Salud Animal

Original

Aplicación informática basada en los principios 3R para el diseño y análisis de estudios preclínicos en ratones de laboratorio

Software application based on the 3R principles for the design and analysis of preclinical studies in laboratory mice

José Aureliano Betancourt-Bethencourt *, Lidyce Quesada-Leyva *, Zaddys Ahimara Ruiz Hunt *, Alexis Ricardo Gregori Caballero *

* Universidad de Ciencias Médicas de Camagüey, Centro de Inmunología y Productos Biológicos. Camagüey, Cuba.
Correspondencia: betanster@gmail.com

Recibido: Julio, 2026; Aceptado: Septiembre, 2026; Publicado: Noviembre, 2026.

RESUMEN

Antecedentes: El cumplimiento de los principios de las 3R (reemplazo, reducción y refinamiento) en la experimentación con animales de laboratorio requiere herramientas accesibles que no comprometan el rigor estadístico. **Objetivo.** Desarrollar una aplicación informática interactiva que integre estos principios en el diseño y el análisis de estudios preclínicos con ratones (*Mus musculus*). **Métodos:** La aplicación se implementó en R mediante el paquete Shiny y se organizó en seis módulos: inicio con lista de verificación 3R; cálculo del tamaño muestral por análisis de potencia; simulación de estudios piloto con distribuciones normal, log-normal, exponencial y t de Student; comparación estadística con pruebas paramétricas y no paramétricas y tamaños del efecto; registro de indicadores de bienestar animal con alertas automáticas; y generación de un reporte descargable. La validación funcional se realizó con escenarios representativos, cuyos resultados se contrastaron con cálculos manuales en R y valores de referencia de la literatura. **Resultados:** Para un tamaño del efecto de 0,80, una potencia de 0,80, un nivel de significación de 0,05 y una pérdida esperada de 10 %, la aplicación calculó 26 animales por grupo (52 en total) y 58 animales tras el ajuste por pérdidas, valores coincidentes con los de referencia. El reporte generado documentó las decisiones metodológicas, lo que facilita la justificación del tamaño muestral ante los comités de ética. **Conclusiones:** La aplicación constituyó un aporte metodológico de bajo costo, reproducible y de código abierto, adoptable en instituciones de investigación agropecuaria y biomédica.

Como citar (APA) Betancourt-Bethencourt, JA, Quesada-Leyva, L., Ruiz Hunt, ZA y Gregori Caballero, AR (2026). Aplicación informática basada en los principios 3R para el diseño y análisis de estudios preclínicos en ratones de laboratorio. *Revista de producción Animal , 38* . https://rpa.reduc.edu.cu/index.php/rpa/article/view/e346

©El (los) autor (es), Revista de Producción Animal 2020. Este artículo se distribuye bajo los términos de la licencia internacional Attribution-NonCommercial 4.0 (https://creativecommons.org/licenses/by-nc/4.0/), asumida por las colecciones de revistas científicas de acceso abierto, según lo recomendado por la Declaración de Budapest, la que puede consultarse en: Budapest Open Access Initiative's definition of Open Access.

Palabras clave: bienestar animal, experimentación animal, potencia estadística, simulación por computador, tamaño de la muestra (*Fuente: AIMS*)

ABSTRACT

Background: Compliance with the 3R principles (replacement, reduction, and refinement) in laboratory animal experimentation requires accessible tools that do not compromise statistical rigor. **Aim.** To develop an interactive software application that integrates these principles into the design and analysis of preclinical studies with mice (*Mus musculus*). **Methods:** The application was implemented in R using the Shiny package and organized into six modules: an introduction with a 3R checklist; sample size calculation by power analysis; pilot study simulation with normal, log-normal, exponential, and Student's t distributions; statistical comparison with parametric and non-parametric tests and effect sizes; recording of animal welfare indicators with automatic alerts; and generation of a downloadable report. Functional validation was performed with representative scenarios, and the results were compared with manual calculations in R and reference values from the literature. **Results:** For an effect size of 0.80, a power of 0.80, a significance level of 0.05, and an expected attrition of 10 %, the application calculated 26 animals per group (52 in total) and 58 animals after adjustment for attrition, in agreement with reference values. The generated report documented the methodological decisions, which facilitates sample size justification before ethics committees. **Conclusions:** The application was a low-cost, reproducible, open-source methodological contribution that can be adopted by agricultural and biomedical research institutions.

