

Introducción a los ensayos clínicos por conglomerados

Emiliano Rossi 

Hospital Italiano de Buenos Aires.
Ciudad Autónoma de Buenos Aires. Argentina.

Acta Gastroenterol Latinoam 2026;56(3):261-264

Recibido: 29/06/2026 / Aceptado: 10/08/2026 / Publicado online: 30/09/2026 / <https://doi.org/10.52787/agl.v56i3.668>

Como profesionales de la salud con frecuencia leemos estudios que evalúan nuevas intervenciones. Tradicionalmente, la mayoría de estos estudios son ensayos clínicos controlados aleatorizados, donde los pacientes individuales se asignan al azar a un nuevo tratamiento. Sin embargo, también podemos encontrarnos con un diseño metodológico distinto: el ensayo controlado aleatorizado por conglomerados (cRCT, por su sigla en inglés).

Para poder realizar una lectura crítica de estos estudios, es necesario que comprendamos qué es un cRCT, por qué se utiliza y si se aplicó la metodología adecuada. El objetivo de este artículo es proporcionar las herramientas para evaluar la calidad y validez de estos ensayos.

¿Qué es un conglomerado o clúster?

Los datos agrupados o clústers son una estructura jerárquica o multinivel donde los participantes tienen una

agrupación impuesta por el entorno. Los datos pueden estructurarse en dos niveles (ejemplo: pacientes por médicos tratantes) o en tres niveles (ejemplo: pacientes por médicos tratantes y médicos tratantes por hospitales). Los individuos de un mismo conglomerado comparten características no medidas del mismo (como los protocolos de atención) o características poblacionales (como la gravedad promedio de la patología atendida en un centro de alta complejidad).¹

Las observaciones de los individuos dentro de un mismo clúster tienden a estar correlacionadas, es decir, no son independientes.² Incluso después de ajustar por las características medidas, los resultados de dos pacientes tratados en un mismo centro de endoscopia, por ejemplo, tenderán a ser más similares entre sí que los resultados de dos pacientes tratados en centros diferentes. Esto puede deberse a características no medidas, como la habilidad de los endoscopistas del centro, los protocolos de sedación o las características socioeconómicas de la población que acude al mismo.¹

El coeficiente de correlación intraclass (ICC, por su sigla en inglés) mide la proporción de la varianza total de los resultados que puede atribuirse a las diferencias entre los conglomerados. Sus valores oscilan entre 0 (no hay similitud o efecto de grupo) y 1 (todos los individuos en el conglomerado tienen el mismo resultado). Reportar el ICC permite cuantificar el efecto del agrupamiento.¹

¿Por qué un cRCT en lugar de un ensayo clínico individual?

Si los ensayos clínicos individuales son el *gold stan-*

Correspondencia: Emiliano Rossi
Correo electrónico: emiliano.rossi@hospitalitaliano.org.ar

dard, ¿por qué un investigador decidiría aleatorizar conglomerados enteros en lugar de pacientes individuales? Algunos motivos que lo justifican son:³

1. Evitar la contaminación del tratamiento: esto se da cuando los sujetos del grupo control están también expuestos a la intervención en estudio. Imaginemos que se diseña un estudio para evaluar un nuevo protocolo de asesoramiento dietético para pacientes con síndrome de intestino irritable. Si se aleatoriza a los pacientes de forma individual, dentro de una misma clínica los médicos tratantes tendrían que aplicar el nuevo protocolo a los pacientes del grupo de intervención y la atención estándar a los del grupo control. En la práctica sería dificultoso que los médicos de esta clínica no influencien su conducta por el nuevo enfoque educativo y es probable que terminen aplicando parte del nuevo asesoramiento a los pacientes del grupo control. En cambio, al aleatorizar clínicas enteras (conglomerados) a realizar el nuevo protocolo o la atención habitual se evita este riesgo de contaminación.

2. Naturaleza de la intervención: la aleatorización por conglomerados es el único método factible en ciertos entornos.² Muchas intervenciones operan a nivel grupal, manipulando el entorno físico o social y, por lo tanto, no pueden ser administradas de forma aislada a los individuos. Por ejemplo, implementar un nuevo sistema de historia clínica electrónica para alertas de detección de cáncer colorrectal afectará a todo el hospital por igual.

3. Factibilidad logística: en algunos casos puede resultar más económico y conveniente implementar una intervención por conglomerados, como cuando se utilizan equipos costosos que requieren alta capacitación para su utilización.

¿Cuál es la consecuencia de ignorar la existencia de clústers?

