

AUDIT ONCO^{Orisc}: calidad de la información en el estudio sobre resultados en cirugía oncológica digestiva

Informes de Evaluación de Tecnologías Sanitarias
AATRM Núm. 2007/01

INFORMES, ESTUDIOS E INVESTIGACIÓN

AUDIT ONCOOrisc: calidad de la información en el estudio sobre resultados en cirugía oncológica digestiva

Informes de Evaluación de Tecnologías Sanitarias
AATRM Núm. 2007/01

AUDIT ONCOOrisc: calidad de la información en el estudio sobre resultados en cirugía oncológica digestiva / Mercè Salvat, Cristian Tebé, Cari Almazán, Mireia Espallargues, Emilia Sánchez, Joan M. V. Pons, Roger Pla.-- Madrid: Plan de Calidad para el Sistema Nacional de Salud. Ministerio de Ciencia e Innovación. Barcelona: Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya. 2009.- 52 p; 24 cm.-- (Colección: Informes, estudios e investigación / Ministerio de Ciencia e Innovación / Informes de Evaluación de Tecnologías Sanitarias; AATRM 2007/01)

1. Aparato digestivo -- Cáncer -- Cirugía -- Evaluación -- Cataluña 2. Historias clínicas -- Evaluación -- Cataluña

I. España. Ministerio de Ciencia e Innovación II. Cataluña. Departament de Salut. Generalitat de Catalunya III. Cataluña. Agència d'Avaluació de Tecnologia i Recerca Mèdiques.

Autores:

Mercè Salvat (1), Cristian Tebé (1), Cari Almazán (1), Mireia Espallargues (1), Emilia Sánchez (1), Joan M.V. Pons (1), Roger Pla (2)

Revisor:

Salvador Peiró (3)

(1) Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya

(2) Institut Català de la Salut a les Terres de l'Ebre

(3) Centro Superior de Investigación en Salud Pública

Para citar este informe: Salvat M, Tebé C, Almazán C, Espallargues M, Sánchez E, Pons JMV, Pla R. AUDIT ONCOOrisc: calidad de la información en el estudio sobre resultados en cirugía oncológica digestiva. Madrid: Plan de Calidad para el Sistema Nacional de Salud. Ministerio de Ciencia e Innovación. Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya; 2009. Informes de Evaluación de Tecnologías Sanitarias, AATRM núm. 2007/01.

Edita: Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya

Roc Boronat, 81-95

08005 Barcelona

www.aatrm.net

Fotocomposició: Sintagma, edicions corporatives

Impresió: Milenio Manipulados

Nipo: 477-09-023-8

ISBN: 978-84-393-8071-9

Depósito Legal: B-17013-2009

© Ministerio de Ciencia e Innovación

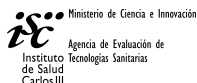
© Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya

Este documento puede ser reproducido parcial o totalmente para uso no comercial, siempre que se cite explícitamente su procedencia.

AUDIT ONCOOrisc: calidad de la información en el estudio sobre resultados en cirugía oncológica digestiva

Informes de Evaluación de Tecnologías Sanitarias
AATRM Núm. 2007/01

Este documento se ha realizado en el marco de colaboración previsto en el Plan de Calidad para el Sistema Nacional de Salud elaborado por el Ministerio de Sanidad y Consumo, al amparo del convenio de colaboración suscrito por el Instituto de Salud Carlos III, organismo autónomo del Ministerio de Ciencia e Innovación, y la Agència d'Avaluació de Tecnologia i Recerca Mèdiques AATRM.



Índice de autores

AUTORES

- **Mercè Salvat.** Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya
- **Cristian Tebé.** Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya
- **Cari Almazán.** Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya
- **Mireia Espallargues.** Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya
- **Emilia Sánchez.** Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya
- **Joan M. V. Pons.** Agència d'Avaluació de Tecnologia i Recerca Mèdiques de Catalunya
- **Roger Pla.** Gerente Territorial - Institut Català de la Salut a les Terres de l'Ebre

REVISIÓN

- **Salvador Peiró.** Jefe del área de investigación de servicios de Salud. Centro Superior de Investigación en Salud Pública

Agradecimientos

Los autores agradecen los comentarios y sugerencias recibidos del Dr. Salvador Peiró, investigador de la Unidad de Investigación en Servicios de Salud y Farmacoepidemiología del Centro Superior de Investigación en Salud Pública (CSIP) de la Conselleria de Sanidad de la Generalitat Valenciana; y de la Dra. Aida Ribera, investigadora de la Unitat d'Epidemiologia Cardiovascular del Hospital Universitario Vall d'Hebron de Barcelona.

Índice

Resumen	11
Introducción	15
Objetivos	23
Metodología	25
Resultados	29
Discusión	37
Conclusiones	41
Bibliografía	43
Anexos	45
Anexo 1. Variables de estudio	45
Anexo 2. Cálculo de la muestra	48
Anexo 3. Análisis por centros	51

Resumen

Las historias clínicas de los pacientes son una de las principales fuentes de obtención de datos en investigación médica y evaluación de servicios sanitarios. La obtención y la transcripción rigurosa de la información aumentan en gran medida la calidad de ésta, a pesar de que la organización de las historias clínicas y los datos que contienen son la principal fuente de problemas y sesgos de información.

L'Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM) y el Plan Director de Oncología llevaron a cabo, en Cataluña, entre los años 2002 y 2005, el estudio *Evaluación del procedimiento terapéutico y sus resultados en oncología digestiva: ONCORisc*, para evaluar la efectividad de la cirugía en los diferentes tipos de cáncer digestivo. Se realizó un estudio de cohortes multicéntrico ambiespectivo cuya principal fuente de información fue la historia clínica. Profesionales de los propios centros participantes realizaron la extracción de los datos, y para asegurar la calidad de éstos, en el protocolo del estudio ONCORisc se utilizaron una serie de estrategias y controles de calidad, entre ellos la realización de una auditoría para validar la calidad de los datos obtenidos en el estudio. Se identificarían los errores en la extracción de datos, entre la historia clínica y la base de datos del estudio ONCORisc; y los errores en la transcripción, entre los cuadernos de recogida de datos y la base de datos del estudio.

Objetivos: Evaluar las coincidencias y discrepancias entre los datos recogidos en la base de datos del estudio ONCORisc y los datos contenidos en la historia clínica.

Método: Se ha revisado una muestra aleatoria de historias clínicas estratificada por región sanitaria. Se ha auditado un subgrupo de 31 variables, entre demográficas y clínicas, del paciente, de proceso y de resultado. La extracción de datos la llevó a cabo personal propio de la AATRM entre abril y noviembre de 2007. Se analizaron las coincidencias y discrepancias entre los datos recogidos en la auditoría y la base de datos del estudio y se realizó un análisis descriptivo del número y del porcentaje de discrepancias detectadas en total, por etapa, sistema y tipo de variable. Se calculó el estadístico Kappa por variable para el estudio de concordancia.

Resultados: Se auditaron 322 historias clínicas, 188 de la etapa retrospectiva y 134 de la etapa prospectiva del estudio ONCORisc. Se incluyeron 14 hospitales de los 49 participantes. El porcentaje de coincidencias entre la base de

datos del estudio y la auditoría fue del 88,8% (rango 51,5%-100%) del total de variables examinadas (9.982). De las 31 variables incluidas, 19 presentaron un número de coincidencias superior al 90%. La mortalidad al alta presentó el 100% de coincidencias y obtuvieron resultados superiores al 95% de coincidencias las siguientes variables: sexo, fecha de muerte al alta, tipo de cáncer, reintervención intrahospitalaria, radioterapia y quimioterapia primarias y antecedentes de diabetes. El porcentaje más bajo de coincidencias (51,5%) fue para el TNM, seguido del número de pruebas, recidiva a los 3 y 6 meses, antecedentes neoplásicos y la American Society of Anesthesiology (escala ASA), con valores situados entre el 70% y el 80% de coincidencias. En los datos concernientes a evaluaciones en el tiempo (mortalidad, recidivas y reintervenciones) disminuyó el porcentaje de coincidencias con el aumento del periodo de seguimiento a los 3 y 6 meses.

El 6,8% de las discrepancias, del total de variables en las que se produjeron discrepancias (1.120), 0,7% (n=9.982) del total de valores comparados, fueron debidas al sistema de lectura y transcripción automatizado de los datos. Por etapas, retrospectiva y prospectiva, se obtuvieron resultados similares. La concordancia interobservador fue importante en el mayor número de variables, 15 de las 25 valoradas.

Conclusión: El control de la calidad de los datos es necesario para conocer la validez de los resultados de la investigación, de ahí la necesidad de incorporar auditorías en los estudios. Mediante la presente auditoría se constató la intervención de diferentes factores que causaron problemas y sesgos de información en la extracción de los datos, en gran medida relacionados con el contenido de las historias clínicas. Se confirma la necesidad de poner la máxima atención en las estrategias de recogida de datos para maximizar su calidad.

Executive summary

The patients' clinical record is one of the main sources of information in medical research and health care services assessment. A sound taking and transcription of the information increase in great measure the quality of clinical records, although the organisation of clinical records and the data contained in them are the main source of problems and information biases.

In Catalonia, the Catalan Agency for Health Technology Assessment and Research (CAHTA) and the Oncology Steering Plan conducted between the years 2002 and 2005 the study *Assessment of the therapeutic procedure and its results in digestive oncology*, to assess the effectiveness of surgery in the different types of digestive cancer. An ambispective, multi-center cohort study was conducted, having clinical records as main information sources. Professionals working in the participating centres carried out the data extraction, and, to ensure the quality of the data, the ONCORisc study protocol used a series of strategies and quality controls, including an audit conducted to validate the quality of the data obtained throughout the study. Errors in data extraction, comparing the clinical records and the ONCORisc study database were identified, as well as transcription errors, comparing case report forms and the ONCORisc study database.

