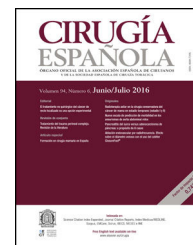




# CIRUGÍA ESPAÑOLA

[www.elsevier.es/cirugia](http://www.elsevier.es/cirugia)



## Original

# Protocolización de estudios clínicos multicéntricos en la era digital. ¿Tiene utilidad la centralización de datos por un data-manager?



Antonio Ríos <sup>a,b,\*</sup>, José Antonio Puñal-Rodríguez <sup>c</sup>, Pablo Moreno <sup>d</sup>, Enrique Mercader-Cidoncha <sup>e</sup>, Eduardo Ferrero-Herrero <sup>f</sup>, Manuel Durán <sup>g,h</sup>, Guadalupe Ruiz-Merino <sup>i</sup>, José Ruiz-Pardo <sup>j</sup>, José Manuel Rodríguez <sup>a,b</sup> y Pedro R Gutiérrez <sup>k,l</sup>

<sup>a</sup> Unidad de Cirugía Endocrina, Servicio de Cirugía General y de Aparato Digestivo, Instituto Murciano de Investigación Bio-Sanitaria (IMIB-Arrixaca), Hospital Clínico Universitario Virgen de la Arrixaca, Servicio Murciano de Salud, Murcia, España

<sup>b</sup> Departamento de Cirugía, Pediatría y Obstetricia y Ginecología, Universidad de Murcia, Murcia, España

<sup>c</sup> Servicio de Cirugía General y Aparato Digestivo, Complejo Universitario de Santiago de Compostela, Santiago de Compostela, La Coruña, España

<sup>d</sup> Cirugía Endocrina, Hospital Universitario de Bellvitge, Hospitalet de Llobregat, Barcelona, España

<sup>e</sup> Sección de Cirugía Endocrino-Metabólica, Hospital Universitario Gregorio Marañón, Instituto de Investigación Biosanitaria Gregorio Marañón, Madrid, España

<sup>f</sup> Servicio de Cirugía General, Aparato Digestivo y Trasplante de Órganos Abdominales, Hospital Universitario 12 de Octubre, Madrid, España

<sup>g</sup> Servicio de Cirugía General y del Aparato Digestivo, Hospital Universitario Rey Juan Carlos, Móstoles, Madrid, España

<sup>h</sup> Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Alcorcón, Madrid, España

<sup>i</sup> FFIS, Fundación para la Formación e Investigación Sanitarias de la Región de Murcia, Murcia, España

<sup>j</sup> Servicio de Cirugía General y del Aparato Digestivo, Hospital Torrecardenas, Almería, España

<sup>k</sup> Servicio de Urología, Complejo Hospitalario Universitario de Canarias (CHUC), La Laguna, Santa Cruz de Tenerife, España

<sup>l</sup> Departamento de Cirugía, Universidad de La Laguna (ULL), San Cristóbal de la Laguna, Santa Cruz de Tenerife, España

## INFORMACIÓN DEL ARTÍCULO

### Historia del artículo:

Recibido el 8 de noviembre de 2022

Aceptado el 21 de febrero de 2023

On-line el 17 de mayo de 2023

### Palabras clave:

Estudio multicéntrico clínico

Carcinoma infrecuente

Data-manager

Protocolización

Concordancia

Base de datos

## RESUMEN

**Introducción:** En los estudios multicéntricos la protocolización de los datos es una fase crítica que puede generar sesgos, sobre todo en estudios clínicos con presupuesto limitado. El objetivo es analizar la concordancia y la confiabilidad de los datos obtenidos en un estudio multicéntrico clínico entre la protocolización del centro de origen y la protocolización centralizada mediante un data-manager.

**Método:** Estudio clínico multicéntrico de prevalencia nacional sobre un carcinoma familiar infrecuente, realizándose una doble protocolización de los datos: a) en el centro de origen, y b) centralizada con un data-manager. La concordancia se analiza para el global de los datos y para los dos subgrupos del proyecto: a) grupo a estudio (carcinoma familiar; protocolizan 30 investigadores) y b) grupo control (carcinoma esporádico; protocolizan 4). Las diferencias interobservador se evalúan mediante el índice de Kappa de Cohen.

\* Autor para correspondencia.

Correos electrónicos: [arriors@um.es](mailto:arriors@um.es), [arriors4@gmail.com](mailto:arriors4@gmail.com) (A. Ríos).

<https://doi.org/10.1016/j.ciresp.2023.02.024>

0009-739X/© 2023 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

**Resultados:** Se incluyen 689 pacientes: 252 del grupo a estudio y 437 del grupo control. Respecto al análisis de concordancia del estadio tumoral, se han objetivado un 2,5% de discordancias, siendo alta la concordancia entre protocolizadores (Kappa = 0,931). Respecto a la valoración del riesgo de recidiva, las discordancias fueron del 7% de los casos, siendo alta la concordancia (Kappa = 0,819). Respecto a la clasificación ecográfica TIRADS, las discordancias son del 6,9% y la concordancia es alta (Kappa = 0,922). Se han detectado un 4,6% de errores de transcripción.

**Conclusiones:** En los estudios multicéntricos clínicos la protocolización centralizada de los datos por un *data-manager* parece presentar resultados similares a la protocolización directa en la base de datos en el centro de origen.

© 2023 AEC. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## Protocolization of multicenter clinical studies in the digital era. Is useful data centralization by a data-manager?

### A B S T R A C T

#### Keywords:

Multicenter clinical study  
Infrequent carcinoma  
Data-manager  
Protocolization  
Concordance  
Database

**Introduction:** In multicenter studies, the protocolization of data is a critical phase that can generate biases. The objective is to analyze the concordance and reliability of the data obtained in a clinical multicenter study between the protocolization in the center of origin and the centralized protocolization of the data by a data-manager.

