



## Cómo hacer investigación en Nutrición Clínica y no morir en el intento ¿Cómo abordar el tan temido plan estadístico?

### *How to tackle the dreaded statistical plan?*

María Ballesteros Pomar

Especialista en Endocrinología y Nutrición. Hospital Universitario de León. León



Uno de los principales obstáculos que se presentan en el desarrollo de investigaciones médicas es el diseño e implementación del plan estadístico. Sin embargo, a pesar de las dificultades que supone, la planificación estadística rigurosa es fundamental para toda investigación médica de calidad.

Siguiendo una serie de pasos básicos, es posible desarrollar un plan estadístico sólido y científicamente válido que guíe a los investigadores y permita obtener resultados confiables y significativos:

1. *Definir el problema y los objetivos.* Es prioritario y esencial identificar el problema, para lo cual se precisa plantear las preguntas o hipótesis que guiarán la investigación. Seguidamente, se deberán establecer los objetivos, es decir, definir qué se busca lograr con el análisis estadístico, ya sea probar hipótesis o describir fenómenos.
2. *Revisión de la literatura.* Se deben investigar antecedentes. Hay que buscar estudios previos relacionados con el tema para entender mejor el contexto y las metodologías utilizadas. Es importante identificar posibles "vacíos", ya que esto determina qué aspectos no han sido abordados o requieren más investigación.
3. *Diseño estadístico.* Hay que hacer una correcta y pormenorizada selección de variables: identifica las variables independientes y dependientes relevantes para tu estudio. Se deben escoger los métodos de muestreo, lo que supone decidir cómo se obtendrán los datos (muestreo aleatorio, estratificado, etc.) y el tamaño de la muestra. También se deben identificar los instrumentos de recolección: define los métodos para recopilar datos (encuestas, experimentos, bases de datos).
4. *Análisis de datos.* Se puede optar por la estadística descriptiva (utiliza medidas de tendencia central y variabilidad para describir las características de las variables), la estadística inferencial (aplica pruebas estadísticas para probar hipótesis o estimar parámetros poblacionales) y/o el análisis exploratorio (utiliza técnicas como gráficos y tablas para explorar relaciones entre variables).

*Conflicto de intereses: la autora declara no tener conflicto de interés.*

*Inteligencia artificial: la autora declara no haber usado inteligencia artificial (IA) ni ninguna herramienta que use IA para la redacción del artículo.*

Ballesteros Pomar M. ¿Cómo abordar el tan temido plan estadístico? Nutr Hosp 2025;42(N.º Extra 2):12-16

DOI: <http://dx.doi.org/10.20960/nh.06348>

©Copyright 2025 SENPE y ©Arán Ediciones S.L. Este es un artículo Open Access bajo la licencia CC BY-NC-SA (<http://creativecommons.org/licenses/by-nc-sa/4.0/>).

5. *Planificación del análisis estadístico.* Proceder a la especificación de métodos, detallando los métodos estadísticos que se utilizarán, incluyendo pruebas de significación y cómo se tratarán los datos faltantes o atípicos. El nivel de significación establece el nivel de significación estadística (*p-value*) para las pruebas. En cuanto a los modelos estadísticos, hay que considerar si se necesitan modelos más complejos para ajustar a los datos.
6. *Implementación y evaluación.* En lo que respecta a la recopilación de datos, hay que asegurarse que los datos sean representativos y estén libres de errores. El análisis de resultados interpreta los resultados en el contexto de las hipótesis y objetivos establecidos. La validación de supuestos es imprescindible, de forma que se verifique que los supuestos de las pruebas estadísticas se cumplan.
7. *Presentación y comunicación.* Se debe realizar un informe detallado; para ello, hay que presentar los resultados en un informe claro y estructurado, incluyendo gráficos y tablas. La discusión e interpretación sirve para analizar los hallazgos en relación con la literatura existente y discutir las implicaciones prácticas.

Asumiendo estas premisas, el siguiente paso es determinar cómo hacer realmente todo este trabajo. Al margen de otras muchas posibilidades, afrontar este reto se puede hacer siendo autónomo (estudiando estadística, que actualmente resulta fundamental para cualquier médico / investigador), buscando ayuda (unidades de investigación) y/o contando con el apoyo de los nuevos recursos de inteligencia artificial (como Julius.ai/chat o Powerdrill AI, una herramienta versátil que combina la automatización de tareas, la interacción en lenguaje natural y la generación de informes profesionales para simplificar y mejorar el análisis estadístico en diversas áreas de investigación).

