



SPECIAL ARTICLE

Critical appraisal (I): Fundamentals, tools, and application to scientific articles on outcomes and harmful effects



Javier González de Dios^{a,*}, María Salomé Albi Rodríguez^b,
Álvaro Gimeno Díaz de Atauri^b, Laura Cabrera Morente^c, Elena Pérez González^d,
Carlos Ochoa Sangrador^e, Comité de Pediatría Basada en la Evidencia¹

^a Servicio de Pediatría, Hospital General Universitario Dr. Balmis, Alicante, Universidad Miguel Hernández, Instituto de Investigación Sanitaria y Biomédica de Alicante (ISABIAL), Alicante, Spain

^b Servicio de Pediatría, Hospital Universitario 12 de Octubre, Madrid, Spain

^c Centro de Salud San Felipe, Distrito Sanitario de Jaén, Jaén, Spain

^d Servicio de Pediatría, Hospital Universitario Virgen Macarena, Sevilla, Spain

^e Servicio de Pediatría, Hospital Virgen de la Concha, Complejo Asistencial de Zamora, Zamora, Spain

Received 6 March 2025; accepted 12 June 2025

Available online 10 October 2025

KEYWORDS

Evidence-based
medicine;
Prognosis;
Observational study;
Pediatrics

Abstract Evidence-Based Medicine relies on five structured steps: formulating a question, searching for evidence, critical appraisal of the evidence, assessing applicability and evaluating performance. While all steps are essential, critical appraisal is the central component, implemented in three stages: analyzing validity or scientific rigor (including both internal and external validity), evaluating clinical relevance or importance (considering quantitative, qualitative, comparative dimensions, and the balance of benefits, risks, and costs), and determining applicability (in routine clinical practice). Among the useful tools for critical appraisal, notable resources include templates and materials from CASPe and Osteba, as well as the epidemiological calculator Calcupedev, developed by the Committee on Evidence-Based Pediatrics.

This article explores the necessary process of critically appraising studies on prognostic and risk factors through the analysis of observational studies (cohort and case-control studies).

© 2025 Asociación Española de Pediatría. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

DOI of original article: <https://doi.org/10.1016/j.anpedi.2025.503941>

* Corresponding author.

E-mail address: javier.gonzalezdedios@gmail.com (J. González de Dios).

¹ Appendix A lists the rest of the members of the Committee on Evidence-Based Pediatrics.

PALABRAS CLAVE

Medicina basada en la evidencia;
Pronóstico;
Estudios observacionales;
Pediatria

Lectura crítica (I): Fundamentos, herramientas y su aplicación en artículos científicos sobre pronóstico y efectos perjudiciales

Resumen La Medicina Basada en la Evidencia se basa en cinco pasos estructurados (pregunta, búsqueda, lectura crítica, aplicabilidad y adecuación). Todos son importantes, pero el eje central es la lectura crítica, que realizamos en tres etapas: analizar la validez o rigor científico tanto la validez interna como externa), la relevancia o importancia clínica (en las dimensiones cuantitativa, cualitativa, comparativa y de relación beneficios-riesgos-costes) y la aplicabilidad (en la práctica clínica habitual). Entre las herramientas útiles para la lectura crítica podemos destacar los recursos y plantillas de CASPe y Osteba, así como la calculadora epidemiológica Calcupedev elaborada desde el Comité de Pediatría Basada en la Evidencia.

En este primer artículo nos hemos adentrado en el necesario camino de la valoración crítica de estudios sobre factores pronósticos y de riesgo, a través del análisis de estudios observacionales (cohortes y caso-control).

© 2025 Asociación Española de Pediatría. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la CC BY-NC-ND licencia (<http://creativecommons.org/licencias/by-nc-nd/4.0/>).

Fundamentals of critical appraisal and available tools

Evidence-based medicine (EBM) is based on five structured steps (ask, acquire, appraise and apply evidence and assess its performance), as discussed in the first installment of this series.¹ The central axis is critical appraisal, which feeds on the two previous steps to answer the two that follow.

