Scientific Tasks in Biomedical and Oncological Research: Describing, Predicting, and Explaining

Víctor Juan Vera-Ponce[1,2,*], Fiorella E. Zuzunaga-Montoya[3], Luisa Erika Milagros Vásquez-Romero[1], Nataly Mayely Sanchez-Tamay[1,2], Joan A. Loayza-Castro[1] and Carmen Inés Gutierrez De Carrillo[1,2]

[1]*Instituto de Investigación de Enfermedades Tropicales, Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM), Amazonas, Perú*

[2]*Facultad de Medicina (FAMED), Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM), Amazonas, Perú*

[3]*Universidad Continental, Lima, Perú*

Abstract: The traditional classification of studies as descriptive and analytical has proven insufficient to capture the complexity of modern biomedical research, including oncology. This article proposes classification based on scientific tasks that distinguish three main categories: descriptive, predictive, and explanatory. The descriptive scientific task seeks to characterize patterns, distributions, and trends in health, serving as a foundation for highlighting disparities and inequities. The predictive scientific task focuses on anticipating future outcomes or identifying conditions, distinguishing between diagnostic (current) and prognostic (future) predictions, and employing multivariable models beyond traditional metrics like sensitivity and specificity. The explanatory scientific task aims to establish causal relationships, whether in etiological studies or treatment effect studies, which can be exploration or confirmatory, depending on the maturity of the causal hypothesis.

Differentiating these scientific tasks is crucial because it determines the appropriate analysis and result interpretation methods. While research with descriptive scientific tasks should avoid unnecessary adjustments that may mask disparities, research with predictive scientific tasks requires rigorous validation and calibration, and study with explanatory scientific tasks must explicitly address causal assumptions. Each scientific task uniquely contributes to knowledge generation: descriptive scientific tasks inform health planning, predictive scientific tasks guide clinical decisions, and explanatory scientific tasks underpin interventions. This classification provides a coherent framework for aligning research objectives with suitable methods, enhancing the quality and utility of biomedical research.

Keywords: Biomedical Research, Causality, Forecasting, Biostatistics, Epidemiology, Research Design (source: Mesh).

INTRODUCTION

Historically, epidemiology has classified studies into two broad categories: descriptive and analytical [1]. However, this dichotomy has proven inadequate to capture the complexity and diversity of modern biomedical research. As Hernán *et al*. [2], note, current biomedical research requires a more sophisticated conceptual framework that better reflects its objectives and methods.

The primary limitation of the traditional classification lies in its failure to recognize that many contemporary studies combine both descriptive and analytical elements and, more importantly, that research objectives can differ significantly even within the same study design [3]. For example, a cohort study can describe a disease's natural history, predict individual clinical outcomes, or estimate the causal effects of exposure.

In response to these limitations, authors like Shmueli [4] and Breiman [5] have proposed a classification based on scientific tasks, distinguishing between explanation, prediction, and description. This framework has since been adapted and refined for specific applications in biomedical research by Wynants *et al*. [6] and van Calster *et al*. [7], demonstrating its usefulness in personalized medicine and clinical decision-making.

This article presents an updated perspective on the classification of biomedical research based on scientific tasks, providing a conceptual framework to help researchers better align their research objectives with appropriate methods and corresponding evaluation metrics. Before delving into this topic, we will also review a recent issue concerning the association or relationship between variables. Thus, the classification has theoretical and practical implications fundamental to biomedical studies' design, analysis, and reporting.

*Address correspondence to this author at the Instituto de Investigación de Enfermedades Tropicales, Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM), Amazonas, Perú;
E-mail: vicvepo@gmail.com

ISSN: 1929-2260 / E-ISSN: 1929-2279/24

ASSOCIATED FACTOR/RISK: MEANING AND CONTROVERSIES

Simply identifying that two variables are associated holds limited clinical value if the context and purpose of the association are not specified [8,9]. When researchers report an association without clarifying the conceptual framework they are working within, their findings may need more practical utility or remain at a purely descriptive level without practical application. This issue is particularly relevant in current biomedical literature, where associations are often reported without clarifying their nature or potential clinical utility [10]. For example, finding that a biomarker is "associated" with a disease reveals little about its real utility: Can it be used for early diagnosis? Does it have predictive value? Does it suggest a causal mechanism that could be targeted for therapeutic interventions? Answers to these questions require a clear conceptual framework from the study design phase onward. In modern biomedical research, we can distinguish four fundamental types of associations, each with different implications and applications: diagnostic, prognostic, etiological, and treatment effects [9].

In the diagnostic context, an association represents a contemporaneous relationship measured at a single time point. For example, when a physician observes ST-segment elevation on an electrocardiogram, this association allows them to identify an ongoing myocardial infarction. Similarly, in oncology, the combination of a significantly elevated PSA with suspicious findings on a digital rectal exam suggests the presence of prostate cancer requiring immediate evaluation [11]. This diagnostic relationship does not imply causation or future prediction; it serves as a tool for immediate identification of a present condition.

Prognostic associations, on the other hand, have a predictive and temporal nature, where a factor precedes the outcome. The glycated hemoglobin (HbA1c) case illustrates the difference between prediction and causation. Elevated HbA1c is a major prognostic factor for amputation in diabetic patients, making it an excellent predictor of future outcomes [12]; however, it is not HbA1c that causes amputation but rather a marker of prolonged poor glycemic control. Similarly, in oncology, elevated lactate dehydrogenase (LDH) levels in patients with metastatic melanoma are a significant prognostic factor for reduced survival. However, LDH itself is not the cause of death but rather a marker of tumor aggressiveness [13]. This example shows how a factor can have strong predictive value without necessarily being the direct cause of the outcome.

Conversely, etiological associations focus on understanding the underlying mechanisms explaining a phenomenon. In diabetes, for example, understanding how sustained hyperglycemia damages vascular endothelium and peripheral nerves, leading to diabetic neuropathy and vasculopathy, represents an etiological association explaining the pathophysiological mechanism of diabetic complications. In oncology, understanding how EGFR mutations in lung cancer lead to uncontrolled cellular proliferation represents an etiological association explaining how this specific type of cancer develops and progresses.

When discussing associations in the context of treatment effects, we specifically refer to causal relationships that can be modified through interventions. This distinction is crucial as it implies that deliberate changes in one variable (the intervention) will produce predictable changes in another (the outcome). For example, using EGFR inhibitors in lung cancer patients with activating mutations causally reduces tumor burden and improves survival. This causal relationship has been established through randomized controlled trials and can be consistently replicated in clinical practice. It is important to note that not all associations imply treatment effects: while EGFR mutations are associated with treatment response both predictively and causally, other markers, such as C-reactive protein, although good prognostic predictors in cancer, are not direct therapeutic targets, as modifying them does not necessarily change the disease course [14].