Keywords: animal experimentation, animal welfare, computer simulation, sample size, statistical power (*Source: AIMS*)

INTRODUCCIÓN

La experimentación animal constituye un pilar de la investigación biomédica y agropecuaria, pero plantea dilemas éticos que han motivado marcos regulatorios internacionales. El principio de las 3R, formulado por Russell y Burch (1959), sigue siendo el referente conceptual más aceptado para orientar el uso ético de animales en investigación: el reemplazo promueve métodos alternativos *in vitro* o *in silico*; la reducción persigue obtener información válida con el menor número posible de animales, lo que exige un cálculo riguroso del tamaño muestral; y el refinamiento busca minimizar el sufrimiento y mejorar el bienestar de los animales empleados (Fenwick *et al*., 2009; Hubrecht y Carter, 2019).

En la producción animal y la investigación veterinaria, el ratón (*Mus musculus*) es el modelo preclínico más utilizado, por su bajo costo de mantenimiento, la disponibilidad de cepas definidas y la extensa caracterización de su fisiología (Hedrich & Bullock, 2004). No obstante, numerosos estudios preclínicos presentan deficiencias en el cálculo del tamaño muestral, lo que conduce al desperdicio de animales y a resultados poco consistentes (Button *et al*., 2013; Ioannidis, 2014; Mazoit y Gonin, 2026). El problema es mayor en instituciones con recursos limitados, donde la falta de herramientas accesibles para el análisis de potencia y la simulación de datos dificulta el cumplimiento efectivo de las 3R.

Los entornos de programación estadística de código abierto, como R (R Core Team, 2023), y los marcos para aplicaciones web interactivas, como Shiny (Chang *et al*., 2021), permiten trasladar metodologías complejas a usuarios no especialistas en programación. En este contexto, el objetivo del trabajo fue desarrollar una aplicación informática interactiva, basada en R y Shiny, que integre los principios de las 3R en un flujo de trabajo único para el diseño y análisis de estudios preclínicos con ratones, como recurso metodológico para instituciones de investigación agropecuaria y biomédica.

MATERIALES Y MÉTODOS

Entorno de desarrollo

La aplicación se desarrolló en R (versión 4.3 o superior; R Core Team, 2023) con el paquete Shiny (Chang *et al*., 2021) para la interfaz web interactiva y shinydashboard para la estructura visual. Como dependencias adicionales se emplearon pwr, para el análisis de potencia (Champely, 2020); ggplot2, para la visualización (Wickham, 2016); DT, para tablas interactivas; dplyr, para la manipulación de datos (Wickham *et al*., 2023); y effectsize, para el cálculo de tamaños del efecto (Ben-Shachar *et al*., 2020). La aplicación instala automáticamente sus dependencias, por lo que su ejecución no requiere conocimientos previos de programación.

Estructura funcional de la aplicación

La aplicación se organizó en seis módulos accesibles desde un menú lateral (Tabla 1). La figura 1 muestra la interfaz del módulo de cálculo del tamaño muestral.

Tabla 1. Módulos de la aplicación informática basada en los principios 3R.

Módulo	Principio 3R	Función y entradas	Salidas principales
1. Inicio/3R	Todos	Fundamentos de las 3R y lista de verificación de seis criterios: modelos *in vitro*/*in silico*, tamaño muestral, análisis estadístico apropiado, puntos finales humanitarios, registro de bienestar, aleatorización y cegamiento.	Porcentaje de cumplimiento 3R en tiempo real
2. Tamaño muestral	Reducción	Análisis de potencia para dos medias independientes, dos medias pareadas, comparación de proporciones y ANOVA de una vía. Entradas: diferencia esperada (Δ), desviación estándar, potencia (1-β), α y tasa de pérdida.	N mínimo por grupo, N total, N total ajustado por pérdidas y curva de potencia
3. Simulación piloto	Reducción	Datos sintéticos para estimar la variabilidad. Diseño de dos grupos o ANOVA; distribución normal, log-normal, exponencial o t de Student; tamaño del piloto, número de réplicas y semilla.	DE combinada mediana con IC 95 %, potencia empírica y tamaño muestral requerido
4. Comparación de grupos	Reducción/ Refinamiento	Datos cargados desde CSV, ingresados manualmente o simulados. Prueba de Shapiro-Wilk y, según el número de grupos, t de Student o Wilcoxon, ANOVA de una vía o Kruskal-Wallis.	Valores de p; d de Cohen y g de Hedges con IC 95 %; η^2 y ω^2 (ANOVA); gráficos

