

ESTRATEGIA DE ANÁLISIS DE DATOS (PARTE 1) : CREACIÓN DE BASES DE DATOS PARA INVESTIGACIONES EN CIENCIAS DE LA SALUD

DATA ANALYSIS STRATEGY (PART 1) : DATABASE CREATION FOR RESEARCH IN HEALTH SCIENCES

Torracchi-Carrasco Esteban^{1*}, Córdova-López Andrea², Chiriboga-Guartambel Gladys², Villavicencio-Caparó Ebingen¹

¹ Catedrático de la Facultad de Odontología de la Universidad Católica de Cuenca. Ecuador

² Odontóloga de consulta particular en Cuenca. Ecuador

*jtorracchic@ucacue.edu.ec

Resumen

Un trabajo de investigación puede ser considerado científicamente válido cuando es el producto de un proceso correcto de obtención, sistematización y análisis de información, la cual, lógicamente, debe ser verificable. Esta información debe surgir de una pregunta de investigación formulada de manera clara y concisa, persiguiendo objetivos que enmarquen el nivel y la profundidad de la información con la que se pretende trabajar. Asimismo, es imperativo que la recolección de datos se sustente en un diseño experimental mediante una adecuada planificación en torno a la investigación en la cual nos hemos embarcado. El objetivo de este trabajo es presentar una propuesta acerca de cómo se debería llevar a cabo el proceso de recolección de datos y construcción de una base de datos para que la información allí depositada resulte confiable al momento de realizar nuestros análisis y especialmente al momento de discutir nuestros resultados.

Palabras clave: Pregunta de investigación, Ciclo de Indagación, Diseño metodológico, Recolección de datos, Análisis de Datos.

Abstract

A research work can be considered scientifically valid when it is the product of a correct process of obtaining, systematizing and analyzing information, which, logically, must be verifiable. This information should come from a research question formulated in a clear and concise manner; Obtaining objectives that frame the level and depth of the information with which it is intended to work. Likewise, it is imperative that the data collection is based on an experimental design through adequate planning around the research in which we have embarked. The objective of this paper is to present a proposal about how the process of data collection and database realization should be carried out so that the extracted information is reliable at the time of our analysis and especially when discussing our results.

Key words: Research Questions, Inquiry Cycle, Research Design, Data Collection, Data Analysis.

1 INTRODUCCIÓN

Desde hace poco más de dos siglos, la investigación científica representa la forma más segura de aproximarse al conocimiento acerca de un determinado fenómeno de la realidad. Este conocimiento debe ser válido y confiable. Sin embargo, determinados aspectos críticos de una investigación pueden vulnerar alguna de estas características. Uno de estos aspectos es el manejo (obtención, procesamiento y análisis) de los datos; por ese motivo, en muchas universidades, los formatos de protocolo de investigación incluyen un acápite en el capítulo -de planteamiento operacional o materiales

y métodos- que se refiere a la estrategia del análisis de datos.¹ En relación con este punto en particular, existe mucha bibliografía acerca del uso y manejo de pruebas estadísticas; sin embargo, hay poca literatura acerca de la técnica más adecuada para crear una buena base de datos.^{1,2} Considerando que se trata de una fase previa al análisis estadístico, es necesario que se realice adecuadamente y que se conozca la forma óptima para su codificación y su organización, con la finalidad de poder controlar la calidad de los datos y tratar de prevenir los sesgos por ingreso incorrecto de los mismos.³

El presente es el primero de tres artículos de contribución didáctica docente, que abordarán el tema del análisis de datos. Para ello, recapitularemos primero todas las etapas del proceso de investigación y explicaremos a continuación la fase pertinente a la construcción de las bases de datos, con la finalidad de ubicar en el contexto este tema tan importante para el desarrollo de un trabajo científico.

