

CALIDAD METODOLÓGICA Y SEGUIMIENTO DE ESTÁNDARES DE
PUBLICACIÓN, EN INTERVENCIONES COGNITIVO CONDUCTUALES DE
CEFALEAS TIPO MIGRAÑA Y TENSIONAL: UNA REVISIÓN SISTEMÁTICA 2000-
2012.

Luis Argenis Osorio Ferrer

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE CIENCIAS SOCIALES

FACULTAD DE PSICOLOGIA

Floridablanca, 2015

CALIDAD METODOLÓGICA Y SEGUIMIENTO DE ESTÁNDARES DE
PUBLICACIÓN, EN INTERVENCIONES COGNITIVO CONDUCTUALES DE
CEFALEAS TIPO MIGRAÑA Y TENSIONAL: UNA REVISIÓN SISTEMÁTICA 2000-
2012.

Luis Argenis Osorio Ferrer

Directora de proyecto de grado

Ana Fernanda Uribe Rodríguez

Proyecto de grado para optar al título de Magister en Psicología

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE CIENCIAS SOCIALES

FACULTAD DE PSICOLOGIA

Floridablanca, 2015

Tabla de Contenido

Tabla de contenido

Lista de Tablas

Lista de figuras

Lista de Apéndices

Resumen	1
Abstrac	2
1. Planteamiento del problema y Justificación	3
2. Objetivos	9
2.1. Objetivo general	9
2.2 Objetivos específicos	9
3. Marco Teórico y estado del arte	10
3.1 Estándares de calidad en el contexto investigativo y de publicación.	10
3.1.1 Conceptos de calidad en el contexto investigativo y de publicación	10
3.1.2 Estándares metodológicos y de publicación en el contexto investigativo	12
3.1.3 Instrumentos para evaluar estándares calidad metodológica y de publicación	18
3.1.4 Valoración del cumplimiento de los estándares calidad metodológica y de	
Publicación	22
3.2. Cefalea	26
3.2.1 Conceptos	26
3.2.2 Epidemiología	27
3.2.3 Clasificación	28
3.2.4 Evaluación	29

3.2.5 Intervención	30
4. Método	34
4.1 Diseño	34
4.2 Universo de estudio	34
4.3 Instrumentos	35
4.4 Definición de variables	39
4.5 Procedimiento	39
4.6 Consideraciones éticas	41
5. Resultados	43
5.1 Validación por jueces del Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01).	43
5.2. Calidad metodológica y de publicación de los artículos evaluados	48
5.3 Evolución de la calidad metodológica y de publicación de los artículos evaluados 2000-2012	63
6. Discusión	88
7. Referencias	107
Apéndices	116

Lista de tablas

Tabla 1. Validez de contenido de indicadores eliminados por categoría y subcategoría	37
Tabla 2. Estructura del instrumento Inventario de valoración de estándares Metodológicos y de publicación (INVEMCION-01)	38
Tabla 3. Índice validez de contenido por jueces Polit y Hungler (2000) en función de categorías y subcategorías.	44
Tabla 4. Índice de validez de contenido por jueces Lawshe (1975) en función de Categorías y subcategorías.	45
Tabla 5. Valores y nivel de significación del coeficiente de concordancia de Kendall, para evaluación de pertinencia de los jueces, en función de categorías y subcategorías.	46
Tabla 6. Valores y nivel de significación del coeficiente de concordancia de Kendall, para evaluación de redacción de los ítems por parte de los jueces, en función de categorías y subcategorías.	47

Lista de Figuras

Figura 1 Diagrama de flujo de los artículos a través del proceso investigativo	35
Figura 2. Porcentaje de artículos que cumplieron con el indicador en la categoría Portada, subcategoría título	49
Figura 3. Porcentaje de artículos que cumplieron con el indicador en la categoría Portada, subcategoría resumen.	50
Figura 4. Porcentaje de artículos que cumplieron con el indicador en la categoría Introducción, subcategoría contexto investigativo	51
Figura 5. Porcentaje de artículos que cumplieron con el indicador en la categoría Introducción, subcategoría propósitos de investigación	52
Figura 6. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría Diseño	52
Figura 7. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría participantes.	53
Figura 8. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría Instrumentos	54
Figura 9. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, Subcategoría definición de variables	54
Figura 10. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría procedimiento.	55

Figura 11. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría procesamiento estadístico	56
Figura 12. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría consideraciones éticas.	57
Figura 13. Porcentaje de artículos que cumplieron con el indicador en la categoría Resultados, subcategoría general.	58
Figura 14. Porcentaje de artículos que cumplieron con el indicador en la Categoría Resultados, subcategoría texto.	59
Figura 15. Porcentaje de artículos que cumplieron con el indicador en la categoría Resultados, subcategoría Tablas.	60
Figura 16. Porcentaje de artículos que cumplieron con el indicador en la categoría Discusión, subcategoría análisis.	61
Figura 17. Porcentaje de artículos que cumplieron con el indicador en la categoría Discusión, subcategoría generalización.	62
Figura 18. Porcentaje de artículos que cumplieron con el indicador en la categoría Conclusiones.	63
Figura 19. Porcentaje de artículos que cumplieron con el indicador en la categoría Referencias.	64
Figura 20. Porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría Portada, subcategoría título.	65
Figura 21. Porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría Portada, subcategoría resumen.	67

Figura 22. Porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría Introducción, subcategoría, contexto investigativo.	68
Figura 23. Porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría Introducción, subcategoría propósitos de investigación.	69
Figura 24. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría diseño.	70
Figura 25. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría sujetos.	72
Figura 26. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría instrumentos.	73
Figura 27. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría definición de variables.	74
Figura 28. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría procedimiento.	76
Figura 29. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría procesamiento estadístico.	77
Figura 30. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría consideraciones éticas.	79
Figura 31. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Resultados, subcategoría general.	80
Figura 32. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Resultados, subcategoría texto.	81
Figura 33. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Resultados, subcategoría Tablas.	82

Figura 34. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Discusión, subcategoría análisis.	83
Figura 35. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Discusión, subcategoría generalización.	84
Figura 36. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Discusión, subcategoría implicaciones.	85
Figura 37. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Conclusiones.	86
Figura 38. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Referencias.	87

Lista de Apéndices

Apéndice A. Referencias de artículos seleccionados para la investigación.	116
Apéndice B. Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01).	122
Apéndice C. Referencias de artículos que presentan estándares metodológicos y de publicación.	134
Apéndice D. Carta para evaluación por los jueces.	136
Apéndice E. Información sobre la prueba (resumen de la investigación, ficha técnica y la estructura de la prueba).	138
Apéndice F. Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01) para validación por jueces.	142

RESUMEN GENERAL DE TRABAJO DE GRADO

TITULO: Calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional: una revisión sistemática 2000-2012.

AUTOR(ES): Luis Argenis Osorio Ferrer

FACULTAD: Maestría en psicología

DIRECTOR(A): Ana Fernanda Uribe Rodriguez

RESUMEN

El objetivo del estudio fue determinar la calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012. El diseño utilizado fue la revisión sistemática. Se analizaron 45 artículos de texto completo extraídos de EBSCO y EMBASE. El instrumento fue el Inventario de Valoración de Estándares Metodológicos y de Publicación (INVEMCION-01), el cual fue construido y validado por los investigadores. La validez de contenido fue de 0.89 por los métodos de de Polit y Hungler (2000), y Lawshe (1975). El nivel de concordancia de Kendall en pertinencia y redacción fue significativo al nivel de $p < 0.01$. Los resultados arrojaron altos niveles de cumplimiento en categorías como Conclusiones y Referencias; y subcategorías como título (portada), contexto investigativo (Introducción), General (Resultados), análisis e implicaciones (Discusión); pero se evidenciaron dificultades en mencionar el diseño de investigación, enmascaramiento, instrumentos y descripción de las características de los sujetos. Al valorar la evolución de la calidad metodológica y seguimiento de estándares de publicación no se encontraron asociaciones significativas, entre el periodo evaluado y el porcentaje de cumplimiento de los indicadores al nivel de $p > 0.05$ utilizando el Test Exacto de Fischer. Se motiva a continuar con la validación del instrumento, y promover la difusión y utilización de estándares metodológicos y de publicación como APA (2008) y AERA (2006) por parte de la comunidad académica y de editores de revistas. Se hace necesario el desarrollo de investigaciones similares en otras áreas de aplicación de la Psicología. También se hace un cuestionamiento a la validez de las investigaciones en psicología, y se evidencia los efectos negativos de este hecho a nivel de la práctica clínica y social

PALABRAS CLAVES:

Estándares metodológicos y publicación, cefalea tensional, migraña, Terapia cognitivo conductual

GENERAL SUMMARY OF WORK OF GRADE

TITLE: Methodological quality and monitoring standards publication in cognitive behavioral interventions migraine and tension-type headache: a systematic review from 2000 to 2012.

AUTHOR(S): Luis Argenis Osorio Ferrer

FACULTY: Maestría en psicología

DIRECTOR: Ana Fernanda Uribe Rodríguez

ABSTRACT

The aim of the study was to determine the methodological quality and monitoring standards publication in cognitive behavioral interventions migraine and tension-type headaches in the years 2000-2012. The design used was the systematic review. Were analyzed 45 full-text articles extracted from EMBASE and EBSCO. The instrument was the Inventory Valuation Methodology and Standards Publication (INVEMCION-01), which was constructed and validated by researchers. Content validity was 0.89 Polit methods and Hungler (2000), and Lawshe (1975). The level of concordance of Kendall in relevance and drafting was significant at $p < 0.01$. The results showed high levels of performance in categories such Conclusions and References; and subcategories as title (cover), research context (Introduction), General (Results), analysis and implications (discussion); but difficulties were evident to mention the research design, masking, instruments and description of the characteristics of the subjects. To assess the evolution of methodological quality and monitoring standards of publication, no significant associations were found between the evaluated period and the percentage of compliance with the indicators at the level of $p < 0.05$ using the Fisher exact test. You are encouraged to continue the validation of the instrument, and promote the dissemination and use of methodological standards and publication as APA (2008) and AERA (2006) by the academic community and magazine publishers. The development of similar research in other areas of application of psychology is needed. Besides criticizes the validity of research in psychology and the negative effects of this fac in clinical practice and in society in general.

KEYWORDS:

Methodological standards and publication, tension headache, migraine, cognitive behavioral therapy

1. Planteamiento del Problema y justificación

En esta apartado se presentarán las razones por las cuales se tomaron las investigaciones de intervenciones en cefalea tensional y migraña como uno de los ejes de este estudio; y por otra parte, los motivos que llevaron al desarrollo de una investigación en calidad metodológica y seguimiento de estándares de publicación en el tópico de la cefalea.

Existen varias razones para el desarrollo de investigaciones psicológicas en dolor y específicamente en las cefaleas tipo migraña y tensional. En primera instancia la alta frecuencia con la cual se presenta la cefalea y el impacto negativo en la vida de las personas; se afirma que la cefalea es uno de los problemas más comunes que afecta a las personas (International Association for the Study of Pain [IASP], 2011a; Trillos, 2010; World Health Organization [WHO], 2004). Los datos epidemiológicos muestran que el 50% de las personas presenta algún tipo de dolor de cabeza en un lapso de un año y más del 90% afirma que lo ha padecido alguna vez en su vida. En el caso de la migraña, el promedio de prevalencia en algún momento de la vida de una persona es del 18%; y del 7,7% en niños y adolescentes (IASP, 2011b). La prevalencia estimada de la cefalea en la niñez es de 58,4% (Ferracini, 2014). La IASP (2011a) considera la cefalea un problema tan grave que el año 2012, fue declarado como el año global contra esta problemática.

Covarrubias y Guevara (2006) afirman que el dolor afecta negativamente la capacidad de relación con otras personas, el desarrollo de las actividades diarias y la economía; además influye el núcleo familiar y social debido a los sentimientos generados por el sufrimiento del paciente con dolor. De forma similar, Vallejo (2005) plantea que el dolor crónico produce un alto grado de incapacidad y disminución de actividad. En lo referente a la cefalea tipo migraña y/o tensional existen diferentes fuentes que evidencian sus consecuencias adversas en la calidad de vida, desempeño laboral y en la inversión económica para su tratamiento (IASP, 2011a; Navarro, 2006; WHO, 2004).

La migraña crónica también se ha relacionado con problemas psiquiátricos, altos costos, uso frecuente de los sistemas de salud, e impacta la calidad de vida de la persona que la sufre (Diener, Solbach, Holle & Gaul, 2015). Además la frecuencia de la migraña, tiene relación directa con el daño psicosocial de la persona que la padece (Ruscheweyh, 2014).

En el caso de los niños la migraña genera efectos negativos en las actividades de la vida diaria, en el rendimiento escolar y aumenta el ausentismo (Ferracini, 2014). En un estudio desarrollado en España sobre el número de incapacidades laborales se evidencio que solo la migraña, generaba 58.295 días de incapacidad en periodo de un año. Vicente, Terradillos, Ramírez, Capdevila & López (2014). Desafortunadamente La migraña esta subestimada como enfermedad, lo cual se podría deber entre otros aspectos a que la discapacidad generada no es visible, ni reduce la esperanza de vida (Mas Sala & Arnall, 2015).

También se ha evidenciado problemas relacionados como el uso excesivo y continuado de los analgésicos, que podría generar consecuencias como la nefropatía, alteraciones cardiovasculares, digestivas, hematológicas, cutáneas; como también la presencia de tolerancia, habituación y dependencia; además los pacientes necesitan cada vez mayores dosis para conseguir el mismo efecto analgésico y experimentan cefalea de rebote cuando se suprimen los analgésicos de forma repentina (Fontanillas & Pascual, 2009).

Un segundo aspecto es que el número de psicólogos que participan en la intervención del dolor es limitado (Benito, Nadador, Fernández, Hernández, Ruiz, & Riquelme, 2006; Moix & Casado, 2011). El anterior fenómeno, puede generar que otras disciplinas desarrollen investigaciones e intervenciones con componentes psicológicos; este es el caso de un estudio realizado por LeForta, Gray, Rowatc y Jeansd (1998) en donde se emplearon enfermeras, para implementar un programa psico-educativo grupal que incluía técnicas cognitivo conductuales para el automanejo de dolor crónico, los resultados mostraron una mejoría significativa de los sujetos al terminar la intervención. Sin embargo, los autores reconocieron como falencias la ausencia de medidas de seguimiento y que los participantes que abandonaron el estudio se caracterizaban por tener mayores niveles de dolor y depresión; como también una funcionalidad más limitada, por lo cual recomendaban intervenciones más especializadas. Una posible consecuencia de estos hechos, sería una restringida realización de investigaciones por parte de psicólogos en este campo.

Un tercer aspecto, son las dificultades observadas tanto en la metodología como en el seguimiento de estándares de publicación en las investigaciones que involucran procesos de intervención en cefalea. Estudios metaanalíticos de intervenciones médicas, han mostrado dificultades en cuanto a la comparabilidad de los grupos de estudio al inicio de la investigación, el proceso de asignación al azar, los test estadísticos y los criterios de inclusión y exclusión de la muestra, entre otros (Bronfort et al., 2008). En el caso de intervenciones psicológicas, Nestoriuc, Martin, Rief y Andrasik (2008) plantean problemas metodológicos como la utilización de muestras pequeñas, dificultades en la descripción del tratamiento, de las características de los pacientes; así como también la utilización de sistemas de diagnóstico inadecuados.

Por su parte, Verhagen, Damen, Berger, Passchier, y Koes (2009) desarrollaron una investigación de tipo metaanalítico, dirigida a evaluar la eficacia de tratamientos conductuales en pacientes con cefalea tensional. El análisis de las investigaciones que conformaron la muestra final, evidenció problemas en el “enmascaramiento” de pacientes y terapeutas, utilización de muestras pequeñas, falta de claridad sobre el diagnóstico de los participantes y deficiencias en la presentación de datos. Los autores recomiendan la realización de estudios con muestras de mayor tamaño, mayor rigurosidad metodológica, optimización de criterios diagnóstico y la utilización de la declaración Consolidated Standards of Reporting Trials (CONSORT), la cual consiste en una serie de lineamientos para la publicación de pruebas controladas aleatorias (Verhagen et al, 2009).

Acevedo, López & Cárdenas (2014) exponen que en los estudios psicológicos los tamaños de muestra son bajos con respecto investigaciones en Biomedicina o genética. También plantean que la falta de confiabilidad en las investigaciones de diferentes disciplinas, parte de que los resultados están basados en análisis e ideas incorrectas de interpretación, problemas en el diseño, en el análisis de procedimientos estadísticos y en el bajo nivel de replicación. Los autores agregan que esto ha generado un deterioro de la imagen de las ciencias del comportamiento sustentada en sesgos dentro de las investigaciones, en la obtención de datos en la publicación y encuentro de fraudes (Acevedo, López & Cárdenas, 2014).

Una posible consecuencia de los problemas metodológicos referidos anteriormente, es la limitada claridad sobre los efectos de las intervenciones psicológicas en el campo del dolor y la cefalea. Mientras algunos autores defienden su efectividad (Moix & Casado, 2011; Vallejo, 2008); otros plantean dudas en este aspecto (Barton & Blanchard, 2001; Navarro, 2006; Silver, 2007; Verhagen et al., 2009). Caso similar ocurre con la terapia de bioretroalimentación en la cual algunos autores defienden sus efectos positivos (Kapitz, Passie, Bernateck, & Karst, 2010; Vallejo, 2008); sin embargo también se han planteado dudas sobre su efectividad (Barton & Blanchard, 2001 ; Navarro, 2006; Verhagen et al., 2009).

Una alternativa de solución para superar las dificultades metodológicas planteadas es la optimización de los estándares de publicación. Un aporte importante en este aspecto se puede evidenciar en el campo de las ciencias médicas, en donde se creó el CONSORT; que como se decía anteriormente, son una serie de lineamientos para la publicación de pruebas controladas aleatorias (Moher, Schulz, & Altman, citados por Boutron, Moher, Altman, Schulz, & Ravaud, 2008). Una de las motivaciones para el desarrollo del CONSORT es que el seguimiento de sus pautas influya de manera indirecta en la calidad del diseño y ejecución de las investigaciones (Schulz, Altman, Moher & the CONSORT group, 2010). Se afirma que los problemas o dificultades en las publicaciones siguen siendo comunes (Boutron et al., 2008) a pesar de que un número considerable de revistas científicas han asumido las recomendaciones del CONSORT (Schulz et al., 2010). Es importante agregar que estos lineamientos, se han ampliado a ensayos o experimentos no farmacológicos dentro de los cuales se encuentran las intervenciones conductuales (Boutron et al., 2008); experimentos con un solo sujeto (Shamseer et al. (2015) y reporte de resultados de pacientes Calvert et al. (2013).

Debido a la existencia de gran variedad de protocolos para los ensayos clínicos, que varían en contenido y calidad, se han desarrollado el SPIRIT (Standard Protocol Items: recommendations for interventional trials) (Chan et al. (2013); lo cual equivaldría a lineamientos estándar para ensayos clínicos controlados. Este documento pretende que el seguimiento de los lineamientos mejorara la transparencia y la presentación completa de los datos, lo cual beneficiara a todos los actores comprometidos con los procesos investigativos como investigadores, pacientes, patrocinadores entre otros.

Desde el contexto de la Psicología se han evidenciado diferentes esfuerzos en esta línea; uno de los más reconocidos, es el liderado por la Asociación Psicológica Americana (APA), la cual ha desarrollado una serie de normas o estándares para la presentación escrita de trabajos (APA, 2010). Además, la APA ha incursionado en el establecimiento de estándares específicos para la publicación de diferentes tipos de estudios, como los de tipo experimental y metaanálisis; afirmando que la utilización de estos estándares permite valorar las fortalezas y debilidades de una investigación, mejora la forma en que las investigaciones se implementan, facilita la replicación y la realización de estudios metaanalíticos. También recomiendan la realización de trabajos dirigidos a la revisión de los efectos que tienen los estándares en la publicación de artículos (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008).

Sobre el establecimiento de estándares de publicación, existen diversos autores que han generado aportes para diferentes tipos de estudios, como por ejemplo, normas para revisión de artículos en Ciencias de la Salud (Bobenrieth, 2002); lineamientos para la revisión de estudios instrumentales (Carretero & Pérez, 2007); criterios para la preparación y revisión de artículos de investigación experimental y cuasi-experimental (Ramos, Moreno, Valdés, & Catena, 2008); y preparación y escritura de artículos de revisión bibliográfica en psicología (Fernández & Buela, 2009).

Acevedo, López & Cárdenas (2014) afirman que en la publicación de las investigaciones se deben reportar los datos positivos, negativos; como también facilitar la realización de metaanálisis por medio del reporte de aspectos como intervalos de confianza, tamaño de la muestra entre otros.

Una forma de saber si estos lineamientos y/o recomendaciones se están cumpliendo de una manera confiable, es por medio de una revisión sistemática de la literatura; al respecto, Marín, Sánchez y López (2009) destacan la importancia de este tipo de estudios, afirmando que debido al aumento de la investigación, la revisión y síntesis de información, permite conocer e integrar el estado de conocimiento de un tema particular. Un ejemplo de estudios en esta línea fue

desarrollado en el campo de la medicina por Tao et al (2011), en el cual realizaron una revisión sistemática en investigaciones, que tenían como objetivo hacer revisiones bibliográficas sobre la aplicación de anestesia, tomando como base las declaraciones QUOROM (Quality of Reporting of Meta-analyses) y PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) que son una serie de lineamientos para la publicación de revisiones sistemáticas y metaanálisis; los autores encontraron que la calidad de las revisiones no es óptima y que la adherencia a los lineamientos de publicación mejora la utilidad de las revisiones sistemáticas y su publicación; además recomienda el desarrollo de más investigaciones en ésta línea.

En el campo psicológico, Vera y Mustaca (2006) implementaron un estudio bibliométrico de publicaciones científicas en psicología clínica, en revistas de Argentina y Chile, encontrando que en estos países el mayor número de investigaciones empíricas se desarrollaban con modelos cognitivos conductuales, y la rigurosidad metodológica de estas investigaciones eran superiores a las que se utilizaban otros enfoques teóricos. En concordancia con esto, se encuentra que en la actualidad el más alto porcentaje de las terapias con apoyo empírico reconocidas por la APA, pertenecen al enfoque cognitivo conductual (Mustaca, 2011).

Según lo tratado, se puede concluir que es de gran importancia la realización de investigaciones en el área del dolor en general y la cefalea en particular, debido a que son problemáticas muy comunes en el ámbito de la salud física, y generan consecuencias negativas importantes a nivel personal, familiar y social. Por otra existe falta de claridad sobre los efectos de las intervenciones psicológicas ocasionados probablemente por los problemas metodológicos existentes en los estudios; lo cual incide negativamente en la validez de los resultados y genera dudas sobre los efectos obtenidos con las intervenciones.

También, se evidencia que una forma de afrontar y superar estas dificultades, es asumiendo una tendencia de trabajo dirigida al establecimiento y seguimiento de estándares de publicación, que tiene como fin optimizar la calidad del diseño y ejecución de las investigaciones. La literatura también muestra que las terapias cognitivo conductuales son las más avanzadas en rigurosidad metodológica y destaca la importancia de la realización de revisiones sistemáticas de la literatura; como también de estudios dirigidos a valorar el impacto de los estándares de publicación en la literatura científica. Por último, es importante destacar

que ya se han empezado a desarrollar estudios en esta línea, que han arrojado información importante sobre la rigurosidad y seguimiento de estándares de las publicaciones; pero estos Han partido principalmente del área de la medicina.

Es por lo expuesto anteriormente, que el presente trabajo pretende generar un aporte metodológico en el diseño y validación de un instrumento que permita valorar la calidad metodológica y de lineamientos de publicación en estudios experimentales; además se dirige a **evaluar** los niveles en calidad metodológica y seguimiento de estándares de publicación, en artículos que presenten intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012 y conocer si ha existido evolución en cuanto a calidad metodológica y de publicación en este periodo de tiempo. Por lo anterior, se plantean los siguientes problemas e hipótesis:

¿Cuál es el nivel de calidad metodológica y seguimiento de estándares de publicación, en artículos que presenten intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012?

Al realizar una valoración de la calidad metodológica y seguimiento de estándares de publicación estándares de publicación, en artículos que presenten intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012, se encontrará que se presentan bajos porcentajes de cumplimiento de estas variables.

¿Cuál ha sido la evolución de la calidad metodológica y seguimiento de estándares de publicación, en artículos que presenten intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012?

Al realizar una valoración de los cambios en calidad metodológica y seguimiento de estándares de publicación, en artículos que presenten intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional, en diferentes periodos de tiempo comprendidos entre los años 2000-2012, se encontrará que presentan porcentajes de cumplimiento similares de estas variables, en los diferentes periodos evaluados.

2. Objetivos

2.1 Objetivo General

Determinar la calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012, por medio de una revisión sistemática.

2.2 Objetivos específicos

Diseñar y validar por jueces un instrumento para valoración de estándares metodológicos y de publicación.

Evaluar los niveles en calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012.

Conocer la evolución de la calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012.

3 Marco teórico y estado del arte

Esta sección está dirigida a presentar los elementos teóricos que dan sustento a la investigación; para lo cual se ha dividido en dos grandes apartados. El primero contiene información sobre los conceptos de calidad en el contexto investigativo y de publicación, los estándares que se han desarrollado en esta área, los instrumentos que se han creado para valorar los estándares y las investigaciones en las que se han utilizado, lo cual respalda teóricamente el interés de esta investigación en la valoración de la calidad metodológica y seguimiento de estándares de publicación

En la segunda parte se definirán los conceptos básicos relacionados con el dolor y la cefalea, su epidemiología, clasificación y la forma en que han sido evaluados e intervenidos, lo cual permite dar un contexto teórico a la temática desarrollada por las investigaciones estudiadas, específicamente en lo referente a la intervención en cefalea y migraña tensional.

3.1 Estándares de calidad en el contexto investigativo y de publicación.

Como se mencionaba anteriormente, en este apartado se presentaran los conceptos de calidad en el contexto investigativo y de publicación, los estándares que se han desarrollado; como también los instrumentos que se han creado para valorar los estándares, y las investigaciones en las que se han aplicado.

3.1.1 Conceptos de calidad en el contexto investigativo y de publicación

En la literatura científica se hace referencia constantemente al concepto de calidad; pero existen pocos autores que se arriesgan a presentar una definición o las características esenciales de esta. Un primer abordaje, está relacionado con el término calidad científica de la investigación, la cual se logra cuando se “evita el sesgo y los hallazgos reflejan o muestran la verdad” (Oxman & Guyatt, 1991 p. 1272). Estos mismos autores agregan que excluyeron de los criterios la relevancia de la pregunta y la calidad de la presentación.

Un concepto de calidad estrechamente relacionado con el anterior es el de calidad de la investigación, Buela-Casal (2003), afirma que en muchos países la calidad de las investigaciones, se determina por medio de índices bibliométricos; entre los cuales se encuentran el factor de impacto y el factor de prestigio. El autor sostiene que estos

indicadores, deberían ser utilizados mejor como datos de difusión, más que de calidad, debido a que están basados en número de citaciones realizadas de los artículos de las revistas. Como aporte a esta dificultad, presenta a consideración 20 criterios para evaluación de trabajos científicos como relevancia de los resultados, validez de la investigación, calidad de los instrumentos utilizados entre otros.

Otra definición de calidad de la investigación es “la capacidad para para lograr conocimientos verdaderos” (Gutiérrez & Jiménez, 2011 p 366). Las autoras complementan esta definición, diciendo que la calidad de la investigación que se desarrolla en un contexto determinado es “el grado de certeza que tiene el conocimiento, con respecto a la realidad que requiere la sociedad” (Gutiérrez & Jiménez, 2011 p. 366). Esta última, definición es de vital importancia, teniendo en cuenta que el desarrollo de la investigación, tiene que estar alineado con los requerimientos de la sociedad y en beneficio de esta.

Gutiérrez & Jiménez (2011), aseveran que con alta frecuencia se presenta una baja calidad de las investigaciones, este fenómeno involucra a todos los continentes y genera preocupación en diversas instituciones y organizaciones en el ámbito mundial; entre las razones posibles está el desconocimiento de los autores sobre algunos estándares investigativos y de las personas que publican los resultados.

Un tercer concepto de gran relevancia para la presente investigación es el de calidad metodológica, la cual ha sido definido como “la probabilidad de que el diseño del experimento genere resultados no sesgados, que sean suficientemente precisos y permitan la aplicación en la práctica clínica” (Verhagen, de Vet, de Bie, Boers, van den Brandt, 2001, p. 651). Un aspecto a destacar en esta definición, es que la calidad metodológica no se toma como un fin en sí mismo, sino que resalta la importancia de que los resultados tengan la posibilidad de ser llevados a la práctica clínica cotidiana, y de esta manera generar un beneficio para la sociedad.

Haciendo referencia también a calidad metodológica, Jüni, Altman & Egger, (2001), exponen que la calidad de los ensayos controlados, es muy importante para la realización de revisiones sistemáticas, pues si el material de base se encuentra distorsionado, las Conclusiones de las revisiones sistemáticas pierden confianza. Estos autores colocan como ejemplo un metaanálisis, en donde encontraron que los estudios que presentaban dificultades

para realizar un adecuado enmascaramiento, normalmente inflaban los efectos del tratamiento; pero al realizar un análisis más detallado, se encontraba que las intervenciones generadas por estudios en donde había problemas metodológicos, eran menos benéficas que las surgidas de investigaciones con una adecuada metodología.

Otra vertiente relacionada con la calidad en el contexto investigativo, es el de calidad de publicación; como se planteaba anteriormente, existen pocas definiciones o aproximaciones al concepto. Un primer acercamiento a su definición, expone que la calidad metodológica de una investigación esta entrelazada o estrechamente relacionada con la calidad del reporte; la cual se define como “la medida en la cual la publicación presenta información sobre el diseño, la forma en que se desarrolló el estudio, y el análisis que se hace de este (Jüni, Witschi, Bloch, Egger, 1999, citado en Jüni, Altman & Egger, 2001 p. 44). Otra propuesta para este término puede ser extractada de lo propuesto por el International Committee of Medical Journal Editors [ICMJE] (2013), esta institución se ha encargado de generar recomendaciones de publicación en revistas del área médica; su objetivo es facilitar la creación y publicación de artículos exactos, claros y no sesgados.

En esta misma línea las recomendaciones propuestas por CONSORT (2014) pretenden que los autores, escriban los artículos de una manera completa y transparente; con el fin de facilitar la interpretación y valoración crítica por parte de los lectores.

3.1.2 Estándares metodológicos y de publicación en el contexto investigativo

Como respuesta a las inquietudes propuestas anteriormente y para generar una materialización de estas definiciones, se ha desarrollado y fortalecido una tendencia en los últimos años, que consiste en el establecimiento de estándares metodológicos y de publicación en el contexto investigativo, los cuales son una serie de lineamientos o normas, que tiene como fin homogeneizar la forma en que son presentados los artículos, y de esta manera mejorar la calidad de las publicaciones e indirectamente la metodología con la cual son presentados.

En el momento existen diferentes fuentes de trabajo que apuntan a la estructuración e implementación de estándares de publicación científica; una de éstas es el Consolidated Standards of Reporting Trials (CONSORT), el cual se refiere a normas o lineamientos que

tienen como objetivo mejorar la publicación de pruebas controladas aleatorias; el primer documento fue publicado por primera vez en 1996 (Moher, et al., 2001) citados por Boutron et al. (2008); la última revisión de esta guía se desarrolló en el 2010 (Schulz et al., 2010). Este estándar está compuesto por una lista de chequeo de 25 ítems que están distribuidos en seis categorías en título y resumen, introducción, método, resultados, discusión e información adicional. Un ejemplo de estos criterios en la categoría de introducción es hacer explícito los objetivos y/o hipótesis de investigación. De forma complementaria se han desarrollado otros estándares como el de resúmenes de investigación de los ensayos clínicos controlados, debido a que este apartado es básico para la evaluación inicial de un experimento, y para tomar la decisión de hacer o no una lectura completa del texto (Hopewell, et al., 2008). También existe una versión ensayos o experimentos no farmacológicos (Boutron et al., 2008), dentro de los cuales se pueden incluir las intervenciones conductuales; experimentos con un solo sujeto (Shamseer et al. (2015) y reporte de resultados de pacientes Calvert et al. (2013).

En este punto es pertinente aclarar, que las pruebas controladas aleatorias, también se han denominado ensayos clínicos controlados aleatorizados (ECCA), y se caracterizan principalmente por la posibilidad de establecer relaciones causa efecto, debido a la asignación aleatoria de los sujetos a los grupos, la existencia de un grupo control y el ocultamiento sobre la identidad de los sujetos que pertenecen al grupo control o experimental (Lazcano, Salazar, Gutiérrez, Ángeles, Hernández, & Viramontes, 2004).

El estándar CONSORT, pretende que los autores sigan sus lineamientos, para promover la realización de reportes claros, transparentes y completos, lo cual facilita la lectura de los informes. Se aclara que no incluyen recomendaciones para diseñar, implementar o analizar los estudios; sino que se centra en el reporte de lo que se hizo y de los resultados arrojados (Schulz et al., 2010). A su vez afirman que el establecimiento del CONSORT, parte de diferentes motivaciones; entre estas, contar con una información clara y completa de la metodología y de los resultados de los estudios, y optimizar la publicación de las pruebas controladas aleatorias. Lo anterior facilita la evaluación de las investigaciones, el reconocimiento de las fallas existentes, y mejora la calidad en el diseño y ejecución de las investigaciones.

El estándar CONSORT, está basado en evidencia empírica; por lo cual los cambios o modificaciones están soportadas en un grupo de trabajo, conformado por investigadores clínicos, estadísticos, epidemiólogos y editores biomédicos de diversos países, los cuales se dedican a explorar la literatura existente con el fin de actualizar los lineamientos establecidos. Con el paso del tiempo un número considerable de revistas científicas han asumido las recomendaciones del CONSORT (Schulz, et al., 2010); Sin embargo, aún es común encontrar problemas o dificultades en las publicaciones (Boutron et al., 2008).

Debido a la existencia de gran variedad de protocolos para los ensayos clínicos, que varían en contenido y calidad, se han desarrollado el SPIRIT (Standard Protocol items: recommendations for interventional trials) (Chan et al. (2013); lo cual equivaldría a lineamientos estándar para ensayos clínicos controlados. El documento incluye una lista de chequeo de 33 ítems que están estructurados en cuatro categorías información administrativa, introducción, método y apéndices. Este documento pretende que el seguimiento de los lineamientos mejore la transparencia y la presentación completa de los datos, lo cual beneficiara a todos los actores comprometidos con los procesos investigativos como investigadores, pacientes, patrocinadores entre otros.

Otro estándar, ha partido del International Committee of Medical Journal Editors [ICMJE], el cual ha sido construido por un grupo de editores de revistas ; que tiene como fin generar recomendaciones para implementar, reportar, editar y publicar trabajos académicos en revistas especializadas en el área de la medicina (ICMJE, 2014). Se afirma que el primer antecedente de este estándar, fue el documento denominado Uniform Requirements for Manuscripts Submitted to Biomedical Journals (URMs), el cual fue publicado en 1978. El documento original ha experimentado revisiones en 1999, 2000, 2001, 2003, 2010 y la última edición se presentó en 2013 (ICMJE, 2014). Propone 11 lineamientos dirigidos a estandarizar los apartados de página inicial o del título, resumen, introducción, método, resultados, discusión, referencias, tablas, figuras, unidades de medida y símbolos.

En el área de la medicina, también se han realizado aportes por autores no vinculados a instituciones. García & Sánchez (2011), enfatizan en la necesidad de abordar la literatura científica desde una perspectiva crítica y propone una guía de diez puntos para leer un ensayo

clínico aleatorizado; entre estos se encuentran la aleatorización de los sujetos, la rigurosidad en la selección de la muestra y la precisión de los resultados.

Ya desde el contexto de la Psicología se han evidenciado diferentes esfuerzos en esta línea; uno de los más reconocidos es el liderado por la Asociación Psicológica Americana (APA), la cual ha generado documentos para estandarizar la forma de escritura de los científicos sociales y conductuales en estudios de caso, empíricos, teóricos entre otros (APA, 2014). La primera publicación de un manual realizado por la APA, para presentar estos estándares, fue en 1952 (APA Publications and Communications Board Working Group on Journal Article Reporting Standards [JARS], 2008). En la actualidad ha presentado seis ediciones, la última fue publicada en 2009 en idioma inglés y en 2010 en español.

En los últimos años la APA, ha generado un interés especial por el contenido de los artículos científicos, que son publicados en sus revistas. Esta motivación parte en gran medida por la necesidad de que las intervenciones actuales tengan respaldo en evidencia científica; lo cual implica el tener claridad sobre la forma en que son desarrolladas las investigaciones y sobre sus resultados. Se afirma, que debido a la inminente necesidad de generar una nueva edición del manual, el Consejo de Publicaciones y Comunicaciones de la APA, conformó el grupo de trabajo en estándares de reporte para artículos de revistas (JARS). Dentro de las Conclusiones a las que llegó el grupo, después de su proceso investigativo, se encuentra el hecho de que existe una necesidad a ser más detallados en el reporte de las investigaciones, es decir, ser más específicos en presentar la forma en que se desarrolló la intervención y cuáles fueron los resultados encontrados. Lo anterior parte de dos fuentes, la primera es la nueva tendencia a utilizar intervenciones basadas en la evidencia, y la segunda es la importancia que se le ha empezado a adjudicar a los estudios metaanalíticos, los cuales permiten entre otros aspectos la generación de hipótesis y guiar los nuevos problemas a estudiar; y lo cual no se puede lograr sin un reporte completo de los métodos y resultados de los estudios (JARS, 2008).

Otro punto de gran relevancia tratado por este grupo, fue el dilema entre la transparencia del reporte y las limitaciones de espacio impuestas por los medios impresos. El grupo JARS comenta que este problema tiende a desaparecer, debido a que las revistas poseen sitios WEB disponibles para incluir material adicional a los artículos; esto llevó al grupo JARS a

recomendar el uso y estandarización de sitios WEB, para ubicar información adicional a la de los estudios. Es posible que cierta información requerida en los estándares de reporte no aparezca en el artículo publicado; pero puede ser localizada en una página WEB. Un ejemplo de esta información puede ser las instrucciones, el diagrama de flujo de los participantes, Tablas de gran tamaño, audios, video clips, programas de computador entre otros. Es necesario aclarar que este material también debe ser revisado en conjunto con el artículo original (JARS, 2008).

Entre los beneficios adicionales, a los expuestos anteriormente por la utilización de los estándares, es que facilitarían el diálogo interdisciplinario y la réplica de los estudios. Sin embargo, también es importante llamar la atención sobre las posibles dificultades de asumir estas normativas; entre estas se encuentran la excesiva estandarización, lo cual podría llevar a reportar información irrelevante o desechar información importante que no es solicitada en los estándares; baja adecuación de las normativas a algunos métodos utilizados, problemas específicos o áreas de conocimiento particulares; también llama la atención sobre la obsolescencia de estos estándares, debido a los rápidos cambios o avances en aspectos metodológicos y en las temáticas que son abordadas por los investigadores (JARS, 2008).

Dentro de los estándares generados por el grupo JARS están los dirigidos a artículos generales, estudios experimentales con y sin asignación al azar y metaanálisis. Este estándar está compuesto por una lista de chequeo de 44 ítems que están distribuidos en seis categorías en título y resumen, introducción, método, resultados y discusión. Afirman que la utilización de estos estándares permiten valorar las fortalezas y debilidades de una investigación, mejora la forma en que se implementan, facilita la replicación y la realización de estudios metaanalíticos; también recomiendan la realización de trabajos dirigidos a la revisión de los efectos que tiene la aplicación de los estándares en la publicación de los artículos (JARS, 2008).

Otra asociación de gran relevancia en el contexto de la Psicología, es la American Educational Research Association [AERA], la cual también ha presentado unos lineamientos para realizar reportes de investigación en educación, enmarcada en estudios empíricos dentro de las ciencias sociales. El objetivo principal de estos estándares es apoyar a los investigadores en la preparación de los artículos; a los editores y revisores para juzgar la

posibilidad de publicación de un artículo, y a los lectores para aprender a construir una publicación. Afirman que los estándares son una guía sobre la información esencial que debe tener un escrito, para comprender la naturaleza de la investigación y la importancia de los resultados (AERA, 2006).

Estos estándares están compuestos por 48 ítems divididos en ocho categorías: formulación del problema, diseño y lógica del estudio, fuentes de evidencia, medición y clasificación, análisis e interpretación, generalización, ética del reporte y título, resumen y encabezados. Los autores agregan que existen dos principios que subyacen a estas normativas, el primero es la suficiencia, consistente en que debe existir evidencia adecuada para justificar los resultados y Conclusiones; y la segunda, es la transparencia, referida a que en el reporte debe hacerse explícita la lógica de la investigación y las actividades que conducen desde el problema de investigación, siguiendo con la recolección y análisis de datos; hasta los resultados articulados o condensados del estudio (AERA, 2006).

Fuera de estas instituciones, existen diversos autores en el campo de la psicología que han generado aportes relacionados con estándares de publicación y con la calidad metodológica, uno de los primeros es Seligman (1995), quien propone ocho características, que debe poseer un estudio experimental ideal en el campo clínico entre las cuales se encuentran la asignación al azar, un control riguroso, la utilización del enmascaramiento entre otros. También se encuentra Bobenrieth (2002), quien afirma que la meta final de la realización de investigaciones en ciencias de la salud, es que el conocimiento derivado de este proceso, sea utilizado en la práctica profesional; es por esto que las investigaciones deben ser evaluadas en cuanto a validez, relevancia, novedad y utilidad. Este mismo autor, propone la existencia de diversos mitos con respecto a las publicaciones científicas; el primero consiste en pensar que los artículos científicos publicados en revistas prestigiosas garantiza su validez, lo cual no es tan cierto en la realidad. Al respecto dice que al hacer una lectura detallada de las publicaciones, se evidencian dificultades de validez; debido a que presentan problemas conceptuales y metodológicos; sumados a esto, están las dificultades en redacción, precisión y orden o secuencia lógica. El segundo mito es considerar que un respaldo suficiente para la calidad de una publicación, es el éxito profesional del autor. La realidad es que la investigación requiere formación específica, diferente a la requerida en el campo particular

de desempeño de la persona (Bobenrieth, 2002). En tercera instancia, se plantea el mito de que personas que han presentado artículos de “buen nivel”, replicarán este mismo patrón en las siguientes publicaciones. Con respecto a este tópico, se puede pensar que existen mayores probabilidades de que el autor siga haciendo buenas publicaciones; pero no hay garantía de que esto ocurra, debido a que cada trabajo requiere de unas condiciones particulares de metodología, estructura, estilo entre otros. El cuarto mito es considerar que un investigador, por el hecho de ejercer esta profesión es imparcial, lo cual debe ser tomado con precaución, debido a que ellos son seres humanos y como tal están sujetos a prejuicios, que podrían viciar los resultados de una investigación (Bobenrieth, 2002).

Con el fin de optimizar la valoración de los estudios, Bobenrieth (2002) ha propuesto que los artículos experimenten una lectura crítica, la cual es definida como una evaluación objetiva de las fortalezas y las debilidades de una investigación, para lo cual el autor propone 138 pautas que cubren desde las partes preliminares como el título, el resumen, hasta las Conclusiones y la bibliografía.

Otros autores que trabajan en la misma línea han sido Ramos et al. (2008) quienes afirman que la publicación de conocimiento científico, se sustenta en criterios bien organizados para la preparación y revisión de artículos de investigación. Los autores expresan que en la revisión de artículos por parte de pares, se han encontrado diversas dificultades; ante lo cual plantean una propuesta de realización y revisión de artículos, que se encuentra sustentada en dos criterios, la estructura y la validez; con respecto a la estructura destacan tres aspectos: el desarrollo teórico, que incluye la introducción y la discusión; el método y los resultados. En lo concerniente a la validez expone que la parte conceptual del artículo está relacionada con la validez de constructo, el método involucraría la validez interna y externa, y los resultados y discusión estarían sustentados en la validez estadística. Teniendo las dificultades expresadas por los autores con respecto a la revisión de artículos presentan 97 criterios para la preparación y revisión de estudios de investigación experimental y cuasi-experimental, que evalúan, entre otros aspectos, la calidad del artículo en general y la calidad técnico metodológica (Ramos et al., 2008).

En concordancia con lo anterior, Fernández y Buela (2009), proponen 44 normas para redactar y realizar análisis crítico de artículos de revisión bibliográfica divididas en ocho

categorías que incluyen aspectos como organización del artículo, comprensión histórica del tema, revisión bibliográfica entre otros. Por su parte, Carretero & Pérez (2007), teniendo en cuenta algunos aspectos como la utilización común de test en la investigación psicológica, la influencia directa del uso de un instrumento en los resultados de una investigación, la limitada información que se da en las publicaciones sobre las pruebas y las dudas sobre la idoneidad de las mismas entre otros, sugiere 82 normas que deben ser tenidas en cuenta en el desarrollo y revisión de estudios instrumentales. Estas normas, se encuentran divididas en seis categorías justificación del estudio, delimitación conceptual del constructo a evaluar, construcción y evaluación cualitativa de los ítems, análisis estadístico de los ítems, estudio de la dimensionalidad del instrumento y estimación de la fiabilidad.

