# **Artificial intelligence (AI) driven scientific evidence in the HTA:** barriers and recommendations to address them

**Antal Zemplenyi**
**Syreon Research Institute, University of Pécs, Hungary**

**ISPOR Europe 2023**
**November 13, Copenhagen**

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# AI in HTA context

A data analytics tool supporting evidence generation for HTA

Machine learning:

- cohort selection, feature selection, causal inference, or developing economic models, etc.

Natural language processing:

- evidence synthesis from the literature, or extracting  key parameters from textual notes (EMR), etc.

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Perspective and Aim

Central and Eastern European (CEE) countries

- lag behind Western Europe in HTA implementation
- have single-payer healthcare systems and large claims databases
- face substantial barriers to implementing AI-driven evidence generation in HTA

Aim

- to provide recommendations to support healthcare decision-makers in properly integrating AI into the HTA methodologies and processes

# Methods

Next Generation Health Technology Assessment

## Identifying barriers

- Literature review

## Selecting top barriers

- Survey of 77 respondents from 11 countries

## Developing recommendations

- Workshop with 23 experts representing 12 countries

ORIGINAL RESEARCH article
Front. Public Health, 14 July 2022
Sec. Digital Public Health
Volume 10 - 2022 | https://doi.org/10.3389/fpubh.2022.921226

### Barriers to Use Artificial Intelligence Methodologies in Health Technology Assessment in Central and East European Countries

Konstantin Tachkov[1] Antal Zemplenyi[2,3] Maria Kamusheva[1] Maria Dimitrova[1] Pekka Siirtola[4] Johan Pontén[5] Bertalan Nemeth[2] Zoltan Kalo[2,6] Guenka Petrova[1*]

ORIGINAL RESEARCH article
Front. Public Health, 26 April 2023
Sec. Digital Public Health
Volume 11 - 2023 | https://doi.org/10.3389/fpubh.2023.1088121

### Recommendations to overcome barriers to the use of artificial intelligence-driven evidence in health technology assessment

Antal Zemplényi[1,2*] Konstantin Tachkov[3] Laszlo Balkanyi[4] Bertalan Németh[2] Zsuzsanna Ida Petykó[2] Guenka Petrova[3] Marcin Czech[5] Dalia Dawoud[6,7] Wim Goettsch[8,9] Inaki Gutierrez Ibarluzea[10] Rok Hren[11] Saskia Knies[9] László Lorenzovici[12,13] Zorana Maravic[14] Oresta Piniazhko[15] Alexandra Savova[3,16] Manoela Manova[3,16] Tomas Tesar[17] Spela Zerovnik[18] Zoltán Kaló[2,19]

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Barriers identified in the review[3]

| | |
|---|---|
| **Data related barriers (11)** | Systemic bias in the data (e.g. due to upcoding)<br>Issues with reliability, validity and accuracy of data (e.g. due to the lack of quality assessment of data entry or self-reporting)<br>Raw fragmented or unstructured data (e.g. electronic medical records, imaging reports), which are difficult to aggregate and analyze<br>Data acquisition and cleansing is not feasible<br>Analysis of multicenter data is limited due to the lack of interoperability across systems (e.g. electronic medical records of different service providers)<br>Lack of well-described patient level health databases<br>Multinational data collection and analysis is limited due to differences in coding system across countries, and the lack of mapping methods to standardize the vocabulary<br>Data that are relevant for research purposes are missing from databases built for the healthcare financing or provision (e.g. important clinical endpoints)<br>The database is incomplete to fully track patient pathways, leading to inconsistent, unreliable findings<br>Sample size of the available databases are low (e.g. databases of health care providers are not linked) |
| **Methodological barriers (6)** | Potential bias of AI to favor some subgroups based on having more or better information<br>Lack of transparency of protocols for data collection methods<br>Text mining and natural language processing algorithms cannot be applied due to the lack of standardized medical terms in the local language<br>Limited reproducibility due to the complexity of the methods<br>Lack of methodological transparency of deep learning models (“black box” phenomenon)<br>Complexity of the diseases and co-morbidities |
| **Technological barriers (3)** | Lack of capacity to build and maintain IT infrastructure to support AI process<br>High costs associated with securing and storing data for research purposes<br>High cost of improving data validity (e. g. data abstracters to evaluate unstructured data) |
| **Regulatory and policy related barriers (5)** | Regulatory compliance issues in the process of managing high volume of sensitive information<br>Lack of awareness and openness on the part of decision-makers to rely on AI based real-world evidence<br>Lack of political commitment (e.g. no health digitization strategy in the country to establish relevant databases)<br>Acceptance and consent by patients and medical professionals<br>Lack of access to patient-level databases due to data protection regulations |
| **Human factor related barriers (5)** | Lack of knowledge in data governance: data ownership and data stewardship<br>Lack of appropriate skills for applying AI methods (natural language processing, machine learning etc.) in outcomes research<br>Lack of adequate education to generate AI driven scientific evidence<br>Lack of decision-makers’ expertise about the methods and use of AI driven scientific evidence |

