DOI: https://doi.org/10.11144/Javeriana.umed59-3.icrc

Investigación cualitativa reproducible: un camino que debemos empezar a recorrer

Qualitative reproducible research: A path we must begin to walk

Recepción: 25/10/2017 | Aceptación: 14/12/2017

Óscar Mauricio Muñoz Velandia[a]
Pontificia Universidad Javeriana, Colombia

RESUMEN

El movimiento conocido como investigación reproducible (IR) ha venido ganando espacio dentro de la investigación cuantitativa como una forma de aumentar la confianza en sus resultados. El movimiento propone que los datos primarios y la codificación usada en el análisis de una investigación estén disponibles para que otros investigadores puedan verificar los resultados publicados y conducir análisis alternativos. Las ventajas asociadas a la IR aplican muy bien dentro del trabajo de los investigadores cualitativos, quienes no solo gozarían de mayor confianza y aceptación de los resultados de su investigación, sino que podrían beneficiarse al permitir que las personas interesadas entiendan sus hallazgos e, incluso, al favorecer la formación de redes de investigación y la construcción de conocimiento de forma colaborativa. Las limitaciones más grandes para que se pueda realizar investigación cualitativa (IC) reproducible están dadas por las restricciones éticas intrínsecas a hacer disponible de forma abierta los datos primarios y la codificación de los estudios. En el presente artículo se presentan algunas propuestas para superar estas dificultades, en lo referente al consentimiento informado y medidas para garantizar la confidencialidad de los datos, reconociendo que aún falta mucho camino superar estas dificultades. Así mismo, se presentan algunos pasos que ya se han avanzado para permitir que se desarrolle IC reproducible incluyendo el desarrollo de programas informáticos para el manejo de datos cualitativos, la disponibilidad de repositorios para hacer disponibles esos datos y los avances que se han desarrollado en la implementación de IR cuantitativa.

Palabras clave

investigación cualitativa; investigación reproducible; consentimiento informado; ética de la investigación.

ABSTRACT

The movement known as reproducible research (RR) has won space in the quantitative research field as a way to increase the confidence in their results. The RR movement proposes that primary data and computer codes used in the research must be publically available to let other researchers verify published results and to conduct alternative analysis. The advantages related to RR apply well to the qualitative researcher's work, and will increase the confidence and approval of their results, and additionally could provide advantages like the deeper understanding of their findings, the chance to create investigation networks and groups for collaborative knowledge building. The biggest limitations to produce qualitative reproducible research are ethical considerations associated with making sensitive information openly available. The present article

[a] Correspondencia: o.munoz@javeriana.edu.co

[a] Médico Internista. Magister en Epidemiología Clínica. PhD (c) en Epidemiología Clínica. Profesor asistente, Departamento de Medicina Interna, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, Colombia.

Cómo citar: Muñoz Velandia OM. Investigación cualitativa reproducible: un camino que debemos empezar a recorrer. Univ Med. 2018;59(3). 1-7. doi: http://dx.doi.org/10.11144/Javeriana.umed59-3.icrc

Conflictos de interés: no

Fuentes de financiación: ninguna

presents some proposals to overcome these difficulties associated with the process of obtaining informed consent and measures to guarantee the confidentiality of data, recognizing that there is much work to be done in this area. Additionally, some steps have been taken to facilitate reproducible qualitative research, including the development of software packages to manage qualitative data, the availability of repositories to make data available and advances developed in reproducible quantitative research implementation.

Keywords
qualitative research; reproducible research; informed consent; research ethics.

Introducción

En los últimos años ha venido creciendo un movimiento que busca que la investigación cuantitativa publicada en las revistas científicas sea reproducible por otros expertos, a fin de aumentar la confianza que se tiene en los resultados de la investigación. Para ello se requiere que, como mínimo, las bases de datos y la codificación usada en el *software* de análisis estadístico estén disponibles para que otros investigadores puedan verificar los resultados publicados y conducir análisis alternativos (1). A este movimiento se le conoce como investigación reproducible (IR).