Analizar un cRCT comparando a todos los individuos tratados contra todos los controles como si no estuvieran agrupados es un error. Ignorar el efecto del ICC subestima la varianza y sobreestima el significado de las diferencias. Es de destacar que aún valores de ICC pequeños son suficientes para dar intervalos de confianza falsamente angostos y valores de p artificialmente bajos. Por ejemplo, un estudio de simulación que utilizó datos ya publicados, demostró que un ICC de 0,02 aumentó la probabilidad de cometer un error de tipo I al 9% (en lugar del 5% asumido por los autores).^{1, 4} Es así que un investigador

podría concluir que una intervención es estadísticamente significativa y clínicamente relevante, cuando en realidad el resultado fue producto del azar y de la aplicación de un método estadístico inapropiado que no tuvo en cuenta el ajuste por clústers.

¿Qué considerar para el cálculo de tamaño muestral de un cRCT?

Debido a que las observaciones de un mismo conglomerado no son independientes, un cRCT es estadísticamente menos eficiente que un ensayo clínico convencional. Lo que implica que para alcanzar el mismo poder estadístico, se necesita reclutar a un mayor número de participantes. Si m es el tamaño del clúster (asumiendo igual número de individuos para todos) y ρ (rho) es el ICC, entonces el factor de inflación o efecto de diseño asociado a la aleatorización por clústers será $1 + (m-1) \rho$. Al evaluar el tamaño muestral hay que tener en cuenta que agregar más participantes dentro de un mismo clúster tiene un límite de utilidad. Por lo tanto, para mejorar el poder del estudio, siempre es más eficiente aumentar el número de clústers que aumentar el número de individuos por clúster.³⁻⁵

¿Cuál es la unidad de aleatorización en un cRCT?

Para que funcionen como unidades válidas de asignación al azar, los clústers deben tener características y límites identificables, ya sean estructurales, jurisdiccionales o geográficos.⁴

¿Cómo debe analizarse correctamente un cRCT?

Existen procedimientos estadísticos que tienen en cuenta el agrupamiento y evitan resultados erróneos. Cuando vemos la sección métodos de un artículo, debemos buscar qué enfoque se utilizó. Tenemos dos grandes aproximaciones, el análisis a nivel de clúster y el análisis a nivel individual.⁶

1. Análisis a nivel de clúster: es una técnica sencilla, a menudo preferible cuando el número total de clústers en el ensayo es muy pequeño (por ejemplo, menos de 15 por rama). Consiste en resumir primero los datos individuales, calculando valores promedio o proporciones para cada clúster (ejemplo: tasa de erradicación de *H. pylori* observada en cada centro) y, segundo, comparar esas medidas de efecto agregadas utilizando pruebas estadísticas tradicionales (como la prueba t de Student para muestras independientes).

Su ventaja es que no requiere modelar el ICC y su desventaja es la pérdida de flexibilidad del análisis, ya que es difícil ajustar el modelo por covariables (como la edad o las comorbilidades).³

2. Análisis a nivel individual: debido a su flexibilidad, el enfoque más común es utilizar todos los datos de los pacientes aplicando modelos de regresión que consideren la estructura jerárquica de los datos. Para ello existen dos grandes familias de modelos:

Modelos de ecuaciones de estimación generalizada (GEE, por su sigla en inglés) o modelos marginales: estiman los coeficientes de regresión tratando a la variación entre los clústers como un parámetro que estorba y debe corregirse. El resultado de un modelo GEE tiene una interpretación de efecto marginal o promedio poblacional. Es decir, que nos da cuál es el efecto promedio de la intervención en toda la población estudiada. Una de sus limitaciones es que solo puede manejar un único nivel de agrupación (por ejemplo: agrupar por hospital, pero no por médico dentro del hospital simultáneamente). Además, pueden subestimar el error estándar si el número de clústeres es pequeño.¹

Modelos de regresión multinivel (jerárquicos, de efectos mixtos o efectos aleatorios): parten de una regresión convencional y luego incluyen efectos aleatorios específicos de clúster, que consideran la homogeneidad de los resultados dentro de cada grupo. A diferencia de los modelos GEE, los modelos multinivel calculan el efecto condicional o específico del clúster. Esto significa que el efecto estimado representa el cambio esperado en el resultado al comparar dos individuos que pertenecen al mismo conglomerado. Asimismo, estos modelos permiten descomponer la varianza total calculando qué porcentaje de la variación se debe a las diferencias entre los conglomerados frente a la variación entre los individuos. Otra ventaja es que permiten manejar múltiples niveles de agrupamiento (por ejemplo: pacientes por médico tratante, que a su vez se agrupan por hospitales). Sin embargo, son modelos más complejos y requieren supuestos sobre la distribución de los efectos aleatorios específicos de clúster.

3. Modelos de análisis de covarianza (ANCOVA): se pueden utilizar tanto a nivel de clúster como a nivel individual. Permiten ajustar el resultado final por una covariable; en este caso, los valores basales o previos a

la aleatorización de la variable de resultado (siempre y cuando esta variable sea continua).