3.1.3 Instrumentos para evaluar estándares calidad metodológica y de publicación

Como consecuencia de la preocupación en el contexto investigativo sobre la calidad de los estudios y por otra parte la generación de normas y estándares de calidad metodológica y de publicación, se han empezado a generar una serie de instrumentos que tiene como fin la medición de estos estándares. Un primer estudio relacionado con este aspecto es el desarrollado por Armijo et al. (2008), quienes realizaron una revisión sistemática, que tenía entre sus fines condensar información referente a las escalas utilizadas para evaluar la calidad de los ensayos controlados en la investigación en salud, de los estudios consultados se extrajeron 21 escalas y sus respectivas modificaciones como es el caso de Jadad, Maastricht, lista Delphi y PEDro entre otras. En el análisis de los estudios, se pudo concluir que en la mayoría de los casos el proceso de construcción de escalas no fue riguroso, y estas difieren en muchos aspectos como el área de interés, número de ítems, tipo de ítems entre otros. Además se evidenció que se han hecho adaptaciones de las escalas para evaluar otras áreas de interés, pero las escalas adaptadas no han experimentado una nueva validación, lo cual demerita las características psicométricas del nuevo instrumento. Desafortunadamente estas escalas han sido utilizadas para hacer valoraciones de calidad metodológica en otras investigaciones; lo anterior, sugiere que cuando se utiliza una escala para evaluar la calidad metodológica de un artículo, se deben tener en cuenta sus limitaciones al momento de hacer las interpretaciones. Otro resultado arrojado por el estudio, fue que los tópicos para evaluar calidad metodológica se centraban en aleatorización, enmascaramiento y el cálculo del tamaño de la muestra (Armijo et al. ,2008).

Losada et al. (2009), crearon un instrumento para evaluar calidad metodológica de los estudios de pronóstico, la escala está compuesta por seis categorías y 25 ítems, y su confiabilidad calculada por medio del alfa de Cronbach es de 0,89. Entre los aspectos evaluados por el instrumento se encuentran el diseño, las variables de confusión y el tamaño de la muestra.

Por su parte Burgos, Manterola y Sanhueza (2011) crearon un instrumento para evaluar calidad metodológica de investigaciones sobre pruebas diagnósticas. La escala final quedo compuesta por nueve ítems distribuidos en tres categorías, que hacían referencia al tipo de diseño de la investigación, tamaño de la población y la descripción de la metodología. La confiabilidad de la prueba se analizó por medio del coeficiente de correlación intraclase, arrojando para la escala general un índice de 0.96. Los autores afirman que ha aumentado la necesidad de poseer instrumentos válidos y confiables para diseñar una investigación, escribir un artículo o evaluar la calidad metodológica de las publicaciones, y estos instrumentos deben ser comprensibles y fáciles de aplicar, para que se puedan utilizar de una forma masiva.

Coherente con lo expuesto anteriormente, Cascaes, Valdivia, da Rosa, Barbosa & da Silva (2013) afirman que la evaluación de la calidad de las investigaciones es esencial en la producción y selección de la literatura científica en el campo de la salud. Desafortunadamente se presentan diversas dificultades en la calidad de las investigaciones como incongruencias entre los objetivos, el procedimiento y los resultados, o problemas de tipo ético. Dentro de este contexto, los autores realizaron una revisión sistemática para identificar escalas y listas de verificación de evaluación de calidad metodológica en artículos científicos en salud. En el estudio se identificaron 14 escalas y 11 listas de verificación. Entre las escalas observadas se encuentran Jadad, Maastrich, Single-Case Experimental Design Scale (SCED), PEDro entre otras. En lo referente a listas de verificación se ubicaron instrumentos como la Lista Delphi, lista de Maastricht Amsterdam (MAL), registro MOOSE y Lista Andrew, entre otras. Las principales Conclusiones del estudio es que en general un número limitado de escalas y listas de verificación presentan de manera explícita las propiedades psicométricas, en segunda instancia se han realizado modificaciones de las escalas originales; pero los instrumentos modificados no han pasado por estudios psicométricos, por lo cual se hace

necesario este proceso para determinar su capacidad de establecer la calidad metodológica de un estudio. Como tercer aspecto, destacan que la evaluación de la calidad de los estudios es esencial en la producción y publicación científica, y debe existir un grado adecuado de conocimiento sobre la validez y confiabilidad de las escalas utilizadas para ser cautelosos al momento de hacer interpretaciones de los resultados. Por último, afirman que los instrumentos valorados difieren en número de ítems, validez, confiabilidad y parámetros o normas de puntuación (Cascaes et al., 2013).

En el campo específico de la psicología, Jarde, Losilla, y Vives (2012) plantean que en las revisiones sistemáticas, la variable de sesgo que tiene mayor dificultad para ser controlada es la baja calidad metodológica de las investigaciones y la existencia de estos problemas en los artículos ocasiona que las conclusiones de las revisiones sistemáticas no sean confiables. Teniendo en cuenta la importancia de las revisiones sistemáticas, se hace necesario lograr un acuerdo sobre un instrumento para evaluar calidad metodológica, específicamente en las investigaciones no experimentales. Con este fin realizaron una revisión sistemática sobre artículos en el campo de las ciencias de la salud y particularmente en psicología, que incluyeran instrumentos de evaluación de calidad en estudios no experimentales de cohorte, casos y controles y transversales. La búsqueda arrojó un total de 74 instrumentos. Entre las Conclusiones más relevantes, se encuentran que la mayoría de los instrumentos analizados no han sido contruidos de una forma rigurosa; en segundo lugar, al evaluar lo referente a la construcción del instrumento y al contenido, solo uno superó los estándares básicos; sin embargo, esta herramienta presentó falencias en cuanto a la ausencia del concepto de calidad y la valoración de las fuentes de financiación y conflictos de intereses. En tercera instancia no existe acuerdo en los ítems que son utilizados para evaluar la calidad, lo cual se ve reflejado en la heterogeneidad de contenidos de los instrumentos.

El interés de la realización de este tipo de instrumentos, también se ha generado en el contexto de los estudios clínicos aleatorizados en dolor; existen dos escalas reportadas por la literatura que se han desarrollado en esta área, la primera de estas es construida por Jadad, et al. (1996), quienes aseveran que establecer la efectividad de una intervención médica es un aspecto complejo, principalmente si los resultados de los estudios sobre una misma problemática son contradictorios. La escala final quedó conformada por tres ítems que

valoran la utilización de la aleatorización, la utilización de doble enmascaramiento y la descripción de los retiros o abandonos de los participantes. El estudio concluyó que debido a que los ítems son muy generales, el instrumento puede ser utilizado en cualquier área de la medicina. Un problema que presenta esta escala es que adolece de un gran número de elementos que en la actualidad son considerados de vital importancia para valorar la calidad de una investigación, como es el caso de los análisis estadísticos y las implicaciones éticas.

El segundo instrumento que ha enfatizado en la valoración los ensayos clínicos controlados en el área de dolor es la escala de Yates, un valor agregado es que se ha especializado en las investigaciones psicológicas. Yates, Morley, Eccleston, & Williams (2005) proponen que la validez y utilidad de los resultados de un ensayo clínico controlado dependen de la calidad de este, y que la cuantificación de la calidad puede ser utilizada en las revisiones sistemáticas y en los metaanálisis para excluir estudios que no cumplen con los criterios de calidad, o para determinar la influencia de las normas de calidad en los resultados de las investigaciones. Los autores afirman que existen diferencias entre las investigaciones médicas y las psicológicas, por lo cual proponen que deben existir diferencias en los parámetros de calidad en cada uno de estos campos; para cumplir con este fin se dirigieron a construir un instrumento para evaluar la calidad metodológica de artículos con intervenciones psicológicas y desarrollaron una escala compuesta por seis ítems dirigidos a evaluar la calidad del tratamiento y una segunda sección de 20 ítems en donde se valora la calidad del método y del diseño. Las principales Conclusiones del estudio se refieren a que el instrumento presenta una buena confiabilidad, y debido a sus características puede ser utilizado en investigaciones en áreas diferentes a las del dolor; sin embargo, una dificultad que presenta es que los ítems propuestos no cubren aspectos que podrían tener impacto en la calidad de un experimento, como la adherencia del terapeuta al tratamiento, la credibilidad de la terapia y la presentación de efectos secundarios (Yates et al., 2005).

3.1.4 Valoración del cumplimiento de los estándares calidad metodológica y de publicación

Una forma de saber si los estándares de calidad se están cumpliendo, es por medio de una revisión de investigaciones, la forma más reconocida de este tipo de estudios es el metaanálisis que es definido por Guallar, Damián, y Martín (1997) como “la utilización de

técnicas óptimas para la realización de búsquedas sistemáticas de la bibliografía, la selección de los estudios originales, la extracción de las medidas de efecto de los estudios y la síntesis de resultados en estimadores combinados” (p. 345).

En los últimos años se han planteado diferentes maneras de hacer revisiones de investigaciones; la primera es la tradicional, en la cual los estudios seleccionados dependen en gran parte de la subjetividad del investigador, lo cual genera sesgos y dificulta la replicabilidad (Marín, Sánchez, & López, 2009). La segunda forma, se ha denominado revisión sistemática, la cual se caracteriza porque se hacen evidentes “los criterios de inclusión y exclusión de los estudios, las estrategias de búsqueda de la literatura, los procedimientos para la codificación sistemática de la información contenida en los estudios y el tipo de análisis aplicado sobre esta información” (Marín et al, 2009, p. 108). La tercera opción es el metaanálisis, el cual, adicionalmente a la revisión sistemática, realiza análisis estadístico de los resultados de las investigaciones seleccionadas (Marín et al., 2009).

De esta manera una tendencia de la investigación en el área de la calidad metodológica y de publicación ha sido la realización de estudios dirigidos a valorar la calidad metodológica y de publicación en diferentes áreas y temáticas de conocimiento; un ejemplo es el estudio desarrollado por Losada et al. (2009) quienes crearon un instrumento para evaluar calidad metodológica de los estudios de pronóstico, el cual fue aplicado a 121 artículos, las principales dificultades encontradas en las publicaciones, en su orden fueron que no realizan cálculo de tamaño de la muestra (97%), no mencionan el diseño de investigación (68%), no presentan criterios de exclusión (52%) y no presentan criterios de inclusión (20%).

Otro estudio generado en esta línea, fue desarrollado en el campo de la medicina por Tao et al (2011) en el cual realizaron una revisión sistemática de investigaciones, que tenían como objetivo hacer revisiones bibliográficas sobre la aplicación de anestesia, tomando como base las declaraciones QUOROM (Quality of Reporting of Meta-analyses) y PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) que son una serie de lineamientos para la publicación de revisiones sistemáticas y metaanálisis. El estudio evaluó un total de 13 revisiones sistemáticas, teniendo en cuenta una lista de chequeo en donde se registraba la presencia o ausencia de un total de 18 características de normas de publicación seleccionadas para el estudio. Los autores encontraron que un porcentaje relativamente bajo

(27%) de las revistas hacen referencia explícita a la utilización de normas de publicación como el Prisma; también se evidenció una buena adherencia a las normas del Quorum obteniéndose un promedio de 15 ítems cumplidos, sobre los 18 posibles. Un dato interesante es que el seguimiento de las normas mejoraba cuando la revista hacía explícita la necesidad de utilizarlas; por lo cual, los autores sugieren promover la utilización de listas de chequeo como el PRISMA por parte de escritores y jurados de las revisiones sistemáticas.

García, Parera, Olléc, & Bonfill (2007) desarrollaron un estudio que tenía como fin evaluar la calidad metodológica de los ensayos clínicos controlados (ECCs) publicados en la Revista Española de Anestesiología y Reanimación, la búsqueda y selección arrojó un número de 640 artículos, las principales dificultades a nivel metodológico de las investigaciones revisadas fueron la aleatorización, el ocultamiento de la asignación de los participantes, el enmascaramiento y el análisis por intención de tratamiento. Los autores sugieren la adopción de las normas CONSORT, para el mejoramiento de las publicaciones.

Otra investigación médica de este corte fue realizada por Borg et al. (2012), quiénes realizaron un metaanálisis de 23 ensayos clínicos controlados en manejo de dolor post operatorio, el objetivo era determinar la calidad de estas publicaciones, tomando como base la declaración CONSORT; para esto se seleccionaron 15 características, en las cuales se evaluaba la ausencia o presencia de estas en cada uno de los artículos, la calidad del reporte en los estudios estuvo entre baja y moderada. Para los autores había tres aspectos que eran claves a nivel metodológico como la asignación al azar, el análisis de resultados según intención de tratamiento y el enmascaramiento; sin embargo el reporte de estos aspectos en los artículos fue deficiente. Las Conclusiones de la investigación evidencian que la utilización del CONSORT mejora potencialmente la calidad de los reportes, los cuales normalmente son utilizados como fuente de información para la aplicación de estos conocimientos y para la realización de metaanálisis. Los autores motivan a las revistas a promover publicaciones completas y claras de los ensayos clínicos controlados.

Un estudio similar fue implementado por Calvo, Gómez & Sánchez (2012), los cuales pretendían evaluar la calidad metodológica de los artículos sobre tratamientos eficaces para el dolor lumbar en niños y adolescentes. Seleccionaron un total de 8 estudios y se concluyó que estos presentaban deficiencias metodológicas como la no utilización de grupos control,

la ausencia de enmascaramiento de los evaluadores, la falta de confiabilidad de los instrumentos y la ausencia de periodos de seguimiento.

Vera y Mustaca (2006) implementaron un estudio bibliométrico de publicaciones científicas en revistas de Argentina y Chile, que presentaran intervenciones en psicología clínica; en el estudio se pudieron identificar 23 publicaciones que cumplían con este requerimiento. Entre los resultados observados se encuentra que el mayor número de investigaciones empíricas se desarrollaban con modelos cognitivos conductuales; otro aspecto que fue analizado en la investigación fue si los experimentos cumplían con las condiciones propuestas por Seligman (1995) para considerarse como tratamiento eficaz, arrojando que el enfoque con mayor cumplimiento de estos criterios fue el cognitivo conductual. Con respecto a este punto, el criterio con mayor cumplimiento en las investigaciones revisadas fue el relacionado con la fijación de un número determinado de sesiones; sin embargo se presentaron de manera recurrente dificultades en lo referente al tamaño de la muestra, la asignación al azar, la ausencia de grupo control y de una fase de seguimiento.

Otro estudio en el contexto psicológico fue el desarrollado por Solar (2010), quien tenía entre sus objetivos centrales evaluar la calidad metodológica de los ensayos controlados aleatorios (ECA), en intervenciones cognitivas dirigidas a pacientes que padecen esclerosis múltiple. El número final de artículos seleccionados para el estudio fue de doce. Las principales Conclusiones encontradas fue que existían dificultades en la selección de la muestra, en la medición de los resultados, tamaño de la muestra y el enmascaramiento; por lo cual recomiendan la utilización de las normas CONSORT para mejorar la calidad de los reportes.

En el caso específico de la intervención en cefalea, también se han desarrollado investigaciones para valorar problemas metodológicos y de estándares de publicación, uno ejemplo de este tipo de estudios es el de Bronfort et al. (2008) quienes desarrollaron una revisión documental para cuantificar y comparar los efectos a corto y a largo plazo de tratamientos físicos no invasivos para cefaleas crónicas/recurrentes; con este fin realizaron búsquedas en bases de datos especializadas, en sistemas de referencia de medicina y en revistas en formato de papel. Los autores observaron que solo 22 estudios cumplían con los

criterios metodológicos requeridos, entre estos el más importante era que el diseño estuviera enmarcado dentro de ensayos controlados aleatorios y cuasi aleatorios. Al hacer un análisis de las investigaciones seleccionadas, se evidenciaron fallas metodológicas en cuanto a la comparabilidad de los grupos de estudio al inicio de la investigación, el proceso de asignación al azar, los test estadísticos y el reporte sobre criterios de inclusión y exclusión de la muestra, entre otros.

De forma similar, Nestoriuc et al. (2008) adelantaron un estudio de revisión documental sobre la eficacia de la bioretroalimentación en las cefaleas tipo tensional y migraña. La muestra final estuvo compuesta por 94 investigaciones, en las cuales se encontraron problemas metodológicos como la utilización de muestras pequeñas, sistemas de diagnóstico inadecuado; como también limitaciones en la descripción del tratamiento y/o las características de los pacientes

Un último estudio de tipo metaanalítico fue desarrollado por Verhagen (2009), el cual pretendía evaluar la eficacia de tratamientos conductuales en pacientes con cefalea tensional. El número final de artículos valorados fue de 44. El análisis de las investigaciones que conformaron la muestra final, evidenció que un gran número de los artículos presentan problemas metodológicos entre los cuales se encuentran presentación insuficiente de datos, muestras pequeñas, dificultades en la selección de los participantes problemas en el “enmascaramiento” de pacientes y terapeutas. Los autores recomiendan la utilización de criterios diagnósticos para la selección de la muestra y utilizar la declaración CONSORT.

Finalmente Acevedo, López & Cárdenas (2014) exponen que en los estudios psicológicos los tamaños de muestra son bajos con respecto investigaciones en Biomedicina o genética.

En la segunda parte del marco teórico y estado del arte, se pretende información sobre la cefalea, incluyendo la cefalea tensional y migraña desde, lo cual es de gran relevancia para este proceso investigativo debido a que todas las investigaciones abordadas están dirigidas a presentar intervenciones psicológicas en este tipo de cefaleas.

3.2. Cefalea

En este apartado se definirán los conceptos básicos relacionados con el dolor y la cefalea, su epidemiología, clasificación y la forma en que han sido evaluados e intervenidos.

3.2.1 Conceptos

El dolor es considerado por la IASP (1994) como “una experiencia emocional y sensorial displacentera asociada a un daño de tejido real o potencial, o descrito en términos de tal daño” (sección pain párr.1); por su parte el dolor crónico es definido como “cualquier etiología no relacionada directamente con neoplasia, asociada con una condición médica crónica o que se extiende en duración más allá del límite temporal esperado de herida del tejido y la curación normal, que afecta adversamente la función o bienestar del individuo” (American Society of Anesthesiologists Task Force on Chronic Pain Management [ASA] & the American Society of Regional Anesthesia and Pain Medicine [ASRA], 2010, p 810).

Por su parte la cefalea es considerada como “dolor en la cabeza, el cuero cabelludo o el cuello” (University of Iowa Hospitals and Clinics, 2011.párr 1). A nivel general las cefaleas se pueden dividir en primarias y secundarias; la diferencia se encuentra principalmente en que las cefaleas primarias no tienen enfermedades o circunstancias identificables que originen el dolor; mientras que en las secundarias sí, como sería el caso de traumatismos o infecciones (Bringas, 2010). En el presente estudio son de interés la migraña y la cefalea tensional, que están clasificadas dentro de las cefaleas primarias, las cuales según Bringas (2010) “constituyen la mayor parte de las consultas médicas ambulatorias por dolor de cabeza como síntoma principal” (p. 129).

3.2.2 Epidemiología

El dolor en general y la cefalea en particular se presentan muy comúnmente en la consulta médica; Trillos (2010) afirma que el diagnóstico e intervención de la cefalea, especialmente la migraña, es una de las tareas más frecuentes de los neurólogos. La IASP (2011a) asevera que “la cefalea es una de las condiciones más comunes, dolorosas e incapacitantes que afecta a las personas a través de sus vidas” (sección carta del presidente de la IASP párr.1). Esta misma institución agrega que la cefalea es una de las problemáticas más frecuentes a nivel neurológico y de medicina general.

Los datos epidemiológicos muestran que el 50 % de las personas presenta algún tipo de dolor de cabeza en un lapso de un año y más del 90% afirma que lo ha padecido alguna vez en su vida. En el caso de la migraña, el promedio de prevalencia en algún momento de la

vida de una persona es del 18%; y del 7,7% en niños y adolescentes (IASP, 2011b). La prevalencia estimada de la cefalea en la niñez es de 58,4% (Ferracini, 2014).

La WHO (2004) afirma que la cefalea la tensional y la migraña son unas de las más comunes, la cefalea tensional en países desarrollados es experimentada por dos tercios de los hombres y el 80% de las mujeres. En el caso de la cefalea tensional episódica, podría presentarse hasta en un 70% en algunas poblaciones. Con respecto a la migraña, los estudios realizados en Europa y Norteamérica arrojan que en el periodo de un año entre el 6 y 8% de los hombres la padecen, y entre el 15 y 18% de las mujeres; lo cual es similar a los datos encontrados en Centro y Suramérica. Datos adicionales muestran que en un día se presentan 3000 episodios de migraña por cada millón de habitantes y que la cefalea crónica diaria (CCD) es experimentada en 1 de cada 20 adultos (WHO, 2004).

En un estudio a nivel nacional, realizado por la Asociación Colombiana para el Estudio del Dolor [ACED] (2008) en personas de 18 a 65 años, encontró que el 53% de los sujetos reportaban haber padecido algún tipo de dolor en el último mes; también se evidenció que la cefalea, es uno de los dolores con mayor presencia arrojando que un 38% de los encuestados manifestaba padecer algún tipo de dolor de cabeza. Otros estudios epidemiológicos realizados en Colombia, han mostrado datos preocupantes relacionados con la cefalea como el de Rueda, Mantilla, Solano y Ortiz (2005) quienes encontraron que el 9,49% de los usuarios de un servicio de urgencias en la ciudad de Bucaramanga (Colombia) presentaban como motivo de consulta la cefalea.

De igual forma e otra investigación realizada en Bucaramanga, con 1505 participantes entre 18 y 65 años, arrojó que el 58% reportaban haber tenido al menos un episodio de cefalea durante su vida y se presentaba con mayor frecuencia en mujeres (64,2%), con respecto a los hombres (44%). El 13,7% de los participantes presentaba migraña y el 8,45 cefalea crónica diaria (CCD); con una edad promedio de inicio de 21,2 años. Las variables que se asociaron con cefalea episódica fueron “género femenino, hipertensión arterial, trauma craneal, alcoholismo, insomnio, ronquidos y depresión; por su parte la edad avanzada fue un factor protector. En el caso de la CCD se encontraron factores adicionales como el uso de hipnóticos y tener padres con CCD” (Rueda & Díaz, 2008 p. 218).

3.2.3 Clasificación

La más reciente clasificación de las cefaleas, ha sido desarrollada por la sociedad internacional de la cefalea en el año 2013. Este documento permite orientar al clínico en el proceso de diagnóstico, y establecer un lenguaje universal en este tópico, de gran utilidad en la investigación básica (Belvis, Mas & Roig, 2015). Para la presente investigación es de interés la clasificación de las cefaleas primarias, específicamente la migraña y la cefalea tensional.

Con respecto a la migraña, se ha firmado que es un desorden primario del cerebro, originado por un problema en los canales iónicos aminérgicos, los cuales están relacionados con las entradas sensoriales e influyen en las venas craneales, generando vasodilatación e inflamación neurogénica, por lo cual se puede afirmar que la migraña tiene un origen neurovascular (Roceanu & Bajenaru, 2014). En coherencia con lo anterior, se ha afirmado que los sujetos con migraña crónica, evidencian cambios morfológicos y funcionales dentro del sistema nervioso central, que evidenciarían alteraciones en la percepción del dolor (Visens, 2014). Otros autores que han respaldado la posición de que la migraña se origina por alteraciones del sistema nervioso son (Panerai, 2013; Rathier & Roth, 2015).

También se ha propuesto que el origen de la migraña está relacionado con un componente genético y con factores precipitantes como el estrés, la alimentación, estímulos sensoriales, cambios de temperatura, cambios hormonales en mujeres, dificultades en el ciclo de sueño y el ejercicio físico intenso (Bringas, 2010).

La migraña se ha clasificado en seis subtipos: migraña sin aura, migraña con aura, migraña crónica, complicaciones de la migraña, Migraña probable y Síndromes episódicos que pueden asociarse a la migraña. En el caso de la migraña sin aura se caracteriza por ser recurrente con episodios de 4 a 72 horas. El dolor es unilateral, pulsátil, con intensidad de moderada a severa y podría estar relacionada con náuseas, fotofobia y empeorarse con actividad física (Headache Classification Committee of the International Headache Society [IHS], 2013).

La migraña con aura, se presenta con episodios frecuentes de varios minutos de duración, acompañada de manifestaciones sensitivas o relacionadas con el sistema nervioso central. La sintomatología es progresiva y se presentan antes de una cefalea tipo migraña (IHS, 2013).

El siguiente tipo de migraña es la crónica, que se caracteriza por tener una duración igual o superior a 15 días en un mes, durante más de tres meses, y que durante 8 días al mes tenga características de migraña (IHS, 2013).

En el caso de complicaciones de la migraña se pueden identificar dependiendo de cuatro subtipos: Estado migrañoso (presenta una duración igual o superior a 72 horas), aura persistente sin infarto (Aura que se prolonga por una semana o mas sin evidencia de infarto), infarto migrañoso (síntomas de aura migrañosa con lesión isquémica cerebral) y crisis epiléptica desencadenada por aura migrañosa (epilepsia desencadenada por migraña con aura) (IHS, 2013).

El siguiente tipo de migraña es la probable, su denominación parte de que no cumple con los criterios de otra migrañas; ni con los de otros tipos de cefalea (IHS, 2013). La última categoría de la migraña son los síndromes episódicos que pueden asociarse a la migraña como trastorno gastrointestinal recurrente, vértigo paroxístico benigno, tortícolis paroxística benigna (IHS, 2013)

Por su parte la cefalea tensional se ha clasificado en cuatro subtipos: cefalea tensional episódica infrecuente, cefalea tensional episódica frecuente, cefalea tensional crónica y cefalea tensional probable. La cefalea tensional episódica infrecuente, se caracteriza por baja frecuencia de aparición, dolor opresivo bilateral, intensidad leve o moderada, duración entre minutos y días. El dolor no empeora con la actividad física y se puede presentar fonofobia o fotofobia. En el caso de la cefalea tensional episódica frecuente, se diferencia de la anterior en su frecuencia mayor de aparición (IHS, 2013).

La cefalea tensional crónica se caracteriza por tener una frecuencia diaria, y la cefalea tensional probable se clasifica de esta manera, debido a que no cumple los criterios de otras cefaleas tensionales, ni de ninguna otra cefalea (IHS, 2013).

Con respecto al El origen de la cefalea tensional, esta se ha relacionado con tensión muscular y con estrés; pero todavía los datos no son concluyentes (Madrid & Vidal, 2008). Por su parte Velázquez, Jurado, Zermeno, & Sánchez (2013) en un estudio que tenía como objetivo diseñar un instrumento para identificar pacientes que padecieran cefalea tensional, arrojó dos hallazgos interesantes, en primera instancia algunos sujetos que padecían este tipo de cefalea empeoraban con el ejercicio físico y por otra parte un gran porcentaje de las

personas diagnosticadas con este problema carecían de estrategias de afrontamiento del estrés.

3.2.4 Evaluación

La evaluación del dolor desde la perspectiva cognitivo conductual, se puede abordar de diversas maneras; Rodríguez y Blanco (2005) proponen que se deben tener en cuenta aspectos como: características del dolor, funcionalidad física, estado emocional, percepción sobre el tratamiento recibido, síntomas acompañantes del dolor, motivación hacia la intervención y el nivel de funcionamiento familiar, social, académico y laboral.

Una forma de obtener esta información ha sido la entrevista clínica (Olivares & Cruzado, 2008; Serrano et al., 2002) en la cual se obtiene información sobre el nivel funcional premórbido, historia educativa laboral, actividad social y aspectos generadores de estrés (Serrano et al., 2002). La segunda forma de evaluación son los cuestionarios de experiencia de dolor (Olivares & Cruzado, 2008; Rodríguez & Blanco, 2005; Serrano et al., 2002), los cuales se dirigen a evaluar principalmente los aspectos cualitativos y emocionales del dolor, un cuestionario de frecuente utilización es el de Dolor de McGill, que obtiene información sobre la localización del dolor, descripción del dolor (ej. palpita, marea) y factores que intensifican o alivian el dolor (Serrano et al., 2002).

También se han utilizado las escalas de intensidad del dolor, como las numéricas, descriptivas simples, verbales y las visuales análogas (Serrano et al., 2002) y de forma complementaria los inventarios estructurados, dirigidos a indagar síntomas, antecedentes médicos, factores que influyen en la conducta de dolor, entre otros (Serrano et al., 2002).

Por último, se han empleado pruebas psicométricas, que evalúan aspectos de la persona que pueden estar relacionados con la conducta de dolor, como la depresión y la ansiedad (Rodríguez & Blanco, 2005; Serrano et al., 2002) como también la observación directa y la autoobservación (Olivares & Cruzado, 2008; Navarro, 2006)

3.2.5 Intervención

Navarro (2006) plantea que la terapia cognitiva-conductual en el campo del dolor crónico, se sustenta en cinco supuestos a) las personas son procesadores activos de información, b) existe una permanente interacción entre los aspectos fisiológicos, comportamentales y cognitivos, c) El ambiente y el individuo actúan de una manera recíproca, d) la intervención debe enfocarse a modificar los aspectos problemáticos fisiológicos, comportamentales o cognitivos, e) Los pacientes o consultantes son activos en el proceso de su tratamiento.

Por otra parte se plantea que un abordaje bioconductual de la cefalea incluye terapia cognitiva conductual, relajación bioalimentación y manejo del estrés. Estas técnicas se dirigen a controlar los componentes fisiológicos, cognitivos y afectivos del estrés. (Rathier & Roth, 2015).

La intervención cognitivo conductual de la migraña incrementa las respuestas adaptativas, modifica el afrontamiento conductual y los pensamientos acerca de la migraña. Las estrategias de afrontamiento conductuales para la migraña pueden ser preventivas como el manejo de la ansiedad o sintomáticas con el fin de disminuir la intensidad, como sería permanecer lejos de fuentes de luz (Seng, 2014).

Existen diversas fuentes que han respaldado la efectividad de las terapias cognitivo conductuales por ejemplo Rathier & Roth, (2015) proponen que las intervenciones como la bioalimentación, la relajación y la terapia cognitivo conductual tienen efectos positivos en el control de la migraña. Una meta de las terapias cognitivo conductuales para la migraña es reducir el pensamiento catastrófico (Seng, 2014). En una intervención cognitivo conductual realizada a pacientes con migraña, se encontró que se dio un decremento en las ideas catastrofizantes, y esto generó una disminución significativa en la incapacidad ocasionada por la migraña. (Seng, 2014).

Existen diferentes formas para el abordaje del dolor desde la perspectiva cognitivo conductual, la primera es la respiración; esta técnica parte del supuesto de que una respiración inadecuada puede generar tensión muscular, y finalmente propiciar la respuesta de dolor (Moix & Casado, 2011). La Relajación, también ha sido mencionada para el manejo de migraña (Navarro, 2006; ASA & ASRA, 2010; Vallejo, 2005), y cefalea tensional (Bendtsen, Evers, Linde, Mitsikostas, Sandrini, & Schoenen, 2010). La relajación está basada en técnicas

para el control fisiológico de la ansiedad, que consisten en la realización de ejercicios físicos y cognitivos (Benito et al., 2006).

Una tercera forma es la modificación de conductas relacionadas con el dolor (Benito et al., 2006; Vallejo, 2005). El objetivo es la reducción o eliminación de las conductas de dolor como quejas, postura, expresiones faciales, conductas de evitación; y a la restauración de las actividades diarias y ejercicio físico (Moix & Casado, 2011). Otra forma de abordaje del dolor ha sido la intervención cognitiva (Bendtsen et al., 2010; Navarro, 2006; Vallejo, 2005). La técnica se basa en el supuesto de que las personas hacen interpretaciones distorsionadas de la realidad, y esto conlleva a la generación de emociones negativas, las cuales están relacionadas con el inicio y o mantenimiento del dolor; por lo tanto la intervención se dirige a modificar los pensamientos distorsionados y de esta manera disminuir o eliminar las emociones negativas que propician el dolor (Moix & Casado 2011).

Un ejemplo de una intervención cognitivo conductual, que agrupa varias de las técnicas mencionadas, es un protocolo de intervención en dolor propuesto por Moix & Casado (2011), los autores proponen 10 sesiones grupales, en las que deben participar entre 8 y 10 personas, y dos sesiones individuales. Cada sesión tiene una duración aproximada de 2 horas y está conformada por tres partes; discusión de las de las tareas realizadas en casa, implementación teórico-práctica del contenido, y asignación de tareas a ser desarrolladas en el espacio entre las sesiones. Las técnicas utilizadas son respiración y relajación, manejo de la atención, reestructuración cognitiva, solución de problemas, manejo de emociones y asertividad, valores y establecimiento de objetivos; organización del tiempo y actividades reforzantes, ejercicio físico, higiene postural y del sueño; y prevención de recaídas. De forma complementaria, se realizan sesiones individuales, para tratar aspectos particulares del paciente.

Otra forma de intervención citada en la literatura es la bioetroalimentación (Bendtsen et al., 2010; Navarro, 2006; Vallejo, 2005). Según Navarro (2006) consiste en brindarle información al consultante sobre sus respuestas fisiológicas; con el fin de que aprenda a modificarlas. Usualmente en este tipo de intervención se utilizan aparatos electrónicos que captan la señal fisiológica, la amplifican y presentan de forma visual o auditiva al paciente. De forma similar Rathier & Roth (2015) exponen que la bioetroalimentación monitorea los

procesos fisiológicos de los cuales los pacientes no tiene conciencia. El paciente recibe información visual o auditiva de esos procesos. Los pacientes deben practicar para disminuir el estrés y la tensión.

Navarro agrega que existen diferentes respuestas que han sido utilizadas para el manejo de la cefalea, como la electromiografía y la electroencefalografía, que miden tensión muscular y actividad eléctrica cerebral, respectivamente.

Vallejo (2008) afirma que la biofeedback genera información sensorial, la cual facilita la acomodación a un control natural del dolor. Los efectos positivos de la biofeedback en la cefalea tensional y la migraña han sido documentados en diferentes respuestas fisiológicas, con varios grupos poblacionales entre estas se encuentran; temperatura de la piel en mujeres adultas con migraña (Kang, Park, Chung, & Yu, 2009); respiración diafragmática en adultos con migraña (Kaushik, Mahajan, & Rajesh, 2005); pulso en la zona temporal en adultos con migraña y cefalea tensional (Martin, Forsyth, & Reece, 2007); temperatura de la piel en niños con migraña (Scharff, Marcus, & Masek, 2002); potenciales corticales en niños con migraña (Siniatchkin, 2000).

Un componente esencial de las terapias cognitivo conductuales es la psicoeducación. Las intervenciones conductuales incluyen dentro de sus metas la modificación de comportamientos que mantienen o intensifican las cefaleas, para reemplazarlos por estrategias de bienestar como sería el caso de abuso de medicación o de cafeína. De esta manera se educa a los pacientes en actividades de bienestar, con el fin de mejorar el manejo de la cefalea. Los pacientes pueden modificar los hábitos alimenticios, realizar actividad física y reducir el uso de analgésicos entre otros. La terapia cognitivo conductual provee al paciente con estrategias para evitar los estímulos provocadores de la cefalea, mejora las estrategias de afrontamiento y ayuda al manejo de síntomas comorbidos como la depresión y ansiedad. Rathier & Roth, 2015).

En los últimos años se han desarrollado nuevas formas de intervención del dolor como la terapia de aceptación, la expresión emocional escrita y terapias mediadas por la WEB. En el caso de la terapia de aceptación, se asume que el paciente debe asumir el dolor, el malestar y las limitaciones que genera, para luego empezar a retomar las actividades cotidianas, a pesar de sentir el malestar o no estar motivado para realizarlas; este cambio de actitud potencia las estrategias de afrontamiento y la autoestima (Vallejo, 2008). Un ejemplo de este tipo de

intervención fue desarrollado por Mo'tamedi, Rezaemaram y Tavallaie (2012) en un estudio, que pretendía reducir la experiencia sensorial de dolor, la angustia y el nivel de discapacidad en mujeres que padecían cefalea tensional y migraña crónica, los resultados evidenciaron reducciones significativas en los niveles de angustia y discapacidad pero no en la experiencia sensorial del dolor.

Otra terapia de utilización relativamente reciente para el manejo del dolor y la cefalea, ha sido denominada expresión emocional escrita, al respecto Correché y Labiano (2003) afirman que existe evidencia de que la escritura de experiencias traumáticas en un contexto clínico, mejora la salud física. Un estudio en donde se aplicó este tipo de intervención, fue desarrollado por D'Souza, Lumley, Kraft y Dooley (2008), en la investigación se le solicitaba a los participantes escribir sobre experiencias que les hubieran generado un alto grado de estrés, enfatizando en los sentimientos experimentados y en la forma en que estos eventos afectaron las relaciones sociales, la salud y/o la cefalea experimentada. En el caso específico de esta investigación, no se presentaron diferencias significativas con respecto al grupo control en las variables de frecuencia e intensidad del dolor; nivel de discapacidad y sentimientos negativos; los investigadores proponen el desarrollo de nuevos estudios que permitan determinar las variables que potencian los efectos de expresión emocional escrita en la intervención del dolor, como podría ser el caso de las características de la muestra utilizada.

Otra tendencia que se ha instaurado en los procesos de terapia de dolor, incluyendo la cefalea y la migraña, es la utilización de tecnologías tradicionales y de nuevas tecnologías, como mediadoras del proceso de intervención; entre estos casos se puede mencionar la investigación de Andersson, Lundström y Ström (2003) en donde implementaron un proceso de intervención, en modalidad autoaplicada a un grupo de pacientes que padecían migraña, cefalea tensional y cefalea en racimos; los pacientes debían utilizar información ubicada en la Web, la cual era complementada por asesorías telefónicas por parte de los terapeutas. Los resultados del estudio mostraron reducciones significativas en el nivel la discapacidad ocasionada por la cefalea, los índices de depresión, las estrategias de afrontamiento desadaptativas y se presentó un leve mejoramiento en el índice de cefalea. Otro resultado arrojado por el estudio, es que la intervención telefónica no generó diferencias significativas con respecto a la intervención mediada por la WEB únicamente. Otras intervenciones

exitosas mediadas por WEB han sido reportadas por (Bromberg et al., 2012; Devineni & Blanchard, 2005; Trautmann & Kröner-Herwig, 2008).

4. Método

4.1 Diseño

El diseño del presente proyecto se enmarca dentro de una revisión de investigaciones. Existen tres formas de esta metodología la tradicional, la revisión sistemática y el metaanálisis (Marín et al., 2009).

La revisión sistemática, se caracteriza porque se hacen evidentes “los criterios de inclusión y exclusión de los estudios, las estrategias de búsqueda de la literatura, los procedimientos para la codificación sistemática de la información contenida en los estudios y el tipo de análisis aplicado sobre esta información” (Marín et al, 2009, p. 108).

4.2 Universo de estudio

Debido a que el presente estudio, se enmarca dentro de una revisión de investigaciones, el universo y la muestra de estudio fueron artículos que cumplieran con los siguientes criterios de inclusión, 1) artículos en texto completo que incluyeran intervenciones psicológicas en cefalea tensional y migraña; 2) Estar publicados entre el 1 de enero de 2000 y el 31 de diciembre de 2012 en las bases de datos Excerpta Medica dataBASE (EMBASE) y Elton B. Stephens Co (EBSCO); en esta última, los artículos deberían encontrarse en las

bases de datos de ciencias de la salud MedicLatina, Health Source: Nursing /Academic Edition y fuente académica premier; o en las de Psicología /sociología: Psychology & Behavioral Sciences Collection, Academic Search Complete y Fuente Académica Premier. Las bases de datos fueron seleccionadas teniendo en cuenta que incluyen revistas de alto impacto en el área de la medicina y de la psicología y de gran pertinencia para el tema específico de investigación del presente proyecto como Headache, Journal of Headache & Pain, Pain Medicine, Behavior Research and Therapy Annals of Behavioral Medicine, Pain, Journal of Medical Sciences, Behavior Therapy Cephalalgia, Applied Psychophysiology & Biofeedback, Journal of Pain, Behavioural and Cognitive Psychotherapy, Behaviour Research & Therapy.

Para la búsqueda, se utilizó la combinación de los términos clave cefalea, migraña, cefalea tensional, tratamiento, intervención, terapia, cognitivo, conductual, psicológico y psicología. La búsqueda de los estudios, la selección, la extracción de datos y la valoración del cumplimiento de estándares en las investigaciones seleccionadas fueron desarrolladas por uno de los autores con formación de pregrado en Psicología, se encuentra desarrollando estudios de maestría en psicología y con experiencia en investigación a nivel aplicado y en el área de docencia. El proceso fue asesorado por un coinvestigador con formación de pregrado en Psicología, posee título de doctorado y con amplia experiencia en el campo de la investigación aplicada y la docencia.

La Figura 1 ilustra el proceso llevado a cabo para la selección de los artículos; como se puede ver en primera instancia se consultaron las bases de datos teniendo en cuenta los términos clave, lo cual arrojó un total de 9827 artículos; posteriormente se hizo el primer proceso de filtraje teniendo en cuenta la información del título y resumen, dejando un saldo de 53 estudios, los artículos fueron excluidos por no cumplir con alguno de los criterios de inclusión expresados anteriormente como es el caso de no estar disponible en texto completo o no incluir intervenciones de tipo psicológico. Estos estudios, fueron revisados en su totalidad para hacer el último proceso de filtraje, en donde se presentaron varias razones para la exclusión de artículos, como tener un diseño correlacional, que la intervención fuera en una cefalea diferente a migraña o tensional entre otros. Finalmente la muestra quedó compuesta por 45 artículos (ver Apéndice A).

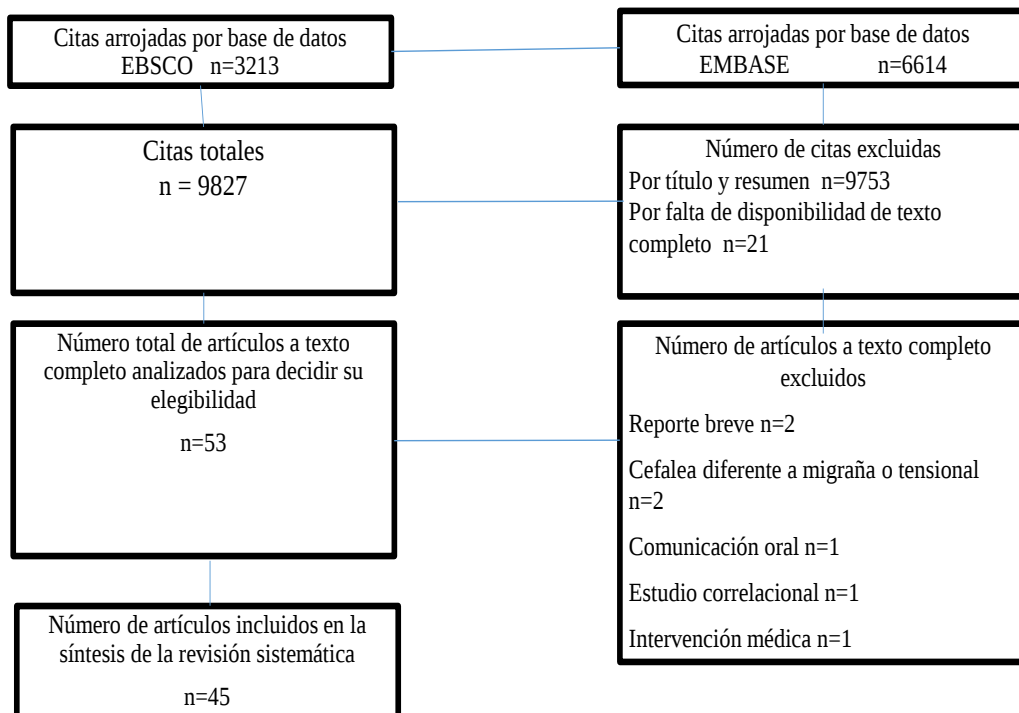


Figura 1. Diagrama de flujo de los artículos a través del proceso investigativo.

4.3 Instrumentos

Para la investigación se utilizó el **Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01)**. El inventario es un instrumento en formato digital, el cual se encuentra en Excel (ver Apéndice B). El instrumento está diseñado de tal forma, que el evaluador al revisar el estudio o artículo, utiliza una calificación dicotómica para cada uno de los indicadores (Cumple con el indicador: 1, no cumple con el indicador: 0). El instrumento fue diseñado, construido y validado por los investigadores, basado en información de artículos que presentaban estándares metodológicos y de publicación, principalmente para investigaciones de corte experimental (ver referencias apéndice C). Posterior a su construcción, se realizó un proceso de validación por once jueces con experiencia en publicación de estudios de corte experimental en psicología o con conocimientos en diseño y construcción de instrumentos o pruebas psicológicas; también se realizó una prueba piloto con dos artículos; en esta última se encontró que los indicadores del instrumento funcionaban adecuadamente al realizar la evaluación en los artículos.

Una vez desarrollado este proceso se eliminaron nueve ítems que no cumplían con los criterios de validez de contenido propuestos por Lawshe y Tristán (ver Tabla 1). En el caso de Lawshe (1975), se plantea que para 11 jueces (número de jueces utilizados en el presente proyecto) la validez mínima de contenido requerida es de 0.59; por su parte para Tristán (2008) la validez mínima requerida de un ítem, independiente del número de jueces es de 0.5823; al ser eliminados los ítems que no cumplían con los criterios mínimos de validez de contenido, la prueba final quedó conformada por siete categorías, 18 subcategorías y 157 indicadores (ver Tabla 2). La validez de contenido obtenida fue de 0.89 tanto por el método de Polit y Hungler (2000), como por el Lawshe (1975). Por su parte el nivel de concordancia de Kendall tanto para pertinencia como para redacción fue significativo con un $p > 0.01$. Es importante aclarar que en el presente estudio, no se presentaron las figuras que muestran los resultados de porcentaje de cumplimiento de los indicadores incluidos en la subcategoría gráficas, correspondiente a la categoría resultados; debido a que esta es una forma de presentación opcional de información; por lo cual se podría generar un sesgo en los datos; lo cual no ocurre en el caso de las tablas, en donde uno de los estándares es la presentación de una Tabla con la línea de base de las variables demográficas (R3.1), este estándar ha sido

propuesto por diversos autores como (Borg et al., 2012; Boutron et al., 2008; Schulz et al., 2010). Por lo anterior los resultados de gráficas se presentarán en formato en texto, y la recomendación es utilizar los estándares de gráficas, solo en artículos que contengan esta forma de presentación de datos.