## DEL DISEÑO, EL ANÁLISIS Y LA PLANIFICACIÓN ESTADÍSTICA

El diseño estadístico es crucial para asegurar que la investigación sea rigurosa y que los datos recopilados sean relevantes y útiles. Esta tarea se inicia con la selección de variables, identificando las variables independientes y dependientes que se estudiarán. Las variables independientes son aquellas que se manipulan o controlan, mientras que las variables dependientes son las que se miden como resultado de la manipulación. Es importante definir bien las variables antes de empezar y ajustarlas a los objetivos.

La elección de los métodos de muestreo supone decidir cómo se obtendrán los datos. Esto puede ser mediante muestreo aleatorio, estratificado, sistemático, etc. El método elegido depende del diseño de la investigación y de la población objetivo.

El cálculo del tamaño muestral es otro de elementos básicos que conforman el diseño estadístico, siendo precisa su correcta determinación en los momentos iniciales de diseño del estudio (antes de registrar al primer paciente, y no cuando ya está en plena fase de desarrollo).

Los instrumentos de recogida de datos a emplear determinan cómo se recopilarán los datos. Esto puede incluir encuestas, entrevistas, experimentos, observaciones, o el uso de bases de datos existentes. En este sentido, resulta fundamental entender que no hay estadística capaz de arreglar unos datos mal recogidos.

El análisis de datos es un paso crucial en cualquier investigación médica. Se divide en tres categorías principales: estadística descriptiva, estadística inferencial y análisis exploratorio.

La estadística descriptiva es fundamental para resumir y describir los datos recopilados. En Medicina, se utiliza para analizar resultados de estudios clínicos o epidemiológicos, lo que permite tomar decisiones informadas y contribuir al avance de la investigación en salud. Las medidas de tendencia central (media, mediana y moda) son esenciales para entender la ubicación central de un conjunto de observaciones, como los resultados de pruebas diagnósticas o la eficacia de tratamientos. Las medidas de dispersión (varianza y desviación estándar) ayudan a comprender la variabilidad de los datos, lo cual es crucial para evaluar la consistencia y confiabilidad de la información médica.

La estadística inferencial permite hacer generalizaciones sobre una población basándose en una muestra. En Medicina se utiliza para probar hipótesis sobre la efectividad de tratamientos o la asociación entre factores de riesgo y enfermedades. Las pruebas de hipótesis se utilizan para validar conclusiones sobre una población (por ejemplo, comparar la efectividad de dos tratamientos diferentes). Por su parte, los intervalos de confianza permiten estimar parámetros poblacionales, como la proporción de pacientes que responden a un tratamiento, con un margen de error conocido.

El análisis exploratorio es una etapa previa al análisis estadístico formal. Entre otras aportaciones, ayuda a identificar errores en los datos, detectar valores atípicos, y explorar relaciones entre variables. En esta fase es importante la visualización de datos, por lo que se aconseja utilizar gráficos y tablas para resumir y visualizar los datos (facilita la identificación de patrones y relaciones entre variables). Por su lado, la detección de errores y valores atípicos es crucial para asegurar que los datos sean precisos y consistentes antes de realizar análisis más complejos.

La planificación del análisis estadístico es un paso indispensable en cualquier investigación médica. Este proceso garantiza que los métodos estadísticos utilizados sean adecuados para responder a las preguntas de investigación y probar las hipótesis planteadas. La especificación de métodos, el nivel de significación y los modelos estadísticos son tres aspectos clave a considerar en la planificación del análisis estadístico para una audiencia de médicos.

Respecto a la especificación de métodos, se debe asumir la importancia que tiene determinar las pruebas estadísticas que se utilizarán para analizar los datos (por ejemplo, si se compara la efectividad de dos tratamientos, se podría utilizar la prueba *t* de Student o ANOVA —análisis de varianza—) si hay más de dos grupos). Dependiendo de la complejidad de los datos, podrían ser necesarios modelos más avanzados como la regresión lineal o logística para ajustar a las variables predictoras y la variable de

respuesta. En relación al análisis multivariante, si hay múltiples variables predictoras, se pueden utilizar técnicas como la regresión múltiple para evaluar su impacto conjunto sobre la variable de interés (Fig. 1).

