

BRIEF REPORT

The importance of data: Monitoring variables in causal inference with medical imaging



C.M. Baracaldo Silva^{a,*}, P.M. Martínez Gironés^b, A. Galiana Bordera^b,
L. Cerdá-Alberich^b, D. Martí-Aguado^c, C. García Villar^d, L. Martí-Bonmatí^{a,b}

^a Servicio de Radiología, Área Clínica de Imagen Médica, Hospital Universitario y Politécnico La Fe, Valencia, Spain

^b Grupo de Investigación Biomédica en Imagen GIBI2³⁰, Instituto de Investigación Sanitaria La Fe, Valencia, Spain

^c Servicio de Medicina Digestiva, Hospital Clínico Universitario, Instituto de Investigación Sanitaria INCLIVA, Valencia, Spain

^d Servicio de Radiología, Hospital Puerta del Mar, Cádiz, Spain

Received 23 July 2024; accepted 23 August 2024

Available online 17 March 2025

KEYWORDS

Explanatory variables;
Causation;
Confounding bias;
Observational study;
Bradford Hill criteria

Abstract

Background and objectives: In scientific studies with medical imaging, it is important that the process involved in selecting the explanatory variables is closely monitored, as this can lead to significant confounding biases in its methodology. The aim of this paper is to explore whether it is feasible to use criteria for causation (such as those outlined by Bradford Hill) when selecting variables in articles published in the journal *RADIOLOGÍA*.

Material and methods: We selected the July-August issue of *Radiología*, issue 4 of 2023. Four original articles with sample sizes of at least 100 patients were chosen and their follow-up of causation was analysed.

Results: The criteria of temporality, consistency, coherence, plausibility and analogy were observed. The criteria of specificity, experiment, strength and biological gradient were not consistently applied. The methodological robustness and quality of these studies was also noted.

Conclusion: Bradford Hill's criteria for causation are useful for the selection of variables in medical imaging studies. The most relevant criteria in medical imaging are temporality, consistency, coherence, plausibility and analogy.

© 2024 SERAM. Published by Elsevier España, S.L.U. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

* Corresponding author.

E-mail addresses: cm.baracaldo2551@gmail.com, cm.baracaldo2551@uniandss.edu.co, baracaldo_car@gva.es (C.M. Baracaldo Silva).

PALABRAS CLAVE

Variables explicativas;
Causalidad;
Sesgo de confusión;
Estudio observacional;
Criterios de Bradford Hill

La importancia de los datos: control de las variables en la inferencia de causalidad con imagen médica**Resumen**

Antecedentes y objetivo: En los estudios científicos con imagen médica es crucial el control del proceso de selección de variables explicativas, ya que puede ser una fuente importante de sesgos metodológicos de confusión. El objetivo de este trabajo es analizar la viabilidad de emplear los criterios de causalidad (como los enunciados por Bradford Hill) en la selección de variables en artículos publicados en la revista *RADIOLOGÍA*.

Material y método: Se seleccionó el número cuatro de 2023 de la revista Radiología, que corresponde a Julio-Agosto. Se eligieron cuatro publicaciones originales observacionales con un tamaño muestral de al menos 100 pacientes para su análisis de control de causalidad.

Resultados: Se observó cumplimiento de los criterios de temporalidad, consistencia, coherencia, plausibilidad y analogía. Los criterios de especificidad, experimento, fuerza y dosis/respuesta no se aplicaron de forma consistente. Se constató la robustez y calidad metodológica de estos estudios.

Conclusión: Los criterios de causalidad de Bradford Hill resultan útiles para la selección de variables en estudios científicos de imagen médica. Los criterios más relevantes en imagen médica son los de temporalidad, consistencia, coherencia, plausibilidad y analogía.

© 2024 SERAM. Publicado por Elsevier España, S.L.U. Se reservan todos los derechos, incluidos los de minería de texto y datos, entrenamiento de IA y tecnologías similares.