In public health and social epidemiology research, another crucial type of association deserves special attention: associations revealing health disparities and inequities. This type of association examines how health outcomes differ across social, ethnic, geographic, or economic groups. Unlike other types of associations, the primary objective here is not diagnostic, prognostic, etiological, or treatment-related but to make visible and quantify health disparities among various population groups. For example, the association between socioeconomic level and maternal mortality or differences in diabetes prevalence among ethnic groups are associations revealing structural inequities within the healthcare system [5]. Another example would be the marked differences in breast cancer survival between patients with and without access to targeted therapies or disparities in colorectal cancer screening rates among different ethnic and socioeconomic groups, associations that reveal structural inequities within the healthcare system [15].

These associations are fundamental to informing public policy and corrective actions to reduce health inequities.

Given these distinctions, researchers should be explicit about their objectives and use precise terminology that reflects their true research goals. When causality is the aim, especially in observational studies, researchers can use explicit causal language (e.g., 'causal effect', 'etiologic factor', 'treatment effect') or alternative terms that maintain causal intent while acknowledging methodological limitations, such as 'effect', 'impact', 'determinant', 'contributing factor', or 'influence'. These terms avoid ambiguous language like 'association' or generic 'risk factor' while appropriately reflecting methodological constraints. The methodology should clearly explain proposed causal mechanisms, and the discussion should address causal implications [16-18].

When the goal is diagnostic, prognostic, or descriptive, this should also be clearly stated using appropriate terms ('diagnostic factor', 'prognostic factor', 'health disparity indicator'). This clarity enables better evaluation of methodology and findings, more appropriate statistical analyses, and clearer translation of results into practice. For instance, a study investigating whether obesity causes cardiovascular disease should explicitly state this causal objective rather than describing it as an 'investigation of associations'. Similarly, a study developing a prognostic model should clearly state its predictive aim rather than using ambiguous terminology about 'risk factors'. This transparency about research objectives allows readers to properly assess whether the chosen methods and analyses align with the stated goals [16-18].

Descriptive Scientific Tasks in Health Research: Definition, Scope, Objectives, and Analysis

Epidemiological studies with a descriptive scientific task are fundamental in biomedical research, although their value and complexity have often been underestimated [19]. Beyond mere data collection, these studies constitute a systematic approach to characterizing biomedical phenomena, identifying

Table 1: Types of Associations in Biomedical Research: Characteristics and Applications

Type of Association	Temporality	Main Objective	Example	Practical Utility	Key Characteristics
Diagnostic	Cross-sectional (current moment)	Immediate identification of conditions	ST elevation → Current acute myocardial infarction Pathogenic BRCA1/2 + Family history → Hereditary breast and ovarian cancer syndrome	Diagnostic	- Does not imply causality or future prediction - Instantaneous relationship - Useful for immediate decisions
Prognostic	Temporal (factor precedes outcome)	Anticipate future outcomes	Elevated HbA1c → Future risk of amputation Elevated Ki-67 in breast cancer → Higher risk of recurrence	Risk stratification Follow-up planning Prevention	- Not necessarily causal - Predictive approach - Anticipatory value
Etiological	Mechanistic (explains processes)	Understand etiopathogenic mechanisms	Sustained hyperglycemia → Endothelial damage and neuropathy EGFR mutation → Constitutive activation of signaling pathways → Uncontrolled proliferation	Understanding pathogenesis Intervention development Basic research	- Explains causal processes - Basis for interventions - Scientific foundation
Treatment Effect	Modifiable (intervention → outcome)	Evaluate intervention impact	ACE inhibitor → Blood pressure reduction Osimertinib → EGFR mutation inhibition → Tumor reduction	Treatment guidance Therapeutic decisions Intervention evaluation	- Modifiable causal relationship - Action-oriented - Basis for interventions
Disparities and Inequities	Structural (systematic differences)	Highlight health gaps	Socioeconomic level → Maternal mortality Socioeconomic status → Access to innovative cancer therapies	Inform public policy Identify inequities Social intervention planning	- Avoids adjustments that obscure disparities - Focus on group differences - Basis for corrective actions

emerging patterns, and generating hypotheses for future research [20].

The descriptive scientific task is a systematic investigation designed to characterize health phenomena' distribution, magnitude, and patterns within specific populations [21]. However, their scope extends beyond simple case enumeration or frequency calculation. As noted by Conroy and Murray [22], this approach can provide crucial information on disease burden, identify vulnerable groups, and reveal temporal or geographical patterns that may suggest underlying causal factors. Additionally, they highlight how adjustment for confounders in descriptive studies is unnecessary and potentially detrimental. Lesko and Zalla [23] also emphasize the importance of conducting descriptive studies continuously, highlighting the study by Tordoff *et al*. [24] as a representative example of this practice.

The objectives of descriptive scientific tasks include the following:

1. Characterization of patterns and data structure: The first objective involves identifying and quantifying patterns in biomedical data. This includes estimating frequency measures (such as prevalence and incidence), central tendency and dispersion measures, and characterizing distributions [10]. For example, the Global Burden of Disease study represents one of the most comprehensive descriptive efforts, providing detailed estimates of the global distribution of diseases and risk factors [25].

2. Identification of temporal and spatial trends: A second crucial objective is to describe how health phenomena vary over time and space. Time series and spatial analyses enable the identification of patterns essential for public health planning and hypothesis generation. Epidemiological surveillance systems are paradigmatic examples of this type of descriptive analysis.

3. Exploration of associations and hypothesis generation: Exploratory data analysis is a core component of descriptive research. This process enables the identification of potential associations between variables and generating hypotheses that can later be tested in predictive or explanatory studies. However, it is essential to distinguish between identifying associations and inferring causality [9].

4. Identification and quantification of disparities and inequities: An often overlooked but important objective is visualization, allowing descriptive studies to reveal health disparities among different population groups. Braveman *et al*. [26] argue that this type of analysis is crucial for identifying systematic health inequities based on social, economic, geographic, or ethnic characteristics. For example, studies that describe maternal mortality differences between racial groups or variations in access to healthcare services across socioeconomic regions provide critical evidence for developing equitable health policies [27].

Research Designs with Descriptive Scientific Tasks

Research designs commonly used for descriptive scientific tasks include classic descriptive studies such as cross-sectional and longitudinal designs for estimating prevalence or incidence, ecological studies for population-level patterns, and economic studies for resource utilization analysis. However, it is a common misconception to assume that certain study designs are inherently linked to specific objectives. Analytical cross-sectional, case-control, and cohort studies can also be used for descriptive purposes, depending on the primary research objective. For instance, a case-control study on type 2 diabetes may be used descriptively to characterize the clinical-metabolic profile of patients with and without microvascular complications without aiming to establish causality. Similarly, a cohort study on breast cancer can be employed to describe the natural history of the disease and recurrence patterns over time without focusing on estimating causal effects. In another example, a classic analytical cross-sectional study in mental health could detail the distribution and patterns of psychiatric comorbidities across socioeconomic groups without attempting to establish predictions or causal relationships.