El nivel de significación estadística (generalmente representado por  $\alpha$ ) es la probabilidad de rechazar la hipótesis nula cuando es verdadera. En la mayoría de los estudios médicos, se establece en 0,05, aunque también se pueden utilizar 0,01 o 0,001 para mayor rigor.

En resumen, la planificación del análisis estadístico es fundamental para garantizar que los métodos utilizados sean adecuados y que los resultados sean confiables y relevantes en el contexto médico.

IMPLEMENTACIÓN Y EVALUACIÓN

La implementación del análisis estadístico y evaluación de resultados involucran la ejecución del plan de análisis y la interpretación de los resultados.

En lo que respecta a la implementación, se deben tener en cuenta tres aspectos fundamentales: la recopilación de datos (que debe efectuarse de forma correcta y asegurando que sean representativos de la población objetivo), la aplicación de métodos estadísticos, y el manejo de datos faltantes y valores atípicos. En cuanto al análisis de resultados, el objetivo debe ser aplicar las técnicas estadísticas planificadas para analizar los datos y responder a las preguntas de investigación; se tienen que inter-

pretar los resultados (incluye evaluar la significación estadística y la relevancia clínica de los hallazgos), validar los supuestos (verificar que los supuestos de las pruebas estadísticas se cumplan) y posteriormente proceder a la discusión e interpretación clínica (discutir los resultados en relación con la literatura existente y su impacto potencial en la práctica clínica; esto incluye considerar las limitaciones del estudio y las implicaciones para futuras investigaciones).

Un aspecto importante a tener en cuenta es el manejo de datos faltantes (es crucial definir cómo se manejarán), sesgos y valores atípicos. Se deben aplicar los métodos definidos para tratar los datos faltantes y los valores atípicos, como imputación o eliminación, según sea necesario. Los valores atípicos pueden influir significativamente en los resultados; se deben definir métodos para identificarlos y decidir si se eliminan o se ajustan mediante transformaciones de datos; hay que evitar eliminar casos completos; es aconsejable emplear métodos de imputación múltiple para preservar potencia estadística. Los sesgos de selección pueden comprometer la validez externa; en este sentido, se aconseja controlar este posible problema mediante aleatorización y criterios de elegibilidad claros. Los factores de confusión distorsionan la relación exposición-resultado; en este caso, la recomendación es afrontar este posible problema mediante el análisis estratificado o multivariado.

Entre los errores más comunes que se suelen cometer en el plan de análisis estadístico, destacan ocho:

- 1. *Falta de claridad en la presentación de datos estadísticos.* No proporcionar suficiente detalle sobre los métodos esta-

| 5. PLANIFICACIÓN DEL ANÁLISIS ESTADÍSTICO: ESPECIFICACIÓN DE MÉTODOS |                                              |                                          |                                        |
|----------------------------------------------------------------------|----------------------------------------------|------------------------------------------|----------------------------------------|
| Tipo de variable                                                     | Número de grupos                             | Prueba estadística                       | Condiciones                            |
| Continua vs. nominal                                                 | 2 grupos                                     | t-test (independiente o emparejado)      | Datos normales                         |
|                                                                      | 2 grupos                                     | Mann-Whitney U (Wilcoxon-Mann Whitney U) | Datos no normales                      |
|                                                                      | 3+ grupos                                    | ANOVA (1 vía)                            | Datos normales                         |
|                                                                      | 3+ grupos                                    | Kruskal-Wallis                           | Datos no normales                      |
| Continua vs. continua                                                | 2 variables                                  | Pearson Correlation                      | Relación lineal                        |
|                                                                      | 2 variables                                  | Spearman's Rank Correlation              | Relación no lineal o datos no normales |
| Nominal vs. nominal                                                  | 2 variables                                  | Chi-squared Test                         | Asociación entre variables             |
|                                                                      | 2 variables (pequeñas frecuencias esperadas) | Fisher's Exact Test                      | Asociación entre variables             |
| Nominal (dependiente) vs. continua o nominal (independientes)        | 2 categorías                                 | Logistic Regression (binaria)            | Predicción de probabilidad             |
|                                                                      | 3+ categorías                                | Multinomial Logistic Regression          | Predicción de probabilidad             |
| Ordinal vs. nominal                                                  | 2 grupos                                     | Mann-Whitney U                           | Comparación de medias                  |
|                                                                      | 3+ grupos                                    | Kruskal-Wallis                           | Comparación de medias                  |

Figura 1. Especificación de métodos de análisis estadístico.

distintos utilizados puede dificultar la replicación del estudio y la interpretación de los resultados.