La idea surgió a partir de demostrarse que se ha publicado investigación fraudulenta en revistas de alto impacto científico. Tal vez el caso más reconocido fue la publicación en la revista *Science* del estudio "Patient-Specific Embryonic Stem Cells Derived from Human SCNT Blastocysts" (2), del grupo de trabajo dirigido por Woo Suk Hwang, que debió ser retirado de la revista una vez que el comité de investigaciones de la Universidad Nacional de Seúl encontró, en un análisis interino, que los datos sometidos a la revista "fueron basados en una manipulación deliberada, o en otras palabras la fabricación de resultados de investigación" (3). Este caso, y otros, impulsaron entre un amplio grupo de editores de revistas científicas la idea de promover la reproducibilidad de la investigación, bajo la premisa de que solo se puede llegar a la verdad bajo la verificación independiente de las nuevas observaciones (4). La idea de realizar investigación que sea reproducible tiene cabida incluso cuando los autores no tienen abiertamente una intención deliberada de fraude, si se tiene en consideración que los análisis modernos han llegado a ser cada vez más complicados y la posibilidad de errores inadvertidos que lleven a conclusiones equivocadas es cada vez más grande.

Los defensores de realizar IR cuantitativa mencionan una serie de ventajas de seguir estos lineamientos (5), que incluyen:

* Mayor transparencia acerca de la calidad y la integridad de la investigación.
* Mayor diseminación y entendimiento de los resultados.
* Disponibilidad de los datos completos para que se puedan realizar metanálisis con otros estudios.
* Implementación más fácil de los hallazgos y recomendaciones provenientes de la investigación.
* El progreso más rápido de la ciencia.
* Un mejor aprovechamiento del dinero invertido en la investigación, especialmente si se considera que el dinero proveniente de fondos públicos sería capaz de producir datos aprovechables por múltiples investigadores.

¿Pueden las ventajas propuestas de la investigación reproducible ser aplicables a la investigación cualitativa?

Hasta el momento se ha reconocido la importancia de mantener un rigor suficiente y unos criterios mínimos de calidad a la hora de conducir, evaluar y publicar una investigación cualitativa (IC), si bien se reconoce que mantener, evaluar y demostrar ese rigor es más difícil para la IC que para la investigación cuantitativa. Dentro de los criterios a la hora de determinar la calidad una IC se propone una evaluación de la validez y la confiabilidad de la investigación. Por validez se entiende la medida en que los resultados son una representación adecuada del fenómeno que se pretende medir; mientras que nos referimos a confiabilidad como la reproducibilidad y estabilidad de los datos (6). Así, demostrar que los datos recogidos y analizados dentro de una IC son reproducibles es

clave para demostrar la confiabilidad y, por tanto, la calidad de una IC.

Las motivaciones que generaron la corriente de IR también aplican a la IC, en la cual —similar a lo visto para la investigación cuantitativa— llegan a publicarse estudios de baja calidad, con datos primarios incompletos o con una codificación y análisis inadecuados o susceptibles de error. Incluso también podría haber manipulación deliberada de los datos o fraude.

La idea de la IR ha empezado a tener cabida entre los investigadores biomédicos que realizan investigación cuantitativa (7), e incluso entre aquellos que proponen métodos mixtos o múltiples (es decir, aquellos que utilizan simultáneamente métodos de investigación cuantitativa y cualitativa) (8); sin embargo, la aplicación de los lineamientos de IR, entendida como el hecho de hacer disponibles los datos primarios y la codificación usada en el análisis para que otros investigadores puedan verificar los resultados publicados, no ha tenido aún amplia acogida dentro de aquellos que trabajan con IC. Para algunos investigadores cualitativos, incluso, no debería ser un tema para tratar, bajo la premisa de que la IC es subjetiva, y por tanto se asume que no existe una sola verdad, sino verdades múltiples e individuales para cada individuo o para cada comunidad, que no son y no deben ser reproducibles a otros individuos o comunidades. Estos investigadores probablemente preferirían hacer una validación de su trabajo de forma individual y cerrada, y presentar solo los resultados de su investigación dentro de círculos específicos.

A pesar de ello, muchos investigadores cualitativos tienen una opinión diferente, y asumen que el resultado de la IC puede y debe ser extrapolable, en alguna medida, a poblaciones diferentes a aquellas en las que se desarrolló (9). Esto, siempre y cuando se describa suficientemente el contexto de la investigación, las características de los participantes en el estudio, la índole de sus interacciones con el investigador y el entorno físico para que otros puedan decidir cuán transferibles a otros contextos son los resultados. Para estos investigadores, la idea de hacer su investigación más transparente puede tener sentido y estarán más abiertos a adoptar conductas que favorezcan la IR, como serían el acceso a la transcripción de sus datos o su codificación a solicitud de cualquiera que esté interesado en ellas.