**Methods:** National multicenter clinical study about an infrequent carcinoma. A double protocolization of the data is carried out: (i) center of origin; and (ii) centralized by a data manager. The concordance between the data is analyzed for the global data and for the two groups of the project: (i) study group (familiar carcinoma, 30 researchers protocolize); (ii) control group (sporadic carcinoma, 4 people protocolize). Interobserver variability is evaluated using Cohen's kappa coefficient.

**Results:** The study includes a total of 689 patients with carcinoma: 252 in the study group and 437 in the control group. Regarding the concordance analysis of the tumor stage, 2.5% of disagreements were observed and the concordance between people who protocolize was near perfect (Kappa = 0.931). Regarding the evaluation of the recurrence risk, disagreements occurred in 7% of the cases and the concordance was near perfect (Kappa = 0.819). Regarding the sonography evaluation (TIRADS), the disagreements were 6.9% and the concordance was near perfect (Kappa = 0.922). Also, 4.6% of transcription errors were detected.

**Conclusions:** In multicenter clinical studies, the centralized data protocolization by a data-manager seems to present similar results to the direct protocolization in the database in the center of origin.

© 2023 AEC. Published by Elsevier España, S.L.U. All rights reserved.

## Introducción

En las patologías poco frecuentes se está potenciando la realización de estudios multicéntricos clínicos que permitan conseguir series suficientemente grandes en relativamente poco tiempo para poder sacar conclusiones representativas<sup>1,2</sup>. Sin embargo, estos estudios también presentan inconvenientes, ya que los pacientes son manejados en centros diferentes y con protocolos no siempre iguales, lo cual va generando una heterogeneidad que puede condicionar los resultados. Además, el trabajo a distancia y por diferentes investigadores puede condicionar diferentes sesgos<sup>3</sup>.

Uno de los pasos donde se pueden producir sesgos es la protocolización de los datos<sup>3</sup>. Actualmente, la tendencia es a generar una base de datos online en la nube que permita a todos los centros la protocolización simultánea. Sin embargo,

esta protocolización conlleva la implicación de diferentes investigadores que puede condicionar una mayor heterogeneidad a la hora de cumplimentar la base. Por otro lado, la protocolización centralizada, generalmente por un *data-manager*, conlleva mayor gasto y no se ha mostrado una mayor utilidad<sup>3</sup>.

En la mayoría de ensayos clínicos, sobre todo farmacéuticos, se ha instaurado la recopilación de datos electrónicos directamente en una base mediante sistemas de captura de datos con el objetivo de conferir una mayor integridad de los datos y acortar el tiempo del estudio. Sin embargo, estos sistemas no son fáciles de implantar<sup>4</sup>, ya que precisa de unas instalaciones previas del centro investigador, un aprendizaje por parte de los usuarios, suele ser requerida la participación del paciente y conlleva unos costos económicos. Por ello, estos sistemas son poco coste-efectivos para proyectos multicéntricos clínicos<sup>1,4</sup> en los que resulta inasumible su coste y la

infraestructura que precisa. Además, la recopilación de datos electrónicos mediante sistema de captura no está exenta de riesgos, y sus ventajas pueden perderse sin una adecuada utilización<sup>5</sup>.

Lo que sí se ha generalizado en los proyectos multicéntricos clínicos, la mayoría sin financiación específica, es la utilización de bases de datos online en la nube para la protocolización, si bien no se ha testado qué sesgos o errores puede conllevar esta situación de recogida de datos en un sistema online frente a la protocolización por un *data-manager*<sup>6,7</sup>, sobre todo si se tiene en cuenta que un sistema de almacenamiento de datos confiable no garantiza la confiabilidad de la entrada de los mismos. Así, se han descrito errores en la transcripción de datos, de tal manera que las entradas cerca del final de la hoja de inserción de datos se correlacionan con un mayor número de errores. Sin embargo, aun en estos casos hay estudios que indican que el control por un administrador de datos o *data-manager* puede considerarse un costo evitable y, en todo caso, hay autores que sugieren el control electrónico para monitorear la confiabilidad de la entrada de datos<sup>6</sup>. Así, hay estudios de costes sobre el proceso de recopilación electrónica de datos en los ensayos clínicos que muestran una reducción de los costes de recopilación superiores al 50%, con un rango de ahorro según el proyecto entre el 49 y el 62%<sup>8</sup>.

El objetivo de este estudio es analizar en un estudio multicéntrico clínico la concordancia entre la cumplimentación del protocolo de recogida de datos en el centro de origen y la protocolización centralizada de los datos mediante un *data-manager*, para valorar la confiabilidad de los datos.

## Material y método

### Diseño del estudio

Estudio multicéntrico de prevalencia nacional avalado por la Asociación Española de Cirujanos, siendo este estudio el primer objetivo principal de dicho proyecto. La población a estudio la constituyen los pacientes con un carcinoma familiar de baja incidencia poblacional.

Se diseñó un protocolo de recogida de datos específico para esta patología, donde se recogen variables socio-personales, clínicas, terapéuticas y de seguimiento. Dicho protocolo conlleva un informe anexo que especifica los datos a recoger en cada una de las variables. Además, se incluye el contacto para consultar cualquier duda al cumplimentarlo.

### Recogida de datos

Los datos son recopilados en el centro de origen por cada uno de los investigadores que colabora en el proyecto, cumplimentando el protocolo del proyecto. Dicho protocolo puede ser digital o en papel, pero no vierte directamente a la base de datos del proyecto.

Se realiza una doble protocolización:

- 1) *Protocolización 1 - Centro de origen*. Se registran los datos tal y como se obtienen en el protocolo remitido desde el centro de origen.

- 2) *Protocolización 2 - Centralizada por un data-manager*. Se registran los datos de manera centralizada, por un mismo investigador *data-manager*, que puede consultar cualquier duda respecto a las variables analizadas con el centro de origen.