Módulo	Principio 3R	Función y entradas	Salidas principales
5. Bienestar animal	Refinamiento	Registro de peso inicial y actual, score de bienestar (0-5), actividad y pelaje. Alertas si la pérdida de peso supera 15 %, el score es ≥ 3 o hay alteraciones de actividad o pelaje.	Clasificación del estado (aceptable, monitorizar, alerta) y alertas automáticas
6. Reporte 3R	Todos	Compilación de las decisiones metodológicas y evaluaciones realizadas durante la sesión.	Reporte descargable en formato de texto

Nota: 3R = Reemplazo, Reducción, Refinamiento; DE = desviación estándar; IC 95 % = intervalo de confianza del 95 %; N = número de animales; CSV = valores separados por comas; ANOVA = análisis de varianza; α = nivel de significación; Δ = diferencia esperada.



Figura. 1. Interfaz de la aplicación en el módulo de tamaño muestral (Reducción): parámetros de entrada y resultados del cálculo para dos medias independientes. *Nota:* Δ = diferencia esperada; α = nivel de significación; N = número de animales.

Validación funcional

La aplicación se sometió a pruebas de validación funcional con escenarios representativos: cálculo del tamaño muestral para dos medias independientes con un tamaño del efecto de 0,80, potencia de 0,80, α de 0,05 y pérdida esperada de 10 %; simulación de datos piloto con n = 5 por grupo bajo distribución normal; y comparación de grupos a partir de los datos simulados. Los resultados se contrastaron con cálculos manuales en R y con valores de referencia de la literatura (Cohen, 1988/2013).

Análisis estadístico y disponibilidad de datos

Todos los análisis se realizaron en R. La validación no incluyó animales: se utilizaron datos simulados, generados con semilla definida por el usuario para garantizar su reproducibilidad. El código fuente de la aplicación está disponible en https://zenodo.org/records/23582244.

RESULTADOS Y DISCUSIÓN

Validación del módulo de tamaño muestral

Para dos medias independientes con un tamaño del efecto de 0,80, una potencia de 0,80 y un α de 0,05, la aplicación calculó un mínimo de 26 animales por grupo (52 en total). Con una pérdida esperada de 10 %, el tamaño muestral total recomendado fue de 58 animales (52/0,90 = 57,8, redondeado al entero superior; Fig. 1). Estos valores coincidieron con los obtenidos mediante cálculo manual en R y con los de referencia (Cohen, 1988/2013). Los restantes escenarios de simulación piloto y comparación de grupos arrojaron resultados coherentes con los cálculos de verificación.

Integración de las 3R en un flujo de trabajo único

Muchas calculadoras de tamaño muestral disponibles en línea se limitan al principio de reducción. La aplicación desarrollada vincula el cálculo del tamaño muestral con la simulación piloto, la comparación estadística posterior y el registro del bienestar animal, en línea con la recomendación de tratar las 3R como dimensiones interdependientes de un mismo diseño experimental (Fenwick *et al*., 2009; Hubrecht y Carter, 2019). El principio de Replacement se aborda mediante la lista de verificación del módulo inicial, que induce al investigador a evaluar modelos *in vitro* o *in silico* antes de recurrir a animales.

Reducción y refinamiento del análisis estadístico

El módulo de reducción aborda una de las críticas más frecuentes a la investigación preclínica: la justificación insuficiente del tamaño muestral. Button *et al*. (2013) mostraron que la baja potencia estadística es frecuente en estudios de neurociencias, incluidos los realizados en modelos animales, y que reduce la confiabilidad de los resultados. Revisiones recientes subrayan que reducir no equivale a emplear el menor número posible de animales, sino el número justificado por el diseño y el análisis, y que los animales que se prevé excluir deben sumarse al tamaño calculado (Mazoit y Gonin, 2026; Zhang y Hartmann, 2023), lo que respalda el ajuste por pérdidas que realiza la aplicación. La simulación piloto permite explorar la variabilidad esperada a partir de datos sintéticos generados bajo distintas distribuciones, en lugar de asumir una desviación estándar arbitraria. Este enfoque es coherente con el uso de estudios de simulación para evaluar métodos estadísticos (Morris *et al*., 2019) y amplía la aplicabilidad a variables biológicas con asimetría o colas pesadas.