Therefore, while epidemiological design traditionally categorizes studies as descriptive or analytical, considering the scientific task approach helps us better understand the research objective and how data should be analyzed and interpreted. This perspective allows investigators to leverage various study designs to address questions with descriptive scientific tasks, regardless of whether the design is traditionally considered descriptive or analytical [20].

Statistical Analysis in Descriptive Scientific Tasks

As previously noted, statistical analysis in descriptive studies is often affected by a problematic tendency toward unnecessary variable adjustment,

stemming from the mistaken belief that more sophisticated analysis is necessarily better [22]. Statistical adjustment should have a clear purpose based on the study objective. In descriptive research that characterizes health disparities, adjustment can paradoxically obscure the inequities that must be highlighted. For example, suppose a study on racial differences in access to health services adjusts for socioeconomic status. In that case, it may "statistically explain" a disparity that reflects a structural injustice that should be recognized and addressed.

In line with this, Hernán *et al*. [28] argue that using adjusted models can distort the reality communities face when the goal is to describe the disease burden in different population groups. An illustrative case is maternal mortality studies: adjusting for factors such as access to prenatal care could mask actual differences between geographical regions or social groups—differences crucial for public health planning and resource allocation. When describing population characteristics, unadjusted analyses are typically more transparent and appropriate, as they reflect the actual reality experienced by different groups.

Moreover, it is important to clarify that this approach is not limited to simple univariate or bivariate analyses. Regression models can be valuable tools for estimating association measures and characterizing relationships between variables when pursuing descriptive objectives. For example, a generalized linear model with a log-binomial link can characterize diabetes prevalence across different age groups, or a linear regression model can be applied to describe blood pressure patterns across the body mass index spectrum. The key is to maintain coherence between the analytical methods and the intended descriptive purpose.

Thus, using regression models in research with descriptive scientific tasks is entirely valid, provided the model's complexity aligns with the descriptive objective. While regression models can be valuable tools for describing patterns and relationships, researchers should avoid creating unnecessarily complex models or performing adjustments that could obscure the very patterns they aim to describe. The key is maintaining methodological coherence with the descriptive scientific task.

LIMITATIONS OF STUDIES WITH A DESCRIPTIVE SCIENTIFIC TASK

Despite their essential role in biomedical research, research with descriptive purposes faces specific methodological challenges that require careful consideration. One key limitation in cross-sectional data collection is the challenge of temporal ordering - while not aiming to establish causality, the inability to determine the sequence of events can lead to ambiguous interpretations of observed patterns. Additionally, while the general principle is to avoid unnecessary adjustments that could mask disparities, researchers sometimes face legitimate questions about potential confounding that require a balanced approach. For instance, when characterizing health outcomes across different populations, adjustment decisions should consider the risk of obscuring important disparities against the need to account for fundamental demographic differences. Furthermore, changes in disease definitions or diagnostic criteria may affect the descriptive approach over time, particularly in longitudinal designs, which can impact trend analyses and pattern identification. Understanding these limitations helps ensure the appropriate implementation and interpretation of research with descriptive objectives within its intended scope.

PREDICTIVE SCIENTIFIC TASK IN HEALTH RESEARCH: FUNDAMENTALS AND DIFFERENCES

Predictive scientific tasks represent a paradigm in biomedical research, aiming to develop and validate tools that allow for accurate predictions of present or future events [29]. Unlike explanatory studies that seek to understand causal mechanisms, predictive studies focus on the ability to anticipate outcomes with the highest possible accuracy, regardless of the underlying mechanistic understanding.

A fundamental feature distinguishing predictive studies from other approaches is their adherence to the "black box" principle, a concept popularized by Leo Breiman [5]. In this paradigm, the emphasis is on predictive accuracy rather than interpretability or understanding of the underlying mechanism. While explanatory models aim to estimate parameters representing causal relationships, predictive models may include variables that do not have a direct causal relationship with the outcome as long as they enhance predictive accuracy [4]. For example, a predictive model for sepsis may include variables like heart rate and body temperature, not because they "cause" sepsis but because their combination improves the ability to identify at-risk patients.

This distinction is crucial as it frees predictive studies from the constraints imposed by the need for causal interpretation. As van Calster *et al*. [7] and Steyerberg [29] argued, a successful predictive model does not need to explain why it works; it only needs to

function consistently and reliably within the target population. This perspective has facilitated the adoption of complex machine learning techniques in predictive medicine, where algorithms like neural networks can capture subtle predictive patterns that would be challenging to model using traditional theory-based approaches.

Types of Predictive Scientific Task: Diagnostic and Prognostic

Prediction in biomedicine can be oriented along two distinct temporal dimensions. The first is contemporary prediction, focused on screening and diagnosis, where the goal is to identify a condition that is present but not directly observable. For example, diagnostic prediction models for deep vein thrombosis combine various signs, symptoms, and risk factors to estimate the current probability of the disease. The second dimension is future or prognostic prediction, which aims to anticipate events or outcomes that have not yet occurred. The 10-year cardiovascular risk prediction model exemplifies this prognostic scientific tasks [30].

Prognostic research, following the PROGRESS framework (PROGnosis RESearch Strategy), can be classified into four fundamental types [31]:

1. Overall Prognosis (Type I): This describes the natural course and typical outcomes of a health condition in a defined population, such as the 5-year survival rate in patients with lung cancer.

2. Prognostic Factors (Type II): Identifies specific characteristics influencing outcomes, ranging from simple variables (age, body mass index) to complex molecular markers.

3. Prognostic Models (Type III): Combines multiple prognostic factors to predict an individual's risk of a future outcome. The Framingham model is a classic example.

4. Stratified Prognosis (Type IV): Investigates how prognostic factors or models can guide therapeutic decisions by identifying which patients may benefit most from specific treatments.

