A Case For Clarity: Knowledge Engineering And Its Evolving Role In Fundamental And Applied Medical Sciences*

Henri Van Overmeire[1,2,*], Patrick Wouters[2]

[1]PointCaré, Research & Development Sint Pietersnieuwstraat 11, 9000, Ghent, Belgium
[2]Dept. of Basic and Applied Medical Sciences, Ghent University, Corneel Heymanslaan 10, 9000, Ghent, Belgium

Abstract

Biomedical science often accumulates data without theoretical convergence because key constructs remain semantically under-specified. Using the central venous pressure (CVP)–preload debate as a case study, this review shows how definitional drift creates "semantic pseudoreplication," where studies reuse the same terms while testing non-equivalent hypotheses, preventing decisive falsification. We argue that knowledge-engineering tools (ontologies, knowledge graphs, computational semantics) could provide machine-interpretable hypothesis definitions that stabilize meaning and improve epistemic efficiency. Finally, we propose an Incremental Epistemic Effectiveness Ratio (IEER) to quantify knowledge gained per research investment once hypotheses are formally specified.

Keywords

physiology, medicine, knowledge engineering, epistemology, central venous pressure, IEER

1. Introduction

Biomedical science is data-rich but conceptually fragile. The long-standing debate over whether central venous pressure is a clinically relevant predictor of cardiac preload and fluid-responsiveness exemplifies how decades of data can accumulate without theoretical convergence when key terms lack fixed meaning. We argue that problem lies not in measurement but in the absence of formal semantics at the level of hypothesis definition.

Unlike physics, which anchors concepts in mathematical formalism, medicine relies on natural language and shifting usage. Constructs such as "preload" or "volume status" blur theoretical and operational boundaries, producing apparent progress without conceptual closure.

Knowledge engineering offers a remedy. Through formal logic, ontologies and knowledge graphs, it enables machine-interpretable representations of biomedical entities and relations, though these tools rarely extend to hypothesis formulation. This review argues that semantic formalization can clarify biomedical reasoning itself. The CVP–preload case shows how definitional vagueness generates pseudo-cumulative research—and how, with semantic precision, an Incremental Epistemic Effectiveness Ratio (IEER) could quantify the knowledge gained per research investment.

2. Methodology

2.1. Objectives and Scope

This review investigates how semantic ambiguity in biomedical research constrains theoretical progress and how formal semantic frameworks could enhance conceptual precision, falsifiability, and epistemic

AI4SciSci'25: Workshop on the Artificial Intelligence and the Science of Science, December 15, 2025, DeKalb, Illinois, USA
*This short paper represents an initial exploration of the topic and is currently being expanded into a more comprehensive publication as research continues to develop.
*Corresponding author.
henri@pointcare.be (H. Van Overmeire); patrick.wouters@ugent.be (P. Wouters)
0000-0003-0840-925X (H. Van Overmeire); 0000-0002-3086-7137 (P. Wouters)

efficiency. Using the controversy over central venous pressure (CVP) and preload as a case study, it examines how definitional drift sustains pseudo-cumulative inquiry and how semantic formalization could realign biomedical science with the clarity characteristic of the physical sciences.

2.2. Research Questions

- How has semantic vagueness shaped the formulation, testing, and persistence of the CVP–preload hypothesis in biomedical research?
- What can the comparison with fields like physics reveal about the role of formal semantics in enabling decisive falsification and theoretical convergence?
- Which knowledge-engineering and formalization approaches (e.g., ontologies, computational semantics) could provide the infrastructure for precise, machine-interpretable biomedical hypotheses?
- What are the socio-economic costs associated with persistent semantic ambiguity in biomedical research, exemplified by the CVP–preload debate?