La concordancia entre la protocolización realizada con los datos del centro de origen y la protocolización centralizada mediante un *data-manager* se analiza en tres grupos diferentes:

- 1) *Grupo 1 - Global del estudio*. Se analiza la concordancia en el total de la muestra a estudio.
- 2) *Grupo 2 - Grupo a estudio del carcinoma familiar*. Se analiza la concordancia en el grupo del carcinoma familiar. En este grupo participan 30 investigadores en la recogida de los datos en el centro de origen, que corresponden a las diferentes unidades hospitalarias que participan en el proyecto.
- 3) *Grupo 3 - Grupo control del carcinoma esporádico*. Se analiza la concordancia en el grupo control que corresponde al carcinoma esporádico. En este grupo participan 4 investigadores en la protocolización en el centro de origen, ya que estos datos proceden solo de 4 unidades hospitalarias.

### Variables analizadas

Del análisis de concordancia se excluyen la mayoría de las variables socio-personales, clínicas y terapéuticas, donde es excepcional la presencia de una discordancia, excepto error de transcripción. Se analizan tres grupos de variables:

- 1) Clasificación pronóstica según el sistema TNM del American Joint Committee on Cancer. Dentro de este sistema, se analizan:
  - a) Tamaño (T).
  - b) Afectación ganglionar (N).
  - c) Afectación sistémica (M).
  - d) Estadío TNM.
- 2) Valoración del riesgo de recidiva según la ATA (American Thyroid Association), que diferencia tres grupos de riesgos:
  - a) Bajo riesgo.
  - b) Riesgo intermedio.
  - c) Alto riesgo.
- 3) Valoración ecográfica y clasificación en TIRADS (Thyroid Imaging Reporting and Data System), según la clasificación TIRADS coreana<sup>9</sup>. En ella se diferencian siete grupos de riesgo:
  - a) TIRADS 1.
  - b) TIRADS 2.
  - c) TIRADS 3.
  - d) TIRADS 4a.
  - e) TIRADS 4b.
  - f) TIRADS 4c.
  - g) TIRADS 5.

Por último se valoran los errores de transcripción detectados durante el proceso de protocolización.

## Análisis estadístico

El análisis estadístico se realizó mediante el uso de un paquete de software (SPSS, versión 19.0 para Windows, SPSS, Chicago, Illinois). Las diferencias interobservador se evaluaron mediante el índice de Kappa de Cohen. El coeficiente de concordancia tiene un valor que oscila entre 0 y 1, y su interpretación se corresponde con los siguientes rangos:

- No existe concordancia.
- 0,01-0,19 - Concordancia muy baja.
- 0,20-0,39 - Concordancia baja.
- 0,40-0,59 - Concordancia moderada.
- 0,60-0,79 - Concordancia buena.
- 0,80-0,99 - Concordancia alta.
- Concordancia perfecta.

Los valores de  $p < 0,05$  se consideraron estadísticamente significativos. Existe concordancia entre los protocolizadores solo si el coeficiente de correlación es superior a 0,00, con o sin significación estadística.

Para la comparación entre los grupos 2 y 3 se aplicó el test de t de Student. Se aplicó el test exacto de Fisher cuando las tablas de contingencia tenían celdas con una frecuencia esperada  $< 5$ .

## Resultados

### Aspectos generales del proyecto

En el estudio se incluyen un total de 689 pacientes con carcinoma, que corresponden al grupo a estudio de carcinoma

familiar ( $n = 252$ ) y al grupo control de carcinoma esporádico ( $n = 437$ ).

### Análisis de concordancia según la clasificación pronóstica del sistema TNM

#### 1. Valoración del componente T del sistema TNM

La protocolización del componente T del sistema TNM ha presentado un 4,2% ( $n = 29$ ) de no concordancias. En el grupo 2 del carcinoma familiar, con 30 protocolizadores, las discrepancias son del 6% ( $n = 15$ ), mientras que en el grupo 3 control del carcinoma esporádico, con 4 protocolizadores, las discrepancias representan el 3,2% ( $n = 14$ ) de los casos (tabla 1). Dichas diferencias no son significativas ( $p = 0,125$ ).

Tanto a nivel global ( $Kappa = 0,939$ ) como en el grupo 2 del carcinoma familiar ( $Kappa = 0,920$ ) y en el grupo 3 control ( $Kappa = 0,950$ ), la concordancia entre protocolizadores es alta (tabla 2).

#### 2. Valoración del componente N del sistema TNM

La protocolización del componente N de afectación ganglionar del sistema TNM ha presentado un 0,7% ( $n = 5$ ) de no concordancias, todas ellas en el grupo 2 del carcinoma familiar (2%;  $n = 5$ ) ( $p = 0,003$ ) (tabla 1).

Tanto a nivel global ( $Kappa = 0,979$ ) como en el grupo 2 del carcinoma familiar ( $Kappa = 0,949$ ) la concordancia entre protocolizadores es alta, y en el grupo 3 control ( $Kappa = 1,000$ ) la concordancia es perfecta (tabla 3).

#### 3. Valoración del componente M del sistema TNM

La protocolización del componente M de afectación sistémica del sistema TNM ha presentado una concordancia perfecta en los tres grupos (tablas 1 y 4).