The critical appraisal of research reports requires the acquisition of specific skills and abilities. It comprises three steps: assessing whether they are *valid* (close to the truth and scientifically rigorous), deciding whether they are *important or relevant* (and therefore potentially valuable to the reader) and determining whether they are *applicable* in everyday clinical practice. This is what we, from the Committee on Evidence-Based Pediatrics, have defined as applying the VARA approach, an acronym for scientific VALIDity, clinical RElevance (or importance), and Applicability in clinical practice.

Fig. 1 presents the key aspects of these three steps.²⁻⁶ Different tools, worksheets and checklists are available to facilitate the adequate critical appraisal of published research. Table 1 lists the Spanish-language tools that are most widely used at the moment.⁷

Research reports should aim to answer clinical questions that arise in everyday practice. To this end, it is essential that the study design is appropriate for the type of question it addresses (Table 2). We invite readers to deepen and broaden their knowledge about every type of study.⁸⁻¹² In this chapter on critical reading, we will address studies on prognosis and harm, which chiefly consist of observational cohort and case-control studies. Table 3 outlines the advantages and disadvantages of these types of studies and key points regarding their validity that are discussed later in this text.

Establishing the prognosis of a health condition is one of the key purposes of medicine: knowing its possible outcomes and the frequency with which they can be expected to occur.

Table 1 Critical appraisal support tools in Spanish.**Critical Appraisal Skills Programme-Español (CASPe):**

Education on critical appraisal of scientific evidence for different learner profiles (clinicians, researchers, administrators and patients) and offers online learning resources.

<https://redcaspe.org/>

Department of Health Technology Assessment of the Basque Country (Osteba):

Free app, validated by experts in health technology assessment, that allows users to create and share their own evidence tables from several sources

<http://www.lecturacritica.com/es/>

Calcupedev:

Epidemiological calculator developed by the Committee on Evidence-Based Pediatrics

<https://aepap.org/calculadora-estudios-pbe/#/>

Table 2 Most appropriate study designs to address different types of clinical questions (from most to least appropriate).

Type of question	Study designs
Diagnosis	Prospective, blind comparison with gold standard
Treatment	RCT (prospective, blind comparison with gold standard)
Etiology/harm	RCT > cohort > case-control > case series
Prognostic	Cohort > case-control > case series
Cost-benefit	Economic evaluation

Abbreviation: RCT, randomized clinical trial.

Step 1. **VALIDITY** or scientific rigor

INTERNAL **VALIDITY**



Are the conclusions of the study correct for the patients under study?

Minimization

- Random error (due to chance)
- Systematic error/bias
 - Factors affecting the selection of patients or the exposure or intervention under study
 - Factors affecting measurement
 - Factors affecting the analysis and publication of results

Were possible confounders (known and unknown) taken into account?

Measures to minimize confounding:

- Correct randomization
- Covariate adjustment

EXTERNAL **VALIDITY**



Can the conclusions be generalized to settings different from those of the original study?

Are the conclusions applicable to my patients?

(Based on the basic elements: PICO(T))

Step 2. **RELEVANCE** or clinical significance

QUANTITATIVE assessment

Descriptive statistics of the most important outcome
Calculation of the size of the effect
(using the most appropriate impact measure)

- Therapy or prevention studies:
 - Impact measures for discrete dichotomous data: ARR, NNT, NNH
 - Impact measures for continuous data: MD
- Diagnostic studies:
difference in pre-test and post-test probability,
likelihood ratio (LR+, LR-)
- Etiological studies: AFe, PAF, impact numbers (CIN, ECIN)
- Predictive model studies: difference in a priori vs a posteriori
probability, percentage of patients that would benefit from the
results of the model (diagnostic or therapeutic decision-making)

COMPARATIVE assessment

- Search for results of other published studies
(ideally, systematic reviews or high-quality studies)
- Comparison of:
 - Primary outcome
 - Alternative outcomes

QUALITATIVE assessment

Assessment of the following:

- CLINICAL SIGNIFICANCE of the primary outcome (GRADE)
 - Critical
 - Important
 - Not important
- IMPORTANCE of the outcome. Effect size.
 - In relation to what is considered clinically significant:
 - Large
 - Moderate
 - Small
 - In relation to clinical decision-making/benefit to the
patient/impact on health care system and society

Costs assessment

- Balance of
 - Benefits
 - Risks
 - Costs

Step 3. **APPLICABILITY** of scientific evidence to clinical practice

Are the differences between our patient(s) and the study patients too significant for the findings to be useful in our practice?