Asumir esta postura podría traer para ellos múltiples ventajas. La primera, y tal vez la más importante, es una mayor confianza en los resultados de la investigación. Ejemplos de ello partiendo de diferentes posturas paradigmáticas se presentan a continuación:

* Los investigadores que se basan en la teoría crítica (10) parten de la idea de generar conocimiento que pueda identificar y demostrar las estructuras de opresión existentes y, a partir de dicho conocimiento, empoderar a las comunidades para generar un cambio. Los resultados de investigadores de esta vertiente pueden, sin embargo, sufrir de menor credibilidad si se asume de entrada que están sesgados hacia demostrar dichas disparidades. Cabe aclarar que para los investigadores que realizan IC crítica no se considera que sus resultados son sesgados, sino más bien enfocados; pero que esta misma teoría crítica, basada en la historiografía, reconoce la importancia de medir adecuadamente, e incluso cuantificar estas disparidades. Para los investigadores cualitativos críticos hacer disponibles sus datos, de tal forma que puedan ser evaluados y reproducidos por otros, podría fortalecer sus conclusiones y facilitar el cambio que están buscando, al servir como punto de partida para el desarrollo de estrategias que respalden a las personas más vulnerables, dando voz a sus experiencias y perspectivas, y haciendo visibles las fuentes de opresión y corrupción en la sociedad.
* Los investigadores que se basan en paradigmas constructivistas (10) estarán de acuerdo en que la realidad se construye por medio de la interacción y el diálogo entre diferentes actores, y aceptarán, por tanto, que puede haber diferentes interpretaciones de esa realidad. Ofrecer los resultados de investigaciones basadas en epistemologías constructivistas (por ejemplo, transcripciones o códigos) a una comunidad más abierta de investigadores u otros usuarios

de esa información será de gran ayuda si se desea un entendimiento más completo del problema, en especial si se pretende que dichas interpretaciones puedan ser extrapoladas, al menos en cierta medida, a otras comunidades semejantes.

Otras ventajas potenciales de realizar IC reproducible incluyen un mejor entendimiento de los resultados de la investigación, al permitir a los usuarios que quieran reinterpretar la información e implementar las medidas que se desprendan de su análisis lo hagan de una forma informada, consciente y más completa. También es posible que sea más fácil la consecución de recursos, bajo la premisa de que la información abierta y ampliamente disponible producirá más y mejores resultados, además de beneficiar a más personas y comunidades.

Por último, hacer reproducible la IC facilitaría la consecución de colaboradores, promovería la retroalimentación de otras personas interesadas en el tema e invitaría a nuevos investigadores a trabajar en el problema.

Limitaciones de la investigación reproducible

Existen también potenciales limitaciones a la hora de realizar IC reproducible. El solo hecho de poder reproducir el análisis de los datos o, incluso, de repetir la misma investigación obteniendo resultados similares no corrige los problemas estructurales de la investigación. Si en la investigación no se tuvieron en cuenta posibles variables de confusión, si el diseño del estudio es pobre o si existe una gran pérdida de información (por presentar solo algunos ejemplos aplicables a la investigación cuantitativa), no habrá forma de solucionarlos solo por tener disponibles los datos o los códigos que permitan reproducir el análisis (11). De igual forma, si se está conduciendo IC, el hecho de hacer disponibles los datos obtenidos (entrevistas, documentos, fotografías, grabaciones de audio, etc.) o los resultados de los procesos de análisis y reducción de la información (fundamentalmente la codificación) no soluciona los potenciales problemas estructurales de la investigación cualitativa: si la metodología seleccionada era inadecuada para responder las preguntas planteadas, si no se obtuvo una muestra suficiente y representativa de entrevistas, si en dichas entrevistas no se abordaron los temas con profundidad, si la información obtenida no era confiable o si la codificación pasó por alto información relevante para cumplir con los objetivos del estudio. El hecho de hacer disponibles los datos primarios o los resultados de la codificación de esta información no solucionará dichos problemas.

Una segunda limitación, y probablemente la más importante, a la hora de aplicar los principios de la IR a la IC proviene, sin embargo, de consideraciones éticas para tener en cuenta en el momento de hacer públicos los datos. Dos consideraciones resultan centrales en esta discusión: la necesidad de consentimiento informado y la confidencialidad de los datos (12).

