OPEN

Do Neuroprognostic Studies Account for Self-Fulfilling Prophecy Bias in Their Methodology? The SPIN Protocol for a Systematic Review

Fernanda J. P. Teixeira, MD[1,2]
Bakhtawar Ahmad, MBBS[1]
Viktoriya Gibatova, MS[1]
Pouya A. Ameli, MD, MS[1,3]
Ivan da Silva, MD, PhD[1]
Thiago Carneiro, MD, MPH[1]
William Roth, MD[1,3]
Jenna L. Ford, MD[1,3]
Terry Kit Selfe, PhD[4]
David M. Greer, MD, MA[5]
Katharina M. Busl, MD, MS[1,3]
Carolina B. Maciel, MD, MSCR[1,3,6,7]

BACKGROUND: Self-fulfilling prophecy bias occurs when a perceived prognosis leads to treatment decisions that inherently modify outcomes of a patient, and thus, overinflate the prediction performance of prognostic methods. The goal of this series of systematic reviews is to characterize the extent to which neuroprognostic studies account for the potential impact of self-fulfilling prophecy bias in their methodology by assessing their adequacy of disclosing factors relevant to this bias.

METHODS: Studies evaluating the prediction performance of neuroprognostic tools in cardiac arrest, malignant ischemic stroke, traumatic brain injury, subarachnoid hemorrhage, and spontaneous intracerebral hemorrhage will be identified through PubMed, Cochrane, and Embase database searches. Two reviewers blinded to each other's assessment will perform screening and data extraction of included studies using Distiller SR and following Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. We will abstract data pertinent to the methodology of the studies relevant to self-fulfilling prophecy bias.

RESULTS: We will conduct a descriptive analysis of the data. We will summarize the reporting of mortality according to timing and mode of death, rates of exposure to withdrawal of life-sustaining therapy, reasoning behind limitations of supportive care, systematic use of standardized neuroprognostication algorithms and whether the tool being investigated is part of such assessments, and blinding of treatment team to results of neuroprognostic test being evaluated.

CONCLUSIONS: We will identify if neuroprognostic studies have been transparent in their methodology to factors that affect the self-fulfilling prophecy bias. Our results will serve as the foundation for standardization of neuroprognostic study methodologies by refining the quality of the data derived from such studies.

KEY WORDS: confirmation bias; neuroprognostication; prognosis; reference standards; reproducibility of results; self-fulfilling prophecy bias

The word "prognosis" is derived from the Greek roots of "*pro-*" meaning before, and "*-gnosis*" meaning knowledge. As such, "prognostication" in medicine indicates an effort to predict the course of a disease, or outcomes, without having full knowledge. In neurocritical care, prognostication is of cardinal importance, as many diseases may require a prolonged recovery process, lead to considerable disability, or even be fatal. Prognostication is aimed at facilitating decision-making, thereby guiding families, patients, and surrogates on what to expect moving forward in the journey following a severe neurologic injury. The ongoing process of predicting the course of recovery and ultimate outcome following severe neurologic injury is ubiquitous in clinical practice, and occurs with large gaps in knowledge such that prognostication may resemble an estimation rather than knowledge-based prediction (1). One such gap in knowledge—widely recognized as a major threat to

DOI: 10.1097/CCE.0000000000000943

KEY POINTS

Question:

Do studies evaluating the prediction performance of neuroprognostic tools in severe acute brain injury consistently report and account for factors associated with self-fulfilling prophecy bias?

Findings:

We will characterize the methodology of neuroprognostic studies. Specifically, we will describe reporting practices pertaining to mortality according to timing and mode of death, rates of exposure to withdrawal of life-sustaining therapy, reasoning behind limitations of supportive care, systematic use of standardized neuroprognostication algorithms and whether the tool being investigated is part of such assessments, and blinding of treatment team to results of neuroprognostic test being evaluated.