Objectives: To assess the coincidences and discrepancies between the data collected in the ONCORisc study database and the data contained in the clinical records.

Method: A random sample of clinical records stratified by health care region was reviewed. A subgroup of 31 patient's variables, both demographical and clinical, and belonging both to the procedure and the results, were audited. Data extraction was conducted by the CAHTA's staff between April and November 2007. The coincidences and discrepancies between the data collected in the audit and the ONCORisc study database were analysed, and a descriptive analysis of the number and percentage of total discrepancies, and by stage, system and type of variable, was conducted. The Kappa statistic by variable was calculated for the concordance study.

Results: Three hundred and twenty-two clinical records were audited, 188 in the retrospective stage, and 134 in the prospective stage of the ONCORisc study. The 14 hospitals of the 49 participants were included. The percentage of coincidences between the ONCORisc study database and the audit was 88.8% (range 51,5%-100%) of the total number of variables examined

(9,982). Of the 31 variables included, 19 showed over 90% coincidences. Discharge mortality showed 100% coincidences, and the following variables showed a percentage of coincidences higher than 95%: sex, date of death at discharge, type of cancer, intra-hospital re-intervention, primary radiotherapy and chemotherapy, and history of diabetes. The lowest percentage of coincidences (51.5%) was for TNM, followed by number of test, relapse at 3 and 6 months, previous neoplasm, and ASA (American Society of Anaesthesiology) scale, with coincidence values ranging between 70% and 80%. In the data regarding timeline assessments (mortality, relapses and re-intervention) had a decreased percentage of coincidences with increasing follow-up periods at 3 and 6 months. From all discrepancies, 6.8% (n=1,120) of the total variables and 0.7% (n=9,982) of the total compared values, were due to the automated reading and transcription of the data. Similar results were obtained by stage, both retrospective and prospective. Intra-observer concordance was significant in most variables, with 15 out of 25 variables observed.

Conclusion: The data quality control is necessary to know the validity of the research results; thus, there is a need to include audits in the studies. This audit showed that there are different factors causing problems and information biases in data extraction, in great part related to the contents of the clinical records. There is a real need to underscore the application of accurate collection strategies to optimise data quality.

Introducción

La calidad de una investigación está estrechamente relacionada con los datos recogidos y con el modo de obtenerlos. Dicha calidad determinará el grado de correspondencia entre las conclusiones de un estudio y la realidad. Etapas como el diseño y el análisis podrían relegar a un segundo plano la atención que merece la recogida de datos, pero de ésta dependen sus resultados¹.

En investigación médica y evaluación de servicios sanitarios, una de las principales fuentes de obtención de datos son las historias clínicas (HC) de los pacientes. Aunque su utilización supone numerosas ventajas respecto a otras fuentes (mayor accesibilidad, elevado volumen de información, menores costes), la calidad de los datos que contienen es cuestionable^{2,3}. En general, la estructura de las HC está definida y estandarizada, no obstante, su contenido es variable entre centros y servicios. La ausencia de datos, la subjetividad en la interpretación de los resultados, la cumplimentación de las HC por parte de profesionales de distintas especialidades, entre otros factores, son algunas de las causas de la variabilidad en los datos registrados y, en consecuencia, pueden ocasionar problemas de interpretación y transcripción, al mismo tiempo que dificultan la verificación de la información^{2,4}.

Además, se trata de información sobre el proceso asistencial cuya función primaria es el soporte a los cuidados y tratamientos del paciente y no a la investigación^{2,4,6}. De ahí que algunos autores hayan destinado esfuerzos al control de la calidad de los datos obtenidos mediante HC, evaluando cómo éstos se obtienen^{7,8} y elaborando guías y recomendaciones para la correcta extracción de datos de las HC^{2,4,9,10}. A pesar de que la calidad de los datos recogidos depende, en gran medida, de la persona responsable de su recogida, siendo claves su experiencia, formación y conocimiento del protocolo de investigación, también existen otros factores, como el método de recogida de datos, la fuente de obtención de éstos o su transcripción, que influyen estrechamente en que la información sea de calidad¹.

Estrategias y control de calidad de los datos

El contenido de las HC es la principal fuente de problemas y sesgos de información de la recogida de datos retrospectivos. A pesar de ello, sólo está al alcance de los investigadores identificar en qué medida los posibles sesgos de información pueden afectar a los objetivos concretos de un estudio, y concentrar el máximo de esfuerzos en la obtención y el registro de los datos para aumentar la calidad de éstos.

En toda investigación en la que se requiere una recogida de datos, las principales estrategias para aumentar la calidad de éstos se centran en:

Tabla 1. Estrategias para asegurar la calidad de la obtención de datos de HC

1. Elaboración y utilización de una herramienta de recogida de datos
2. Elaboración de un manual de procedimientos
3. Selección y entrenamiento de los investigadores
4. Transcripción y control continuo de los datos
5. Evaluación de la calidad de los datos

1. La utilización de una herramienta de recogida de datos (instrumento, *dosier* o cuaderno de recogida de datos) adecuadamente diseñada. Su finalidad, además del registro de la información, es la estandarización de ésta², por lo que debe tratarse de un documento o soporte electrónico estructurado, claro, simple y ordenado de forma lógica^{4,11}, que facilite la obtención de datos y la comprensión de la información. Se recomienda maximizar la utilización de ítems de respuesta simple y objetiva^{3,4}. Se han señalado entre sus ventajas la reducción de errores de transcripción, así como evitar interpretaciones inapropiadas de los evaluadores².

2. La elaboración detallada de un manual de procedimientos estrechamente vinculado a la herramienta de recogida de datos, que incluya todas las instrucciones necesarias sobre cómo debe ser recogida la información y dónde encontrarla. Se trata de un manual de guía e instrucción con el objetivo de uniformizar la extracción de datos, minimizando la variabilidad entre los investigadores. Además de una breve explicación del protocolo de investigación, el manual deberá contener los criterios de inclusión y exclusión, la definición detallada de las variables y la definición operativa de éstas para la obtención de los datos y su registro posterior^{2,4}.

3. La selección y el entrenamiento de los investigadores. La formación en el campo de la salud, la experiencia clínica en el área de interés y la experiencia previa en investigación y revisión de HC son elementos fundamentales en la selección del equipo investigador^{2,4,8,9}. El intenso entreno de los investigadores responsables de la extracción de datos permite, principalmente, reducir la variabilidad entre evaluadores y evitar la libre interpretación de los datos que se recogen. La formación del equipo debe incluir el conocimiento del manual de procedimientos, de la herramienta de recogida de datos y la utilización práctica de ésta⁸⁻¹⁰, así como el entreno en el examen de la información del paciente, las notas de los clínicos y los procesos médi-

cos. Además, las reuniones periódicas con el equipo investigador favorecen la resolución de dudas, la discusión de ambigüedades^{2,4} y, en caso necesario, la introducción de modificaciones en las herramientas de obtención de datos⁹. En estudios de envergadura y multicéntricos se recomienda emplear un único equipo de investigadores para los diferentes lugares de estudio²; no obstante, los colaboradores de los centros tendrán mayores habilidades en el manejo de la documentación. Otros factores que se deben tener en cuenta son, por un lado, mantener el anonimato de las hipótesis del estudio para minimizar la subjetividad y los potenciales sesgos de información de los revisores y, por otro, es conocido que informar de la realización de controles de calidad aumenta la atención en los procesos y, en consecuencia, el rigor en la obtención de los datos⁴.

4. La transcripción meticulosa y el control continuo de los datos. La lectura e introducción de los datos de investigación en una base de datos informatizada, para su posterior análisis, se realiza de forma manual o mediante procesos informatizados como los sistemas de lectura óptica que traducen, interpretan y verifican los datos exportándolos a bases electrónicas (teleForm, MedQuest) para el posterior análisis de los datos^{9,12,13}. También pueden emplearse cuestionarios de recogida de datos en formato electrónico mediante sistemas directos de introducción de datos (ordenadores portátiles, PDA). El proceso de verificación de los datos se centra en la detección continua de errores de obtención e introducción de datos (ausencia de datos relevantes, comprobación de incoherencias) y la corrección de éstos contactando con los investigadores^{8,11}. Entre las medidas de control de los datos se utilizan la transcripción, mediante la técnica de la doble entrada de los datos al sistema informático para resolver las discrepancias entre introductores, y el muestreo de algunos casos contrastando su contenido campo por campo con la fuente original. El desarrollo de nuevas tecnologías para la transcripción y gestión de datos permite agilizar el proceso, reducir los posibles errores de transcripción^{11,13} y mantener la calidad de los datos, especialmente cuando la obtención de éstos implica un elevado número de variables y la revisión de muchas HC⁹. A pesar de todo, se reconoce como ventajosa la obtención de datos mediante un cuestionario en formato papel dado que permite validaciones posteriores².

5. Estrategias de validación/monitorización de los datos. Se recomienda la realización de auditorias o evaluaciones al inicio, durante i/o al final del proceso de obtención de datos para asegurar su fiabilidad, determinar y resolver dificultades o como medida de calidad. Mediante la concordancia entre colaboradores se mide el porcentaje de acuerdo en la extracción de datos cuando dos o más personas obtienen los datos de una misma fuente. Con la

concordancia intracolaboradores una misma persona obtiene los mismos datos en dos tiempos. Según el tipo de variable, una valoración de Cohen's Kappa o un coeficiente de correlación intraclases determina estadísticamente la concordancia^{3,4,8,9}.