A targeted literature search examined how central venous pressure (CVP) and preload have been defined and investigated in biomedical research, focusing on conceptual evolution rather than exhaustive coverage. The recent narrative review by Lloyd-Donald et al. [1] served as a guide for identifying key themes and debates. Complementary searches in PubMed, Web of Science, and Google Scholar combined terms such as "*CVP*" "*preload*," "*fluid responsiveness*". Representative papers were selected to illustrate four phases in the construct's trajectory—from early physiological formulations to modern critiques. A critical interpretative synthesis [2] was used to examine how each study defined and operationalized key terms, revealing how shifting semantics sustained the CVP–preload hypothesis despite recurring empirical contradictions.

3. Results and Discussion

3.1. Semantic divergence in the CVP–preload literature

Definitional drift clouds studies of fluid responsiveness, a concept itself variably defined. Central venous pressure (CVP) and cardiac preload have long been conflated: CVP measured variably as static pressure, dynamic response, or mechanically influenced value, while cardiac preload—physiologically the ventricular end-diastolic stretch that determines stroke volume via the Frank-Starling relationship—has often been interpreted loosely as right ventricular end-diastolic volume or wall stress. This shifting terminology created the illusion of coherent evidence. Meta-analyses concluding that CVP fails to predict fluid responsiveness reveal a deeper design flaw: studies were chasing vague relationships from the outset. Decades of research thus produced pseudo-replications—experiments linked by words but divided by meaning.

The CVP–preload debate exemplifies what Yarkoni [3] terms the generalizability crisis: the systematic failure of scientific claims to accumulate meaningfully when their core constructs lack stable semantic mappings to empirical measures. Without formalized definitions of physiological concepts, biomedical research risks reproducing data rather than knowledge.

3.2. The epistemic parallel: from CVP to the aether

The CVP–preload controversy echoes the 19th-century search for the luminiferous aether. Like CVP, the aether was conceived as an invisible intermediary. The Michelson–Morley experiment initially produced a null result, but its impact was decisive because the aether hypothesis had precise mechanical definitions—its falsification was semantically unambiguous. When Dayton Miller later reported positive fringe shifts in the 1920s, physicists could conclusively rebut his findings because the underlying definitions and expectations were unchanged; the dispute could be resolved through numerical analysis of the statistical results rather than conceptual reinterpretation. [4] By contrast, the CVP–preload debate persists because biomedical terms lack this semantic stability.

In medicine, no such semantic infrastructure exists. "CVP" and "preload" remain elastic, allowing each negative result to be reinterpreted rather than resolved. Physics advanced by redefining its ontology; medicine, lacking formal semantics, continues to accumulate exceptions instead of discarding untenable constructs. Physics could abandon the aether because its constructs—velocity, mass, field—were defined within a formal mathematical language that fixed meaning and enabled falsification. Medicine, however, still relies on natural language: key terms such as "preload," "tone," or "compliance" remain metaphorical and context-dependent.

3.3. Do we have the tools to resolve semantic ambiguity?

Today, we possess tools that could close this gap. Ontologies like the OBO Foundry and SNOMED CT, systems biology models, and knowledge graphs can encode physiological entities and relations with logical precision. Yet these frameworks are rarely applied to hypothesis definition, leaving medicine rich in data but poor in semantic discipline. The CVP–preload debate shows that without formal semantics, negative findings cannot decisively falsify ambiguous hypotheses. Neurosymbolic AI offers a promising avenue to bridge this divide: by combining symbolic reasoning over formal ontologies with statistical learning from large biomedical datasets, these hybrid systems can both encode explicit conceptual definitions and infer novel patterns. This approach could automate hypothesis testing, detect inconsistencies in conceptual frameworks, and guide experimental design, effectively translating vast data into semantically coherent and actionable biomedical knowledge.