El módulo de comparación de grupos verifica el supuesto de normalidad y ofrece alternativas no paramétricas cuando no se cumple, y calcula tamaños del efecto con intervalos de confianza, que facilitan la interpretación biológica de las diferencias y responden a las guías de reporte de investigación con animales (Percie du Sert *et al*., 2020).

Refinamiento y documentación

El módulo de bienestar animal operacionaliza el principio de *refin*amiento mediante el registro sistemático de indicadores y la emisión automática de alertas cuando se superan umbrales predefinidos (pérdida de peso mayor de 15 % y score de bienestar ≥ 3). Estos criterios se apoyan en los enfoques de valoración del dolor, el distrés y el bienestar descritos en la literatura (Hawkins *et al*., 2011; Morton y Griffiths, 1985). Estudios recientes indican que la pérdida de peso es un indicador útil, pero difícil de interpretar de forma aislada, que su umbral debe ajustarse al modelo experimental y que conviene combinarla con otros criterios clínicos en puntuaciones compuestas (Brochut *et al*., 2024; Talbot *et al*., 2020, 2023). Por ello, los umbrales de la aplicación deben adaptarse a cada protocolo y ser aprobados por el comité de ética. El reporte descargable documenta las decisiones metodológicas y puede adjuntarse a las solicitudes de aprobación ética y a los manuscritos científicos, lo que contribuye a cerrar la brecha entre la declaración de cumplimiento de las 3R y su evidencia documental (Percie du Sert *et al*., 2020). Asimismo, coincide con la recomendación reciente de planificar y registrar a priori el cálculo del número de animales (Mazoit y Gonin, 2026).

Limitaciones y perspectivas

El cálculo del tamaño muestral se basa en el análisis de potencia tradicional, que asume normalidad y homocedasticidad; aunque la simulación piloto permite explorar escenarios no normales, el cálculo definitivo depende de supuestos paramétricos (Mazoit y Gonin, 2026). Además, los estudios piloto con pocos animales no resultan útiles para estimar el tamaño muestral cuando se desconocen el tamaño del efecto y la varianza poblacional (Bird *et al*., 2025); por ello, los resultados del módulo de simulación piloto deben interpretarse como una exploración de la incertidumbre y complementarse con datos previos o de la literatura. La herramienta no sustituye la consulta con un especialista en estadística ni la revisión del comité de ética, y su validación se limitó a escenarios representativos con datos simulados. Resulta necesario validarla con datos reales de estudios preclínicos, e incorporar módulos para datos longitudinales, la integración con sistemas de gestión de bioterios y una versión de la interfaz en inglés. Por funcionar en equipos modestos y sin licencias comerciales, la aplicación es especialmente pertinente para la investigación agropecuaria en instituciones cubanas con recursos limitados.

CONCLUSIÓN

Se desarrolló una aplicación informática interactiva, basada en R y Shiny, que integra los principios de las 3R en un flujo de trabajo único para el diseño y análisis de estudios preclínicos con ratones: calcula el tamaño muestral mediante análisis de potencia, simula datos piloto, ejecuta comparaciones estadísticas con tamaños del efecto, registra indicadores de bienestar animal y genera un reporte documentado de cumplimiento. La herramienta constituye un aporte

metodológico de bajo costo, reproducible y de código abierto. Se recomienda validarla con datos reales de instituciones cubanas, incorporarla a la formación de posgrado en bienestar animal y metodología de la investigación, y difundirla entre los comités de ética institucionales.

REFERENCIAS

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software, 5*(56), 2815. https://doi.org/10.21105/joss.02815

Bird, A. D., Jedlicka, P., & Wilhelm, J. (2025). Animal pilot studies should not be used to estimate sample size if effect size and population variance are unknown. *ALTEX, 42*(3), 531-535. https://doi.org/10.14573/altex.2408141

Brochut, M., Heinonen, T., Snäkä, T., Gilbert, C., Le Roy, D., & Roger, T. (2024). Using weight loss to predict outcome and define a humane endpoint in preclinical sepsis studies. *Scientific Reports, 14*, Article 21150. https://doi.org/10.1038/s41598-024-72039-1

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14*(5), 365-376. https://doi.org/10.1038/nrn3475

Champely, S. (2020). *pwr: Basic functions for power analysis* (R package version 1.3-0) [Software]. https://CRAN.R-project.org/package=pwr

