BRAIN. Broad Research in Artificial Intelligence and Neuroscience
e-ISSN: 2067-3957 | p-ISSN: 2068-0473
Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)
2026, Volume 17, Issue 3, pages: 32-42
Submitted: June 27th, 2026 | *Accepted for publication:* August 14th, 2026

Preparing the Health Workforce for AI: Linking Health-System AI Adoption to Pre-Service and In-Service Training across the WHO European Region

Muhammet Berigel
Prof. Dr., Department of Management Information Systems, Karadeniz Technical University, Trabzon, Türkiye.
Email: berigel@ktu.edu.tr
https://orcid.org/0000-0001-5682-8956

İslam Suiçmez
Department of Education, University of Kyrenia, Lefkoşa, via Mersin, Türkiye.
Email: islam.suicmez@kyrenia.edu.tr
https://orcid.org/0009-0003-8159-5236

Zehra Altinay
Societal Research and Development Centre, Faculty of Education, Near East University, Nicosia, Mersin 10, Turkey.
zehra.altinaygazi@neu.edu.tr,
https://orcid.org/0000-0002-6786-6787

Gokmen Dagli
Faculty of Education, University of Kyrenia, Kyrenia, Mersin 10, Turkey.
gokmen.dagli@kyrenia.edu.tr,
https://orcid.org/0000-0002-4416-8310

Abstract: *Artificial intelligence (AI) is being integrated into health systems at different speeds, raising questions about whether health-professional education is developing alongside technological implementation. This cross-sectional, country-level study examined associations between health-system AI characteristics and pre-service and in-service AI training across the WHO European Region using 2025 data from the World Health Organization Regional Office for Europe's Artificial Intelligence for Health in the WHO European Region (AIRA) dataset. The overall dataset included 50 countries, with analytical sample sizes varying by indicator because of unknown or missing responses. Pre-service and in-service training status were examined in relation to three health-system measures: AI Application Implementation, AI Opportunity, and AI Adoption Barrier Burden. Composite indicators were derived from conceptually related AIRA items, with equal weighting within each index and "Don't know" responses treated as unknown rather than negative. Descriptive statistics and Kruskal-Wallis tests were used, with Holm correction for multiple testing, Dunn-Holm post-hoc comparisons where appropriate, and epsilon-squared effect sizes. Established pre-service training was reported by 20% of countries and established in-service training by 24%. The clearest association was observed for in-service training: AI Application Implementation differed significantly across training levels after Holm correction ($H(2)=10.21$, adjusted $p=.0182$, $\varepsilon^2=.265$), with established-training countries showing higher implementation scores than both the No and Under development groups. For pre-service training, implementation scores showed an ordered descriptive pattern, but the Holm-adjusted omnibus result did not meet the .05 threshold (adjusted $p = .0501$). AI Opportunity and AI Adoption Barrier Burden did not significantly distinguish training levels after correction. The findings indicate that established in-service AI training is associated with broader reported AI implementation at the health-system level, while the cross-sectional design does not permit causal inference.*

Keywords: *artificial intelligence; health professions education; health workforce; pre-service training; in-service training; digital health; WHO European Region.*

How to cite: Berigel, M., Suiçmez, İ., Altinay, Z., & Dagli, G. (2026). *Preparing the health workforce for AI: Linking health-system* AI adoption to pre-service and in-service training across the WHO European region. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 17*(3), 32-42. https://doi.org/10.70594/brain/17.3/2

1. Introduction

Artificial intelligence (AI) is increasingly used in healthcare for clinical decision support, diagnosis and imaging, patient monitoring, service delivery, and administrative processes. The World Health Organization (WHO) has emphasised that realising the potential of AI requires not only technological capacity but also governance, regulation, implementation capability, and safeguards for ethical and equitable use (World Health Organization [WHO], 2021, 2024).

These developments are also changing the competencies expected of the health workforce. As AI systems move into routine health-system functions, professionals need sufficient knowledge to use, evaluate, and question AI-supported decisions safely. Education is therefore an important component of AI readiness and can be considered across two complementary stages: preparation before professional practice and continuing development after entry into the workforce.

Accordingly, this study distinguishes between pre-service AI training and in-service AI training.