Figure 1 Steps in the critical appraisal of scientific reports.

Internal validity (certainty of results based on scientific rigor, according to the study design); Abbreviations: AFe, attributable fraction among the exposed; ARR, absolute risk reduction; CIN, case impact number; ECIN, exposed cases impact number; LR, likelihood ratio; MD, mean difference, NNH, number needed to harm; NNT, number needed to treat; PAF, population attributable fraction.

Prognostic factors are patient characteristics that can be used to try to predict an outcome. Among these, risk factors are those associated with the development and onset of disease. Since prognostic factors are characteristics or habits of individuals, the studies conducted to identify them rarely have an experimental design, as allocation is usually impossible or unethical. Instead, observational designs are used under these circumstances. A correctly designed and imple-

mented cohort study is a valuable strategy for obtaining a valid estimate of the association between a prognostic factor and an outcome. Case-control studies can also be used, but the selection biases inherent in their design and their retrospective nature limit the reliability of the conclusions that may be drawn.

When the objective of a study is to investigate the potential *harmful effects* of a treatment or the risks associated

Table 3 Strengths and weaknesses of cohort studies and case-control studies. Key aspects in the assessment of validity.**Cohort study****Strengths**

- Very useful for the study of rare exposures
- Can measure multiple outcomes for a single exposure
- Can establish the temporal sequence of exposure and disease
- If the follow-up is prospective, it limits bias in the establishment of exposure
- Allows direct calculation of incidence in exposed and unexposed groups

Weaknesses

- Not suitable for investigation of rare diseases, only if attributable risk is high
- If prospective, it is costly and requires a long follow-up
- If retrospective, it depends on adequate past documentation/records
- The validity of results can be seriously affected by losses to follow-up
- The validity of results may be questioned based on biases in case ascertainment
- May be impractical when exposure varies significantly over time and has a short-lived effect on the risk of disease

Key aspects in the assessment of validity

- a) Clear definition of the population, the exposure and the outcome under study
- b) Cohorts representative of the exposed and unexposed populations, with nondifferential and sufficiently complete ($\geq 80\%$) follow-up
- c) Valid independent/blind measure of exposure
- d) Assessment of temporal sequence of events (exposure-effect) and the relationship between the level of exposure and the magnitude of the effect (dose-response)
- e) Adequate analysis (control of confounders and effect modifiers)

Case-control study**Strengths**

- Useful for diseases with a low incidence or long latency period
- Less costly than a cohort study
- Assesses several risk factors
- Does not require as long a follow-up as a cohort study
- May be used to study preventive measures

Weaknesses

- Susceptible to biases
- Cannot be used to directly estimate incidence or prevalence
- Cannot establish causation

Key aspects in the assessment of validity

- a) Clear definition of the population, the exposure and the outcome under study
- b) Case group representative of the case population and control group representative of the level of exposure in the population from which the cases originate (without the disease/effect of interest, but at risk of having it)

Table 3 (Continued)**Cohort study**

- c) Valid independent/blind measure of exposure
- d) Assessment of temporal sequence of events (exposure-effect) and the relationship between the level of exposure and the magnitude of the effect (dose-response)
- e) Adequate analysis (control of confounders and effect modifiers)

with certain exposures, questions regarding the cause-effect association can be addressed with different study designs: randomized controlled trial (RCT), cohort study, case-control study, isolated case reports, systematic review, etc. Depending on the study design chosen to determine whether a treatment is or is not harmful, the level of evidence varies based on the design-specific “threats” to the validity of the findings. To avoid rash decision-making, caution must be exerted when inferring causation.