Chang, W., Cheng, J., Allaire, J. J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., & Borges, B. (2021). *shiny: Web application framework for R* (R package version 1.7.1) [Software]. https://CRAN.R-project.org/package=shiny

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences* (2.ª ed.). Routledge. (Obra original publicada en 1988). https://doi.org/10.4324/9780203771587

Fenwick, N., Griffin, G., & Gauthier, C. (2009). The welfare of animals used in science: How the "Three Rs" ethic guides improvements. *The Canadian Veterinary Journal*, *50*(5), 523-530. https://pmc.ncbi.nlm.nih.gov/articles/PMC2671878

Hawkins, P., Morton, D. B., Burman, O., Dennison, N., Honess, P., Jennings, M., Lane, S., Middleton, V., Roughan, J. V., Wells, S., & Westwood, K. (2011). A guide to defining and implementing protocols for the welfare assessment of laboratory animals: Eleventh report of the BVAAWF/FRAME/RSPCA/UFAW Joint Working Group on Refinement. *Laboratory Animals*, *45*(1), 1-13. https://doi.org/10.1258/la.2010.010031

Hedrich, H. J., & Bullock, G. R. (Eds.). (2004). *The laboratory mouse*. Elsevier Academic Press.

Hubrecht, R. C., & Carter, E. (2019). The 3Rs and humane experimental technique: Implementing change. *Animals, 9*(10), 754. https://doi.org/10.3390/ani9100754

Ioannidis, J. P. A. (2014). How to make more published research true. *PLoS Medicine, 11*(10), e1001747. https://doi.org/10.1371/journal.pmed.1001747

Mazoit, J. X., & Gonin, P. (2026). Reduce or not reduce: A practical guide to calculate sample size for various experimental design in studies involving animals. *Frontiers in Pharmacology*, *17*, Article 1786234. https://doi.org/10.3389/fphar.2026.1786234

Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine, 38*(11), 2074-2102. https://doi.org/10.1002/sim.8086

Morton, D. B., & Griffiths, P. H. M. (1985). Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Veterinary Record, 116*(16), 431-436. https://content.wellcomecollection.org/works/hefrkyft

Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., . . . Würbel, H. (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLoS Biology*, *18*(7), e3000410. https://doi.org/10.1371/journal.pbio.3000410

R Core Team. (2023). *R: A language and environment for statistical computing* (Version 4.3) [Software]. R Foundation for Statistical Computing. https://www.R-project.org/

Russell, W. M. S., & Burch, R. L. (1959). *The principles of humane experimental technique*. Methuen. https://norecopa.no/textbase/the-principles-of-humane-experimental-technique/

Talbot, S. R., Biernot, S., Bleich, A., van Dijk, R. M., Ernst, L., Häger, C., Helgers, S. O. A., Koegel, B., Koska, I., Kuhla, A., Miljanovic, N., Müller-Graff, F.-T., Schwabe, K., Tolba, R., Vollmar, B., Weegh, N., Wölk, T., Wolf, F., Wree, A., . . . Zechner, D. (2020). Defining body-weight reduction as a humane endpoint: A critical appraisal. *Laboratory Animals*, *54*, 99-110. https://doi.org/10.1177/0023677219883319

Talbot, S. R., Kumstel, S., Schulz, B., Tang, G., Abdelrahman, A., Seume, N., Wendt, E. H. U., Eichberg, J., Häger, C., Bleich, A., Vollmar, B., & Zechner, D. (2023). Robustness of a multivariate composite score when evaluating distress of animal models for gastrointestinal diseases. *Scientific Reports*, *13*, Article 2605. https://doi.org/10.1038/s41598-023-29623-8

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2.ª ed.). Springer. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (R package version 1.1.2) [Software]. https://CRAN.R-project.org/package=dplyr

Zhang, X., & Hartmann, P. (2023). How to calculate sample size in animal and human studies. *Frontiers in Medicine, 10*, Article 1215927. https://doi.org/10.3389/fmed.2023.1215927

CONTRIBUCIÓN DE LOS AUTORES

Concepción y diseño de la investigación: JABB, LQL, ZARH, ARGC; análisis e interpretación de los datos: JABB, LQL, ZARH, ARGC; redacción del artículo: JABB, LQL, ZARH, ARGC.

CONFLICTO DE INTERESES

Los autores declaran que no existe conflicto de intereses.