3.4. Socio-economic cost of semantic vagueness

The decades-long debate over whether central venous pressure (CVP) reliably reflects preload has not only stalled conceptual progress but also incurred substantial societal costs. A meta-review by Eskesen et al. [5] identified at least 51 individual studies addressing this question; using order of magnitude conservative estimates for study costs (200,000 USD per study; [6]), direct research expenditures likely exceed 10 million USD. This figure underestimates the total burden because unpublished studies, misdirected clinical practices, and opportunity costs are not captured. Crucially, this analysis pertains only to the CVP–preload use case. If similar semantic ambiguity pervades other domains of biomedical research—which evidence suggests is widespread—the cumulative economic and societal cost could be orders of magnitude higher, representing hundreds of millions, if not billions, of dollars diverted to conceptually fragile or pseudo-replicated investigations. Formalizing biomedical concepts could therefore yield not only epistemic clarity but also massive societal savings across the research ecosystem. An Incremental Epistemic Effectiveness Ratio (IEER)—analogous to the ICER in health economics—could quantify the cost per unit of reliable knowledge produced by a study. With formal semantic definitions, each experiment's contribution to reducing uncertainty or falsifying hypotheses becomes measurable. The CVP–preload case shows that decades of research generated abundant data but limited conceptual progress. Semantic clarity would enable funding allocation based on epistemic return on investment, turning biomedical research into a cost-effective engine of reliable discovery.

4. Conclusion

The case of central venous pressure illustrates a broader epistemological weakness in biomedical science: the absence of formal semantics at the level of hypothesis definition. Where physics evolved mathematical frameworks that enforce internal coherence and permit decisive falsification, medicine remains reliant on intuitive, metaphorical constructs whose meanings shift with context.

The CVP–preload debate has persisted not because physiology is mysterious, but because language is. Researchers have been testing different phenomena under the same name, generating an illusion of cumulative evidence. The result is a form of semantic pseudoreplication—the repetition of experiments that cannot, even in principle, converge on a shared truth condition.

If biomedical science is to attain the semantic stability that made physics cumulative, it must treat hypothesis definition as a formal act. Ontological modeling, computational semantics, and explicit definitional contracts could supply the missing infrastructure. Only then will negative experiments become not just failures to predict, but steps toward theoretical clarity. Under such semantic rigor, metrics like an Incremental Epistemic Effectiveness Ratio (IEER) could quantify the epistemic value of research, guiding funding toward studies that maximize knowledge gain relative to cost and transforming the allocation of resources into a principled, evidence-driven process.

Acknowledgments

The authors have no external funding to declare for the current research.

Declaration on Generative AI

During the preparation of this work, the author(s) used GPT5 in order to: perform a grammar and spelling check, and provide alternative formulations for selected sentences. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] P. Lloyd-Donald, M. Fujino, B. Waldman, L. F. Miles, Measurement and interpretation of central venous pressure: a narrative review, Anaesthesia 80 (2025) 1093–1102.

[2] J. Depraetere, C. Vandeviver, I. Keygnaert, T. V. Beken, The critical interpretive synthesis: an assessment of reporting practices, International Journal of Social Research Methodology 24 (2021) 669–689. URL: https://doi.org/10.1080/13645579.2020.1799637. doi:10.1080/13645579.2020.1799637. arXiv:https://doi.org/10.1080/13645579.2020.1799637.

[3] T. Yarkoni, The generalizability crisis, Behavioral and Brain Sciences 45 (2022) e1. doi:10.1017/S0140525X20001685.

[4] R. Lalli, Hunting for the luminiferous ether: The american revival of the michelson–morley experiment in the 1920s, in: Ether and Modernity: The recalcitrance of an epistemic object in the early twentieth century, Oxford University Press, 2018. URL: https://doi.org/10.1093/oso/9780198797258.003.0009. doi:10.1093/oso/9780198797258.003.0009. arXiv:https://academic.oup.com/book/0/chapter/195472443/chapter-ag-pdf/44615134/book$_2$6682$_s$*ection*$_1$95472443.*ag.pdf*.

[5] T. G. Eskesen, M. Wetterslev, A. Perner, Systematic review including re-analyses of 1148 individual data sets of central venous pressure as a predictor of fluid responsiveness, Intensive Care Med. 42 (2016) 324–332.

[6] M. Michelson, K. Reuter, The significant cost of systematic reviews and meta-analyses: A call for greater involvement of machine learning to assess the promise of clinical trials, Contemp. Clin. Trials Commun. 16 (2019) 100443.