1.1. AI in Health Professions Education and Pre-Service Preparation

Evidence from health-professions education consistently identifies a gap between positive attitudes toward AI and formal educational preparation. A systematic review found that healthcare students generally viewed AI positively but often reported limited knowledge and practical skills (Mousavi Baigi et al., 2023). Similarly, a global mixed-methods study of medical students from 48 countries found strong support for including AI in core medical curricula, while formal AI training remained uncommon (Ejaz et al., 2022).

Large international samples reinforce this educational gap. In a survey covering 192 medical, dental, and veterinary faculties across 48 countries, 76.3% of participants reported that they had not taken an AI-related course despite substantial interest in further AI education (Busch et al., 2024). A multi-country study of health-professions students likewise reported limited AI knowledge and substantial reliance on sources outside formal curricula (Bani Issa et al., 2024).

Reviews of AI education further show that consensus on the need for curricula has developed faster than agreement on curriculum content, timing, pedagogy, and assessment. Existing programmes remain heterogeneous and are often insufficiently described in terms of educational frameworks and delivery models (Lee et al., 2021; Tolentino et al., 2024).

Recent competency-focused work similarly concludes that recommendations for future physicians remain fragmented and supports more structured, competency-based approaches spanning technical foundations, critical appraisal, ethics, regulation, and clinical use (Hunt et al., 2026).

1.2. In-Service AI Training and Continuing Professional Development

For professionals already in practice, continuing education may be especially important because health-system technologies can change more rapidly than formal curricula. A scoping review of AI education programs for healthcare professionals identified deficiencies in AI literacy as an important challenge to workforce preparedness and implementation (Charow et al., 2021).

In-service education can therefore provide a mechanism for updating competencies as AI applications are introduced into operational settings. However, evidence linking the institutionalisation of such training to country-level patterns of AI implementation remains limited.

1.3. The Gap Between AI Adoption and Educational Preparedness in Health Systems

The integration of AI into health systems depends on multiple dimensions, including technological infrastructure, governance, financing, regulatory capacity, data readiness, and workforce capability. The WHO Regional Office for Europe has documented substantial cross-country variation in AI readiness across these domains (WHO Regional Office for Europe, 2025b). Yet most research on AI education has focused on students' or professionals' knowledge, attitudes, and educational preferences rather than on whether training status is associated with health-system implementation at the country level. This study addresses that gap by using standardised WHO/Europe data from 50 countries to compare pre-service and in-service AI training separately against three distinct health-system characteristics: reported AI application implementation, perceived AI opportunities, and perceived adoption barriers. This country-level comparison extends the literature by examining educational preparedness alongside operational implementation rather than treating AI education only as an individual-level learning outcome.

2. Method

2.1. Study Design and Data Source

A cross-sectional, country-level secondary data analysis was conducted to examine associations between health-system AI characteristics and the development status of AI training for health professionals. Data were obtained from the World Health Organization Regional Office for Europe's Artificial Intelligence for Health in the WHO European Region (AIRA) dataset, available through the European Health Information Gateway (WHO Regional Office for Europe, 2025a). All indicators used in the analysis referred to the 2025 reporting period. The country was the unit of analysis.

2.2. Study Sample

The AIRA source files use a standard country framework for the WHO European Region. Usable observations for at least one of the indicators retained in this study were available for 50 countries. Bosnia and Herzegovina, Monaco, and Turkmenistan contained no usable observations across the selected analytical indicators and were not included. Because the unit of analysis was the country, the findings describe health-system-level patterns rather than individual health professionals' knowledge, skills, attitudes, or behaviours.

2.3. Data Integration, Cleaning, and Refinement

Individual AIRA indicator files were downloaded separately and merged using country code and reference year as common identifiers. Country names, codes, response categories, years, and missing values were checked for consistency before integration. The broader merged dataset initially contained 27 AIRA indicators covering AI strategy and governance, data readiness, applications, opportunities, barriers, ecosystem characteristics, and professional training. The integration workflow followed the standard secondary-data sequence of data understanding, preparation, cleaning/transformation, and validation described for structured health datasets (Areco et al., 2021).

To align the analytical dataset with the study question, only directly relevant indicators were retained. The final model retained 20 raw AIRA indicators: two AI training indicators (AIRA_73 and AIRA_74), seven AI application indicators (AIRA_47-53), four AI opportunity indicators (AIRA_54, AIRA_56-58), and seven AI adoption-barrier indicators (AIRA_60, AIRA_62-64, AIRA_66, AIRA_68-69). 'Don't know' responses were retained as unknown information and were not recorded as 'No'. Derived measures were created only for conceptually related indicator groups, and the original item-level responses were retained in the analysis-ready dataset to support reproducibility.