Tabla 1

Validez de contenido de indicadores eliminados por categoría y subcategoría

Categoría	Subcategoría	Indicador	Valor validez de contenido
Portada	Título	1. El título está compuesto por 15 palabras o menos.	0.54
	Resumen	2. En el resumen se presentan los criterios de inclusión y exclusión de los participantes.	0.54
		3. En el resumen se presentan las características centrales de los terapeutas de investigación, según lo expuesto en el apartado de método.	0.54
		4. en el resumen se especifica la forma en la cual se hizo cegamiento (enmascaramiento) a los participantes de la investigación	0.45
		5. En el resumen se presentan los efectos secundarios en los sujetos, como consecuencia de la intervención implementada.	0.54

		6. En el resumen se presenta el nombre y número de registro del experimento.	0.54
		7. En el resumen se nombran las instituciones que financiaron la investigación.	0.54
		8. En el resumen se especifica la información que ha sido presentada en forma oral y/o escrita, en artículos o eventos anteriores.	0.36
Método	Participantes	9. En el apartado de los sujetos, se hace explícito que la muestra es representativa de la población	0.54

Tabla 2

Estructura del instrumento Inventario de Valoración de Estándares Metodológicos y de Publicación (INVEMCION-01).

Categorías	Subcategorías	Indicadores	Total indicadores por categoría
Portada	Título	6	26
	Resumen	20	
Introducción	Contexto investigativo	5	11
	Propósitos de la investigación	6	
Método	Diseño	2	75
	Participantes	27	
	Instrumentos	2	
	Definición de variables	3	
	Procedimiento	24	

	Procesamiento estadístico	5	
	Consideraciones éticas	12	
Resultados	General	14	21
	Texto	3	
	Tablas	2	
	Gráficas	2	
Discusión	Análisis	7	16
	Generalización	8	
	Implicaciones	1	
Conclusiones		2	2
Referencias		6	6
Total categorías: 7	18		157

4.4 Definición de variables

Las variables centrales del presente estudio serán definidas tanto en el campo conceptual como operacional.

Definiciones conceptuales

Estándares de publicación: Normas para reportar una investigación de una manera exacta, clara y sin sesgo. (ICMJ, 2013).

Estándares metodológicos: Normas que regulan el diseño de un experimento, con el fin de que generen resultados no sesgados, que sean suficientemente precisos y permitan la aplicación en la práctica clínica (Verhagen, de Vet, de Bie, Boers, Van den Brandt, 2001).

Definiciones Operacionales:

Estándares de publicación: Número de indicadores rotulados o definidos dentro de esta categoría en cada uno de los apartados del Inventario de Valoración de Estándares Metodológicos y de Publicación (ver Apéndice B).

Estándares metodológicos: Número de indicadores rotulados o definidos dentro de esta categoría en cada uno de los apartados del Inventario de Valoración de Estándares Metodológicos y de Publicación (ver Apéndice B).

4.5 Procedimiento

El desarrollo de la investigación estuvo conformado por dos fases, la primera consistió en el Diseño, construcción y validación del instrumento; y la segunda, aplicación del instrumento y análisis de resultados.

Fase I. Diseño, construcción y validación del instrumento

Esta fase de la investigación estuvo compuesta por cinco momentos que estaban comprendidos entre la revisión de la literatura, hasta el análisis de los resultados de la validación del instrumento.

Revisión de literatura: En esta fase se realizó la búsqueda de literatura actualizada, en lo referente a estándares metodológicos y de publicación en investigaciones de tipo experimental, principalmente en el contexto de la psicología y de la medicina.

Categorización de estándares: Teniendo en cuenta lo propuesto por la literatura, se organizaron los estándares en categorías y subcategorías; por ejemplo, las categorías obtenidas fueron portada, introducción, método, resultados, discusión, Conclusiones y referencias bibliográficas.

Construcción de ficha técnica y estructura de la prueba: Se procedió a hacer la ficha técnica y la estructura de la prueba.

Redacción de indicadores y construcción del formato del instrumento: Se redactaron los indicadores que pertenecían a cada una de las categorías; y se diseñó el formato de la prueba que incluía los datos generales, las instrucciones, la forma de diligenciamiento y el cuerpo de la prueba.

Validación por jueces y prueba piloto: En este punto se realizó la validación por jueces. Para este fin, se generaron tres documentos, el primero fue la carta para evaluación por los jueces (ver Apéndice D), en donde se solicitaba la evaluación y se presentaban las instrucciones para la realización de la validación. El segundo documento contenía la información sobre la prueba en donde se incluía el resumen de la investigación, la Ficha Técnica y la estructura de la prueba (ver Apéndice E). Finalmente, el tercer documento fue el Inventario de Valoración

de Estándares Metodológicos y de Publicación (INVEMCION-01) para validación por jueces (ver Apéndice F) que estaba conformado por la prueba y un espacio adicional para que los evaluadores registraran las valoraciones sobre pertinencia y redacción de cada uno de los indicadores o ítems, sus observaciones y su percepción sobre clasificar el indicador como estándar Metodológico (EM) o estándar de Publicación (}EM) o ambos, lo cual será tema de una investigación posterior. Estos documentos fueron enviados a 40 profesionales en el ámbito nacional e internacional, expertos en psicometría y/o investigación de los cuales 11 accedieron voluntariamente a su evaluación. y la prueba piloto con dos artículos

Análisis de resultados de validación: Los análisis realizados a la prueba fueron validez de contenido por el método de Polit y Hungler (2000), como por Lawshe (1975). Los resultados de la validez de contenido, permitieron eliminar nueve ítems que no cumplían con los criterios mínimos de validez de contenido propuestos por Lawshe (1975) y Tristán (2008) (Ver apartado 4.3 Instrumentos). También se utilizó el coeficiente de concordancia de Kendall para las valoraciones de los jueces tanto en la pertinencia, como en la redacción del instrumento para lo cual se utilizó el Paquete Estadístico para las Ciencias Sociales (SPSS) 22,

Fase II Aplicación del instrumento y análisis de resultados

En esta segunda fase se desarrolló el proceso de revisión sistemática propiamente dicho, el cual estaba compuesto por tres momentos que incluían la búsqueda de artículos, la valoración de estos y el análisis de resultados.

Búsqueda en bases de datos y selección de artículos: Teniendo en cuenta las palabras clave y los criterios de inclusión de los artículos mencionados en el apartado 4.2, se realizó la búsqueda en las Bases de Datos EBSCO y EMBASE y se eligieron los estudios que finalmente cumplían con los criterios propuestos.

Valoración de artículos: Los estudios que cumplieron con los criterios establecidos fueron 45, los cuales fueron valorados de manera individual, registrando el cumplimiento o no en cada uno de los indicadores propuestos en las diferentes categorías y subcategorías.

Análisis de resultados de los indicadores: Los datos arrojados fueron básicamente dos, el porcentaje de cumplimiento del indicador en cada una de las categorías y subcategorías, tanto a nivel general (Periodo 2000-2012), como en cada uno de los subperiodos evaluados (2000-2004, 2005-2008, 2009-2012). También se utilizó un análisis inferencial no paramétrico con

el Test Exacto de Fisher, el cual permite determinar asociaciones entre dos variables dicotómicas, cuando la muestra a estudiar es pequeña y por lo tanto no se cumplen las condiciones para la aplicación de la prueba Chi-Cuadrado (Peralta, 2005). El nivel mínimo de significación aceptado es $p < 0.05$. En el caso particular de este estudio se intenta determinar si existe asociación entre el porcentaje de cumplimiento de cada indicador y los periodos evaluados, a saber 2000-2004, 2005-2008 y 2009-2012. Los datos fueron procesados con el Paquete Estadístico para las Ciencias Sociales (SPSS) 22

4.6 Consideraciones éticas

El presente estudio se desarrolló en el marco de la Ley 1090 de 2006, por la cual se reglamenta el ejercicio de la profesión de Psicología, se dicta el Código Deontológico y Bioético y otras disposiciones (Congreso de la República de Colombia, 2006). Es importante aclarar que la investigación no requiere contacto y/o manipulación de seres humanos y animales.

El anteproyecto fue valorado por dos Psicólogos con experiencia en investigación desde las perspectivas conceptual, metodológica y ética; esta evaluación fue avalada por el comité de la Maestría de Psicología de la Universidad Pontificia Bolivariana de Bucaramanga (Colombia). También se declara que el presente estudio fue patrocinado financieramente por la Dirección de investigaciones de esta Universidad, como apoyo al desarrollo de un trabajo de grado de la maestría en Psicología con Código 086-0413-3100. Información adicional sobre la investigación puede ser solicitada a Luis Argenis Osorio Ferrer (luis.osorio@upb.edu.co) o Ana Fernanda Uribe Rodríguez (anafernanda.uribe@upb.edu.co).

5. Resultados

En esta sección se presentarán los resultados en tres apartados, en el primero estarán los resultados obtenidos de la validación de jueces, por contenido y coeficiente de concordancia del instrumento **INVEMCION-01**. En el segundo, se presentarán los análisis de calidad metodológica y de publicación de los artículos evaluados en cada una de las categorías y subcategorías. Por último, la evolución de la calidad metodológica y de publicación, de los artículos evaluados en el periodo 2000-2012, teniendo en cuenta el porcentaje de cumplimiento de cada uno de los indicadores en las diferentes categorías y subcategorías.

5.1 Validación por jueces del Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01).

En este apartado los resultados se presentarán, teniendo en cuenta dos tipos de análisis, la validez de contenido del inventario, como lo propuesto por (Lawshe, 1975; Tristán, 2008) y el coeficiente de concordancia de Kendall para las valoraciones realizadas por los jueces como lo propuesto por (Escobar & Cuervo ,2008).

El primer tipo de análisis es la validez de contenido, la cual fue valorada por dos métodos diferentes, el primero fue por medio de la prueba de expertos descrita por Polit y Hungler (2000), cuyos resultados se presentan en la tabla 3. Como se puede observar la validez total

del instrumento fue de 0.89, al respecto Lawshe (1975) plantea que para 11 jueces (número de jueces utilizados en el presente proyecto) la validez mínima de contenido requerida es de 0.59; por su parte Tristán (2008), propone que la validez mínima requerida para un instrumento o un ítem, independiente del número de jueces es de 0.5823, ante lo cual se puede afirmar que el instrumento supera los dos criterios propuestos. En cuanto a las categorías del instrumento, la gran mayoría presentó valores superiores a 0.9. Los valores menores se observaron en portada (0.82) y la subcategoría título (0.77).

Tabla 3

Índice validez de contenido por jueces Polit y Hungler (2000) en función de categorías y subcategorías.

Categoría	Subcategoría	Índice de validez
Portada	Título	0.77
	Resumen	0.84
	Total categoría	0.82
Introducción	Contexto investigativo	0.95
	Propósitos de la investigación	0.94
	Total categoría	0.94
Método	Diseño	0.95
	Participantes	0.86
	Instrumentos	1
	Definición de variables	0.85
	Procedimiento	0.92
	Procesamiento estadístico	0.96
	Consideraciones éticas	0.9
	Total categoría	0.9
	General	0.9
	Texto	0.91
Resultados	Tablas	0.91
	Gráficas	0.91
	Total categoría	0.9
	Análisis	0.97

	Generalización	0.89
	Implicaciones	0.91
	Total categoría	0.92
Conclusiones		1
Referencias		0.91
Total Prueba		0.89

El segundo método utilizado para calcular la validez de contenido fue el propuesto por Lawshe (1975). En la Tabla 4, se puede observar que la validez de contenido obtenida por este método (0,89) coincide exactamente con el anterior, lo cual corrobora los resultados obtenidos. Las categorías oscilaron 0.82 en portada y 1 en conclusiones. En el caso de las subcategorías, el valor mínimo se presentó en título (0,79) y las restantes presentaron valores superiores a 0.83. De lo anterior se puede afirmar que tanto el instrumento como las categorías y subcategorías superaron los valores mínimos de validez de contenido propuestos por Lawshe (1975) y Tristán (2008).

Tabla 4

Índice de validez de contenido por jueces Lawshe (1975) en función de categorías y subcategorías.

Categoría	Subcategoría	Índice de validez
Portada	Título	0.79
	Resumen	0.83
	Total categoría	0.82
Introducción	Contexto investigativo	0.94
	Propósitos de la investigación	0.96
	Total categoría	0.95
Método	Diseño	0.91
	Participantes	0.84
	Instrumentos	1
	Definición de variables	0.84
	Procedimiento	0.89
	Procesamiento estadístico	0.96
	Consideraciones éticas	0.91
	Total categoría	0.88
	General	0.90
	Texto	0.91
Resultados	Tablas	0.91
	Gráficas	0.91
	Total categoría	0.91

Discusión	Análisis	0.97
	Generalización	0.89
	Implicaciones	1
	Total categoría	0.93
Conclusiones		1
Referencias		0.91
Total prueba		0.89

El segundo tipo de análisis realizado al inventario, consistió en la utilización del coeficiente de concordancia de Kendall, el cual básicamente mide la relación entre las valoraciones de diferentes jueces sobre un aspecto particular. En la Tabla 5, se puede apreciar que el inventario en su totalidad presentó una concordancia significativa entre los jueces con un W de Kendall= 0,145 ($p<0.01$), con respecto a la pertinencia de los ítems. Al hacer el análisis por categorías, tres de las siete presentaron valores significativos con un $p<0.01$ que fueron portada, método y conclusiones. En este punto es importante anotar que el Paquete Estadístico para las Ciencias Sociales (SPSS) 22 no arrojó valores para la categoría conclusiones, porque todos los jueces presentaron la misma calificación “5” en los indicadores existentes. En cuanto a las subcategorías, el título presentó un nivel de significancia de $p<0.05$ y en el caso de instrumentos el Paquete Estadístico para las Ciencias Sociales (SPSS) 22 no arrojó valores, porque el nivel de concordancia entre los jueces fue demasiado alto, debido a que 10 de los 11 jueces acordaron en un valor de “5”. Por su parte las subcategorías resumen, procedimiento y consideraciones éticas presentaron niveles de significancia importantes al nivel de $p>0.1$

Tabla 5

Valores y nivel de significación del coeficiente de concordancia de Kendall, para evaluación de pertinencia de los ítems por parte de los jueces, en función de categorías y subcategorías.

Categoría	Subcategoría	Kendall P	Chi	Gl	Sig. Asintótica
Portada	Título	0,313	17,189	5	0,004***
	Resumen	0,144	30,077	19	0,051*
	Total categoría	0,176	48,467	25	0,003***
Introducción	Contexto investigativo	0,14	6,182	4	0,186
	Propósitos de la investigación	0,034	1,857	5	0,869

Método	Total categoría	0,082	8,98	10	0,534
	Diseño	0,273	3	1	0,083*
	Participantes	0,118	33,66	26	0,144
	Instrumentos			1	
	Definición de variables	0,182	4	2	0,135
	Procedimiento	0,138	34,876	23	0,054*
	Procesamiento estadístico	0,109	4,8	4	0,308
	Consideraciones éticas	0,144	17,439	11	0,096*
Resultados	Total categoría	0,147	119,637	74	0,001***
	General	0,094	13,489	13	0,411
	Texto	0,145	3,2	2	0,202
	Tablas	0	0	1	1
	Gráficas	0,182	2	1	0,157
Discusión	Total categoría	0,094	20,681	20	0,416
	Análisis	0,126	8,333	6	0,215
	Generalización	0,028	2,165	7	0,95
	Implicaciones	No es posible calcularlo, por tener un solo indicador.			
Total categoría Conclusiones	Total categoría	0,06	9,885	15	0,827
Referencias		0,063	8,953	5	0,111
Total prueba		0,145	294,094	156	0***

$p < 0.1^*$, $p < 0.05^{**}$, $p < 0.01^{***}$

Para valorar la concordancia de los jueces con respecto a la redacción de los ítems del inventario, se utilizó de igual forma, el coeficiente de concordancia de Kendall, cuyos resultados se presentan en la Tabla 6. Se puede observar que nuevamente el instrumento en su totalidad presentó una concordancia significativa entre los jueces con un W de Kendall = 0,121 ($p < 0.01$). Al hacer el análisis por categorías, sólo la categoría conclusiones presentó valores significativos con un $p < 0.01$; en este punto es importante anotar que el Paquete Estadístico para las Ciencias Sociales (SPSS) 22 no arrojó valores para esta categoría, debido a que todos los jueces presentaron la misma calificación “5” en los dos indicadores existentes. Por su parte, las categorías de introducción y resultados presentaron valores significativos con un $p < 0.1$. En cuanto a las subcategorías resumen, procedimiento y resultados generales presentaron un nivel de significancia al nivel de $p > 0.1$.

Tabla 6

Valores y nivel de significación del coeficiente de concordancia de Kendall, para evaluación de redacción de los ítems por parte de los jueces, en función de categorías y subcategorías.

Categoría	Subcategoría	Kendall P	Chi	Gl	Sig. Asintótica
Portada	Título	0,028	1,565	5	0,905
	Resumen	0,133	27,866	19	0,086*
	Total categoría	0,097	26,554	25	0,379
Introducción	Contexto	0,091	4	4	0,406
	investigativo				
	Propósitos de la investigación	0,155	8,529	5	0,129
Método	Total categoría	0,146	16,087	10	0,097*
	Diseño	0,091	1	1	0,317
	Participantes	0,124	35,478	26	0,102
	Instrumentos	0,091	1	1	0,317
	Definición de variables	0,091	2	2	0,368
	Procedimiento	0,132	33,452	23	0,074*
	Procesamiento estadístico	0,145	6,4	4	0,171
	Consideraciones éticas	0,096	11,563	11	0,397
	Total categoría	0,11	89,231	74	0,11
	General	0,148	21,109	13	0,071*
Resultados	Texto	0,209	4,588	2	0,101
	Tablas	0,182	2	1	0,157
	Graficas	0,182	2	1	0,157
	Total categoría	0,143	31,352	20	0,051*
	Análisis	0,091	6	6	0,423
Discusión	Generalización	0,1	7,696	7	0,36
	Implicaciones	No es posible calcularlo, hay un solo indicador			
	Total categoría	0,088	14,511	15	0,487
Conclusiones				1	
Referencias		0,067	3,681	5	0,596
Total prueba		0,121	206,896	156	0,004***

p< 0.1*, p< 0.05**, p< 0.01***

Con base en los análisis realizados para la validación por jueces se puede establecer que el inventario a nivel general presenta un alto nivel tanto de validez de contenido como de concordancia en las puntuaciones de los jueces teniendo en cuenta pertinencia y redacción.

5.2. Calidad metodológica y de publicación de los artículos evaluados

En este apartado se registran los resultados relacionados con la calidad metodológica y de publicación de los 45 artículos que componen la muestra documental. Específicamente se presentará el porcentaje de cumplimiento de cada uno de los indicadores, teniendo en cuenta las categorías y subcategorías del instrumento aplicado.

Para la descripción de los datos se utilizará como base el estadístico de posición denominado cuartil, el cual divide un conjunto de datos en cuatro partes iguales, cada una compuesta por un 25% (Ross, 2005). Los cuartiles han sido utilizados como medidas de los índices de impacto de las revistas, en donde las de mayor impacto se ubican en el primer cuartil, el segundo y tercer cuartil son para revistas de impacto medio y el cuarto cuartil es el de más bajo impacto (Universidad Nacional de Educación a Distancia [UNED], 2015). Para el caso de este estudio se retomara una clasificación similar en donde se propone las siguientes categorías para valorar el porcentaje de cumplimiento de los artículos.

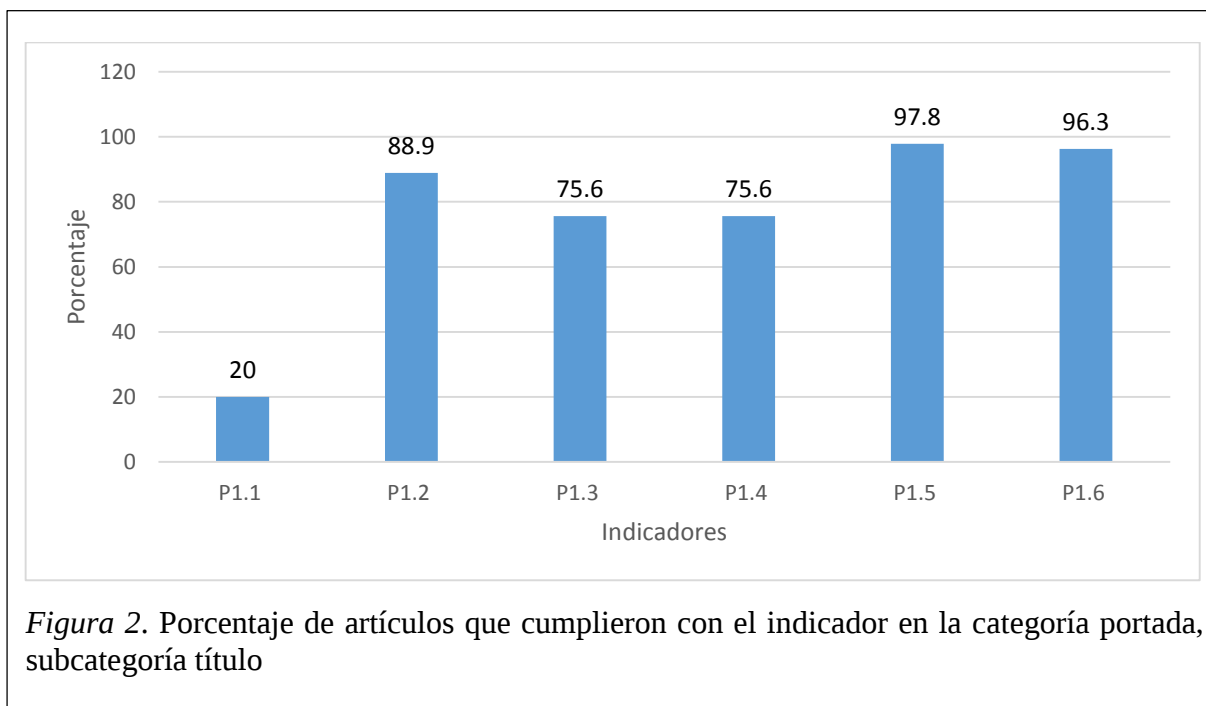
Cuartil 1 76-100%: Porcentaje de cumplimiento alto.

Cuartil 2 51-75%: Porcentaje de cumplimiento medio alto.

Cuartil 3 26-50%: Porcentaje de cumplimiento medio bajo.

Cuartil 4 0-25%: Porcentaje de cumplimiento bajo.

La Figura 2 presenta la frecuencia y porcentaje de artículos que cumplieron con el indicador en la categoría portada, en la subcategoría título, se destaca que la gran mayoría de los indicadores el porcentaje de cumplimiento superó el 75% y dos indicadores superaron el 90%. Los mayores valores estuvieron relacionados con la utilización de palabras completas en el título (97,8%) y la presentación del título de manera afirmativa (96.3%). La dificultad en esta subcategoría se dio en lo indicador relacionado con nombrar la palabra aleatoria o equivalente en el título (P1.1) en donde el porcentaje de cumplimiento solo alcanzó el 20%.



Con respecto a la subcategoría resumen, que se presenta en la Figura 3, ocho de los 20 indicadores lograron un porcentaje de cumplimiento superior al 66% destacándose los indicadores relacionados con la facilidad de comprensión (P2.1); la presentación de las

características esenciales del artículo (P2.3); y la presentación de las principales Conclusiones e implicaciones (P2.20), estos con un porcentaje de cumplimiento del 95,6%. Por otra parte, cuatro indicadores presentaron un porcentaje de cumplimiento inferior al 10%, en donde el menor valor se dio en el que hace referencia a la presentación del número de participantes que fueron objeto de análisis estadístico en los grupos control y experimental (P1.11), con un 4.4%

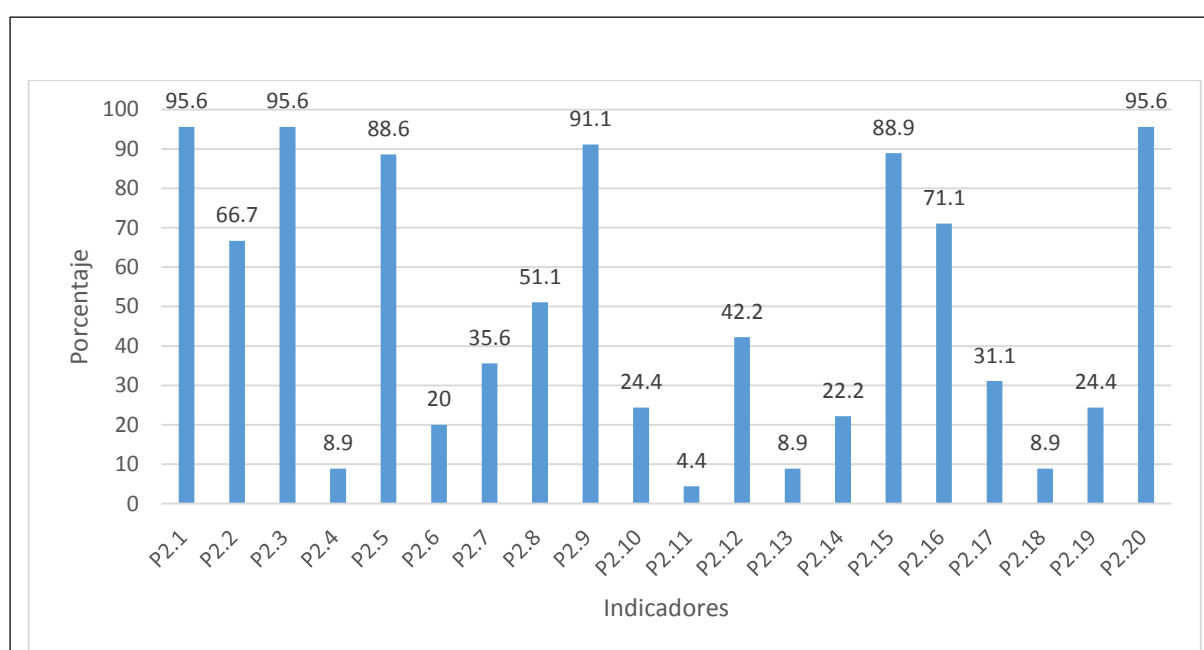
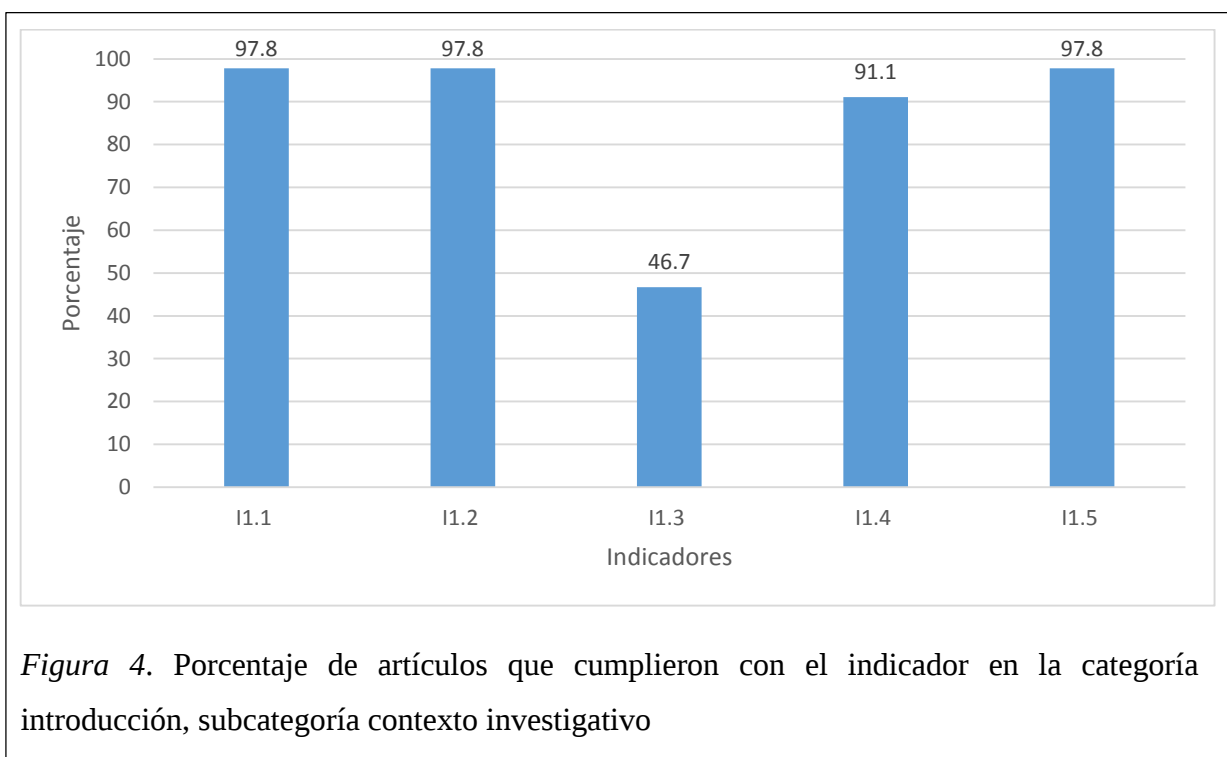


Figura 3. Porcentaje de artículos que cumplieron con el indicador en la categoría portada, subcategoría resumen.

En la Figura 4 se presenta la frecuencia y porcentaje de artículos que cumplieron con el indicador en la categoría introducción, subcategoría contexto investigativo. Se observa un nivel de cumplimiento superior al 90% en cuatro de los cinco indicadores propuestos, el único indicador que presentó un porcentaje medio bajo de cumplimiento, fue el que hace referencia a la presentación del marco conceptual que sustenta las hipótesis de investigación (46.7%); lo cual se debió principalmente a que en los artículos no se muestra de manera explícita las hipótesis que sustentan los estudios.



De otra parte, en la subcategoría propósitos de investigación, Figura 5, el indicador referido a la presentación explícita del objetivo y/o hipótesis de la investigación (I1.1) logró un porcentaje de cumplimiento del 93.3%, ubicándose en un nivel alto. Los otros presentaron porcentajes de cumplimiento equivalentes a medios bajos y bajos; encontrándose entre estos últimos los relacionados con la presentación explícita del problema de investigación (2.2%) y la relación entre la hipótesis y el problema de investigación (0%). Estas dificultades se fundamentaron en que en la gran mayoría de los artículos, no se presentaba de manera explícita el problema de investigación que sustentaba el estudio.

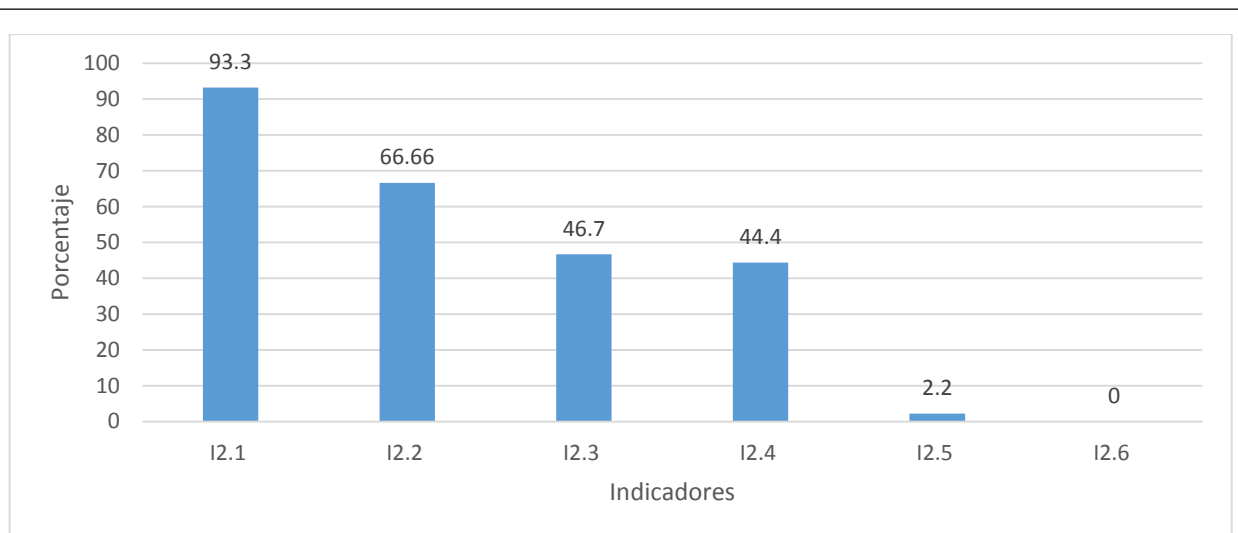


Figura 5. Porcentaje de artículos que cumplieron con el indicador en la categoría introducción, subcategoría propósitos de investigación

En la Figura 6 se presenta la frecuencia y porcentaje de artículos que cumplieron con el indicador en la categoría método, subcategoría diseño. Al observar los resultados, se puede evidenciar que el indicador relacionado a la descripción del diseño para que pueda realizarse una réplica de la investigación (M1.2, 77.8%) fue alto; sin embargo, en el caso del indicador que evalúa la presentación explícita del diseño de investigación (M1.1, 28,9%), se ubica en el nivel medio bajo.

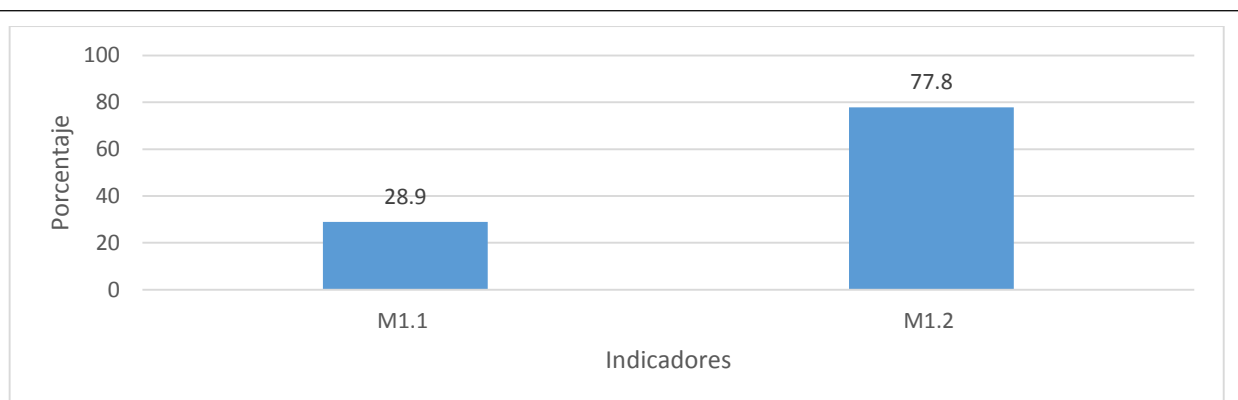


Figura 6. Porcentaje de artículos que cumplieron con el indicador en la categoría método, subcategoría diseño.

Al analizar los resultados obtenidos en la subcategoría de participantes, presentados en la Figura 7, se puede notar que existe una distribución aproximadamente equivalente entre los índices de cumplimientos altos, medios y bajos. El valor de cumplimiento más alto, se dio en el indicador que evalúa la presentación del número de participantes del estudio (M2.5, 95,6%), El porcentaje más bajo se dio en el indicador 21 referido a la presentación de la formación y experiencia de los evaluadores (M2.1, 0%).

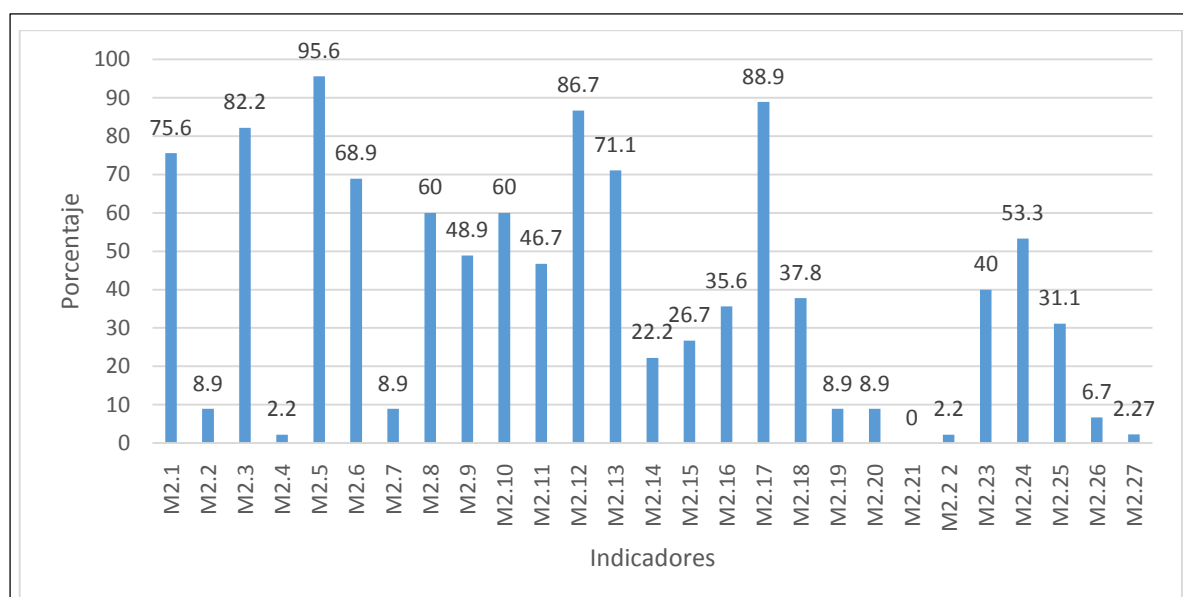


Figura 7. Porcentaje de artículos que cumplieron con el indicador en la categoría método, subcategoría participantes.

Al abordar la subcategoría instrumentos, presentadas en la Figura 8, se puede encontrar que, el cumplimiento del indicador donde se nombran y describen las pruebas utilizadas en el proceso investigativo (M3.1, 64,4%), fue medio alto; sin embargo, en el caso del indicador que evalúa si se presentan las propiedades psicométricas de los instrumentos (M3.2), alcanzó un porcentaje menor de cumplimiento del 31,1 %.

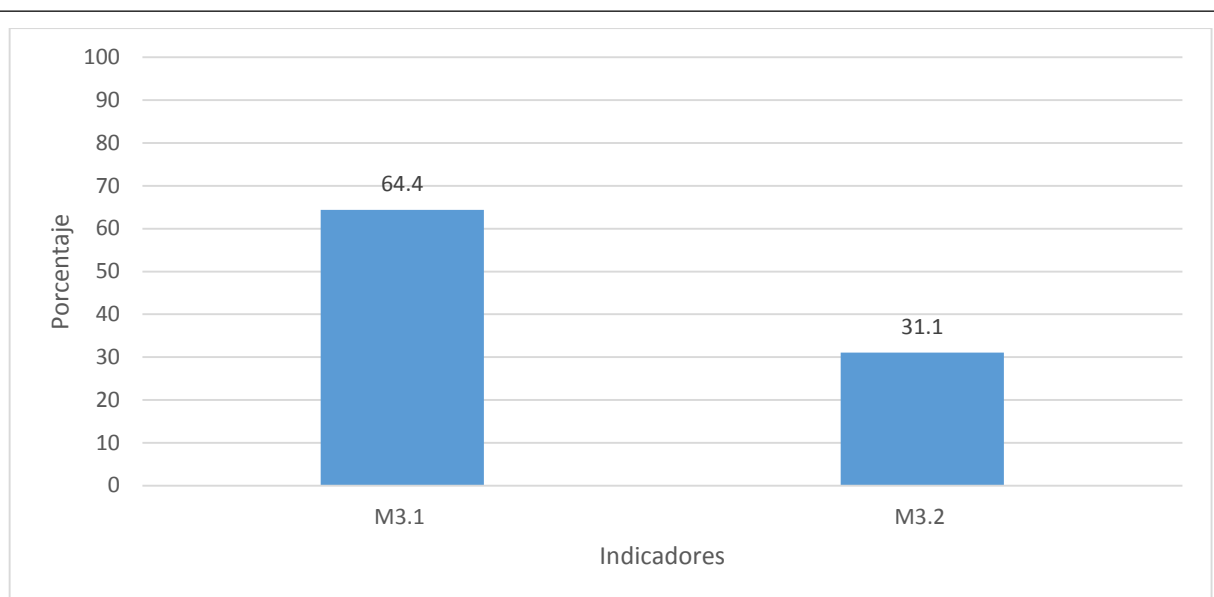


Figura 8. Porcentaje de artículos que cumplieron con el indicador en la categoría método, subcategoría instrumentos.

La subcategoría de definición de variables, Figura 9, presentó un nivel alto en el indicador referido a que se presentan de manera explícita las variables dependientes e independientes (M4.1, 97,8%); lo cual contrastó con los indicadores restantes en donde se dieron porcentajes bajos como el referido a la claridad en la definición de las variables (M4.3, 24,4%), y la definición operacional de las variables el (M4.2, 15,6%).

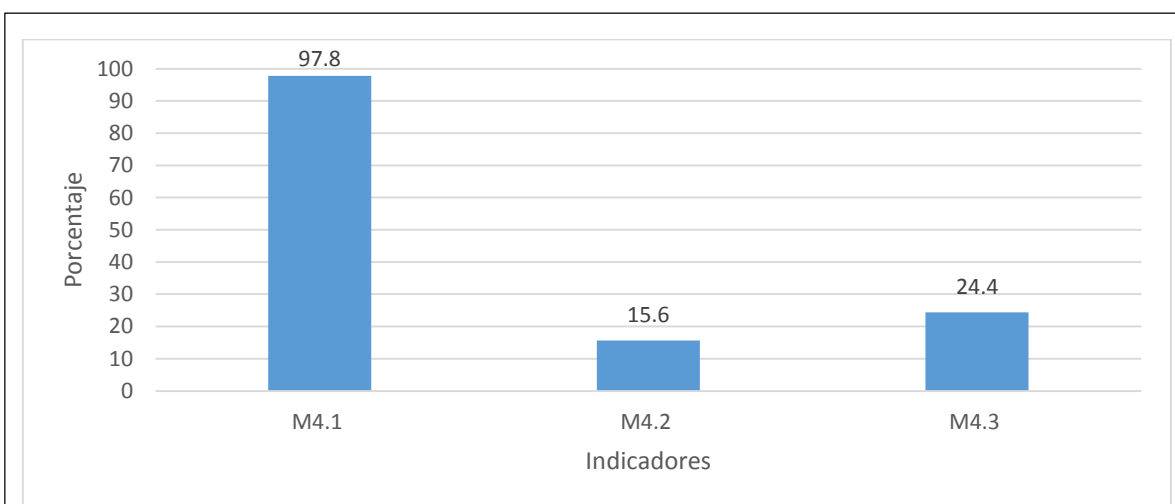
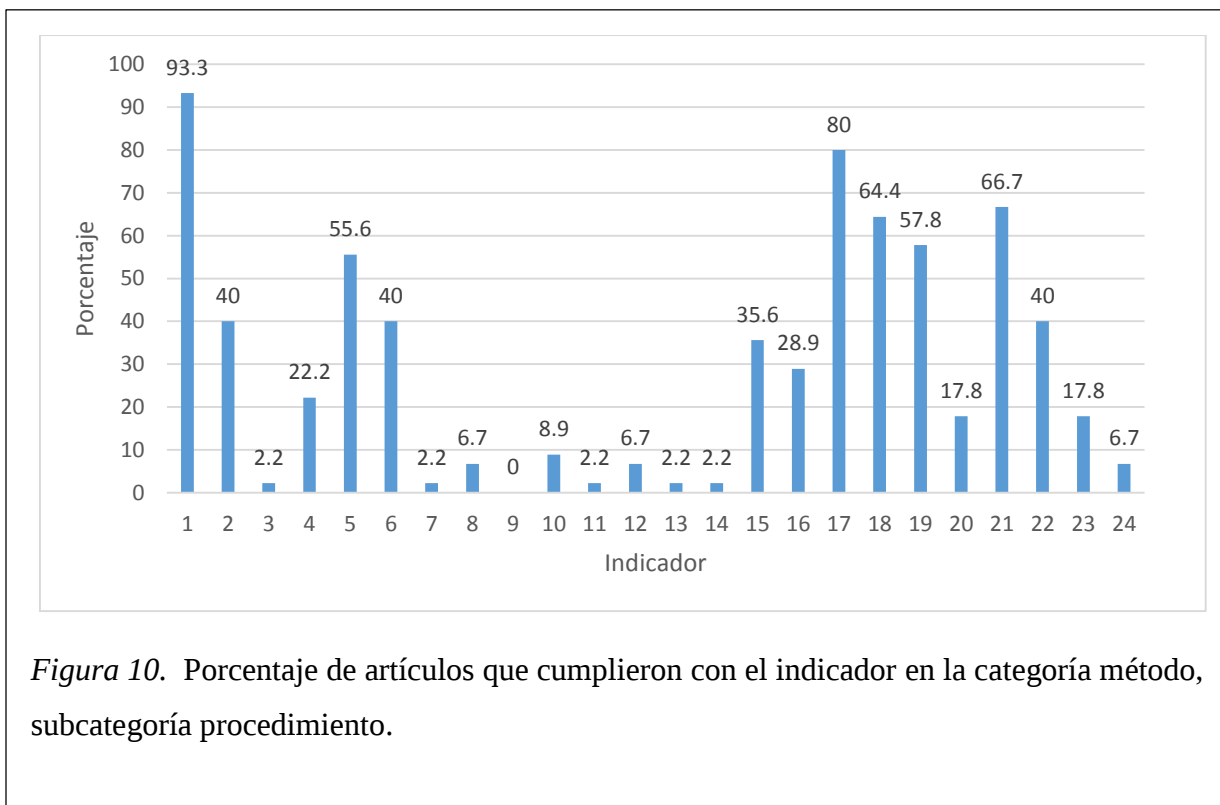


Figura 9. Porcentaje de artículos que cumplieron con el indicador en la categoría Método, subcategoría definición de variables.