Meaning:

Our findings will serve as the foundation for standardizing the study methodology in neuroprognostication while furnishing data for the development of an assessment tool specific to self-fulfilling prophecy bias, and eventually resulting in improved rigor and quality of data derived from studies investigating neuroprognostic tools.

accurate prognostic impressions—is how to account for the impact of self-fulfilling prophecy bias when interpreting results on the prediction performance of neuroprognostic tools (2–6). Self-fulfilling prophecy bias is a type of confirmation bias that occurs when the results of outcome prediction methods under investigation for their prediction performance influence the outcomes in a cohort. The result is usually an overinflated prediction performance of poor outcomes, with falsely high specificity and positive prediction values. The self-fulfilling prophecy bias has long been recognized to influence prognostication and mortality in the neurocritically ill. Despite this recognition, standards are lacking for assessment and reporting of factors that help gauge the extent to which self-fulfilling prophecy bias has affected the results in neuroprognostication studies.

While standardization of prognostic studies improved following the publication of Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines in 2015 (7)—curated by the Enhancing the QUAlity and Transparency Of health Research network—these guidelines do not account for factors reflecting the impact imparted by the bias of self-fulfilling prophecy. For example, TRIPOD guidelines do not explicitly require reporting on the use of standardized neuroprognostication algorithms at study institutions, or if the treatment team was blinded to the results of the investigational neuroprognostic tool despite the possibility for significant skewing of prediction performance. Another important aspect in considering the impact of self-fulfilling prophecy bias is the reporting of mode of death in neuroprognostic cohorts: brain death, death due to refractory cardiac arrest despite resuscitation attempts and aggressive care, death due to withdrawal of life-sustaining therapies (WLSTs) due to perceived poor neurologic prognosis, or death due to WLST due to medical condition. Of these, self-fulfilling prophecy bias is presumably highest in cohorts with a high proportion of deaths related to care limitations, particularly if due to a perceived poor neurologic prognosis in neuroprognostic studies. In addition, the timing of final neuroprognostic assessments and WLST in relation to when the injury occurred is important. This is because the evolution of the patient's trajectory over time is a key aspect in neuroprognostication—for many reasons, from gauging cerebral resilience and capacity for recovery, accounting for the potential confounding effect of drugs and seizures, and to evaluate for the burden of secondary injury. The lack of uniformity in reporting factors most relevant to self-fulfilling prophecy bias severely limits the interpretation of data on neuroprognostic tools, and thus, hinders our ability to provide the most accurate prognostic impressions.

OBJECTIVES

To determine whether the methodology employed by neuroprognostication studies evaluated for and reported factors influencing the potential extent of self-fulfilling prophecy bias when evaluating the prediction performance of neuroprognostic tools in the setting of cardiac arrest, malignant ischemic stroke, traumatic brain injury, subarachnoid hemorrhage, and

spontaneous intracerebral hemorrhage. We aim to objectively and quantitatively demonstrate how much of the neuroprognostication literature has been impacted by self-fulfilling prophecy bias, and lay the foundation to quantify the degree to which results were impacted by such bias in each study.

METHODS AND ANALYSIS

Study Design

We will conduct a series of systematic reviews of published studies evaluating the prediction performance of neuroprognostic tools and transparency of reporting of Self-Fulfilling Prophecy in Neuroprognostication (SPIN) studies. The study is guided by the statement: a) population: cohort or systematic review studies of critically ill adult patients (i.e., ≥ 17 yr old) with one of the diagnoses of interest (i.e., cardiac arrest, malignant ischemic stroke, traumatic brain injury, subarachnoid hemorrhage, and spontaneous intracerebral hemorrhage); b) intervention: evaluation of the prediction performance of neuroprognostic tools; c) comparison: not applicable; d) outcome: reporting of factors associated with self-fulfilling prophecy bias and preliminary overall assessment of such risk for bias. We will use Distiller SR software (Ottawa, ON, Canada) to screen references and abstract data from relevant studies. The software handles dual reviewer screening, conflict resolution, capturing exclusion reasons, risk of bias assessments, duplicate detection, multiple database searches, and reporting templates such as Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). The results of this study will be reported according to the PRISMA Protocols guidelines (8) (**Supplemental Table 2**, http://links.lww.com/CCX/B219). Each individual systematic review was registered in the International Prospective Register of Systematic Reviews database: SPIN cardiac arrest CRD42021271923; SPIN intracerebral hemorrhage CRD42021276539; SPIN malignant ischemic stroke CRD42021276543; SPIN subarachnoid hemorrhage CRD42021276343; and SPIN traumatic brain injury CRD42022312805.