Introduction

One of the main challenges for radiologists when designing studies that correlates image findings with specific medical events are the biases or systematic errors that can arise from the failure to effectively control the selection of variables and their potential confounding factors.

To avoid inconsistencies and non-reproducibility, the methodology design should control for possible biases that generate uncertainty.¹ One of the most critical factors in this process is the selection of the explanatory variables to be collated and analysed. The selection of variables is particularly important in the retrospective and prospective observational studies used in medical imaging research.² Although retrospective studies are often associated with a greater methodological bias, this is not necessarily the case. The quality of the data is directly impacted by how well the variables are defined and the level of standardisation employed in the data collection process.

Confounding bias is a systematic error associated with the presence of confounding variables that can distort the conclusions drawn in terms of causality or cause errors in the calculation of associations between different selected variables used to make a causal inference, the process by which cause-effect relations are identified.^{2–6} To avoid these biases, we must therefore analyse what data should be collected at the beginning of any study.

The literature includes various checklists that help us design a good quality scientific investigation. In the case of medical imaging research, checklists like STROBE (STrengthening the Reporting of OBservational studies in Epidemiology) are particularly useful for observational studies,⁷ and STARD (STAndards for Reporting Diagnostic

accuracy studies) for calculating the diagnostic accuracy of a test.⁸ STROBE mentions the variables (section 7) and descriptive data (section 14) relating to the study population, including confounding factors, but does not define the criteria for the selection of explanatory variables. STARD discusses how to describe the subjects (sections 19–22) and recommends mentioning the possible biases and limitations of the study (section 26), although it does not define how to select and evaluate these variables.

Our objective is to use the Bradford Hill criteria, otherwise known as Hill's criteria for causation,⁹ to evaluate the methodological quality of studies published in the *RADIOLOGÍA* journal to verify the applicability of the criteria and demonstrate how they can be used in the selection of explanatory variables when designing a medical imaging study.

Material and methods

We have evaluated publications from issue four (July/August) 2023 of the *RADIOLOGÍA* journal, which was the most recent issue available at the time this study was performed. In selecting the studies, we applied the following inclusion criteria: original studies, observational studies (prospective and retrospective) and sample size of at least 100 patients.

The Bradford Hill criteria⁹ were adjusted as a guide for establishing causal relations (Annex 1 of the supplementary material). One of the authors (CMBS) undertook the critical reading, applying these criteria to check where they had been used and followed. Where there were doubts as to

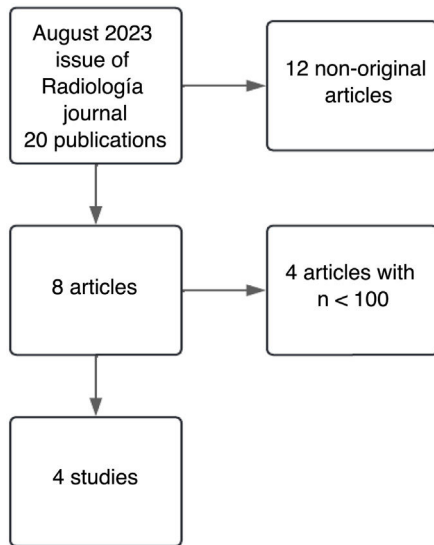


Figure 1 Study selection process flow diagram.

compliance, all authors discussed the causal inference for each study.

Results

Four articles complied with all the inclusion criteria^{10–13} (Fig. 1). Table 1 summarises the analysis in the form of a causality compliance matrix. The details of this analysis are provided below for each article.

Study 1 (Serrano et al.)¹⁰ is a retrospective observational study. It aims to analyse the efficacy of inferior vena cava filter removal and the clinical and radiological factors associated with a difficult removal. It uses multiple variables to define a difficult removal and correlates them with the different patient characteristics. They explain in detail the selection of variables, based principally on temporality, coherence, plausibility and analogy. The biological gradient criterion is applied to correlate the complexity of the procedure with the severity of patient-related factors and the time elapsed. Experiment and specificity criteria are not applied because the sample was not randomised and they find more than one factor that influences the medical event studied.

