


THE ECOSYSTEM OF EVIDENCE

Navigating the future
of the ecosystem of evidence

Artificial intelligence and digital transformation
in research and healthcare

11th EBHC International Conference
Taormina, 21st – 24th October 2026

#EBHC2026

# Enhancing Systematic Reviews through Neuro-Symbolic AI

## Logic Augmented Generation for Evidence-Based Health Care

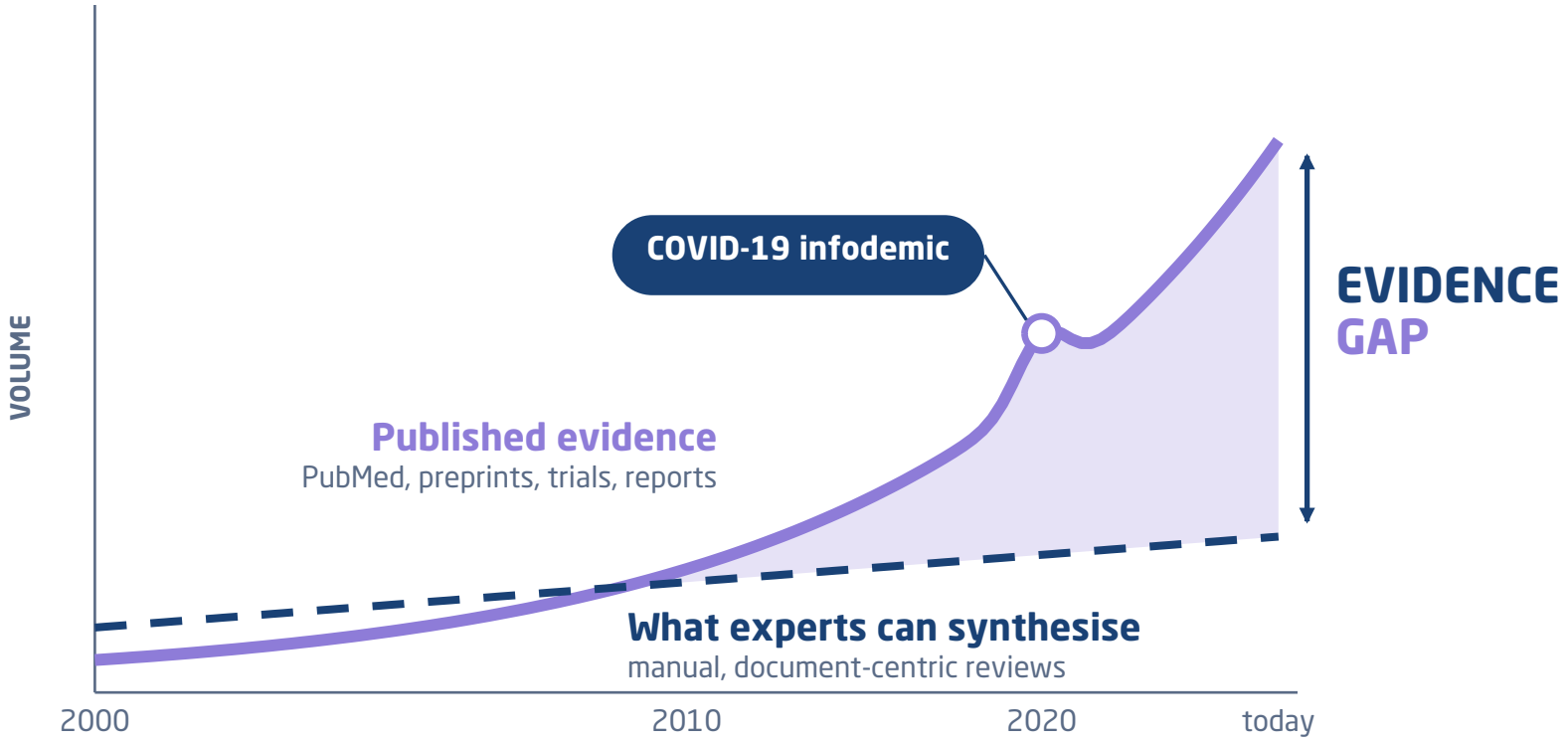
L. Bulla, D. Golinelli, M. Mongiovì, A. Russo, V. Trianni, A. Gangemi, A.G. Nuzzolese

**Luana Bulla**
University of Bologna
ISTC – CNR

30th GIMBE 1996-2026 EVIDENCE FOR HEALTH

# The Evidence Gap

**Evidence grows exponentially. Our capacity to synthesise it does not.**




**1M+**
new PubMed records every year — about 3,000 a day

**6-7 wk**
to complete and publish a single systematic review

**88K**
COVID-19 papers in 2020 alone

**The gap is structural:** manual synthesis alone cannot give EBHC a timely, traceable foundation.

# Background

Evidence synthesis is no longer keeping pace with the literature it has to summarise.

### Literature overload

Exponential growth of publications. COVID-19 exposed the limits of document-centric repositories and manual workflows.

### Large Language Models

Scale to unstructured text, but offer no guarantees of consistency or traceability.

### Semantic Knowledge Graphs

Provide logical rigour and explicit semantics, but have limited adaptability.

**Bridging neural and symbolic paradigms** is essential for next-generation EBHC infrastructures

# Aims

Support experts in synthesising evidence from **multimodal, heterogeneous** literature through a **dynamic, computable and traceable** ecosystem for hybrid human-AI EBHC.

| P | I | C | O |
|---|---|---|---|
| Population | Intervention | Comparison | Outcome |

PICO formalised as an ontological scaffold

## Neuro-symbolic framework

Logic Augmented Generation (LAG): Knowledge Graphs + LLMs under explicit logical constraints

## Biomedical ontologies

Grounded in SNOMED CT, UMLS and ICD-11 terminologies

## PICO as scaffold

Structures and synthesises evidence around Population, Intervention, Comparison, Outcome

# Methods

Three integrated components, with experts in the loop at every stage.