2 LA BASE DE DATOS EN EL CONTEXTO DE UNA INVESTIGACIÓN CIENTÍFICA

La secuencia de pasos correctos y sistematizados, para crear una base de datos, está basada en un sistema estructurado para lograr resultados y conclusiones válidas, que aporten credibilidad al trabajo realizado. Esta secuencia está compuesta por cuatro pasos²⁻⁴ (Figura 1^{5,6}):

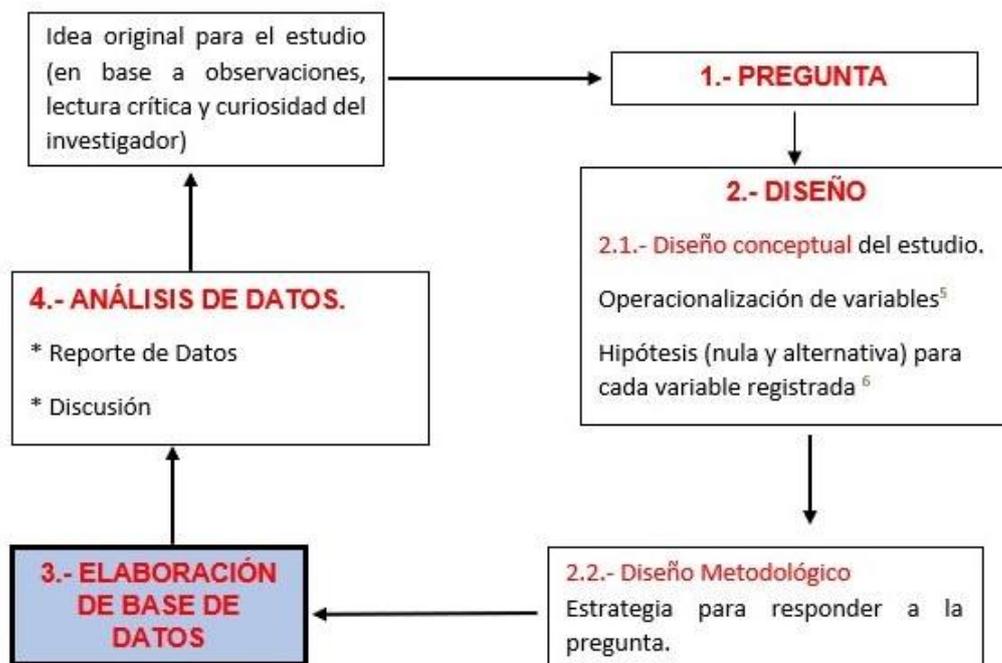


Fig. 1. Ciclo de Indagación científica (Tomado y Modificado de Feinsiger).⁷

2.1 FORMULACIÓN DE LA PREGUNTA DE INVESTIGACIÓN

Una pregunta de investigación debe ser formulada de forma precisa⁸ y clara de tal manera que no exista confusión con respecto a la respuesta esperada.⁹ Esta pregunta nace de una interrogante que se plantea al identificar un vacío de información ("brecha en el conocimiento") o un problema en una determinada área.^{1,9} La interrogante formulada debe tener en consideración si la respuesta a la misma aportará nueva información.¹⁰

La pregunta de investigación debe ser:¹¹

- Concisa: utilizar lenguaje claro con frases cortas y directas
- Viable: el estudio y la recolección de datos debe ser posible.
- Relevante: la investigación debe ser importante desde el punto de vista teórico, científico y social.
- Ética: si la investigación es en seres vivos, deben res-

tarse los principios bioéticos.

2.2 DISEÑO DEL ESTUDIO

2.2.1 Diseño Conceptual

Se encuentra representado en la Figura 1.