El caso particular de los estudios basados en la obtención de datos retrospectivos merece especial atención. A pesar de la infrautilización de esta técnica cuya calidad ha sido cuestionada, representa una de las mayores fuentes de datos en investigación dada la riqueza de la información obtenida, su disponibilidad y su bajo coste. Además, permite obtener datos para la comparación y validación de hipótesis en estudios prospectivos.

Gearing *et al.* proponen una metodología para llevar a cabo investigaciones retrospectivas mediante la revisión de HC⁴. Han elaborado una guía metodológica, basada en nueve pasos, para el desarrollo de un protocolo de estudio y tratan de forma más detallada la extracción efectiva de datos haciendo especial énfasis en recomendaciones destinadas a garantizar la calidad de éstos: 1) concepción; 2) revisión de la bibliografía; 3) desarrollo del protocolo; 4) instrumento de extracción de datos; 5) desarrollo de protocolos y guías para la extracción; 6) extracción de datos; 7) muestra; 8) ética, y 9) estudio piloto. Destaca la importancia de la realización de un estudio piloto, no sólo como un ensayo y una medida de validación de los instrumentos de recogida de datos, sino como parte de las principales medidas de control de calidad. Se señala este momento como la mejor oportunidad para medir o analizar la fiabilidad de la extracción de datos puesto que permite aplicar las medidas correctoras necesarias. En la Tabla 2 se resumen las principales estrategias para asegurar la calidad de los datos.

Estudio ONCORisc

En Cataluña, l'Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM) y el Plan Director de Oncología llevaron a cabo el estudio *Evaluación del procedimiento terapéutico y sus resultados en oncología digestiva: ONCORisc*, para evaluar la efectividad de la cirugía en los diferentes cánceres digestivos: de esófago, estómago, recto, páncreas y metástasis hepáticas, en los hospitales, tanto públicos como privados, de Cataluña¹⁴. Con el objetivo de estudiar la variabilidad en el proceso y los resultados de la cirugía oncológica digestiva en Catalunya, así como los factores asociados a ésta, se realizó un estudio de cohortes multicéntrico ambispectivo (etapas retrospectiva y prospectiva). Retrospectivamente, se estudiaron todos los pacientes sometidos a un tratamiento quirúrgico (curativo o paliativo) durante el año 2002. Prospectivamente, fueron incluidos, de forma consecutiva,

Tabla 2. Resumen de las principales estrategias para la revisión de historias clínicas				
Documento	Instrumento de recogida de datos (4)	Protocolo y guía de extracción (5)	Extracción de datos (6)	Estrategias para asegurar la calidad de los datos
Gearing RE <i>et al.</i> (2006) ⁴	<p>Características: Organización, simplicidad y claridad Orden lógico de los datos Sección de respuesta simple libre de ambigüedades Formato papel (menor coste efectividad, fácil aplicación) o electrónico (mayor coste efectividad, reducción de errores, facilita centralización y acceso a los datos)</p> <p>Transcripción y control de los datos: Utilización de un software de extracción de datos para la traducción a formato electrónico, control de calidad y gestión de los datos (análisis e informe de resultados)</p>	<p>Elaboración de un manual para la instrucción de los revisores (cuándo y cómo extraer los datos de HC)</p> <p>Características: Lista todas las variables Explica cómo se registra cada variable en el instrumento de recogida de datos Describe la localización de cada variable en las HC Determina criterios para aumentar la concordancia interobservadores: directrices para la toma de decisiones en situaciones ambiguas Describe la gestión de datos ausentes Revisable según necesidades (p. ej., después del estudio piloto)</p>	<ol style="list-style-type: none"> 1. Conocimiento de los requerimientos específicos de cada institución 2. Selección y entreno de los investigadores <p>Recomendaciones: Preferible con experiencia en investigación retrospectiva y/o en área de interés Profesionales de la salud con nivel de formación avanzado Determinar el número y si son internos o externos de los centros. Preferible un equipo de revisores externos. Alternativa: uno o dos investigadores clave asistiendo, entrenando y auditando a los colaboradores de los centros</p> <p>Características y entreno: Familiarización con las HC Conocer la ubicación de la información necesaria Objetividad Entreno en la utilización del instrumento de recogida de datos y de los protocolos y guías Procedimientos de resolución de conflictos con reuniones de práctica y resolución de dudas y ambigüedades</p>	<p>Mínimo 2 evaluadores y máximo 4 para asegurar la fiabilidad entre colaboradores</p> <p>Desconocimiento de los revisores de las hipótesis para reducir la subjetividad y las interpretaciones</p> <p>El conocimiento de la realización de controles de calidad aumenta la precisión de los revisores</p> <p>Determinar la fiabilidad del instrumento y de la extracción de datos mediante la realización de un estudio piloto y/o controles aleatorizados. Dependiendo del nivel de medida de las variables, una valoración de Cohen Kappa o un coeficiente de correlación intraclass (ICC)</p>

Basado en las recomendaciones metodológicas de Gearing RE *et al.*

los pacientes operados en cada centro durante un periodo de 6 meses (entre noviembre de 2003 y julio de 2005). En ambas etapas se excluyeron aquellos pacientes sometidos a reintervenciones e intervenciones en dos tiempos (con una primera intervención previa al periodo de estudio).

Se recogió información sobre el paciente, el proceso y los resultados de la atención previamente a la cirugía, al alta hospitalaria y a los 3 y 6 meses (Tabla 3). Asimismo, se recopiló información relacionada con la estructura de los centros (Anexo 1). Todos los servicios de cirugía general de los hospitales públicos (N=61) y privados (N=10) de Cataluña fueron invitados a participar, de los que accedieron a participar en total 49 centros (69%), 43 públicos y 3 privados. Se identificaron 3.108 pacientes en total, entre los que no participaron 69 (2%), siendo los motivos más frecuentes el no cumplimiento de los criterios de inclusión, la no localización de la historia clínica o el no haber recibido la información de los centros participantes.

Tabla 3. Relación de variables del estudio ONCORisc

Características del paciente	Proceso de atención	Resultados
Datos sociodemográficos y antropométricos	Tratamiento comorbilidad	Mortalidad (supervivencia)
Antecedentes	Tratamientos primarios y adyuvantes (radioterapia y quimioterapia)	Tiempo libre de enfermedad (recidiva)
Tipo, localización, estadiaje, histología del cáncer	Complicaciones del tratamiento adyuvante	Complicaciones (generales, técnica y herida)
Marcadores tumorales	Características de la cirugía (técnica, tipo, intención, tiempos)	Reintervenciones
Comorbilidad	Profilaxis antibiótica y antitrombótica	Estado funcional
Signos y síntomas	Reintervenciones	
Analítica	Datos administrativos (circunstancia de ingreso y alta, servicios, estancia, reingresos)	
Pruebas complementarias		
Estado funcional previo al tratamiento		
Riesgo quirúrgico		

En la etapa retrospectiva, la principal fuente de información fue la HC. En la prospectiva, los datos se obtuvieron, en la medida de lo posible, directamente del cirujano y a partir de la HC. Para su registro se elaboró un cuaderno de recogida de datos (CRD) diseñado ad hoc para el estudio. La recogida de datos de las HC en las dos etapas la realizaron profesionales de los propios centros (mayoritariamente cirujanos con la especialidad o en formación) específicamente entrenados. Los datos recogidos en los CRD fueron tratados con el programa CardiffTeleForm®.

Estrategias para el control de la calidad de los datos del estudio

Para asegurar la calidad de los datos, en el protocolo del estudio ONCORisc se incluyeron los siguientes controles:

- **Estandarización de la recogida de datos.** Se diseñaron CRD acompañados de un manual de definiciones operativas de las variables y el manual de instrucciones para la recogida de datos de los pacientes. El diseño de los CRD incorporaba cuatro formularios diferenciados, ordenados cronológicamente según la sucesión de las etapas en el procedimiento quirúrgico, con los datos relativos al preoperatorio, episodio clínico de la intervención y seguimiento a los 3 y 6 meses, respectivamente. Respecto al formato, el CRD reunía los cuatro formularios impresos en páginas de tres hojas autocopiables (para el equipo investigador, para la transcripción y para el investigador que realizaba la extracción de datos) que permitirían la realización de controles y reevaluaciones.
- **Actividades de formación.** Se llevaron a cabo reuniones informativas y de entreno con las personas que llevarían a cabo el trabajo de campo. Se entregó y explicó detalladamente el protocolo del estudio, el CRD y los manuales de definiciones e instrucciones. Se realizó una prueba piloto, seleccionando las HC que cumplían los criterios de inclusión, el mes previo al inicio del estudio. Durante el periodo de realización del estudio de campo, los investigadores mantuvieron el contacto con las personas que realizaban el trabajo de campo para resolver las dudas o problemas que podían surgir.
- **Transcripción automatizada de los datos.** Se diseñaron los CRD de forma que permitían la lectura óptica de los datos mediante el sistema CardiffTeleForm®. Este programa permite administrar un centro de captura masiva de datos en formato papel para convertirlos en datos digitales. Asimismo, incorpora reglas de validación automáticas, lo que agiliza y reduce el tiempo de entrada de datos, además de minimizar los errores en este proceso. Este sistema proporciona una base de datos con SPSS y la digitalización de los CRD en formato PDF. Además, genera un informe de errores realizando una validación interna de la calidad y la consistencia de la información introducida en la base de datos. Se establecieron criterios de validación de los datos y se compararon los informes de errores con los formularios para la detección de datos incumplidos o ilegibles e incoherencias.