En la Figura 10, se presentan los resultados de la subcategoría procedimiento en la cual, el caso más extremo se dio en el indicador que mide si los participantes fueron enmascarados (cegados) a la asignación de las intervenciones (M5.9, 0%); este ítem evidencia una falencia observada en la mayoría de los estudios y es la ausencia de procesos de enmascaramiento para los participantes. Por otra parte, los porcentajes de cumplimiento más altos se presentaron en el indicador que evalúa la presentación del procedimiento de investigación (M5.1, 93.3%) y en el referido a la descripción de la forma en que se realizó la evaluación (M5.17, 80%).



Los resultados de la subcategoría procesamiento estadístico, se pueden observar en la Figura 11, la gran mayoría de los ítems se localizaron en un nivel medio y solo un indicador se ubicó en el nivel bajo, en donde se evalúa si el estudio informa sobre circunstancias del análisis de datos que pueden afectar la validez de los resultados (M6.5, 8,9%).

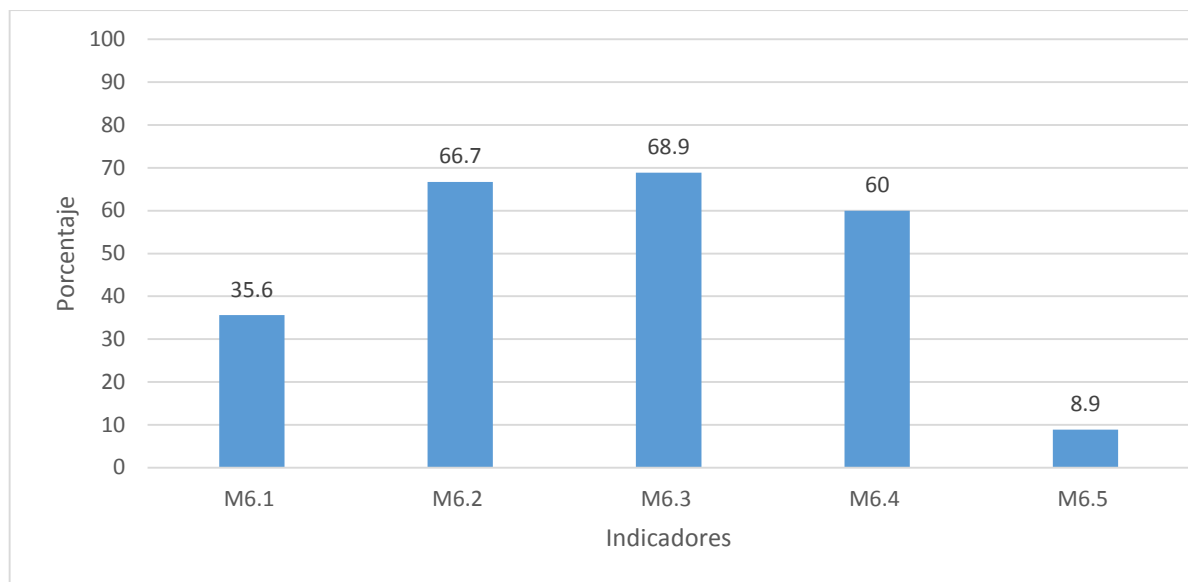


Figura 11. Porcentaje de artículos que cumplieron con el indicador en la categoría método, subcategoría, procesamiento estadístico.

En la subcategoría consideraciones éticas, Figura 12, se observó que casi la totalidad de los ítems se ubicaron en las categorías medio bajo y bajo, lo cual era previsible si se tiene en cuenta que ha sido un aspecto que solo en los últimos años ha empezado a adquirir relevancia y a ser considerada como algo central y esencial en el proceso investigativo. El único indicador que presentó un nivel medio alto fue el relacionado con la utilización del consentimiento informado (M7.2, 60%).

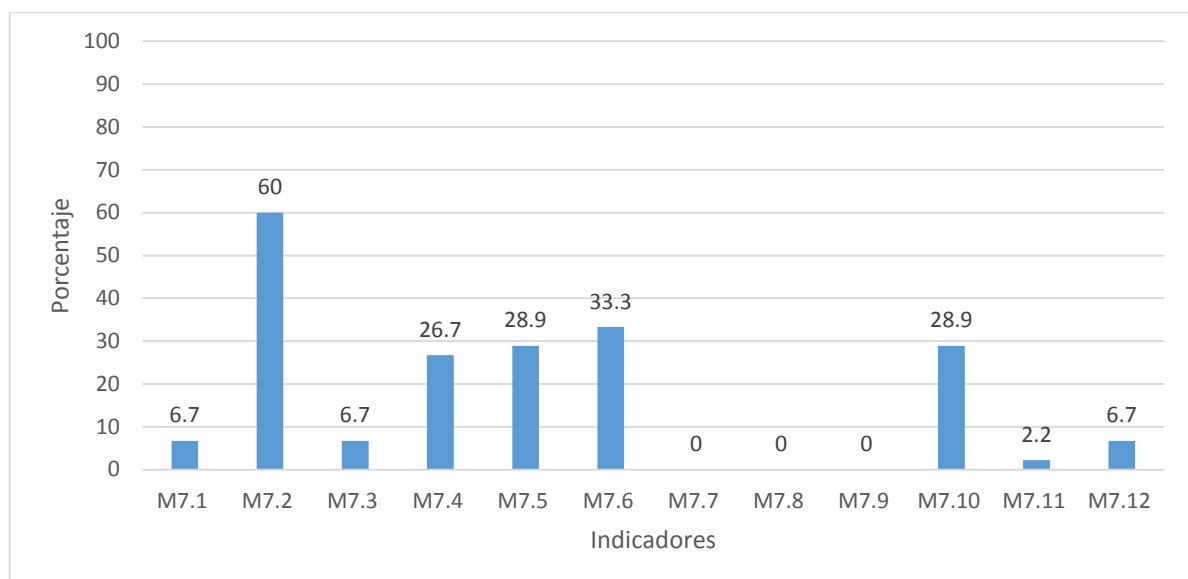


Figura 12. Porcentaje de artículos que cumplieron con el indicador en la categoría método, subcategoría consideraciones éticas.

En la Figura 13 se presentan los datos obtenidos en la categoría de resultados, subcategoría general. Esta subcategoría se encuentra compuesta por 14 indicadores, de los cuales 11, se localizaron en los niveles medio alto y alto. Los indicadores con valores más altos fueron los datos se presentan de forma organizada (R1.1 93.3%) y la presentación del valor de error estadístico (R1.8, 95,6%). El indicador de cumplimiento más bajo se dio lo referente al reporte de los grados de libertad para el análisis estadístico realizado (R1.10, 4.4%). Evidenciando una gran falencia en este aspecto.

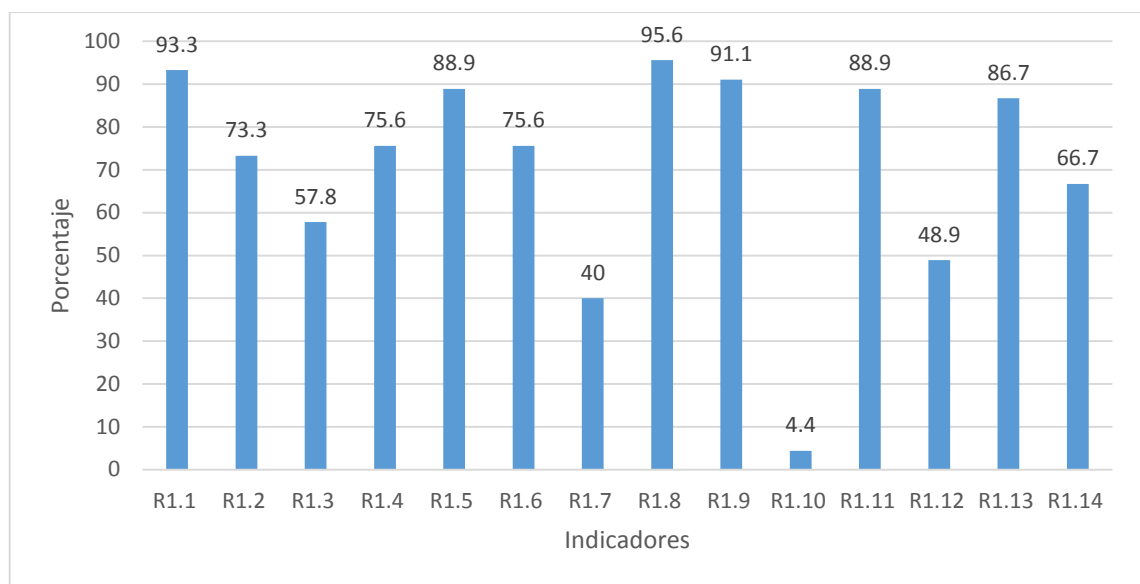


Figura 13. Porcentaje de artículos que cumplieron con el indicador en la categoría resultados, subcategoría general.

Los datos de la subcategoría texto, pueden ser observados en la Figura 14, los indicadores relacionados con la claridad del texto (R2.1) y la justificación de las causas para que existan valores perdidos en la investigación. (R2.2), presentaron un nivel de cumplimiento medio alto y medio bajo respectivamente; por su parte el indicador referido a la complementariedad entre el texto y las Tablas (R2.3) logró un nivel alto (93,3%).

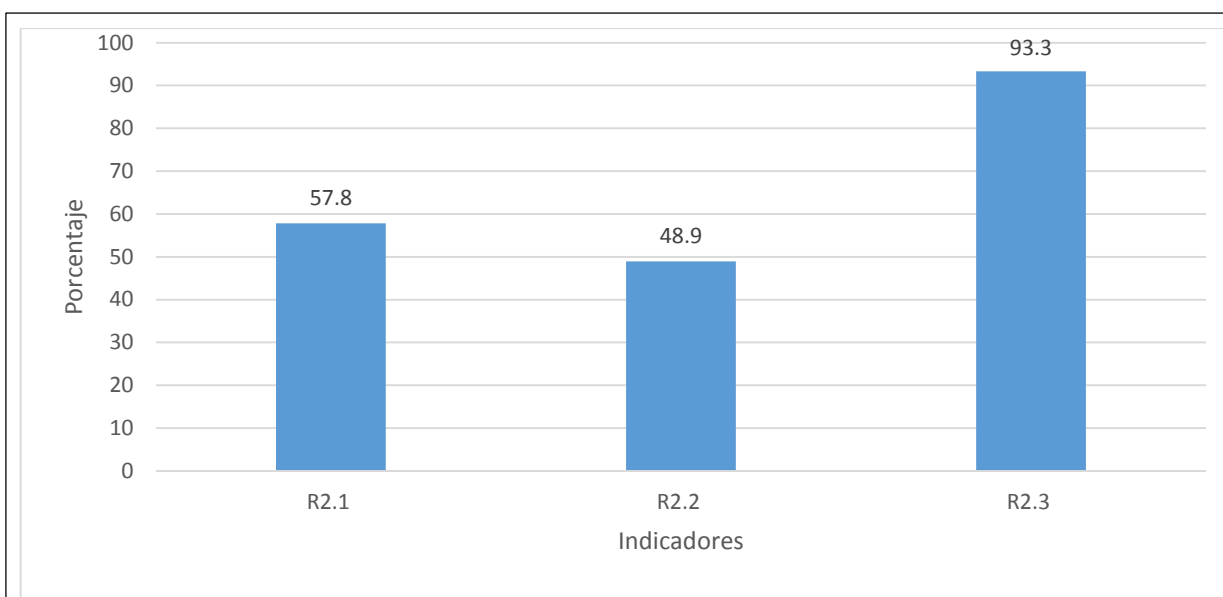


Figura 14. Porcentaje de artículos que cumplieron con el indicador en la categoría resultados, subcategoría texto.

Por su parte la subcategoría, tablas, Figura 15, presentó un indicador en nivel medio alto, relacionado con la presentación de la Tabla con la línea de base de variables demográficas y clínicas (R3.1, 66,7%), y el otro en alto, referido a la claridad de las tablas (R3.2, 82,2%). En lo referente a gráficas los porcentajes de cumplimiento fueron medio alto (R4.1, 60%) y medio bajo (R4.2, 47%).

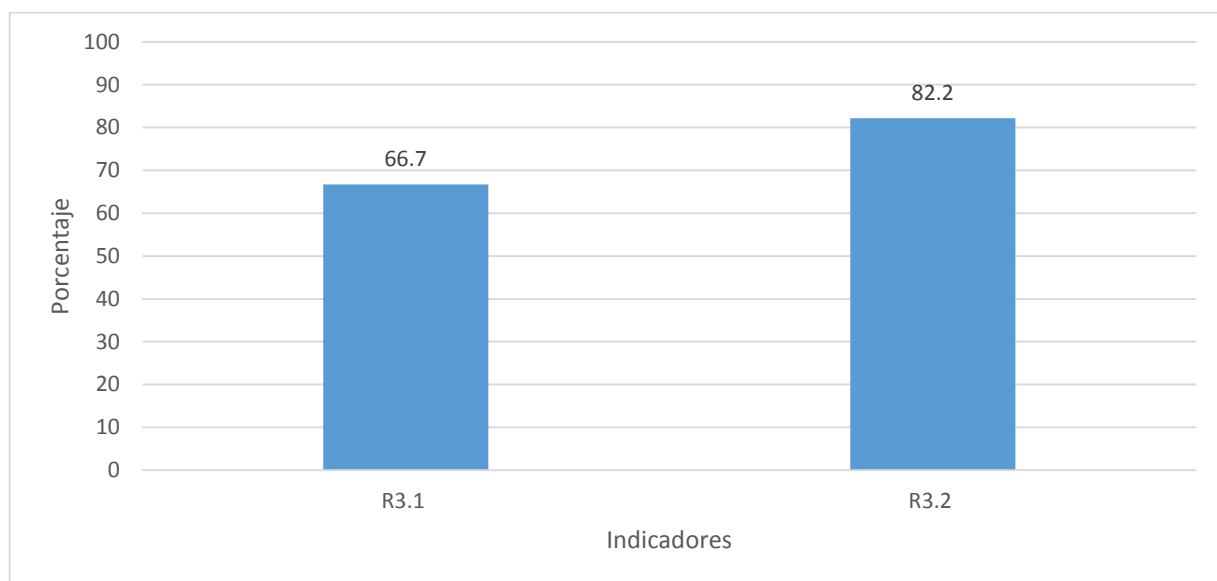
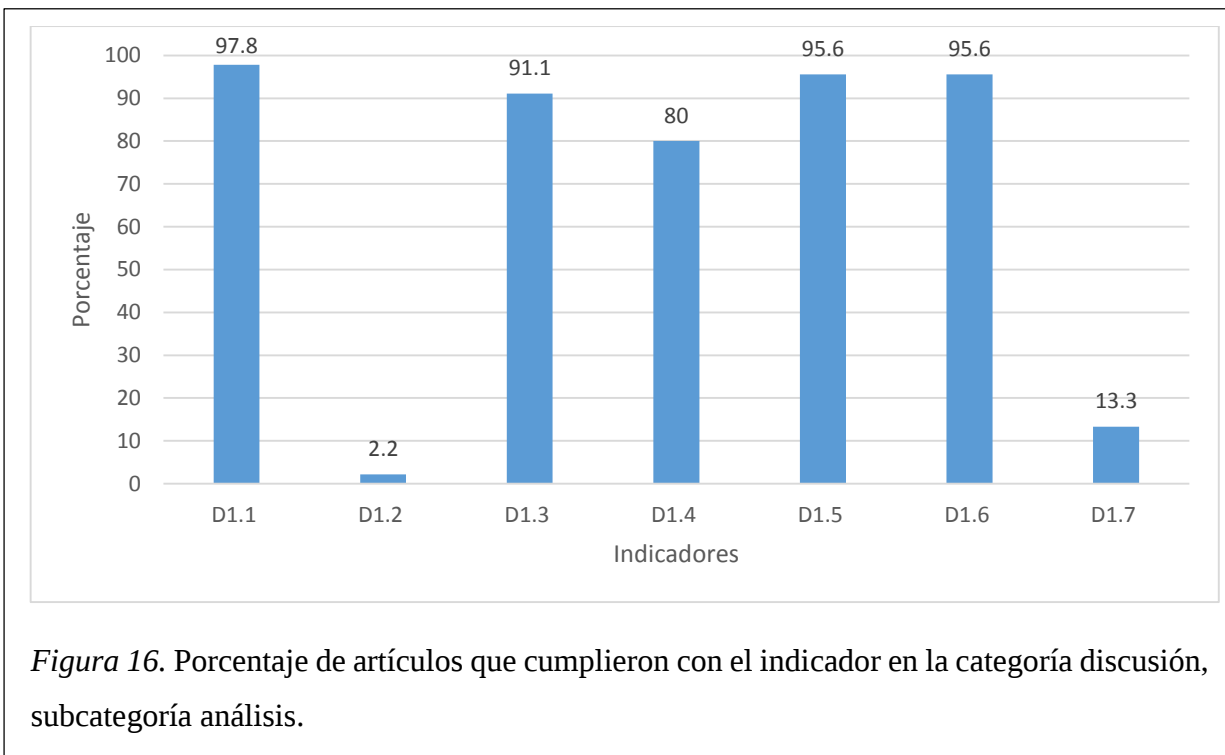


Figura 15. Porcentaje de artículos que cumplieron con el indicador en la categoría resultados, subcategoría Tablas.

La categoría Discusión, subcategoría análisis, presenta los resultados en la Figura 16. Se puede observar un comportamiento polarizado; por ejemplo, dos ítems se ubicaron en nivel bajo, el referido a si los análisis están relacionados con el problema de investigación (D1.2, 2.2%) y la presentación de los efectos secundarios de las intervenciones (D1.7, 13,3%). Los otros indicadores por el contrario presentaron un alto nivel de cumplimiento alto; teniendo mayor nivel el referido a la relación entre el análisis de los datos y los objetivos y/o hipótesis del estudio (D1.1, 97,8%).



Un comportamiento similar a la anterior subcategoría, se presentó en la de generalización, en donde nuevamente los datos presentaron valores altos y bajos. En la Figura 17, se observa que el porcentaje más alto se presentó en el indicador relacionado con justificar las razones por las cuales se hacen las generalizaciones (D2.8, 97,8%) y el más bajo se dirigía a evaluar la presentación de los métodos de recolección de datos a los cuales puede ser generalizado el estudio (D2.5, 8.88%). En lo referente a la subcategoría de implicaciones de la investigación, el único ítem arrojó un valor alto (88.9%), lo cual evidencia una fortaleza, en este aspecto de los estudios evaluados.

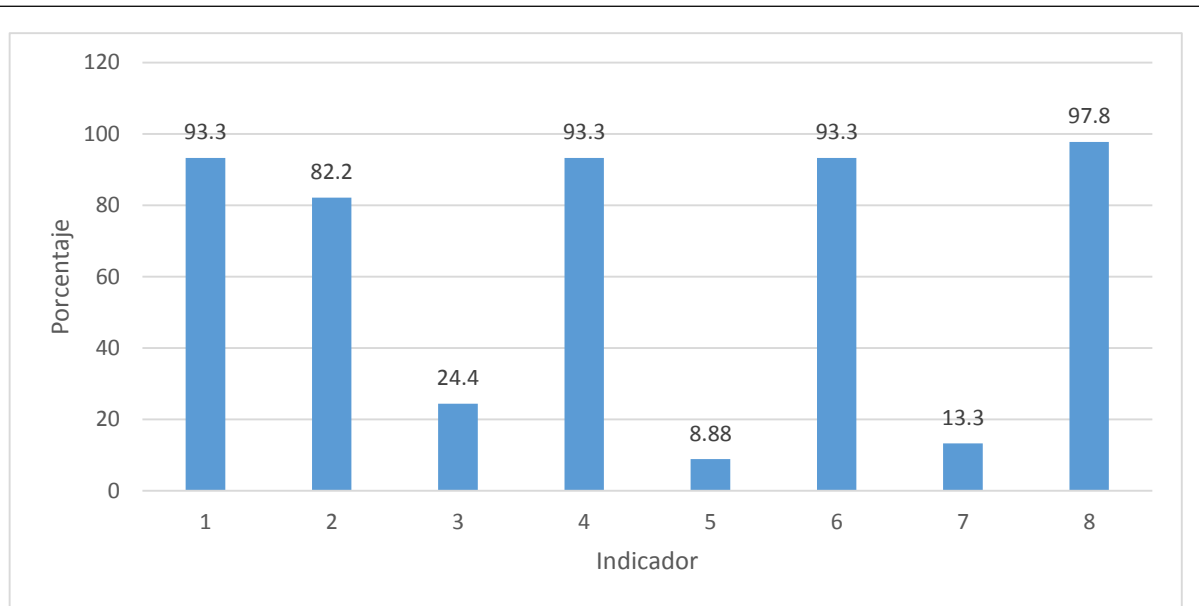
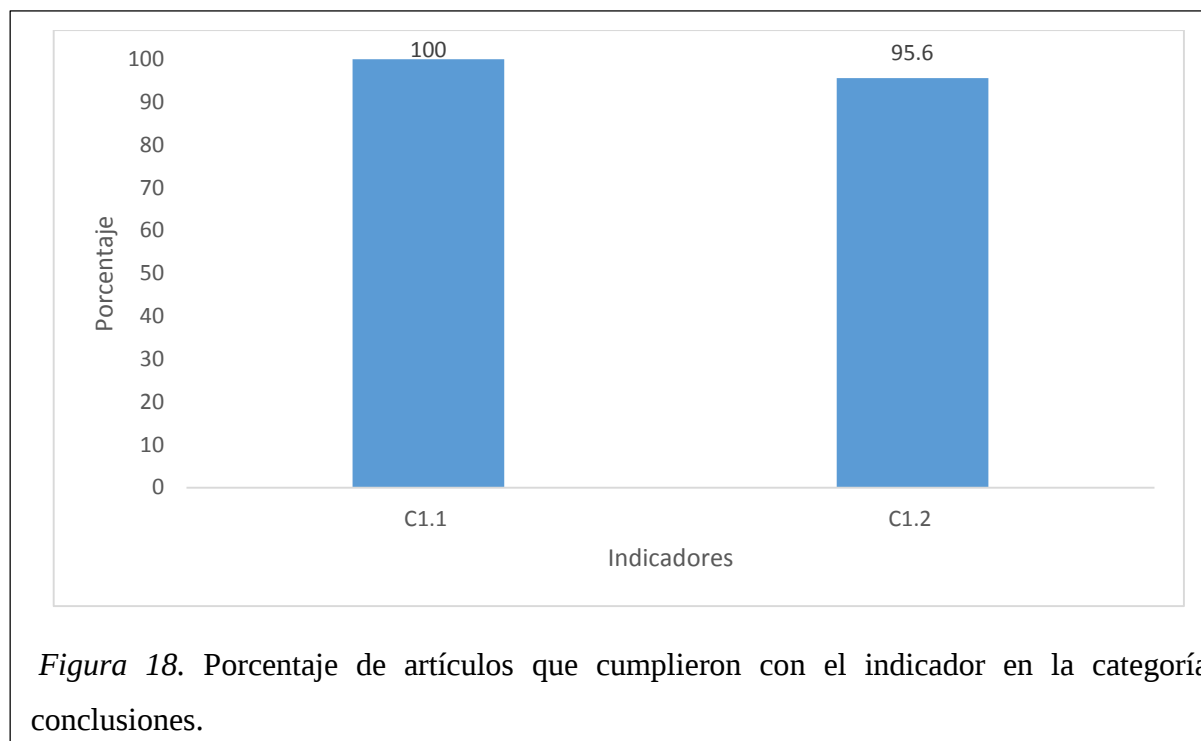


Figura 17. Porcentaje de artículos que cumplieron con el indicador en la categoría discusión, subcategoría generalización.

La categoría conclusiones, como se puede ver en la Figura 18, presentó porcentajes de cumplimiento altos en los dos ítems que la componen. El primero evalúa si las Conclusiones están basadas en los resultados (C1.1, 100%), y el segundo la relación entre las Conclusiones y los objetivos (95,6%), lo cual evidencia una fortaleza en este aspecto, en lo referente al reporte de las investigaciones.



La última categoría del instrumento, referencias, presenta los datos en la Figura 19, Se evidencia que la gran mayoría de los indicadores mostraron porcentajes altos, lográndose un cumplimiento del 100% en tres indicadores, el referido a la relación de las referencias con la temática central del artículo (RF1.1), el que evalúa el seguimiento de normas (RF1.5), como también la utilización de fuentes confiables (RF1.6). Sin embargo una falencia que se evidenció en esta categoría, fue la del indicador que evalúa la presencia de referencias actualizadas (RF1.2), en donde solo se logró un porcentaje de cumplimiento del 8.9%.

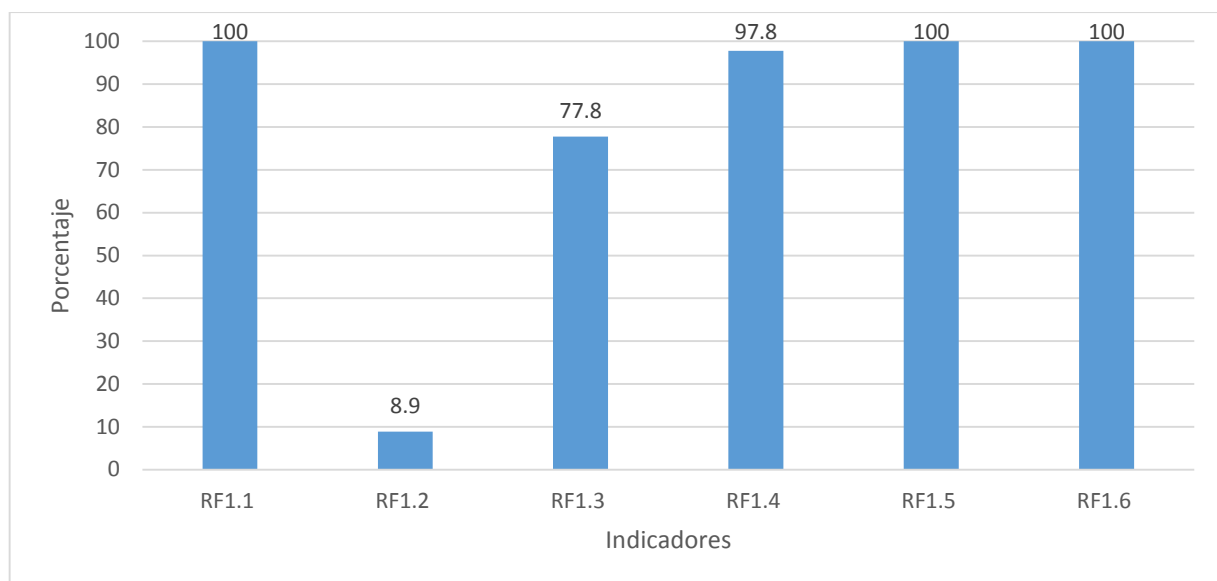


Figura 19. Porcentaje de artículos que cumplieron con el indicador en la categoría Referencias.

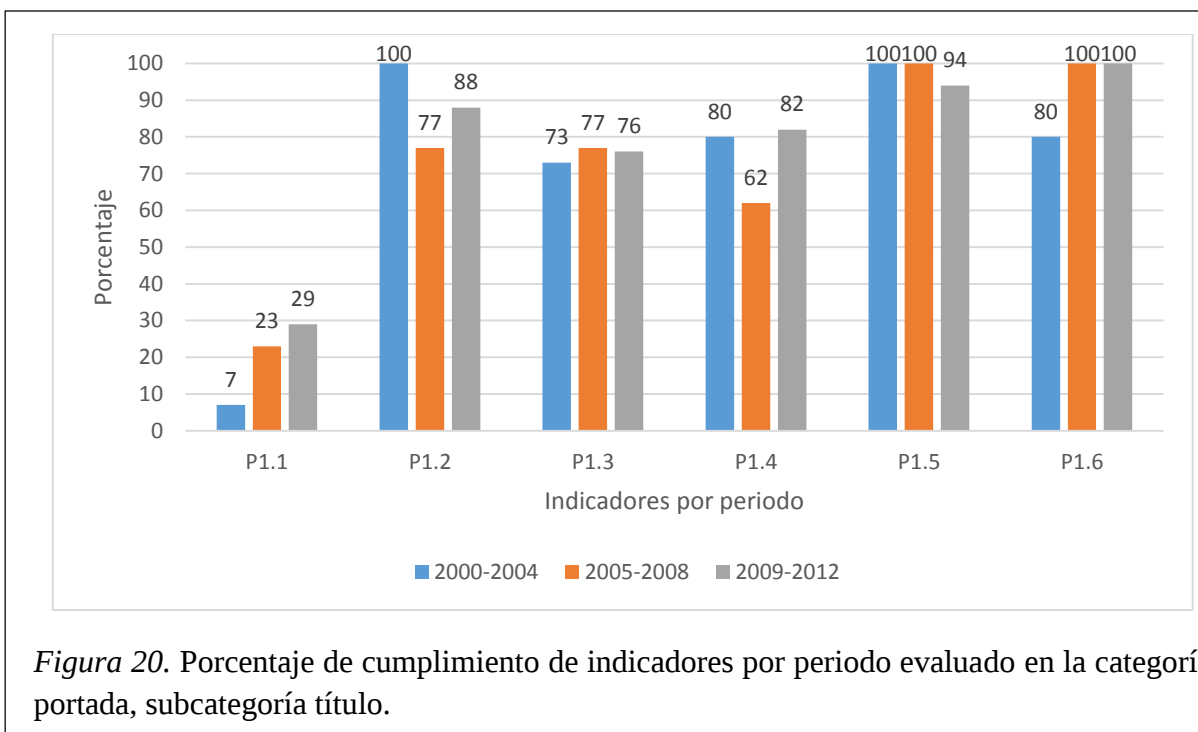
5.3 Evolución de la calidad metodológica y de publicación de los artículos evaluados 2000-2012.

En esta apartado se presentará el porcentaje de cumplimiento de los indicadores por periodo evaluado en cada una de las categorías y subcategorías del instrumento. También se utilizará un análisis inferencial no paramétrico con el Test Exacto de Fisher, el cual permite determinar asociaciones entre dos variables dicotómicas, cuando la muestra a estudiar es pequeña y por lo tanto no se cumplen las condiciones para la aplicación de la prueba Chi-Cuadrado (Peralta, 2005). El nivel mínimo de significación aceptado es $p < 0.05$. En el caso particular de este estudio se intenta determinar si existe asociación entre el porcentaje de cumplimiento de cada indicador y los periodos evaluados, a saber 2000-2004, 2005-2008 y 2009-2012. Teniendo en cuenta lo anterior se plantean las siguientes hipótesis

H0 El cumplimiento o no del indicador, es independiente o no tiene relación con el periodo evaluado

H1 Existe asociación entre el cumplimiento o no del indicador y el periodo evaluado

La Figura 20 muestra el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría portada, subcategoría, título, como se puede observar a nivel general el porcentaje de cumplimiento los indicadores en esta subcategoría es variable, pues a nivel general no se evidencia una tendencia clara de predominio, de alguno de los periodos. Como aspectos a destacar se encuentran los altos porcentajes de cumplimiento del indicador relacionado con la utilización de palabras completas en el título (P1.5) y la presentación del título de manera afirmativa (P1.6). En contraste el indicador relacionado nombrar en el titulo la palabra aleatorio (P1.1) presentó un porcentaje de cumplimiento bajo; pero llama la atención el hecho de que muestra una tendencia de aumento progresivo del porcentaje de cumplimiento desde el periodo más antiguo (2000-2004) al más reciente (2004-2009). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



La Figura 21 muestra el porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría Portada, subcategoría resumen, como se puede evidenciar los datos arrojados

por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en seis de los 20 indicadores relacionados con facilidad en la comprensión del resumen (P2.1); presentación de las características esenciales del artículo (P2.3); presentación del objetivo y/o hipótesis de la investigación (P2.5); presentación del número de participantes (P2.9); presentación de la intervención realizada al grupo experimental (P2.15) y presentación de las Conclusiones e implicaciones (P2.20).

Por otra parte en cuatro de los indicadores se evidenció un aumento progresivo del porcentaje de cumplimiento, desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales estaban relacionados con hacer evidente el marco conceptual que sustenta la investigación (P2.6); presentar el diseño de investigación (P2.7); nombrar los instrumentos utilizados (P2.14) y presentar el procedimiento utilizado para obtener la información (P2.18). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.

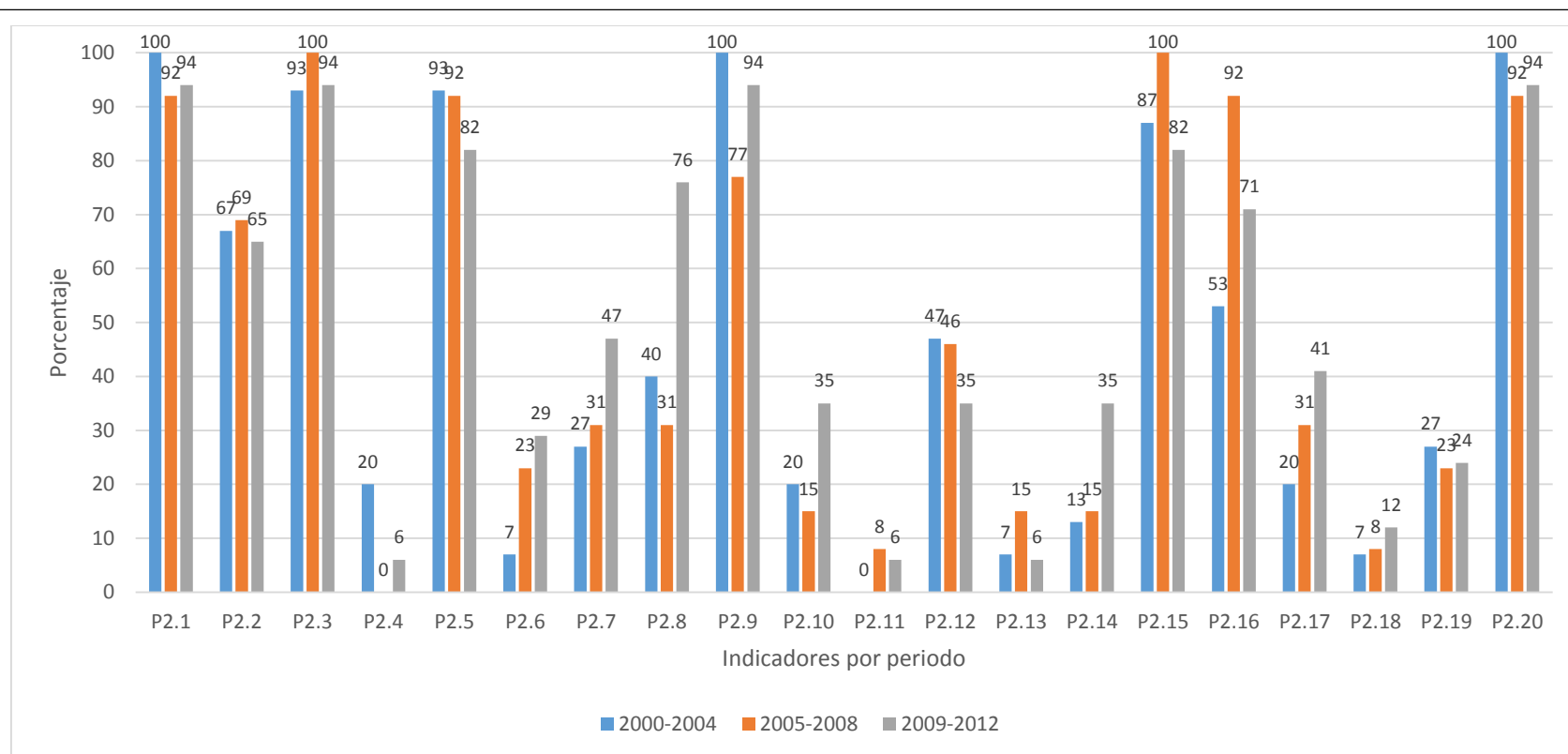


Figura 21. Porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría portada, subcategoría resumen.

La Figura 22 muestra porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría Introducción, subcategoría contexto investigativo, como se puede observar el comportamiento de los indicadores en esta subcategoría es variable, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Se encontraron altos porcentajes de cumplimiento en todos los indicadores a excepción del relacionado con la presentación de un marco conceptual que sustente las hipótesis de investigación (I1.3). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.

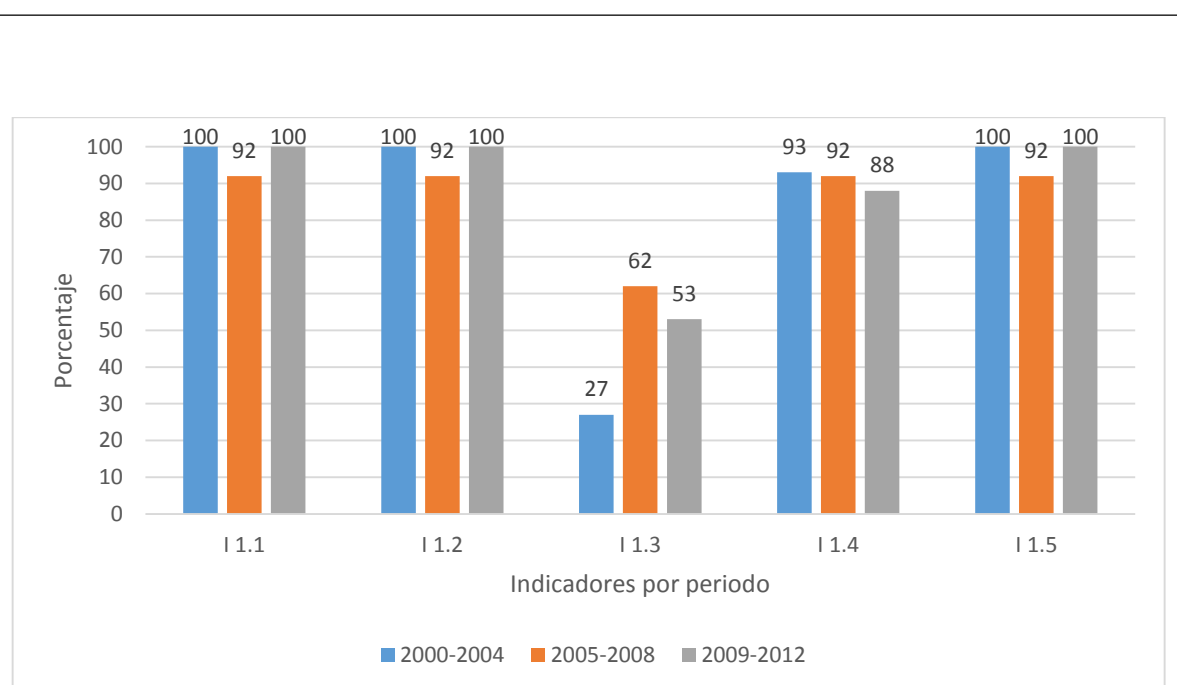
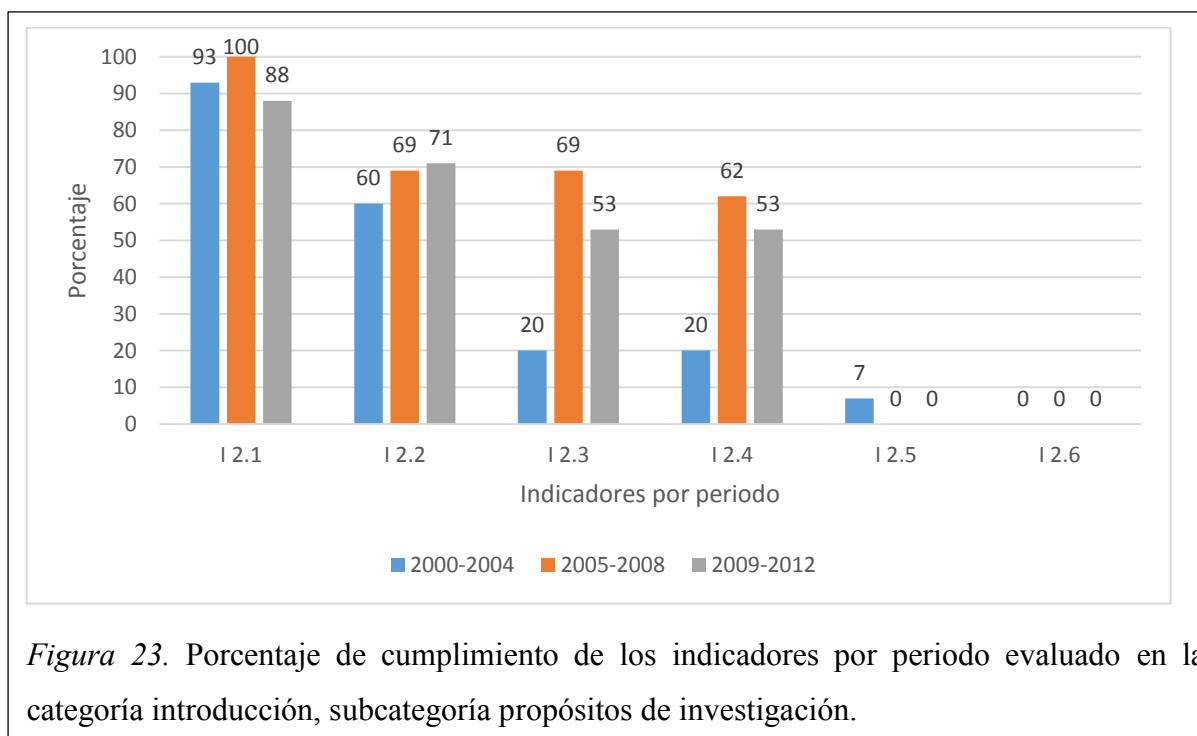


Figura 22. Porcentaje de cumplimiento de indicadores por periodo evaluado en la categoría introducción, subcategoría, contexto investigativo.

La Figura 23 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría introducción, subcategoría propósitos de investigación, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar los altos porcentajes de cumplimiento en el indicador relacionado con la presentación del objetivo y/o hipótesis de la investigación (I2.1) y el bajo porcentaje de cumplimiento de los indicadores

referidos a la presentación del problema de investigación (I2.5) y la relación entre la hipótesis y el problema de investigación (I2.6). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



En la Figura 24, se puede observar el porcentaje de cumplimiento por periodo evaluado en la categoría método, subcategoría diseño, se puede evidenciar que el porcentaje de cumplimiento del indicador referido a la presentación del diseño de investigación (M1.1) fue más bajo en todos los periodos, con respecto al indicador relacionado con la descripción del diseño, para que pueda replicarse la investigación (M1.1). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en los dos indicadores de la subcategoría.

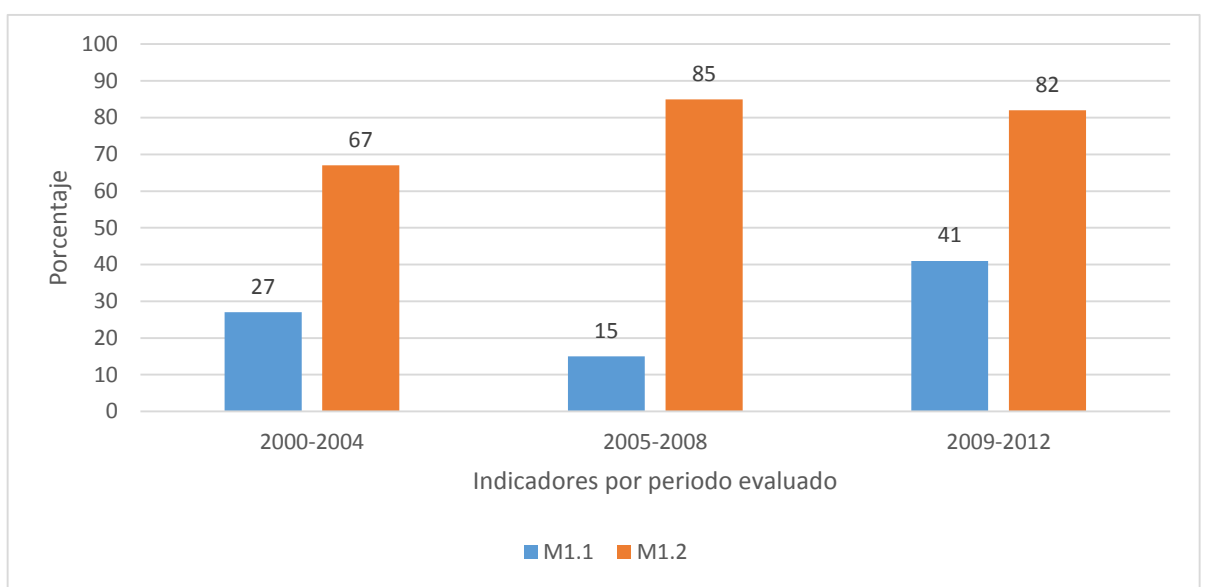


Figura 24. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría método, subcategoría diseño.