2.3 ÁMBITO

Para empezar el diseño conceptual de un estudio es necesario conocer el ámbito en que se desarrollará el mismo, es decir, el espacio en el que se efectuará el estudio.^{12,13} En relación con este aspecto se pueden diferenciar los estudios experimentales y los observacionales.¹⁴

Los primeros se centran en intervenir sobre el sujeto de investigación,¹⁵ es decir, el investigador manipula o controla las condiciones de la investigación; los segundos se realizan sin intervención alguna de parte del investigador, quien actúa como un observador de la realidad, limitándose a medir y analizar las variables definidas en el estudio.⁶ Este tipo de estudios se clasifican, a su vez, en descriptivos y analíticos.¹⁶

2.3.1 Temporalidad

En las investigaciones de corte epidemiológico es necesario establecer la temporalidad en la cual van a desarrollarse.¹⁷ Esta temporalidad puede ser de cuatro tipos: retrospectiva (registro de datos que ocurrieron en el pasado); prospectiva (los hechos se registran a medida que suceden); transversal (obtiene información de una muestra determinada de elementos de una población específica en un espacio de tiempo limitado); y longitudinal (incluye una muestra fija de elementos de la población que son evaluados más de una vez a lo largo de un período de tiempo determinado).^{11,18,19}

2.3.2 Definición de variables

Una vez establecida la pregunta de investigación, el ámbito y la temporalidad del estudio, el punto siguiente es definir las variables³ y su rol en la investigación. Se trata de características medibles u observables (por ejemplo, de una persona u objeto), que puedan ser expresadas en términos numéricos o categóricos.⁵

Existen cuatro tipos estadísticos de variables:

- 1) Cualitativa nominal (cuando la variable tiene categorías que no pueden ser ordenadas).⁶
- 2) Cualitativa ordinal (cuando la variable tiene categorías que pueden ser ordenadas por intensidad).²⁰
- 3) Cuantitativa discreta (cuando la variable numérica no admite decimales).
- 4) Cuantitativa continua (cuando la variable numérica admite decimales).

Para más información sobre este aspecto, consultar Villavicencio y cols, 2019.⁵

2.3.3 Escala

Para cada variable se requiere de una escala de medición, la cual se clasifica en dos grandes grupos:⁵

- 1) Escalas no métricas: se definen como categorías que describen al fenómeno por estudiar, utilizado en variables cualitativas.²¹ A su vez, las escalas no métricas se clasifican en: nominales y ordinales. La escala nominal es aquella que no representa ninguna jerarquía en particular, mientras que la escala ordinal es aquella que mantiene un orden, descendente o ascendente, permitiendo jerarquizar los datos.²²
- 2) Escalas métricas: identifican y describen al fenómeno por la cantidad o grado de una característica específica, que se utiliza en variables cuantitativas.² Estas escalas se clasifican en: escala de intervalo (en la cual puede existir el cero como punto arbitrario e inclusive valores negativos); y escala de razón (en donde el cero asume un valor absoluto, y no se admite como valor de una variable).

2.3.4 Formulación de hipótesis

La hipótesis es un enunciado que plantea el investigador luego de analizar a fondo el tema de interés, siendo coherente con la pregunta de investigación y las variables a relacionar.^{17,23,24} Ésta hipótesis puede plantearse como la respuesta esperada a la pregunta inicial y transformarse en una guía hacia lo que se quiere estudiar.²⁵

Una hipótesis bien estructurada está compuesta por la unidad de observación y las variables que serán evaluadas.¹⁰ Además, en la hipótesis se puede indicar cómo se espera que se relacionen estos dos elementos, es decir, establecer la direccionalidad de la hipótesis.¹⁸ Respecto a la lógica existen tres tipos de hipótesis: las de comparación de promedios, las de comparación de proporciones y las de correlación de variables. En cuanto a la naturaleza estadística existen Hipótesis nula (H_0) e Hipótesis alternativa (H_1).²⁶

2.3.5 Diseño Metodológico

Es la estrategia que se adoptará para responder a la pregunta de investigación. Existen diseños de estudios clínicos y no clínicos. Los diseños clínicos más usados son: Caso-control, Descriptivo, de Sensibilidad y Especificidad, Ensayos clínicos aleatorizados y de Cohortes.²⁷