- **Verificación de los datos.** Siempre que se creyó oportuno, se contactó con los revisores para la verificación de los errores con la HC. Se estableció un circuito de recuperación de las HC que serían revisadas de nuevo, devolviendo las validaciones que se incorporaron a la base de datos.
- Dadas las características del estudio multicéntrico y las limitaciones temporales y presupuestarias del mismo, que no permitían la realización del estudio mediante un único equipo de investigadores para los distintos centros, se creyó relevante la **realización de una auditoría.** Se auditaría la información del estudio a partir de la revisión de HC mediante una muestra aleatoria de los pacientes incluidos en el estudio ONCORisc, que corresponde al trabajo que aquí se presenta.

A pesar de la importancia de la calidad de los datos en los resultados de los estudios, pocos trabajos hacen referencia explícita a los métodos de control de calidad utilizados para minimizar al máximo los errores y para controlar la calidad de los datos obtenidos en el propio estudio^{15, 16}. No obstante, parece existir, en los últimos años, una tendencia al reconocimiento de la necesidad de controlar y evaluar la calidad de los datos de los estudios en nuestro contexto¹⁷⁻¹⁹.

En este informe se presentan los resultados de la auditoría realizada para validar la calidad de los datos obtenidos en el estudio ONCORisc mediante la identificación de los errores en la extracción de datos, entre la HC y la base de datos del estudio ONCORisc (BdO); y de los errores en la transcripción, entre los CRD y la BdO.

Objetivos

Objetivo general

Evaluar las coincidencias y discrepancias de los datos recogidos en la Base de Datos del estudio ONCORisc con los datos contenidos en la HC.

Objetivos específicos

- Estimar el porcentaje de discrepancias global, por etapa y variable.
- Analizar si existen diferencias en el porcentaje global de discrepancias entre las fases prospectiva y retrospectiva.
- Analizar el porcentaje de errores atribuibles al sistema de transcripción utilizado (CardiffTeleForm®).

Metodología

Muestra

En una primera instancia se estimó que era necesaria una muestra de 420 HC para poder estimar el porcentaje de discrepancias para el total, por región sanitaria y tipo de centro, con una precisión de entre el 1% y el 2,2%, esperando una tasa de reposición no superior al 10%. Aunque finalmente diferentes problemas logísticos y de coordinación con los centros seleccionados al azar limitaron el número de HC auditadas a 322, un número insuficiente para estimar el porcentaje de discrepancias por región sanitaria o tipo de centro con una precisión aceptable, sí fue posible estimarlas por el total, por etapa y por tipo de variable con un intervalo de confianza (IC) del 95%. De todos modos, la asignación muestral se realizó mediante un muestreo aleatorio estratificado por región sanitaria (RS Barcelona, RS Costa de Ponent, RS Maresme, RS Centre, RS Girona, RS Lleida y RS Tarragona) y nivel de complejidad del centro (aislado geográficamente/comarcal/red complementaria; hospital de referencia/alta tecnología/privado). En el Anexo 2 se describen el cálculo y la distribución de la muestra.

Variables

A partir de los cuestionarios elaborados para el estudio ONCORisc se identificó el siguiente subgrupo de variables clave que se deberían auditar:

- Variables demográficas y clínicas del paciente:
 - fecha de nacimiento (dd/mm/aaaa),
 - sexo (hombre/mujer),
 - tipo de cáncer,
 - estadiaje del cáncer (TNM clínico),
 - antecedentes de enfermedad pulmonar obstructiva crónica (EPOC) (sí/no),
 - antecedentes de hipertensión arterial (HTA) (sí/no),
 - antecedentes de diabetes (sí/no),
 - antecedentes neoplásicos (sí/no),
 - clasificación del estado físico preoperatorio según la American Society of Anesthesiology (escala ASA).

- Variables de proceso:
 - tipo de ingreso (urgente/programado),
 - fecha de ingreso (dd/mm/aaaa),
 - fecha de intervención (dd/mm/aaaa),
 - número de pruebas diagnósticas para el estadiaje,
 - radioterapia primaria (sí/no),
 - quimioterapia primaria (sí/no),
 - consumo de bolsas de hematíes (sí/no),
 - fecha de alta hospitalaria (dd/mm/aaaa).

- Variables de resultado:
 - complicaciones respiratorias (sí/no),
 - complicaciones cardiocirculatorias (sí/no),
 - complicaciones renales (sí/no),
 - complicaciones de la herida quirúrgica (sí/no)
 - complicaciones de la técnica quirúrgica (sí/no),
 - reintervención intrahospitalaria (sí/no),
 - muerte al alta hospitalaria (sí/no),
 - fecha de la muerte al alta hospitalaria (dd/mm/aaaa),
 - reintervención a los 3 meses (sí/no),
 - recidiva a los 3 meses (sí/no),
 - muerte a los 3 meses (sí/no),
 - reintervención a los 6 meses (sí/no),
 - recidiva a los 6 meses (sí/no),
 - muerte a los 6 meses (sí/no).

Recogida de datos

La auditoría fue realizada por personal propio de la AATRM, mediante la revisión de las HC de los pacientes. La recogida de datos de la auditoría se hizo entre los meses de abril y noviembre de 2007 mediante una aplicación en MS Access diseñada específicamente para esta tarea.

Análisis

Se analizaron las coincidencias y las discrepancias entre los datos recogidos en la auditoría y la BdO. Se definió coincidencia cuando aparecía el mismo valor en el dato recogido en la auditoría y la BdO. Se consideró que había

discrepancia cuando los datos de la auditoría no coincidían con la BdO. Los valores de las variables discrepantes fueron verificados con la imagen escaneada (IE) de los CRD. Las discrepancias se clasificaron en: 1) discrepancias audit/BdO cuando los valores de las variables de la BdO y de la IE coincidían y ambos diferían del valor de la auditoría, y 2) discrepancias CardiffTeleForm® cuando coincidían los valores de las variables de la auditoría y de la IE y diferían del valor de la BdO.

En las variables continuas (fecha de ingreso, fecha de intervención, fecha de alta y fecha de muerte), se estableció un margen razonable de respuesta para determinar la coincidencia: se consideraron datos coincidentes cuando las diferencias entre las fechas no superaban las 48 horas. Para la variable TNM (estadaje del cáncer), se consideró coincidencia cuando coincidían exactamente los tres valores que la componen.

Para resolver los objetivos del estudio se evaluó el número y el porcentaje de discrepancias detectadas en total, por etapa, sistema y variable, estimando también el IC del 95% de este porcentaje con el método binomial exacto. Por variable, además, se calculó el estadístico Kappa para el estudio de la concordancia más allá de lo esperable por azar entre las variables comparadas. Finalmente, para analizar las diferencias en el porcentaje global de discrepancias por etapas, se realizó el test de la χ^2 , fijando el nivel de significación al 5%.

Consideraciones éticas-confidencialidad

Por razones de confidencialidad, los números de las HC no se detallaron en documentos de presentación y solicitud de colaboración, en correos electrónicos, ni por vía telefónica, y fueron comunicadas in situ en las visitas a los centros. Se cumplieron los requisitos establecidos por cada hospital con relación a los circuitos de extracción y utilización de las HC y firma de documentos de confidencialidad. En uno de los centros se procedió a la “anonimización” de los documentos necesarios previo acceso del auditor.

Resultados

Fueron auditadas 322 HC, de las cuales 188 correspondieron a la etapa retrospectiva y 134 a la prospectiva. Se incluyeron 14 hospitales, de los 49 participantes en el estudio ONCORisc: 3 hospitales comarcales (n=27 casos), 5 hospitales de referencia (n=135), 3 hospitales de alta tecnología (n=139) y 3 centros privados (n=21).

Del total de variables examinadas (9.982), el porcentaje de coincidencias entre la BdO y la auditoría (BdAudit) fue del 88,8% (8.862). De las 1.120 variables en las que se produjeron discrepancias, 1.044 (93,2%; n=1.120) correspondieron a discrepancias entre los datos de la BdAudit y los datos la BdO y 76 (6,8%; n=1.120) a la lectura del TeleForm®. Estas últimas representan el 0,7% (n=9.982) del total de valores comparados (Tabla 4).

Tabla 4. Resultados globales			
	N	%	IC 95%
Discrepancias BdAudit/BdO	1.044	10,5	[9,90-11,10]
Discrepancias TeleForm®	76	0,7	[0,65-0,75]
Coincidencias	8.862	88,8	[88,2-89,4]
TOTAL	9.982	100,0	

La Tabla 5 muestra los porcentajes de coincidencias y discrepancias de las 31 variables incluidas, entre las que 19 presentaron un número de coincidencias superior al 90%, 6 variables entre el 80% y el 90% y otras 6 variables por debajo del 80%. La mortalidad al alta mostró el 100% de coincidencias, seguida del sexo (99,1%) y la fecha de muerte al alta (98,8%). Las variables “tipo de cáncer”, “reintervención intrahospitalaria”, “radioterapia y quimioterapia primarias” y “antecedentes de diabetes” obtuvieron coincidencias superiores al 95%.