Research Designs in Predictive Scientific Tasks

The timing of the prediction fundamentally determines the choice of study design in predictive research. For diagnostic and screening models, where the objective is to predict a present but not directly observable condition, analytical cross-sectional studies are the most appropriate design [32]. In these studies, potential predictors and the outcome of interest are measured simultaneously at a specific time. However, there may be a short technical interval between predictor measurement and diagnostic confirmation via the gold standard. For example, in developing a model for predicting acute appendicitis, data on symptoms, signs, and laboratory tests would be collected simultaneously with surgical or pathological confirmation of the diagnosis. However, cross-sectional studies for diagnostic prediction have important limitations. They cannot capture temporal changes in clinical manifestations that might affect diagnostic accuracy, potentially missing early disease markers or dynamic progression patterns. Additionally, these designs may not account for variations in test performance across different disease stages. For instance, in early cancer detection, diagnostic accuracy often varies depending on tumor progression, a limitation that cross-sectional designs cannot adequately address. These constraints highlight the importance of complementary longitudinal validation when developing diagnostic prediction tools.

Diagnostic prediction has diverse practical applications across clinical specialties. For instance, in emergency medicine, models combining clinical symptoms, vital signs, and biomarkers help identify high-risk sepsis patients requiring immediate intervention. In oncology, diagnostic prediction models integrate imaging features, molecular markers, and clinical data to classify suspicious lesions, such as the BI-RADS system for breast imaging or the Lung-RADS for pulmonary nodules. These tools support standardized risk assessment and guide clinical decision-making regarding the need for invasive procedures

On the other hand, predictive models require cohort studies, as their goal is to predict future events [33]. These longitudinal designs allow for a clear temporal sequence between baseline predictor measurement and the occurrence of the event of interest during follow-up. A classic example is the development of cardiovascular risk prediction models, where risk factors are measured at baseline, and follow-up is conducted to document cardiovascular events. However, it is crucial to recognize that the validity of these studies depends not only on the chosen design but also on the quality of its implementation, including sample representativeness, measurement standardization, and appropriate handling of follow-up and attrition [34].

Predictive Versus. Explanatory Scientific Task

Observing how predictive models can reverse the biological causal sequence is interesting, illustrating the distinction between prediction and causation. A paradigmatic example is the relationship between jaundice and gallbladder cancer. From a pathophysiological and causal perspective, we know that cancer causes jaundice through bile duct obstruction. However, from diagnostic predictive scientific tasks, jaundice becomes a valuable predictor for detecting cancer, even though it is a late manifestation of the disease [35]. This apparent "reversal" of the temporal-causal relationship is seen in many other clinical contexts: dyspnea is caused by heart failure, but in predictive models, dyspnea helps identify heart failure; a brain tumor causes a headache, but predictively, headache helps to suspect and detect cancer.

This fundamental distinction between prediction and causation underscores why predictive models can and should include variables that are consequences rather than causes of the condition of interest if they improve the model's predictive ability. This flexibility in predictor selection, free from causal temporal constraints, is one of the reasons predictive models can achieve high diagnostic performance, even when including late manifestations of the disease.

Statistical Analysis in Scientific Tasks

Predictive models are typically constructed using various regression techniques, chosen based on the nature of the outcome of interest. Logistic regression is the most commonly used method for diagnostic models, where the result is generally binary (presence/absence of disease) [29]. For prognostic models, where time to event is crucial, Cox regression is the gold standard, allowing for the incorporation of censoring and variable follow-up times. For example, Cox models can integrate multiple prognostic variables in cancer survival prediction while properly handling incomplete follow-up for some patients [36].

However, these traditional models are increasingly complemented or, in some cases, replaced by more advanced techniques. Generalized additive models allow for modeling nonlinear relationships between predictors and outcomes. Machine learning techniques, such as random forests, gradient boosting, or neural networks, are gaining popularity, especially when large datasets are available or when relationships between variables are complex [37]. Nevertheless, as Wynants *et al*. [6] note, the choice of statistical method should be based not only on predictive performance but also on practical considerations, such as model interpretability and ease of implementation in clinical practice.

Multivariable models address key limitations of traditional single-variable approaches by capturing complex interactions between predictors and accounting for their relative contributions to risk assessment. These models can better handle patient heterogeneity and provide more nuanced risk stratification than individual markers, leading to more accurate and personalized predictions.

Additionally, traditional predictive evaluation concepts such as sensitivity, specificity, predictive values, and likelihood ratios, though widely used, have significant limitations highlighted by modern research. The issue is that these indicators mistakenly assume that the performance of a test is constant across all patients when it varies based on individual patient characteristics and clinical context [38]. For example, the sensitivity and specificity of troponin for diagnosing myocardial infarction vary according to patient age, time since pain onset, and renal disease presence [39]. This reality has driven a paradigm shift toward multivariable predictive models integrating multiple predictors. This updated approach of using models rather than individual variables allows for more accurate and personalized prediction by considering multiple patient characteristics simultaneously. For instance, instead of relying solely on D-dimer for diagnosing deep vein thrombosis, current models combine this biomarker with clinical features, risk factors, and other findings to provide more accurate and personalized probability estimates. This shift from fixed performance measures to dynamic predictive models represents an essential evolution toward more precise, patient-centered medicine.

Furthermore, model performance assessment requires a multidimensional approach examining three fundamental aspects: discrimination, calibration, and overall clinical performance.

1. Discrimination assesses the model’s ability to distinguish between individuals who will or will not experience the event of interest. The C-statistic (equivalent to the area under the ROC curve) is the most common measure, where 0.5 indicates random prediction and 1.0 indicates perfect discrimination. However, as Steyerberg EW [29], notes, discrimination alone is insufficient, as a model may discriminate well but be poorly calibrated.

2. Calibration, often considered the "Achilles' heel" of clinical prediction, evaluates the concordance between predicted probabilities and observed event frequencies. A well-calibrated model predicting a 20% risk should observe the event in approximately 20 out of every 100 patients with that prediction. Calibration can be assessed using calibration plots, Hosmer-Lemeshow tests (though increasingly discouraged), and more modern measures such as calibration slope and intercept [40].
3. Overall Clinical Performance is evaluated using measures such as net benefit and decision curves, which consider the clinical impact of model-based decisions. These measures are crucial because a model may have excellent discrimination and calibration but still be clinically useless if its decision thresholds do not align with real-world clinical practice [41]

EXPLANATORY STUDIES: THE OFTEN-UNSTATED SEARCH FOR CAUSALITY

Research with explanatory scientific tasks seeks to establish causal relationships between health exposures and outcomes. Unlike descriptive and predictive approaches, explanatory research aims to answer the "why" behind health phenomena, aspiring to understand the underlying mechanisms connecting causes and effects. In clinical practice, this approach often falls into two major areas: etiological research, which aims to identify disease causes, and treatment effect evaluation, which examines the causal impact of therapeutic interventions [8,9].