La Figura 25, ilustra el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría método, subcategoría sujetos, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en cinco de los 27 indicadores relacionados con la descripción del proceso de selección de la muestra (M2.3); presentación del número de participantes del estudio (M2.5); presentación de las características de la muestra (M2.12); especificación del número de pacientes tratados en cada lugar donde se hace la intervención (M2.17) y la presentación del número de lugares en donde se hace la intervención para cada grupo (M2.27).

En contraste, se presentaron bajos porcentajes de cumplimiento en ocho indicadores relacionados con la presentación de las características de la población investigada (M2.2); descripción de la forma en que se calculó el número de sujetos de la muestra (M2.4); especificación del número de personas en los grupos experimental y control que recibieron el tratamiento planeado (M2.7); identificar la persona que generó la secuencia de asignación al azar (M2.19); identificar la persona que asignó los participantes a las intervenciones o

grupos (M2.20); presentación de la formación y experiencia de los evaluadores (M2.21), presentación de los criterios de selección, para los terapeutas (M2.22) y presentación de los criterios de selección de los lugares donde se realizó la investigación (M2.26).

Por su parte en siete indicadores se evidencia un aumento progresivo del porcentaje de cumplimiento desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales estaban relacionados con descripción de los sujetos, para que pueda replicarse la investigación (M2.1); presentación de las características de la población investigada (M2.2); presentación del número de participantes de los grupos experimental y control (M2.6); presentación del número de personas en los grupos experimental y control que completaron la intervención (M2.8); presentación de las razones de abandono después de la asignación al azar, por parte de las personas los grupos experimental y control (M2.11); presentación de las características de la muestra (M2.12) y presentación del el diagrama de flujo de participantes (M2.15). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.

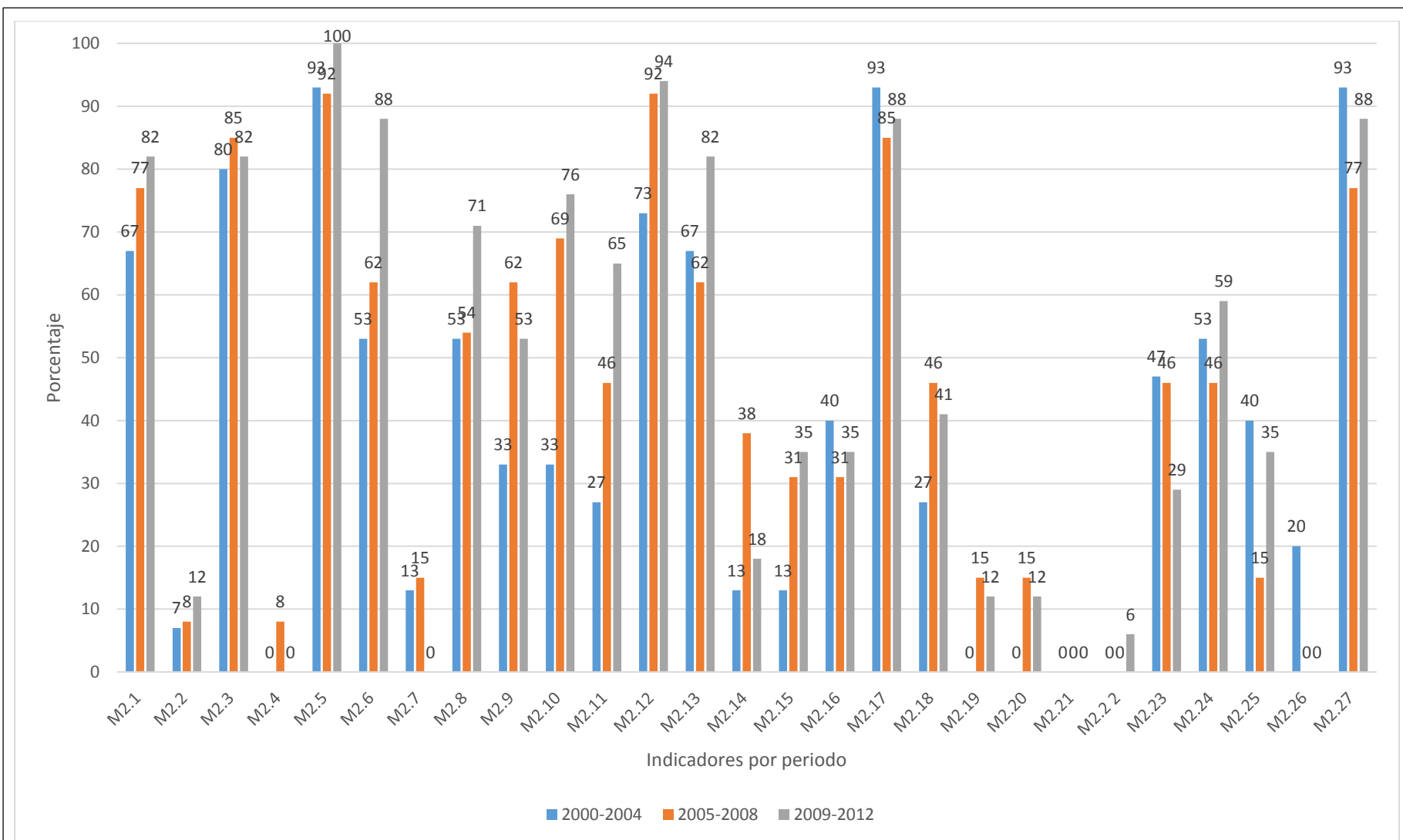
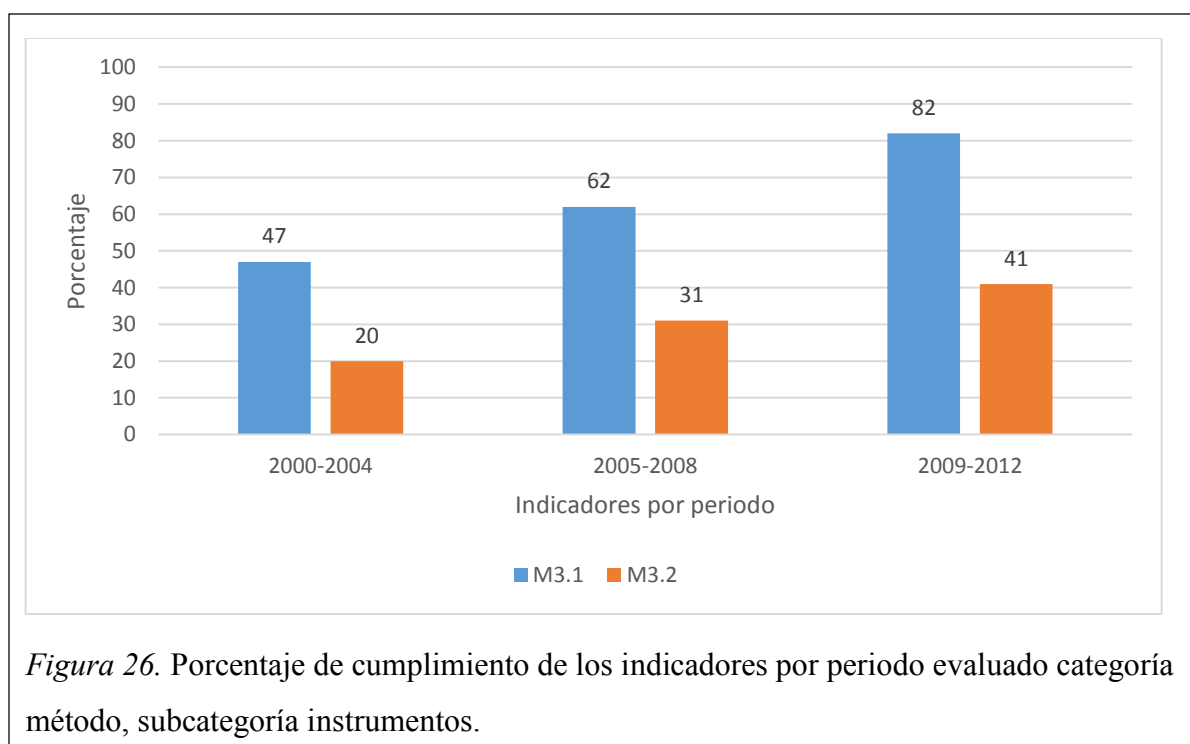


Figura 25. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría método, subcategoría sujetos.

La Figura 26 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría instrumentos, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría presentan una tendencia ascendente desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), tanto en el indicador relacionado con nombrar y describir las pruebas y/o aparatos utilizados en el proceso investigativo (M3.1) como en el indicador que se refiere a la presentación de las propiedades psicométricas de las pruebas y/o aparatos utilizados en el proceso investigativo (M3.2). Otro resultado de relevancia es que los porcentajes de cumplimiento más altos se presentaron sistemáticamente en el indicador M3.1 con respecto al indicador M3.2. Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría sujetos.



La Figura 27 ilustra el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría método, subcategoría definición de variables, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría a nivel general no evidencian una tendencia clara de predominio en alguno de los periodos evaluados. Entre los

resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en el indicador relacionado con la presentación de las variables dependientes e independientes (M4.1), y por el contrario se encontraron bajos porcentajes de cumplimiento en el indicador referido a la definición operacional de las variables y sus grados (M4.2). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.

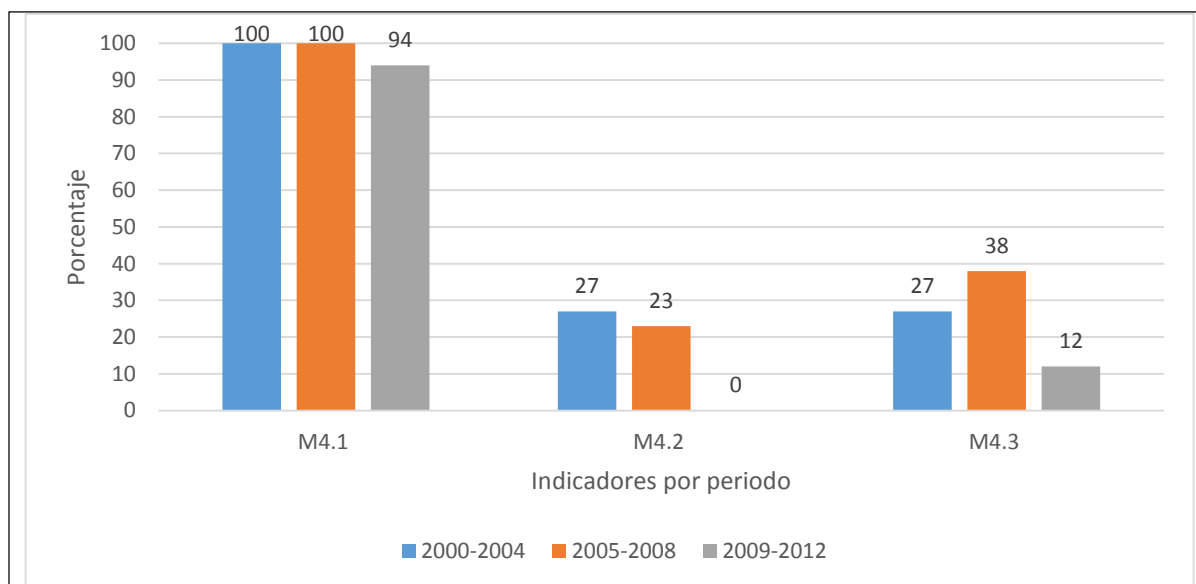


Figura 27. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría Método, subcategoría definición de variables.

El porcentaje de cumplimiento de los indicadores por periodo evaluado categoría método, subcategoría procedimiento, como se puede evidenciar en la Figura 28 los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontró un alto porcentaje de cumplimiento en el indicador relacionados con presentación del procedimiento de la investigación (M5.1).

Por otra parte, se presentaron bajos porcentajes de cumplimiento en diez indicadores los cuales hacen referencia a la presentación del número de registro y el nombre de registro del experimento (M5.3); descripción de la forma en que se ocultó la secuencia de asignación al azar, hasta el momento en que se asignan las intervenciones (M5.7); evidencia de enmascaramiento de los participantes con respecto a las intervenciones (M5.8); descripción de la forma en que los participantes fueron enmascarados con respecto a las intervenciones (M5.9); evidencia que los terapeutas fueron enmascarados con respecto a las intervenciones (M5.10), explicación de la forma en que los terapeutas fueron enmascarados con respecto a las intervenciones (M5.11); evidencia de que los evaluadores fueron enmascarados con respecto a las intervenciones (M5.12); explicación de cómo evaluadores fueron enmascarados o cegados con respecto a las intervenciones (M5.13); la forma en que fueron asignados los terapeutas a los grupos control y experimental (M5.14) y presentación de los aspectos que pueden afectar la validez de los resultados (M5.24). Se puede notar que un grave problema que se presenta en el procedimiento es el enmascaramiento de los diferentes actores que participan en el proceso investigativo.

En cinco indicadores se evidencia un aumento progresivo del porcentaje de cumplimiento desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales están relacionados la presentación de los periodos de reclutamiento de los sujetos (M5.4); se especifica cómo se hizo el control de variables extrañas (M5.15); se evidencia que el control de variables extrañas fue riguroso (M5.16); se describe el qué momento en que se implementó la evaluación (M5.18) y se hace explícita la forma en que se estandarizaron las intervenciones (M 5.22). Al realizar el análisis con el Test Exacto de Fisher, se encontraron asociaciones significativas entre los periodos 2000-20004 y 2005-2008, con respecto al porcentaje de cumplimiento con un $p > 0,05$, en los indicadores referidos a que se la forma en que se hizo el control de variables extrañas M5.15 y se evidencia que el control de variables extrañas fue riguroso (M5.16).

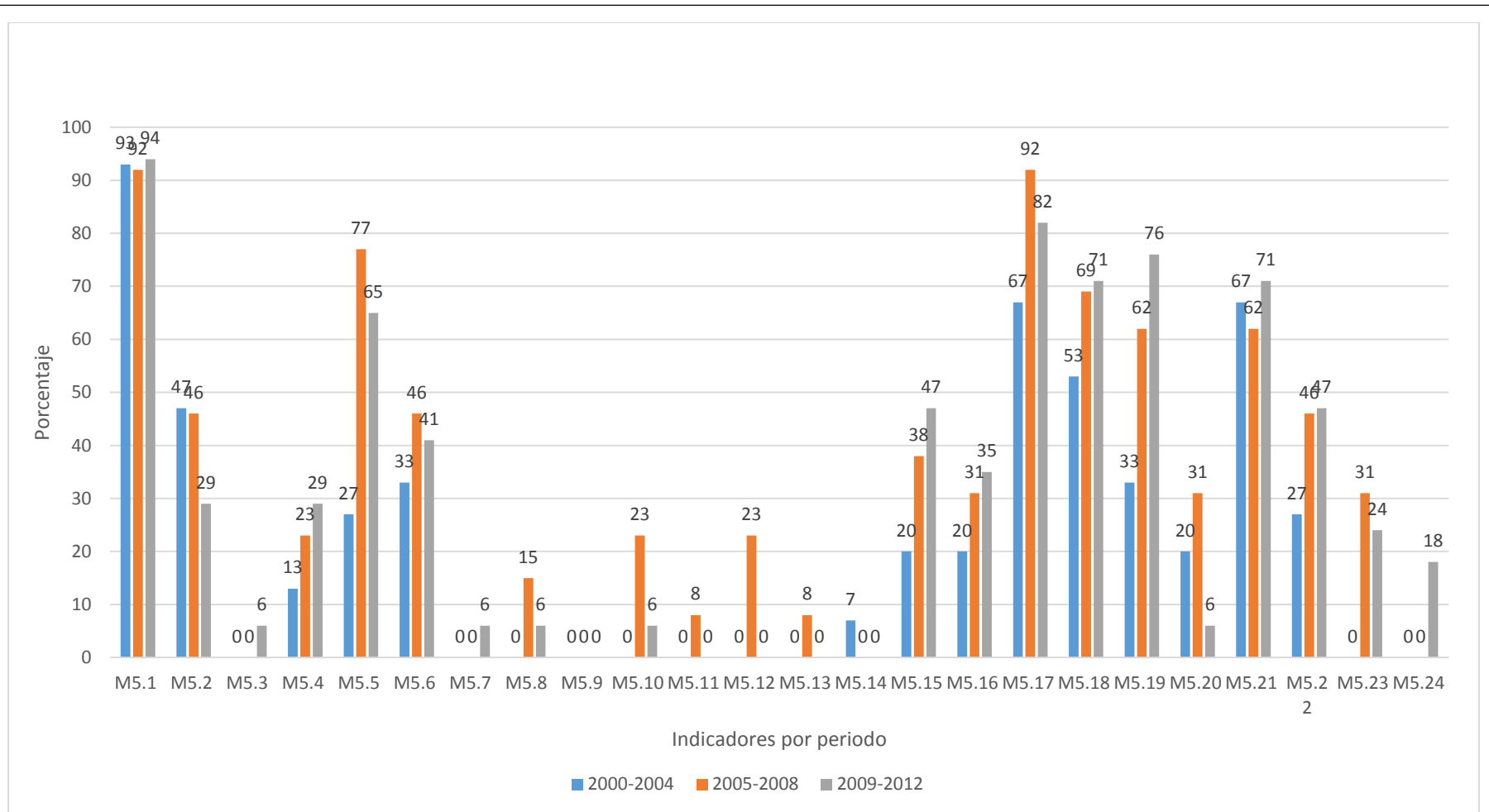
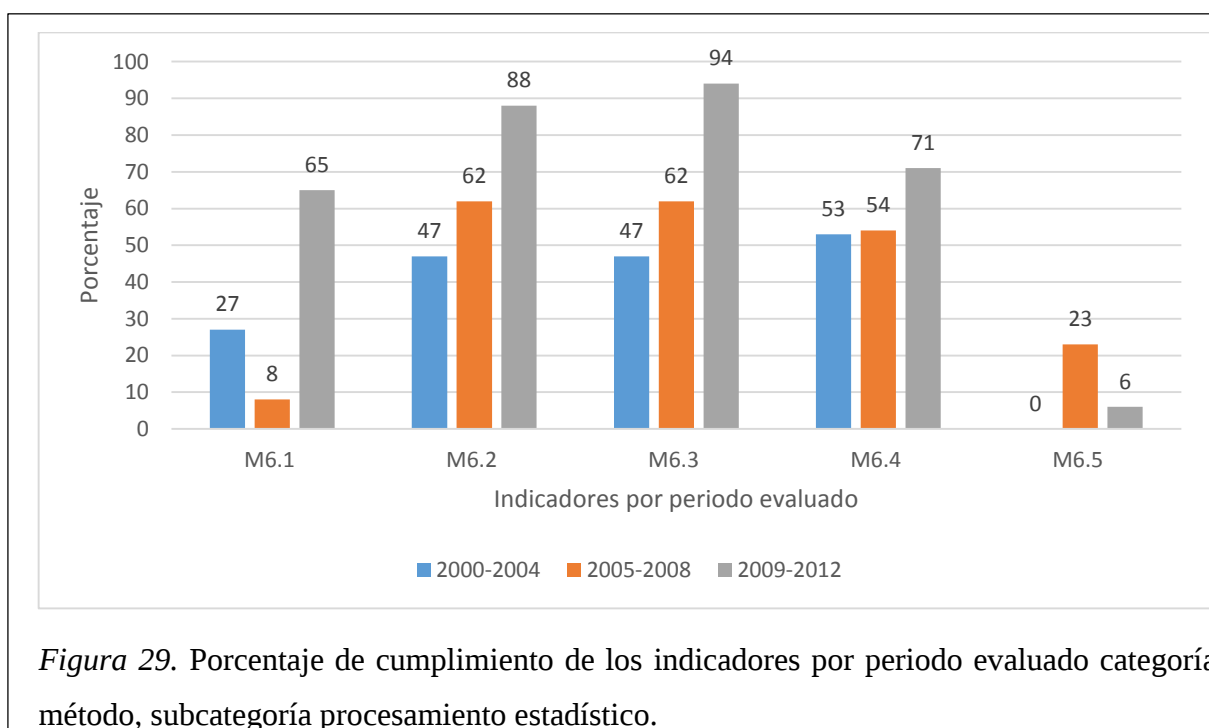


Figura 28. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría método, subcategoría procedimiento

La Figura 29 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría método, subcategoría procesamiento estadístico, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontró un bajo porcentaje de cumplimiento en el indicador relacionado con informar sobre las circunstancias del análisis de datos afectan la validez de los resultados (M6.5). Por su parte en dos indicadores se evidencia un aumento progresivo del porcentaje de cumplimiento del indicador desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales están relacionados con la presentación de los estadísticos utilizados para comparar los grupos (M6.3) y descripción del procesamiento estadístico para que pueda replicarse la investigación (M6.4). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



El porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría método, subcategoría consideraciones éticas, como se puede evidenciar los datos arrojados

por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados (ver Figura 30). Entre los resultados obtenidos es importante destacar que se encontraron bajos porcentajes de cumplimiento en siete de los doce indicadores relacionados con se hace explícito que la investigación se está guiando por un referente ético (M7.1); descripción de la forma en que se obtuvo el consentimiento informado (M7.3); es evidente que las normas éticas se cumplieron (M7.7); se describe cómo se hizo el monitoreo de cumplimiento de las normas éticas (M7.8); se describe la forma en que se garantiza la confidencialidad de los datos en la publicación (M7.9); descripción de la forma en que se manejaron los conflictos de interés (M7.11) y disponibilidad de datos empíricos y/o materiales para reproducir los análisis (M7.12).

Por su parte en cuatro indicadores se evidencia un aumento progresivo del porcentaje de cumplimiento del indicador desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales hacen referencia a utilización del consentimiento informado (M7.2); la aprobación de la investigación por un comité de ética institucional (M7.6); presentación de los conflictos de interés (M7.10) y disponibilidad de datos empíricos y/o materiales para reproducir los análisis (M7.12). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.

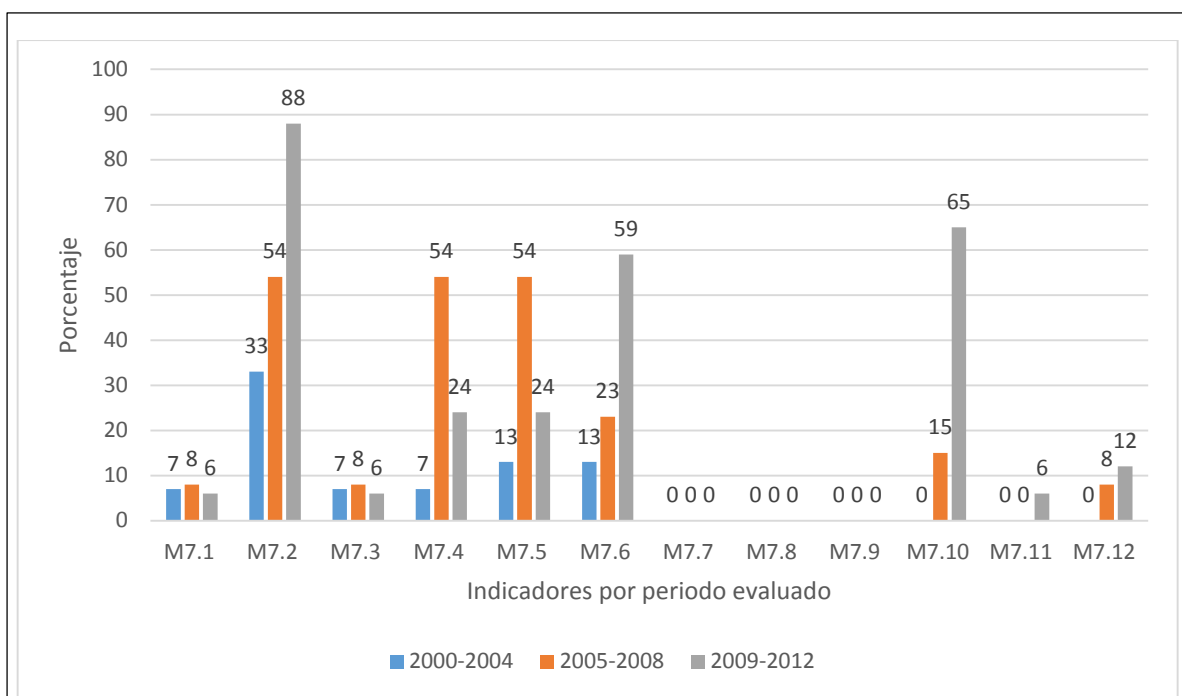
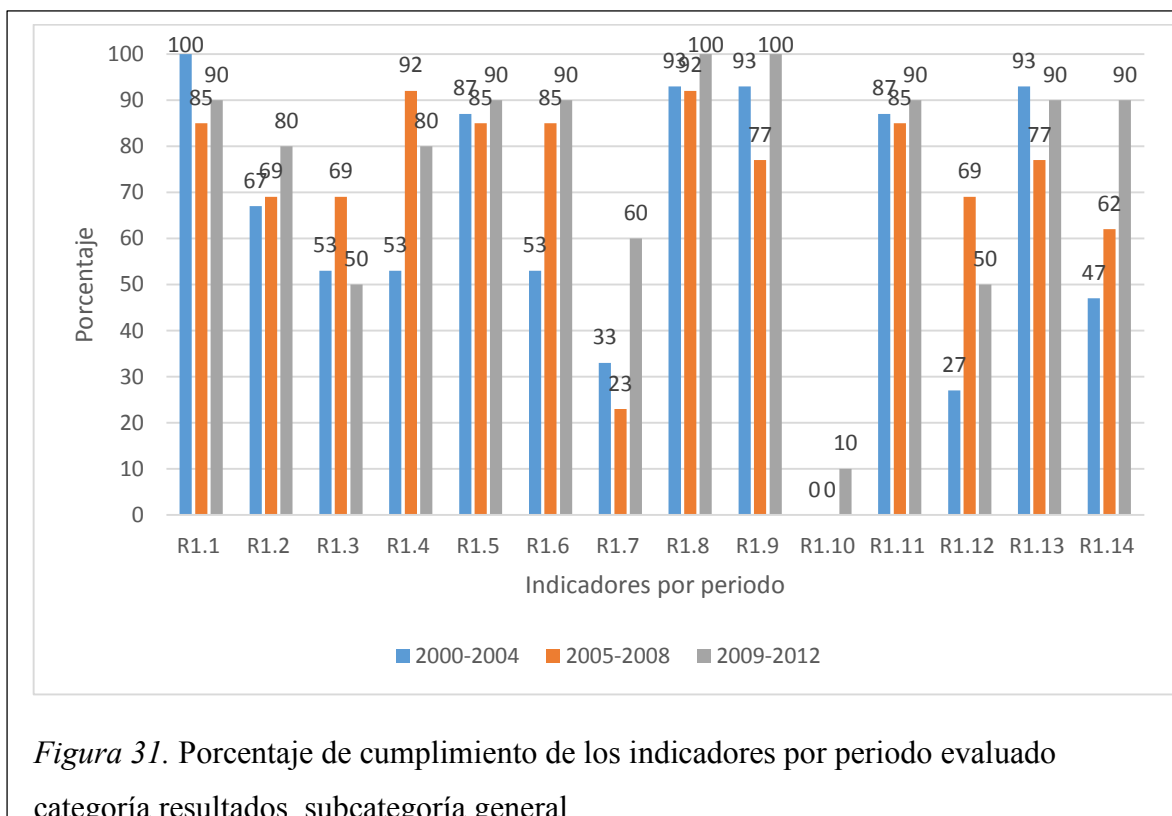


Figura 30. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría método, subcategoría consideraciones éticas.

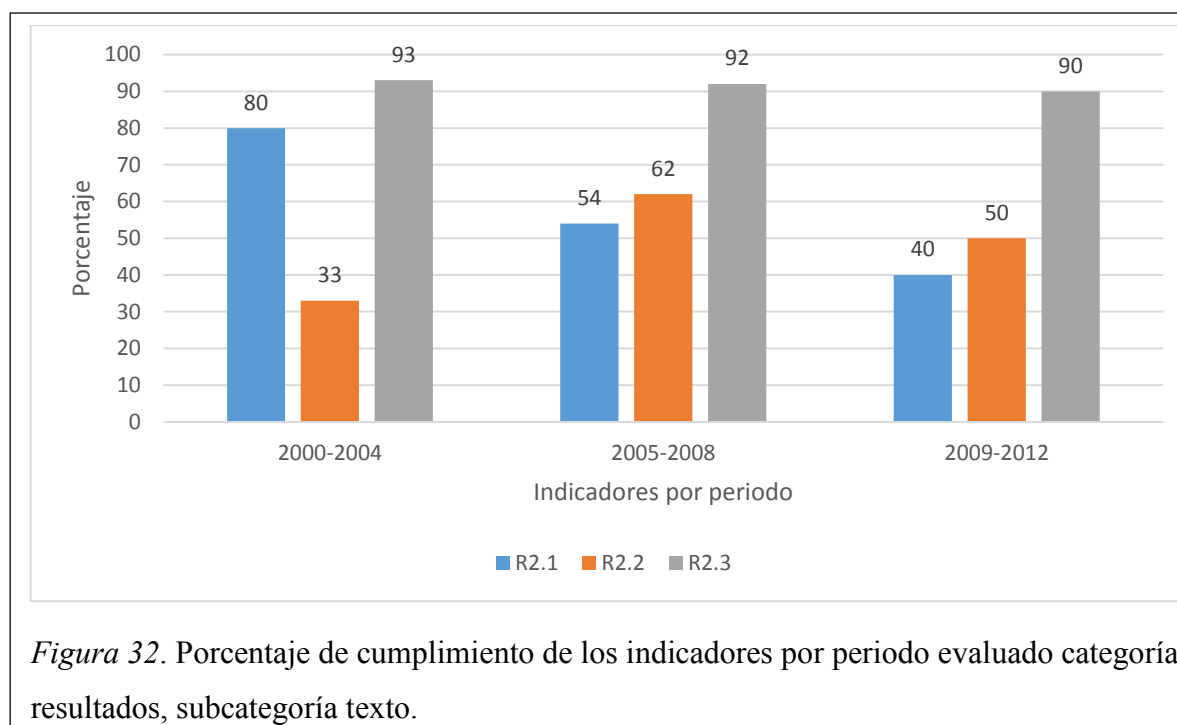
La Figura 31 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría resultados, subcategoría general, como se puede observar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en siete de los 14 indicadores relacionados con presentación organizada de los datos (R1.1); reporte del tamaño de muestra por celda o subgrupo evaluado (R1.4); presentación de los resultados principales y complementarios para cada grupo (R1.5); presentación del valor de error estadísticos (R1.8), reporte de la dirección de "p" del análisis estadístico realizado (R1.9); aplicación correcta de los procedimientos estadísticos para el nivel de medición de los datos (R1.11) y presentación los datos con precisión (R1.13)

En contraste se presentaron bajos porcentajes de cumplimiento en el indicador relacionado con el reporte de los grados de libertad (R1.10). Por su parte en tres indicadores se evidencia un aumento progresivo del porcentaje de cumplimiento del indicador desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales están relacionados con la

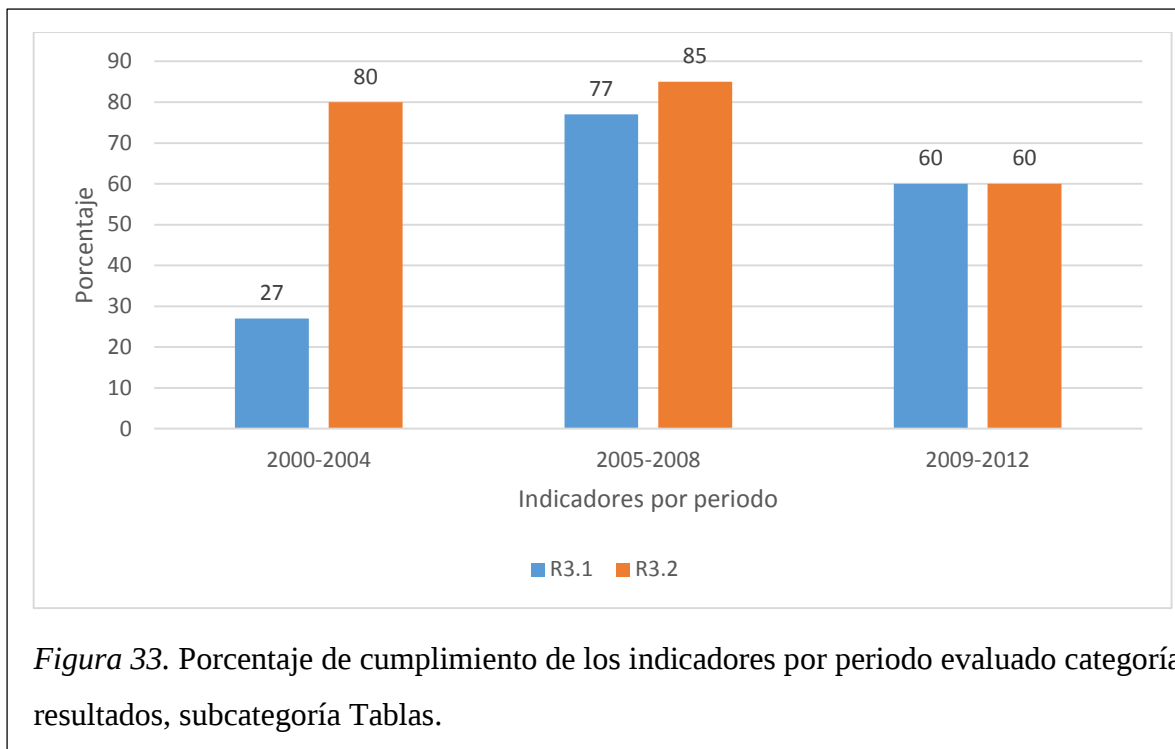
presentación de la frecuencia y/o porcentaje de datos valores perdidos (R1.2); presentación de los tamaños del efecto y los intervalos de confianza (R1.6) y hacer explícito que el grupo control y experimental son similares en las variables relevantes, antes de implementarse la intervención (R1.14). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



La Figura 32 muestra el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría Resultados, subcategoría texto, como se puede observar de los tres indicadores, el que sistemáticamente presentó un mayor porcentaje fue el relacionado con complementariedad entre la información del texto y la de las Tablas (R2.3). También se destaca que el porcentaje de cumplimiento de este indicador fue superior al 90% en los tres periodos. Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



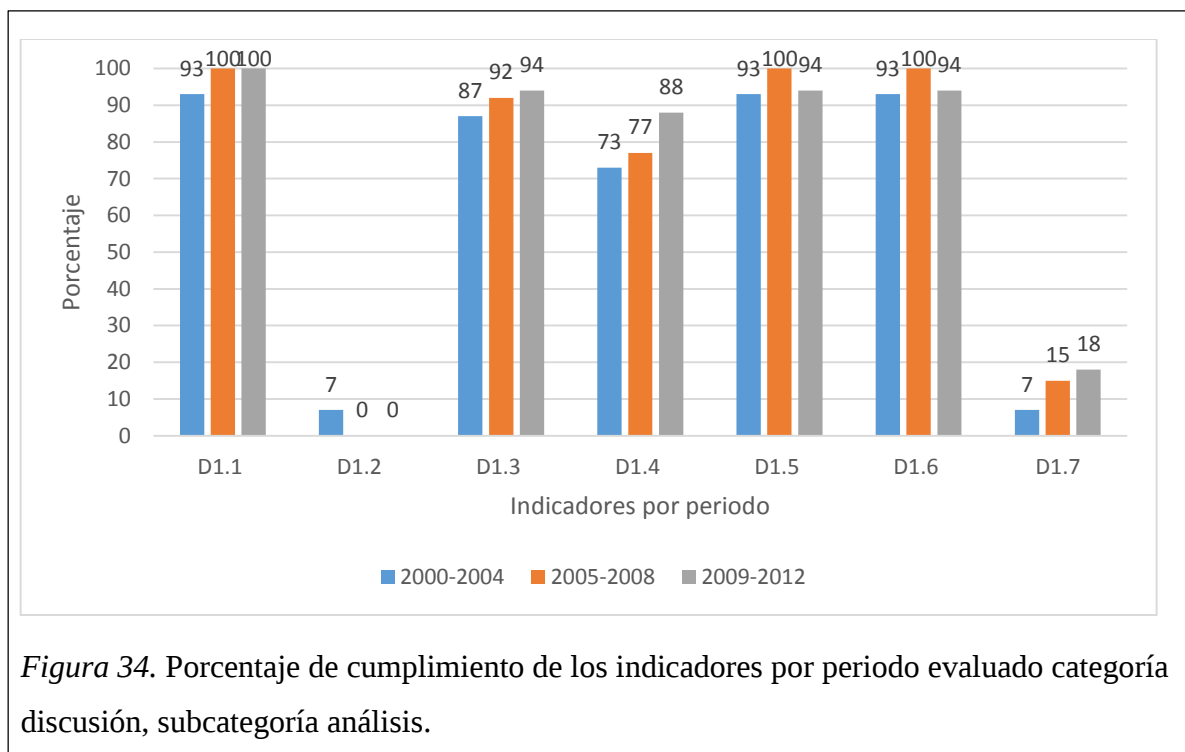
En la categoría resultados se presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado, subcategoría Tablas, como se puede observar el periodo que presentó mayor porcentaje de cumplimiento fue 2005-2008 en los dos indicadores relacionados con la presentación de la Tabla de línea de base de las variables (R3.1) y las Tablas son claras (R3.2) (ver figura 33). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría. En el caso de las gráficas, los porcentajes de cumplimiento en los diferentes periodos se ubicaron preferencialmente en medio bajo y bajo, presentándose los porcentajes mas bajos d cumplimiento en el periodo 2005-2008.



En la categoría discusión, subcategoría análisis, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en cuatro de los siete indicadores referidos a la relación entre el análisis de los datos con los objetivos y/o hipótesis (D1.1); comparación de los resultados de la investigación con otros estudios (D1.3); retomar los resultados, sin repetirlos (D1.5) y relación con los resultados (D1.6) (ver Figura 34).

Por otra parte, se presentaron bajos porcentajes de cumplimiento en los indicadores referidos a la relación entre el análisis de los datos y el problema de investigación (D1.2) y la presentación de los efectos secundarios de las intervenciones (D1.7). Por su parte en tres indicadores se evidencia un aumento progresivo del porcentaje de cumplimiento desde el periodo más antiguo (2000-2004) al más reciente (2004-2009), los cuales están relacionados con la comparación de los resultados con otros estudios (D1.3); presentación de las limitaciones del estudio que afectan la validez interna (D1.4) y presentación de los efectos secundarios (D1.7).

Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



La Figura 35 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría discusión; subcategoría generalización, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en cuatro de los ocho indicadores relacionados con las generalizaciones están basadas en los resultados (D2.1; especificación de las problemáticas a las cuales puede ser generalizado el estudio (D2.4); especificación de las intervenciones a las cuales puede ser generalizado el estudio (D2.6) y se justificación de las generalizaciones propuestas (D2.8).

De otra parte se presentaron bajos porcentajes de cumplimiento en los indicadores relacionados con la especificación de la generalización en cuanto a los métodos de

recolección de datos (D2.5) y características de los terapeutas (D2.7). Por su parte en el indicador que evalúa la generalización a las intervenciones (D2.6), se evidencia un aumento progresivo del porcentaje de cumplimiento del indicador desde el periodo más antiguo (2000-2004) al más reciente (2004-2009). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.

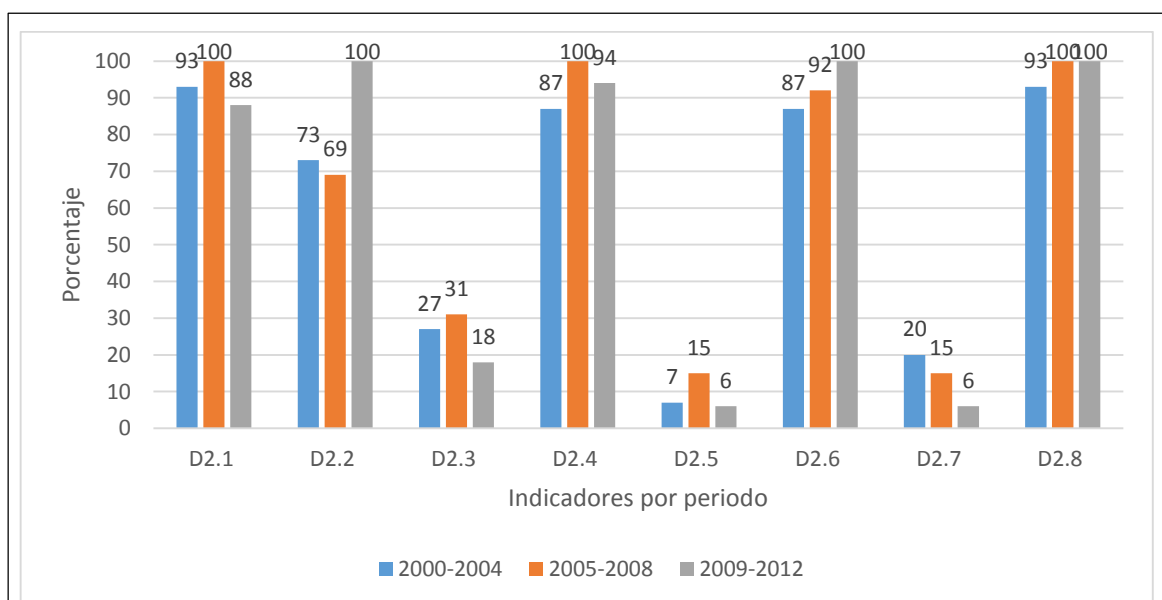
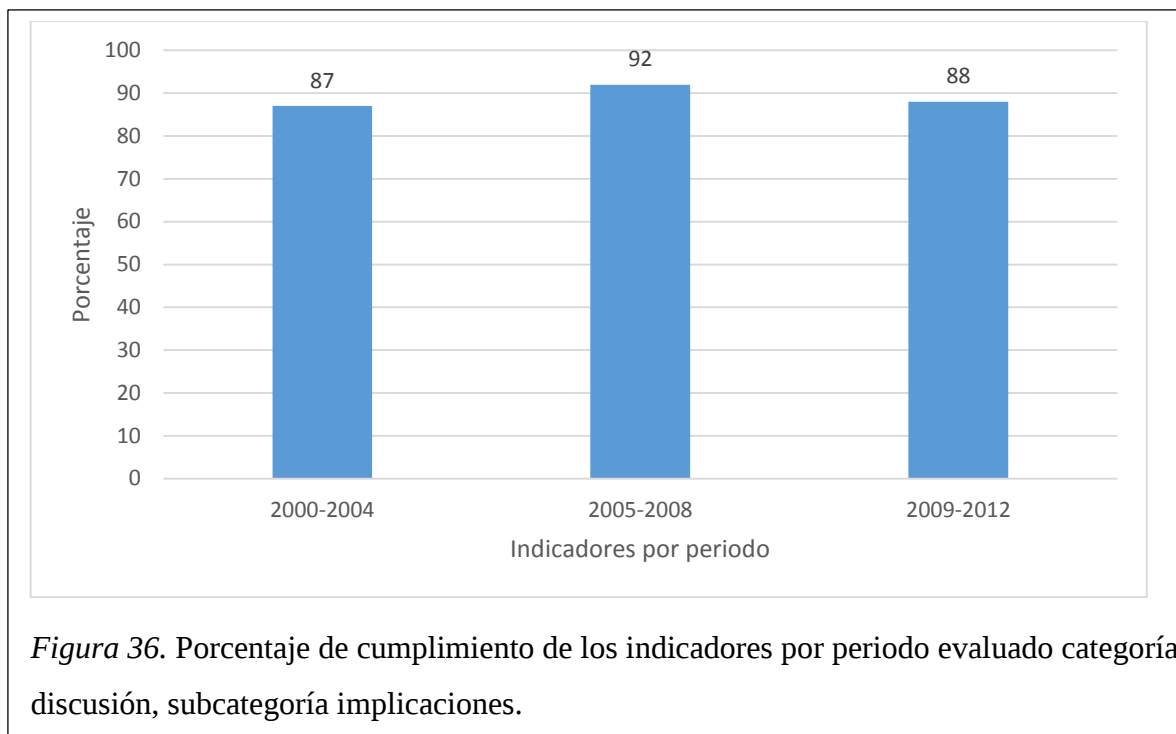
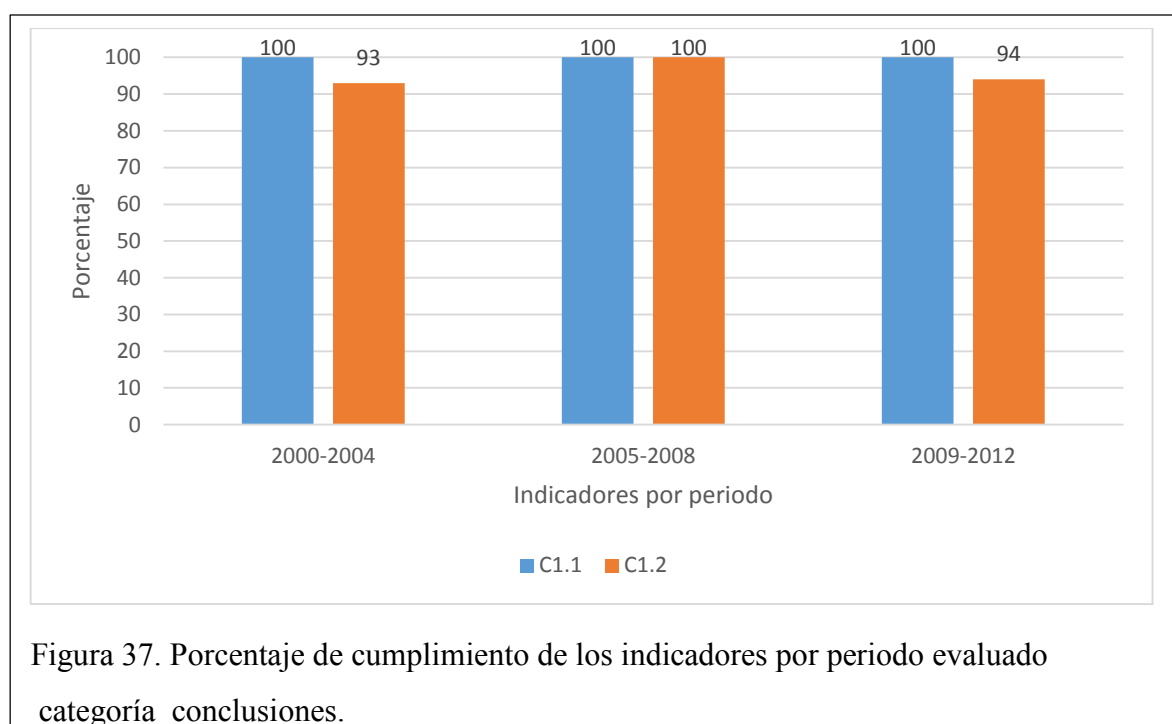


Figura 35. Porcentaje de cumplimiento de los indicadores por periodo evaluado categoría discusión, subcategoría generalización.

