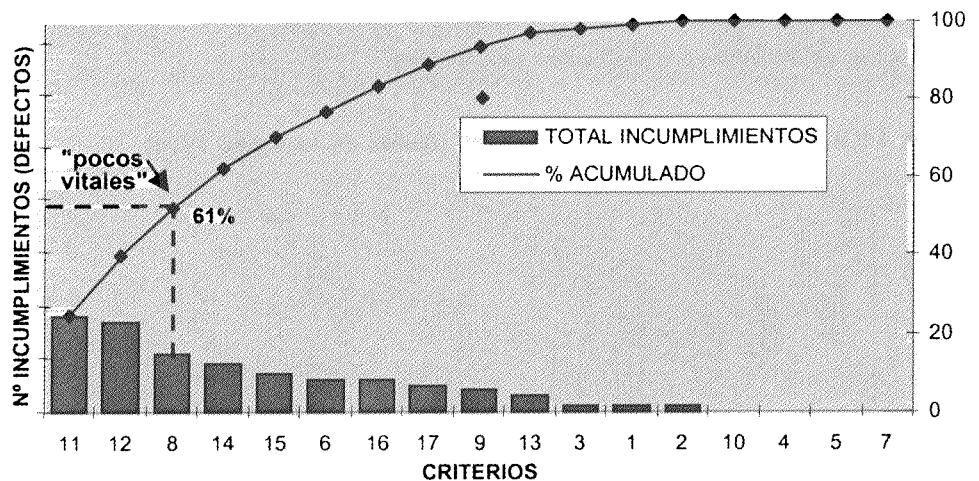
El gráfico "box and whiskers" es una representación estadísticamente más completa que el diagrama de barras, y puede ser mejorada, como el diagrama de barras, ordenando los criterios de mayor a menor cumplimiento, tal como se ha hecho en la Figura 11.8. Es posible, igualmente, combinar ambos gráficos superponiendo el de "caja y bigotes" al diagrama de barras.

De todas formas, aunque todos estos gráficos son útiles para comunicar adecuadamente los resultados de la evaluación, no hay que olvidar que nuestro objetivo al evaluar es saber qué debemos mejorar; en este sentido nuestra preocupación son fundamentalmente los defectos de calidad o incumplimientos encontrados (las causas del nivel de calidad deficiente), y nos conviene enfocar nuestro análisis gráfico en esta dirección. Para ello, la mejor representación y análisis gráfico de los datos es la realización de un diagrama de Pareto sobre los incumplimientos.

5. ANÁLISIS GRÁFICO DE LOS DATOS PARA PRIORIZAR INTERVENCIONES: DIAGRAMA DE PARETO

El diagrama de Pareto (Figura 11.9) es un gráfico de barras que representa la frecuencia de incumplimientos de cada uno de los criterios evaluados, ordenados de mayor a menor número de incumplimientos y completado con una curva de frecuencia porcentual acumulada. El objetivo es visualizar conjuntamente qué criterios aportan mayor número de defectos de calidad (cada incumplimiento es un "defecto de calidad") y señalar en la curva de frecuencia acumulada cuáles son los que representan la mayoría de defectos, y que por tanto serían prioritarios como objeto de intervención para mejorar.

FIGURA 11.9. Diagrama de Pareto (esquema)



Este enfoque de buscar los pocos criterios (o en general causas) que son la base de la mayoría de los defectos de calidad es lo que se conoce como “principio de Pareto” desde que lo propuso Juran a finales de los años cuarenta.

Juran observó que los estudios de Pareto (un economista italiano de finales del XIX) sobre la distribución de la riqueza señalando su desigualdad y acumulación en unas pocas manos respondían en realidad a una ley universal que se repetía en todos los ámbitos; llevado al terreno de la calidad, indicó que lo más común es que la mayoría de los defectos de calidad se deban a unas pocas causas o tipos de defecto, principio que bautizó como “principio de Pareto” o de “los pocos vitales y muchos triviales”, señalando la importancia de buscar unos pocos defectos o causas de defecto “vitales” que serían la base o acumulación de la mayoría de los defectos de calidad del producto o servicio a mejorar. Para su representación gráfica, aparte de proponer como norma que se ordenaran las barras en el gráfico de mayor a menor importancia en cuanto a número de defectos encontrados y no por cualquier otro orden, propuso completar el análisis con la curva de frecuencia de causas acumuladas (basada en las curvas que propuso Lorenz, también sobre distribución de la riqueza) para señalar en ellas donde se alcanza la mayoría de los defectos.