2.3.6 Recolección u obtención de datos

Para obtener los datos de una investigación de corte epidemiológico, se debe tomar en cuenta el universo de estudio, que es el total de personas u objetos con características específicas que deberían intervenir en el estudio.^{28,29} Generalmente no se evalúa a todo el universo, pues resultaría excesivamente costoso o tomaría demasiado tiempo hacerlo.²¹ Es por ello por lo que se debe trabajar con una muestra representativa de este universo, para lo cual es necesario realizar un cálculo de tamaño muestral, tomando en cuenta que a mayor tamaño muestral menor error.²⁷ La selección de la muestra es crucial para obtener resultados válidos y confiables. Existen dos tipos de muestras:¹²

- Las muestras probabilísticas se establecen de la siguiente manera:³⁰
 - 1) Muestreo simple: se elige la población al azar.
 - 2) Muestreo estratificado: se divide a la población en grupos y se aplica un muestreo simple dentro de cada grupo.
 - 3) Muestreo por segmentos o racimos: se elige la población usando criterios de proximidad.⁵
- Las muestras no probabilísticas establecidas según el criterio subjetivo del investigador, es decir, que no son representativas de todo el universo. Este tipo de muestras poblacionales son:³¹
 - 1) Muestreo por conveniencia.
 - 2) Muestreo por demanda de atención
 - 3) Muestreo por conglomerados.

- 4) Muestreo de sujetos voluntarios (que se usan en diseños de estudios de series de casos y controles.)

2.4 CREACIÓN DE UNA BASE DE DATOS

Una vez que se han definido el diseño de investigación apropiado y la muestra adecuada en base a la pregunta de investigación y la hipótesis, la siguiente etapa consiste en recolectar los datos y elaborar un plan detallado de proce-

dimientos que nos permitan organizarlos con el fin de analizarlos de manera correcta.²² Para ello es necesario crear o construir una base de datos.³² Recibe este nombre el conjunto de datos que están relacionados entre sí, los cuales son agrupados y estructurados en forma de campos y registros que son archivados para su posterior análisis.³³ Una base de datos puede ser una hoja de cálculo (por ejemplo, en formato Excel) que se organiza bajo la forma de filas y columnas²³ (Figura2).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	UniqueKey	NUMERO FICHA	NOMBRE	SEXO	CANTNDENACIMIENTO	PARROQUIA	FECHA	COLEGIO	EXAMINADOR	C18	C17	C16	C15	C14	C13
1	1	5,318		0	Cuenca	1	6/6/2016	Cuenca	WILSON	8	8	1	0	0	0
2	2	5,002		1	Cuenca	1	3/6/2016	Febres Cordero	Jessica	8	1	0	0	0	0
3	3	5,003		1	Cuenca	1	3/6/2016	Febres Cordero	Jessica	8	0	0	0	0	0
4	4	5,004		1	Cuenca	1	3/6/2016	Febres Cordero	LUCY	8	0	1	0	0	0
5	5	5,005		1	Cuenca	1	3/6/2016	Febres Cordero	LUCY	8	8	0	0	0	0
6	6	5,006		1	Cuenca	1	3/6/2016	Febres Cordero	DANIELA	8	8	3	8	8	8
7	7	5,007		1	Cuenca	1	3/6/2016	Febres Cordero	DANIELA	8	8	1	1	0	0
8	8	5,008		1	Cuenca	1	3/6/2016	Febres Cordero	DANIELA	8	8	0	0	0	1
9	9	5,009		1	Cuenca	0	3/6/2016	Febres Cordero	GEOVANY	8	0	0	0	0	0
10	10	5,011		1	Cuenca	0	3/6/2016	Febres Cordero	RAUL	8	0	1	0	0	0

Fig. 2. Anatomía de una base de datos.