**Tabla 1 – Porcentaje de no concordancias entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data-manager**

Variables	Concordancia	No concordancia	Porcentaje
<b>Grupo 1 – Serie global</b>			
T	660	29	4,2%
N	684	5	0,7%
M	689	0	0%
Estadio	672	17	2,5%
Valoración riesgo ATA	641	48	7%
TIRADS	549	41	6,9%
Errores de transcripción	657	32	4,6%
<b>Grupo 2 – Carcinoma familiar</b>			
T	237	15	6%
N	247	5	2%
M	252	0	0%
Estadio	242	10	4%
Valoración riesgo ATA	234	18	7,1%
TIRADS	201	14	6,5%
Errores de transcripción	240	12	4,8%
<b>Grupo 3 – Control</b>			
T	423	14	3,2%
N	437	0	0%
M	437	0	0%
Estadio	430	7	1,6%
Valoración Riesgo ATA	407	30	6,9%
TIRADS	348	27	7,2%
Errores de transcripción	417	20	4,6%

**Tabla 2 – Concordancia entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data-manager en la valoración del componente T (tamaño) del sistema TNM**

Valoración inicial en el centro de origen	Valoración centralizada por <i>data-manager</i>					
	T1a	T1b	T2	T3	T4	Total
Grupo 1 – Serie global						
Coeficiente de concordancia de Kappa 0,939						
$p < 0,0001$						
T1a	306	3	0	5	0	314
T1b	7	144	6	1	0	158
T2	0	2	99	1	0	102
T3	0	0	2	105	0	107
T4	0	0	2	0	6	8
Total	313	149	109	112	6	689
Grupo 2 – Carcinoma familiar						
Coeficiente de concordancia de Kappa 0,920						
$p < 0,0001$						
T1a	74	3	0	3	0	80
T1b	4	65	1	0	0	70
T2	0	2	48	0	0	50
T3	0	0	2	48	0	50
T4	0	0	0	0	2	2
Total	78	70	51	51	2	252
Grupo 3 – Control						
Coeficiente de concordancia de Kappa 0,950						
$p < 0,0001$						
T1a	232	0	0	2	0	234
T1b	3	79	5	1	0	88
T2	0	0	51	1	0	52
T3	0	0	0	57	0	57
T4	0	0	2	0	4	6
Total	235	79	58	61	4	437

#### 4. Valoración del estadio TNM

La protocolización del estadio del sistema TNM ha presentado un 2,5% (n = 17) de no concordancias (n = 17). Dicho porcentaje es mayor en el grupo 2 del carcinoma familiar (4%; n = 10) que

en el grupo 3 control del carcinoma esporádico (1,6%; n = 7), si bien dichas diferencias no son significativas ( $p = 0,094$ ) (tabla 1).

Tanto a nivel global (Kappa = 0,931) como en el grupo 2 del carcinoma familiar (Kappa = 0,904) y en el grupo 3 control

**Tabla 3 – Concordancia entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data-manager en la valoración del componente N (afectación ganglionar) del sistema TNM**

Valoración inicial en el centro de origen	Valoración por <i>data-manager</i>			
	N0	N1a	N1b	Total
<b>Grupo 1 – Serie global</b>				
Coeficiente de concordancia de Kappa 0,979				
$p < 0,0001$				
N0	541	0	0	541
N1a	5	95	0	100
N1b	0	0	48	48
Total	546	95	48	689
<b>Grupo 2 – Carcinoma familiar</b>				
Coeficiente de concordancia de Kappa 0,949				
$p < 0,0001$				
N0	189	0	0	189
N1a	5	39	0	44
N1b	0	0	19	19
Total	194	39	19	252
<b>Grupo 3 – Control</b>				
Coeficiente de concordancia de Kappa 1,000				
$p < 0,0001$				
N0	352	0	0	352
N1a	0	56	0	56
N1b	0	0	29	29
Total	352	56	29	437

**Tabla 4 – Concordancia entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data-manager en la valoración del componente M (afectación sistémica) del sistema TNM**

Valoración inicial en el centro de origen	Valoración por data-manager		
	M0	M1	Total
<b>Grupo 1 – Serie global</b>			
Coeficiente de concordancia de Kappa 1,000			
$p < 0,0001$			
M0	687	0	687
M1	0	2	2
Total	687	2	689
<b>Grupo 2 – Carcinoma familiar</b>			
Coeficiente de concordancia de Kappa 1,000			
$p < 0,0001$			
M0	251	0	251
M1	0	1	1
Total	251	1	252
<b>Grupo 3 – Control</b>			
Coeficiente de concordancia de Kappa 1,000			
$p < 0,0001$			
M0	436	0	436
M1	0	1	1
Total	436	1	437

(Kappa = 0,950), la concordancia entre protocolizadores es alta (tabla 5).

#### Análisis de concordancia según la valoración del riesgo de recidiva según la ATA

La protocolización de la valoración del riesgo de recidiva según la ATA ha presentado un 7% (n = 48) de no concordancias. En el grupo 2 del carcinoma familiar las discrepancias son del 7,1%

**Tabla 5 – Concordancia entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data-manager en la valoración de la estadificación según el sistema TNM**

Valoración inicial en el centro de origen	Valoración por data-manager				
	I	II	III	IV	Total
<b>Grupo 1 – Serie global</b>					
Coeficiente de concordancia de Kappa 0,931					
$p < 0,0001$					
I	540	4	3	0	547
II	1	44	1	0	46
III	3	0	59	5	67
IV	0	0	0	29	29
Total	544	48	63	34	689
<b>Grupo 2 – Carcinoma familiar</b>					
Coeficiente de concordancia de Kappa 0,904					
$p < 0,0001$					
I	186	2	1	0	189
II	1	24	0	0	25
III	2	0	24	4	30
IV	0	0	0	8	8
Total	189	26	25	12	252
<b>Grupo 3 – Control</b>					
Coeficiente de concordancia de Kappa 0,950					
$p < 0,0001$					
I	354	2	2	0	358
II	0	20	1	0	21
III	1	0	35	1	37
IV	0	0	0	21	21
Total	355	22	38	22	437

(n = 18), mientras que en el grupo 3 control del carcinoma esporádico las discrepancias representan el 6,9% (n = 30) de los casos (tabla 1). Dichas diferencias no son significativas ( $p = 0,890$ ).