Explanatory research employs various methodological approaches depending on the research question and context. Randomized Controlled Trials (RCTs) are considered the gold standard for causal inference due to their ability to control for known and unknown confounders through randomization. However, well-designed observational studies, particularly prospective cohort studies, can also provide strong causal evidence when RCTs are not feasible or ethical. While less robust for causal inference, case-control studies may be valuable in rare outcomes or initial causal exploration. Each design offers different strengths for causal investigation: RCTs excel in establishing treatment effects, cohort studies are particularly useful for studying long-term exposures and multiple outcomes. In contrast, case-control studies can efficiently explore potential causal factors in rare diseases.

The distinction between etiology and treatment effects is critical, as it implies different methodological approaches and practical considerations. While etiological studies often investigate nonmodifiable risk factors or long-term exposures (such as the relationship between smoking and lung cancer), treatment effect studies focus on specific, modifiable interventions (such as the efficacy of a new drug). However, both share the fundamental goal of establishing causal relationships that can inform clinical practice and public health policies [8,9].

Reluctance to Discuss Causality

Explanatory scientific task represent a unique paradigm where a peculiar dichotomy exists: while most researchers implicitly seek to establish causal relationships, there is widespread reluctance to state these intentions explicitly [2,9]. More concerning is that many researchers appear unaware that they are conducting causal research, even though their methods suggest it: they use directed acyclic graphs (DAGs), adjust for confounders, and assess interactions and mediation—all tools from the causal inference toolkit. These disconnects between the methods employed and the true research objective can lead to inadequate interpretations and ambiguous conclusions. Unfortunately, this lack of conceptual clarity affects research quality and complicates the critical evaluation of generated evidence, as researchers conduct causal analyses without acknowledging or fully understanding it.

The reasons why most researchers avoid using the term "causality" in their studies are varied. First, there is a deeply ingrained caution around causal language, so much so that some scientific journals explicitly prohibit causal terminology in observational studies, reserving it exclusively for RCTs [16,17]. However, this stance can be counterproductive, as observational designs, backed by strong arguments and methodologies, can serve as the first step in causal exploration, which may be unfeasible in experimental studies for economic or ethical reasons. Thus, limiting causality claims to RCTs would confine causal understanding to studies using this design alone, potentially excluding critical findings [8,9].

Another reason this conceptual confusion manifests in practice is the selection of study design when determining relationships between variables. Researchers frequently resort to cross-sectional designs, which, while useful, have inherent limitations for causal inference, including the phenomenon known as "reverse causation" [10]. Faced with these methodological limitations, authors often take refuge in conservative

terms like "association" or "risk," even when their true objective is to explore causal relationships (especially since, as noted initially, they use tools typically associated with an explanatory scientific task).

For this reason, while many researchers may want to avoid pursuing causality, two methodologically coherent paths can be considered. The first is to remain within the descriptive realm, refraining from causal language and the tools specific to causal inference. Alternatively, a second option is to openly acknowledge the study's causal intentions, employing appropriate methodological tools while explicitly discussing the design's limitations for causal inference. Though more challenging, this second scientific tasks allows for a more transparent and critical evaluation of the evidence generated [2,8].

Classification of the Explanatory Scientific Task

The explanatory approach can be approached from two complementary perspectives, exploratory and confirmatory, which reflect different stages in building causal knowledge [42].

The exploratory scientific tasks represent an initial investigation into potential causal relationships. Researchers examine multiple possible relationships without a strongly preconceived hypothesis in this context, searching for patterns that suggest plausible causal mechanisms. For example, studies that initially identified the association between oral contraceptive use and venous thromboembolism began as exploratory investigations [43]. Although these studies have inherent limitations due to their exploratory nature, they are essential for developing new causal hypotheses that warrant further research.

The exploratory scientific tasks typically rely more on observational data, such as cohort or case-control studies, where researchers can simultaneously examine multiple potential causal relationships. In contrast, the confirmatory approach often employs more structured designs like randomized controlled trials or carefully controlled observational studies with pre-specified hypotheses and rigorous control of potential confounders. This progression from observational exploration to controlled confirmation reflects the natural maturation of causal evidence. For example, the investigation of smoking's health effects began with observational studies exploring multiple potential outcomes before moving to more focused controlled studies examining specific causal mechanisms.

On the other hand, the confirmatory approach represents a more mature stage in causal research, where the goal is to verify specific, previously formulated causal hypotheses. These studies are characterized by pre-specified hypotheses, clearly defined assessment criteria, and methods to control known confounding factors. A classic example would be a study specifically designed to confirm the causal role of HPV in cervical cancer, with precise measurements of exposure, control of known confounders, and an appropriate timeframe [44].

However, as Greenland warns [45], the distinction between these approaches is not always clear in practice. Many studies combine exploratory and confirmatory elements, and the current trend is to acknowledge this duality rather than enforce a dichotomous classification. Thus, it is essential to highlight that the path from exploration to causal confirmation typically follows natural progression in biomedical research. In the initial stages, the approach tends to be exploratory when using classical epidemiological designs such as cross-sectional or case-control studies (as discussed below) and recognizing their inherent limitations. Although these studies cannot establish definitive causality, they are fundamental in generating hypotheses and indicating promising directions for research. On the other hand, when a solid theory is supported by preliminary evidence and more robust designs such as RCTs or observational studies with advanced causal inference methods (such as instrumental variables or sensitivity analyses) are employed, we can speak of a confirmatory approach. This natural progression from exploration to confirmation reflects the maturation of scientific knowledge and the gradual accumulation of causal evidence.

Research Designs with a Explanatory Scientific Task

Causal-focused studies can be implemented through various research designs, both observational and experimental. Among observational designs, cross-sectional studies, although widely used for their efficiency and feasibility, have significant limitations for causal inference due to the inability to establish temporality between exposure and outcome. Therefore, they can be useful in initiating the search for potential etiological factors, albeit with noted limitations. Case-control studies can go a bit further and are especially useful for studying rare events and more resource-efficient [10].

Cohort studies, on the other hand, represent one of the most robust observational designs for causal inference, as they establish temporality and allow for the evaluation of multiple outcomes. A paradigmatic example is the study of the causal relationship between smoking and lung cancer, where observational evidence was so convincing that it established causality without the need for RCTs. This case illustrates how well-designed observational studies can provide solid causal evidence, especially when findings are consistent across multiple studies and populations [46].

RCTs are considered the gold standard for causal inference due to their ability to control known and unknown confounding factors through randomization. However, they also present limitations: they are costly, may face external validity issues, and are often unethical or infeasible. Additionally, randomization alone does not guarantee the absence of other biases, such as loss to follow-up or non-compliance with the assigned treatment [8].

It is important to recognize that each design has strengths and weaknesses for causal inference. The design should be based on methodological considerations and practical, ethical, and study-specific aspects. The strongest evidence often emerges from triangulating results from different study designs [47].