2.4. AI Training Outcome Variables

Two AIRA indicators represented AI training readiness for the health workforce. AIRA_73 (Pre-service AI training) captures AI-related training before entry into professional practice, whereas AIRA_74 (In-service AI training) captures AI-related education or professional development for health professionals already working in the health system. Both use four response categories: 'No', 'Under development', 'Yes', and 'Don't know'. Inferential comparisons used the three substantive categories ('No', 'Under development', and 'Yes'). 'Don't know' responses were excluded from the relevant comparison because unknown training status cannot be interpreted as absence of training.

2.5. Health-System AI Indicators

Three health-system dimensions were selected because they represented conceptually distinct aspects of AI readiness directly relevant to the research question: operational implementation of AI applications, perceived opportunities associated with AI, and perceived barriers to adoption. Their composition, scoring, and interpretation are summarised in Table 1.

Table 1. Health-system AI indicators included in the analysis

Construct	AIRA indicators	Scoring
AI Application Implementation	AIRA_47-53: logistics/administration, chatbots, AI-assisted surgery, diagnostics, prognosis prediction, symptom checkers, remote monitoring	Yes = 1; No = 0; Don't know = missing. Score = implemented applications / known responses × 100. At least 5 of 7 valid responses required.
AI Opportunity	AIRA_54, 56, 57, 58: reduce inequalities, reduce workforce pressure, increase efficiencies, advance research and discovery	No relevance = 0; Minor = 1; Moderate = 2; Major = 3. Summed score: 0-12.
AI Adoption Barrier Burden	AIRA_60, 62, 63, 64, 66, 68, 69: legal uncertainty, infrastructure, data quality/standards, financial affordability, capacity, trust, cultural barriers	No importance = 0; Minor = 1; Moderate = 2; Major = 3. Summed score: 0-21.

The AI Application Implementation measure was conceptualised as a formative breadth index because its seven components represent distinct AI use cases rather than interchangeable indicators of a single latent trait. Each known application domain therefore contributed equally to the index. Differential weights were not applied because the AIRA framework does not provide validated weights indicating that one application domain should contribute more strongly than another to implementation breadth; applying such weights would therefore introduce an additional arbitrary assumption. The AI Opportunity and AI Adoption Barrier Burden measures were also calculated using equal item weights because their component items share the same response metric and were intended to provide transparent summary indices of conceptually related domains. Internal consistency was examined as a descriptive check of aggregation rather than as evidence of unidimensional scale validity; Cronbach's α was .789 for AI Opportunity and .708 for AI Adoption Barrier Burden. This distinction is consistent with recommendations that aggregation should be guided by both conceptual coherence and the measurement role of the component indicators (Boateng et al., 2018).

2.6. Statistical Analysis

Descriptive analyses summarised pre-service and in-service AI training using frequencies and percentages, and the three health-system AI measures using medians, interquartile ranges (IQRs), minima, and maxima. Because the groups were small and unequal and the scores were bounded or derived from ordinal responses, non-parametric comparisons were used consistently. Kruskal-Wallis H tests compared the three independent country groups defined by training status ('No', 'Under development', and 'Yes') (Nahm, 2016).

For each training outcome, AI Application Implementation, AI Opportunity, and AI Adoption Barrier Burden were compared across the three groups, producing three omnibus tests for pre-service training and three for in-service training. Holm's sequential correction was applied separately within the pre-service and in-service families to control familywise type I error while retaining greater power than simple Bonferroni correction (Chen et al., 2017). When a Holm-adjusted omnibus test remained significant, Dunn pairwise comparisons with Holm adjustment were used to identify the differing groups. Epsilon-squared (ε^2) was reported as the effect-size measure. Statistical significance was evaluated at two-sided adjusted $p < .05$.

For AI Application Implementation, a sensitivity analysis repeated the Kruskal-Wallis comparison using only countries with complete responses for all seven application indicators. The analysis assessed whether the direction and magnitude of the implementation pattern depended on allowing limited item-level missingness in the primary score. It was not used to replace the multiplicity-adjusted primary inference. Analyses were conducted in Python 3.13 using SciPy and statsmodels.