**Tabla 6 – Concordancia entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data-manager en la valoración del riesgo de recidiva según la ATA**

Valoración inicial en el centro de origen	Valoración por data-manager			
	Bajo riesgo	Riesgo intermedio	Alto riesgo	Total
<b>Grupo 1 – Serie global</b>				
Coeficiente de concordancia de Kappa 0,819				
$p < 0,0001$				
Bajo riesgo	498	40	1	539
Riesgo intermedio	1	127	6	134
Alto riesgo	0	0	16	16
Total	499	167	23	689
<b>Grupo 2 – Carcinoma familiar</b>				
Coeficiente de concordancia de Kappa 0,829				
$p < 0,0001$				
Bajo riesgo	177	13	0	190
Riesgo intermedio	1	46	4	51
Alto riesgo	0	0	11	11
Total	178	59	15	252
<b>Grupo 3 – Control</b>				
Coeficiente de concordancia de Kappa 0,813				
$p < 0,0001$				
Bajo riesgo	321	27	1	349
Riesgo intermedio	0	81	2	83
Alto riesgo	0	0	5	5
Total	321	108	8	437

Tanto a nivel global (Kappa = 0,819) como en el grupo 2 del carcinoma familiar (Kappa = 0,829) y en el grupo 3 control (Kappa = 0,813), la concordancia entre protocolizadores es alta (tabla 6).

#### Análisis de concordancia según la valoración de la clasificación ecográfica del TIRADS

La protocolización de la valoración de la clasificación ecográfica del TIRADS ha presentado un 6,9% (n = 41) de no concordancias. En el grupo 2 del carcinoma familiar las discrepancias son del 6,5% (n = 14), mientras que en el grupo 3 control del carcinoma esporádico las discrepancias representan el 7,2% (n = 27) de los casos (tabla 1). Dichas diferencias no son significativas (p = 0,739).

Tanto a nivel global (Kappa = 0,922) como en el grupo 2 del carcinoma familiar (Kappa = 0,921) y en el grupo 3 control (Kappa = 0,922), la concordancia entre protocolizadores es alta (tabla 7).

#### Errores de transcripción

Se han detectado errores de transcripción en el 4,6% (n = 32) de los casos, porcentaje similar en el grupo 2 (4,8%; n = 12) y en el grupo 3 (4,6%; n = 20) (p = 0,939) (tabla 1).

#### Discusión

En un ensayo clínico o estudio multicéntrico uno de los documentos básicos es el cuaderno de recogida de datos, donde el investigador incorpora los datos, clínicos y no clínicos, de los pacientes incluidos en el estudio. Estos datos deben integrarse en una base de datos electrónica para poder realizar el análisis de los mismos<sup>10,11</sup>. En la actualidad la recogida de datos tiende a realizarse directamente en una base de datos para agilizar el proyecto<sup>8</sup>. Además, los formularios electrónicos de recogida de datos suelen contener rutinas de verificación que reducen la entrada de datos erróneos y

**Tabla 7 – Concordancia entre la valoración inicial realizada en el centro de origen y la valoración final centralizada por data manager en la valoración del TIRADS ecográfico**

valoración inicial en el centro de origen	Valoración por <i>data-manager</i>							Total
	TIRADS 1	TIRADS 2	TIRADS 3	TIRADS 4a	TIRADS 4b	TIRADS 4c	TIRADS 5	
Grupo 1 – Serie global								
Coeficiente de concordancia de Kappa 0,922								
<i>p</i> < 0,0001								
TIRADS 1	0	1	0	0	0	0	0	1
TIRADS 2	0	34	3	0	0	0	0	37
TIRADS 3	0	2	27	0	0	0	0	29
TIRADS 4 <sup>a</sup>	0	0	0	104	2	0	0	106
TIRADS 4b	0	0	0	1	155	19	0	175
TIRADS 4c	0	0	1	0	8	211	4	224
TIRADS 5	0	0	0	0	0	0	18	18
Total	0	37	31	105	165	230	22	590
Missing data								
99								
Grupo 2 – Carcinoma familiar								
Coeficiente de concordancia de Kappa 0,921								
<i>p</i> < 0,0001								
TIRADS 1	0	0	0	0	0	0	0	0
TIRADS 2	0	12	0	0	0	0	0	12
TIRADS 3	0	0	1	0	0	0	0	1
TIRADS 4a	0	0	0	43	0	0	0	43
TIRADS 4b	0	0	0	1	45	9	0	56
TIRADS 4c	0	0	0	0	3	91	1	95
TIRADS 5	0	0	0	0	0	0	9	9
Total	0	12	1	44	48	100	10	215
Missing data								
37								
Grupo 3 – Control								
Coeficiente de concordancia de Kappa 0,922								
<i>p</i> < 0,0001								
TIRADS 1	0	1	0	0	0	0	0	1
TIRADS 2	0	22	3	0	0	0	0	25
TIRADS 3	0	2	26	0	0	0	0	28
TIRADS 4a	0	0	0	61	2	0	0	63
TIRADS 4b	0	0	0	0	110	10	0	120
TIRADS 4c	0	0	1	0	5	120	3	129
TIRADS 5	0	0	0	0	0	0	9	9
Total	0	25	30	61	117	130	12	375
Missing data								
62								



ofrecen a los administradores una visión continua de los datos y del proceso de recopilación<sup>4,12</sup>, pero presentan un coste difícil de asumir en muchos de los estudios multicéntricos clínicos, si bien hay que destacar que cada vez hay más plataformas informáticas y es de esperar que los costes irán bajando.

