

3

DEL RR DEL ECG AL RR DEL RCT

Una aproximación a la interpretación de los resultados del RCT para clínicos no epidemiólogos

Jorge Hernando Donado Gómez M.D. MSc*.

RESUMEN

El Ensayo clínico aleatorio (Randomized Controlled Trial – RCT-) es el diseño metodológico de mayor valor para la evaluación de medidas preventivas o terapéuticas en el área de la salud. En el presente artículo se expone en forma clara y resumida la interpretación de los resultados de los ensayos clínicos para el personal de la salud sin profundos conocimientos en epidemiología.

Palabras clave: Pruebas Aleatorias Controladas; Ensayo clínico; Intervalo de confianza; Riesgo.

* *Profesor Asociado U.P.B. – F.U.S.M. Medellín .Colombia.S.A.*

Separatas: Dr. Jorge Donado, A.A. 56006, Medellín - Colombia, S.A.

ABSTRACT

The Randomized Controlled Trial (RCT) is the most valuable methodological design used for the evaluation of preventive or therapeutic actions in health.

In this article, I expose clearly and briefly the interpretation of the outcomes of the Randomized Controlled Trials to the health personnel without deep knowledge in epidemiology.

Key Words: Randomized Controlled Trials; Clinical Trial; Confidence Interval; Risk.

Durante las básicas en medicina, se estudian los elementos fisiológicos del electrocardiograma (ECG) y posteriormente en Medicina Interna o Cardiología se enfrenta a su interpretación, donde se debe tener presente el ritmo, la frecuencia, el eje del QRS, y las diferentes ondas del registro. Si la frecuencia está aumentada y el intervalo QRS estrecho, se debe determinar si el intervalo RR es regular o irregular para pensar en una fibrilación auricular, taquicardia auricular multifocal u otra taquiarritmia supraventricular. (1)

Existe “otro” RR, el riesgo relativo, medida de asociación epidemiológica. Al igual que el fonendoscopio, el tensiómetro o el termómetro, instrumentos utilizados por los médicos, la información de la literatura y su adecuada interpretación son una herramienta indispensable en el quehacer profesional. La calculadora que se emplea para la determinación exacta de la frecuencia cardíaca o la duración del intervalo QTc en el ECG, ahora se utilizará también para calcular el riesgo relativo (RR) de los RCT (Randomized controlled trial).

En el siglo XXI es fundamental para todos los clínicos analizar e interpretar adecuadamente la información científica publicada periódicamente en la literatura mundial. El experimento o ensayo clínico controlado (RCT) es un diseño epidemiológico que ofrece el nivel más alto de evidencia en medidas de tratamiento o prevención, cuando es llevado con todo el rigor metodológico. (2)

En resumen, el RCT es un diseño donde se comparan dos (o más) grupos seleccionados en forma aleatoria de una muestra de pacientes. Estos grupos son lo

más parecido posible en sus características sociodemográficas, severidad de la enfermedad y cointervención. A un grupo se le asigna la intervención (medicamento activo o nuevo, procedimiento a evaluar) y el otro grupo actúa como control (placebo, medicamento o procedimiento estándar); ambos grupos se siguen en el tiempo durante un período predeterminado y se observa la incidencia (casos nuevos) del desenlace (resultado, outcome, end point) que se desea valorar (mortalidad, morbilidad, calidad de vida, días hospitalarios, etc). Debido a que en teoría los grupos son semejantes excepto en la intervención aplicada, cualquier diferencia en los desenlaces son atribuibles a la intervención. (3)

Los resultados del RCT se deben analizar una vez se ha estudiado la validez interna del mismo, para lo cual existen diferentes guías que ayudan a realizarla. Si el estudio pasa la fase de validez interna, viene el análisis de los resultados. Ellos pueden presentarse con base en el *valor de P* o con estimados puntuales e intervalos de confianza. (4-6)

El *valor de P*, desafortunadamente todavía muy utilizado, se define como la probabilidad de obtener un resultado igual o más extremo que el observado, asumiendo que no hay efecto o diferencia (Hipótesis nula cierta). La crítica mayor es que la medida de la evidencia *no* tiene en cuenta el tamaño del efecto observado. Eso significa que diferencias en el desenlace de dos grupos de tratamiento, por ejemplo el medicamento A disminuye la mortalidad en un 15% y el medicamento B en un 5%, con un valor de $p < 0.05$, puede ser similar si el medicamento C disminuye la mortalidad

en 7.2% y el medicamento D en un 5.9% con un valor $p < 0.05$ (7)