Study 2 (Villanueva et al.)¹¹ is a retrospective observational study, which aims to describe the appearance of pleural appendages (PA) in computed tomography (CT) and their relationship with body mass index (BMI). They analyse variables such as presence, size and localisation of PA. They mainly employ the criteria of coherence, plausibility and analogy to infer the hypothesis of the study. They use the biological gradient criterion to infer a greater probability of presenting with PA (outcome) as BMI increases (exposure). In contrast with Study 1, they apply the criterion of specificity to establish a causal relationship between BMI and the presence of PA. The experiment criterion is not used as they do not apply any form of randomisation.

Study 3 (Castro-García et al.)¹² is a retrospective diagnostic performance study. It analyses the diagnostic performance of CT pulmonary angiography by comparing different D-dimer cut-off values in the diagnosis of acute pulmonary embolism in patients with and without SARS-CoV-2 infection. They use two cohorts of patients, one during the pandemic (group A, subdivided into patients positive and negative for SARS-CoV-2) and one prior to the pandemic (group B). They analysed the sensitivity, specificity, ROC curves and cut-off points, correlating it with the CT pulmonary angiography results, the D-dimer values, the SARS-CoV-2 PCR results, the need for admission to intensive care unit (ICU) and the pulmonary embolism pattern. They employed temporality, coherence, analogy and plausibility criteria. The criteria of strength and biological gradient are not applicable as the study is designed to identify diagnostic performance with no analysis of exposure and outcome causality, while the criteria of specificity and experiment were not applied as the sample was not randomised, nor was a single cause-effect factor sought.

Study 4 (Hajiahmadi et al.)¹³ is a prospective observational study and determines the predictors of pulmonary hypertension and right heart dysfunction caused by pulmonary embolism. They use the pulmonary artery obstruction index calculated according to the Qanadli score, the dilatation of the right ventricle and diameter of main pulmonary artery during the acute phase of the illness using CT pulmonary angiography. They correlate this with signs of right ventricular dysfunction and chronic pulmonary hypertension in echocardiography six months after the acute

Table 1 Analysis results for causality criteria.

	Study 1 (Serrano et al.)	Study 2 (Villanueva et al.)	Study 3 (Castro-García et al.)	Study 4 (Hajiahmadi et al.)
Temporality	Yes	Yes	Yes	Yes
Strength	Yes	Yes	Not applicable	Yes
Biological gradient	Yes	Yes	Not applicable	Yes
Consistency	Yes	Yes	Yes	Yes
Coherence	Yes	Yes	Yes	Yes
Plausibility	Yes	Yes	Yes	Yes
Specificity	No	Yes	No	No
Experiment	No	No	No	No
Analogy	Yes	Yes	Yes	Yes

event. The principle variables are temporality (CT angiography during the acute phase and echocardiography six months after), strength, coherence and plausibility. The biological gradient criteria is employed when correlating the severity of the acute pulmonary embolism with the risk of suffering chronic pulmonary hypertension. The experiment and specificity criteria were not used as no randomisation was performed on the sample nor was single-factor causality sought.

Discussion

In medical imaging we make inferences in all our radiological reports. Many publications seek to use predictive models in which a combination of observations and data indicate a level or risk (such as the prediction of microvascular invasion in hepatocellular carcinoma). To trust the associations identified, we must establish a methodological design that controls for possible confounding factors that generate uncertainty in the results obtained.¹ The selection of variables is a critical factor in this methodological design. In our study, we have used the Bradford Hill criteria^{5,9} as a useful quality control tool in the design of observational studies, frequently used in medical imaging.

Our analysis of the four studies published in the *RADIOLOGÍA* journal reveals that these causality criteria can be very useful in defining the quality of a study, even if not all apply to every study. To be precise, the criteria of specificity, experiment, strength and biological gradient were not applied in every study assessed. This is due principally to the type of studies analysed, as the applicability of each criterion depends on the objectives and methodology used, so the relevant criteria should be defined according to the characteristics of each investigation. For example, none of the four studies involved randomisation of data or patients, as it was not relevant or necessary for the specific objectives of each study.