Independientemente de la forma de protocolización, es importante garantizar que la información se introduzca en la base de datos de forma coherente y precisa<sup>13-15</sup>, ya que incluso una pequeña proporción de errores puede tener un gran impacto en los resultados de un estudio<sup>16</sup>. Los errores aleatorios y sistémicos inadvertidos en la base de datos son fuentes de sesgo. Además, la disponibilidad de sistemas electrónicos confiables no es suficiente para garantizar la validez de los estudios transversales de base poblacional. De hecho, la relevancia del estudio depende en gran medida de dos pasos principales: la calidad de la recopilación de datos en las historias clínicas y la fidelidad de la transferencia de datos desde las historias clínicas al sistema electrónico. Cualquier debilidad en estas dos etapas invalida el estudio<sup>16-19</sup>. En este sentido, hay que resaltar la importancia de la capacitación de los investigadores clínicos, en especial la formación en buenas prácticas clínicas<sup>20-22</sup>.

Por todo ello, aunque los estudios multicéntricos ofrecen la ventaja de obtener información sobre patologías poco frecuentes en un período de tiempo relativamente corto, son estudios más heterogéneos y con más riesgos de presentar sesgos durante su realización<sup>1,3,22-24</sup>. Por ello, hay que buscar el equilibrio entre disminuir al máximo los potenciales sesgos que puede presentar el estudio, y no incrementar los costes de los estudios que los hagan poco viables. En este sentido, conceptualmente la protocolización centralizada de los datos por un *data-manager* implica una protocolización más homogénea de los mismos. Sin embargo, implica un coste económico a cargo del proyecto de investigación, a veces no asumible<sup>1,3,22-24</sup>, y suele implicar un mayor tiempo de protocolización que la protocolización independiente por los diferentes grupos.

Nuestro proyecto multicéntrico, desarrollado en el contexto de una sociedad científica nacional, muestra como con un desarrollo adecuado y un protocolo detallado con una explicación completa de lo que implica cada variable la concordancia es alta, y, por lo tanto, aceptable. El análisis detallado de los datos muestra que aunque en casi todas las variables analizadas la concordancia es alta y en algunos casos perfecta, existe un porcentaje de discordancias que en algunas variables llega hasta el 7%. Además, dichas discordancias se acentúan cuando el número de protocolizadores es más alto. Así, hay más discordancias en el grupo 2 del carcinoma familiar, donde hay 30 protocolizadores, que en el grupo 3 del grupo control del carcinoma esporádico, donde hay 4.

Además, las discrepancias se centran sobre todo en variables que tienen algún tipo de matización en su determinación. En este sentido, en variables como la afectación linfática (N) o la presencia de metástasis (M) la concordancia es casi perfecta, y el porcentaje de no concordancia es del 0,7% y del 0%, respectivamente. Sin embargo, en variables como la T, donde no solo es el tamaño, sino que factores como la afectación de estructuras vecinas, etc., la condicionan, el

porcentaje de no concordancia es superior al 4%. Esta situación se acentúa en variables donde la clasificación es aún más compleja. Así, hay un 7% de discordancia en la valoración del riesgo de la ATA y un 6,9% en la clasificación ecográfica del TIRADS.

Por todo ello, aunque la concordancia entre la protocolización directa en el centro de origen y centralizada por un *data-manager* es alta, existe un porcentaje de discordancias en variables pronósticas importantes que oscila entre el 2,5% en el estadio tumoral TNM y el 7% en la valoración del riesgo según la ATA. Esto implica un porcentaje de error o sesgo solo en el proceso de protocolización de los datos del estudio multicéntrico. Así, en estas variables, aunque la concordancia es alta, está en el límite bajo de dicha concordancia, cerca de una concordancia solo buena. A esto hay que añadir un 4,6% de errores de transcripción detectados por el *data-manager*. Hay que recordar que la protocolización es un proceso lento y a veces tedioso, y no siempre se encarga el investigador principal en el centro de origen, sino que con frecuencia se delega. Todo ello conlleva errores de transcripción. Hay que indicar que el objetivo de este estudio no es detectar errores de transcripción, lo cual hace presuponer que posiblemente ese porcentaje sea mayor, ya que solo se pueden detectar errores de protocolización evidentes. Para detectarlos todos habría que revisar todas las variables del protocolo.

Por ello, la recogida de datos debe estar bien supervisada para mejorar la calidad de los estudios, ya que aunque el tener o no un *data-manager* no presenta diferencias, casi un 5% presenta errores de transcripción, y por ello hay que hacer bases de datos claras y de fácil cumplimentación, y es aconsejable que un miembro del equipo de recogida, con experiencia, revise los datos de las variables principales a estudio; o en la fase de diseño del estudio, incluir menos variables en la base pero de mayor calidad e interés para el análisis.

Se coincide con otros autores en que si el personal que protocoliza los datos está bien formado y supervisado, este no es el eslabón más débil de la cadena de gestión de datos<sup>6,25,26</sup>, si bien conlleva un margen de error, como se ha visto. Un aspecto que no se ha corroborado en nuestro proyecto, posiblemente por ser un protocolo corto, es que la posición en el campo de inserción en el protocolo juegue un papel importante en la proporción de errores. Las últimas posiciones están asociadas con más errores que las iniciales, especialmente cuando se consideran campos numéricos. Esto se ha atribuido a la fatiga del administrador de datos cuando los cuestionarios tienen demasiadas entradas<sup>6</sup>. Por lo tanto, estos resultados sugieren que, para crear cuestionarios más efectivos, la información más importante debe recopilarse en los primeros campos, debe reducirse el número de campos y deben usarse cuadros combinados o cuadros de texto en lugar de campos con inserción numérica directa (especialmente en la última parte del cuestionario). Una limitación de este estudio es que se basó en un único gestor de datos, por lo que es difícil generalizar nuestras conclusiones. Sin embargo, debe recordarse que la decisión de utilizar un solo administrador de datos se introdujo para mejorar la confiabilidad de la entrada de datos al reducir el sesgo entre administradores de datos. En este sentido hay que indicar que



sería interesante poder realizar la comparación con otros estudios multicéntricos en otros contextos, para poder realizar una validación externa de los resultados. Por último, habría que valorar la calidad de los datos en los estudios multicéntricos sin *data-manager*, respecto a quién realiza la recogida de los datos (personal administrativo, residentes, etc.) y bajo qué supervisión, ya que todo ello puede condicionar los resultados, en cuyo caso la presencia de un *data-manager* sí que podría presentar diferencias.