“El principio de Pareto” o de “los pocos vitales y muchos triviales” (ambos en realidad de Juran) se conoce también en la jerga de la calidad en la industria como “principio del 80/20”, indicando que, numéricamente el “principio de Pareto” quiere decir que el 80% de los defectos se debe a un 20% de las posibles causas.

Si consideramos la construcción de un diagrama de Pareto paso a paso, podemos distinguir dos fases: (i) preparación de los datos; y (ii) construcción del gráfico en sí. Los pasos que implica cada una de las fases se explicitan en la Tabla 11.5. Vamos a realizarlos utilizando los datos de nuestro ejemplo (Tabla 11.4).

TABLA 11.5. Construcción de un diagrama de Pareto

A. PREPARACIÓN DE LOS DATOS

1. Hacer una tabla con el listado de criterios y la frecuencia absoluta (número de veces) relativa y (porcentaje sobre el total de incumplimientos de la muestra) de incumplimientos, por cada criterio.
2. Ordenar los criterios de mayor a menor número de incumplimientos.
3. Calcular la frecuencia relativa (%) acumulada.

B. CONSTRUIR GRÁFICO

4. Construir el gráfico de barras, con dos ejes (frecuencia absoluta y relativa) o uno solo (frecuencia relativa).
5. Construir la curva de frecuencia relativa acumulada.
6. (Opcional). Señalar el punto que acumula el 60-80% de los incumplimientos encontrados.

El diagrama de Pareto es un gráfico de barras para representar y analizar las causas de los defectos de calidad (incumplimientos de criterios), aplicando el principio de las “pocas vitales y muchas triviales”, de forma que los criterios se representan ordenadamente de mayor a menor número de incumplimientos y se completa el gráfico con la curva de frecuencia de incumplimientos acumulada para señalar en ella el punto que resume los “pocos vitales”.

5.1. PREPARACIÓN DE LOS DATOS

Para el diagrama de Pareto, la frecuencia relativa (%) de incumplimientos se calcula sobre el total de incumplimientos encontrados en la muestra, no sobre el total de casos (n) de la muestra.

- **Construir una tabla con la frecuencia absoluta (nº de incumplimientos) y relativa (% sobre el total de incumplimientos) encontrados en la muestra.** Según los datos de la Tabla 11.4, la frecuencia absoluta y relativa de incumplimiento de cada criterio es la que figura en la Tabla 11.6. Estos datos los tendremos normalmente (igual que los de los cumplimientos) como resultado en bruto de la recogida de datos que hayamos hecho. No obstante, también pueden derivarse a partir de las proporciones de cumplimiento como las que contienen la Tabla 11.4: obsérvese que “incumplimientos” es el complemento de “cumplimientos” y que, por ejemplo, un cumplimiento del 90%, significa un incumplimiento del 10%; si la muestra es de 60 casos, un 10% de incumplimientos es un total de 6 incumplimientos (un 10% de 60) para ese criterio. De esta forma pueden derivarse criterio a criterio los datos de la Tabla 11.6, a partir de las proporciones de cumplimiento que figuran en la Tabla 11.4.

Conviene subrayar que el porcentaje de incumplimientos NO se calcula sobre el número de casos de la muestra (como haríamos para estimar el porcentaje de cumplimientos de cada criterio) sino *sobre el total de incumplimientos* (defectos de calidad) encontrados; ésta es la distinción fundamental del diagrama de Pareto *sobre los defectos de calidad* en relación con el diagrama de barras representando *el nivel de cumplimientos*; en nuestro ejemplo, cada caso puede contribuir a este total de incumplimientos desde un mínimo de cero (si cumple todos los criterios) hasta un máximo de 12 (si los incumple todos).