A	B	C	D	E	F	G	H	I
UniqueKey	número de paciente	del 001 hasta el 682						
NUMERODEFICHA	número de ficha	de 5318 hasta el 6000						
NOMBRE	nombre del paciente codificado	datos codificado						
SEXO	Sexo del paciente	0= Femenino 1= Masculino						
CANTNDENACIMIENTO	Cantón de nacimiento	Dato que refiere el paciente						
PARROQUIA	Parroquia donde vive actualmente	Una de las 21 parroquias de la ciudad de Cuenca						
FECHA	Fecha que se tomó el dato	del 21/04/2016 al 10/ 12/2016						
COLEGIO	Nombre del centro educativo	Uno de los centros educativos de la muestra						
EXAMINADOR	Nombre del examinador	Uno de los 4 examinadores que levantó el dato en el estudio						
C18	diagnóstico de caries de la pieza 1.8	0= Sana	1= Cariada	2= Obturada y con caries	3= Obturada sana	4= Extraída por caries	5= Perdida por otro motivo	6= Sellant
C17	diagnóstico de caries de la pieza 1.7	idem						
C16	diagnóstico de caries de la pieza 1.6	idem						
C15	diagnóstico de caries de la pieza 1.5	idem						
C14	diagnóstico de caries de la pieza 1.4	idem						
C13	diagnóstico de caries de la pieza 1.3	idem						
C12	diagnóstico de caries de la pieza 1.2	idem						
C11	diagnóstico de caries de la pieza 1.1	idem						
C21	diagnóstico de caries de la pieza 2.1	idem						
C22	diagnóstico de caries de la pieza 2.2	idem						

Fig. 3. Hoja adicional - Diccionario.

Dentro de lo que se puede llamar la anatomía de una base de datos, las filas corresponden a cada unidad de estudio (por ejemplo, un paciente) y las columnas corresponden a cada uno de los campos (variables evaluadas) que se han considerado en la unidad de estudio.^{29,34} Se llama registro a cada fila y se denomina campo a cada columna.²⁴ La intersección entre la fila y la columna delimita una celda y esta contiene un dato. En el ejemplo planteado, la celda L5 contiene un dato que es el número 1, que por sí solo no tiene ningún significado.²² Sin embargo, en el contexto de esta base de datos, este valor adquiere un significado único, es decir: 0 (no presenta caries), 1 (presenta caries); así también, en el caso del paciente número 5, el valor correspondiente a la variable reportada en el campo C16 que significa el código para caries de la pieza 2.6, el valor 0 significa diente sano.³⁵

Para construir una base de datos, es recomendable codificar el nombre de las variables (campos) y guardar el significado en una hoja adicional denominada Diccionario³¹ (Figura 3). En esta hoja puede verse que en la columna A están codificados los nombres de las variables; en la columna B, se presenta el nombre extenso descriptivo de cada variable; en la columna C se muestran los datos que son admitidos en la celda (válidos en el registro) (nótese que a partir de la variable C18 en adelante se admiten los mismos códigos; por ese motivo, a partir de C11 se pone idem. que significa idéntico al anterior); en la columna D (de la base de datos), Figura 2 se presenta la variable sexo (la cual ha sido codificada con 0 para el sexo femenino y 1 para el sexo masculino (lo cual está explicado en el Diccionario

en la fila 4),⁶ Figura 3. Presentamos este ejemplo porque es la forma más conveniente de registrar los datos de una variable cualitativa dicotómica (como el sexo) y que puede ser exportada a cualquier paquete estadístico.²⁰

Otra ventaja de codificar las variables es que se evitan los errores dactilográficos.²⁸ Por ejemplo, en el caso del dato "Sexo Femenino", este puede ser escrito de 3 formas distintas (femenino, Femenino, FEMENINO), lo cual podría originar un conflicto en la base de datos al momento de hacer los cálculos de frecuencias, ya que la computadora contará cada una de estas formas como si fueran datos distintos.³⁶