El uso de ANCOVA a nivel de análisis individual es una estrategia recomendada cuando existen limitaciones de poder estadístico, ya que permite reducir el tamaño de la muestra requerida.

Otro punto a tener en cuenta es la estrategia de análisis. Se debe definir de antemano qué hacer cuando los participantes o los clústers completos no cumplen con el tratamiento asignado. Lo recomendable en ensayos de superioridad es realizar un análisis por intención de tratar: esto significa analizar a todos los participantes en función del tratamiento asignado durante la aleatorización, independientemente si hubo cumplimiento efectivo. El análisis por intención de tratar tiende a disminuir el error tipo I comparado con el análisis por protocolo.

¿Cómo reportar o evaluar la publicación de un cRCT?

Diversas revisiones de la literatura han revelado deficiencias críticas en la publicación de cRCT. Por tal motivo se elaboró la extensión de la declaración CONSORT 2010 para ensayos por conglomerados, que fue publicada en 2012. Allí se incluyó una lista de verificación de los elementos que deben estar presentes en el reporte de estos estudios. Se debe incluir información adicional que no se pide en los ensayos clínicos individuales, tal como los métodos de análisis que contempla el agrupamiento, el número de clústers, el tamaño de los mismos y, sobre todo, la estimación del ICC con su grado de incertidumbre.⁵

Para concluir, podemos afirmar que la validez de los resultados de un ensayo clínico por conglomerados descansa en que se haya aplicado la metodología adecuada. Al responder las preguntas planteadas, estamos en condiciones de realizar una lectura crítica de este tipo de estudios.

Propiedad intelectual. *El autor declara que los datos presentes en el manuscrito son originales y se realizaron en su institución perteneciente.*

Financiamiento. *El autor declara que no hubo fuentes de financiación externas.*

Conflictos de interés. El autor declara no tener conflictos de interés en relación con este artículo.

Aviso de derechos de autor



© 2026 *Acta Gastroenterológica Latinoamericana*. Este es un artículo de acceso abierto publicado bajo los términos de la Licencia Creative Commons Attribution (CC BY-NC-SA 4.0), la cual permite el uso, la distribución y la reproducción de forma no comercial, siempre que se cite al autor y la fuente original.

Cite este artículo como: Rossi E. Introducción a los ensayos clínicos por conglomerados. *Acta Gastroenterol Latinoam*. 2026;56(3):261-264. <https://doi.org/10.52787/agl.v56i3.668>

Referencias

1. Austin PC, Kapral MK, Vyas MV, Fang J, Yu AYY. Using Multilevel Models and Generalized Estimating Equation Models to Account for Clustering in Neurology Clinical Research. *Neurology*. 2024 Nov 12;103(9):e209947. DOI: [10.1212/WNL.0000000000209947](https://doi.org/10.1212/WNL.0000000000209947). Epub 2024 Oct 11. PMID: 39393031; PMCID: PMC11469681.
2. Campbell MK, Elbourne DR, Altman DG; CONSORT group. CONSORT statement: extension to cluster randomised trials. *BMJ*. 2004 Mar 20;328(7441):702-8. DOI: [10.1136/bmj.328.7441.702](https://doi.org/10.1136/bmj.328.7441.702). PMID: 15031246; PMCID: [PMC381234](https://pubmed.ncbi.nlm.nih.gov/PMC381234/).
3. Campbell MJ, Walters SJ. How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research. Chichester: John Wiley & Sons; 2014.
4. Brown AW, Li P, Bohan Brown MM, Kaiser KA, Keith SW, Oakes JM, Allison DB. Best (but oft-forgotten) practices: designing, analyzing, and reporting cluster randomized controlled trials. *Am J Clin Nutr*. 2015 Aug;102(2):241-8. DOI: [10.3945/ajcn.114.105072](https://doi.org/10.3945/ajcn.114.105072). Epub 2015 May 27. PMID: [26016864](https://pubmed.ncbi.nlm.nih.gov/26016864/); PMCID: [PMC4515862](https://pubmed.ncbi.nlm.nih.gov/PMC4515862/).
5. Campbell MK, Piaggio G, Elbourne DR, Altman DG; CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. *BMJ*. 2012 Sep 4;345:e5661. DOI: [10.1136/bmj.e5661](https://doi.org/10.1136/bmj.e5661). PMID: 22951546.
6. Billot L, Copas A, Leyrat C, Forbes A, Turner EL. How should a cluster randomized trial be analyzed? *J Epidemiol Popul Health*. 2024 Feb;72(1):202196. DOI: [10.1016/j.jep.2024.202196](https://doi.org/10.1016/j.jep.2024.202196). Epub 2024 Feb 10. PMID: 38477477; PMCID: [PMC7616648](https://pubmed.ncbi.nlm.nih.gov/PMC7616648/).