1

## PICO sentence classification

Automated systematic-review support: abstract sentences are labelled with PICO elements

2

## Knowledge graph construction

LLM-assisted generation linked to biomedical ontologies, validated by domain experts

3

## Logic Augmented Generation

LLM outputs constrained by KG-based logical boundaries, preserving ontological coherence

**Human-in-the-loop:** experts filter, validate and analyse the structured evidence

# Methods

PICO classifier: a deep neural architecture that labels each sentence of an abstract.




## Six sentence labels

| | |
|---|---|
| A | Aim |
| P | Participants |
| I | Intervention + Comparison |
| M | Method |
| O | Outcome + Results |
| C | Conclusion |

Trained on 24,668 RCT abstracts from Embase (automatically annotated); an LLM then builds the KG from the labelled sentences.

# Results

Case study: early COVID-19 literature on pre-exposure prophylaxis (Embase).

**50**
abstracts randomly selected

**752**
sentences extracted

**584**
sentences validated by two experts

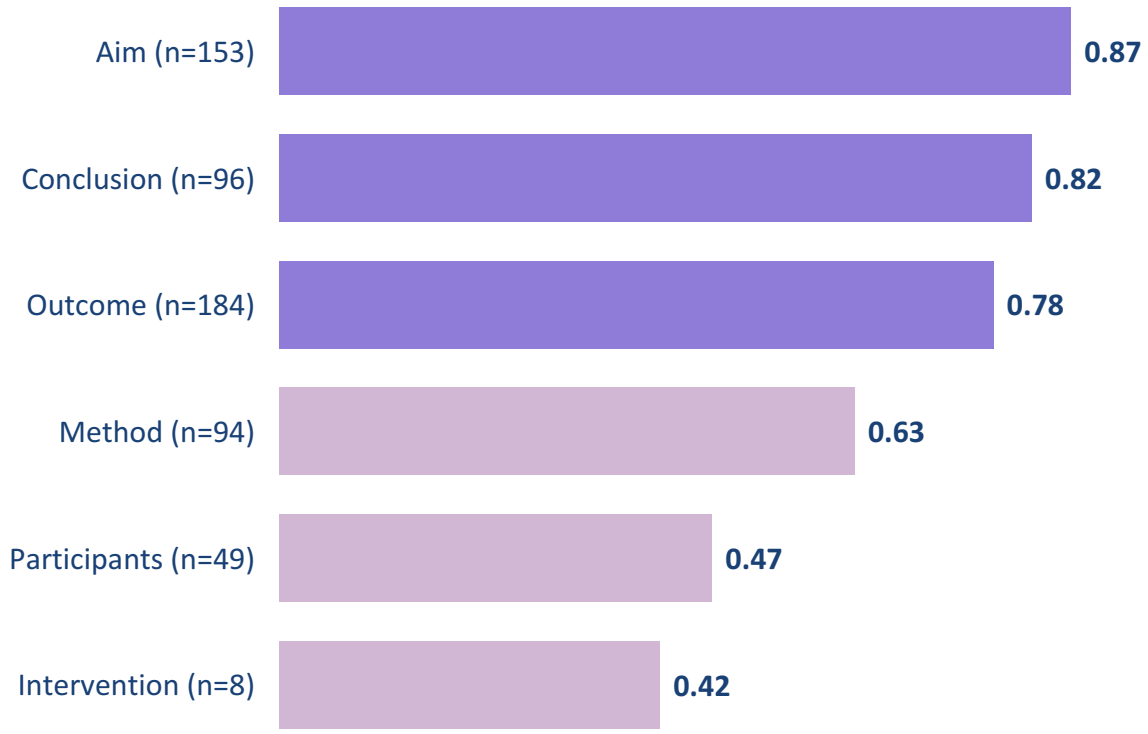
**76%**
overall accuracy

**Robust to domain shift:** trained on randomised controlled trials, tested on heterogeneous pandemic publications.

Weighted F1 = 0.76 → semi-automated PICO structuring can support rapid evidence screening in emergencies.

# Results

## Per-label performance (F1) and complementary results of the framework.




### Strong on frequent labels

F1 ≥ 0.78 for Aim, Conclusion, Outcome; weaker on rare ones

### LLM-assisted ontology construction

Reduces modelling effort under expert validation

### Logic Augmented Generation

Controlled KG extension aligned with biomedical terminologies

# Limits

## Uneven PICO performance

Lower scores for under-represented elements (Intervention, Participants)

## Probabilistic LLM outputs

Generated content still requires expert validation

## Ontology alignment

Interoperability challenges across institutional contexts

# Conclusions

**Static document repositories**

→

**Computable, traceable evidence ecosystems**

- PICO embedded in ontology-driven knowledge graphs
- Generative models constrained by formal logic (LAG)
- Transparent, scalable systematic reviews

**Next step**

Combine symbolic rigour with controlled generative intelligence for adaptive, accountable, evidence-informed decision-making.

ebhc CONFERENCE 2026

THE ECOSYSTEM OF EVIDENCE

Navigating the future
of the ecosystem of evidence

Artificial intelligence and digital transformation
in research and healthcare

11th EBHC International Conference
Taormina, 21st – 24th October 2026

#EBHC2026

30th 1996-2026 GIMBE EVIDENCE FOR HEALTH

# Thank you!

luana.bulla@unibo.it

**Luana Bulla**
University of Bologna
ISTC – CNR