Following the recommendations of the Evidence-Based Medicine Working Group (EBMWG), we discuss the critical appraisal of the validity, importance and applicability of prognostic and harm/risk studies, formulating a series of questions in each step (Critical Appraisal Skills Program-Español [CASPe], Spanish translation and adaptation of the original CASP checklist¹³).

Critical appraisal of prognostic studies

Assess validity

It requires asking a series of questions with two elements in common: a qualitative element (what are the possible outcomes?) and a temporal element (how long does it take them to occur?).¹⁴ We will answer screening questions, or assess primary criteria (if they are not met, it may not be worth reading the entire report) and secondary criteria (more in-depth questions asked once the primary criteria have been met).

a) Primary criteria:

Was the defined representative sample of patients assembled at a common point in the course of their disease?

Since it is impossible to include the entire population that has a given disease in a study, it is very important that the selected sample be as similar as possible to this immeasurable universe (and that it manifests the spectrum of disease that we encounter in clinical practice).

The report must specify the criteria applied to ascertain that the included patients had the disease whose prognosis was being investigated and how the sample was selected, as the findings may be distorted by biases in these processes. The selected patients should all be at a similar stage in the course of disease, which needs to be clearly defined in the report,¹⁵ and the sample size must be adequate for the study of the outcome of interest.

Was patient follow-up sufficiently long and complete?

In case-control studies, determining whether the exposure occurred prior to the target outcome is key. It is important to assess whether the researchers implemented strategies to minimize recall bias or interviewer bias. In the case of cohort studies and clinical trials, the reader must carefully consider whether strategies were implemented to ensure that the occurrence of the outcome has been ascertained through the same method and with equal probability in both groups.

The follow-up must be sufficiently long to allow time for the outcome of interest to occur and sufficiently complete in all patients. However, losses to follow-up are common. The number of losses (which can affect the accuracy of the risk estimate) and its relationship to the proportion of patients who experience the outcome of interest need to be taken into account. If the reasons for the losses are not related to the phenomenon under study and the characteristics of patients lost to follow-up are similar to those of the patients who completed it, the results are probably valid.

If the answer to any of the above questions is "no", one can be fairly certain that the study will not provide accurate prognostic estimates, whereas if the answer to all of them was "yes," one can be fairly confident that the study will provide good prognostic information.

b) Secondary criteria:

Were outcome criteria objective and applied in a 'blind' or unbiased fashion?

The authors should have defined the outcomes of interest clearly before the start of the study: these outcomes may be objective and easily measurable (death) or subjective and difficult to measure (pain intensity). It is important that outcome assessors are blinded in order to avoid observer bias.

Did adjustment for all important prognostic factors take place?

When data from patient subgroups are reported, the reader must verify that adjustments were made for other known relevant prognostic factors. Adjustments can be simple (stratified analysis) or more complex and powerful (multivariate analysis, in which the authors decide which variables to include). One of these methods must be applied to be able to tentatively conclude that the prognosis is better or worse in a given subgroup. We say "tentatively" because the statistical techniques used to determine the prognosis of subgroups are predictive, not explanatory.

Were results validated in an independent group of 'test-set' patients?

The group in which the putative prognostic factor is first identified is called the training set (derivation set). When multiple prognostic factors are evaluated, it is possible that some may have been identified by chance alone, so these findings must be validated in an independent group of subjects (the validation set).

Table 4 Prognostic studies. Methods for generation of survival curves.

Cumulative incidence	Proportion of candidates who develop the disease outcome during the follow-up
Actuarial curve	Based on life tables, which are used when the exact date of the event is unknown; used in very large samples; measures time in predetermined intervals.
Product limit or Kaplan-Meier method	Uses all the follow-up data; it is used in small samples and gives the probability of survival for each specific time point. Kaplan-Meier curves are the clearest and most comprehensive way to present prognostic data.

Assess the importance of the results

The next step is to assess the results in terms of the potential clinical significance of the findings. To this end, the reader needs to look for the probability of the event of interest, its precision and the risk associated with prognosis-modifying factors.

Did the authors specify the likelihood of outcomes in a specified period of time?