En base a los datos de nuestro estudio, no es precisa la existencia de un *data-manager* dada la concordancia alta entre los dos sistemas de protocolización, y ello conlleva ahorro económico y de tiempo<sup>8</sup>. Los inconvenientes que puede presentar la protocolización online se pueden resolver, como en nuestro estudio, con bases de datos estándar que no precisan coste de línea directa o mantenimiento del sistema online, si bien no presenta las ventajas que ofrecen estos productos online con control de calidad, pero para proyectos multicéntricos como el que se evalúa en este trabajo pueden ser suficientes. Todos estos aspectos son importantes en estudios multicéntricos de patologías infrecuentes, donde la aportación de casos es importante, pero a la vez la financiación económica es reducida o inexistente. En un ensayo clínico en fase II los costos de administración de datos pueden estimarse como casi el 30% de los costos de los ensayos clínicos de fase III<sup>8</sup>, y posteriormente dichos costos van a repercutir en los costos de comercialización de los medicamentos. En los estudios multicéntricos clínicos dichos costos van a condicionar la viabilidad del proyecto en sí<sup>27</sup>. Por último, indicar que, aunque no hay diferencias y podría no ser necesario un *data-manager* para protocolizar los datos, hay que mencionar que los *data-manager* suelen tener otras funciones en el desarrollo de un proyecto, como la coordinación de las pruebas de los pacientes, solucionar problemas que puedan surgir en los centros participantes, servir de enlace entre los investigadores, etc. Por lo tanto, si el presupuesto del proyecto lo permite, su contratación suele agilizar la realización de un proyecto.

En conclusión, podemos decir que en los estudios clínicos multicéntricos la protocolización centralizada de los datos por un *data-manager* parece presentar resultados similares a la protocolización directa en la base de datos en el centro de origen.

## Comité ético

El protocolo del estudio fue aprobado por el Comité Ético de Centro. Código del Comité de Ética: 2021-2-13-HCUVA.

Este estudio en seres humanos ha sido revisado por el Comité de Ética correspondiente y, por lo tanto, se ha realizado de acuerdo con las normas éticas establecidas en la Declaración de Helsinki de 2000, así como en la Declaración de Estambul de 2008.

## Aspectos relevantes a destacar en el proyecto

No hay aspectos relevantes desde el punto de vista de la investigación que precisen ser destacados.

## Contribución de los autores

1. Concepción y diseño: Ríos A.
2. Adquisición de la parte sustancial de los datos: Ríos A, Ferrero-Herrero E, Puñal-Rodríguez JA, Durán M, Mercader-Cidoncha E, Ruiz-Pardo J, Rodríguez JM.
3. Análisis e interpretación de los datos: Ríos A, Gutiérrez PR.
4. Redacción del borrador del artículo: Ríos A.
5. Revisión crítica del artículo: Ríos A, Gutiérrez PR, Mercader-Cidoncha E.
6. Análisis y revisión estadística: Ríos A.
7. Obtención de fondos para la financiación del proyecto: Ríos A.
8. Supervisión: Ríos A, Ferrero-Herrero E, Puñal-Rodríguez JA, Gutiérrez PR, Durán M, Mercader-Cidoncha E, Ruiz-Pardo J, Rodríguez JM.
9. Aprobación final de la versión a publicar: Ríos A, Ferrero-Herrero E, Puñal-Rodríguez JA, Gutiérrez PR, Mercader-Cidoncha E, Ruiz-Pardo J, Rodríguez JM.

## Financiación

No se han utilizado fuentes de financiación específicas para este estudio.

## Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

## Agradecimientos

Agradecer a la Asociación Española de Cirujanos y a la Sección de Cirugía Endocrina su apoyo en la realización de este trabajo.

Agradecer a las 24 unidades de cirugía endocrina españolas que participan en este estudio haber colaborado en la realización de este proyecto y facilitado los datos que lo han hecho posible.

## BIBLIOGRAFÍA

1. Serra-Aracil X, Pascua-Sol M, Badia-Closa J, Navarro-Soto S. Cómo poner en marcha y desarrollar un estudio multicéntrico prospectivo, controlado y aleatorizado. *Cir Esp*. 2020;98:119-26. <http://dx.doi.org/10.1016/j.ciresp.2019.11.012>.
2. Hemminki K, Eng C, Chen B. Familial risks for nonmedullary thyroid cancer. *J Clin Endocrinol Metab*. 2005;90:5747-53. <http://dx.doi.org/10.1210/jc.2005-0935>.
3. Pannucci CJ, Wilkins EG. Identifying and avoiding bias in research. *Plast Reconstr Surg*. 2010;126:619-25. <http://dx.doi.org/10.1097/PRS.0b013e3181de24bc>.
4. Welker JA. Implementation of electronic data capture systems: Barriers and solutions. *Contemp Clin Trials*. 2007;28:329-36. <http://dx.doi.org/10.1016/j.cct.2007.01.001>.
5. Schmier JK, Kane DW, Halpern MT. Practical applications of usability theory to electronic data collection of clinical trials.