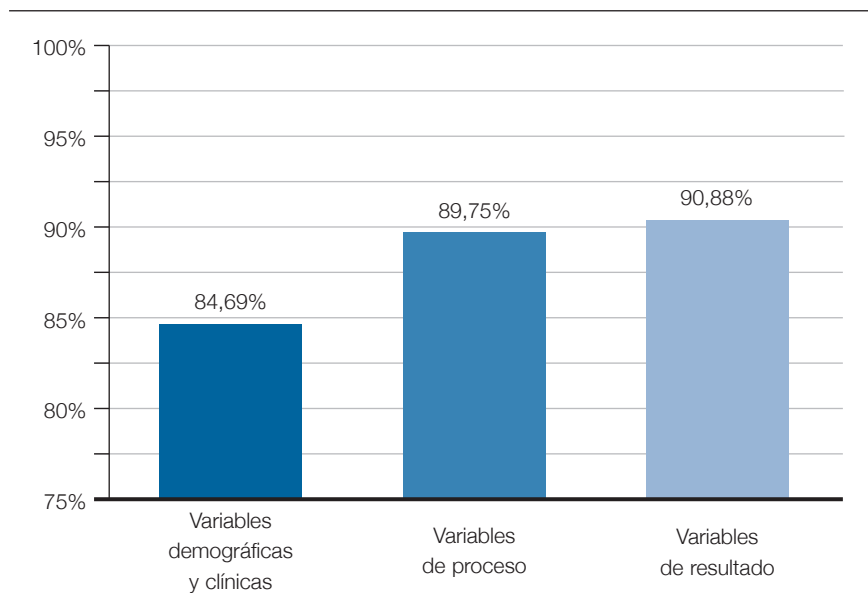
El porcentaje más bajo de coincidencias (51,5%) fue para el TNM, seguido del número de pruebas, la recidiva a los 3 y 6 meses, los antecedentes neoplásicos y el ASA, con valores situados entre el 70% y el 80%.

Los datos concernientes a las fechas mostraron valores próximos al 90% de coincidencias, a excepción de la fecha de muerte (98,7%), siendo la fecha de alta la de menor número de coincidencias.

En relación con las variables de resultado en las que se realizaron evaluaciones en el tiempo (mortalidad, recidivas y reintervenciones), se apreció una disminución en el porcentaje de coincidencias conforme aumentaba el periodo de seguimiento a los 3 y 6 meses. En el caso de la mortalidad, se pasó del 100% de coincidencias respecto a la mortalidad intrahospitalaria, al 93,4% a los 3 meses y al 90,7% a los 6 meses. En cuanto a las reintervenciones, las intrahospitalarias presentaron un 98,4% de coincidencias, pero se redujeron a un 90,7% a los 3 meses y a un 89,7% a los 6 meses. La variación fue mínima en el caso de las recidivas, 75,8% y 74,2% a los 3 y 6 meses, respectivamente.

En general, las variables de resultado son las que tuvieron mayor número de coincidencias (90,8%), seguidas de las variables de proceso (89,7%), y las demográficas y clínicas en menor medida (84,6%) (Figura 1).

Figura 1. Porcentaje global de coincidencias según categoría de las variables



Los resultados del estadístico de Kappa mostraron valores superiores a 0,81 en 8 variables, lo que indica muy buena y casi perfecta concordancia entre observadores, según los criterios de interpretación propuestos por Landis y Koch²⁰. Entre ellas, la concordancia fue perfecta en la variable mortalidad al alta. Las variables correspondientes al sexo, los antecedentes de hipertensión arterial, diabetes y EPOC, radioterapia, quimioterapia y reintervención intrahospitalaria pertenecen al grupo de las mejor puntuadas. La concordancia fue moderada en la variable de seguimiento de la recidiva a los 6

meses y débil en la variable de la recidiva a los 3 meses. En las 15 variables restantes, de las 25 valoradas, la concordancia fue buena, obteniéndose puntuaciones entre 0,61 y 0,80. El estadístico de Kappa no fue calculado para las variables compuestas TNM y fechas (n=6 variables).

Entre las discrepancias, cabe destacar que en 9 variables no se produjo ningún error de lectura, y en otras 11 el porcentaje de error era inferior al 0,7% (Tabla 5).

Las Tablas 6 y 7 muestran el número de discrepancias y coincidencias por etapa, retrospectiva y prospectiva. Los porcentajes totales de coincidencias, tanto de la etapa retrospectiva como de la prospectiva del estudio, fueron similares, del 88,3% y 89,4%, respectivamente. Por variables, destaca la reducción del número de discrepancias en la etapa prospectiva respecto a los resultados obtenidos en la etapa retrospectiva en las variables correspondientes a fecha de nacimiento, antecedentes neoplásicos, recidivas a los 3 y 6 meses y EPOC. Entre las variables que aumentaron el número de discrepancias destacan los resultados del ASA, fecha de alta, mortalidad y reintervención a los 3 meses.

Tabla 5. Número, porcentaje e intervalo de confianza de coincidencias, concordancia y tipo de discrepancias en cada variable estudiada según categoría de las variables (N=322)											
	Coincidencias						Discrepancias				
	n	%	IC 95%	K	Discrepancia BdAudit / Bdo		Discrepancia TeleForm®				
					n	%	n	n	%		
Variables demográficas y clínicas	Sexo	319	99,1	0,97-1,00	0,98	3	0,9	0	0	0,0	
	Tipo de cáncer	317	98,4	0,96-0,99	0,97	5	1,6	0	0	0,0	
	Diabetes	307	95,3	0,92-0,97	0,91	12	3,7	3	3	0,9	
	HTA	304	94,4	0,91-0,97	0,92	17	5,3	1	1	0,3	
	EPOC	298	92,5	0,89-0,95	0,76	16	5,0	8	8	2,5	
	Fecha de nacimiento	282	87,6	0,83-0,91	...	34	10,6	6	6	1,9	
	Antecedentes neoplásicos	234	72,7	0,67-0,77	0,70	85	26,4	3	3	0,9	
	ASA	226	70,2	0,65-0,75	0,79	95	29,5	1	1	0,3	
	TNM	166	51,5	0,46-0,57	...	56	48,4	0	0	0,0	
	Radioterapia	315	97,8	0,96-0,99	0,94	6	1,9	1	1	0,3	
	Quimioterapia	311	96,6	0,94-0,98	0,89	8	2,5	3	3	0,9	
	Fecha de intervención	298	92,5	0,89-0,95	...	18	5,6	6	6	1,9	
	Tipo de ingreso	285	88,5	0,85-0,92	0,68	37	11,5	0	0	0,0	
Fecha de ingreso	285	88,5	0,85-0,92	...	30	9,3	7	7	2,2		
Transfusión de hematies	283	87,9	0,84-0,91	0,77	38	11,8	1	1	0,3		
Fecha de alta	279	86,6	0,82-0,90	...	34	10,6	9	9	2,8		
Número de pruebas	256	79,5	0,75-0,84	0,61	58	18,0	8	8	2,5		

Continúa en la página siguiente

Tabla 6. Número y porcentaje de discrepancias en cada variable estudiada según etapa y categoría de las variables (N=322)

	Retrospectiva (n=188)		Prospectiva (n=134)	
	n	%	n	%
Sexo	2	1,06	1	0,75
Tipo de cáncer	1	0,53	4	2,99
Diabetes	9	4,79	6	4,48
HTA	11	5,85	7	5,22
EPOC	18	9,57	6	4,48
Fecha de nacimiento	30	15,96	10	7,46
Antecedentes neoplásicos	71	37,77	17	12,69
ASA	48	25,53	48	35,82
TNM	91	48,40	65	48,51
Radioterapia	4	2,13	3	2,24
Quimioterapia	6	3,19	5	3,73
Fecha de intervención	17	9,04	7	5,22
Tipo de ingreso	19	10,11	18	13,43
Fecha de ingreso	19	10,11	18	13,43
Transfusión de hemáties	21	11,17	18	13,43
Fecha de alta	17	9,04	26	19,40
Número de pruebas	41	21,81	25	18,66
Mortalidad alta	0	0,00	0	0,00
Fecha de muerte	2	1,06	2	1,49
Reintervención intrahospitalaria	4	2,13	1	0,75
Complicaciones respiratorias	10	5,32	10	7,46
Mortalidad 3 meses	7	3,72	14	10,45
Complicaciones locales	15	7,98	10	7,46
Complicaciones de la técnica	13	6,91	12	8,96
Complicaciones renales	16	8,51	11	8,21
Complicaciones cardiológicas	18	9,57	11	8,21
Reintervención 3 meses	12	6,38	18	13,43
Mortalidad 6 meses	15	7,98	16	11,94
Reintervención 6 meses	18	9,57	15	11,19
Recidiva 6 meses	63	33,51	15	11,19
Recidiva 3 meses	63	33,51	20	14,93

ASA: escala de la American Society of Anesthesiology; EPOC: enfermedad pulmonar obstructiva crónica; HTA: hipertensión arterial; TNM: estadiaje del tumor.

Tabla 7. Coincidencias y discrepancias según etapa

	Retrospectiva (n=188)		Prospectiva (n=134)		Total
	n	%	n	%	N
Discrepancia	681	11,7	439	10,6	1.120
Coincidencia	5.147	88,3	3.715	89,4	8.862
Total variables	5.828	100,0	4.154	100,0	9.982

χ^2 3,037; p valor = 0,081

Discusión

El presente estudio es una de las primeras auditorías sobre la calidad de los datos propios de un estudio de evaluación, realizada en Cataluña, en el campo de la oncología digestiva. Si bien en registros de cáncer²¹, cirugía²² y otras especialidades^{7,23} es frecuente realizar auditorías para comprobar los datos que contienen, no es habitual el detalle de los controles y medidas de calidad en el ámbito de los estudios de investigación. La realización de una auditoría para validar la calidad de los datos del estudio ONCORisc permitió analizar la extracción de datos de las HC y así identificar el tipo de datos más conflictivos, los de más difícil obtención o los que podían propiciar interpretaciones individuales y sus posibles causas.