The quantification of results is based on the number of events that occur during the follow-up of a cohort and is usually expressed as a cumulative frequency at a given time point or the median survival. There are three methods for estimating this (Table 4).

How precise are the prognostic estimates?

Even when it is valid, a prognostic study only provides an estimate of the true risk. The next step is to appraise the precision of this estimate by means of the confidence interval (CI), which is a measure of the uncertainty due to random variability. The greater the precision, the more reliable the results. In most survival curves, earlier follow-up periods tend to include results from a larger number of patients compared to later periods (decreasing number of subjects available for follow-up, as patients are not included in the study at the same time); this means that survival curves are usually more precise in earlier follow-up periods, with narrower CIs at the left hand side of the curve.

Which factors can modify the prognosis?

One of the aspects that doctors, patients and families are most interested in is which factors can affect the prognosis. Once a diagnosis has been established, multiple factors related to the patient's previous health status, disease, treatment or socioeconomic circumstances can have an impact on the final outcome. Clinicians, in particular, are most interested in determining whether they can intervene

to modify these factors (through changes in diet, lifestyle, etc.). Prognostic factors are usually reported with the corresponding relative risk (RR) in cohort studies or odds ratio (OR) in case-control studies. The RR and the OR represent the change in the probability of the unfavorable outcome. Negative values of the RR or OR indicate that the factor has a protective effect.

Assess the *applicability* to clinical practice

Are the results applicable to patient care?

First and foremost, clinicians need to assess the degree of similarity between their own patients and the subjects included in the study and whether the differences are large enough to doubt whether they can apply the results to make adequate prognoses for their patients.

Are the results useful for making decisions regarding treatment and for counseling or reassuring patients or families?

Prognostic data can be very helpful in guiding therapeutic decision-making. In addition, valid, precise and generalizable results supporting a uniformly good prognosis for a disease or condition can be very useful in reassuring affected patients. On the other hand, prognostic results supporting a uniformly poor prognosis provide clinicians a starting place for discussion with the patient and family and planning care. The results of a prognostic study are applicable to the extent that the study subjects are similar to the local patients and that they are clinically significant.

Critical appraisal of harm studies

Once again, we describe the three steps to follow in the critical appraisal of studies that investigate the harmful effects of interventions and/or exposures.¹⁶

Assess validity

Did a given intervention cause a given adverse effect in some patients? To answer this, one must start by asking a series of screening questions to identify the sources that merit more in-depth reading.

a) Primary criteria:

Were patient groups similar for prognostic factors that are known to be associated with the outcome other than the exposure?

While RCTs provide the highest-quality evidence of a cause-effect relationship, they can rarely be used to evaluate potential harmful exposures due to the ethical issues it would entail (however, if a clinical trial observes an association between an exposure and an adverse effect, the result is very trustworthy).

Cohort studies are often used as an alternative, in which case it is very important that the comparability of the cohorts has been evaluated to ensure that the subjects exposed to the putative harmful factor are similar to the unexposed subjects. If differences are detected in any varia-

bles with a known association with the outcome of interest, a stratified or multivariate analysis should be performed to control for potential confounding due to their influence; but it is important to keep in mind that, these adjustments notwithstanding, there could always be confounding due to an unknown or uncontrolled factor, so the results of a cohort study should always be interpreted with caution.

When the harmful effect is very rare or takes a long time to occur, case-control studies may be used, but they are more susceptible to bias and confounding, so their results should be interpreted with even greater caution. Harmful effects can also be reported in isolated case reports or case series, but evidence from these reports is very weak. As is the case of other aspects of health care, the best evidence is obtained through the systematic review of all pertinent studies.

Were the outcomes and exposures measured in the same way in the groups being compared?

In case-control studies, the ascertainment of previous exposure is the key issue; readers should look for any strategies used to minimize recall bias or biases related to the researchers. In contrast, in cohort studies and RCTs, readers must pay attention to the measures taken to ensure that the methods and circumstances of exposure ascertainment were similar and the risk of exposure the same in both groups.

Was the follow-up sufficiently long and complete?