TABLA 11.6. Frecuencia de incumplimientos de los criterios evaluados

| CRITERIO | Nº INCUMPLIMIENTOS (Frecuencia absoluta) | % (frecuencia relativa) |
|----------------------------|---|-------------------------|
| 1. Diagnóstico correcto | 15 | 4,4 |
| 2. Antecedentes familiares | 12 | 3,5 |
| 3. Antecedentes personales | 6 | 1,8 |
| 4. Consumo de tabaco | 42 | 12,3 |
| 5. Consumo de alcohol | 54 | 15,8 |
| 6. Peso y talla y/o IMC | 9 | 2,6 |
| 7. Auscultación cardiaca | 15 | 4,4 |
| 8. Exploración abdominal | 36 | 10,5 |
| 9. Pulsos periféricos | 30 | 8,8 |
| 10. Soplos carotídeos | 27 | 7,9 |
| 11. Edemas | 45 | 13,2 |
| 12. Fondo de ojo | 51 | 14,9 |
| | 342 | 100,0 |

- **Ordenar los criterios de mayor a menor número de incumplimientos.** Con esta necesaria ordenación la tabla queda como sigue (Tabla 11.7):

TABLA 11.7. Frecuencia de incumplimientos de los criterios evaluados ordenados de mayor a menor

| CRITERIO | Nº INCUMPLIMIENTOS (Frecuencia absoluta) | % (frecuencia relativa) |
|----------------------------|---|-------------------------|
| 5. Consumo de alcohol | 54 | 15,8 |
| 12. Fondo de ojo | 51 | 14,9 |
| 11. Edemas | 45 | 13,2 |
| 4. Consumo de tabaco | 42 | 12,3 |
| 8. Exploración abdominal | 36 | 10,5 |
| 9. Pulsos periféricos | 30 | 8,8 |
| 10. Soplos carotídeos | 27 | 7,9 |
| 1. Diagnóstico correcto | 15 | 4,4 |
| 7. Auscultación cardiaca | 15 | 4,4 |
| 2. Antecedentes familiares | 12 | 3,5 |
| 6. Peso y talla y/o IMC | 9 | 2,6 |
| 3. Antecedentes personales | 6 | 1,8 |
| | 342 | 100,0 |

- **Calcular la frecuencia relativa (%) acumulada.** Es la base para representar la curva de Lorenz y poder identificar claramente los "pocos vitales" (aquellos criterios responsables de la mayoría de los defectos encontrados). Este cálculo implica solamente ir sumando (*acumulando*) al porcentaje de incumplimientos de cada criterio el porcentaje de incumplimientos que suponen todos los anteriores a él en la tabla. La frecuencia acumulada del último criterio de la lista debe ser, naturalmente, 100. La tabla final de nuestros datos para análisis de Pareto quedaría así (Tabla 11.8)

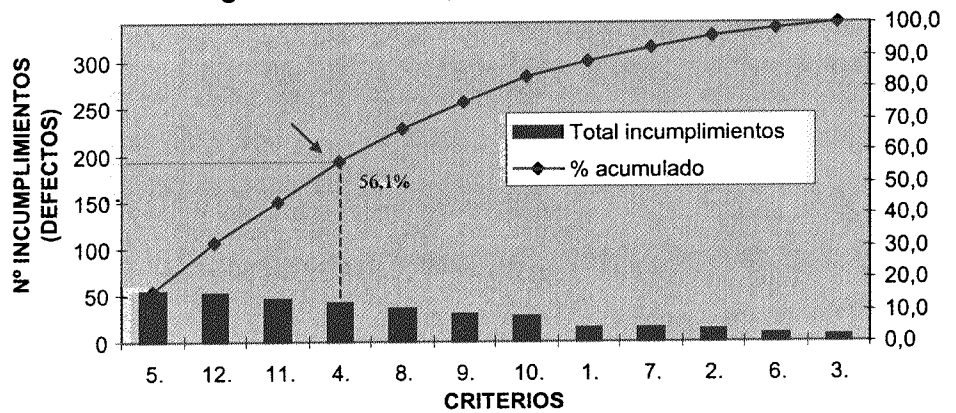
TABLA 11.8. Frecuencia absoluta, relativa (ordenadas) y acumulada de los incumplimientos de los criterios evaluados