Los datos relacionados con los resultados de la cirugía fueron los que proporcionaron mayor número de coincidencias, como la mortalidad al alta, la fecha de la muerte y las reintervenciones intrahospitalarias. También merece atención el porcentaje de coincidencias superior al 90% de los datos relativos a las complicaciones cuya obtención parecía más difícil y más propensa a variadas interpretaciones. Le siguieron las características del proceso asistencial y las demográficas y clínicas del paciente con menor porcentaje de coincidencias, en contraste con otros estudios^{3,7} para los cuales al auditar la extracción de datos de la HC, los datos demográficos fueron los de mayor coincidencia y menos problemáticos. Aunque algunos datos demográficos y del paciente, como son el sexo, el tipo de cáncer y los tratamientos adyuvantes, fueron los que mejor puntuaron, cabe observar que este grupo de variables incluía tres de las que obtuvieron mayor porcentaje de discrepancias: estadiaje TNM clínico, escala de ASA y antecedentes neoplásicos.

El estadiaje TNM fue la variable que ofreció el menor porcentaje de coincidencias (51,5%). Presentaba un elevado número de datos ausentes, como se había constatado previamente en el estudio ONCORisc y, cuando se encontraba, la fuente de obtención era variable. En los centros que utilizaban dicho estadiaje, se hallaba habitualmente en el curso clínico, en el informe de anatomía patológica o en el de alta. También se obtenía este dato de las anamnesis realizadas por el servicio de oncología previas a la intervención y en los informes de las pruebas ecográficas endorrectales cuando se realizaron. Diferentes ubicaciones en la HC y la interpretación por diferentes clínicos, o el hecho de no estar transcrita en la HC, pudieron ser las causas de identificar distintos valores y omisiones en la obtención. El estadiaje TNM se compone de tres valores que describen tres características del tumor (T=extensión del tumor primario, N=afectación de ganglios linfáticos

con metástasis, M=presencia o ausencia de metástasis a distancia), pudiendo discrepar cada una de ellas por separado, lo que aumenta la probabilidad de cometer errores de interpretación y de transcripción. Asimismo, cabe destacar que su evaluación se realizó con el máximo nivel de exigencia, comparando uno por uno los tres valores que lo definen.

Con respecto a la variable ASA, un elevado número de discrepancias fue ocasionado por la falta de datos detectada en la auditoría, en uno de los centros donde la proporción de casos era más elevada. En numerosos casos, los antecedentes neoplásicos, sobre todo en la etapa retrospectiva, no fueron registrados en el estudio ONCORisc, por lo que causaron discrepancia. Este dato no siempre se encontraba descrito explícitamente en la recopilación de antecedentes del paciente, lo que pudo condicionar la interpretación y el registro del mismo.

Otras variables en las que el número de diferencias fue mayor a lo que cabía esperar fueron las relacionadas con las fechas, en coincidencia con los resultados de otro estudio⁷. A pesar de que se estableció un margen de 48 horas en las variables continuas (fecha de ingreso, fecha de intervención y fecha de alta), los datos relacionados con las fechas, también en nuestro estudio, provocaron menos coincidencias de las que nos parecía razonable esperar para este tipo de datos. Aunque dichas variables no suelen presentar problemas de obtención ni de ausencia de datos, en general, las fechas parecen ocasionar errores de transcripción.

Algunos pacientes realizaron un ingreso previo al de la intervención para la realización de pruebas para completar el diagnóstico, siendo dados de alta y reingresando en un breve plazo de tiempo. Administrativamente, se trataba de dos ingresos independientes, pero en algunos casos en el curso clínico se identificaban como el mismo episodio asistencial, hecho que pudo ser otra causa de discrepancia.

La localización de la información relativa al seguimiento de los pacientes después del episodio de intervención era diferente de unos centros a otros. En general, los datos de seguimiento presentaban más ausencias de información y más dificultades de localización. En los centros privados no se realizaban controles posteriores a la intervención en el mismo centro y en algún caso la información se registraba en una HC independiente, única para la consulta externa. Dichas incidencias pueden estar relacionadas con la variación sufrida en las variables de mortalidad, recidivas y reintervenciones, respecto a la disminución de coincidencias en correspondencia al periodo de seguimiento.

En general, según la escala de valoración cualitativa del estadístico de Kappa mayormente utilizada²⁰, la concordancia de las variables del estudio es buena. A pesar de ello, se han considerado aceptables valores entre 0,60 y 0,80, y no aceptables los valores inferiores.

La concordancia entre observadores fue muy buena en la mortalidad y las reintervenciones intrahospitalarias, las variables demográficas y clínicas, el sexo, el tipo de cáncer, los antecedentes de diabetes y HTA, y la variable de proceso radioterapia. Algunas de las variables que obtuvieron mayor porcentaje de coincidencias no puntuaron los máximos valores del índice de Kappa. Es el caso de las de EPOC y las de resultado de complicaciones respiratorias, renales, cardiológicas, de la técnica y de la herida quirúrgica, mortalidad a los 3 y 6 meses y reintervención a los 3 meses, para las cuales el elevado porcentaje de coincidencias obtenido fue debido no sólo al acuerdo entre observadores, sino también al azar ($\kappa=0,70-0,80$).

También se han podido comprobar los errores producidos por el sistema CardiffTeleForm® para la transcripción de los datos. El valor de las discrepancias encontradas en la presente auditoría es superior al de otros estudios que obtuvieron valores entre 0,041% y 0,4%. Dichos estudios fueron específicamente diseñados para la comprobación del sistema, concentrando los esfuerzos al máximo en la validación previa al escaneado de los datos^{12,13}. A pesar de implicar una revisión detallada de los cuestionarios previa al escaneado o durante éste, se confirma la utilidad de este sistema de entrada y gestión de datos para la reducción de errores de transcripción. Además, la familiarización con el sistema, el entreno y el cuidado en el registro de los CRD aumenta la efectividad de la transcripción.

La diferencia en el porcentaje global de coincidencias/discrepancias de los datos obtenidos en la revisión de las HC por etapas retrospectiva y prospectiva del estudio ONCORisc, no ofreció resultados estadísticamente significativos, aunque se apreciaron diferencias considerables en algunas variables. Los estudios retrospectivos se basan exclusivamente en el contenido de la HC y, por lo tanto, la fuente de información es la misma para el investigador del centro y para el auditor, hecho que explicaría resultados mejores en dicha etapa. Yoon *et al.*, a pesar de coincidir con el resultado global por etapas, identifican un mayor número de coincidencias en la etapa retrospectiva, lo cual explican por estar presentes los datos en las HC⁷. Un menor número de discrepancias en la etapa prospectiva respecto a los valores obtenidos en la etapa retrospectiva podría ser indicativo de la efectividad de las medidas de control llevadas a cabo a lo largo del estudio.

El difícil acceso a los datos podía ser una causa común de discrepancias. Se constató que las HC de los pacientes oncológicos contienen un gran volumen de información. Tal y como ya se había considerado en las limitaciones y conclusiones del estudio ONCORisc, se verificó que en la mayoría de los casos se daba un importante desorden de los registros. En general, la falta de estandarización de las historias clínicas en algunos casos, diferentes modelos de estandarización en otros y la variabilidad de contenidos definidos en otros contextos^{2,5} fueron un problema recurrente en la mayoría de los centros incluidos en el presente estudio.

Además, aunque se recomienda emplear un solo equipo de investigadores para la extracción de datos en los diferentes centros², el hecho de tratarse de un estudio de gran envergadura, con limitaciones temporales y presupuestarias, supuso la colaboración de diferentes profesionales (una o más personas por centro) para la recogida de datos. La participación de numerosos investigadores en las tareas de extracción de datos, en el estudio ONCORisc, sin duda, pudo introducir variabilidad en la forma de recogerlos, convirtiéndose en una amenaza a la calidad de la información recogida, a pesar de su conocimiento de la documentación y de los requerimientos de cada institución.

Entre las limitaciones de la auditoría, razones logísticas y el hecho de que algunos hospitales declinaran la participación ocasionaron la reducción de los casos que se auditaron. Para este estudio se calculó cuál era el número de historias que debían ser auditadas para poder realizar una estimación con una precisión del 1% del porcentaje de errores imputables a la revisión de las HC (Anexo 2). Se consideró, a priori, que un 10% de las HC seleccionadas para ser auditadas deberían ser repuestas por otras, debido a pérdidas de casos, pero finalmente este número fue superior. El motivo principal fue la falta de acuerdo con determinados centros (especialmente comarcales) para acceder de nuevo a consultar las HC de sus pacientes, basando su decisión en la Ley de protección de datos. Este hecho tuvo un impacto sobre la representatividad de la muestra, dificultó cualquier análisis estratificado por región sanitaria y/o tipo de centro, aunque no sobre el global. En este sentido, la distribución de pacientes por sexo, edad, tipo de cáncer y riesgo quirúrgico no fue distinta en la muestra auditada respecto a la del estudio.

La auditoría de datos obtenidos prospectivamente en el estudio ONCORisc, realizada a posteriori, supuso otra limitación, dado que la obtención de datos una vez finalizado el estudio no permitía el acceso a las mismas fuentes de información. Asimismo, el resultado podría reflejar en mayor medida la calidad de las HC y no la calidad de la recogida de datos.