Una diferencia significativa estadísticamente *no* es siempre equivalente a importancia clínica. Con tamaños de muestra suficientemente grandes se pueden demostrar pequeñas diferencias estadísticamente significativas, pero quizás clínicamente no. En un estudio publicado recientemente, se encuentra que el pH en el grupo de terapia estándar vs el grupo de intervención en el período de 7 a 72 horas fue 7.36 ± 0.12 vs 7.40 ± 0.12 y valor $p < 0.001$, lo cual es estadísticamente significativo pero no clínicamente importante(8)

La segunda forma de analizar los resultados es con la estimación de la magnitud (tamaño, intensidad) del efecto de la intervención con sus intervalos de confianza. Esto lo podemos realizar de cuatro maneras, conociendo la incidencia del desenlace entre los expuestos a la intervención y los no expuestos: (2,5,9-12):

Riesgo Absoluto (RA): Es la probabilidad de desarrollar un desenlace nuevo durante un período de tiempo. La incidencia es la expresión básica del riesgo.

Riesgo Relativo (RR): Es la relación entre dos riesgos absolutos. Es el resultado de la división de la proporción de incidencia (o riesgo absoluto) en el grupo expuesto entre el grupo no expuesto. $(I_e / I_{ne}) = (R_a e / R_a ne)$

Su valor está entre $0 \rightarrow 1 \rightarrow \alpha$

Un valor de 1 significa que ambas incidencias son iguales, y por lo tanto no hay efecto de la intervención sobre el desenlace.

Valores menores de 1 indican que el grupo de la intervención protege, o desarrolla menos eventos que el grupo control (Factor protector), siempre y cuando el desenlace medido sea “negativo”, tal como mortalidad, morbilidad, recaída, efectos adversos

Valores mayores de 1 indican que el grupo de la intervención es de mayor riesgo para el evento que el grupo control (Factor de riesgo), siempre y cuando el desenlace sea “negativo”.

Para eventos positivos (curación, supervivencia, etc.), valores mayores de 1 indican que el grupo de la intervención es más efectivo en conseguir tales desenlaces. Los valores menores de 1 indican que la intervención es inefectiva.

Si se tiene, en un ejemplo hipotético, una incidencia de mortalidad en expuestos a la aspirina del 40% y en los no expuestos (placebo) del 50%, el $RR = 40 / 50 = 0.8$, significa que por cada paciente del grupo control (placebo) que desarrolla el evento de interés (mortalidad), 0.8 pacientes lo hacen en el grupo de la intervención (aspirina), o lo que es más comprensible, por 100 pacientes que se mueren en el grupo del placebo, 80 pacientes lo hacen en el grupo de la aspirina. (Efecto protector)

Si se tiene, en otro ejemplo hipotético, una incidencia de cáncer de mama en expuestos a estrógenos del 20% y en los no expuestos (placebo) del 12.5%, el $RR = 20 / 12.5 = 1.6$, significa que por cada paciente del grupo control (placebo) que desarrolla el evento de interés (cáncer de mama) 1.6 pacientes lo hacen en el grupo de la

intervención (estrógenos) o por 100 pacientes que desarrollan cáncer de mama en el grupo del placebo, 160 pacientes lo hacen en el grupo que recibió estrógenos. (Factor de riesgo)

Reducción del Riesgo Relativo (RRR): se obtiene restando el RR de la unidad ($1 - RR$). Lo cual expresa en forma porcentual cuanto disminuye (o aumenta) el riesgo del desenlace evaluado al relacionarlo con lo que ocurre en los pacientes del grupo control. No distingue efectos de tratamiento importantes de triviales, disminuir la mortalidad del 60 al 30% produce el mismo RRR que disminuirla de 0.6% a 0.3%.

Su valor está entre $-\alpha \rightarrow 0 \rightarrow 1$

$RRR = 1 - 0.8 = 0.2$ lo cual significa que la intervención (aspirina) disminuye en un 20% el desenlace (mortalidad) en el grupo de intervención con relación al grupo control (placebo)

$RRR = 1 - 1.6 = -0.6$ significa un incremento del riesgo (IRR) del 60% de desarrollar el desenlace (cáncer de mama) en el grupo de intervención (estrógenos) referente al grupo control (placebo)

Reducción Absoluta Riesgo (RAR): es la diferencia (resta) de la incidencia del evento en los no expuestos contra la incidencia del evento en los expuestos a la intervención. Indica en forma absoluta el porcentaje en que se disminuye el evento en el grupo de la intervención.