Classical Statistical Analyses Used in Explanatory Scientific Tasks

Choosing effect measures is crucial for correctly interpreting results in causally focused studies. Relative Risk (RR) is one of the most intuitive and directly interpretable measures, representing how many times more likely the event is to occur in the exposed group compared to the non-exposed group. Hazard Ratios (HR) are especially useful in longitudinal studies where time to event is important, allowing for the incorporation of censored data and variable follow-up times [10].

The Odds Ratio (OR) is often used, particularly in case-control studies, although its interpretation requires caution. While OR approximates RR when the event is rare (less than 10%), it may overestimate the association when the event is common. In a causal context, it is essential to remember that these association measures can only be interpreted as causal effects when the fundamental assumptions of consistency, exchangeability, and positivity are met [10].

Absolute effect measures, such as Risk Difference and Number Needed to Treat (NNT), are particularly valuable for public health decision-making and clinical practice. NNT, which indicates how many individuals need to be treated to prevent an additional event, provides a more tangible measure of an intervention's impact. However, it is important to consider that NNT can vary significantly depending on the baseline risk of the studied population [10].

While these traditional effect measures have proven useful in clinical fields, advances in epidemiological methodology have opened new horizons in explanatory approaches. Advances in causal inference theory have provided more sophisticated tools for rigorously analyzing cause-effect relationships, representing technical advancement and a fundamental shift in our understanding of causality in epidemiology, as discussed below.

Modern Assumptions in Explanatory Scientific Tasks

In recent decades, traditional epidemiological research approaches have been complemented by more rigorous causal frameworks to strengthen the validity of causal inferences. This evolution has led to the development of fundamental assumptions that must be met for valid causal effect estimation, providing a more solid conceptual framework to evaluate our conclusions [8].

In modern causal inference, three assumptions go beyond the classical Bradford Hill criteria: consistency, exchangeability, and positivity. Consistency establishes that the observed outcome under a specific treatment should correspond to the potential outcome under that same treatment, a crucial concept for linking observed outcomes with counterfactuals. Exchangeability, modernizing the traditional idea of confounding, implies that comparison groups are similar in all relevant aspects except for the exposure of interest. Positivity requires that each individual has a non-zero probability of receiving any treatment level, a rarely considered assumption in traditional approaches [48].

In addition to these fundamental assumptions, the modern approach incorporates new methodological tools such as advanced statistical methods like the g-formula and inverse probability weighting. These methods enable the estimation of different types of causal effects, both at the individual and population levels, including the average treatment effect (ATE), effect among treated individuals (ATT), and controlled and natural causal effects in mediation analyses [47]. Additionally, in this context, these traditional estimators

are complemented by more sophisticated methods such as standardization, inverse probability weighting (IPTW), and G-methods, which allow for better control of confounding and more precise causal effect estimation in the presence of time-dependent confounders or mediation [47].

The validity of these assumptions must be carefully evaluated in each research context. While some assumptions are met by design in RCTs, observational studies require more detailed consideration and possibly additional methodological adjustments. However, RCTs also face important challenges in meeting causal assumptions. Treatment non-compliance can violate consistency assumptions, differential loss to follow-up may compromise exchangeability, and strict inclusion/exclusion criteria can affect positivity. Additionally, protocol deviations, crossover effects, and missing data can threaten the validity of causal estimates even in randomized designs. Understanding these limitations is crucial for proper causal inference in experimental studies [49].

An Additional Challenge in Explanatory Scientific Tasks: Manipulable vs. Non-manipulable Variables

The distinction between manipulable and non-manipulable variables is fundamental in causal inference, as it directly impacts how we study and understand causal relationships. Manipulable variables can be directly controlled or modified in a study (e.g., administering a medication). In contrast, non-manipulable variables are characteristics or states that cannot be randomly assigned (e.g., obesity, gender, or age) [8].

Non-manipulable variables challenge causal inference because they cannot be directly controlled or assigned in a study. Obesity perfectly illustrates this complexity: while we can observe different Body Mass Index (BMI) levels in a population, we cannot "assign" people to be obese or non-obese as we would with a traditional medical intervention [8].

To understand this better, consider the following example: imagine conducting three hypothetical RCTs

Table 2: Biomedical Research Scientific Tasks: Comparison of Descriptive, Predictive, and Explanatory

Characteristic	Descriptive	Predictive	Explanatory
Main objective	Characterize patterns, distributions, and trends	Anticipate future outcomes or identify current conditions	Establish causal relationships
Key questions	What is happening? How is it distributed? Where and when does it occur?	What is the probability of occurrence? Who is at risk?	Why does it occur? What causes it?
Examples	- Diabetes prevalence by region - COVID-19 temporal trends - Characterization of health inequities - Molecular distribution of breast cancer by subtypes	- 10-year cardiovascular risk - Sepsis prediction - Venous thrombosis diagnosis - Predictors of colorectal cancer recurrence	- Effect of statins on mortality - Causality between smoking and cancer - Impact of health interventions
Typical methods	- Frequency measures - Trend analysis - Data visualization - Descriptive statistics	- Predictive regression models - Machine learning - Validation and calibration - Predictive performance metrics	- Causal analysis - Directed Acyclic Graphs (DAGs) - Confounder adjustment - Mediation analysis
Common designs	- Cross-sectional studies - Case series - Ecological studies	- Cohort studies - Analytical cross-sectional studies - Diagnostic studies	- Clinical trials - Cohort studies - Case-control studies
Special considerations	- Avoid unnecessary adjustments - Highlight disparities - Do not infer causation	- External validation - Calibration - Periodic updates	- Confounding control - Temporality - Causal assumptions
Primary application	- Health planning - Trend monitoring - Problem identification	- Clinical decision-making - Risk stratification - Early diagnosis	- Therapeutic interventions - Health policy - Prevention
Limitations	- Does not establish causation - Does not predict future outcomes - Possible ecological fallacy	- Requires ongoing validation - May lack generalizability - Data quality dependent	- Causation difficult to establish - Residual confounding - Selection bias

to reduce BMI—one based on intensive exercise, another on a restrictive dietary intervention, and a third combining moderate exercise and diet. Although all three studies achieved the same reduction in BMI, each showed different effects on mortality. This occurs because each weight reduction method can have direct health effects beyond those mediated solely through BMI change [50]. This leads us to an important conclusion: when studying the "effect of obesity," we investigate the effects of different mechanisms and interventions that lead to a particular BMI. Therefore, it is not obesity that we can manipulate but interventions that lead to body weight changes. This has crucial implications for research and public health policy, suggesting that we should focus on studying the effects of specific, well-defined interventions rather than trying to estimate the general effect of obesity [50].