El uso de los filtros en Excel proporciona una forma rápida de buscar y analizar los datos en un rango o tabla.²⁸ Al filtrar una lista, se ocultan temporalmente algunos datos para enfocarse en aquellos que necesitamos; existen filtros para un rango de datos, datos de una tabla, de color y uso de criterios avanzados.³³

Existen herramientas informáticas para gestionar el correcto ingreso de datos y obtener una adecuada base de datos; por ejemplo, en estudios de tipo epidemiológico, es útil utilizar el Sistema Epi Info ® el cual es un programa gratuito, especializado y público para la introducción y análisis de datos.³⁷ Mediante el empleo de esta herramienta se pueden procesar cuestionarios o formularios, minimizando la posibilidad de cometer errores al ingresar los datos, definir patrones o codificaciones, seleccionar registros, crear y modificar datos y operaciones durante el análisis, e importar y exportar archivos.³²

1	A	B	C	D	E	F	G	H	I	J	K	L		M		N		O
												SEXO		PROCEDENCIA				
2	ID	CI	SEXO	EDAD	PROC	RMD	RMI	RCD	RCI		CÉDULA	FEMENINO	MASCULINO	URBANA	RURAL			
3	_001	_0151646995	1	22	1	1	1	1	1		_0151646995		X	X				
4	_002	_0151646996	0	24	0	2	2	2	2		_0151646996	X				X		
5	_003	_0151646997	1	26	1	3	3	3	3		_0151646997		X	X				
6	_004	_0151646998	0	28	0	1	1	1	1		_0151646998	X						X
7	_005	_0151646999	1	30	1	2	2	2	2		_0151646999		X	X				
8	_006	_0151647000	0	32	0	3	3	3	3		_0151647000	X						X
9	_007	_0151647001	1	34	1	1	1	1	1		_0151647001		X	X				
10	_008	_0151647002	0	36	0	2	2	2	2		_0151647002	X						X
11	_009	_0151647003	1	38	1	3	3	3	3		_0151647003		X	X				
12	_010	_0151647004	0	40	0	1	1	1	1		_0151647004	X						X
13	_011	_0151647005	1	42	1	2	2	2	2		_0151647005		X	X				
13	_012	_0151647006	0	44	0	3	3	3	3		_0151647006		X	X				

Fig. 4. Formas alternativas de ingreso de datos

En la Figura 4 presentamos dos formas alternativas de ingresar datos como sexo y procedencia. En el lado izquierdo tenemos una base de datos en la cual el sexo está codificado como 1 y 0, significando que 1 corresponde al sexo masculino y 0 al femenino; de la misma forma, la procedencia también está codificada como 1 (urbano) y 0 (rural). Al lado derecho en esta misma figura se puede apreciar la forma incorrecta de introducir los datos: el sexo se reporta en dos columnas (una para masculino y otra para femenino) donde usualmente se coloca X o 1 para realizar el conteo. De esta forma los paquetes estadísticos no reconocen la información. Por eso, lo consideramos como una manera incorrecta de ingresar los datos. Algo similar sucede con la procedencia.

Una vez ingresados los datos en las bases, es imperativo realizar un control de calidad de los mismos.²⁵ En otras palabras, es estrictamente necesario monitorear y valorar la calidad de los datos durante el progreso del estudio.²⁹ El proceso de control de calidad consta de: preparar instrumentos específicos, desarrollar procedimientos para colección de datos y calibrar a los investigadores sobre la toma de datos y su incorporación a la base de datos en proceso de creación. Una vez creada la base de datos, el análisis estadístico de la información allí contenida permitirá responder la pregunta de investigación planteada. El análisis estadístico permite que los datos se conviertan en información organizada y útil.³⁴

ANÁLISIS DE DATOS

Es la etapa en la que se analizan los datos para llegar a conclusiones firmes y robustas. Para lograrlo se puede utilizar estadística descriptiva y/o estadística inferencial.³² La primera se refiere a la adecuada presentación de tablas y gráficos estadísticos para reportar los datos. La segunda por su lado se refiere a la correcta manera de extrapolar los datos de una muestra a una población en general (inferencia mediante intervalos de confianza) y permite la comprobación de las hipótesis.³⁷ Estos temas se desarrollarán en los próximos dos artículos de esta misma serie.