Su valor está entre $-\alpha \rightarrow 0 \rightarrow 1$

$RAR = 50 - 40 = 10$ lo cual significa que la intervención (aspirina) disminuye en un 10% el desenlace (mortalidad) en forma absoluta.

$RAR = 12.5 - 20 = -7.5$ significa un incremento del riesgo absoluto (IAR) del 7.5% de desarrollar el desenlace (cáncer de mama) en el grupo de la intervención (estrógenos).

Número Necesario a Tratar (NNT) = es el inverso de la RAR ($1 / RAR$) o ($100 / RAR$ en %). Significa el número de pacientes que es necesario tratar para disminuir un (1) evento o desenlace, de la misma forma y durante el mismo tiempo que se hizo en el estudio. Mientras mayor sea la RAR, menor será el NNT

$NNT = 1 / 0.1 = 10$ y significa que se debe tratar a 10 pacientes con la intervención del estudio (aspirina) y durante el tiempo de seguimiento para prevenir un desenlace adverso (mortalidad)

$NNH = 1 / -7.5 = -13$ y significa que se requiere tratar 13 pacientes con la intervención del estudio (estrógenos) para producir un efecto dañino (Harm) o adverso (cáncer de mama).

El NNT también puede ser obtenido mediante la fórmula $1 / (RRR \times I_{ne})$, la cual es muy útil e importante para el clínico en su práctica diaria, ya que le permite conocer al frente de su paciente, que tiene un riesgo puntual (I_{ne} o riesgo absoluto en no expuesto) y con la RRR obtenida del RCT, cuál es en NNT de su paciente específico. Mientras mayor sea el riesgo absoluto del paciente, menor será el NNT(13,14).

Si se tiene que la RRR de una intervención es del 20% y estoy al frente de dos pacientes, uno con un riesgo basal del 50% del evento y otro del 1%, necesito tratar solamente 10 pacientes ($1 / 0.2 \times 0.5$) para

beneficiar a uno con riesgo del 50%, mientras tengo que tratar 500 pacientes ($1 / 0.2 \times 0.01$) con riesgo basal de 1% para beneficiar a uno.

INTERVALOS DE CONFIANZA

Lo que se acaba de calcular son los estimados (aproximaciones) puntuales de la magnitud o tamaño del efecto de la intervención con respecto al grupo control. El estimado puntual nos indica que el valor verdadero está cercano a éste, pero es improbable que sea exactamente dicho valor. Estos estimados puntuales, *siempre* deben ir acompañados con sus respectivos intervalos de confianza (90, 95, 99%), siendo lo más usual del 95%, ya que nos informan acerca del grado de incertidumbre y nos dicen con una confianza determinada (por ejemplo 95%) entre que valores se encuentra el efecto verdadero o real de la intervención. En otras palabras, si se repite el estudio en idénticas condiciones en 95 de 100 veces el estimado puntual se encontrará entre los límites del intervalo calculado. Es el rango entre el cual probablemente se encuentra el efecto del tratamiento verdadero(15)

Para la interpretación del Intervalo de Confianza es necesario tener en cuenta las siguientes tres preguntas:

1. ¿El intervalo atraviesa la unidad?
Cuando se trata del RR o OR

¿El intervalo atraviesa el cero? Cuando se trata del RAR, RRR o NNT (Si no lo hace los resultados son estadísticamente significativos y se asocian con un valor de $p < 0.05$)

2. ¿Que tan preciso es el intervalo? (Si los límites superior e inferior son próximos o lejanos del estimado puntual. La precisión está directamente relacionada con el tamaño de la muestra, a mayor tamaño de muestra, intervalos más precisos y viceversa)
3. Si el intervalo es no significativo, o sea que atraviesa la unidad, se debe observar la "tendencia", de cual de los límites del intervalo (superior o inferior) está más cerca de la unidad y si es probable que con un tamaño de muestra mayor se obtendría significación estadística.

En estudios "positivos", o sea aquellos que muestran significación estadística, se debe analizar el límite inferior del intervalo de confianza y si este es aún importante o clínicamente significativo, entonces los investigadores han reunido en número de pacientes suficientes.