Siempre que se desee hacer pública la información, será necesario un proceso de consentimiento informado en el cual se explicite la intención de hacer disponibles los datos primarios de una forma abierta y fácil. Sin embargo, las condiciones que debe cumplir dicho consentimiento informado son aún un tema que amerita discusión. Se ha planteado que la IC requiere ajustes a dicho proceso para hacerlo acorde con las características propias de este tipo de investigación (13). Primero, la IC es realizada dentro de un contexto temporal determinado, en un periodo específico de la vida de una persona, con características que pueden cambiar de forma súbita. Segundo, los participantes pueden tener dificultades para entender y anticipar las consecuencias de un determinado estudio cuando estas se explican únicamente al inicio de la investigación. Será necesario realizar un proceso dinámico de consentimiento informado que tenga en cuenta las características cambiantes de la investigación a través del tiempo y que actualice la voluntad expresada por los pacientes (13). Más aún, es posible que un participante conozca totalmente las implicaciones de un estudio y autorice compartir la información en un determinado momento del

tiempo; pero decida cambiar de opinión tiempo después de haber terminado la investigación, en la medida en que cambie su realidad. El proceso de consentimiento informado deberá ser capaz de adaptarse a esta nueva demanda.

Otras consideraciones deben plantearse dentro del proceso de consentimiento informado para la publicación, a nivel individual, de los datos producto de la IC. ¿Qué hacer cuando la persona que da el consentimiento no es directamente el sujeto que entró en la investigación, sino su representante (por ejemplo, en el caso de menores de edad o participantes con alteraciones en su salud mental)? ¿Qué conducta seguir cuando existen discrepancias entre la voluntad expresada por el participante y la opinión de los investigadores u otro actor (por ejemplo, personal de salud) con respecto a si es o no conveniente la publicación de esta información? Las respuestas a estas preguntas están por ser definidas y ameritarán discusiones bioéticas profundas.

La confidencialidad de los datos es también fundamental y tiene algunas particularidades aplicables a la IC. En general, los investigadores cualitativos se caracterizan por hacer hincapié en temas sensibles, por lo que las características particulares de los participantes son importantes si se pretende hacer disponible, por cualquier vía, los datos obtenidos. Hacer públicos los datos no será posible si los participantes en la investigación son susceptibles a consecuencias negativas asociadas a ello. Vale la pena pensar que con frecuencia se tratan temas muy personales, traumáticos (por ejemplo, abuso). Hacer públicos esos datos puede ser muy invasivo emocionalmente y generar vergüenza o revictimización. De igual forma, hacer pública información obtenida para una IC puede generar consecuencias negativas desde el punto de vista político o social, haciendo a los participantes vulnerables a consecuencias políticas (por ejemplo, marginalización) o sociales, como estigmatización y rechazo (14).

Aun en los casos en los que sea posible hacer pública la información, deberán hacerse todos los esfuerzos para impedir la identificación individual del participante en la investigación, proceso conocido como *anonimización*. El riesgo de no realizarla es alto cuando existe una cantidad pequeña de participantes o cuando ellos experimenten una condición poco frecuente que los hagan distinguibles del resto de miembros de su comunidad. Existen herramientas informáticas que facilitan llevar a cabo este proceso que ya vienen utilizándose en las bases de datos publicadas de estudios cuantitativos. En ciertas ocasiones, los participantes no tendrán reparos en permitir que otros los identifiquen; por el contrario, pueden desear que eso ocurra, lo que debe quedar claramente expresado en el documento de consentimiento informado.

Facilitadores para la llegada de la reproducibilidad a la investigación cualitativa

A pesar de las limitaciones presentadas, existen algunos pasos que facilitarían la realización de IR dentro de los estudios cualitativos.

Diferentes *softwares* están disponibles en la actualidad que permiten guardar de forma organizada tanto los datos primarios como la codificación asignada a ellos, lo que facilita compartir la información tanto entre los diferentes autores de una investigación como con investigadores externos, tal como se necesita en IR. Dentro de los programas desarrollados para estudios cualitativos se destacan el *software* Nvivo, el programa QSR N6 e, incluso, algunos *softwares* gratuitos como el desarrollado por los Centros de Control y Prevención de la Enfermedad de los Estados Unidos (CDC EZ test), disponible a través de su página electrónica (https://www.cdc.gov/hiv/library/software/index.html).