Conclusiones

La medida de la calidad de los datos es necesaria, ya que nos proporciona una idea sobre la calidad y la validez de los resultados de la investigación, en particular cuando se utilizan datos de HC o información retrospectiva. De ahí la necesidad de incorporar auditorías para conocer la calidad de los datos obtenidos. La realización de auditorías al inicio y durante el estudio permite establecer medidas correctoras. Es ineludible y necesaria la estructuración y estandarización de las HC, que supondrá claros beneficios tanto para el paciente como para los profesionales de la salud, siendo además una de las principales fuentes de datos para la investigación y la evaluación de resultados. En este mismo sentido, la informatización de las HC representa un importante avance y una oportunidad para alcanzar dicho objetivo: la estandarización. Además, la utilización de métodos informatizados de recogida y transcripción de datos minimiza los errores y aumenta la calidad de los datos. Se confirma la efectividad de los controles de calidad de los datos y la necesidad de poner la máxima atención en las estrategias de recogida de datos para maximizar su calidad.

Bibliografía

1. Argimon Pallàs J, Jiménez Villa J. Recogida de datos. Métodos de investigación aplicados a la atención primaria de salud. En: Argimon Pallàs J, Jiménez Villa J, (Ed.) Barcelona: Doyma; 1991.
2. Jansen AC, van Aalst-Cohen ES, Hutten BA, Buller HR, Kastelein JJ, Prins MH. Guidelines were developed for data collection from medical records for use in retrospective analyses. *J Clin Epidemiol.* 2005; 58(3):269-74.
3. Pan L, Fergusson D, Schweitzer I, Hebert PC. Ensuring high accuracy of data abstracted from patient charts: the use of a standardized medical record as a training tool. *J Clin Epidemiol.* 2005;58(9):918-23.
4. Gearing RE, Mian IA, Barber J, Ickowicz A. A methodology for conducting retrospective chart review research in child and adolescent psychiatry. *J Can Acad Child Adolesc Psychiatry.* 2006;15(3):126-34.
5. Mann R, Williams J. Standards in medical record keeping. *Clin Med.* 2003;3(4):329-32.
6. Hess DR. Retrospective studies and chart reviews. *Respir Care.* 2004;49(10):1171-4.
7. Yoon SS, George MG, Myers S, Lux LJ, Wilson D, Heinrich J, et al. Analysis of data-collection methods for an acute stroke care registry. *Am J Prev Med.* 2006;31(6 Suppl 2):S196-S201.
8. Cassidy LD, Marsh GM, Holleran MK, Ruhl LS. Methodology to improve data quality from chart review in the managed care setting. *Am J Manag Care.* 2002;8(9):787-93.
9. Allison JJ, Wall TC, Spettell CM, Calhoun J, Fargason CA Jr., Kobylinski RW, et al. The art and science of chart review. *Jt Comm J Qual Improv.* 2000;26(3):115-36.
10. Reisch LM, Fosse JS, Beverly K, Yu O, Barlow WE, Harris EL, et al. Training, quality assurance, and assessment of medical record abstraction in a multisite study. *Am J Epidemiol.* 2003;157(6):546-51.
11. Bolúmar Montrull F. Planificación de una investigación. Estudios descriptivos. En: Doménech J, Ed. Diseño de estudios sanitarios. Barcelona: Signo; 2002.
12. Jinks C, Jordan K, Croft P. Evaluation of a computer-assisted data entry procedure (including Teleform) for large-scale mailed surveys. *Comput Biol Med.* 2003;33(5):425-37.
13. Quan KH, Vigano A, Fainsinger RL. Evaluation of a data collection tool (Teleform) for palliative care research. *J Palliat Med.* 2003; 6(3):401-8.

14. Avaluació del procediment terapèutic i els seus resultats en oncologia. Protocol d'estudi [document d'ús intern]. Barcelona: Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM). Servei Català de la Salut. Departament de Salut. Generalitat de Catalunya; 2003.
15. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. 1991. *Qual Saf Health Care*. 2004;13(2):145-51.
16. Thomas EJ, Orav EJ, Brennan TA. Hospital ownership and preventable adverse events. *J Gen Intern Med*. 2000;15(4):211-9.
17. Ribera A, Ferreira-González I, Cascant P, Pons JM, Permanyer-Miralda G. Evaluación de la mortalidad hospitalaria ajustada al riesgo de la cirugía coronaria en la sanidad pública catalana. Influencia del tipo de gestión del centro (estudio ARCA). *Rev Esp Cardiol*. 2006; 59(5):431-40.
18. Abilleira S, Ribera A, Sánchez E, Roquer J, Duarte E, Tresserras R, et al. La atención hospitalaria en el paciente con ictus en Cataluña. Resultados del «Primer Audit Clínic de l'Ictus. Catalunya, 2005/2006». *Gac Sanit*. 2008;22:565-73.
19. Ferreira-González I, Permanyer-Miralda G, Marrugat J, Heras M, Cunat J, Civeira E, et al. MASCARA (Manejo del Síndrome Coronario Agudo. Registro Actualizado). Resultados globales. *Rev Esp Cardiol*. 2008;61(8):803-16.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
21. Kelly M, Lamah M. Evaluating the accuracy of data entry in a regional colorectal cancer database: implications for national audit. *Colorectal Dis*. 2007;9(4):337-9.
22. Milburn JA, Driver CP, Youngson GG, King PM, Macaulay E, Krukowski ZH. The accuracy of clinical data: a comparison between central and local data collection. *Surgeon*. 2007;5(5):275-8.
23. Mack MJ, Herbert M, Prince S, Dewey TM, Magee MJ, Edgerton JR. Does reporting of coronary artery bypass grafting from administrative databases accurately reflect actual clinical outcomes? *J Thorac Cardiovasc Surg*. 2005;129(6):1309-17.

Anexos

Anexo 1. Variables de estudio

a) Variables referentes al enfermo:

- Datos identificativos: nombre y apellidos, número CIP (código de identificación personal), número de historia clínica, teléfono, hospital.
- Datos sociodemográficos y antropométricos: edad, sexo, peso, talla.

a.1) Datos relacionados con el estado basal y preoperatorios:

- Antecedentes: intervalo diagnóstico-tratamiento.
- Tipo de cáncer: estadiaje (TNM clínico), histología, marcadores tumorales (CEA [antígeno carcinoembrionario] en el caso del cáncer colorrectal), pruebas complementarias (radiografía de tórax, electrocardiograma, fibroscopia).
- Comorbilidades (trastornos crónicos), signos y síntomas específicos, y tratamientos médicos.
- Tratamientos primarios: radioterapia y quimioterapia (tipo, dosis, tandas, número de sesiones, fechas de inicio y final del tratamiento).
- Estado funcional (escala de Karnofsky y Glasgow Coma Score).
- Riesgo quirúrgico: ASA (clasificación del estado físico preoperatorio según la American Society of Anesthesiology).
- Analítica (hemoglobina, leucocitos, plaquetas, urea, creatinina, sodio, potasio, albúmina, proteínas totales, tiempos de protrombina).

a.2) Datos perioperatorios:

- Fecha de la intervención.
- Intención de la cirugía (curativa y/o paliativa) después de la intervención y justificación.
- Tipo de cirugía (limpia, limpia-contaminada, contaminada y sucia).
- Localización anatómica del tumor.
- Técnica quirúrgica.
- Bolsas de hematíes y de plasma.
- Tiempo quirúrgico.
- Profilaxis antibiótica y antitrombótica.
- Complicaciones intraoperatorias generales y las debidas al acto quirúrgico.
- Complicaciones postoperatorias inmediatas.
- Muerte.

a.3) Datos postoperatorios:

- Complicaciones postoperatorias.
- Recidiva.
- Muerte.
- Reintervenciones.
- Estadaje pTN, tipo histológico del cáncer, número de ganglios resecados, analizados y resecados positivos.
- Tratamiento adyuvante: farmacológico (no citostáticos), quimioterapia y radioterapia (tipo, dosis, tandas, número de sesiones, fechas de inicio y final del tratamiento), y sus complicaciones.
- Estado funcional (escala de Karnofsky).
- Salud percibida por el paciente (instrumento específico de calidad de vida relacionada con la salud para enfermos oncológicos).
- Participación en algún ensayo clínico (farmacológico, quirúrgico, etc.).

Se consideran variables de resultado las siguientes:

- Muerte (en las diferentes evaluaciones o momentos de interés/global o relacionada con el cáncer); es la variable de resultado principal.
- Supervivencia (en las diferentes evaluaciones o momentos de interés/global, libre de enfermedad).
- Recidiva del cáncer (en las diferentes evaluaciones o momentos de interés).
- Complicaciones (las intraoperatorias y las postoperatorias en las diferentes evaluaciones o momentos de interés).
- Estado funcional.
- Salud percibida.

a.4) Otros datos (administrativos):

- Circunstancia al ingreso (programado, urgente, traslado de otro hospital de agudos).
- Servicio de ingreso.
- Ingreso en un servicio/sala de críticos o semicríticos.
- Estancia hospitalaria.
- Circunstancia al alta (domicilio, traslado a un hospital de agudos, traslado a un centro o unidad de cuidados paliativos, traslado a un centro sociosanitario).
- Códigos diagnósticos y de procedimientos de la CIM-9 (Clasificación Internacional de Enfermedades, 9ª rev.) del informe de alta (se pedirá copia del informe de alta).
- Reingresos.
- Fecha de cumplimentación del formulario.
- Cirujano/médico responsable del enfermo.

b) Variables referentes al cirujano y al equipo:

- Características demográficas, académicas y profesionales: edad, sexo, año de finalización de la carrera y de la especialidad (cirugía general u otros), formación complementaria, tipo de práctica habitual (oncológica u otros tipos), ámbito/centros de práctica habitual (sector de práctica, atención hospitalaria y extrahospitalaria, etc.), entre otros.
- Responsabilidad: jefe de servicio, adjunto, residente, asistente, etc.
- Actividad realizada: volumen de intervenciones (semestral o anual) por cada tipo de cáncer, etc.

c) Variables referentes al centro:

- Sector: público o privado (dependencia patrimonial y funcional).
- Nivel de complejidad: alta tecnología, referencia y comarcal.
- Docencia de pregrado y postgrado (hospitales universitarios, unidades docentes MIR, etc.).
- Volumen de actividad del centro.
- Prácticas organizativas o de gestión (existencia de comité de tumores, registro de tumores, visita preoperatoria del anestesiólogo, equipos multidisciplinares/unidades funcionales, sesiones conjuntas, etc.).
- Otras características de habilidad o pericia de la organización (uso de guías de práctica clínica, vías clínicas *clinical pathways*, etc.).
- Instalaciones, equipos y recursos disponibles.
- Otras características estructurales: número de camas, número de camas de cuidados intensivos/camas para enfermos críticos o semicríticos, número de cirujanos, número de cirujanos por equipo, etc.