2. *No considerar la calidad de los datos.* Ignorar la calidad de los datos puede llevar a conclusiones incorrectas. Es crucial asegurarse de que los datos sean completos, precisos y representativos de la población estudiada.
3. *Mala selección de pruebas estadísticas.* Utilizar pruebas estadísticas que no se ajustan a la pregunta de investigación o a la naturaleza de los datos puede resultar en resultados engañosos.
4. *No manejar correctamente los datos faltantes.* Ignorar o eliminar datos faltantes sin un método adecuado puede introducir sesgos y reducir la potencia estadística del estudio.
5. *Sobreajuste de modelos.* Crear modelos que se ajustan demasiado bien a los datos disponibles puede resultar en un sobreajuste, lo que reduce la capacidad predictiva del modelo en nuevos datos.
6. *Cherry-picking de datos* (selección sesgada de resultados). Presentar solo los resultados que apoyan la hipótesis mientras se ignoran los que la contradicen puede llevar a conclusiones engañosas.
7. *No ajustar por comparaciones múltiples.* No corregir por comparaciones múltiples puede aumentar el riesgo de encontrar resultados estadísticamente significativos por casualidad.
8. *No validar supuestos estadísticos.* Muchas pruebas estadísticas requieren que ciertos supuestos se cumplan (como la normalidad o la igualdad de varianzas). Ignorar estos supuestos puede invalidar los resultados.

Para evitar estos errores, es crucial desarrollar un plan de análisis estadístico bien estructurado, considerar la calidad de los datos, y asegurarse de que los métodos estadísticos sean adecuados para la pregunta de investigación. Además, es importante ser transparente sobre cualquier desviación del plan original y justificar las razones para tales cambios.

La evaluación de la efectividad de una metodología estadística implica varios criterios que garantizan que los métodos utilizados sean adecuados para responder a las preguntas de investigación y probar las hipótesis planteadas. Entre estos criterios, se incluyen: la alineación con el objetivo de la investigación, el tipo y distribución de los datos, el nivel de significación y potencia del estudio, el manejo de datos faltantes y valores atípicos, la validación de supuestos estadísticos, la interpretación de resultados y los criterios de evaluación (precisión, sensibilidad y especificidad). En resumen, la evaluación de una metodología estadística debe considerar tanto la adecuación del método a la pregunta de investigación como la calidad y relevancia de los resultados obtenidos (Fig. 2).

## CONCLUSIONES

En el desarrollo de una investigación válida, resulta fundamental efectuar una planificación estadística rigurosa. Es importante contar con colaboración multidisciplinaria, estableciendo una alianza entre médicos y bioestadísticos. Actualmente, se dispone de numerosas herramientas y guías (SPSS, R, CONSORT, STROBE)



**Figura 2.**

Criterios para evaluar la efectividad de una metodología estadística (AUC-ROC: área bajo la curva).

que ayudan en la implementación de esta labor estadística, siendo muy prometedora y sugerente la aportación que ofrecen los nuevos recursos de inteligencia artificial (IA). Y, como consejo final, hay que considerar la opción de consultar con bioestadísticos desde las fases iniciales del diseño de cualquier investigación. Sin duda, un plan estadístico sólido mejora la validez y el impacto de una investigación médica.

Como mensajes finales, recordar:

1. Quien no sabe lo que busca, no entiende lo que encuentra.
2. Es preciso desarrollar un plan de análisis estadístico bien estructurado.
3. Hay que considerar la calidad de datos y asegurarse que métodos estadísticos sean adecuados para la pregunta de investigación.

4. Se deben analizar los resultados en el contexto de la pregunta de investigación.
5. Y estamos de enhorabuena, porque el futuro ya está aquí, y viene marcado fundamentalmente por el creciente papel que va a tener la IA.

#### **BIBLIOGRAFÍA RECOMENDADA**

---

1. Herzog MH, Francis G, Clarke A. Understanding Statistics and Experimental Design. How to Not Lie with Statistics. Springer Cham; 2019. DOI: 10.1007/978-3-030-03499-3
2. Díaz Portillo J. Guía Práctica del Curso de Bioestadística Aplicada a las Ciencias de la Salud. Instituto Nacional de Gestión Sanitaria. Disponible en: <https://bit.ly/3DAxBxw>