El uso de diferentes paquetes informáticos de forma simultánea puede, sin embargo, dificultar la reproducción del análisis de los estudios. Se deberán, entonces, diseñar algunos estándares mínimos de obligatorio cumplimiento si se quiere publicar IC reproducible. Estos estándares deberán definir de forma clara qué datos reportar, bajo qué formato, con cuáles métodos de

anonimización y dónde quedarán depositados, especificando algunas medidas para garantizar la seguridad de esta información. Algunas lecciones se podrían aprender de revistas que publican investigación cuantitativa reproducible, que ya han estandarizado algunas reglas para facilitar este proceso (1,15).

Algunos sitios web se han desarrollado para guardar los datos asociados a un determinado estudio o publicación. Dichos repositorios podrían ser también de utilidad para hacer disponibles para otros investigadores o a la comunidad en general los datos primarios y la codificación correspondientes a investigación cualitativa. Un ejemplo de ellos es el proyecto Dataverse, disponible a través de su página electrónica (https://dataverse.org).

El proceso de reproducir IC puede hacerse siguiendo pequeños pasos. Ante las dificultades éticas planteadas para compartir datos a nivel individual, es posible empezar por compartir únicamente los códigos y los datos primarios que se relacionen exclusivamente con ellos, sin hacer disponible la totalidad de las entrevistas o documentos primarios. Solo para casos muy seleccionados se dejaría disponible la totalidad de la información. El campo de IC reproducible solo progresará en la medida en que entendamos sus ventajas y estemos dispuestos a trabajar para lograr sus beneficios.

Conclusiones

El concepto de IR ha venido ganando espacio dentro de las ciencias biomédicas. Las ventajas asociadas aplican muy bien dentro del trabajo de los investigadores cualitativos, siendo las más importantes una mayor confianza en los resultados de la investigación, además de favorecer la formación de redes y la construcción de conocimiento de forma colaborativa. Queda aún mucho camino por recorrer para que pueda aplicarse de forma generalizada. Las limitaciones más grandes están dadas por las restricciones éticas que implican hacer disponibles de forma abierta los datos primarios y la codificación de los estudios. El consentimiento informado y las medidas para garantizar la confidencialidad de los datos son temas en los que debemos trabajar para superar estas dificultades.

Referencias

1. Peng RD. Reproducible research and Biostatistics. Biostatistics. 2009;10(3):405-8.

2. Hwang WS, Roh S Il, Lee BC, Kang SK, Kwon DK, Kim S, et al. Patient-specific embryonic stem cells derived from human SCNT blastocysts. Science. 2005;308(5729):1777-83.

3. Kennedy D. Editorial expression of concern. Science. 2006;311(5757):36.

4. Laine C, Goodman SN, Griswold ME, Sox HC. Reproducible research: Moving toward research the public can really trust. Ann Intern Med. 2007;146(6):450-3.

5. Groves T, Godlee F. Open science and reproducible research. BMJ. 2012;344(June):e4383.

6. Anderson C. Presenting and evaluating qualitative research. Am J Pharm Educ. 2010;74(8):1-7.

7. Iqbal SA, Wallach JD, Khoury MJ, Schully SD, Ioannidis JPA. Reproducible research practices and transparency across the biomedical literature. PLoS Biol. 2016;14(1):1-13.

8. Dunning T, Rosenblatt F. Transparency and reproducibility in multi-method research. Rev Cienc Política. 2016;36(3):773-83.

9. Ulin P, Robinson E, Tolley E. El análisis de los datos cualitativos. En: Investigación aplicada en salud pública: Métodos cualitativos. Washington: Organización Panamericana de la Salud; 2006. p. 137-70.

10. Guba EG, Lincoln YS. Paradigmatic controversies, contradictions, and

emerging confluences. En: Denzin NK, Lincoln YS, editores. The Sage handbook of qualitative research. 3rd. ed. Thousand Oaks: Sage; 2005. p. 191-215.

11. Leek JT, Peng RD. Opinion: Reproducible research can still be wrong: Adopting a prevention approach. Proc Natl Acad Sci. 2015;112(6):1645-6.

12. Williams G, Pigeot I. Consent and confidentiality in the light of recent demands for data sharing. Biometrical J. 2017;59(2):240-50.

13. Oye C, Sorensen NO, Glasdam S. Qualitative research ethics on the spot: Not only on the desktop. Nurs Ethics. 2016;23(4):455-64.

14. Peter E. The ethics in qualitative health research: special considerations. Cien Saude Colet. 2015;20(9):2625-30.

15. Hofner B, Schmid M, Edler L. Reproducible research in statistics: A review and guidelines for the Biometrical Journal. Biometrical J. 2016;58(2):416-27.