Conclusions, Challenges, and Perspectives in Scientific Task-Oriented Approaches

Classifying biomedical research into descriptive, predictive, and explanatory scientific tasks represents a significant advancement over the traditional descriptive-analytical dichotomy. However, each approach faces challenges that merit attention. In research with descriptive scientific tasks, the primary challenge lies in avoiding unnecessary adjustments that could obscure important disparities, especially when the goal is to highlight health inequities. In research with predictive scientific tasks, the challenge lies in continuously validating and updating models across different populations and effectively integrating these tools into clinical practice. In research with explanatory scientific tasks, the tension persists between the need to establish causality, the inherent limitations of observational designs, and the historical reluctance to declare causal objectives explicitly.

All three approaches share the common challenge of methodological transparency. Researchers must be clear about their objectives and methods, recognizing that each scientific task requires different analytical techniques. For example, a cohort study can be used for descriptive, predictive, or explanatory purposes, but the analysis methods and interpretation of results will differ substantially depending on the primary objective.

In the current context of evidence-based health application, it is crucial to recognize that these approaches are not mutually exclusive but complementary. Descriptive scientific tasks can generate hypotheses later evaluated in explanatory scientific tasks, while findings from predictive studies can inform the description of phenomena and the investigation of causal mechanisms. This interrelationship underscores the importance of maintaining methodological rigor specific to each approach.

Looking ahead, advances in statistical methodology and the availability of large databases are likely to continue expanding the possibilities of each approach. However, the key to success will remain the proper alignment between research objectives and the methods employed. Researchers should resist the temptation to make inferences beyond the scope of their study design, whether attempting to establish causality from purely descriptive scientific tasks or extrapolating predictions beyond the populations in which the models were developed.

The scientific community should work towards developing specific methodological guidelines for each scientific task, recognizing their unique requirements and particularities. This will facilitate the planning and evaluation of biomedical research, contributing to more transparent and reproducible science. The goal is to generate evidence that is methodologically sound and useful for decision-making in public health and clinical practice.

ACKNOWLEDGEMENTS

A special thanks to the members of Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas (UNTRM), Amazonas, Peru for their support and contributions throughout the completion of this research.

FINANCIAL DISCLOSURE

This study was financed by Vicerectorado de Investigación de la Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

INFORMED CONSENT

This study is a systematic review therefore informed consent is not required.

DATA AVAILABILITY

Data are available upon request to the corresponding author.

AUTHORS' CONTRIBUTION

Víctor Juan Vera-Ponce: Conceptualization, Investigation, Methodology, Resources, Writing - Original Draft, Writing - Review & Editing.

Fiorella E. Zuzunaga-Montoya: Methodology, Software, Data Curation, Formal analysis, Writing - Review & Editing.

Luisa Erika Milagros Vásquez-Romero: Validation, Visualization, Investigation, Project administration, Writing - Original Draft, Writing - Review & Editing.

Joan A. Loayza-Castro: Investigation, Project administration, Writing - Original Draft, Writing - Review & Editing.

Nataly Mayely Sanchez-Tamay: Investigation, Methodology, Writing - Original Draft, Writing - Review & Editing.

Carmen Inés Gutierrez De Carrillo: Methodology, Resources, Supervision, Funding acquisition, Writing - Review & Editing.

REFERENCES

[1] Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. Lancet 2002; 359(9301): 145-9. https://doi.org/10.1016/S0140-6736(02)07373-7

[2] Hernán MA, Hsu J, Healy B. A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. CHANCE 2019; 32(1): 42-9. https://doi.org/10.1080/09332480.2019.1579578

Modern Epidemiology [Internet]. [citado el 21 de octubre de 2024]. Disponible en: https://www.wolterskluwer.com/en/solutions/ovid/modern-epidemiology-4634

[3] Shmueli G. To Explain or to Predict? Statist Sci [Internet] 2010 [citado el 26 de octubre de 2024]; 25(3). https://doi.org/10.1214/10-STS330

[4] Breiman L. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). Statist Sci [Internet] 2001 [citado el 26 de octubre de 2024]; 16(3). https://doi.org/10.1214/ss/1009213726

[5] Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, *et al*. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ 2020; 369: m1328. https://doi.org/10.1136/bmj.m1328

[6] Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, *et al*. Calibration: the Achilles heel of predictive analytics. BMC Medicine 2019; 17(1): 230. https://doi.org/10.1186/s12916-019-1466-7

[7] Hernan MA, Robins JM. Causal Inference: What If. CRC Press; 2024. 312 p.

[8] Huitfeldt A. Is caviar a risk factor for being a millionaire? BMJ 2016; 355: i6536. https://doi.org/10.1136/bmj.i6536

[9] Rothman KJ. Modern Epidemiology. 4th ed. Wolters Kluwer, 2021.

[10] Leenen RCA, Venderbos LDF, Helleman J, Gómez Rivas J, Vynckier P, Annemans L, *et al*. Prostate Cancer Early Detection in the European Union and UK. European Urology [Internet] 2024 [citado el 9 de noviembre de 2024]; https://doi.org/10.1016/j.eururo.2024.07.019

[11] Suckow BD, Newhall KA, Bekelis K, Faerber AE, Gottlieb DJ, Skinner JS, *et al*. Hemoglobin A1c Testing and Amputation Rates in Black, Hispanic, and White Medicare Patients. Ann Vasc Surg 2016; 36: 208-17. https://doi.org/10.1016/j.avsg.2016.03.035

[12] Long GV, Grob J-J, Nathan P, Ribas A, Robert C, Schadendorf D, *et al*. Factors predictive of response, disease progression, and overall survival after dabrafenib and trametinib combination treatment: a pooled analysis of individual patient data from randomised trials. Lancet Oncol 2016; 17(12): 1743-54. https://doi.org/10.1016/S1470-2045(16)30578-2

[13] Leonetti A, Sharma S, Minari R, Perego P, Giovannetti E, Tiseo M. Resistance mechanisms to osimertinib in EGFR-mutated non-small cell lung cancer. Br J Cancer 2019; 121(9): 725-37. https://doi.org/10.1038/s41416-019-0573-8

[14] Singh GK, Jemal A. Socioeconomic and Racial/Ethnic Disparities in Cancer Mortality, Incidence, and Survival in the United States, 1950-2014: Over Six Decades of Changing Patterns and Widening Inequalities. J Environ Public Health 2017; 2017: 2819372. https://doi.org/10.1155/2017/2819372

[15] Haber NA, Wieten SE, Rohrer JM, Arah OA, Tennant PWG, Stuart EA, *et al*. Causal and Associational Language in Observational Health Research: A Systematic Evaluation. Am J Epidemiol 2022; 191(12): 2084-97. https://doi.org/10.1093/aje/kwac137

[16] Hernán MA. The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. American Journal of Public Health 2018; 108(5): 616. https://doi.org/10.2105/AJPH.2018.304337

[17] Vickers AJ, Assel M, Dunn RL, Zabor EC, Kattan MW, van Smeden M, *et al*. Guidelines for Reporting Observational Research in Urology: The Importance of Clear Reference to Causality. Urology 2023; 177: 1-5. https://doi.org/10.1016/j.urology.2023.04.007

[18] Ellenberg JH. Selection bias in observational and experimental studies. Stat Med 1994; 13(5-7): 557-67. https://doi.org/10.1002/sim.4780130518

[19] Fox MP, Murray EJ, Lesko CR, Sealy-Jefferson S. On the Need to Revitalize Descriptive Epidemiology. Am J Epidemiol 2022; 191(7): 1174-9. https://doi.org/10.1093/aje/kwac056

[20] Bhopal RS. Concepts of Epidemiology: Integrating the ideas, theories, principles, and methods of epidemiology. Third Edition, Third Edition. Oxford, New York: Oxford University Press; 2016. 480 p.