La Figura 36 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría discusión, subcategoría implicaciones. Como se puede observar los porcentajes de cumplimiento fueron estables a través de los tres periodos, siendo ligeramente superior 2005-2008. Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en el indicador.

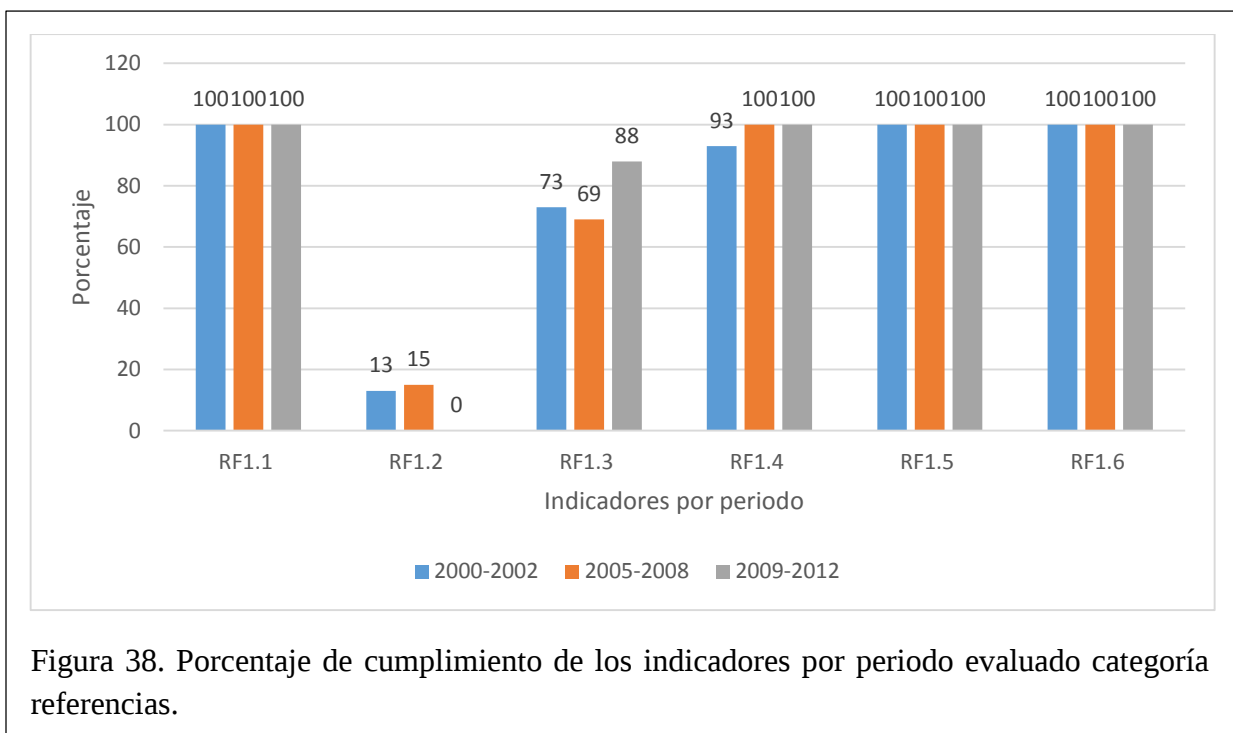


El porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría conclusiones, como se puede observar en la Figura 37, los porcentajes fueron estables a través de los tres periodos, siendo ligeramente superior 2005-2008 en los dos indicadores relacionados con las Conclusiones están fundamentadas en los resultados (C1.1) y las Conclusiones están relacionadas con los objetivos de investigación (C1.2). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en el indicador.



La Figura 38 presenta el porcentaje de cumplimiento de los indicadores por periodo evaluado en la categoría referencias, como se puede evidenciar los datos arrojados por los indicadores en esta subcategoría son variables, pues a nivel general no se evidencia una tendencia clara de predominio en alguno de los periodos evaluados. Entre los resultados obtenidos es importante destacar que se encontraron altos porcentajes de cumplimiento en cuatro de los seis indicadores referidos a la relación entre las referencias y la temática central del artículo (RF1.1); la originalidad de las referencias (RF1.4); adecuarse a las normas utilizadas (RF1.5) y confiabilidad de las referencias (RF1.6).

Por otra parte se evidencio un bajos porcentaje de cumplimiento en el indicador relacionado con referencias son actualizadas (RF1.2). Al realizar el análisis con el Test Exacto de Fisher, no se encontraron asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en ninguno de los indicadores de la subcategoría.



Como se puede notar, a nivel general existen dificultades en el porcentaje de cumplimiento de un gran número de indicadores, y por otra parte se puede afirmar que no se presentó un proceso de evolución a través de los periodos evaluados.

Discusión

En el presente apartado se pretende realizar un análisis de los resultados del estudio, teniendo en cuenta cada uno de los objetivos propuestos. Con relación a los objetivos se discutirán básicamente tres aspectos: los puntos de encuentro y discrepancia con respecto a la teoría en la cual se basó el estudio; también se pretende realizar un análisis de la validez de los datos y establecer las implicaciones de los resultados obtenidos.

El primer objetivo de la investigación estaba dirigido a diseñar y validar por jueces un instrumento para valoración de estándares metodológicos y de publicación; existen diferentes razones que sustentan la creación en este estudio de un instrumento para la valoración de este tópico. La primera consiste en las dificultades observadas tanto en la metodología como en el seguimiento de estándares de publicación en las investigaciones que involucran procesos de intervención, específicamente en el campo de la cefalea. Lo expuesto anteriormente está respaldado en investigaciones de tipo metaanalítico en área médica, en donde se han evidenciado problemas en aspectos como la comparabilidad de los grupos de estudio al inicio de la investigación, el proceso de asignación al azar, los test estadísticos y los criterios de inclusión y exclusión de la muestra, entre otros (Bronfort et al., 2008).

De forma similar, se han encontrado falencias en el campo metodológico y de publicación en investigaciones que involucran intervenciones psicológicas. Entre las dificultades evidenciadas se encuentran utilización de muestras pequeñas, limitaciones en la descripción del tratamiento, de las características de los pacientes; así como también la utilización de sistemas de diagnóstico inadecuados (Nestoriuc et al., 2008). Por su parte, Verhagen et al. (2009) han encontrado dificultades en lo referente a “enmascaramiento” de pacientes y terapeutas, utilización de muestras pequeñas, falta de claridad sobre el diagnóstico de los participantes y deficiencias en la presentación de datos.

La segunda razón para la creación de un instrumento en valoración de estándares metodológicos y de publicación, es que una de las opciones que se han desarrollado y fortalecido en respuesta a estas dificultades, es el establecimiento de estándares metodológicos y de publicación en el contexto investigativo como es el caso del CONSORT en el área de la medicina (Moher et al., 2001, en Boutron et al., 2008), y las normas APA en el campo de Psicología (APA, 2010); como también las normas AERA en psicología

educativa (AERA, 2006), lo cual permite tener una base teórica sólida para el desarrollo de un instrumento en este tópico.

El tercer aspecto que justifica la generación de un instrumento para la valoración de estándares metodológicos y de publicación, es que debido a los aspectos mencionados anteriormente, se han empezado a generar una serie de instrumentos que tiene como fin la medición de estos estándares de calidad; pero que en muchos casos su construcción no se ajusta a los parámetros psicométricos establecidos. Un ejemplo de esta aseveración se puede encontrar en un estudio desarrollado por Armijo et al. (2008), quienes realizaron una revisión sistemática sobre escalas utilizadas para evaluar la calidad de los ensayos controlados en la investigación en salud. De los resultados obtenidos se pudo concluir que en la mayoría de los casos el proceso de construcción de escalas no fue riguroso y que las adaptaciones de las escalas no han experimentado una nueva validación. Coherente con lo expuesto anteriormente, Cascaes et al. (2013) realizaron una revisión sistemática de escalas y listas de verificación de evaluación de calidad metodológica en artículos científicos en salud; las principales Conclusiones del estudio son que en general un número limitado de escalas y listas de verificación presentan de manera explícita las propiedades psicométricas, y se han realizado modificaciones de las escalas originales, sin haber realizado procesos de validación.

En el campo específico de la psicología, Jarde et al. (2012) realizaron una revisión sistemática sobre artículos en el campo de las ciencias de la salud y particularmente en psicología, que incluyeran instrumentos de evaluación de calidad en estudios no experimentales. Entre las Conclusiones más relevantes, se encuentran que la mayoría de los instrumentos analizados no han sido contruidos de una forma rigurosa. Esto ha llevado a que autores como Burgos et al. (2011) afirmen que se ha aumentado la necesidad de poseer instrumentos válidos y confiables para diseñar una investigación, escribir un artículo o evaluar la calidad metodológica de las publicaciones.

En la presente investigación se fijó como meta la construcción de un instrumento que pudiera superar las falencias reportadas en los instrumentos existentes en el área de interés del proyecto. De esta manera, el diseño del instrumento de esta investigación partió de la revisión en la literatura sobre normas y estándares metodológicos y de publicación en el

campo de la medicina y de la psicología. A partir de esta información se elaboró en primera instancia la ficha técnica y estructura de la prueba; para posteriormente realizar el formato inicial del instrumento. Este proceso es coherente con lo propuesto por Carretero-Dios & Pérez (2005), al referirse a los pasos a seguir en la construcción de pruebas psicométricas, los autores afirman que para la creación de un instrumento es necesario diseñar la Tabla de especificaciones del test, que incluye entre otros aspectos el constructo a ser evaluado y sus componentes.

Otro aspecto de gran relevancia en la construcción de una prueba es la validez de contenido, al respecto se ha afirmado que la validez de contenido es una parte integral de la validez de constructo y ésta se convierte en una prueba de que la definición conceptual está representada en los ítems realizados; para este proceso, se utilizan jueces, que tienen como función valorar la pertinencia de los ítems y adicionalmente la claridad de la redacción (Carretero-Dios & Pérez 2005). Por su parte, Escobar y Cuervo (2008), plantean que el juicio de expertos o jueces es una de las formas para medir la validez de contenido, afirman que el juicio de expertos es una opinión de personas con experiencia en el tema y de esta forma están capacitados para hacer valoraciones sobre éste. En el caso de la prueba INVEMCION-01 para la validación de contenido, se realizó una valoración de la pertinencia y redacción de cada uno de los ítems o indicadores basados en la ficha técnica y estructura de la prueba por parte de once jueces expertos en psicometría y/o investigación.

Una vez obtenidas las valoraciones de los jueces se hace necesario el cálculo del índice de validez de contenido, al respecto existen varias propuestas como la de Lawshe (1975) y la de Polit y Hungler (2000), las cuales tienen como elemento común que establecen una proporción entre las valoraciones de los jueces que consideran un ítem como pertinente, con las valoraciones totales realizadas por los jueces. Adicionalmente, se han propuesto criterios para determinar el índice mínimo de validez que debe arrojar un ítem o una prueba para ser considerado(a) válido(a). Entre las más conocidas se encuentran la de Lawshe (1975), quien establece valores mínimos en relación al número de jueces que participaron en la evaluación del instrumento; para el caso de la presente prueba, en donde se utilizaron once jueces, el valor mínimo debería ser 0.59. Otra propuesta es la desarrollada por Tristan (2008), quien propone que la validez mínima requerida para un instrumento o un ítem, independiente del

número de jueces es de 0.5823. En concordancia con lo propuesto por la teoría de construcción de pruebas la validez de contenido del presente instrumento se evaluó por el Método de Lawshe (1975) y por el de Polit y Hungler (2000). Los resultados fueron contrastados con los índices mínimos de validez de contenido propuestos por Lawshe (1975) y Tristan (2008), encontrándose que el instrumento superó los dos criterios propuestos al obtener una validez de contenido de 0.89 por los dos métodos utilizados. Esto permite afirmar que a nivel general los ítems de la prueba están reflejando o representando el concepto de estándares metodológicos y de publicación de un artículo.

Otro de los análisis estadísticos propuestos cuando se realiza un juicio de expertos o jueces es el Coeficiente de concordancia de Kendall; este tipo de análisis permite ver el acuerdo entre los diferentes jueces cuando las valoraciones se encuentran en una escala ordinal (Escobar & Cuervo, 2008). En coherencia con lo propuesto por estos autores en el presente estudio se utilizó esta prueba estadística, para determinar la concordancia de los evaluadores tanto en pertinencia como en la redacción de los indicadores. En lo referente a la pertinencia, la prueba en general arrojó un nivel de acuerdo significativo con un $p < 0.01$, lo cual muestra que los jueces estaban utilizando estándares similares en los indicadores o ítems para realizar las valoraciones. Al hacer el análisis de concordancia con respecto a la redacción de los indicadores, se encontró nuevamente una concordancia significativa en la totalidad de la prueba.

Una razón adicional para la construcción del instrumento de este estudio, es el limitado número de indicadores que evalúan las pruebas de calidad metodológica y de publicación, en el contexto de investigaciones experimentales en dolor. Uno de estos instrumentos es referenciado por Jadad, et al. (1996), el autor expone que la Escala Jadad está conformada por tres ítems que valoran la utilización de la aleatorización, la utilización de doble enmascaramiento y la descripción de los retiros o abandonos de los participantes. Se afirma que debido a que los ítems son generales, el instrumento puede ser utilizado en cualquier área de la medicina; un problema de la escala es que no valora elementos importantes de la calidad de una investigación como los análisis estadísticos y las implicaciones éticas (Jadad, et al., 1996). A su vez se creó la Escala de Yates, que según los autores es específico para ser utilizado en investigaciones psicológicas en dolor; pero podría ser utilizado en otras áreas

(Yates et al., 2005). El instrumento presenta una buena confiabilidad; sin embargo los ítems no evalúan elementos importantes de la calidad de un experimento, como la adherencia del terapeuta al tratamiento, la credibilidad de la terapia y la presentación de efectos secundarios (Yates et al., 2005).

En el caso del diseño y construcción del INVEMCION-01, se pensó inicialmente como un instrumento para evaluar estándares metodológicos y de publicación en artículos que reporten en intervenciones psicológicas; pero al hacer una revisión más detallada podría existir la viabilidad de ser utilizado en otro tipo de disciplinas, como sería el caso de la medicina. Una de las falencias que pretendía mejorar este instrumento con respecto a los anteriores, era la ausencia de indicadores de gran importancia para la evaluación de la calidad de los estudios; por lo cual, se intentó explorar de una manera amplia los estándares metodológicos y de publicación presentes en el contexto de la medicina (Moher et al., 2001, en Boutron et al., 2008) de la Psicología (APA, 2010) y de la psicología educativa (AERA, 2006).

Con respecto al diseño y validación por jueces del instrumento utilizado en el presente estudio, se puede afirmar que ha sido construido teniendo en cuenta la literatura psicométrica para estos procesos (Carretero-Dios y Pérez, 2005; Escobar & Cuervo, 2008; Lawshe, 1975; Tristán, 2008) y se ha pretendido superar dificultades manifestadas en la investigaciones anteriores como procesos no rigurosos de construcción (Armijo et al., 2008), o utilización de un limitado número de indicadores (Jadad, et al., 1996; Yates et al., 2005).

Para el futuro inmediato, se hace necesario hacer los ajustes pertinentes al instrumento, basados en la experiencia de la presente investigación y continuar con los procesos de validación; con el fin de que pueda ser utilizado en diferentes tópicos de investigación experimental tanto en el contexto de la psicología como en el de la medicina. Por ejemplo un aspecto a ser discutido es la utilización de la subcategoría gráficas (resultados), como un apartado opcional para ser utilizado con artículos que incluyan esta forma de presentación en sus resultados.

El segundo objetivo de la investigación pretendía evaluar los niveles en calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012. Hay varias

razones que justifican el desarrollo de este objetivo, en primera instancia existe una gran preocupación por la calidad de las investigaciones; por ejemplo Gutiérrez & Jiménez (2011), aseveran que con alta frecuencia se presenta una baja calidad de las investigaciones, este fenómeno involucra a todos los continentes y genera preocupación en diversas instituciones y organizaciones en el ámbito mundial. El problema de la calidad de las investigaciones no solo se circunscribe a que estén bien o mal realizadas, sino que también involucra efectos negativos en el medio social en donde se desarrolla. Al respecto, Jüni et al. (2001), afirman que las intervenciones originadas de estudios en donde existen problemas metodológicos, son menos benéficas que las surgidas de investigaciones con una adecuada metodología.

Acevedo, López & Cárdenas (2014), plantean que la falta de confiabilidad en las investigaciones de diferentes disciplinas, parte de que los resultados están basados en análisis e ideas incorrectas de interpretación, problemas en el diseño, en el análisis de procedimientos estadísticos y en el bajo nivel de replicación. Los autores agregan que esto ha generado un deterioro de la imagen de las ciencias del comportamiento sustentada en sesgos dentro de las investigaciones, en la obtención de datos en la publicación y encuentro de fraudes.

En segundo lugar, estas dificultades también se presentan en tópicos específicos como es el caso del dolor y de la cefalea tensional y migraña, temáticas que son, el interés principal del presente estudio. Lo anterior ha motivado a que diversos investigadores se hayan dirigido a explorar la calidad metodológica y de investigación en el área del dolor y la cefalea; por medio de metaanálisis y revisiones sistemáticas, en las cuales han encontrado un gran número de falencias en aspectos como muestreo de sujetos (Losada et al., 2009; Nestoriuc et al., 2008; Verhagen, 2009), no se menciona el diseño de investigación (Losada et al., 2009), ausencia de criterios de inclusión y exclusión (Bronfort et al., 2008; Losada et al., 2009), problemas de aleatorización (Borg et al., 2012; Bronfort et al., 2008; García et al., 2007), ausencia de enmascaramiento (Borg et al., 2012; Calvo et al., 2012; García et al., 2007; Verhagen, 2009), ausencia de análisis por intención de tratamiento (Borg et al., 2012; García et al., 2007), ausencia de grupo control y de periodos de seguimiento (Calvo et al., 2012), dificultades con los instrumentos (Calvo et al., 2012; Nestoriuc et al., 2008), problemas comparabilidad de los grupos de estudio al inicio de la investigación y dificultades en los test estadísticos (Bronfort et

al., 2008), limitaciones en la descripción del tratamiento y de las características de los pacientes (Nestoriuc et al., 2008).

El presente proyecto pretendió hacer una revisión sistemática de investigaciones en el contexto de los niveles en calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional, teniendo como base un instrumento diseñado y construido por los autores. Los resultados del análisis metodológico y de estándares se presentaron en porcentaje de cumplimiento de los indicadores en cada una de las categorías y sub categorías estipuladas.

La primera categoría revisada se refiere a portada, a nivel general se presentó un alto porcentaje de cumplimiento en el título, a excepción de la mención de aleatorización, lo cual puede ocasionarse en parte por la inexistencia de aleatorización en la muestra, (Allen, Elliott & Arndorfer, 2002; Ciannella, Capobianco, D'Alessio & Feleppa, 2004; Davis, Reeves, Graff-Radford, Hastie, & Naliboff, 2003); o por un probable desconocimiento del estándar (Andersson, Lundström & Ström, 2003; Calhoun & Ford, 2007; Cottrell, Drew, Gibson, Holroyd & O'Donnell, 2007). Investigaciones anteriores han llamado la atención sobre problemas de aleatorización en los estudios (Borg et al., 2012; Bronfort et al., 2008; García et al., 2007). Por su parte en el resumen los porcentajes de cumplimiento se ubicaron entre medio y bajo. Al respecto Hopewell et al. (2008) destaca la importancia de que los resúmenes sean claros y detallados, debido a que los lectores frecuentemente evalúan un estudio con base en la información del resumen. Este hecho los ha motivado a la creación de los estándares CONSORT para resúmenes; sin embargo los autores destacan que las restricciones de espacio de las revistas limitan la información que se puede registrar en el resumen. Lo planteado por Hopewell et al. (2008) podría ser el responsable de las limitaciones de información que se encontraron en los estudios evaluados.

En la introducción se presentó un alto porcentaje de cumplimiento en contexto investigativo a excepción de presentar la teoría que sustenta las hipótesis de investigación, la cual obtuvo un porcentaje de cumplimiento medio; lo cual se debió principalmente a que en los artículos no muestran de manera explícita las hipótesis lo cual se refleja en investigaciones (Ciannella et al., 2004; Cottrell et al., 2007; Fichtel & Larsson, 2001). En estrecha relación con la dificultad expuesta en contexto investigativo, en el caso de propósitos de investigación la deficiencia se

dio en la presentación explícita del problema de investigación (Abdoli, Rahzani, Safaie & Sattari, 2012; Allen et al., 2002; Andersson et al., 2003). Con respecto al problema de investigación Acevedo, López & Cárdenas (2014) afirman que un aspecto que genera un considerable impacto es la formulación y elección del problema investigativo. En investigaciones anteriores no se habían reportado estas dificultades de una manera específica; a pesar de ser puntos centrales de un proceso investigativo tanto es su proceso de planeación, como en el de ejecución y análisis.

En la categoría de método, se evidenciaron dificultades en las diferentes subcategorías que la componen. La primera fue en el diseño, en donde los estudios evaluados presentaron en un bajo porcentaje, el rotulo o el nombre del diseño de investigación que fue utilizado; esto es coherente con lo propuesto por Losada et al. (2009), quien asevera que en diversos estudios no se menciona el diseño de investigación. Lo anterior se convierte en una gran falencia, en un aspecto tan central de un proceso investigativo, pues como lo afirma Montero, y León, (2007), el diseño es una guía para la realización del proyecto de investigación, se encuentra en estrecha relación con la forma de responder a la pregunta problema y además permite prever las posibles variables extrañas asociadas al tipo de diseño y de esta manera generar los mecanismos para controlarlas. En lo referente a participantes se presentaron porcentajes bajos de cumplimiento en aspectos como presentación de las características de la población, lo cual está acorde con lo propuesto por Nestoriuc et al. (2008), en cuanto las limitaciones presentadas en los estudios, referidas a la descripción de las características de los participantes. Otro problema en los participantes estuvo relacionado con la descripción del equipo de investigación y de los lugares donde se desarrolló el estudio, lo cual ha sido considerado de importancia en estándares como APA (2010) y AERA (2006). La ausencia de esta información dificulta las posibilidades de réplica y comparación de los resultados de las investigaciones. Las réplicas son importantes en la formación de investigadores y generan impacto en las investigaciones originales (Acevedo, López & Cárdenas, 2014).

En el método, en lo referente a también se presentaron dificultades en la presentación las propiedades psicométricas; esto es coherente con la advertencia de algunos investigadores sobre las dificultades en las pruebas utilizadas en procesos investigativos. (Calvo et al., 2012; Nestoriuc et al., 2008). Llama la atención esta deficiencia, teniendo en cuenta que las pruebas

son centrales para los procesos de investigación, al respecto Carretero y Pérez (2007), aseveran que la medición en psicología, adquiere gran relevancia, pues gran parte trabajos de investigación en esta disciplina utilizan pruebas, y de estas dependen las garantías científicas que pueden ofrecer.

En el caso del procedimiento, se presentó un predominio de indicadores con bajo índice de cumplimiento, localizados principalmente en dificultades enmascaramiento tanto en los participantes de la investigación como en los evaluadores y terapeutas. Lo anterior respalda lo propuesto por (Borg et al., 2012; Calvo et al., 2012; García et al, 2007; Verhagen., 2009), en donde se ha señalado el enmascaramiento como una de las falencias más comunes en los estudios experimentales. Este hecho genera implicaciones graves en los resultados de los estudios, pues facilita los sesgos de las personas que están involucradas en el proceso investigativo y por ende distorsiona los resultados del estudio, relacionado con este aspecto, Restrepo y Gómez (2004) plantean que el enmascaramiento es útil para evitar sesgos del sujeto o del observador.

En el método, en lo referente a consideraciones éticas, se presentaron en porcentajes de cumplimiento medio bajo y bajo, estos últimos están relacionados principalmente con monitorear y garantizar el cumplimiento de las normas éticas. Los datos evidencian una gran debilidad de las investigaciones revisadas en este aspecto, lo cual va en detrimento de la calidad de la investigación y de los participantes. Sobre la ética en la investigación González (2002), asevera que en la investigación científica la conducta no ética, debe ser erradicada, pues deteriora la ciencia y a sus productos. La importancia de este aspecto en las investigaciones, es respaldada también por AERA (2006) y Bobenrieth (2002).

En la categoría resultados, los indicadores presentaron porcentajes de cumplimiento medios y altos constituyéndose de esta manera en una gran fortaleza de las investigaciones revisadas, lo cual se refleja en algunos estudios como (Mérelle, Sorbi, Van Doornen, & Passchier (2008) y Mo'tamedi et al., 2012). Al respecto se ha afirmado que la presentación en detalle de los resultados es relevante para justificar las Conclusiones (APA, 2010).

En la categoría discusión, en lo que hace referencia al análisis, se observó que la mayoría de indicadores presentaron porcentajes de cumplimiento altos. Un porcentaje bajo a destacar, es el referido a la relación entre los análisis y el problema de investigación, este resultado es

explicable debido a que un gran número de artículos no hacían explícito el problema de investigación (Abdoli et al., 2012; Allen et al., 2002; Andersson et al., 2003). La subcategoría generalización, presentó un comportamiento similar a la anterior, pues la mayoría de los indicadores presentaron porcentajes de cumplimiento alto, exceptuando los que evalúan la generalización hacia el equipo de investigación y a los lugares donde se desarrolló el estudio; estos resultados son coherentes con lo presentado en la subcategoría participantes en donde se hace una limitada referencia a aspectos relacionados con los investigadores y sitios donde se desarrolló el estudio, lo cual se puede evidenciar en artículos como (Abdoli et al., 2012), Bromberg, 2012; Calhoun, & Ford, 2007). En lo referente a la subcategoría de implicaciones, el único ítem arrojó un valor alto, lo cual evidencia una fortaleza, en este aspecto de los estudios (Abdoli et al., 2012; Allen et al., 2002; Andersson et al., 2003).

En la categoría conclusiones, se presentaron porcentajes de cumplimiento altos en los dos ítems que la componen, lo cual se muestra en artículos como (Abdoli et al., 2012; Bromberg, 2012; Calhoun & Ford, 2007). Lo anterior evidencia una fortaleza en este aspecto, lo cual es central dentro de un estudio, pues es una de las partes en donde se hacen aportes como resultado del proceso investigativo.

La última categoría del instrumento es referencias, la gran mayoría de los indicadores mostraron porcentajes altos de cumplimiento, lo cual permite afirmar que es un área de fortaleza en las investigaciones evaluadas. La única falencia que se evidenció en esta categoría, fue la utilización de referencias actualizadas. Es importante recordar que consideraba actualizada cuando más del 50% de las referencias se ubicaban en los 5 años previos a la publicación del estudio (Bobenrieth, 2002; Ramos et al., 2008). Este se convierte en un punto a discutir en el contexto académico en donde se debe dirimir si el estándar es muy alto o son los autores los que han manejado una baja rigurosidad investigativa en este aspecto. De cualquier manera la relevancia de las referencias en un proceso investigativo es indiscutible, al respecto APA (2010) asevera que las referencias sustentan las afirmaciones basadas en la literatura y por lo tanto estas deben ser precisas, completas y útiles para las personas que implementan y valoran el proceso investigativo.

Un aspecto relevante de los resultados obtenidos en cuanto a porcentaje de cumplimiento de los indicadores en este estudio, es el poder destacar aspectos positivos en los artículos

evaluados, debido a que tradicionalmente los estudios se han centrado en las falencias existentes (Borg et al., 2012; Bronfort et al., 2008; Calvo et al., 2012; García et al., 2007; Losada et al., 2009). En la presente investigación se encontraron potencialidades en diferentes categorías como conclusiones y referencias; y subcategorías como título (portada), contexto investigativo (introducción), general (resultados), análisis e implicaciones (discusión).

Por otra parte se evidenciaron dificultades en varios aspectos que habían sido reportados por la literatura como mencionar el diseño de investigación (Losada et al., 2009), ausencia de enmascaramiento (Borg et al., 2012; Calvo et al., 2012; García et al., 2007; Verhagen, 2009), dificultades con los instrumentos (Calvo et al., 2012; Nestoriuc et al., 2008), limitaciones en la descripción de las características de los pacientes (Nestoriuc et al., 2008).

Otro aspecto de importancia es que debido a que el instrumento cuenta con un gran número de indicadores, que le han permitido evaluar aspectos que no se habían tenido en cuenta en investigaciones anteriores, se pudieron encontrar falencias adicionales como la falta de presentación explícita del problema, descripción del equipo de investigación y de los lugares donde se realizó el estudio; como también la falta de utilización de referencias actualizadas.

Probablemente estas dificultades encontradas tengan relación con la limitada claridad sobre los efectos de las intervenciones psicológicas en el campo del dolor y la cefalea, en donde algunos autores sustentan la efectividad (Moix & Casado, 2011; Vallejo, 2008), y otros sus resultados limitados (Barton, & Blanchard, 2001; Navarro, 2006; Silver, 2007; Verhagen et al., 2009). De forma similar, ocurre con la terapia de biofeedback en la cual algunos autores defienden sus efectos positivos (Kapitz et al., 2010; Vallejo, 2008), y otros no han encontrado resultados satisfactorios (Barton, & Blanchard, 2001; Navarro, 2006; Verhagen et al., 2009).

Es necesario aclarar que los resultados solo pueden ser generalizados a los artículos evaluados y que se debe tener en cuenta que el instrumento utilizado a pesar de que se ha alineado con los procesos psicométricos estipulados para la creación de este tipo de pruebas, aún se encuentra en un proceso preliminar de validación. De cualquier manera los datos arrojados se convierten en objeto de reflexión para los investigadores, editores de revistas y comunidad académica, específicamente en el área de dolor y cefalea, para plantear estrategias

de reestructuración de los estándares que exigen; y de seguimiento y verificación en el cumplimiento de estos. Otra línea de trabajo que puede surgir es la aplicación del instrumento utilizado en esta investigación, en otros tópicos de la psicología diferentes al dolor y la cefalea que permitan determinar el estado en que se encuentran estas áreas en cuanto a los estándares metodológicos y de publicación.

El tercer objetivo fue conocer la evolución de la calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012. La razón para fijarse este objetivo, ha partido de la motivación por parte de algunas instituciones e investigadores de que los estándares metodológicos y de publicación mejoren la calidad de los artículos y de las investigaciones; una muestra de este hecho se puede observar en lo propuesto por Schulz et al (2010), quienes afirman que el establecimiento del CONSORT, tiene entre sus motivaciones contar con información clara y completa de la metodología y de los resultados de los estudios, y optimizar la publicación de las pruebas controladas aleatorias. Lo anterior facilita la evaluación de las investigaciones, el reconocimiento de las fallas existentes, y mejora la calidad en el diseño y ejecución de las investigaciones. Schulz et al (2010) agregan que cada vez más un número considerable de revistas científicas han asumido las recomendaciones del CONSORT (Schulz, et al., 2010).

La tendencia de mejorar la calidad de los estudios y las publicaciones, ha hecho que algunas instituciones, hayan generado y/o actualizado estándares metodológicos y de publicación en el campo de la medicina y de la psicología (AERA, 2006; APA, 2010; Moher et al., 2001, en Boutron et al., 2008); de forma complementaria ha existido la preocupación en proponer lineamientos para los procesos de escritura y/o revisión de artículos médicos y psicológicos (Bobenrieth, 2002; Carretero & Pérez, 2007; Fernández y Buela, 2009; Ramos et al., 2008; Seligman, 1995). En coherencia con lo anterior algunos autores recomiendan la utilización de los estándares como CONSORT, para mejorar la calidad de las publicaciones (García et al., 2007; Solar, 2010; Verhagen, 2009). También se motiva a las revistas a realizar publicaciones completas y claras de la investigaciones (Borg et al., 2012). A pesar de esta tendencia de optimización de la metodología de los estudios y de las publicaciones, se ha afirmado que los problemas o dificultades se siguen presentando (Boutron et al., 2008).

Por lo anterior, se considera necesario valorar si ha existido un proceso de cambio o evolución entre los años 2000 y 2012, en lo referente a la calidad metodológica y el seguimiento de estándares de publicación de los artículos en el contexto de la intervención psicológica, específicamente en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional. En esta parte, los resultados del instrumento se presentaron en porcentaje de cumplimiento de los indicadores en cada una de las categorías y sub categorías estipuladas, teniendo en cuenta tres periodos de tiempo 2000-2004, 2005-2008 y 2009-2012. Adicionalmente, se utilizó el Test exacto de Fisher para determinar si existe asociación entre el porcentaje de cumplimiento de cada indicador y los periodos evaluados.

La primera categoría portada en lo que hace referencia al título, a nivel general los indicadores en los diferentes periodos presentaron porcentajes de cumplimiento medio alto y alto; sin embargo, no presentaron tendencia a aumento de porcentaje de cumplimiento desde el año 2000 al 2012. La única excepción se presentó en el indicador relacionado con nombrar la palabra aleatoria en el título, en donde presentó un bajo porcentaje de cumplimiento en los diferentes periodos evaluados; pero fue el único indicador que mostró una tendencia ascendente de cumplimiento a través de los periodos evaluados. El bajo porcentaje de este indicador, puede estar relacionado con la dificultad referida a la aleatorización en las investigaciones (Borg et al., 2012; Bronfort et al., 2008; García et al., 2007). Por su parte, en el resumen los porcentajes de cumplimiento de los indicadores en los diferentes periodos fueron preferencialmente medio bajo y bajo, y sólo cuatro de los indicadores presentaron una tendencia ascendente de cumplimiento a través de los periodos evaluados, lo anterior refleja una falencia general en este indicador, que como se había dicho anteriormente puede tener relación con las normativas de las revistas, en donde se limita el número máximo de palabras para los resúmenes (Hopewell et al., 2008).

En la categoría introducción, en lo referente a contexto investigativo, se presentó un alto porcentaje de cumplimiento en todos los indicadores, durante los diferentes periodos a excepción del relacionado con la presentación del marco conceptual que sustenta las hipótesis de investigación, el cual presentó porcentajes medios de cumplimiento en los diferentes periodos. La razón principal parte de que los artículos no muestran de manera explícita las hipótesis que sustentan los estudios como es el caso de Ciannella et al., 2004,

Cottrell et al. (2007) y Fichtel & Larsson (2001). Ningún indicador presentó una tendencia ascendente de cumplimiento a través de los periodos evaluados, lo cual se puede deber en parte a que la mayoría de los periodos mostraron porcentajes de cumplimiento superiores al 93% y varios de estos lograron el nivel del 100%, lo cual aumenta la dificultad para superar estos porcentajes de cumplimiento. Por su parte, en propósitos de investigación, se puede destacar que dos indicadores presentaron porcentajes de cumplimiento bajo en todos los periodos, los cuales están relacionados con la presentación explícita del problema de investigación (Abdoli et al., 2012; Allen et al., 2002; Calhoun & Ford, 2007). En esta subcategoría, el único indicador que presentó una tendencia ascendente de cumplimiento a través de los periodos está relacionado con la claridad de los objetivos.

En el método, en lo referente al diseño, se evidenció en todos los periodos porcentajes de cumplimiento medio bajo en el cumplimiento del indicador referido a nombrar el diseño de investigación; además este no evidencio una tendencia ascendente de cumplimiento a través de los periodos, la falencia en la presencia en problemas relacionados con nombrar el diseño de investigación (Losada et al., 2009) y su importancia como guía para la realización del proyecto de investigación ha sido destacada por Montero, y León (2007). En la subcategoría participantes se puede afirmar que un número importante de indicadores presentaron un porcentaje de cumplimiento bajo en los periodos evaluados, los cuales están relacionados con la descripción del equipo de investigación y lugares donde se desarrolló el estudio, además no evidenció tendencia ascendente de cumplimiento a través de los periodos, esta dificultad se puede ilustrar en artículos de años recientes que siguen mostrando este problema (Bromberg et al., 2012; Gerber et al., 2010; Kang, Park, Chung & Yu, 2009). Las dificultades en descripción de participantes habían sido advertidas por Nestoriuc et al. (2008); además se ha insistido en su relevancia en la descripción del proceso investigativo (APA, 2008; AERA, 2006).

En los instrumentos, se evidenció en todos los periodos porcentajes de cumplimiento medio bajo en el cumplimiento del indicador referido a presentar las propiedades psicométricas de las pruebas utilizadas, la problemática referida a la descripción de los instrumentos utilizados en las investigaciones (Calvo et al., 2012; Nestoriuc et al., 2008); por su parte, Carretero y Pérez (2007) aseveran que las pruebas son de gran ayuda para los

procesos investigativos en Psicología. Es importante destacar que dos indicadores pertenecientes a esta subcategoría presentaron una tendencia ascendente de cumplimiento a través de los periodos, a pesar de que no se alcanzaron a lograr asociaciones significativas entre los periodos evaluados y el porcentaje de cumplimiento en estos indicadores.

En lo referente al procedimiento, se presentó un predominio de los indicadores con bajo índice de cumplimiento en los diferentes periodos evaluados, encontrando que gran parte de estos indicadores están relacionados con el proceso de enmascaramiento en participantes, evaluadores y terapeutas. Lo anterior respalda las falencias reportadas en estudios anteriores sobre el enmascaramiento (Borg et al., 2012; Calvo et al., 2012; García et al, 2007; Verhagen, 2009). Un aspecto de gran relevancia en esta subcategoría fue el hecho de que los únicos indicadores del instrumento, que presentaron asociaciones significativas entre los periodos 2000-20004 y 2005-2008 y el porcentaje de cumplimiento se encontraban en esta subcategoría, los cuales estaban relacionados con el control de variables extrañas, y los avances en estos indicadores (Martin et al., 2007; Mérelle et al, 2008; Thorn et al., 2007). De esta manera se puede afirmar con alta seguridad, que estos indicadores en los artículos evaluados, realmente experimentaron un proceso de cambio positivo entre estos periodos.

En lo referente a consideraciones éticas, en casi la totalidad de los indicadores el porcentaje de cumplimiento en los diferentes periodos se ubicó en las categorías medio bajo y bajo, predominando este último; los indicadores con las principales falencias estaban relacionados con el control del cumplimiento de las normas éticas. Lo anterior permite inferir que en los artículos evaluados, aún no se refleja la importancia que se le debe asignar a los aspectos éticos investigación; a pesar de que existen diversas fuentes que destacan su relevancia (AERA, 2006; Bobenrieth, 2002; González, 2002).

En la categoría resultados, a nivel general los indicadores presentaron porcentajes de cumplimiento medios y altos en los periodos evaluados; sin embargo, un número limitado de indicadores presentaron una tendencia ascendente de cumplimiento a través de los periodos, esto se puede ilustrar en artículos de años recientes que siguen mostrando dificultades en esta categoría como (Kang, Park, Chung, & Yu, 2009; Sharma, & Sharma, 2012)

En la discusión, en lo relacionado con el análisis, la mayoría de indicadores presentaron porcentajes de cumplimiento altos en los periodos evaluados; además tres indicadores

presentaron una tendencia de aumento de porcentaje de cumplimiento desde el año 2000 al 2012. Un porcentaje bajo a destacar, es el referido a la relación entre los análisis y el problema de investigación, este resultado es explicable debido a que un gran número de artículos no hacían explícito el problema de investigación (Abdoli et al., 2012; Allen et al., 2002; Andersson et al., 2003). La subcategoría generalización, presentó porcentajes de cumplimiento medio alto y altos en la mayoría de indicadores en los diferentes periodos evaluados. Los porcentajes bajos se presentaron en los indicadores que evalúan la generalización hacia los contextos en donde se desarrolló el estudio y hacia los terapeutas, (Abdoli et al., 2012; Bromberg, 2012; Calhoun & Ford, 2007). En lo referente a la subcategoría de implicaciones, arrojó un valor alto en todos los periodos evaluados, sin embargo no presentó una tendencia ascendente de cumplimiento, lo cual puede deberse en parte a que el periodo 200-2004 mostró un porcentaje de cumplimiento del 87%, lo cual aumenta la dificultad para superar estos porcentajes de cumplimiento a niveles significativos.

En las conclusiones, se presentaron porcentajes de cumplimiento altos durante los periodos evaluados; sin embargo, no presentó una tendencia ascendente de cumplimiento a través de los periodos, lo cual se puede deber en parte a que los porcentajes de cumplimiento son superiores al 93% lo cual hace muy complicado superar este estándar.

La última categoría del instrumento es referencias, la gran mayoría de los indicadores mostraron porcentajes altos de cumplimiento en los tres periodos evaluados convirtiéndose en una categoría de gran fortaleza. La única falencia que se evidenció en esta categoría, fue la del indicador que evalúa la presencia de referencias actualizadas en donde se presentaron porcentajes bajos de cumplimiento en los tres periodos (Devineni, & Blanchard, 2005; Dindo et al., 2012; D'Souza et al., 2008). La importancia de este indicador ha sido propuesta por autores como Bobenrieth (2002) y Ramos et al. (2008).

Al hacer un análisis de los resultados obtenidos con los artículos evaluados, en cuanto a la evolución de la calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012, en lo referente a porcentaje de cumplimiento, se puede afirmar que no existe evidencia para sustentar que ha existido una evolución. Lo que se ha podido observar es que algunos indicadores, han mostrado una tendencia al aumento como nombrar la palabra

aleatorio o su equivalente en el título (P1.1); hacer evidente el marco conceptual que sustenta la investigación (P2.6); presentar el diseño de investigación (P2.7), entre otros; sin embargo, estos cambios no alcanzan a ser estadísticamente significativos usando el Test exacto de Fisher. En casos aislados una posible explicación podría estar relacionada con porcentajes de cumplimiento altos en el primer periodo (2000-2004), lo cual dificulta la superación de estos niveles, como es el caso del indicador relacionado con la claridad de la introducción (I1.4), el cual iniciaba en el primer periodo con un porcentaje de cumplimiento del 93%. De esta que la pretensión de que estándares como el CONSORT optimicen la publicación de las investigaciones, la calidad de su diseño y ejecución como lo referido por Schulz et al (2010), lo cual no se ha reflejado, en los artículos valorados en el presente estudio.

La única excepción a esta regla se ubica en la categoría método, en donde dos indicadores relacionados con control de variables extrañas (M5.15 y M5.16), presentaron asociaciones significativas entre los periodos 2000-2004 y 2005-2008 y el porcentaje de cumplimiento del indicador; en otras palabras se evidenció un proceso de aumento del porcentaje de cumplimiento de los indicadores entre estos periodos; lo cual estuvo respaldado en artículos como los de (Martin et al. (2007), Mérelle et al. (2008) y Thorn et al. (2007).

En la literatura revisada no se encuentran antecedentes de investigaciones que hayan tenido como objetivo la evaluación de la evolución de artículos en cuanto a calidad metodológica y seguimiento de estándares de publicación, en unos periodos determinados y en consecuencia tampoco se ha valorado si los posibles cambios han sido significativos estadísticamente. Una referencia que podría estar tangencialmente relacionada es la afirmación de Boutron et al., (2008) sobre la persistencia de problemas metodológicos y de publicación en los artículos, a pesar de la presencia y divulgación de normas como CONSORT.