En estudios "negativos", o sea aquellos que no muestran significación estadística, se debe analizar el límite superior del intervalo de confianza y si éste es aún importante o clínicamente significativo, entonces los investigadores han fallado al excluir un efecto importante del tratamiento por no incluir el número suficiente de pacientes en el experimento.(5)

Ejemplo # 1

Si en un grupo de 1000 pacientes expuestos a la intervención (vitamina E), 80 desarrollan el evento (enfermedad de Alzheimer) durante el seguimiento del estudio, la incidencia será $80/1000 = 0.08$ (8%)

Si en un grupo de 1000 pacientes del grupo control (placebo) 190 desarrollan el evento (enfermedad de Alzheimer) durante el seguimiento del estudio, la incidencia será $190/1000 = 0.19$ (19%)

		Estimado puntual	IC 95%
Incidencia en expuestos (Ie)	80 / 1000	0.08 (8%)	
Incidencia en no expuestos (Ine)	190 / 1000	0.19 (19%)	
RR (Ie / Ine)	0.08 / 0.19	0.42	0.33 - 0.54
RRR (1 - RR)	1 - 0.42	0.58 (58%)	0.46 - 0.67
RAR (Ine - Ie)	0.19 - 0.08	0.11 (11%)	0.08 - 0.14
NNT (1 / RAR)	1 / 0.11	9	7 - 12

$$RR = 0.08 / 0.19 = 0.42$$

El RR de 0.42 significa que por cada paciente del grupo placebo que desarrolla enfermedad de Alzheimer, 0.42 pacientes lo hacen en el grupo de la vitamina E, o lo que es más comprensible, por 100 pacientes que desarrollan el evento de interés en el grupo control, 42 pacientes lo hacen en el grupo de la intervención.

El intervalo nos dice con una confianza del 95% que el valor verdadero o real se encuentra entre un mínimo de 33 a un máximo 54 pacientes del grupo de la vitamina E que desarrollan la enfermedad por cada 100 del grupo control. Como no atraviesa la unidad, es un resultado significativo estadísticamente y moderadamente preciso.

$$RRR = 1 - 0.58 = 0.42$$

La RRR de 0.58 significa que la vitamina E disminuye en un 58% el riesgo de la enfermedad de Alzheimer con respecto a los pacientes del grupo placebo.

El intervalo nos dice con una confianza del 95% que la reducción relativa del evento puede ser tan alta como un 67% o tan baja como un 46% en la realidad.

$$RAR = 0.19 - 0.08 = 0.11 \text{ (11\%)}$$

La RAR de 0.11 significa que la vitamina E disminuye en forma absoluta en un 11% el riesgo de la enfermedad de Alzheimer

El intervalo nos dice con una confianza del 95% que la reducción absoluta puede ser tan alta como un 14% o tan baja como un 8% en la realidad.

$$NNT = 1 / 0.11 = 9 \text{ (100 / 11\% = 9)}$$

El NNT de 9 significa que se deben tratar a 9 pacientes con vitamina E durante para beneficiar el desarrollo de la enfermedad de Alzheimer a un (1) paciente. El intervalo nos dice con una confianza del 95% que en el mejor de los casos sólo tengo que tratar a 9 o en el peor a 12 pacientes, para beneficiar a uno.

Ejemplo # 2

Si en un grupo de 100 pacientes hipertensos expuestos a captopril 10 desarrollan enfermedad cerebrovascular (ECV) durante el seguimiento del estudio, la incidencia será $10/100 = 0.1$ (10%)

Si en un grupo control de 100 pacientes hipertensos expuestos a verapamilo 12 desarrollan ECV durante el seguimiento del estudio, la incidencia será $12/100 = 0.12$ (12%)

		Estimado puntual	IC 95%
Incidencia en expuestos (Ie)	10 / 100	0.1 (10%)	
Incidencia en no expuestos (Ine)	12 / 100	0.12 (12%)	
RR (Ie / Ine)	0.1 / 0.12	0.83	0.38 - 1.84
RRR (1 - RR)	1 - 0.83	0.17 (17%)	-0.8 - 0.62
RAR (Ine - Ie)	0.12 - 0.1	0.02 (2%)	-0.1 - 0.11
NNT (1 / RAR)	1 / 0.02	50	9.3 - (-15)

$$RR = 0.10 / 0.12 = 0.83$$

El RR de 0.83 significa que por cada paciente del grupo verapamilo que desarrolla ECV 0.83 pacientes lo hacen en el grupo del captopril o lo que es más comprensible, por 100 pacientes que desarrollan el evento de interés en el grupo control, 83 pacientes lo hacen en el grupo de la intervención.