| CRITERIO | Nº INCUMPLIMIENTOS (Frecuencia absoluta) | % (frecuencia relativa) | Frecuencia acumulada |
|----------------------------|---|-------------------------|----------------------|
| 5. Consumo de alcohol | 54 | 15,8 | 15,8 |
| 12. Fondo de ojo | 51 | 14,9 | 30,7 |
| 11. Edemas | 45 | 13,2 | 43,9 |
| 4. Consumo de tabaco | 42 | 12,3 | 56,1 |
| 8. Exploración abdominal | 36 | 10,5 | 66,7 |
| 9. Pulsos periféricos | 30 | 8,8 | 75,4 |
| 10. Soplos carotídeos | 27 | 7,9 | 83,3 |
| 1. Diagnóstico correcto | 15 | 4,4 | 87,7 |
| 7. Auscultación cardiaca | 15 | 4,4 | 92,1 |
| 2. Antecedentes familiares | 12 | 3,5 | 95,6 |
| 6. Peso y talla y/o IMC | 9 | 2,6 | 98,2 |
| 3. Antecedentes personales | 6 | 1,8 | 100,2 |
| | 342 | 100,0 | |

5.2. CONSTRUCCIÓN DEL GRÁFICO

- **Construcción del gráfico de barras con dos ejes (frecuencia absoluta y relativa) o uno solo (frecuencia relativa).** El diagrama de Pareto es un gráfico para comparar tipos de defectos de calidad encontrados y señalar cuáles son los “pocos vitales” para diseñar la intervención para mejorar. Para cumplir este objetivo es suficiente con representar las barras con la frecuencia relativa de incumplimientos de cada criterio (porcentaje sobre total de incumplimiento en la muestra) y la curva de la frecuencia porcentual acumulada, con lo cual sólo necesitaríamos un eje para representar los porcentajes. Sin embargo no cuesta mucho construir el gráfico más completo, incluyendo la información sobre el número absoluto de defectos que corresponde a cada criterio. Para ello, construimos el gráfico con dos ejes: uno a la izquierda para representar la frecuencia absoluta (nº de incumplimientos) y otro a la derecha para representar los porcentajes (frecuencia relativa); el eje para la frecuencia absoluta tiene como rango de 0 al total de incumplimientos de la muestra, y el de la frecuencia relativa de 0 a 100. Ambos deben tener el mismo tamaño. Cada barra, representando a cada uno de los criterios, tiene la altura que corresponde, a la vez, a su número de incumplimientos (en el eje de la izquierda) y del porcentaje que supone sobre el total (en el eje de la derecha). El eje de la derecha sirve a su vez para referencia de los puntos de la curva de frecuencia acumulada.
- **Completar el diagrama de barras con la curva de frecuencia relativa acumulada.**
- **Señalar el punto que acumula la mayoría (en torno a 60% ó más) de los incumplimientos.** El diagrama de Pareto que corresponde a nuestros datos es el de la Figura 11.10:

FIGURA 11.10. Diagrama de Pareto (Datos de la Tabla 11.8)



CRITERIOS

- | | |
|----------------------------|--------------------------|
| 1. Diagnostico correcto | 7. Auscultación cardiaca |
| 2. Antecedentes familiares | 8. Exploración abdominal |
| 3. Antecedentes personales | 9. Pulsos periféricos |
| 4. Consumo de tabaco | 10. Soplos carotídeos |
| 5. Consumo de alcohol | 11. Edemas |
| 6. Peso y talla y/o IMC | 12. Fondo de ojo |

El diagrama de Pareto más completo tiene dos escalas y reúne dos gráficos: la escala y eje de la izquierda va de 0 hasta el total de incumplimientos de la muestra y sirve de referencia para el gráfico de barras de la frecuencia absoluta (nº de incumplimientos) de cada criterio; la escala de 0 a 100 en el eje de la derecha sirve de referencia para representar la curva de frecuencia relativa acumulada. Ambas escalas deben tener idéntica altura.

Como vemos, hay 4 criterios que representan el 56,1% de los defectos de calidad (incumplimientos) encontrados y serían los “pocos vitales” sobre los que debemos actuar prioritariamente en nuestro esfuerzo para mejorar.