ELIGIBILITY CRITERIA

Neuroprognostication Studies

Peer-reviewed studies evaluating the prediction performance of any tool used in the prediction of mortality or neurologic outcome in patients with the diagnosis of interest will be included, without limitation in year of publication or original language in which the article was written. Study design will be limited to human studies that include at least 20 relevant subjects, constituting a small cohort. Abstracts without full text and non-English articles in a format that precludes digital translation to English will be excluded. All included articles were published and available via stated search criteria by the final search date of December 31, 2022.

Population

Neurocritically ill, adult patients (i.e., ≥ 17 yr old) with one of the diagnoses of interest (i.e., cardiac arrest, malignant ischemic stroke, traumatic brain injury, subarachnoid hemorrhage, and spontaneous intracerebral hemorrhage) in whom the neurologic outcome is in question. Studies that included patients less than 17 years old or defined their population as adult from age 16 years or older will be included, if a minimum of 20 subjects are included.

Outcome Measure

We have defined the primary outcome as the reporting of factors associated with self-fulfilling prophecy bias and preliminary overall assessment of such risk for bias. This preliminary assessment will be done using the questions presented in **Supplementary Table 1** (http://links.lww.com/CCX/B219), which were selected by the authors based on the perceived main contributors to self-fulfilling prophecy bias when evaluating the prediction performance of a neuroprognostic tool based on experience and available literature (9). The first question in the assessment table is the report of self-fulfilling prophecy bias by the original article. The results of this question will be reported separately. Each question will be answered as "Yes," "No," or "Uncertain." All questions are phrased so that "Yes" indicates reduced risk of self-fulfilling bias. Any question answered as "No" or "Uncertain" has a potential risk for self-fulfilling prophecy bias. Two senior, board-certified neurointensivists (K.M.B., C.B.M.) will independently review all questions for each study included in the systematic review. Any disagreements will be reconciled following a discussion between the two reviewers until consensus can be reached.

Data Sources and Search Strategy

Searches for published studies meeting the eligibility criteria was performed by a health science librarian (T.K.S.) in the following electronic databases: Embase, PubMed, and Cochrane. Automated search alerts were weekly following the initial upload, with records added through December 31, 2022. Initial deduplication was performed using EndNote software (Clarivate, London, United Kingdom).

STUDY SELECTION, DATA MANAGEMENT, AND DATA COLLECTION

Study Selection

Studies identified through the search strategy were uploaded to Distiller SR after removal of duplicates. During level 1 screening, two independent reviewers will review title and abstract of references and categorize as "include," "exclude," or "uncertain" in a blinded fashion. References not available in English will be screened by study member fluent in the original language in which the article was written or following the use of Google Translate (Mountain View, CA). All references marked "include" or "uncertain" will progress to level 2 screening after discrepancies are resolved by an independent third, senior reviewer (F.J.P.T., B.A., P.A.A., I.S., T.C., W.R., J.L.F., K.M.B., C.B.M.). During level 2 screening, full text of all articles will be screened by two independent reviewers in a blinded fashion, with discrepancies resolved by a third, independent, senior reviewer (F.J.P.T., B.A., P.A.A., I.S., T.C., W.R., J.L.F., K.M.B., C.B.M.). In case of unavailability of full-text of an article, the study team will request the article from the corresponding author. All references of included articles will be screened for inclusion of any additional, relevant studies.