2.7. Missing-Data Approach

No numerical imputation was performed. 'Don't know' was treated as unknown information rather than absence. For the primary AI Application Implementation Score, a country was retained when at least five of the seven application indicators were known. This criterion ensured that each score represented information from a substantial majority of application domains (at least approximately 70%) while avoiding unnecessary exclusion of countries with limited item-level missingness. The score was standardised as the proportion of implemented applications among known responses. The complete-case sensitivity analysis required valid responses for all seven application indicators. Analytical sample sizes therefore differed across comparisons: 'Don't know' responses for the relevant training indicator were excluded, and Application Implementation analyses additionally required sufficient application-level information. This approach follows recommendations that missing or uncertain values should be handled according to their substantive meaning because inappropriate replacement or deletion can affect bias, efficiency, and statistical validity (Kwak & Kim, 2017).

3. Findings

3.1. AI Training Readiness and Health-System AI Profile

AI-related training status varied across the 50 countries (Table 2). Established pre-service AI training was reported by 10 countries (20.0%) and established in-service training by 12 (24.0%). Pre-service training was under development in 16 countries (32.0%) and in-service training in 12 (24.0%); 13 countries (26.0%) reported no pre-service training and 15 (30.0%) reported no in-service training. Eleven countries (22.0%) reported 'Don't know' for each training indicator.

Table 2. AI training readiness and health-system AI characteristics across the WHO European Region

Measure	Category / Valid n	n (%) / Median	IQR	Range
Panel A. AI training readiness				
Pre-service AI training	No	13 (26.0%)		
	Under development	16 (32.0%)		
	Yes	10 (20.0%)		
	Don't know	11 (22.0%)		
In-service AI training	No	15 (30.0%)		
	Under development	12 (24.0%)		
	Yes	12 (24.0%)		
	Don't know	11 (22.0%)		
Panel B. Health-system AI profile				
AI Application Implementation Score	n = 38	57.1	17.9-100.0	0-100
AI Opportunity Score	n = 50	9.0	7.0-11.0	0-12
AI Adoption Barrier Burden	n = 50	14.0	12.0-16.0	0-21

Note. AI Application Implementation was calculated only for countries with valid information for at least five of seven application indicators. Higher Application and Opportunity scores indicate broader AI implementation and stronger perceived opportunities, respectively; higher Barrier Burden scores indicate greater perceived barriers.

Among the 38 countries meeting the minimum data requirement for AI Application Implementation, the median score was 57.1 (IQR = 17.9-100.0). Median AI Opportunity was 9.0 of 12 (IQR = 7.0-11.0), while median AI Adoption Barrier Burden was 14.0 of 21 (IQR = 12.0-16.0). Thus, relatively high perceived opportunities coexisted with substantial reported barriers across the region.

3.2. Health-System AI Characteristics Across AI Training Levels

Health-system AI characteristics were compared across the three substantive training groups ('No', 'Under development', and 'Yes'), with 'Don't know' responses excluded. Analytical N varied by indicator because Application Implementation additionally required at least five known application responses. Table 3 therefore reports both the total analytical N and the group-specific n values for each comparison.

3.2.1. Pre-Service AI Training

AI Application Implementation showed an ordered descriptive pattern across pre-service training categories, with the highest median in countries reporting established training. The omnibus test was significant before multiplicity correction ($H(2) = 8.19$, raw $p = .0167$), but did not remain statistically significant after Holm correction (adjusted $p = .0501$, $\varepsilon^2 = .206$). The pattern is therefore interpreted descriptively rather than as statistically supported evidence of an association in the primary analysis.

AI Opportunity did not differ across pre-service training groups ($H(2) = 2.58$, Holm-adjusted $p = .2753$, $\varepsilon^2 = .016$). AI Adoption Barrier Burden also did not remain significant after correction ($H(2) = 6.57$, adjusted $p = .0750$, $\varepsilon^2 = .127$), although established-training countries had a lower descriptive median than the other groups (Table 3).

3.2.2. In-Service AI Training

For in-service training, AI Application Implementation differed significantly across the three groups ($H(2) = 10.21$, raw $p = .0061$, Holm-adjusted $p = .0182$, $\varepsilon^2 = .265$). Countries with established in-service training had the highest implementation scores (Table 3).