[21] Conroy S, Murray EJ. Let the question determine the methods: descriptive epidemiology done right. Br J Cancer 2020; 123(9): 1351-2. https://doi.org/10.1038/s41416-020-1019-z

[22] Lesko CR, Zalla LC. Rigorous Descriptive Epidemiology for Health Justice. Epidemiology 2023; 34(6): 838-40. https://doi.org/10.1097/EDE.0000000000001658

[23] Tordoff DM, Fernandez A, Perry NL, Heberling WB, Minalga B, Khosropour CM, *et al*. A Quantitative Intersectionality Analysis of HIV/STI Prevention and Healthcare Access Among Transgender and Nonbinary People. Epidemiology 2023; 34(6): 827-37. https://doi.org/10.1097/EDE.0000000000001669

Global Burden of Disease (GBD) [Internet]. [citado el 26 de octubre de 2024]. Disponible en: https://www.healthdata.org/research-analysis/gbd

[24] Braveman PA, Kumanyika S, Fielding J, LaVeist T, Borrell LN, Manderscheid R, *et al*. Health Disparities and Health

Equity: The Issue Is Justice. American Journal of Public Health 2011; 101(Suppl 1): S149. https://doi.org/10.2105/AJPH.2010.300062

[25] VanderWeele TJ, Robinson WR. On causal interpretation of race in regressions adjusting for confounding and mediating variables. Epidemiology (Cambridge, Mass) 2014; 25(4): 473. https://doi.org/10.1097/EDE.0000000000000105

[26] Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol 2002; 155(2): 176-84. https://doi.org/10.1093/aje/155.2.176

[27] Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating [Internet]. Cham: Springer International Publishing; 2019 [citado el 26 de octubre de 2024]. (Statistics for Biology and Health). https://doi.org/10.1007/978-3-030-16399-0

[28] D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, *et al*. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008; 117(6): 743-53. https://doi.org/10.1161/CIRCULATIONAHA.107.699579

[29] The PROGRESS framework [Internet]. Prognosis Research. [citado el 26 de octubre de 2024]. Disponible en: https://www.prognosisresearch.com/progress-framework

[30] Knottnerus JA, Muris JW. Assessment of the accuracy of diagnostic tests: the cross-sectional study. J Clin Epidemiol 2003; 56(11): 1118-28. https://doi.org/10.1016/s0895-4356(03)00206-3

[31] Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. BMC Medicine 2015; 13(1): 1. https://doi.org/10.1186/s12916-014-0241-z

[32] Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, *et al*. Calculating the sample size required for developing a clinical prediction model. BMJ 2020; 368: m441. https://doi.org/10.1136/bmj.m441

[33] Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal Evaluation of the Accuracy of a Biomarker Used for Classification or Prediction: Standards for Study Design. JNCI Journal of the National Cancer Institute 2008; 100(20): 1432. https://doi.org/10.1093/jnci/djn326

[34] Harrell , FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis [Internet]. Cham: Springer International Publishing; 2015 [citado el 26 de octubre de 2024]. (Springer Series in Statistics). https://doi.org/10.1007/978-3-319-19425-7

[35] van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. BMC Medical Research Methodology 2014; 14(1): 137. https://doi.org/10.1186/1471-2288-14-137

[36] Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. J Clin Epidemiol 2016; 69: 245-7. https://doi.org/10.1016/j.jclinepi.2015.04.005

[37] Than M, Herbert M, Flaws D, Cullen L, Hess E, Hollander JE, *et al*. What is an acceptable risk of major adverse cardiac event in chest pain patients soon after discharge from the Emergency Department?: a clinical survey. Int J Cardiol 2013; 166(3): 752-4. https://doi.org/10.1016/j.ijcard.2012.09.171

[38] Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, *et al*. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. JAMA 2017; 318(14): 1377-84. https://doi.org/10.1001/jama.2017.12126

[39] Vickers AJ, Calster BV, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ 2016; 352: i6. https://doi.org/10.1136/bmj.i6

[40] Vandenbroucke JP. When are observational studies as credible as randomised trials? Lancet 2004; 363(9422): 1728-31. https://doi.org/10.1016/S0140-6736(04)16261-2

[41] Thorogood M, Vessey MP. An epidemiologic survey of cardiovascular disease in women taking oral contraceptives. Am J Obstet Gynecol 1990; 163(1 Pt 2): 274-81. https://doi.org/10.1016/0002-9378(90)90567-q

[42] Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, *et al*. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999; 189(1): 12-9. https://doi.org/10.1002/(SICI)1096-9896(199909)189: 1<12: : AID-PATH431>3.0.CO; 2-F

[43] Greenland S. For and Against Methodologies: Some Perspectives on Recent Causal and Statistical Inference Debates. Eur J Epidemiol 2017; 32(1): 3-20. https://doi.org/10.1007/s10654-017-0230-6

[44] Doll R, Hill AB. The Mortality of Doctors in Relation to Their Smoking Habits. British Medical Journal 1954; 1(4877): 1451. https://doi.org/10.1136/bmj.1.4877.1451

[45] VanderWeele TJ. Explanation in causal inference: developments in mediation and interaction. International Journal of Epidemiology 2016; 45(6): 1904. https://doi.org/10.1093/ije/dyw277

[46] Pearl J. Causality: Models, Reasoning, and Inference. Cambridge University Press; 2000. 412 p.

[47] VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med 2017; 167(4): 268-74. https://doi.org/10.7326/M16-2607

[48] Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. Int J Obes (Lond) 2008; 32 Suppl 3: S8-14. https://doi.org/10.1038/ijo.2008.82

Received on 22-09-2024 Accepted on 18-10-2024 Published on 20-11-2024

https://doi.org/10.30683/1929-2279.2024.13.08