Not fulfilling all the criteria does not mean that the assessed investigations present methodological deficiencies, as the criteria should not be interpreted as a strict checklist.

In contrast, the criteria of temporality, consistency, coherence, plausibility and analogy were fundamental to the studies reviewed. The importance of these criteria lies in their close alignment with observational studies, used frequently in our field. It is possible that in other fields of medicine, where other types of studies prevail, the foundations of causality on which investigators base their work differ from our own. As a result, the Bradford Hill criteria can be applied accordingly, adapting and adjusting to each type of study.

The main limitation of our study is that only one issue of the journal was analysed, so it may not be representative. It will be interesting to see how the criteria for causality are applied in other contexts and as time goes on.

To conclude, the Bradford Hill causality criteria are relevant and useful in the methodological design of medical imaging studies. The criteria of temporality, consistency,

coherence, plausibility and analogy should be considered fundamental in the selection of explanatory variables for any study in this field.

Declaration of competing interest

Cristina García Vilar is the editor-in-chief of the *RADIOLOGÍA* journal. The authors declare that they have no conflicts of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.rx.2024.08.003>.

References

- Blackmore CC. The challenge of clinical radiology research. *AJR Am J Roentgenol*. 2001;176:327–31, <http://dx.doi.org/10.2214/ajr.176.2.1760327>.
- Martí-Bonmatí L. Estimaciones de causalidad con imagen médica en oncología. *ANALES RANM*. 2021;138:16–23, <http://dx.doi.org/10.32440/ar.2021.138.01.rev02>.
- Hernán MA, Robins JM. *Causal inference: what if*. Boca Raton: Chapman & Hall/CRC; 2020.
- Rothman KJ, Greenland S. Causation and causal inference in epidemiology. *Am J Public Health*. 2005;95 Suppl 1:S144–50, <http://dx.doi.org/10.2105/AJPH.2004.059204>.
- Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol*. 2015;12:14, <http://dx.doi.org/10.1186/s12982-015-0037-4>.
- Castro DC, Walker I, Glocker B. Causality matters in medical imaging. *Nat Commun*. 2020;11:3673, <http://dx.doi.org/10.1038/s41467-020-17478-w>.
- Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, et al., STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335:806–8, <http://dx.doi.org/10.1136/bmj.39335.541782.AD>.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015. An updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527, <http://dx.doi.org/10.1136/bmj.h5527>.
- Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58:295–300.
- Serrano E, Vila-Trias E, Zarco F, Zamora Martínez C, Moisés J, Gómez F, et al. Retirada difícil de filtro de vena cava inferior (FVCI). Consideraciones técnicas y variables asociadas. *Radiología*. 2023;65:230–8, <http://dx.doi.org/10.1016/j.rx.2022.07.006>.
- Villanueva Campos A, Etxano Cantera J, Patel N, Villanueva Marcos A. Apéndices pleurales: descripción, incidencia y relación con el índice de masa corporal. *Radiología*. 2023;65:106–11, <http://dx.doi.org/10.1016/j.rx.2021.03.008>.
- García JC, Arenas-Jiménez J, Ureña-Vacas A, García-Garrigós E, Sirera PS. Angio-TC pulmonar durante la pandemia de COVID-19: comparación en pacientes con y sin la infección

- y con una serie prepandémica. Radiología. 2023;65:222–9, <http://dx.doi.org/10.1016/j.rx.2022.07.005>.
13. Hajiahmadi S, Tabesh F, Shayganfar A, Shirani F, Ebrahimian S. Índice de obstrucción de la arteria pulmonar, diámetro de la arteria pulmonar y strain del ventrículo derecho como hallazgos pronósticos de la tomografía computarizada en el paciente con embolia pulmonar aguda. Radiología. 2023;65:200–12, <http://dx.doi.org/10.1016/j.rx.2021.02.008>.