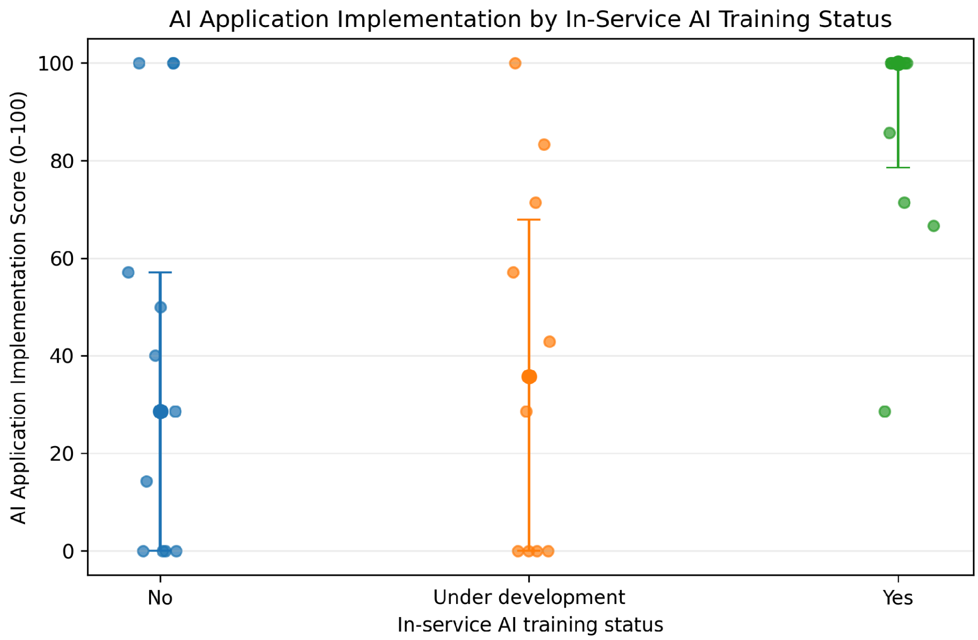

Figure 1. Distribution of AI Application Implementation scores by in-service AI training status.

Dunn-Holm post-hoc comparisons showed higher AI Application Implementation in the established in-service training group than in both the No group (adjusted $p = .0159$) and the Under development group (adjusted $p = .0159$). The No and Under development groups did not differ (adjusted $p = .8367$).

Neither AI Opportunity nor AI Adoption Barrier Burden differed significantly across in-service training levels after Holm correction (AI Opportunity: H(2) = 4.43, adjusted p = .1637, ε^2 = .068; Barrier Burden: H(2) = 5.01, adjusted p = .1637, ε^2 = .084). Figure 1 visualises the distribution of AI Application Implementation scores across in-service training groups.

Figure 1 illustrates the distribution of AI Application Implementation scores across the three in-service AI training categories. Countries with established in-service AI training generally showed higher implementation scores, with most observations concentrated near the upper end of the scale, whereas countries reporting no training or training under development displayed lower and more dispersed scores. This visual pattern is consistent with the statistically significant difference identified in the Kruskal–Wallis analysis.

Table 3. Health-system AI characteristics across pre-service and in-service AI training levels

Indicator	N	No n; Median (IQR)	Under development n; Median (IQR)	Yes n; Median (IQR)	H(2)	Holm-adjusted p	ε²
Panel A. Pre-service AI training							
AI Application Implementation	33	n=11 28.6 (7.1-53.6)	n=13 42.9 (0.0-83.3)	n=9 100.0 (71.4-100.0)	8.19	.0501	.206
AI Opportunity	39	n=13 9.0 (7.0-11.0)	n=16 8.5 (7.0-10.0)	n=10 10.0 (9.3-11.0)	2.58	.2753	.016
AI Adoption Barrier Burden	39	n=13 14.0 (13.0-17.0)	n=16 15.0 (14.0-16.0)	n=10 12.0 (11.0-13.0)	6.57	.0750	.127
Panel B. In-service AI training							
AI Application Implementation	34	n=13 28.6 (0.0-57.1)	n=10 35.7 (0.0-67.9)	n=11 100.0 (78.6-100.0)	10.21	.0182*	.265
AI Opportunity	39	n=15 9.0 (7.5-11.0)	n=12 8.5 (7.0-10.0)	n=12 10.5 (9.8-12.0)	4.43	.1637	.068
AI Adoption Barrier Burden	39	n=15 14.0 (13.0-17.5)	n=12 15.0 (13.8-16.0)	n=12 12.0 (10.8-14.3)	5.01	.1637	.084