[3]Tachkov, et al. (2022). Barriers to use Artificial Intelligence methodologies in Health Technology Assessment in Central and East European countries. Frontiers in Public Health. doi: 10.3389/fpubh.2022.921226

| Rank | Barrier | Mean Likert score | Barrier group |
|---|---|---|---|
| 1 | Lack of decision-makers' expertise about the methods and use of AI driven scientific evidence | 4.03 | H |
| 2 | Lack of appropriate skills for applying AI methods (natural language processing, machine learning etc.) in outcomes research | 3.90 | H |
| 3 | Lack of adequate education to generate AI driven scientific evidence | 3.90 | H |
| 4 | Issues with reliability, validity and accuracy of data (e.g. due to the lack of quality assessment of data entry or self-reporting) | 3.88 | D |
| 5 | Lack of awareness and openness on the part of decision-makers to rely on AI-based real-world evidence | 3.85 | R |
| 6 | Lack of political commitment (e.g. no health digitization strategy in the country to establish relevant databases) | 3.81 | R |
| 7 | Multinational data collection and analysis is limited due to differences in the coding system across countries, and the lack of mapping methods to standardize the vocabulary | 3.68 | D |
| 8 | Lack of resources to build and maintain IT infrastructure to support AI process | 3.68 | T |
| 9 | Regulatory compliance issues in the process of managing a high volume of sensitive information | 3.67 | R |
| 10 | Analysis of multicentre data is limited due to differences in database structures across systems (e.g. electronic medical records database of different service providers) | 3.63 | D |
| 11 | Raw fragmented or unstructured data (e.g. electronic medical records, imaging reports), which are difficult to aggregate and analyse | 3.62 | D |
| 12 | High cost of improving data validity (e.g. data abstracters to evaluate unstructured data) | 3.62 | T |
| 13 | Systemic bias in the data (e.g. due to upcoding) | 3.59 | D |
| 14 | Lack of well-described patient-level health databases | 3.59 | D |
| 15 | Data that are relevant for research purposes (e.g. important clinical endpoints) are missing from databases or are available only on paper. | 3.54 | D |
| 16 | The database is incomplete to fully track patient pathways, leading to inconsistent, unreliable findings | 3.49 | D |
| 17 | High costs associated with securing and storing data for research purposes | 3.47 | T |
| 18 | Data is not transferable across countries for multinational analyses | 3.46 | D |
| 19 | Lack of access to patient-level databases due to data protection regulations | 3.42 | R |
| 20 | Lack of knowledge in data governance: data ownership and data stewardship | 3.42 | H |
| 21 | Lack of transparency of protocols for data collection methods | 3.37 | M |
| 22 | Data cleansing is not feasible | 3.28 | D |
| 23 | Lack of methodological transparency of deep learning models ("black box" phenomenon) | 3.18 | M |
| 24 | Potential bias of AI to favour some subgroups based on having more or better information | 3.15 | M |
| 25 | Sample size of the available databases are low (e.g. databases of health care providers are not linked) | 3.13 | D |
| 26 | Acceptance and consent by patients and medical professionals | 3.13 | R |
| 27 | Text mining and natural language processing algorithms cannot be applied due to the lack of standardized medical terms in the local language | 3.12 | M |
| 28 | The result of analysing complex diseases with AI is difficult to use in health economic models | 3.10 | M |
| 29 | Limited reproducibility due to the complexity of AI methods | 3.06 | M |