Estos resultados motivan a alinearse con autores que recomiendan la utilización de los estándares como CONSORT, para mejorar la calidad de las publicaciones (García et al., 2007; Solar, 2010; Verhagen, 2009) o motivan a las revistas a realizar publicaciones completas y claras de la investigaciones (Borg et al., 2012). Adicionalmente Acevedo, López & Cárdenas (2014) recomiendan reportar los datos positivos y negativos; como también facilitar la realización de metaanálisis por medio del reporte de aspectos como intervalos de confianza,

tamaño de la muestra entre otros. También sugieren la realización de réplicas, el uso correcto de los procedimientos estadísticos, entrenamiento continuo de los investigadores y la publicación y lectura de artículos metodológicos.

En el caso específico de la Psicología se debería fortalecer la divulgación y utilización de estándares como APA (2008) y AERA (2006) tanto en la comunidad académica, como en las editoriales de revistas. De la investigación, también surge la necesidad de continuar realizando este tipo de seguimientos en diferentes áreas de la psicología, que permitan tener referentes del estado en que se encuentran las investigaciones en cuanto a la calidad metodológica y de publicación, permitiendo a la vez valorar los progresos o cambios realizados.

La preocupación por la calidad de las investigaciones no solo tiene repercusiones en el ámbito académico; sino que también tiene impacto social, al respecto Verhagen et al., (2001) afirma que la calidad metodológica de una investigación pretende obtener resultados no sesgados y precisos que permitan la implementación en la práctica clínica; por otra parte, Jüni et al. (2001) aseveran que las intervenciones generadas por estudios en donde hay problemas metodológicos, son menos benéficas que las surgidas de investigaciones con una adecuada metodología; de esta manera se puede afirmar que de la calidad de las investigaciones depende en parte, los beneficios o perjuicios a los usuarios finales, que en este caso serían los consultantes de la práctica psicológica habitual.

A nivel de reflexión general, es necesario advertir que la presente investigación pretende mostrar la realidad de los estudios abordados, y no hacer una generalización a todas las investigaciones de intervención psicológica en cefalea tensional y migraña, que se han publicado en el periodo de tiempo 2000-2012. Lo que se pretende es ofrecer información que permita generar un proceso de autorreflexión en la comunidad académica, que permita sensibilizar sobre la necesidad de hacer modificaciones en la manera de diseñar, implementar y publicar una investigación; lo cual en el mediano y largo plazo puede impactar positivamente en intervenciones de mayor calidad, que permitan en el caso específico de este

estudio disminuir y/o contrarrestar los efectos negativos de la cefalea tensional y migraña, que han sido expuestos anteriormente en este trabajo.

También la investigación presenta a consideración un instrumento que sirva de referente en procesos de autoevaluación y heteroevaluación del seguimiento de estándares metodológicos y de publicación, el cual por encontrarse en una etapa preliminar es susceptible de modificaciones basadas en nuevos procesos de validación en otros contextos y con otros tópicos de interés. El instrumento también se convierte en una herramienta útil para realización de investigaciones que hagan seguimiento a la evolución de las investigaciones en el seguimiento de estándares y de esta manera estar en posibilidad de recibir retroalimentación permanente sobre el estado de desarrollo de las publicaciones en estos tópicos.

7. Referencias

- Abdoli, S., Rahzani, K., Safaie, M., & Sattari, A. (2012). A randomized control trial: the effect of guided imagery with tape and perceived happy memory on chronic tension type headache. *Scandinavian Journal Of Caring Sciences*, 26 (2), 254-261. doi:10.1111/j.1471-6712.2011.00926.x
- Acevedo, C., López, W., & Cárdenas, F. (2014). Recomendaciones en el diseño, la ejecución y la publicación de investigaciones en Psicología y ciencias del comportamiento. *Revista Costarricense de Psicología*, 33(2), 155-177.
- Allen K., Elliott A., & Arndorfer, R. (2002). Behavioral pain management for pediatric headache in primary care. *Children's Health Care*, 31 (3), 175-189.
- American Educational Research Association (2006). Standards for reporting on empirical social science Research In Aera Publications. *Educational Researcher*, 35 (6), 33–40.
- American Psychological Association (2010). Manual de Publicaciones de la American Psychological Association. México: Manual Moderno.
- American Society of Anesthesiologists Task Force on Chronic Pain Management, & The American Society of Regional Anesthesia and Pain Medicine (2010). Practice guidelines for chronic pain management. *Anesthesiology*, 112 (4), 810-833.
- Andersson, G., Lundström, P., & Ström, L. (2003). Internet-based treatment of headache: does telephone contact add anything? *Headache*, 43 (4), 353-361. doi:10.1046/j.1526-4610.2003.03070.x
- APA Publications and Communications Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in psychology why do we need them? What Might They Be? *American Psychologist*, 63 (9), 839–85.
- Armijo, S., Macedo, L., Gadotti I., Fuentes, J., Stanton, T., & Magee, D. (2008). Scales to assess the quality of randomized controlled trials: a systematic review. *Journal of Physical Therapy*, 88, 156–175.

- Asociación Colombiana para el estudio del Dolor (2008). *Cuarta Encuesta Nacional del Dolor*. Recuperado de, http://www.dolor.org.co/index.php?option=com_content&view=article&id=69&Itemid=84
- Barton, K., & Blanchard, E. (2001). The failure of Intensive self-regulatory treatment with chronic daily Headache: a prospective study. *Applied Psychophysiology and Biofeedback*, 26 (4), 311-318.
- Belvís R, Mas N., & Roig C. (2015). Novedades en la reciente Clasificación Internacional de las Cefaleas: clasificación ICHD-III beta. *Revista de Neurología*, 60, 81-89.
- Bendtsen, L., Evers, S., Linde, M., Mitsikostas, D., Sandrini, G., & Schoenen, J. (2010). EFNS guideline on the treatment of tension-type headache –Report of an EFNS task force. *European Journal of Neurology*, 17, 1318–1325.
- Benito, G., Nadador, V., Fernández J., Hernández, J., Ruiz, M., & Riquelme, I. (2006). Intervenciones del psicólogo en las Clínicas del Dolor: Una propuesta desde la experiencia de la Unidad del Dolor del Hospital Universitario Príncipe de Asturias en Alcalá de Henares, Madrid. *Revista de la Sociedad Española del Dolor*, 13 (4), 254 – 262.
- Bobenrieth, M. (2002). Normas para revisión de artículos originales en Ciencias de la Salud. *Revista Internacional de Psicología Clínica y de la Salud*, 2 (3), 509-523.
- Borg, V., Zhang, S., Ye, C., Paul J., Arya, A., Hurlburt, L., Murthy, Y., & Thabane, L. (2012). The quality of reporting of RCTs used within a postoperative pain management meta-analysis, using the CONSORT statement. *BioMed Central Anesthesiology*, 12 (13), 1-11.
- Boutron, I., Moher, D., Altman, D., Schulz, K., & Ravaud, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine*, 148, 295-309.
- Bringas, L. (2010). La migraña. *Acta Médica Peruana*, 27 (2), 129-136.
- Bromberg, J., Wood, M. E., Black, R. A., Surette, D. A., Zacharoff, K. L., & Chiauzzi, E. J. (2012). A randomized trial of a web-based intervention to improve migraine self-management and coping. *Headache*, 52(2), 244-261. doi:10.1111/j.1526-4610.2011.02031.x
- Bronfort, G., Nilsson, N., Haas, M., Evans R., Goldsmith CH., Assendelft W., & Bouter L. (2008). Tratamientos físicos no invasivos para la cefalea crónica/recurrente. *Biblioteca*

- Cochrane Plus*, 4. Recuperado de <http://www.update-software.com/BCP/BCPGetDocument.asp?DocumentID=CD001878>
- Bronfort, G., Nilsson, N., Haas, M., Evans, R., Goldsmith, C. H., & Assendeft, W. J. (2005). Tratamientos físicos no invasivos para la cefalea crónica/recurrente (Revisión Cochrane traducida). *La Biblioteca Cochrane Plus*, (2).
- Buela-Casal, G. (2003). Evaluación de la calidad de los artículos y de las revistas científicas: Propuesta del factor de impacto ponderado y de un índice de calidad. *Psicothema*, 15 (1), 23-35.
- Burgos M., Manterola C., & Sanhueza A. (2011). Diseño de una escala para evaluar calidad metodológica de estudios de pruebas diagnósticas. Estudio piloto. *Revista Chilena de Cirugía*, 63 (5), 493-497.
- Calhoun, A., & Ford, S. (2007). Behavioral sleep modification may revert transformed migraine to episodic migraine. *Headache* 47 (8), 1178-1183.
- Calvert M., Blazeby J, Altman D, et al. (2013). Reporting of Patient-Reported Outcomes in Randomized Trials: The CONSORT PRO Extension. *The journal of the American Medical Association*. 309(8), 814-822. doi:10.1001/jama.2013.879.
- Calvo, I., Gómez, A., & Sánchez, J. (2012). Eficacia de los tratamientos de fisioterapia para el dolor lumbar en niños y adolescentes. Revisión sistemática. *Revista de la Sociedad Española de Dolor*, 19 (3), 138-146.
- Carretero, H., & Pérez, C. (2007). Standards for the development and review of instrumental studies: Considerations about test selection in psychological research. *International Journal of Clinical and Health Psychology*, 7 (3), 863-882.
- Cascaes, F., Valdivia, B., Da Rosa, R., Barbosa, P., & Da Silva, R. (2013). Escalas y listas de evaluación de la calidad de estudios científicos. *Revista Cubana de Información en Ciencias de la Salud*, 24 (3), 295-312.
- Chan A, Tetzlaff J., Altman D., Laupacis A., Gøtzsche P., Krleža-Jerić K., et al. (2013). SPIRIT statement: defining standard protocol items for clinical trials. *Annals of Internal Medicine*. 158, 200-207. doi:10.7326/0003-4819-158-3-201302050-00583

- Ciannella, L., Capobianco, N. N., D'Alessio, A., & Feleppa, M. (2004). An integrated approach to cephalalgic patients. Preliminary results on 64 adult patients with migraine without aura. *Journal of Headache & Pain*, 5, 103-108. doi:10.1007/s10194-004-0121-6
- Congreso de la República de Colombia (2006). *Ley 1090 de 2006 (miércoles 6 de septiembre) por la cual se reglamenta el ejercicio de la profesión de Psicología, se dicta el Código Deontológico y Bioético y otras disposiciones*, Recuperado de <http://www.colpsic.info/resources/Ley1090-06.pdf>
- CONSORT (2014). The CONSORT Statement. Recuperado de <http://www.consort-statement.org/>
- Correché, M., & Labiano, L. (2003). Aplicación de técnicas psicoterapéuticas a un grupo de estudiantes con síntomas de estrés. *Fundamentos en Humanidades Universidad Nacional de San Luis*, 4 (1), 129-147.
- Cottrell, C., Drew, J., Gibson, J., Holroyd, K., & O'Donnell, F. (2007). Feasibility assessment of telephone-administered behavioral treatment for adolescent migraine. *Headache*, 47 (9), 1293-1302. doi:10.1111/j.1526-4610.2007.00804.x
- Covarrubias, A., & Guevara, U. (2006). ¿Que son las clínicas del dolor? *Revista Digital Universitaria*, 7(4), 1-8 recuperado de http://www.revista.unam.mx/vol.7/num4/art33/abr_art33.pdf
- Davis, P., Reeves II, J., Graff-Radford, S., Hastie, B., & Naliboff, B. (2003). Multidimensional subgroups in migraine: differential treatment outcome to a pain medicine program. *Pain Medicine*, 4 (3), 215. doi:10.1046/j.1526-4637.2003.03027.x
- Devineni, T., & Blanchard, E. (2005). A randomized controlled trial of an internet-based treatment for chronic headache. *Behavior Research and Therapy*, 43 (3), 277-292.
- Diener, H., Solbach, K., Holle, D., & Gaul, C. (2015). Integrated care for chronic migraine patients: epidemiology, burden, diagnosis and treatment options. *Clinical Medicine*, 15(4), 344-350.

- Dindo, L., Recober, A., Marchman, J. N., Turvey, C., & O'Hara, M. W. (2012). One-day behavioral treatment for patients with comorbid depression and migraine: A pilot study. *Behaviour Research & Therapy*, 50 (9), 537-543. doi:10.1016/j.brat.2012.05.007;
- D'Souza, P., Lumley, M., Kraft, C., & Dooley, J. (2008). Relaxation training and written emotional disclosure for tension or migraine headaches: a randomized, controlled trial. *Annals of Behavioral Medicine*, 36(1), 21-32. doi:10.1007/s12160-008-9046-7
- Escobar, J., & Cuervo, A. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en Medición*, 6, 27–36.
- Fernández, L., & Buela, G. (2009). Standards for the preparation and writing of Psychology review articles. *International Journal of Clinical and Health Psychology*, 9 (2), 329-344.
- Ferracini, G. G. (2014). Quality of Life and Health-Related Disability in Children With Migraine. *Headache: The Journal Of Head & Face Pain*, 54(2), 325-334.
- Fichtel, A., & Larsson, B. (2001). Does relaxation treatment have differential effects on migraine and tension-type headache in adolescents? *Headache*, 41 (3), 290-296.
- Fontanillas, N., Pascual, J., & Colás, R. (2009). Cefalea crónica diaria con Abuso de Analgésicos. *Revista Clínica de Medicina de Familia*, 2 (8), 400-407.
- García, G., & Sanchez, I. (2011). Perspectiva metodológica análisis crítico de la literatura científica: como analizar un artículo publicado. *Revista Medicina, Salud y Sociedad*, 1(2) 1-16.
- García, J., Parera, A., Olléc, G., & Bonfill, X. (2007). Características y calidad metodológica de los ensayos clínicos publicados en la Revista Española de Anestesiología y Reanimación. *Revista Española de Anestesiología y Reanimación*, 54, 333-339.
- Gerber, W., Petermann, F., Müller, G., Dollwet, M., Darabaneanu, S., Niederberger, U., & Andrasik, F. (2010). MIPAS-Family—evaluation of a new multi-modal behavioral training program for pediatric headaches: clinical effects and the impact on quality of life. *Journal of Headache & Pain*, 11 (3), 215-225. doi:10.1007/s10194-010-0192-5

- González, M. (2002), Aspectos éticos de la investigación cualitativa. *Revista Iberoamericana de Educación*, 29, 85-103
- Guallar, E., Damián J., & Martín, J. (1997). Metaanálisis y revisiones sistemáticas en cardiología. *Revista Española de Cardiología*, 50, 345-354.
- Gutiérrez, A., & Jiménez, R. (2011). ¿Por qué y cómo evaluar la calidad de la Investigación científica en medicina? *Revista Cubana de Cardiología y Cirugía Cardiovascular*, 17 (4), 365-370.
- Headache Classification Committee of the International Headache Society (IHS) 2013. The International Classification of Headache Disorders, 3rd edition (beta version). *Cephalalgia* 33, 629-808.
- Hopewell S., Clarke M., Moher, D., Wager, E., Middleton P., Altman D., Schultz K., & The CONSORT Group (2008). CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet*, 371 (9609), 281-283.
- International Association for the Study of Pain (1994). *Dolor*. Recuperado de <http://www.iasp-pain.org/Content/NavigationMenu/GeneralResourceLinks/PainDefinitions/default.htm>
- International Association for the Study of Pain (2011a). *2011-2012 global year against headache*. Recuperado de <http://www.iasppain.org/Content/NavigationMenu/GlobalYearAgainstPain/GlobalYearAgainstHeadache/default.htm>
- International Association for the Study of Pain (2011b). *Epidemiology of headache*. Recuperado de http://www.iasp-pain.org/AM/Template.cfm?Section=Fact_Sheets4&Template=/CM/ContentDisplay.cfm&ContentID=14452
- International Committee of Medical Journal Editors (2013). Recommendations for the Conduct, reporting, Editing, and Publication of Scholarly Work in Medical Journals. Recuperado de <http://www.icmje.org/icmje-recommendations.pdf>

- International Committee of Medical Journal Editors (2014). About ICMJE. Recuperado de <http://www.icmje.org/about-icmje/>
- Jadad, A., Moore, R., Carroll, D., Jenkinson, C., Reynolds, J., Gavaghan, D., & McQuay, H. (1996). Assessing the Quality of Reports of Randomized Clinical Trials: Is Blinding Necessary? *Controlled Clinical Trials*, 17, 1-12.
- Jarde, A., Losilla J., & Vives, J. (2012). Methodological quality assessment tools of non-experimental studies: a systematic review. *Anales de Psicología*, 28 (2), 617-628.
- Jüni, P., Altman, D., & Egger, M. (2001). Systematic reviews in health care Assessing the quality of controlled clinical trials. *British Medical Journal*, 323, 42–6.
- Kang, E.-H., Park, J.-E., Chung, C.-S., & Yu, B.-H. (2009). Effect of Biofeedback-assisted Autogenic Training in Korean Female Migraine Patients. *Journal of Korean Medical Science*, 24(5), 936–940. [http on Headache Activity and Mood States://doi.org/10.3346/jkms.2009.24.5.936](http://onlinelibrary.wiley.com/doi/10.3346/jkms.2009.24.5.936)
- Kapitz, K., Passie, T., Bernateck, M., & Karst, M. (2010). First non-contingent respiratory biofeedback placebo versus contingent biofeedback in patients with chronic low back pain: a randomized, controlled, double-blind trial. *Applied Psychophysiology and Biofeedback*, 35, 207–217.
- Kaushik, R., Kaushik, R., Mahajan, S., & Rajesh, V. (2005). Biofeedback assisted diaphragmatic breathing and systematic relaxation versus propranolol in long term prophylaxis of migraine. *Complementary Therapies in Medicine*, 13 (3), 165–174. [doi: 10.1016/j.ctim.2005.04.004](https://doi.org/10.1016/j.ctim.2005.04.004)
- Lawshe, C. (1975). A quantitative approach to content validity. *Personnel Psychology*. 28 (4), 563-575.
- Lazcano, E., Salazar, E., Gutiérrez, P., Ángeles, A., Hernández, A., & Viramontes, J. (2004). Ensayos clínicos aleatorizados: variantes, métodos de aleatorización, análisis, consideraciones éticas y regulación. *Salud Pública de México*, 46 (6), 559-584.

- LeForta, S., Gray, K., Rowatc, K., & Jeansd, M. (1998). Randomized controlled trial of a community-based psychoeducation program for the self-management of chronic pain. *Pain*, 74,297–306.
- Losada, H., Manterola, C., Pineda, V., Vial, M., Sanhueza, A., & Grupo MINCIR. (2009). Diseño de una escala para la evaluación de calidad metodológica de estudios de pronóstico. *Revista Chilena de Cirugía*, 61 (1), 59-72.
- Madrid, S., & Vidal, E. (2008). Cefalea tipo tensional. *Revista de Evidencia e Investigación Clínica*, 1 (1), 15-24.
- Marín, F., Sánchez, J., & López, J. (2009). El metaanálisis en el ámbito de las Ciencias de la Salud: una metodología imprescindible para la eficiente acumulación del conocimiento. *Fisioterapia*, 31 (3) ,107–114.
- Martin, P., Forsyth M., & Reece, J. (2007). Cognitive-behavioral therapy versus temporal pulse amplitude biofeedback training for recurrent headache. *Behavior Therapy*, 38 (4), 350-363.
- Mas Sala, N., & Arnall, C. I. (2015). Conceptos básicos clínico terapéuticos. *Dolor*, 30 (1), 7-17.
- Mérelle, S., Sorbi, M., Van Doornen, L., & Passchier, J. (2008). Migraine patients as trainers of their fellow patients in non-pharmacological preventive attack management: short-term effects of a randomized controlled trial. *Cephalalgia*, 28 (2), 127-138. doi:10.1111/j.1468-2982.2007.01472.x
- Moix, J., & Casado M. (2011).Terapias psicológicas para el tratamiento del dolor crónico. *Clínica y Salud*, 22 (1), 41-50.
- Montero, I. & León, O. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7 (3), 847-86.
- Mo'tamedi, H., Rezaiemaram, P., & Tavallaie, A. (2012). The effectiveness of a group-based acceptance and commitment additive therapy on rehabilitation of female outpatients with chronic headache: preliminary findings reducing 3 dimensions of headache impact. *Headache*, 52(7), 1106-1119. doi:10.1111/j.1526-4610.2012.02192.x

- Mustaca, A. (2011). Evaluación objetiva de los tratamientos psicológicos: modelos basados en la ciencia. *Revista Colombiana de Psicología*, 2 (1), 99-106.
- Navarro, M. (2006). Tratamiento cognitivo conductual de la migraña en el adulto. *Actualidades en Psicología*, 20(107), 1-22.
- Nestoriuc, Y., Martin, A., Rief, W., & Andrasik, F. (2008). Biofeedback treatment for headache disorders: a comprehensive efficacy review. *Applied Psychophysiology and Biofeedback*, 33, 125–140.
- Olivares, M., & Cruzado, J. (2008). Evaluación psicológica del dolor. *Clínica y Salud*, 19 (3), 321-341.
- Oxman, A. & Guyatt, G. (1991). Validation Of an index of the Quality of review articles. *Journal of Clinical Epidemiology*, 44 (11), 1271-1278.
- Panerai, A. (2013). Is migraine a disorder of the central nervous system?. *Neurological Sciences*, 3433-3435. doi:10.1007/s10072-013-1363-3
- Peralta, L. (2005). Análisis y descripción de datos en estudios comparativos. *Dermatología Revista Mexicana*, 49 (5), 219-221.
- Polit, D., & Hungler B. (2000). *Investigación científica en Ciencias de la Salud: principios y métodos*. México DF, México: McGraw-Hill Interamericana.
- Ramos, M., Moreno, M., Valdés, B., & Catena, A. (2008) .Criteria of the peer review process for publication of experimental and quasiexperimental research in Psychology: A guide for creating research papers. *International Journal of Clinical and Health Psychology*, 8 (3), 751-764.
- Rathier, L., & Roth, J. (2015). A Biobehavioral Approach to Headache Management. *Rhode Island Medical Journal*, 98(2), 26-28.
- Restrepo, M., & Restrepo, Gómez C. (2004). Sesgos en diseños analíticos. *Revista Colombiana de Psiquiatría*, 33(3), 327-335
- Roceanu, A., & Bajenaru, O. (2014). Current therapies in episodic migraine management. *Romanian Journal Of Neurology*, 13(4), 173-177.

- Rodríguez, L., & Blanco, A. (2005). Evaluación del dolor. *Boletín de Psicología*, 84,77-89.
- Ross, S. (2005). *Introducción a la estadística*. Barcelona: Reverté, S.A.
- Rueda, M., & Díaz, L. (2008). Prevalence and associated factors for episodic and chronic daily headache in the Colombian population. *Cephalalgia*, 28, 216–225.
- Rueda, M., Mantilla, F., Solano, M., & Ortiz, C., (2005). Prevalencia de cefaleas en un servicio de urgencias en Colombia. *Revista de Neurología*, 40 (4), 209-213.
- Ruscheweyh, R. A. (2014). Correlation of Headache Frequency and Psychosocial Impairment in Migraine: A Cross-Sectional Study. *Headache: The Journal Of Head & Face Pain*, 54(5), 861-871.
- Scharff, L., Marcus, D., & Masek B.J. (2002). A controlled study of minimal-contact thermal biofeedback treatment in children with migraine. *Journal of Pediatric Psychology*, 27 (2), 109-119.
- Schulz, K., Altman, D., Moher D., & the CONSORT Group (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Trials*, 11, 1-8.
- Seligman, M. (1995). The effectiveness of psychotherapy the consumer reports study. *American Psychologist*, 50 (1), 965–974.
- Seng, E. A. (2014). Behavioral Migraine Management Modifies Behavioral and Cognitive Coping in People With Migraine. *Headache: The Journal Of Head & Face Pain*, 54(9), 1470-1483.
- Serrano, M., Caballero J., Cañas, A., García, P., Serrano, C., & Prieto J. (2002). Valoración del dolor (II). *Revista de la Sociedad Española del Dolor*, 9, 109-121.
- Shamseer L, Sampson M, Bukutu C, Schmid CH, Nikles J, Tate R, Johnston BC, Zucker D, Shadish WR, Kravitz R, Guyatt G, Altman DG, Moher D, Vohra S; CENT group (2015). CONSORT extension for reporting N-of-1 trials (CENT) 2015:Explanation andelaboration. *British Medical Journal*. doi: 10. 1136/bmj. h1793.

- Sharma, M., & Sharma, V. (2012). Cognitive behavior therapy and relaxation in chronic migraine patients. *SIS Journal of Projective Psychology & Mental Health*, 19 (2), 131-136.
- Silver, N. (2007). Headache (Chronic Tension-Type). *American Family Physician*, 76 (1), 114-116.
- Siniatchkin, M. U. (2000). Self-regulation of slow cortical potentials in children with migraine: an exploratory study. *Applied Psychophysiology & Biofeedback*, 25 (1), 13-32.
- Solar, A. (2010). Methodological aspects of randomized controlled trials on cognitive interventions. *Journal of the Neurological Sciences*, 31 (2), 279–282. doi 10.1007/s10072-010-0375-5
- Tao, K., Li, X., Zhou, Q., Moher, D., Ling C., & Wei, F. (2011). From QUOROM to PRISMA: a survey of high-impact medical journals' instructions to authors and a review of systematic reviews in anesthesia literature. *Plos One*, 6 (11), 1-5.
- Thorn, B., Pence, L., Ward, L., Kilgo, G., Clements K., Cross, T., Davis, A., & Tsui P.W. A (2007). Randomized clinical trial of targeted cognitive behavioral treatment to reduce catastrophizing in chronic headache sufferers. *Journal of Pain*, 8 (12), 938-949.
- Trautmann, E., & Kröner-Herwig, B. (2008). Internet-based self-help training for children and adolescents with recurrent headache: a pilot study. *Behavioural and Cognitive Psychotherapy*, 36, 241-245. doi: 10.1017/S1352465808004219.
- Trillos, C. (2010). Migraña e investigación en epidemiología. *Acta Neurológica Colombiana*, 26 (1), 1-4.
- Tristán A. (2008). Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Avances en Medición*, 6, 37– 48.
- Universidad Nacional de Educación a Distancia (2015). Índices de impacto. Recuperado de http://www.uned.es/biblioteca/guia_impacto/gf2.html#inicio

- University of Iowa Hospitals and Clinics (2011). *Dolor de cabeza*, recuperado de <http://www.uihealthcare.org/adamXml.aspx?product=SpanishMultimedia&type=5&content=003024> HIE
- Urrútia, G., & Bonfill, X. (2010). Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina Clínica*, 135 (11), 507–511.
- Vallejo, M. (2005). Tratamiento psicológico del dolor crónico. *Boletín de Psicología*, 84, 41-58.
- Vallejo, M. (2008). Perspectivas en el estudio y tratamiento psicológico del dolor crónico. *Clínica y Salud*, 19 (3), 417-430.
- Velázquez-Jurado, H. R., Zermeno-Pöhls, J., & Sánchez-Sosa, J. (2013). Detección y valoración integral de la cefalea tensional. *Revista Intercontinental de Psicología y Educación*, 15(1) 61-78. Recuperado de <http://redalyc.org/articulo.oa?id=80225697005>
- Vera, P., & Mustaca, A. (2006). Investigaciones en psicología clínica basadas en la evidencia en Chile y Argentina. *Revista Latinoamericana de Psicología*, 38 (3), 551-565.
- Verhagen, A., Damen, L., Berger, M., Passchier, J., & Koes, B. (2009). Behavioral treatments of chronic tension-type headache in adults: are they beneficial? *CNS Neuroscience & Therapeutics*, 15, 183–205.
- Verhagen, A., de Vet H., de Bie R., Boers M, van den Brandt P. (2001). The art of quality assessment of RCTs included in systematic reviews. *Journal of Clinical Epidemiology*, 54 (7), 651-6544.
- Vicente-Herrero, M. T., Terradillos-García, M. J., Ramírez-Iñiguez de la Torre, M. V., Capdevila-García, L. M., & López-González, Á. A. (2014). Repercusión de las cefaleas sobre la incapacidad laboral en España. (Spanish). *Revista CES Salud Pública*, 5(1), 38-49.
- World Health Organization (2004). *Headache disorders*. Recuperado de <http://www.who.int/mediacentre/factsheets/fs277/en/>.

Yates, S., Morley, S., Eccleston, C., & Williams, A. (2005). A scale for rating the quality of psychological trials for pain. *Pain*, 117 (3), 314–325. doi:10.1016/j.pain.2005.06.018

Apéndices

Apéndice A

Referencias de artículos seleccionados para la investigación

- Abdoli, S., Rahzani, K., Safaie, M., & Sattari, A. (2012). A randomized control trial: the effect of guided imagery with tape and perceived happy memory on chronic tension type headache. *Scandinavian Journal Of Caring Sciences*, 26 (2), 254-261. doi:10.1111/j.1471-6712.2011.00926.x
- Allen K., Elliott A., & Arndorfer, R. (2002). Behavioral pain management for pediatric headache in primary care. *Children's Health Care*, 31 (3), 175-189.
- Andersson, G., Lundström, P., & Ström, L. (2003). Internet-based treatment of headache: does telephone contact add anything? *Headache*, 43 (4), 353-361. doi:10.1046/j.1526-4610.2003.03070.x
- Bromberg, J., Wood, M. E., Black, R. A., Surette, D. A., Zacharoff, K. L., & Chiauzzi, E. J. (2012). A randomized trial of a web-based intervention to improve migraine self-management and coping. *Headache*, 52(2), 244-261. doi:10.1111/j.1526-4610.2011.02031.x
- Calhoun, A., & Ford, S. (2007). Behavioral sleep modification may revert transformed migraine to episodic migraine. *Headache* 47 (8), 1178-1183.
- Ciannella, L., Capobianco, N. N., D'Alessio, A., & Feleppa, M. (2004). An integrated approach to cephalalgic patients. Preliminary results on 64 adult patients with migraine without aura. *Journal of Headache & Pain*, 5, 103-108. doi:10.1007/s10194-004-0121-6
- Cottrell, C., Drew, J., Gibson, J., Holroyd, K., & O'Donnell, F. (2007). Feasibility assessment of telephone-administered behavioral treatment for adolescent migraine. *Headache*, 47 (9), 1293-1302. doi:10.1111/j.1526-4610.2007.00804.x
- Davis, P., Reeves II, J., Graff-Radford, S., Hastie, B., & Naliboff, B. (2003). Multidimensional subgroups in migraine: differential treatment outcome to a pain medicine program. *Pain Medicine*, 4 (3), 215. doi:10.1046/j.1526-4637.2003.03027.x
- Devineni, T., & Blanchard, E. (2005). A randomized controlled trial of an internet-based treatment for chronic headache. *Behavior Research and Therapy*, 43 (3), 277-292.

- Dindo, L., Recober, A., Marchman, J. N., Turvey, C., & O'Hara, M. W. (2012). One-day behavioral treatment for patients with comorbid depression and migraine: A pilot study. *Behaviour Research & Therapy*, 50 (9), 537-543. doi:10.1016/j.brat.2012.05.007
- D'Souza, P., Lumley, M., Kraft, C., & Dooley, J. (2008). Relaxation training and written emotional disclosure for tension or migraine headaches: a randomized, controlled trial. *Annals of Behavioral Medicine*, 36(1), 21-32. doi:10.1007/s12160-008-9046-7
- Fichtel, A., & Larsson, B. (2001). Does relaxation treatment have differential effects on migraine and tension-type headache in adolescents? *Headache*, 41 (3), 290-296.
- Fichtel A. & Larsson B. (2004). Relaxation treatment administered by school nurses to adolescents with recurrent headaches. *Headache*, 44 (6), 545-554.
- Fritsche, G., Frettlöh, J., Hüppe, M., Dlugaj, M., Matatko, N., Gaul, C., & Diener, H. (2010). Prevention of medication overuse in patients with migraine. *Pain*, 151 (2), 404-413. doi:10.1016/j.pain.2010.07.032
- Gerber, W., Petermann, F., Müller, G., Dollwet, M., Darabaneanu, S., Niederberger, U., & Andrasik, F. (2010). MIPAS-Family—evaluation of a new multi-modal behavioral training program for pediatric headaches: clinical effects and the impact on quality of life. *Journal of Headache & Pain*, 11 (3), 215-225. doi:10.1007/s10194-010-0192-5
- Grazzi, L., Andrasik, F., D'Amico, D., Leone, M., Usai, S., Kass, S. J., & Bussone, G. (2002). Behavioral and Pharmacologic Treatment of Transformed Migraine With Analgesic Overuse: Outcome at 3 Years. *Headache*, 42 (6), 483-490. doi:10.1046/j.1526-4610.2002.02123.x
- Gunreben-Stempfle, B., Griesinger, N., Lang, E., Muehlhans, B., Sittl, & R., Ulrich, K. (2009). Effectiveness of an intensive multidisciplinary headache treatment program: Research submission. *Headache*, 49 (7), 990-1000.
- Hansen, J., Bendtsen, L., & Jensen, R. (2007). Predictors of treatment outcome in headache patients with the Millon Clinical Multiaxial Inventory III (MCMI-III). *Journal of Headache & Pain*, 8 (1), 28-34. doi:10.1007/s10194-007-0358-y

- Edberg, K., & Mohr, C. (2011). Multimodal behavioral treatment of migraine: An Internet-administered, randomized, controlled trial Upsala. *Journal of Medical Sciences*, 116 (3), 169-186.
- Holroyd, K. A., Labus, J. S., & Carlson, B. (2009). Moderation and mediation in the psychological and drug treatment of chronic tension-type headache: The role of disorder severity and psychiatric comorbidity. *Pain*, 143(3), 213-222. doi:10.1016/j.pain.2009.02.019
- Holroyd, K., Cottrell, C., O'Donnell, F., Cordingley, G., Drew, J., Carlson, B., & Himawan, L. (2010). Effect of preventive (β blocker) treatment, behavioural migraine management, or their combination on outcomes of optimised acute treatment infrequent migraine: randomized controlled trial. *British Medical Journal*, 341, 1-12. doi:10.1136/bmj.c4871
- Hoodin, F. F., Brines, B. J., Lake, A. E., Wilson, J. J., & Saper, J. R. (2000). Behavioral self-management in an inpatient headache treatment unit: increasing adherence and relationship to changes in affective distress. *Headache*, 40 (5), 377-383.
- Kang, E., Park, J., Chung, C., & Yu, B. (2009). Effect of biofeedback-assisted autogenic training on headache activity and mood states in Korean female migraine patients. *Journal of Korean Medical Science*, 24 (5), 936-940.
- Kaushik, R., Kaushik, R., Mahajan, S., & Rajesh, V. (2005). Biofeedback assisted diaphragmatic breathing and systematic relaxation versus propranolol in long term prophylaxis of migraine. *Complementary Therapies in Medicine*, 13 (3), 165–174. doi: 10.1016/j.ctim.2005.04.004
- Kraft, C. A., Lumley, M. A., D'Souza, P. J., & Dooley, J. A. (2008). Emotional approach coping and self-efficacy moderate the effects of written emotional disclosure and relaxation training for people with migraine headaches. *British Journal of Health Psychology*, 13 (1), 67-71. doi:10.1348/135910707X251144
- Kroener-Herwig, B., & Denecke, H. (2002). Cognitive-behavioral therapy of pediatric headache: Are there differences in efficacy between a therapist-administered group training and a self-help format? *Journal of Psychosomatic Research*, 53 (6), 1107-1114.

- Larsson, B., Carlsson, J., Fichtel, A., & Melin, L. (2005). Relaxation treatment of adolescent headache sufferers: results from a school-based replication series. *Headache*, 45 (6), 692-704. doi:10.1111/j.1526-4610.2005.05138.x
- Lemstra, M., Stewart, B., & Olszynski W.P. (2002). Effectiveness of multidisciplinary intervention in the treatment of migraine: A randomized clinical trial. *Headache*, 42 (9), 845-854.
- Magnusson J., Riess, C., & Becker, W. (2004). Effectiveness of a multidisciplinary treatment program for chronic daily headache. *Canadian Journal of Neurological Sciences*, 31 (1), 72-79.
- Martin, P., Forsyth M., & Reece, J. (2007). Cognitive-behavioral therapy versus temporal pulse amplitude biofeedback training for recurrent headache. *Behavior Therapy*, 38 (4), 350-363.
- Mérelle, S., Sorbi, M., Van Doornen, L., & Passchier, J. (2008). Migraine patients as trainers of their fellow patients in non-pharmacological preventive attack management: short-term effects of a randomized controlled trial. *Cephalalgia*, 28 (2), 127-138. doi:10.1111/j.1468-2982.2007.01472.x
- Mo'tamedi, H., Rezaieamaram, P., & Tavallaie, A. (2012). The effectiveness of a group-based acceptance and commitment additive therapy on rehabilitation of female outpatients with chronic headache: preliminary findings reducing 3 dimensions of headache impact. *Headache*, 52(7), 1106-1119. doi:10.1111/j.1526-4610.2012.02192.x
- Nash, J., Park, E., Walker, B., Gordon, N., & Nicholson, R. (2004). Cognitive-behavioral group treatment for disabling headache. *Pain Medicine*, 5 (2), 178-186. doi:10.1111/j.1526-4637.2004.04031.x
- Nicholson, R., Nash, J., & Andrasik F. (2005). A self-administered behavioral intervention using tailored messages for migraine. *Headache*, 45(9) 1124-1139.
- Scharff, L., Marcus, D., & Masek B.J. (2002). A controlled study of minimal-contact thermal biofeedback treatment in children with migraine. *Journal of Pediatric Psychology*, 27 (2), 109-119.

- Seng, E., & Holroyd, K. (2010). Dynamics of changes in self-efficacy and locus of control expectancies in the behavioral and drug treatment of severe migraine. *Annals of Behavioral Medicine*, 40 (3), 235-247. doi: 10.1007/s12160-010-9223-3
- Sharma, M., & Sharma, V. (2012). Cognitive behavior therapy and relaxation in chronic migraine patients. *SIS Journal of Projective Psychology & Mental Health*, 19 (2), 131-136.
- Siniatchkin, M. U. (2000). Self-regulation of slow cortical potentials in children with migraine: an exploratory study. *Applied Psychophysiology & Biofeedback*, 25 (1), 13-32.
- Thorn, B., Pence, L., Ward, L., Kilgo, G., Clements K., Cross, T., Davis, A., & Tsui P.W. A (2007). Randomized clinical trial of targeted cognitive behavioral treatment to reduce catastrophizing in chronic headache sufferers. *Journal of Pain*, 8 (12), 938-949.
- Trautmann, E., & Kröner-Herwig, B. (2008). Internet-based self-help training for children and adolescents with recurrent headache: a pilot study. *Behavioural and Cognitive Psychotherapy*, 36, 241-245. doi: 10.1017/S1352465808004219.
- Trautmann, E., & Kröner-Herwig, B. (2010). A randomized controlled trial of internet-based self-help training for recurrent headache in childhood and adolescence. *Behaviour Research & Therapy*, 48 (1), 28-37. doi:10.1016/j.brat.2009.09.004
- Wallasch, T., & Hermann, C. (2012). Validation of criterion-based patient assignment and treatment effectiveness of a multidisciplinary modularized managed care program for headache. *Journal of Headache & Pain*, 13 (5), 379-387. doi: 10.1007/s10194-012-0453-6
- Wallasch, T., & Kropp, P. (2012). Multidisciplinary integrated headache care: a prospective 12-month follow-up observational study. *Journal of Headache & Pain*, 13 (7), 521-529. doi: 10.1007/s10194-012-0469-y
- Wallasch, T., Angeli, A., & Kropp, P. (2012). Outcomes of a Headache-Specific Cross-Sectional Multidisciplinary Treatment Program. *Headache*, 52 (7), 1094-1105. doi:10.1111/j.1526-4610.2012.02189.x
- Zsombok, T., Juhasz, G., Budavari, A., Vitrai, J., & Bagdy G. (2003). Effect of autogenic training on drug consumption in patients with primary headache: An 8-month follow-up study. *Headache*, 43 (3), 251-257

Apéndice B

Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01).

UNIVERSIDAD PONTIFICIA BOLIVARIANA
Inventario de valoración de estándares metodológicos y de publicación
Referencia bibliográfica del artículo: Evaluador: Fecha de evaluación:
Instrucciones
En cada una de las páginas de este archivo, encontrará una serie de afirmaciones (indicadores) relacionadas con categorías de estándares metodológicos y de publicación en artículos científicos de tipo experimental. Usted debe hacer una valoración del artículo en cada una de estas afirmaciones teniendo en cuenta la siguiente escala: No cumple con lo propuesto por el indicador (0), Cumple plenamente lo propuesto por el indicador (1). Si considera pertinente no darle calificación al ítem coloque no aplica (NA) por ej. se pretende dar valoración a un elemento que se dijo que no existía y por lo tanto había obtenido una evaluación de "0". Cuando lo requiera consigne sus comentarios en la columna de observaciones.

Categoría Portada: Es la parte inicial de un artículo científico que permite identificar las características centrales de este. Subcategorías Título: Frase que identifica el contenido del artículo. Resumen: Texto breve que sintetiza los aspectos más relevantes de la investigación.			
Subcategoría	Indicadores/ítem	Valoración	Observaciones
Título	P 1.1 En el título se nombra la palabra aleatorio o su equivalente (p.ej. Azar).		
	P 1.2 En el título se nombran las variables independiente y dependiente.		
	P 1.3 El título es coherente con el contenido del artículo.		
	P 1.4 El título se comprende fácilmente (es claro).		
	P 1.5 En el título se utilizan palabras completas (<u>no</u> aparecen abreviaturas).		

	P 1.6 El título se presenta de manera afirmativa (hay ausencia de negaciones o interrogantes)		
Resumen	P 2.1 El resumen se comprende fácilmente (es claro).		
	P 2.2 El resumen está compuesto por 250 palabras o menos.		
	P 2.3 El resumen presenta las características esenciales del artículo (ideas centrales y contenidos de los apartados por los que está compuesto).		
	P 2.4 En el resumen se presenta de manera explícita el problema de investigación.		
	P 2.5 En el resumen se presenta de manera explícita el objetivo y/o hipótesis de la investigación.		
	P2.6 En el resumen se hace evidente el marco conceptual que sustenta el desarrollo de la investigación.		
	P 2.7 En el resumen se presenta de manera explícita el diseño de investigación.		
	P 2.8 En el resumen se presentan las características centrales de los sujetos de investigación, según lo expuesto en el apartado de método.		
	P 2.9 En el resumen se presenta el número de participantes del estudio.		
	P 2.10 En el resumen se presenta el número de participantes asignados al azar al (a los) grupo(s) control y experimental.		
	P 2.11 En el resumen se presenta el número de participantes que fueron objeto de análisis estadístico en el (los) grupo(s) control y experimental.		
	P 2.12 En el resumen se hace explícita la forma en que se asignaron los sujetos al (a los) grupo(s) control y experimental.		
	P 2.13 En el resumen se presentan las características principales de los sitios o centros (ej. Hospital o centro médico de primer, segundo o tercer nivel) donde se realizó la investigación.		
	P 2.14 En el resumen se nombran los aparatos o instrumentos utilizados.		
	P 2.15 En el resumen se presenta el tipo de intervención realizada al grupo experimental.		
	P 2.16 En el resumen se presenta el tipo de intervención realizada al grupo control.		
	P 2.17 En el resumen se presentan los datos estadísticos más relevantes según lo expuesto en el apartado de resultados.		
	P 2.18 En el resumen se presenta el procedimiento utilizado por los investigadores para obtener la información.		

	P 2.19 En el resumen se presenta el tamaño del efecto, intervalos de confianza y niveles de significación estadística.		
	P 2.20 En el resumen se presentan las principales Conclusiones e implicaciones según lo expuesto en el texto del artículo.		

Categoría Introducción: texto en forma de ensayo que se dirige a contextualizar y justificar la realización de la investigación. Subcategorías Contexto investigativo: antecedentes teóricos y razones (justificación) por las cuales se desarrolla el estudio. Propósitos de la investigación: apartado que está conformado por los objetivos a cumplir, la hipótesis a validar y el problema o interrogante a responder.			
Subcategoría	Indicadores/ítem	Valoración	Observaciones
Contexto investigativo	I 1.1 En introducción se presentan referentes conceptuales que argumentan o sustentan el tema de investigación.		
	I 1.2 En la introducción se presenta de manera explícita la relación entre la investigación realizada y estudios anteriores.		
	I 1.3 En la introducción se hace evidente el marco conceptual que sustenta la generación de las hipótesis de investigación.		
	I 1.4 El contenido de la introducción es claro y coherente.		
	I 1.5 En la introducción se justifica la realización de la investigación desde una perspectiva práctica, metodológica, teórica y/o de impacto social.		
Propósitos de la investigación	I 2.1 En la introducción se presenta de manera explícita el objetivo y/o hipótesis de la investigación.		
	I 2.2 Los objetivos son claros y concisos.		
	I 2.3 Las hipótesis de investigación son claras.		
	I 2.4 En la introducción existe una relación entre el objetivo y la hipótesis.		
	I 2.5 En la introducción se presenta de manera explícita el problema de investigación.		
	I 2.6 En la introducción existe una relación entre la hipótesis y el problema de investigación.		