In RCTs and cohort studies, the follow-up should be long enough to allow for the development and detection of the harmful effect.

b) Secondary criteria:

Do the results of the harm study fulfill some of the diagnostic tests for causation?

It is unlikely that all the criteria for causation are met. The more that are met, the greater the confidence with which a causal relationship can be inferred. Clinicians should determine whether the association between the exposure and the harmful effect meets at least some of the tests for reasonable inference of causation, which are outlined in [Table 5](#).

Aside from the experimental intervention, were the groups treated equally?

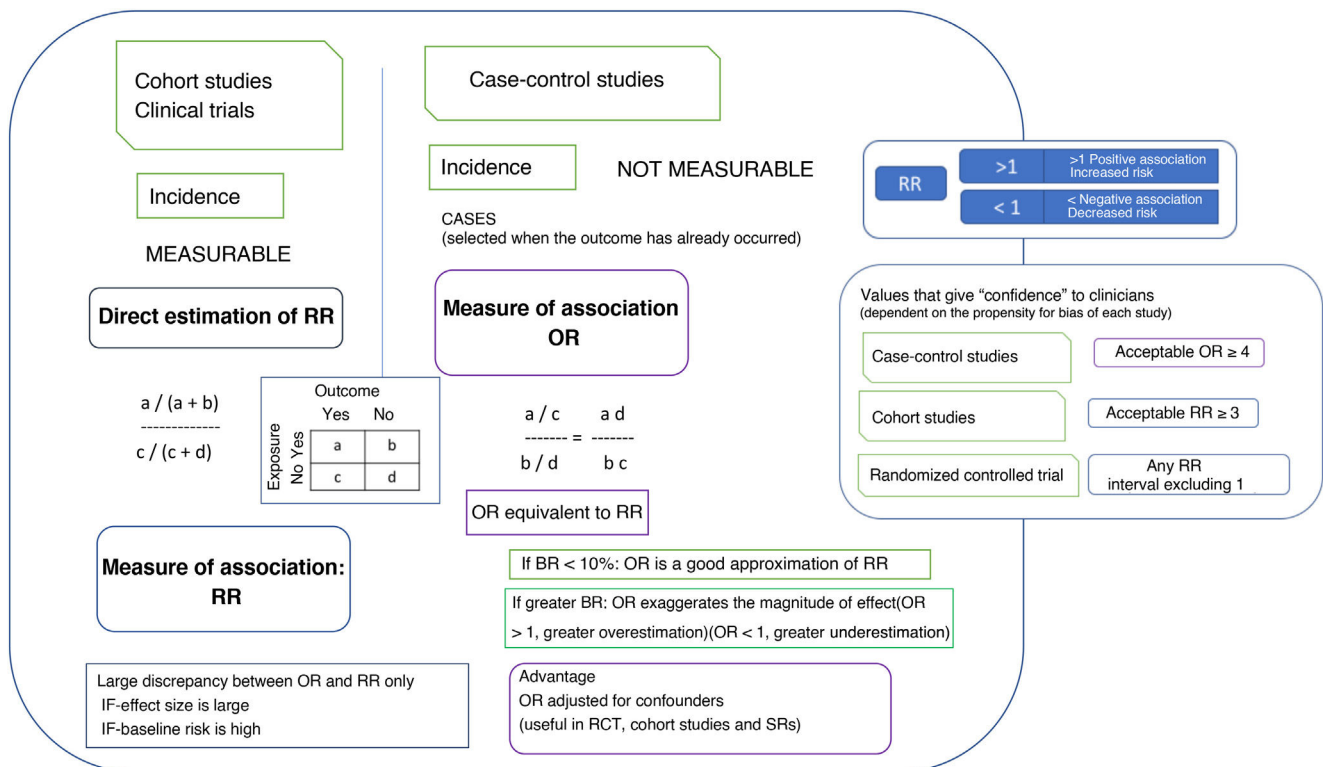
Interventions other than the experimental intervention under study (cointerventions), when implemented differentially in treatment and control groups, can affect the comparability of the groups and pose an even greater risk of bias when the study design is not double-blind or when the use of additional treatments that can affect the toxicity of the intervention under study is allowed.