*Note. Values are group-specific n and medians (IQR). "Don't know" training responses were excluded. AI Application Implementation was available only for countries with at least five known application indicators. Holm adjustment was applied separately to the three pre-service and three in-service Kruskal-Wallis tests. *Holm-adjusted p < .05. Dunn-Holm post-hoc comparisons for in-service AI Application Implementation: Yes vs No, p = .0159; Yes vs Under development, p = .0159; No vs Under development, p = .8367.*

3.2.3. Sensitivity Analysis

The complete-case sensitivity analysis retained the same ordered implementation pattern. For pre-service training (n = 24), H(2) = 7.58, p = .0225, and ε^2 = .266; for in-service training (n = 26), H(2) = 8.01, p = .0182, and ε^2 = .261. These analyses support the robustness of the implementation pattern to the stricter requirement of complete responses across all seven application indicators, but they do not override the Holm-adjusted primary inference.

Overall, the clearest statistical evidence concerned in-service training: established in-service AI training was associated with broader reported implementation of AI applications. Pre-service training showed a similar descriptive ordering, but the primary Holm-adjusted result did not meet the .05 significance threshold. Perceived AI opportunities and adoption barriers did not significantly distinguish training levels after multiplicity correction.

4. Discussion

This country-level analysis across 50 WHO European Region countries found that established in-service AI training was associated with broader reported implementation of AI applications in health systems. This association remained significant after Holm correction and was supported by pairwise and sensitivity analyses. In contrast, neither perceived AI opportunities nor adoption-barrier burden significantly differentiated in-service training levels after correction.

The result should be interpreted as an association rather than a directional effect. Countries with broader AI implementation may develop continuing education because professionals increasingly encounter AI tools in practice; conversely, stronger training systems may support organisational capacity to implement AI safely. A third possibility is that both processes reflect broader health-system readiness. This interpretation is consistent with WHO guidance emphasising that responsible AI deployment depends on technological, governance, regulatory, and implementation capacity rather than on technology alone (WHO, 2024).

The in-service finding also complements literature focused primarily on individual educational preparedness. Charow et al. (2021) identified inadequate AI literacy as a barrier to preparing healthcare professionals for AI-enabled practice. At the health-system level, the present results suggest that institutionalised continuing education tends to coexist with more extensive operational AI use, although the design cannot determine which develops first.

For pre-service training, median implementation scores increased across the ‘No’, ‘Under development’, and ‘Yes’ categories, but the omnibus comparison did not remain statistically significant after Holm correction (adjusted $p = .0501$). Accordingly, this pattern should not be interpreted as evidence of a confirmed association. It may nevertheless be relevant for future longitudinal research examining whether formal curricula respond more slowly than in-service programs to changes in health-system technology. The broader education literature documents persistent gaps between students’ interest in AI and access to formal AI instruction (Busch et al., 2024; Ejaz et al., 2022).

Pre-service and in-service education may therefore operate on different institutional timescales: continuing professional development can be updated in response to operational needs, whereas formal curricula often require longer approval and implementation processes. This explanation remains interpretive and should be tested directly in longitudinal studies.

AI Opportunity did not significantly differentiate either pre-service or in-service training status. This suggests that perceiving AI as beneficial and institutionalising education are not necessarily the same stage of readiness. Similar gaps between positive attitudes and limited formal preparation have been reported among health-professions students (Mousavi Baigi et al., 2023; Bani Issa et al., 2024).

Taken together, the findings support viewing pre-service and in-service education as complementary components of workforce preparedness. Policy implications should nevertheless remain proportionate to the evidence: the data show that established in-service training is associated with broader AI implementation, not that training alone produces implementation. Future research should examine the temporal sequence of these processes and the structural conditions under which training and implementation develop together.

4.1. Limitations

The findings should be interpreted in light of several limitations.

- First, the cross-sectional design does not permit causal inference. The analyses cannot establish whether in-service education supports AI implementation, whether implementation creates demand for education, or whether both are consequences of broader system readiness.
- Second, the country is the unit of analysis. Associations observed at the health-system level should not be generalised directly to individual professionals, institutions, or specific training programs.
- Third, the analytical sample was limited to 50 countries and became smaller in individual comparisons after excluding unknown training status and, for Application

Implementation, countries with insufficient application-level information. Unequal and modest group sizes may have limited statistical power, particularly for results close to the significance threshold.