Categoría Método: apartado que agrupa los elementos que describen cómo se desarrolló la investigación. Subcategorías Diseño: forma o estrategia general para el desarrollo de la investigación. Participantes: Sujetos que participan suministrando información, evaluando, interviniendo o en algún proceso relevante del desarrollo de la investigación. Instrumentos: prueba o aparato utilizado en el proceso de evaluación y/o intervención. Definición de variables: descripción de las propiedades que caracterizan las variables. Procedimiento: descripción de cada uno de los pasos que se implementaron para el desarrollo de la investigación. Procesamiento estadístico: descripción de la forma en que se abordó el análisis cuantitativo de los datos. Consideraciones éticas: apartado que presenta los elementos que se tuvieron en cuenta para cumplir con los requerimientos éticos de la investigación			
Subcategoría	Indicadores/ítem	Valoración	Observaciones
Diseño	M 1.1 Se presenta de manera explícita el diseño de investigación.		
	M 1.2 El diseño se describe de forma suficiente, para que pueda replicarse la investigación.		
Participantes	a) Sujetos		
	M 2.1 Los sujetos se describen de forma suficiente, para que pueda replicarse la investigación.		
	M 2.2 En el apartado de sujetos, se presentan las características centrales de la población investigada (Ej. Número, características sociodemográficas, características psicológicas).		
	M 2.3 En el apartado de sujetos, se describe el proceso de selección de la muestra.		
	M 2.4 En el apartado de sujetos, se describe la forma para calcular el número de sujetos de la muestra.		
	M 2.5 En el apartado de sujetos, se presenta el número de participantes del estudio.		
	M 2.6 En el apartado de sujetos, se presenta el número de participantes en cada uno de los grupos (experimental, control).		
	M 2.7 En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que recibieron el tratamiento planeado.		

M 2.8 En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que completaron el protocolo de intervención.		
M 2.9 En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que se analizaron para obtener el resultado principal (más importante) del estudio.		
M 2.10 En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que después de la asignación al azar, abandonaron (no continuaron) en el proceso investigativo.		
M 2.11 En el apartado de sujetos, se presentan las razones de abandono (no continuación) del proceso investigativo, después de la asignación al azar, por parte de las personas en cada uno de los grupos (experimental, control).		
M 2.12 En el apartado de sujetos, se presentan las características centrales de la muestra investigada (Ej. número, características socio demográficas, características psicológicas).		
M 2.13 En el apartado de sujetos se presentan los criterios de inclusión y exclusión de los participantes.		
M 2.14 En el apartado de sujetos, se presentan de manera explícita las retribuciones (en dinero o especie) recibidas por los sujetos, debido a su participación en el estudio.		
M 2.15 En el apartado de sujetos, se presenta el diagrama de flujo de participantes.		
M 2.16 En el apartado de sujetos, se presenta el número de pacientes tratados por cada terapeuta.		
M 2.17 En el apartado de método, se presenta el número de pacientes tratados en cada centro o lugar donde se hace la intervención.		
b) Equipo de investigación		
Reclutadores		
M 2.18 En el apartado de método, se identifica de manera explícita la persona que estuvo encargada de incorporar las personas al estudio.		
Generadores de asignación al azar		
M 2.19 En el apartado de método, se identifica de manera explícita la persona que generó la secuencia de asignación al azar.		
M 2.20 En el apartado de método, se identifica de manera explícita la persona que asignó los participantes a las intervenciones o grupos.		
Evaluadores		

	M 2.21 En el apartado de método, se presenta la formación y experiencia de los evaluadores (personas que hacen los registros y/o califican las pruebas).		
	Terapeutas		
	M 2.22 En el apartado de método, se presentan los criterios de selección, para las personas encargadas del proceso de intervención.		
	M 2.23 En el apartado de método, se presenta el número de personas encargadas del proceso de intervención en cada grupo.		
	M 2.24 En el apartado de método, se presenta la formación y experiencia de las personas encargadas del proceso de intervención.		
	c) Ambientes y ubicaciones		
	M 2.25 En el apartado de método, se presentan las características principales de los sitios o centros (ej. Hospital o centro médico de primer, segundo o tercer nivel donde se realizó la investigación).		
	M 2.26 En el apartado de método, se presentan los criterios de selección de los ambientes y ubicaciones donde se recolectaron los datos.		
Instrumentos	M 2.27 En el apartado de método, se presenta el número de centros o lugares en donde se hace la intervención para cada grupo (control, experimental)		
	M 3.1 En el apartado de instrumentos, se nombran y describen las pruebas y/o aparatos utilizados en el proceso investigativo.		
Definición de variables	M 3.2 En el apartado de instrumentos, se presentan las propiedades psicométricas (confiabilidad, validez) de las pruebas y/o aparatos utilizados en el proceso investigativo.		
	M 4.1 En el apartado de método, se hacen explícitas las variables dependientes e independientes.		
	M 4.2 En el apartado de método, se definen operacionalmente las variables y sus grados.		
Procedimiento	M 4.3 En el apartado de método, las variables se definen claramente.		
	M 5.1 En el apartado de método, se presenta el procedimiento de la investigación		
	M 5.2 El procedimiento se describe de forma suficiente, para que pueda replicarse la investigación.		
	M 5.3 En el apartado de método, se presenta el número de registro y el nombre de registro del experimento.		
	M 5.4 En el apartado de procedimiento se presentan los periodos o momentos en los cuales fueron incorporados los sujetos a la investigación.		

	M 5.5 En el procedimiento se hace explícito qué participantes se asignaron de manera aleatoria al grupo control y al experimental.		
	M 5.6 En el procedimiento, se especifica como se hizo el proceso de asignación a los grupos (experimental y control).		
	M 5.7 En el procedimiento, se describe la forma en que se ocultó la secuencia (programa) de asignación al azar, hasta el momento en que se asignan las intervenciones a cada uno de los sujetos.		
	M 5.8 En el procedimiento, se hace evidente que los participantes fueron enmascarados (cegados) a la asignación de las intervenciones.		
	M 5.9 En el procedimiento, se describe como los participantes fueron enmascarados (cegados) a la asignación de las intervenciones.		
	M 5.10 En el procedimiento se hace explícito que los terapeutas fueron enmascarados o cegados a la asignación de las intervenciones.		
	M 5.11 En el procedimiento, se explica cómo los terapeutas fueron enmascarados (cegados) a la asignación de las intervenciones.		
	M 5.12 En el procedimiento se hace explícito que los evaluadores fueron enmascarados o cegados a la asignación de las intervenciones.		
	M 5.13 En el procedimiento se describe como los evaluadores fueron enmascarados o cegados a la asignación de las intervenciones.		
	M 5.14 En el procedimiento, se presenta la forma en que fueron asignados los terapeutas al grupo control y experimental.		
	M 5.15 En el procedimiento se especifica cómo se hizo el control de variables extrañas.		
	M 5.16 En el procedimiento se evidencia que el control de variables extrañas es riguroso.		
	M 5.17 En el procedimiento se describe la forma en que se llevó a cabo o se desarrolló la evaluación.		
	M 5.18 En el procedimiento se describe en qué momento o tiempo se implementó la evaluación.		
	M 5.19 En el procedimiento, la evaluación se describe de forma suficiente, para que pueda replicarse la investigación.		
	M 5.20 En el procedimiento, la intervención se describe de forma suficiente, para que pueda replicarse la investigación.		
	M 5.21 En el procedimiento, se describe en qué momento o tiempo se implementó la intervención.		

	M 5.22 En el procedimiento, se hace explícita la forma en que se estandarizaron las intervenciones.		
	M 5.23 En el procedimiento, se hace explícita la forma como se controló la adherencia al protocolo, por parte de los terapeutas.		
	M 5.24 En el procedimiento, se hacen explícitas las circunstancias o aspectos de la investigación que pueden afectar la validez de los resultados.		
Procesamiento estadístico.	M 6.1 En el apartado de método, se nombra el software estadístico utilizado.		
	M 6.2 En el apartado de método, se presentan las mediciones o medidas de los resultados principales y complementarios.		
	M 6.3 En el apartado de método, se presentan los estadísticos utilizados para comparar los grupos en los resultados principales y complementarios.		
	M 6.4 En el apartado de método, se describe el procesamiento estadístico de forma suficiente, para que pueda replicarse la investigación.		
	M 6.5 En el apartado de método, se informa sobre circunstancias del análisis de datos que pueden afectar la validez de los resultados.		
Consideraciones éticas	M 7.1 En el apartado de método, se hace explícito que la investigación se está guiando por un referente ético formal (ej. Código de ética de la Asociación Psicológica Americana).		
	M 7.2 En el apartado de método, se hace explícito que se utilizó el consentimiento informado de los participantes.		
	M 7.3 En el apartado de método, se describe la forma en que se obtuvo el consentimiento informado.		
	M 7.4 En el apartado de método, se hace evidente que el grupo control fue intervenido posterior a la finalización de la investigación.		
	M 7.5 En el apartado de método, se nombran las retribuciones dadas a los participantes, en dinero o especie.		
	M 7.6 En el apartado de método, se hace explícito que la investigación fue aprobada por un comité de ética institucional.		
	M 7.7 En el apartado de método, se hace explícito que las normas éticas se cumplieron.		
	M 7.8 En el apartado de método, se describe cómo se hizo el monitoreo de cumplimiento de las normas éticas.		
	M 7.9 En el apartado de método, se describe la forma en que se garantiza la confidencialidad de los datos en la publicación.		
	M 7.10 En el apartado de método, se presentan de manera explícita los conflictos de interés en el estudio.		

	M 7.11 En el apartado de método, se describe como se manejaron los conflictos de interés.		
	M 7.12 En el apartado de método, se evidencia que los datos empíricos y/o materiales están disponibles para que otro investigador esté en capacidad de utilizarlos y reproducir un análisis que conduzca a las Conclusiones del autor.		

Categoría Resultados: presentación de datos cuantitativos y/o cualitativos arrojados por la investigación. Subcategorías General: Consideraciones genéricas a ser tenidas en cuenta para la presentación de los datos. Texto: consideraciones a ser tenidas en cuenta en la descripción de los datos. Tablas: Consideraciones a ser tenidas en cuenta en la presentación de los datos en formato de Tabla. Gráficas: Consideraciones a ser tenidas en cuenta en la presentación de los datos en formato de imágenes o gráficos.			
Subcategoría	Indicadores/ítem	Valoración	Observaciones
General	R 1.1 En los resultados, los datos se presentan de forma organizada (ordenada).		
	R 1.2 En los resultados, se presenta la frecuencia y/o porcentaje de datos faltantes (valores perdidos)		
	1.3 En los resultados, se presenta el número de casos que se eliminan (no se tiene en cuenta) para cada uno de los análisis realizados.		
	R 1.4 En los resultados, se reporta el tamaño de muestra por celda o subgrupo evaluado.		
	R 1.5 En los resultados, se presentan los resultados principales y complementarios para cada grupo.		
	1.6 En los resultados, se presentan los tamaños del efecto y los intervalos de confianza.		
	R 1.7 En los resultados, se presenta la tasa de error tipo I aceptada previamente.		
	R 1.8 En los resultados, se presenta el valor de error estadístico (ej. $p < 0,05$).		
	R 1.9 En los resultados, se reporta la dirección de "p" para el análisis estadístico realizado.		

	R 1.10 En los resultados, se reportan los grados de libertad para el análisis estadístico realizado.		
	R 1.11 En los resultados, los procedimientos estadísticos se aplican correctamente para el nivel de medición de los datos.		
	R 1.12 En los resultados, los datos están dirigidos a evaluar las hipótesis planteadas.		
	R 1.13 En los resultados, los datos se presentan con precisión (no existe ambigüedad)		
	R 1.14 En los resultados, se hace explícito que el grupo control y experimental son similares en las variables relevantes, antes de implementarse la intervención.		
Texto	R 2.1 En los resultados, el texto es claro.		
	R 2.2 En los resultados, se justifican de forma empírica y/o teórica las causas para que existan "valores perdidos" en la investigación.		
	R 2.3 En los resultados, la información en texto es complementaria a la de las Tablas (no existe información redundante)		
Tablas	R 3.1 En los resultados, se presenta una Tabla con la línea de base de variables demográficas y clínicas de cada grupo.		
	R 3.2 En los resultados, las Tablas son claras (fáciles de entender por un experto).		
Gráficas	R 4.1 En los resultados, las gráficas son claras (fáciles de entender por un experto).		
	R 4.2 En los resultados, las gráficas son pertinentes (adecuadas) para la información que se presenta.		

Categoría

Discusión : Análisis de los resultados obtenidos en el contexto de la evidencia existente

Subcategorías

Análisis: debate entre los resultados de la investigación, el conocimiento en el área de estudio y las características propias de la investigación.

Generalización: posibilidad de que los resultados, sean aplicables a otros sujetos, condiciones y/o metodologías similares a los de la investigación realizada.

Implicaciones: aspectos que se derivan del estudio y pueden guiar la investigación y/o las practicas futuras.

Sub categoría	Indicadores/ítem	Valoración	Observaciones
Análisis	D 1.1 En la discusión, el análisis de los datos se hace en función (está relacionado) con los objetivos y/o hipótesis del estudio.		
	D 1.2 En la discusión, el análisis de los datos se hace en función (está relacionado) con el problema de investigación del estudio.		
	D 1.3 En la discusión, se comparan los resultados de la investigación con otros estudios.		
	D 1.4 En la discusión, se presentan las limitaciones del estudio que afectan la validez interna.		
	D 1.5 En la discusión, se retoman los resultados, pero no se repiten.		
	D 1.6 En la discusión se evidencia coherencia con los resultados.		
	D 1.7 En la discusión, se presentan los efectos secundarios negativos de la(s) intervención(es).		
Generalización	D 2.1 En la discusión, las generalizaciones están basadas en los resultados.		
	D 2.2 En la discusión, se especifica la población a la cual puede ser generalizado el estudio.		
	D 2.3 En la discusión, se especifican los contextos a los cuales puede ser generalizado el estudio.		
	D 2.4 En la discusión, se especifican las problemáticas a las cuales puede ser generalizado el estudio.		
	D 2.5 En la discusión, se especifican los métodos de recolección de datos a los cuales puede ser generalizado el estudio		
	D 2.6 En la discusión, se especifican las intervenciones a las cuales puede ser generalizado el estudio.		
	D 2.7 En la discusión, se especifican las características de los terapeutas a los cuales puede ser generalizado el estudio.		
	D 2.8 En la discusión, se justifica la razón por las cuales se hacen las generalizaciones propuestas.		

Implicaciones	D 3.1 En la discusión, se presentan implicaciones para la investigación futura, programas y/o políticas.		
----------------------	--	--	--

Categoría Conclusiones: Principales deducciones que se generan posterior al análisis y discusión de los resultados.			
Subcategoría	Indicadores/ítem	Valoración	Observaciones
	C 1.1 Las Conclusiones están fundamentadas en los resultados.		
	C 1.2 Las Conclusiones se hacen en función (están relacionadas) con los objetivos de investigación.		

Categoría Referencias: listado ordenado y detallado en base a una metodología específica, de las fuentes de donde se obtuvo la información que sustenta la investigación.			
Subcategoría	Indicadores/ítem	Valoración	Observaciones
	RF 1.1 Las referencias son acordes con el título y/o temática central del artículo.		
	RF 1.2 Las referencias son actualizadas (más del 50% hace referencia a artículos con fechas comprendidas en los cinco años previos a la publicación del estudio).		
	RF 1.3 El número de referencias es superior a 25.		
	RF 1.4 El tipo de referencias es adecuado (más del 50% son artículos originales).		
	RF 1.5 Las referencias poseen los elementos requeridos para las normas utilizadas (ej. APA).		
	RF 1.6 Las referencias son confiables (predominan los artículos de revistas especializadas en el área de conocimiento del artículo).		

Apéndice C

Referencias de artículos que presentan estándares metodológicos y de publicación

- American Educational Research Association (2006). Standards for reporting on empirical social science Research In Aera Publications. *Educational Researcher*, 35 (6), 33–40.
- APA Publications and Communications Board Working Group on Journal Article Reporting Standards (2008). Reporting standards for research in psychology why do we need them? What Might They Be? *American Psychologist* 63 (9), 839–85.
- Bobenrieth, M. (2002). Normas para revisión de artículos originales en Ciencias de la Salud. *Revista Internacional de Psicología Clínica y de la Salud*, 2(3), 509-523.
- Borg, V., Zhang, S., Ye, C., Paul J., Arya, A., Hurlburt, L., Murthy, Y., & Thabane, L. (2012). The quality of reporting of RCTs used within a postoperative pain management meta-analysis, using the CONSORT statement. *BioMed Central Anesthesiology*, 12 (13), 1-11.
- Boutron, I., Moher, D., Altman, D., Schulz, K., & Ravaud, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of Internal Medicine*, 148, 295-309.
- Hopewell S., Clarke M., Moher, D., Wager, E., Middleton P., Altman D., Schultz K. & The CONSORT Group (2008). CONSORT for reporting ramdomised trials in journal and conference abstracs. *Lancet*, 371 (9609), 281-283.
- Ramos, M., Moreno, M., Valdés, B., & Catena, A. (2008) .Criteria of the peer review process for publication of experimental and quasiexperimental research in Psychology: A guide for creating research papers. *International Journal of Clinical and Health Psychology*, 8 (3), 751-764.
- Schulz, K., Altman, D., Moher D., & the CONSORT Group (2010).CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Trials*, 11, 1-8.
- Seligman, M. (1995). The effectiveness of psychotherapy the consumer reports study. *American Psychologist*, 50 (1), 965–974.

Apéndice D

Carta para evaluación por los jueces

Floridablanca, octubre

Doctor (a)

Docente investigador

Cordial Saludo.

Teniendo en cuenta su conocimiento teórico y metodológico en el contexto investigativo, nos dirigimos a usted con el fin de solicitarle su participación como evaluador de contenido del **Inventario de valoración de estándares metodológicos y de publicación** el cual será utilizado para la investigación “Calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional: una revisión sistemática 2000-2012” para la obtención del título de Maestría en Psicología de la Universidad Pontificia Bolivariana Bucaramanga.

Para la evaluación, se anexan dos documentos: **Información de la prueba**, que contiene el resumen de la investigación en la cual se va a utilizar el instrumento, la ficha técnica y estructura del instrumento. Por otra parte, se encuentra el **Inventario**, que está compuesto por varias hojas que van desde encabezado hasta referencias. En cada una de las hojas se encuentra las definiciones de las categorías y subcategorías, los indicadores y dos columnas (valoración y observaciones), las cuales serán diligenciadas por la persona que va a valorar el artículo en referencia a cada uno de los indicadores. Adicional a estas columnas se encuentran los espacios para ser diligenciados por el evaluador o jurado del instrumento (pertinencia y redacción). La forma de diligenciar el formato consiste en registrar para cada indicador (afirmación) una valoración numérica de 1 a 5 (el 1 indica muy bajo cumplimiento del criterio y el 5 muy alto cumplimiento del criterio). En cuanto a pertinencia y redacción según las siguientes definiciones:

Pertinencia: la afirmación es un estándar metodológico y/o de publicación de la categoría en la cual se encuentra ubicado.

Redacción: la afirmación tiene una estructura gramatical adecuada y es fácil de comprender.

Si usted considera pertinente, puede hacer las observaciones del caso, en el espacio dispuesto para ello.

También le solicitamos el favor de revisar la casilla de clasificación del estándar, en donde se categoriza cada indicador en estándar metodológico (EM) y estándar de publicación (EP) y hacer las observaciones pertinentes en cuanto a si el indicador corresponde a estas categorías, teniendo en cuenta las definiciones conceptuales propuestas en la ficha técnica de la prueba.

Agradecemos su colaboración para la evaluación del instrumento y contribuir al desarrollo investigativo en el área de la medición. Le solicitamos el favor de hacernos llegar su evaluación antes del miércoles 6 de noviembre de 2013

Cordialmente,

Luis Argenis Osorio Ferrer
Estudiante Maestría en Psicología
Universidad Pontificia Bolivariana Bucaramanga

Ana Fernanda Uribe Rodríguez
Directora Tesis de Maestría
Universidad Pontificia Bolivariana Bucaramanga

Apéndice E

Información sobre la prueba (resumen de la investigación, ficha técnica y la estructura de la prueba)



Universidad Pontificia Bolivariana

Formato de Evaluación por jueces

Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01).

El Inventario de valoración de estándares metodológicos y de publicación, se diseñó y construyó para ser utilizado en la investigación denominada “**Calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivas conductuales de cefaleas tipo migraña y tensional: una revisión sistemática 2000-2012.** Con el fin de tener una mejor contextualización se presenta a continuación el resumen del anteproyecto.

RESUMEN

El presente estudio tiene como objetivo determinar la calidad metodológica y seguimiento de estándares de publicación, en intervenciones cognitivas conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012, por medio de una revisión sistemática. El tipo de investigación a desarrollar será una revisión sistemática de literatura. Para la investigación se seleccionarán artículos de texto completo de bases de datos generales, de ciencias de la salud y de psicología. El instrumento a utilizar será Inventario de valoración de estándares metodológicos y de publicación. Entre los resultados esperados se encuentran el diseño y validación de un instrumento para la valoración de seguimiento de estándares metodológicos en trabajos y artículos de investigación y de publicaciones; formación a docentes de la universidad en estudios sobre revisión de investigaciones; y la consolidación de una línea de base sobre el nivel metodológico y de seguimiento de estándares de publicación en intervenciones cognitivas conductuales de cefaleas tipo migraña y tensional entre los años 2000-2012.

Ficha técnica

Finalidad: Evaluar estándares metodológicos y de publicación en artículos que reporten en intervenciones psicológicas.

Nombre: Inventario de valoración de estándares metodológicos y de publicación.

Código: INVEMCION-01.

Autores: Luis Argenis Osorio Ferrer y Ana Fernanda Uribe Rodríguez.

Característica evaluada: estándares metodológicos y de publicación

Definiciones conceptuales

Estándares de publicación: Normas para reportar una investigación de una manera exacta, clara y sin sesgo. (ICMJ, 2013).

Estándares metodológicos: Normas que regulan el diseño de un experimento, con el fin de que generen resultados no sesgados, que sean suficientemente precisos y permitan la aplicación en la práctica clínica (Verhagen, de Vet, de Bie, Boers, Van den Brandt, 2001).

Definiciones Operacionales:

Estándares de publicación: Número de indicadores rotulados o definidos dentro de esta categoría en cada uno de los apartados del Inventario de valoración de estándares metodológicos y de publicación (ver formato del inventario).

Estándares metodológicos: Número de indicadores rotulados o definidos dentro de esta categoría en cada uno de los apartados del Inventario de valoración de estándares metodológicos y de publicación (ver formato del inventario).

Tipo de prueba: Inventario

Tiempo de aplicación estimado: Cinco (5) horas

Objeto de evaluación: artículos especializados de intervenciones cognitivo conductuales de cefaleas tipo migraña y tensional

Lugar de realización: Universidad pontificia Bolivariana Bucaramanga

Fecha: 2013

Estructura de la prueba

Las categorías y subcategorías que se presentan en la estructura de la prueba, fueron extraídas de la tabulación de estándares de publicación y/o metodológicos, publicados en artículos especializados en este tópico.

Tabla 1. Estructura del Inventario de valoración de estándares metodológicos y de publicación

Tema/categoría	indicadores	Peso	Tipo de reactivo
Portada	34	2	Dicotómico
Introducción	11	20	Dicotómico
Método	76	30	Dicotómico
Resultados	21	30	Dicotómico
Discusión	16	10	Dicotómico
Conclusiones	2	5	Dicotómico
Referencias	6	3	Dicotómico
Total: 7	166	100	Dicotómico

La Tabla 1 presenta la estructura de la prueba en donde se incluyen las categorías, subcategorías, los indicadores, el peso o importancia asignada a la categoría y el tipo de ítem utilizado para la evaluación del indicador.

Tabla 2. Estructura general del instrumento por categorías, subcategorías e indicadores.

Categoría	Subcategoría	Indicador	Total por categoría
Portada	Título	7	34
	Resumen	27	
Introducción	Contexto investigativo	5	11
	Propósitos de la investigación	6	
Método	Diseño	2	76
	Participantes	28	
	Instrumentos	2	
	Definición de variables	3	
	Procedimiento	24	
	Procesamiento estadístico	5	
	Consideraciones éticas	12	
Resultados	General	14	21
	texto	3	
	Tablas	2	
	Graficas	2	
Discusión	Análisis	7	16
	Generalización	8	

	Implicaciones	1	
Conclusiones		2	2
Referencias		6	6
Total: 7	18		166

Apéndice F

Inventario de valoración de estándares metodológicos y de publicación (INVEMCION-01) para validación por jueces

Inventario de valoración de estándares metodológicos y de publicación
Referencia bibliográfica del artículo:
Evaluador:
Fecha de evaluación:
Instrucciones
En cada una de las páginas de este archivo, encontrará una serie de afirmaciones (indicadores) relacionadas con categorías de estándares metodológicos y de publicación en artículos científicos de tipo experimental. Usted debe hacer una valoración del artículo en cada una de estas afirmaciones teniendo en cuenta la siguiente escala: No cumple con lo propuesto por el indicador (0), Cumple plenamente lo propuesto por el indicador (1). Cuando lo requiera consigne sus comentarios en la columna de observaciones.

Categoría

Portada: Es la parte inicial de un artículo científico que permite identificar las características centrales de este.

Subcategorías

Título: Frase que identifica el contenido del artículo.

Resumen: Texto breve que sintetiza los aspectos más relevantes de la investigación.

Subcategoría	Indicadores/ítem	Valoración	Observaciones	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
Título	En el título se nombra la palabra aleatorio o su equivalente (p.ej. Azar).						EP	
	En el título se nombran las variables independiente y dependiente.						EP	
	El título es coherente con el contenido del artículo.						EP, EM	
	El título se comprende fácilmente (es claro).						EP	
	El título está compuesto por 15 palabras o menos.						EP	
	En el título se utilizan palabras completas (no aparecen abreviaturas).						EP	
	El título se presenta de manera afirmativa (hay ausencia de negaciones o interrogantes)						EP	

Resumen	El resumen se comprende fácilmente (es claro).						EP	
	El resumen está compuesto por 250 palabras o menos.						EP	
	El resumen presenta las características esenciales del artículo (ideas centrales y contenidos de los apartados por los que está compuesto).						EP	
	En el resumen se presenta de manera explícita el problema de investigación.						EP	
	En el resumen se presenta de manera explícita el objetivo y/o hipótesis de la investigación.						EP	
	En el resumen se hace evidente el marco conceptual que sustenta el desarrollo de la investigación.						EP	
	En el resumen se presenta de manera explícita el diseño de investigación.						EP	
	En el resumen se presentan las características centrales de los sujetos de investigación, según lo expuesto en el apartado de método.						EP	
	En el resumen se presentan los criterios de inclusión y exclusión de los participantes.						EP	
	En el resumen se presenta el número de participantes del estudio.						EP	
	En el resumen se presenta el número de participantes asignados al azar al (a los) grupo(s) control y experimental.						EP	
	En el resumen se presenta el número de participantes que fueron objeto de análisis estadístico en el (los) grupo(s) control y experimental.						EP	
	En el resumen se hace explícita la forma en que se asignaron los sujetos al (a los) grupo(s) control y experimental.						EP	
	En el resumen se presentan las características centrales de los terapeutas de investigación, según lo expuesto en el apartado de método.						EP	

	En el resumen se presentan las características principales de los sitios o centros (ej. Hospital o centro médico de primer, segundo o tercer nivel) donde se realizó la investigación.						EP	
	En el resumen se nombran los aparatos o instrumentos utilizados.						EP	
	En el resumen se presenta el tipo de intervención realizada al grupo experimental.						EP	
	En el resumen se presenta el tipo de intervención realizada al grupo control.						EP	
	En el resumen se especifica la forma en la cual se hizo cegamiento (enmascaramiento) a los participantes de la investigación						EP	
	En el resumen se presentan los datos estadísticos más relevantes según lo expuesto en el apartado de resultados.						EP	
	En el resumen se presenta el procedimiento utilizado por los investigadores para obtener la información.						EP	
	En el resumen se presenta el tamaño del efecto, intervalos de confianza y niveles de significación estadística.						EP	
	En el resumen se presentan los efectos secundarios en los sujetos, como consecuencia de la intervención implementada.						EP	
	En el resumen se presentan las principales Conclusiones e implicaciones según lo expuesto en el texto del artículo.						EP	
	En el resumen se presenta el nombre y número de registro del experimento.						EP, EM	
	En el resumen se nombran las instituciones que financiaron la investigación.						EP, EM	
	En el resumen especifica la información que ha sido presentada en forma oral y/o escrita, en artículos o eventos anteriores.						EP,EM	

Categoría
Introducción: texto en forma de ensayo que se dirige a contextualizar y justificar la realización de la investigación.
Subcategorías
Contexto investigativo: antecedentes teóricos y razones (justificación) por las cuales se desarrolla el estudio.
Propósitos de la investigación: apartado que está conformado por los objetivos a cumplir, la hipótesis a validar y el problema o interrogante a responder.

Subcategoría	Indicadores/ítem	Valoración	Observaciones	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
Contexto investigativo	En introducción se presentan referentes conceptuales que argumentan o sustentan el tema de investigación.						EP y EM	
	En la introducción se presenta de manera explícita la relación entre la investigación realizada y estudios anteriores.						EP y EM	
	En la introducción se hace evidente el marco conceptual que sustenta la generación de las hipótesis de investigación.						EP y EM	
	El contenido de la introducción es claro y coherente.						EP	
	En la introducción se justifica la realización de la investigación desde una perspectiva práctica, metodológica, teórica y/o de impacto social.						EP y EM	
Propósitos de la investigación	En la introducción se presenta de manera explícita el objetivo y/o hipótesis de la investigación.						EP y EM	
	Los objetivos son claros y concisos.						EP y EM	
	Las hipótesis de investigación son claras.						EP y EM	

	En la introducción existe una relación entre el objetivo y la hipótesis.						EP y EM	
	En la introducción se presenta de manera explícita el problema de investigación.						EP y EM	
	En la introducción existe una relación entre la hipótesis y el problema de investigación.						EP y EM	

Categoría

Método: apartado que agrupa los elementos que describen cómo se desarrolló la investigación.

Subcategorías

Diseño: forma o estrategia general para el desarrollo de la investigación.

Participantes: Sujetos que participan suministrando información, evaluando, interviniendo o en algún proceso relevante del desarrollo de la investigación.

Instrumentos: prueba o aparato utilizado en el proceso de evaluación y/o intervención.

Definición de variables: descripción de las propiedades que caracterizan las variables.

Procedimiento: descripción de cada uno de los pasos que se implementaron para el desarrollo de la investigación.

Procesamiento estadístico: descripción de la forma en que se abordó el análisis cuantitativo de los datos.

Consideraciones éticas: apartado que presenta los elementos que se tuvieron en cuenta para cumplir con los requerimientos éticos de la investigación.

Subcategoría	Indicadores/ítem	Valoración	Observaciones	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
Diseño	Se presenta de manera explícita el diseño de investigación.						EP y EM	
	El diseño se describe de forma suficiente, para que pueda replicarse la investigación.						EP y EM	
Participantes	a) Sujetos							
	Los sujetos se describen de forma suficiente, para que pueda replicarse la investigación.						EP y EM	

	En el apartado de sujetos, se presentan las características centrales de la población investigada (Ej. Número, características sociodemográficas, características psicológicas).						EP y EM	
	En el apartado de sujetos, se describe el proceso de selección de la muestra.						EP y EM	
	En el apartado de sujetos, se describe la forma para calcular el número de sujetos de la muestra.						EP y EM	
	En el apartado de sujetos, se presenta el número de participantes del estudio.						EP y EM	
	En el apartado de sujetos, se presenta el número de participantes en cada uno de los grupos (experimental, control).						EP y EM	
	En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que recibieron el tratamiento planeado.						EP y EM	
	En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que completaron el protocolo de intervención.						EP y EM	
	En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que se analizaron para obtener el resultado principal (más importante) del estudio.						EP y EM	
	En el apartado de sujetos, se presenta el número de personas en cada uno de los grupos (experimental, control) que después de la asignación al azar, abandonaron (no						EP y EM	

	continuaron) en el proceso investigativo.						
	En el apartado de sujetos, se presentan las razones de abandono (no continuación) del proceso investigativo, después de la asignación al azar, por parte de las personas en cada uno de los grupos (experimental, control).					EP y EM	
	En el apartado de sujetos, se presentan las características centrales de la muestra investigada (Ej. número, características socio demográficas, características psicológicas).					EP y EM	
	En el apartado de sujetos se hace explícito que la muestra es representativa de la población.					EP y EM	
	En el apartado de sujetos se presentan los criterios de inclusión y exclusión de los participantes.					EP y EM	
	En el apartado de sujetos, se presentan de manera explícita las retribuciones (en dinero o especie) recibidas por los sujetos, debido a su participación en el estudio.					EP y EM	
	En el apartado de sujetos, se presenta el diagrama de flujo de participantes.					EP y EM	
	En el apartado de sujetos, se presenta el número de pacientes tratados por cada terapeuta.					EP y EM	
	En el apartado de método, se presenta el número de pacientes tratados en cada centro o lugar donde se hace la intervención.					EP y EM	
	b) Equipo de investigación						

	Reclutadores						
	En el apartado de método, se identifica de manera explícita la persona que estuvo encargada de incorporar las personas al estudio.					EP y EM	
	Generadores de asignación al azar						
	En el apartado de método, se identifica de manera explícita la persona que generó la secuencia de asignación al azar.					EP y EM	
	En el apartado de método, se identifica de manera explícita la persona que asignó los participantes a las intervenciones o grupos.					EP y EM	
	Evaluadores						
	En el apartado de método, se presenta la formación y experiencia de los evaluadores (personas que hacen los registros y/o califican las pruebas).					EP y EM	
	Terapeutas						
	En el apartado de método, se presentan los criterios de selección, para las personas encargadas del proceso de intervención.					EP y EM	
	En el apartado de método, se presenta el número de personas encargadas del proceso de intervención en cada grupo.					EP y EM	
	En el apartado de método, se presenta la formación y experiencia de las personas encargadas del proceso de intervención.					EP y EM	
	c) Ambientes y ubicaciones						
	En el apartado de método, se presentan las características principales de los sitios o centros (ej. Hospital o centro médico de primer, segundo o					EP y EM	

	tercer nivel) donde se realizó la investigación.							
	En el apartado de método, se presentan los criterios de selección de los ambientes y ubicaciones donde se recolectaron los datos.						EP y EM	
	En el apartado de método, se presenta el número de centros o lugares en donde se hace la intervención para cada grupo (control, experimental)						EP y EM	
Instrumentos	En el apartado de instrumentos, se nombran y describen las pruebas y/o aparatos utilizados en el proceso investigativo.						EP y EM	
	En el apartado de instrumentos, se presentan las propiedades psicométricas (confiabilidad, validez) de las pruebas y/o aparatos utilizados en el proceso investigativo.						EP y EM	
Definición de variables	En el apartado de método, se hacen explícitas las variables dependientes e independientes.						EP y EM	
	En el apartado de método, se definen operacionalmente las variables y sus grados.						EP y EM	
	En el apartado de método, las variables se definen claramente.						EP y EM	
Procedimiento	En el apartado de método, se presenta el procedimiento de la investigación						EP y EM	
	El procedimiento se describe de forma suficiente, para que pueda replicarse la investigación.						EP y EM	
	En el apartado de método, se presenta el número de registro y el nombre de registro del experimento.						EP y EM	

	En el apartado de procedimiento se presentan los periodos o momentos en los cuales fueron incorporados los sujetos a la investigación.						EP y EM	
	En el procedimiento se hace explícito qué participantes se asignaron de manera aleatoria al grupo control y al experimental.						EP y EM	
	En el procedimiento, se especifica como se hizo el proceso de asignación a los grupos (experimental y control).						EP y EM	
	En el procedimiento, se describe la forma en que se ocultó la secuencia (programa) de asignación al azar, hasta el momento en que se asignan las intervenciones a cada uno de los sujetos.						EP y EM	
	En el procedimiento, se hace evidente que los participantes fueron enmascarados (cegados) a la asignación de las intervenciones.						EP y EM	
	En el procedimiento, se describe como los participantes fueron enmascarados (cegados) a la asignación de las intervenciones.						EP y EM	
	En el procedimiento se hace explícito que los terapeutas fueron enmascarados o cegados a la asignación de las intervenciones.						EP y EM	
	En el procedimiento, se explica cómo los terapeutas fueron enmascarados (cegados) a la asignación de las intervenciones.						EP y EM	
	En el procedimiento se hace explícito que los evaluadores fueron enmascarados o cegados						EP y EM	

	a la asignación de las intervenciones.							
	En el procedimiento se describe como los evaluadores fueron enmascarados o cegados a la asignación de las intervenciones.						EP y EM	
	En el procedimiento, se presenta la forma en que fueron asignados los terapeutas al grupo control y experimental.						EP y EM	
	En el procedimiento se especifica cómo se hizo el control de variables extrañas.						EP y EM	
	En el procedimiento se evidencia que el control de variables extrañas es riguroso.						EM	
	En el procedimiento se describe la forma en que se llevó a cabo o se desarrolló la evaluación.						EP y EM	
	En el procedimiento se describe en qué momento o tiempo se implementó la evaluación.						EP y EM	
	En el procedimiento, la evaluación se describe de forma suficiente, para que pueda replicarse la investigación.						EP y EM	
	En el procedimiento, la intervención se describe de forma suficiente, para que pueda replicarse la investigación.						EP y EM	
	En el procedimiento, se describe en qué momento o tiempo se implementó la intervención.						EP y EM	
	En el procedimiento, se hace explícita la forma en que se estandarizaron las intervenciones.						EP y EM	
	En el procedimiento, se hace explícita la forma como se controló la adherencia al						EP y EM	

	protocolo, por parte de los terapeutas.							
	En el procedimiento, se hacen explícitas las circunstancias o aspectos de la investigación que pueden afectar la validez de los resultados.						EP y EM	
Procesamiento estadístico.	En el apartado de método, se nombra el software estadístico utilizado.						EP y EM	
	En el apartado de método, se presentan las mediciones o medidas de los resultados principales y complementarios.						EP y EM	
	En el apartado de método, se presentan los estadísticos utilizados para comparar los grupos en los resultados principales y complementarios.						EP y EM	
	En el apartado de método, se describe el procesamiento estadístico de forma suficiente, para que pueda replicarse la investigación.						EP y EM	
	En el apartado de método, se informa sobre circunstancias del análisis de datos que pueden afectar la validez de los resultados.						EP y EM	
Consideraciones éticas	En el apartado de método, se hace explícito que la investigación se está guiando por un referente ético formal (ej. Código de ética de la Asociación Psicológica Americana).						EP y EM	
	En el apartado de método, se hace explícito que se utilizó el consentimiento informado de los participantes.						EP y EM	
	En el apartado de método, se describe la forma en que se obtuvo el consentimiento informado.						EP y EM	

	En el apartado de método, se hace evidente que el grupo control fue intervenido posterior a la finalización de la investigación.						EP y EM	
	En el apartado de método, se nombran las retribuciones dadas a los participantes, en dinero o especie.						EP y EM	
	En el apartado de método, se hace explícito que la investigación fue aprobada por un comité de ética institucional.						EP y EM	
	En el apartado de método, se hace explícito que las normas éticas se cumplieron.						EP y EM	
	En el apartado de método, se describe cómo se hizo el monitoreo de cumplimiento de las normas éticas.						EP y EM	
	En el apartado de método, se describe la forma en que se garantiza la confidencialidad de los datos en la publicación.						EP y EM	
	En el apartado de método, se presentan de manera explícita los conflictos de interés en el estudio.						EP y EM	
	En el apartado de método, se describe como se manejaron los conflictos de interés.						EP y EM	
	En el apartado de método, se evidencia que los datos empíricos y/o materiales están disponibles para que otro investigador esté en capacidad de utilizarlos y reproducir un análisis que conduzca a las Conclusiones del autor.						EP y EM	

Categoría
Resultados: presentación de datos cuantitativos y/o cualitativos arrojados por la investigación.
Subcategorías
General: Consideraciones genéricas a ser tenidas en cuenta para la presentación de los datos.
Texto: consideraciones a ser tenidas en cuenta en la descripción de los datos.
Tablas: Consideraciones a ser tenidas en cuenta en la presentación de los datos en formato de Tabla.
Gráficas: Consideraciones a ser tenidas en cuenta en la presentación de los datos en formato de imágenes o gráficos.

Subcategoría	Indicadores/ítem	Valoración	Observaciones	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
General	En los resultados, los datos se presentan de forma organizada (ordenada).						EP	
	En los resultados, se presenta la frecuencia y/o porcentaje de datos faltantes (valores perdidos)						EP y EM	
	En los resultados, se presenta el número de casos que se eliminan (no se tiene en cuenta) para cada uno de los análisis realizados.						EP y EM	
	En los resultados, se reporta el tamaño de muestra por celda o subgrupo evaluado.						EP y EM	
	En los resultados, se presentan los resultados principales y complementarios para cada grupo.						EP	
	En los resultados, se presentan los tamaños del efecto y los intervalos de confianza.						EP y EM	
	En los resultados, se presenta la tasa de error tipo I aceptada previamente.						EP y EM	
	En los resultados, se presenta el valor de error estadísticos (ej. $p < 0,05$).						EP y EM	
	En los resultados, se reporta la dirección de "p" para el análisis estadístico realizado.						EP y EM	

	En los resultados, se reportan los grados de libertad para el análisis estadístico realizado.						EP EM	y	
	En los resultados, los procedimientos estadísticos se aplican correctamente para el nivel de medición de los datos.						EM		
	En los resultados, los datos están dirigidos a evaluar las hipótesis planteadas.						EM		
	En los resultados, los datos se presentan con precisión (no existe ambigüedad)						EP EM	y	
	En los resultados, se hace explícito que el grupo control y experimental son similares en las variables relevantes, antes de implementarse la intervención.						EP EM	y	
Texto	En los resultados, el texto es claro.						EP		
	En los resultados, se justifican de forma empírica y/o teórica las causas para que existan "valores perdidos" en la investigación.						EP EM	y	
	En los resultados, la información en texto es complementaria a la de las Tablas (no existe información redundante)						EP		
Tablas	En los resultados, se presenta una Tabla con la línea de base de variables demográficas y clínicas de cada grupo.						EP EM	y	
	En los resultados, las Tablas son claras (fáciles de entender por un experto).						EP		
Gráficas	En los resultados, las gráficas son claras (fáciles de entender por un experto).						EP		
	En los resultados, las gráficas son pertinentes (adecuadas) para la información que se presenta.						EP		

Categoría
Discusión : Análisis de los resultados obtenidos en el contexto de la evidencia existente
Subcategorías
Análisis: debate entre los resultados de la investigación, el conocimiento en el área de estudio y las características propias de la investigación.
Generalización: posibilidad de que los resultados, sean aplicables a otros sujetos, condiciones y/o metodologías similares a los de la investigación realizada.
Implicaciones: aspectos que se derivan del estudio y pueden guiar la investigación y/o las prácticas futuras.

Sub categoría	Indicadores/ítem	Valoración	Observaciones	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
Análisis	En la discusión, el análisis de los datos se hace en función (está relacionado) con los objetivos y/o hipótesis del estudio.						EP y EM	
	En la discusión, el análisis de los datos se hace en función (está relacionado) con el problema de investigación del estudio.						EP y EM	
	En la discusión, se comparan los resultados de la investigación con otros estudios.						EP y EM	
	En la discusión, se presentan las limitaciones del estudio que afectan la validez interna.						EP y EM	
	En la discusión, se retoman los resultados, pero no se repiten.						EP	
	En la discusión se evidencia coherencia con los resultados.						EP y EM	
	En la discusión, se presentan los efectos secundarios negativos de la(s) intervención(es).						EP y EM	
Generalización	En la discusión, las generalizaciones están basadas en los resultados.						EP y EM	
	En la discusión, se especifica la población a la cual puede ser generalizado el estudio.						EP y EM	
	En la discusión, se especifican los contextos a los cuales puede ser generalizado el estudio.						EP y EM	

	En la discusión, se especifican las problemáticas a las cuales puede ser generalizado el estudio.						EP y EM	
	En la discusión, se especifican los métodos de recolección de datos a los cuales puede ser generalizado el estudio						EP y EM	
	En la discusión, se especifican las intervenciones a las cuales puede ser generalizado el estudio.						EP y EM	
	En la discusión, se especifican las características de los terapeutas a los cuales puede ser generalizado el estudio.						EP y EM	
	En la discusión, se justifica la razón por las cuales se hacen las generalizaciones propuestas.						EP y EM	
Implicaciones	En la discusión, se presentan implicaciones para la investigación futura, programas y/o políticas.						EP y EM	

Categoría

Conclusiones: Principales deducciones que se generan posterior al análisis y discusión de los resultados.

Subcategoría	Indicadores/ítem	<i>Valoración</i>	<i>Observaciones</i>	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
	Las Conclusiones están fundamentadas en los resultados.						EP y EM	
	Las Conclusiones se hacen en función (están relacionadas) con los objetivos de investigación.						EP y EM	

Categoría

Referencias: listado ordenado y detallado en base a una metodología específica, de las fuentes de donde se obtuvo la información que sustenta la investigación.

Subcategoría	Indicadores/ítem	Valoración	Observaciones	Pertinencia	Redacción	Observaciones	Clasificación del estándar	Observaciones
	Las referencias son acordes con el título y/o temática central del artículo.						EP y EM	
	Las referencias son actualizadas (más del 50% hace referencia a artículos con fechas comprendidas en los cinco años previos a la publicación del estudio).						EP y EM	
	El número de referencias es superior a 25.						EP y EM	
	El tipo de referencias es adecuado (más del 50% son artículos originales).						EP y EM	
	Las referencias poseen los elementos requeridos para las normas utilizadas (ej. APA).						EP	
	Las referencias son confiables (predominan los artículos de revistas especializadas en el área de conocimiento del artículo).						EP y EM	