- Contemp Clin Trials. 2005;26:375–85. <http://dx.doi.org/10.1016/j.cct.2007.01.001>.
6. Mauri D, Karampoiki V, Mauri J, Kamposioras K, Alexiou G, Ferentinos G, et al., SESy Europe Task Force. Double-blind control of the data manager doesn't have any impact on data entry reliability and should be considered as an avoidable cost. BMC Med Res Methodol. 2008;8:66. <http://dx.doi.org/10.1186/1471-2288-8-66>.
7. Howells K. Living without a data management system. IDrugs. 2008;11:733–7.
8. Pavlović I, Kern T, Miklavcic D. Comparison of paper-based and electronic data collection process in clinical trials: Costs simulation study. Contemp Clin Trials. 2009;30:300–16. <http://dx.doi.org/10.1016/j.cct.2009.03.008>.
9. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, et al. Thyroid imaging reporting and data system for US features of nodules: A step in establishing better stratification of cancer risk. Radiology. 2011;260:892–9. <http://dx.doi.org/10.1148/radiol.11110206>.
10. Lee N, Millman A, Osborne M, Cox J. ABC of medical computing. Storing and managing data on a computer. BMJ. 1995;311:562–5. <http://dx.doi.org/10.1136/bmj.311.7004.562>.
11. Millman A, Lee N, Brooke A. ABC of medical computing. Computers in general practice – I. BMJ. 1995;311:800–2. <http://dx.doi.org/10.1136/bmj.311.7008.800>.
12. Brandt CA, Argraves S, Money R, Ananth G, Trocky NM, Nadkarni PM. Informatics tools to improve clinical research study implementation. Contemp Clin Trials. 2006;27:112–22. <http://dx.doi.org/10.1016/j.cct.2005.11.013>.
13. Los RK, van Ginneken AM, Roukema J, Moll HA, Lei van der J. Why are structured data different? Relating differences in data representation to the rationale of OpenSDE. Med Inform Internet Med. 2005;30:267–76. <http://dx.doi.org/10.1080/14639230500367563>.
14. De Lusignan S. The barriers to clinical coding in general practice: A literature review. Med Inform Internet Med. 2005;30:89–97. <http://dx.doi.org/10.1080/14639230500298651>.
15. Marks RG. Validating electronic source data in clinical trials. Control Clin Trials. 2004;25:437–46. <http://dx.doi.org/10.1016/j.cct.2004.07.001>.
16. Mullooly JP. The effects of data entry error: An analysis of partial verification. Comput Biomed Res. 1990;23:259–67. [http://dx.doi.org/10.1016/0010-4809\(90\)90020-D](http://dx.doi.org/10.1016/0010-4809(90)90020-D).
17. Levitt SH, Aeppli DM, Potish RA, Lee CK, Nierengarten ME. Influences on inferences: Effect of errors in data on statistical evaluation. Cancer. 1993;72:2075–82. [http://dx.doi.org/10.1002/1097-0142\(19931001\)72:7<2075](http://dx.doi.org/10.1002/1097-0142(19931001)72:7<2075).
18. Arndt S, Tyrrell G, Woolson RF, Flaum M, Andreasen NC. Effects of errors in a multicenter medical study: Preventing misinterpreted data. J Psychiatr Res. 1994;28:447–59. [http://dx.doi.org/10.1016/0022-3956\(94\)90003-5](http://dx.doi.org/10.1016/0022-3956(94)90003-5).
19. Crombie IK, Irving JM. An investigation of data entry methods with a personal computer. Comput Biomed Res. 1986;19:543–50. [http://dx.doi.org/10.1016/0010-4809\(86\)90028-5](http://dx.doi.org/10.1016/0010-4809(86)90028-5).
20. Saleh M, Naik G, Jester P, Joiner C, Westfall E, Kimberlin DW, et al. Clinical Investigator Training Program (CITP) — A practical and pragmatic approach to conveying clinical investigator competencies and training to busy clinicians. Contemp Clin Trials Commun. 2020;19:100589. <http://dx.doi.org/10.1016/j.conctc.2020.100589>.
21. Swezey T, McGuire FH, Hurley P, Panhuis J, Goldstein K, Chuck T, et al. More than a box to check: Research sponsor and clinical investigator perspectives on making GCP training relevant. Contemp Clin Trials Commun. 2020;19:100606. <http://dx.doi.org/10.1016/j.conctc.2020.100606>.
22. Bechtel J, Chuck T, Forrest A, Hildebrand C, Panhuis J, Pattee SR, et al. Improving the quality conduct and efficiency of clinical trials with training: Recommendations for preparedness and qualification of investigators and delegates. Contemp Clin Trials. 2020;89:105918. <http://dx.doi.org/10.1016/j.cct.2019.105918>.
23. Huang GD, Bull J, McKee KJ, Mahon E, Harper B, Roberts JN. Clinical trials recruitment planning: A proposed framework from the Clinical Trials Transformation Initiative. Contemporary Clinical Trials. 2018;66:74–9. <http://dx.doi.org/10.1016/j.cct.2018.01.003>.
24. Chung KC, Song JW, The WRIST study group. A guide on organizing a multicenter clinical trial. Plast Reconstr Surg. 2010;126:515–23. <http://dx.doi.org/10.1097/PRS.0b013e3181df64fa>.
25. Day S, Fayers P, Harvey D. Double data entry: What value, what price? Control Clin Trials. 1998;19:15–24. [http://dx.doi.org/10.1016/S0197-2456\(97\)00096-2](http://dx.doi.org/10.1016/S0197-2456(97)00096-2).
26. Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. Control Clin Trials. 1992;13:487–94. [http://dx.doi.org/10.1016/0197-2456\(92\)90205-E](http://dx.doi.org/10.1016/0197-2456(92)90205-E).
27. Ríos A, Rodríguez MA, Puñal JA, Moreno P, Mercader E, Ferrero E, et al. Biological behavior of familial papillary thyroid microcarcinoma: Spanish multicenter study. Langenbecks Arch Surg. 2022;407:3631–42. <http://dx.doi.org/10.1007/s00423-022-02704-4>. PMID: 36251077.