3 CONCLUSIONES

La construcción de la base de datos es una fase crítica en el proceso de investigación científica. La anatomía de la base de datos guarda una relación directa con la pregunta de investigación, el diseño del estudio y el desarrollo de las conclusiones. Una correcta estrategia para la codificación y el manejo de los datos ayuda a evitar errores en el análisis estadístico.

AGRADECIMIENTOS:

Los autores agradecen al Dr. Luis Andrés Yarzabal (Departamento de Investigación, Carrera de Odontología, Universidad Católica de Cuenca) por su apoyo en la edición.

Referencias Bibliográficas

- 1 Tapia L, Palomino M, Lucero Y, Valenzuela R. Pregunta, Hipótesis y Objetivos de una Investigación Clínica. *Revista Médica Clínica las Condes*. 2019 Diciembre; 30(1):29-35.
- 2 Burdiles P, Castro M, Simian D. Planificación y factibilidad de un proyecto de investigación clínica. *Revista Médica Clínica Las Condes*. 2019 Diciembre; 30(1):8-18.
- 3 Zhang Y, Qin G, Zhu Z, Xud W. A novel robust approach for analysis of longitudinal data. *Computational Statistics & Data Analysis*. 2019 Abril; 83-95.
- 4 Caro J, Briggs A, Siebert U, Kuntz K. Modeling good research practices - overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Value in Health*. 2012; 32(5):796-803.
- 5 Villavicencio E, Torracchi E, Pariona MdC, Alvear MC. ¿Cómo plantear las variables de una investigación? Operacionalización de las variables. *Odontología Activa Revista Científica*. 2019 Enero - Abril; 4(1):9-14.
- 6 Editorial. The relevance of good research reporting. *The Lancet Neurology*. 2016 Octubre; 15:1103.
- 7 Feinsinger P. Metodologías de la investigación en ecología aplicada y básica: ¿Cuál estoy siguiendo, y por qué? *Revista Chilena de Historia Natural*. 2013; 86(4):385-402.
- 8 Simian D, Martínez M. Experiencia en la creación de registros con fines de Investigación Clínica. *Revista Médica Clínica Las Condes*. 2019 Noviembre; 30(1):76-82.
- 9 Junginger S. Desing Research and Practice for the Public Good: A Reflection. *The Journal of Desing, Economics, and Innovation*. 2017 Febrero; 3(4):290-302.
- 10 Editorial. Doing Good Research Is Difficult, Doing No Research is More Difficult. *Journal of Investigative Dermatology*. 2017; 137(5):993-995.
- 11 Barker L, Rattihalli R, Field D. How to write a good reseach grant proposal. *Paediatrics and Child Health*. 2015; 26(3):105-109.
- 12 Waddington H, White H, Snilstveit B, García J, Vojkova M, Davies P, et al. How to do a good systematic revieix of effects in international development: a tool kit. *Journal of Development Effectiveness*. 2012 Septiembre ; 4(3):359-387.
- 13 Espinoza E. La hipótesis en la investigación. *Mendive Revista de Educación*. 2018 Enero; 16(1):122-139.
- 14 Davies M, Mosdell N. *Practical Research Methods For Media And Cultural Studies: Making People Count*. Edinburgh University Press;2006.
- 15 Nelson-Megan S. Where do we go from here?En: Carlson J,Johnston, Editors. *Data Information Literacy*. *Data Information Literacy*; 2015;323-345
- 16 Arrom LM, Huguer J, Errando A, Palou B, Palou J. Cómo escribir un artículo original. *Actas Urológicas Españolas*. 2018 FEBrero; 42(9):545-550.
- 17 García M. Tópicos en la planificación del personal médico. *Gaceta Sanitaria*. 2009; 23(5):483-484.