Data Extraction

Outcomes of interest will be abstracted from the selected studies in duplicate by two independent reviewers through data extraction forms in a blinded fashion on Distiller. Any conflicting remarks regarding studies will be adjudicated by a third, senior reviewer (F.J.P.T., B.A., P.A.A., W.R., J.L.F., K.M.B., C.B.M.). Qualitative data will be extracted from studies meeting inclusion upon full-text review by two senior reviewers. Data extracted will be specifically those pertinent to the systematic review, including study characteristics such as year of publication, study design, sample size, population characteristics, type of neuroprognostic tool, and primary and secondary outcomes for each study. We will also abstract variables pertaining to factors related to self-fulfilling prophecy bias, such as acknowledgment of the effect of this bias by the authors in the article, use of and adherence to standardized neuroprognostication protocol, report of cultural tendency pertaining to WLST practices including institutional protocols for care limitations, blinding of the treatment team to the prognostic tool being investigated, mode and timing of death, including details on frequency, timing, and circumstances surrounding WLST. Data extraction will be independently cross-checked. Data produced from this systematic review will be uploaded in a data repository and made available to investigators upon reasonable request.

Data Analysis

Descriptive statistics of included articles on the frequency of reporting of each factor listed in Supplementary Table 1 (http://links.lww.com/CCX/B219) will be reported in aggregate. We will report each of the factors listed in Supplementary Table 1 (http://links.lww.com/CCX/B219), for each included article and include the breakdown in percentages where applicable. For example, the breakdown of death and rates of WLST (i.e., percentage or proportion of deaths attributable to limitations of care). No a priori subgroup analysis is planned.

Risk of Bias Assessment

The Risk of Self-fulfilling Prophecy Bias (ROSFPB) is the outcome of this systematic review and will be exposed according to the "Outcome Measure" section. Grading of the ROSFPB in low, medium, or high will be accomplished on a separate study after the outcome measures are weighted by an independent panel of experts using Delphi consensus methodology. The risk of bias will be assessed using the Prediction model Risk Of Bias ASsessment Tool (PROBAST) tool (10).

TRIPOD Guidelines

We will report adherence to TRIPOD guidelines as stated in the original articles. Since TRIPOD guidelines

were originally published in 2015, studies published prior to the year will not be penalized for lack of its utilization.

Ethics and Dissemination

This protocol does not require ethical approval as a systematic review of previously published literature. We plan to publish our findings in a peer-reviewed scientific journal and share the results with the lay press. The results of this systematic review may be presented at conferences and/or academic meetings prior to full publication.

CONCLUSIONS

Neuroprognostication—a key aspect of care in patients with severe neurologic injury—relies on many prognostic tools aiming at characterizing injury burden and potential for recovery. Accuracy in neuroprognostication hinges upon the reliability of the prediction performance of such tools, which is threatened by the pervasive self-fulfilling prophecy bias that tends to overinflate their performance. Thus far, despite wide recognition of the importance of the impact of self-fulfilling prophecy bias in neuroprognostication studies, there are no minimal standards to ensure adequate rigor and transparency in this line of research. SPIN is a series of systematic reviews focused "on the methodology" of outcome prediction studies with the goal of unveiling whether neuroprognostication studies reported factors influencing self-fulfilling prophecy bias when evaluating the prediction performance of neuroprognostic tools in the setting of hypoxic ischemic brain injury post-cardiac arrest, malignant ischemic stroke, traumatic brain injury, subarachnoid hemorrhage, and spontaneous intracerebral hemorrhage. Our findings will serve as the foundation for standardizing study methodology in neuroprognostication, eventually resulting in improved quality of data derived from studies investigating neuroprognostic tools.

ACKNOWLEDGMENTS

We would like to acknowledge the De Castro and Gorman families for their generous support of the scientific mission of the neurocritical care division in addressing this important knowledge gap in neuroprognostication.