El intervalo nos dice con una confianza del 95% que el valor verdadero o real se encuentra entre un mínimo de 38 a un máximo de 184 pacientes del grupo de intervención (captopril) que desarrollan el evento por 100 del grupo control

(verapamilo). Como atraviesa la unidad, es un resultado no significativo estadísticamente.

$$RRR = 1 - 0.83 = 0.17$$

La RRR de 0.17 significa que el captopril disminuye en un 17% el riesgo de ECV con respecto a los pacientes del grupo verapamilo

El intervalo nos dice con una confianza del 95% que la reducción puede ser tan alta como un 62% en el grupo del captopril o reducir el riesgo de ECV en 80% en el grupo verapamilo.

$$\text{RAR} = 0.12 - 0.10 = 0.02 \text{ (2\%)}$$

La RAR de 0.02 significa que el tratamiento con captopril disminuye en forma absoluta en un 2% el riesgo del desenlace (ECV)

El intervalo nos dice con una confianza del 95% que la reducción absoluta puede ser tan alta como un 11% en el grupo captopril o disminuir el riesgo del evento en 10% con verapamilo

$$\text{NNT} = 1 / 0.02 = 50 \text{ (100 / 2\% = 50)}$$

El NNT de 50 significa que se deben tratar 50 pacientes con captopril para beneficiar del desarrollo de ECV a un (1) paciente. El intervalo nos dice con una confianza del 95% que en el mejor de los casos sólo tengo que tratar a 9 pacientes para beneficiar a uno o en el peor de los casos, tratando a 15 produzco un efecto adverso (ECV). Por ser negativo (-15) se denomina NNH (número necesario para producir daño). ■

REFERENCIAS

- (1) Goldberger AL. Electrocardiography. En: Braunwald, Fauci, Kasper, Hauser, Longo, Jameson, editors. Harrison's Principles of Internal Medicine. New York : Mc Graw Hill, 2001.
- (2) Guyatt GH, Sackett DL, Sinclair JC, Hayward R, Cook DJ, Cook RJ. Users' guides to the medical literature. IX. A method for grading health care recommendations. Evidence-Based Medicine Working Group. JAMA 1995; 274(22):1800-1804.
- (3) Lozano JM, Dennis RJ. El experimento clínico. En: Ruiz A, Gómez C, Londoño D, editors. Investigación clínica: Epidemiología clínica aplicada. Bogotá, D.C.: Centro Editorial Javeriano, 2001.
- (4) Guyatt GH, Sackett D, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? JAMA 1993; 270:2598-2601.
- (5) Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? JAMA 1994; 271:59-63.
- (6) Gordon S, Doig. Interpreting and using clinical trials. Critical Care Clinics 1998; 14: 513-523.
- (7) Goodman SN. Toward evidence-based medical statistics. 1: The P value fallacy. Ann Intern Med 1999; 130(12):995-1004.
- (8) Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B et al. Early Goal-Directed Therapy in the Treatment of Severe Sepsis and Septic Shock. N Engl J Med 2001; 345:1368-1377.
- (9) Sackett DL, Cook RJ. Understanding clinical trials. BMJ 1994; 309(6957):755-756.
- (10) McAlister EA, Straus SE, Guyatt GH, Haynes RB. Users' guides to the medical literature: XX. Integrating research evidence with the care of the individual patient. JAMA 2000; 283:2829-2836.
- (11) Sackett DL. Why randomized controlled trials fail but needn't: 1. Failure to gain «coal-face» commitment and to use the uncertainty principle. CMAJ 2000; 162(9):1311-1314.
- (12) Sackett DL. Why randomized controlled trials fail but needn't: 2. Failure to employ physiological statistics, or the only formula a clinician-trialist is ever likely to need (or understand!). CMAJ 2001; 165(9):1226-1237.
- (13) Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ 1995; 310(6977):452-454.
- (14) Chatellier G, Zapletal E, Lemaitre D, Menard J, Degoulet P. The number needed to treat: a clinically useful nomogram in its proper context. BMJ 1996; 312(7028):426-429.
- (15) Guyatt GH, Jaeschke R, Heddke N, Cook DJ, Shannon H, Walter S. Interpreting study results: confidence intervals. CMAJ 1995; 152:196-173.