Top barriers: average score ≥3.5

# Selected barriers by groups

1 (very low importance) – 5 (very high importance)

| Barrier group | Avg score | Selected |
|---|---|---|
| Human factor related barriers | 3.8 | 3/4 |
| Regulatory and policy-related barriers | 3.6 | 3/5 |
| Technological barriers | 3.6 | 2/3 |
| Data related barriers | 3.5 | 7/11 |
| ~~Methodological barriers~~ | 3.2 | 0/6 |
| **Total** | **3.5** | **15/29** |

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Human factor related barriers

Limited AI expertise for decision-makers, lack of skills in methods like NLP and ML, and lack of education to generate AI-driven scientific evidence.

**Recommendations:**

- More **reliance on expertise in academic centers**
- Ensuring diverse representation of WE and CEE countries in international collaborative projects (Horizon Europe, IHI) to facilitate **knowledge transfer in AI methods or database developments**
- Create or adopt a **virtual platform for experience exchange** between countries (e.g., Decide Health Decision Hub)
- Developing **targeted (AI for HTA) educational material**

© The HTx Consortium 2019-2023. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Regulatory and policy related barriers

Next Generation Health Technology Assessment

Limited decision-maker awareness and openness to AI-based real-world evidence, insufficient political commitment for health digitization and data collection, and regulatory compliance challenges in handling sensitive information.

**Recommendations:**

- Advocate **more reliance on claims databases** in health policy decisions, leading to better regulation of their use
- Join "**federated network studies**" to run analysis without the need to share or centralize datasets (e.g., EHDEN, DARWIN EU)
- Establish **trusted research environments** for the safe use of data in research

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Data related barriers

Interoperability issues, incl. coding system and database structure differences

Data incompleteness, unstructured data, reliability, validity and accuracy issues

**Recommendations:**

- **Standardization through common data models** (e.g. OMOP CDM): structural and terminological mapping → allows for **federated network studies**
- **Local coding systems to be harmonized** with standards when updated
- Adapting **data abstraction methods** (e.g., RegExp) to retrieve key variables from texts
- Improve **standardized and structured data entry in routine practice**
- **Link databases** (e.g., claims with EHR)
- A **chapter on limitations of the AI-based evidence (also in lay language)** should be a standard part of the HTA report

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Technological barriers

Lack of resources to build and maintain IT infrastructure

**Recommendations:**

- Develop infrastructure and human **capacities for AI in centers**
- Use **technologies that are scaled to the needs** of the analysis
- **Sustainability plan to** be presented for publicly funded database or AI-related project (e.g., Horizon Europe, IHI)

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

# Key takeaways

- **AI can enhance HTA** by providing data-driven insights, but it **needs rigorous evaluation**
- **CEE countries have unique chances** in AI due to (mostly) single-payer systems but **face barriers**
- Recommendations include **enhancing expertise, international collaboration, regulatory enhancements**, infrastructure development
- **Data standardization** is vital for AI-driven HTA, including common data models and harmonized coding systems. **Participation in data initiatives** is key for influencing EU-level decisions

 This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825162.

Thank you for your kind attention!

antal.zemplenyi@syreon.eu

# References

Tachkov K, Zemplenyi A, Kamusheva M, Dimitrova M, Siirtola P, Pontén J, Nemeth B, Kalo Z and Petrova G (2022) Barriers to Use Artificial Intelligence Methodologies in Health Technology Assessment in Central and East European Countries. Front. Public Health 10:921226. https://doi.org/10.3389/fpubh.2022.921226

Zemplényi A, Tachkov K, Balkanyi L, Németh B, Petykó ZI, Petrova G, Czech M, Dawoud D, Goettsch W, Gutierrez Ibarluzea I, Hren R, Knies S, Lorenzovici L, Maravic Z, Piniazhko O, Savova A, Manova M, Tesar T, Zerovnik S and Kaló Z (2023) Recommendations to overcome barriers to the use of artificial intelligence-driven evidence in health technology assessment. Front. Public Health. 11:1088121. https://doi.org/10.3389/fpubh.2023.1088121