1 *Department of Neurology, University of Florida, Gainesville, FL.*

2 *Department of Neurology, University of Miami, Miami, FL.*

3 *Department of Neurosurgery, University of Florida, Gainesville, FL.*

4 *Academic Research Consulting and Services, University of Florida, Gainesville, FL.*

5 *Department of Neurology, Boston University, Boston, MA.*

6 *Department of Neurology, University of Utah, Salt Lake City, UT.*

7 *Department of Neurology, Yale School of Medicine, New Haven, CT.*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccejournal).

Dr. Maciel conceptualized and devised the study, and provided critical revisions to the article. Dr. Selfe and Ms. Gibatova established the search terms for the individual disease categories, supervised by Drs. Maciel and Busl. Drs. Teixeira, Ameli, and Ahmad drafted the protocol and revised the article following critical input from Drs. Maciel and Busl. Drs. Carneiro, Silva, Greer, Ford, and Roth provided critical revisions to the article. All authors will participate in all reviews; however, primary responsibility for data collection and analysis for each diagnosis will be: Dr. Maciel for cardiac arrest, Dr. Ahmad for malignant ischemic stroke, Dr. Busl for subarachnoid hemorrhage, Dr. Maciel for traumatic brain injury, and Dr. Ameli for spontaneous intracerebral hemorrhage. Each author will draft and guarantee the final article for their assigned diagnosis, with all authors contributing to all initial articles resulting from this review.

The authors have disclosed that they do not have any potential conflicts of interest.

Drs. Teixeira and Ahmad share first authorship.

Drs. Busl and Maciel share senior authorship.

For information regarding this article, E-mail: carolina.maciel@neurology.ufl.edu

International Prospective Register of Systematic Reviews Registration Numbers: Self-Fulfilling Prophecy in Neuroprognostication (SPIN) cardiac arrest: CRD42021271923; SPIN intracerebral hemorrhage: CRD42021276539; SPIN ischemic stroke: CRD42021276543; SPIN subarachnoid hemorrhage: CRD42021276343; and SPIN traumatic brain injury: CRD42022312805.

REFERENCES

1. Wartenberg KE, Hwang DY, Haeusler KG, et al: Gap analysis regarding prognostication in neurocritical care: A joint statement from the German Neurocritical Care Society and the Neurocritical Care Society. *Neurocrit Care* 2019; 31:231–244
2. Caporro M, Rossetti AO, Seiler A, et al: Electromyographic reactivity measured with scalp-EEG contributes to prognostication after cardiac arrest. *Resuscitation* 2019; 138:146–152
3. Geocadin RG, Callaway CW, Fink EL, et al; American Heart Association Emergency Cardiovascular Care Committee: Standards for studies of neurological prognostication in

comatose survivors of cardiac arrest: A scientific statement from the American Heart Association. *Circulation* 2019; 140:e517–e542

4. Maciel CB: Neurologic outcome prediction in the intensive care unit. *Continuum (Minneap Minn)* 2021; 27:1405–1429
5. Souter MJ, Blissitt PA, Blosser S, et al: Recommendations for the critical care management of devastating brain injury: Prognostication, psychosocial, and ethical management: A position statement for healthcare professionals from the Neurocritical Care Society. *Neurocrit Care* 2015; 23:4–13
6. Zahuranec DB, Brown DL, Lisabeth LD, et al: Early care limitations independently predict mortality after intracerebral hemorrhage. *Neurology* 2007; 68:1651–1657
7. Collins GS, Reitsma JB, Altman DG, et al: Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *BMC Med* 2015; 13:1
8. Shamseer L, Moher D, Clarke M, et al: Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ* 2015; 349:g7647
9. Mayli M, Owen CK, Michel JAMP, et al: Can we learn from hidden mistakes? Self-fulfilling prophecy and responsible neuroprognostic innovation. *J Med Ethics* 2022; 48:922
10. Wolff RF, Moons KGM, Riley RD, et al; PROBAST Group†: PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019; 170:51–58