Assess the *importance* of the results

This involves considering the size and precision of the estimated effect of the intervention and/or exposure.

Table 5 Harm studies. Reasonable proof of causation.

Type of question	Type of study
Did the exposure preceded the onset of the outcome?	It is important to verify that the exposure did precede the onset of the adverse outcome
Is there a dose-response gradient?	That is, the greater the intensity or duration of exposure, the greater the risk of the adverse outcome
Is there any positive evidence from a "dechallenge-rechallenge" study?	The strength of the evidence also increases if there are data showing that the adverse effect disappears or decreases if exposure is discontinued and reappears or increases if exposure resumes
Is the association consistent from study to study?	When several studies are available, a causal relationship is supported when results are consistent among studies, despite them being conducted in different places and different populations
Does the association make biological sense?	Confidence in the strength of the association increases when there is a biologically plausible explanation of the observed association that is compatible with the current knowledge of the biological effects of the exposure factor


Figure 2 Measures of association used in studies on harmful effects and their magnitude.

In most cases, interpreting the OR as the RR does not affect the qualitative interpretation of the results, although it may lead to a degree of overestimation of the magnitude of the effect.

BR, baseline risk (frequency of the outcome of interest in the reference population); OR, odds ratio (odds of exposure [probability of being exposed vs not exposed] in case group relative to the odds in the control group); RR, relative risk (incidence in the exposed group divided by incidence in the unexposed group); SR, systematic review.

How strong is the association between exposure and outcome?

The main measure that indicates that a therapeutic intervention causes an adverse outcome is the strength of the association between the implementation of the intervention and the occurrence of a harmful effect. Different methods are used in different types of study to estimate the

strength of the association. The most commonly used measure of association is the RR, which indicates the extent to which the adverse outcome is more likely to develop in the exposed group compared to the unexposed group. In case-control studies, since the incidence cannot be calculated because subjects are selected after the outcome has already occurred, the appropriate measure of association is

the OR. Another measure that can be used in cohort studies and RCTs is the difference in incidence or risk difference (RD) (an absolute measure of the excess risk of the exposed compared to the unexposed attributable to the exposure). Its inverse is a measure that indicates the number of people who need to be exposed for harm or an adverse outcome to occur, known as the number needed to harm (NNH). Another measure, similar to the relative risk reduction (RRR), is the attributable fraction in the exposed (AF_e) (proportion of the events observed in the exposed group that can be attributed to the exposure). Fig. 2 details the significance and magnitude of these measures.

How precise is the estimate of the risk?

The CI offers a measure of the precision with which the population parameter has been estimated based on the simple point estimate obtained from a sample of patients. Confidence intervals can be calculated for most statistical estimates or comparisons (OR, RR, RRR, ARR, NNT, NNH). Although readers can calculate CIs from the results presented in the study, it is preferable that the authors provide them in the report. If the CI contains the null value (1 in the case of the RR and the OR), the estimate of risk may be disregarded.

Assess applicability to clinical practice

Are the results applicable to patient care?

The first key issue is assessing whether our patients differ from studied patients in terms of morbidity, age, race, previous interventions or exposures or other important characteristics, in order to determine whether the study results can be extrapolated to our clinical practice.

What is the magnitude of the risk?

The next step is determining the patient's individual risk of experiencing the harmful effect. A simple way to do this is to express it as a fraction of the risk of the study subjects, assigning the patient a subjective value based on the patient's clinical condition.

Should clinicians attempt to stop exposure?

Once all the information has been obtained, clinicians must decide on the appropriate course of action based on the strength of the inference, the magnitude of the risk if exposure continues, the availability of alternatives, and, obviously, the patient's values and preferences, combined with the expertise of the provider that is making the decision.