Anexo 2. Cálculo de la muestra

Tamaño muestral

El número de pacientes reclutados y su dispersión en el territorio no permitía auditarlos a todos, a pesar de la reducción de variables a auditar, por lo que fue necesaria la selección de una muestra representativa de HC. El objetivo del estudio era estimar el porcentaje de errores de la base de datos a partir del porcentaje observado en una muestra de tamaño n . Asimismo, se pretendía que el error cometido en la estimación del porcentaje global de discrepancias fuera inferior al 1% con una probabilidad 0,95. Se puede suponer que la frecuencia f del número de errores en una muestra de n variables es aproximadamente normal $N(np, \sqrt{npq})$ siendo $p=P/100$, ya que $n>30$. Por lo tanto, el porcentaje de errores ($100 \cdot f/n$) es de, aproximadamente, $N(P, 100 \cdot \sqrt{pq/n})$. Dada la falta de información sobre P , se estudió el caso de máxima incertidumbre, es decir, cuando $pq=p(1-p)$ es máximo ($p=q=0,5$).

$$N = (1,96 \times 100)^2 \times p \times q = (1,96 \times 100)^2 \times 0,5 \times 0,5 = 9.604$$

Teniendo en cuenta que para cada paciente se examinaron 31 variables, el número de historias a revisar fue de $9.604/31 \sim 309,8$. Cabía la posibilidad de que alguna historia clínica no estuviera disponible en el momento de la auditoría y, por lo tanto, se definió una proporción estimada de reposiciones necesarias que quedó fijada en un 10% de forma arbitraria.

$$n = (9.604 \times 100)/90 = 10.671,11 \quad 10.671,11/51 = 209,24 \sim 210$$

Asignación muestral

La asignación de los pacientes se hizo mediante un muestreo aleatorio estratificando por región y nivel del centro.

Región del centro:

- RS Barcelona.
- RS Costa de Ponent, Maresme y Centre.
- RS Girona, Lleida y Tarragona.

Nivel del centro:

- Aislado geográficamente/Comarcal/Red complementaria.
- Hospital de referencia.
- Alta tecnología.
- Privado.

Pacientes por estrato en la población

Región Sanitaria		Nivel de hospital				Total
		Comarcal	Referencia	Alta tecnología	Privado	
Barcelona	Recuento	3	231	656	140	1.030
	% de región	0,3	22,4	63,7	13,6	100,0
	% de nivel de hospital	0,6	17,6	61,4	91,5	33,9
Ponent-Maresme-Centre	Recuento	260	686	412	13	1.371
	% de región	19,0	50,0	30,1	0,9	100,0
	% de nivel de hospital	51,2	52,4	38,6	8,5	45,1
Lleida-Girona-Tarragona	Recuento	245	392	0	0	637
	% de región	38,5	61,5	0,0	0,0	100,0
	% de nivel de hospital	48,2	29,9	0,0	0,0	21,0
Total	Recuento	508	1.309	1.068	153	3.038
	% de región	16,7	43,1	35,2	5,0	100,0
	% de nivel de hospital	100,0	100,0	100,0	100,0	100,0

Estratificación de la muestra (n=420)

Región Sanitaria		Nivel de hospital				Total
		Comarcal	Referencia	Alta tecnología	Privado	
Barcelona	Recuento	0	32	91	19	142
	% de región	0,3	22,4	63,7	13,6	100,0
	% de nivel de hospital	0,6	17,6	61,4	91,5	33,9
Ponent-Maresme-Centre	Recuento	36	95	57	2	189
	% de región	19,0	50,0	30,1	0,9	100,0
	% de nivel de hospital	51,2	52,4	38,6	8,5	45,1
Lleida-Girona-Tarragona	Recuento	34	54	0	0	88
	% de región	38,5	61,5	0,0	0,0	100,0
	% de nivel de hospital	48,2	29,9	0,0	0,0	21,0
Total	Recuento	70	181	148	21	420
	% de región	16,7	43,1	35,2	5,0	100,0
	% de nivel de hospital	100,0	100,0	100,0	100,0	100,0

Conocidos el número de historias por auditar en cada estrato, quedaba por definir cómo serían seleccionados los centros.

Centros por estrato en la población					
Región sanitaria	Nivel de hospital				
	Comarcal	Referencia	Alta tecnología	Privado	Total
Barcelona	2	3	3	3	11
Ponent-Maresme-Centre	10	11	2	0	23
Lleida-Girona-Tarragona	10	5	0	0	15
Total	22	19	5	3	49

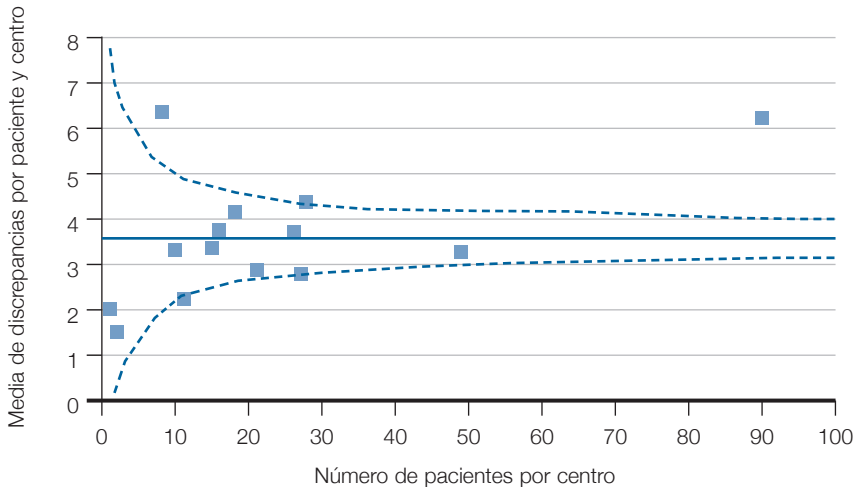
Para cada estrato se planteó realizar un muestreo aleatorio simple de los centros, ponderando por su volumen. Por eso se generaron 5 números aleatorios distribuidos uniformemente entre 0 y 100 por cada estrato, a los cuales se hizo corresponder con la frecuencia acumulada del listado de centros ordenados alfabéticamente. El primer número permitió identificar el centro escogido y ver si su volumen era suficiente para satisfacer las necesidades muestrales del estrato; en caso contrario, el siguiente número aleatorio permitió identificar un segundo centro y volver a verificar si se habían cumplido las necesidades muestrales del estrato, en caso contrario se continuó con el siguiente número, y así sucesivamente.

Si el número de pacientes en un centro superaba el número de pacientes necesarios en aquel estrato, entonces éstos fueron seleccionados al azar. Se ordenó el listado de pacientes por el número del día, el mes del ingreso y el número de registro, y mediante una tabla de números aleatorios fueron seleccionados aquellos pacientes a los que se les asoció un número más pequeño que $k=n/m$ hasta llegar al número necesario. En la fórmula anterior n es el número de pacientes necesarios en el estrato y m el número de pacientes incluidos por el centro.

Anexo 3. Análisis por centros

La estructura de los datos estudiados permitía suponer una cierta dependencia de los mismos. Entre las variables de un mismo paciente y entre las de pacientes de un mismo centro.

Para analizar este punto, se calculó la media de errores por paciente en cada uno de los centros, para luego calcular la media de errores entre centros, ponderada por el inverso de la varianza de la media de cada centro. De este modo fue posible identificar mediante un gráfico de embudo qué centros se escapaban de la norma. Dos centros tuvieron que ser excluidos de este análisis a causa del bajo número de pacientes auditados (<3).



Los centros que participaron en el análisis fueron 12 y el número de historias clínicas auditadas en cada centro fue proporcional al de pacientes incluidos en el estudio ONCORisc (ver Tamaño muestral en el Anexo 2). La media ponderada de discrepancias por paciente en cada centro fue de 3,6 (DE 2,2). Todos los centros exceptuando dos (centros 3 y 6) presentaron una media de discrepancias por paciente dentro de los valores esperados.

Hospital	n	Media	Desv. típ.
Centro 1	18	4,1	2,9
Centro 2	11	2,2	1,8
Centro 3	90	6,2	2,8
Centro 4	49	3,3	2,1
Centro 5	28	4,4	2,5
Centro 6	10	3,3	3,7
Centro 7	15	3,3	2,0
Centro 8	21	2,9	1,9
Centro 9*	2	1,5	0,7
Centro 10	8	6,4	3,1
Centro 11	27	2,7	2,2
Centro 12	16	3,7	2,7
Centro 13	26	3,7	1,4
Centro 14*	1	2,0	--
Global	322	3,6	2,2

* Centros excluidos del análisis



P.V.P.: 10 euros