Qeios

Peer Review

Review of: "Development and Functionality of the ABC Tool (Automated Boolean Composer) to Simplify and Standardize Search Strategy Construction in Bibliographic Databases"

Matthew Grainger[1]

1. Norwegian Institute for Nature Research, Trondheim, Norway

The paper addresses a genuine and well-recognised bottleneck in evidence synthesis: the difficulty, time cost, and error-proneness of constructing high-quality Boolean search strategies. Overall, this is a clear, well-written, and practically valuable paper. The tool is well-motivated, and the validation logic is sensible and grounded in established information-retrieval methods.

However, the paper currently reads more as a product description plus an internal validation report than as a full methods paper. Several claims (efficiency gains, error reduction, reproducibility improvements) are plausible but not empirically demonstrated, and the validation approach, while rigorous, would benefit from clearer reporting boundaries and acknowledgement of its epistemic limits.

Major comments

The use of relative recall against included studies from Epistemonikos systematic reviews is appropriate and well-justified. However, the manuscript should more explicitly acknowledge the closed-loop nature of this validation framework. The reference standard is derived entirely from Epistemonikos-indexed reviews. Strategies are validated against studies already retrieved by prior search processes (which may share biases with the ABC approach). As a result, the validation primarily demonstrates consistency with existing Epistemonikos retrieval practices, rather than absolute sensitivity in an open universe of studies. This does not invalidate the method, but it does constrain the interpretation of “95% sensitivity”. I recommend explicitly framing this as “high sensitivity relative to Epistemonikos systematic review corpora” and not as a general guarantee of recall across all bibliographic contexts.

doi.org/10.32388/I0C2GJ

Claims about efficiency, error reduction, and reproducibility are not empirically tested. The Discussion repeatedly claims that the ABC Tool reduces workload, minimises human error, improves reproducibility, and substantially cuts the time required to build strategies. These claims are highly plausible but currently asserted rather than demonstrated. There is no user study, no comparative timing exercise, no error-rate comparison between manual vs. ABC-assisted strategies, etc.

The threshold choices (95% sensitivity, NNR $\leq$ 50) need stronger justification. These are reasonable choices, but the justification is currently implicit rather than explicit. Why 95% rather than 88% or 94%? Are these thresholds applied uniformly across concept types?

The tool is presented as broadly relevant to "evidence synthesis", but in practice, the taxonomy is health-focused, the validation corpus is biomedical, and the implementation is database-specific. This is fine, but the boundary conditions should be clearer. What would be required to adapt it to other domains (e.g., environmental science, social science)?

Minor comments

The tool is referred to as both ABC Tool and Automated Boolean Strategy Creator / Composer. Standardising the name would reduce minor confusion.

Figures - consider adding one conceptual schematic showing: taxonomy → strategy → validation → user interface, which would help readers grasp the architecture more quickly.

Conflict of interest statement - The declaration is transparent and appropriate. You may wish to explicitly note that open access to the strategies and taxonomy mitigates some developer bias concerns.

About the authors – this is not about the authors; it is about the foundation (which is great, but I would like to learn something about the authors!)

Declarations

Potential competing interests: No potential competing interests to declare.