Appendix A. Rest of members of the Committee on Evidence-Based Pediatrics

Pilar Aiuzpurúa Galdeano, María Aparicio Rodrigo, Nieves Balado Insunza, Albert Balaguer Santamaría, Carolina Blanco Rodríguez, Fernando Carvajal Encina, Jaime Javier Cuervo Valdés, Eduardo Cuestas Montañes, María Jesús Esparza Olcina, Sergio Flores Villar, Garazi Fraile Astorga, Paz González Rodríguez, Rafael Martín Masot, María Victoria Martínez Rubio, Manuel Molina Arias, Eduardo Ortega Páez,

Begoña Pérez-Moneo Agapito, María José Rivero Martín, Juan Ruiz-Canela Cáceres.

Declaration of competing interest

The authors have no conflicts of interest to declare.

References

- Martín-Masot R, Ochoa Sangrador C, Cuervo Valdés JJ, Blanco Rodríguez C, González Rodríguez P, González de Dios J, en representación del Comité de Trabajo de Pediatría Basada en la Evidencia. Evidence-based medicine: 5 steps to navigate uncertainty. *An Pediatr (Engl Ed)*. 2025;103(2):503930, <http://dx.doi.org/10.1016/j.anpede.2025.503930>. PMID: 40784829.
- González de Dios J, Martín Masot R, Ochoa Sangrador C. Tercer paso de la MBE: valoración crítica de la literatura científica. ...eje central de la Medicina basada en la evidencia. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 67–79.
- Gimeno Díaz de Atauri A, Albi Rodríguez MS. Generalidades sobre lectura crítica. Importancia clínica. Herramientas de lectura crítica. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 411–20.
- González de Dios J, Buñuel Álvarez JC, Ochoa Sangrador C. La valoración crítica de documentos científicos y su aplicabilidad a la práctica clínica: aspecto clave en la toma de decisiones basada en las mejores pruebas científicas. *Evid Pediatr*. 2006;2:37.
- Aparicio Rodrigo M, Pérez González E. GRADE. Valoración de la calidad de la evidencia y fuerza de las recomendaciones. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 397–408.
- González de Dios J, Ochoa Sangrador C, Martín Masot R. Cuarto y quinto paso de la MBE: aplicabilidad y adecuación de la evidencia científica a la práctica clínica. ...evitar nadar para morir en la playa frente a la variabilidad sanitaria. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 81–9.
- Pérez-Moneo Agapito B, Ortega Páez E. Calcupedev. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 683–700.
- González de Dios J, Ruiz-Canela Cáceres J, Ochoa Sangrador C. Lectura crítica de artículos científicos sobre pruebas diagnósticas. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 433–43.
- Cabrera Morente L, Gimeno Díaz de Atauri A. Lectura crítica de artículos científicos sobre reglas de predicción clínica. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 481–9.
- Cuervo Valdés JJ, Gimeno Díaz de Atauri A. Lectura crítica de artículos científicos sobre estudios ecológicos. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 491–8.
- Ruiz-Canela Cáceres J, Pérez González E. Lectura crítica de artículos de artículos científicos sobre estudios de evaluación económica. In: *Medicina basada en la evidencia. Lo que siem-*

- pre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento. Madrid: Lúa Ediciones; 2024. p. 499–506.
12. Albi Rodríguez MS, Pérez González E, Molina Arias M. Lectura crítica de artículos de otros tipos de metanálisis. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 471–9.
13. Programa de habilidades en lectura crítica español (CASPe). Redscape [Online] [Accessed 19 February 2025]. Available from: <https://redcaspe.org/materiales/>.
14. González de Dios J, Ibáñez Pradas V, Modesto i Alapont V. Evaluación de artículos científicos sobre pronóstico. *Evid Pediatr*. 2007;3:81.
15. González de Dios J, Albi Rodríguez MS. Lectura crítica de artículos científicos sobre pronóstico. . . entre los factores de riesgo y factores pronóstico. In: *Medicina basada en la evidencia. Lo que siempre quiso saber sobre la evidencia aplicada a la práctica clínica sin morir en el intento*. Madrid: Lúa Ediciones; 2024. p. 453–9.
16. González de Dios J, Cabrera Morente L. Lectura crítica de artículos científicos sobre efectos perjudiciales de intervenciones y/o exposición. . . de la causalidad al número de pacientes a dañar. Madrid: Lúa Ediciones; 2024. p. 445–52.