El diagrama de Pareto es una forma sencilla, brillante e inteligente de presentar y analizar los datos de una evaluación, iluminando donde está el problema, es decir, lo que tenemos que mejorar. Sin embargo, es conveniente hacer algunas puntualizaciones para su construcción y uso:

5.3. OBSERVACIONES SOBRE LA CONSTRUCCIÓN E INTERPRETACIÓN DE UN DIAGRAMA DE PARETO SOBRE DEFECTOS DE CALIDAD

En primer lugar hay que indicar que se asume que todos los criterios y sus correspondientes defectos son igualmente importantes; cualquier valoración sobre la diferente relevancia de uno u otro tipo de defecto de calidad debe hacerse al margen del análisis con gráfico de Pareto, y podrá, obviamente, modificar las prioridades que marca el diagrama.

La segunda observación es de tipo formal: cuando se representan un gran número de criterios, es habitual que varios de ellos tengan un número pequeño de incumplimientos y se agrupen en una sola barra como “otros” al final del diagrama; en estos casos hay que asegurarse siempre que el número de incumplimientos de cualquier criterio agrupado en “otros” es menor que el último representado de forma independiente. Por otra parte, no hace buen efecto una barra de “otros” mucho más alta que la última representada: rompe la imagen del diagrama.

Finalmente, aunque lo más frecuente es que el “principio de Pareto” se cumpla, no siempre es así y podemos encontrarnos con diagramas relativamente “planos”, en los que todos los criterios tienen niveles parecidos de incumplimientos. En estos casos, la priorización para intervenir, si es que se establece alguna, se basará en otras consideraciones (importancia relativa de cada criterio, facilidad de asumir su mejora, etc.) que no son las puramente cuantitativas.

6. ANÁLISIS DE LOS DATOS DE ESTUDIOS DE CAUSAS HIPOTÉTICAS

En las UT 9 y 10 vimos dos tipos de estudios dentro de los ciclos de mejora: estudios del nivel de calidad, en los cuales medimos el nivel de cumplimiento de los requisitos o criterios que sabemos de antemano se relacionan con la calidad (en realidad son la definición operativa de calidad del servicio que se evalúa); y estudios de causas hipotéticas, cuando lo que medimos es, normalmente, sólo un criterio de calidad y buscamos qué causas (frecuentemente de tipo organizacional) pueden hacer que esté a niveles no deseables. Por ejemplo: ¿por qué el tiempo de espera es más largo de lo deseable?, ¿Qué factores estructurales y organizacionales influyen en que el diagnóstico de la hipertensión no se haga correctamente?, etc., etc. El segundo tipo de estudios precisa de la medición simultánea del nivel de calidad (reducido normalmente a un solo criterio) y las causas hipotéticas que influyen en la mala calidad.

El diagrama de Pareto asume que todos los tipos de defectos son igualmente importantes

El “principio de Pareto no siempre se cumple. En estos casos, el diagrama no resulta útil para priorizar intervenciones para mejorar

Como vimos en la UT 10, el análisis de los datos de este tipo de estudios consiste esencialmente en ver el grado de asociación entre las diversas causas hipotéticas y el nivel de calidad. Si las causas han sido "medidas" subjetivamente (por ejemplo mediante encuesta de causas o votación múltiple) su análisis puede hacerse mediante un diagrama de Pareto; si por el contrario las causas hipotéticas se ha medido empíricamente y tenemos datos reales, podemos realizar análisis estadístico y gráfico para evidenciar su asociación con el criterio o indicador de calidad que representa el problema a solucionar.

BIBLIOGRAFÍA

- Briones G: Métodos y técnicas de investigación para las ciencias sociales. México: Trillas; 1990.
- Richards LE, LaCava JJ. Business Statistics. New York: McGraw Hill; 1983.
- Juran JM. La función de la calidad. En: Juran JM, Gryna FN, Bingham RS. Manual de control de la calidad. Vol.1. 4º ed. Madrid: McGraw Hill; 1993. p.2.1-2.13.
- Farnum NR. Modern Statistical quality control and improvement. Belmont: Duxbury Press; 1994.
- Wadsworth HM, Stephens KS, Godfrey AB. Modern methods for quality control and improvement. New York: John Wiley & sons; 1986