- 18 Manterola C. Estudios Observacionales. Los diseños utilizados con mayor frecuencia en investigación clínica. *Revista Médica Clínica Las Condes*. 2009 Junio; 20(4):539-548.
- 19 Domínguez Y. El análisis de información y las investigaciones cuantitativa y cualitativa. *Revista Cubana Salud Pública*. 2007 Mayo; 33(3):1-8.
- 20 Pita Fernández S, Pértegas Díaz S. Tipos de estudios clínico epidemiológicos. *Unidad de Epidemiología Clínica y Bioestadística*. 2002 Septiembre;1-9.
- 21 Begoña Bermejo F. Estudios experimentales. *Matronas Prof*. 2008; 9(1):15-20.
- 22 Osborg Ose S. Using Excel and Word to Structure Qualitative Data. *Journal of Applied Social Science*. 2016 Septiembre;1-16.
- 23 Castro M. Bioestadística aplicada en investigación clínica: conceptos básicos. *Revista Médica Clínica Las Condes*. 2019 Diciembre; 30(1):50-65.
- 24 Spiece K, Colosi J. Redefining the "Scientific Method". *The American Biology Teacher*. 2000; 62(1):32-40.
- 25 Huebner M, Vach W, Le Cessie S. A systematic approach to initial data analysis is good research practice. *The Journal of Thoracic and Cardiovascular Surgery*. 2016; 151(1):25-7.
- 26 Cavada C. Docimasia de hipótesis. *Revista Chilena Endocrinología Diabetes*. 2009; 2(4)256-257.
- 27 Villavicencio Caparó E, Alvear Córdova MC, Cuenca León K, Calderón Curipoma M, Zhunio Ordoñez K, Webster Carrión F. El tamaño muestral para la tesis ¿Cuántas personas debo encuestar? *Odontología Activa Revista Científica*. 2017; 2(1):59-62.
- 28 Madero R, Pérez E, San José B. Estadística para pediatras (III). Análisis de datos. *Hablemos de...* 2007; 5(1):63-7.
- 29 Bruni R, Bianchi G, Dolente C, Claudio L. Logical Analysis of Data as a tool for the analysis of Probabilistic Discrete Choice Behavior. *Computers & Operations Research*. 2019 June; 106:191-201.
- 30 Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics*. 1996; 5(3):299-314.
- 31 Lejeune M, Lozin V, Lozina I, Ragab A, Yacout S. Recent advances in the theory and practice of Logical Analysis of Data. *European Journal of Operational Research*. 2019 Mayo; 275(1):1-15.
- 32 Khezrimotlagh D, Zhu J, Cook WD, Toloo M. Data envelopment analysis and big data. *European Journal of Operational Research*. 2019 Mayo; 274(3):1047-1054.
- 33 Kawulich BB. La observación participante como método de recolección de datos. *Forum: Qualitative*. 2005 Mayo; 6(2).
- 34 Berlin C, Adams C. *Data Collection and Task Production Ergonomics*: Ubiquity Press; 2019:128-137.
- 35 Pita Fernandez S, Pértegas Díaz S. Investigación cuantitativa y cualitativa. *Unidad de Epidemiología Clínica y Bioestadística*. 2002; 9:76-78.
- 36 Otzen T, Manterola C. Técnicas de Muestreo sobre una Población de Estudio. *International Journal of Morphology*. 2017; 35(1):227-232.
- 37 Rémy NM, Martial TT, Clémentin TD. The prediction of good physicians for prospective diagnosis using data. *Informatics in Medicine Unlocked*. 2018; 12:120-127.

Recibido: 1 de febrero del 2019.

Aceptado: 15 de marzo del 2019.