- Fourth, the composite indicators were constructed from selected AIRA items and may not capture every dimension of implementation, opportunity, barriers, or educational preparedness. Equal weighting was chosen for transparency and to avoid unsupported differential weighting, but alternative operationalisations could yield somewhat different estimates.
- Finally, country-level structural characteristics may confound the observed associations. Digital maturity, health-system resources, technological infrastructure, data readiness, governance capacity, regulatory environments, and national investment in digital health may influence both AI implementation and the development of AI training. Recent implementation literature similarly emphasises the importance of organisational, infrastructural, financial, regulatory, and workforce conditions in AI adoption (Lanfranchi et al., 2026). The limited number of countries and the comparative design constrain stable adjustment for multiple structural covariates; future longitudinal studies with harmonised contextual indicators should examine these potential confounders explicitly.

4.2. Data Availability

The original AIRA country-level data are publicly available from the WHO European Health Information Gateway (WHO Regional Office for Europe, 2025a). The cleaned, analysis-ready dataset derived from the selected AIRA indicators and used in this study is provided as supplementary material to support verification and secondary analysis. The dataset contains only aggregated country-level information and no personal identifiers.

4.3. Ethics Statement

This study used publicly available, aggregated country-level secondary data and did not involve recruitment of human participants, direct interaction with individuals, or access to identifiable personal information. Accordingly, institutional ethics committee approval and informed consent were not required for this secondary analysis. The study used the data solely for research purposes and preserved the original meaning of unknown and missing response categories.

5. Conclusion

This study examined country-level associations between health-system AI characteristics and the development status of pre-service and in-service AI training across the WHO European Region. Established in-service AI training was associated with broader reported implementation of AI applications after correction for multiple comparisons. Pre-service training showed a similar descriptive ordering, but the Holm-adjusted result did not meet the .05 threshold. Perceived AI opportunities and adoption-barrier burden did not significantly distinguish training levels. These findings indicate that continuing AI education tends to be more established in health systems reporting broader operational AI use, while the cross-sectional design does not establish direction or causality. Strengthening workforce preparedness remains relevant as AI expands in healthcare, but future longitudinal and multivariable studies are needed to determine whether education facilitates implementation, responds to implementation, or develops alongside broader health-system capacity.

References

Areco, K. N., Konstantyner, T., Bandiera-Paiva, P., Balda, R. C. X., Costa-Nobre, D. T., Sanudo, A., Kiffer, C. R. V., Kawakami, M. D., Miyoshi, M. H., Marinonio, A. S. S., Freitas, R. M. V., Morais, L. C. C., Teixeira, M. L. P., Waldvogel, B., Almeida, M. F. B., & Guinsburg, R. (2021). Operational challenges in the use of structured secondary data for health research. *Frontiers in Public Health, 9*, Article 642163. https://doi.org/10.3389/fpubh.2021.642163

Bani Issa, W., Shorbagi, A., Al-Sharman, A., Rababa, M., Al-Majeed, K., Radwan, H., Ahmed, F. R., Al-Yateem, N., Mottershead, R., Abdelrahim, D. N., Hijazi, H., Khasawneh, W., Ali, I., Abbas, N., & Fakhry, R. (2024). Shaping the future: Perspectives on the integration of artificial intelligence in health profession education: A multi-country survey. *BMC Medical Education, 24*, Article 1166. https://doi.org/10.1186/s12909-024-06076-9

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Busch, F., Hoffmann, L., Truhn, D., Ortiz-Prado, E., Makowski, M. R., Bressem, K. K., Adams, L. C., & COMFORT Consortium. (2024). Global cross-sectional student survey on AI in medical, dental, and veterinary education and practice at 192 faculties. *BMC Medical Education, 24*, Article 1066. https://doi.org/10.1186/s12909-024-06035-4

Charow, R., Jeyakumar, T., Younus, S., Dolatabadi, E., Salhia, M., Al-Mouaswas, D., Anderson, M., Balakumar, S., Clare, M., Dhalla, A., Gillan, C., Haghzare, S., Jackson, E., Lalani, N., Mattson, J., Peteanu, W., Tripp, T., Waldorf, J., Williams, S., . . . Wiljer, D. (2021). Artificial intelligence education programs for health care professionals: Scoping review. *JMIR Medical Education, 7*(4), e31043. https://doi.org/10.2196/31043

Chen, S.-Y., Feng, Z., & Yi, X. (2017). A general introduction to adjustment for multiple comparisons. *Journal of Thoracic Disease, 9*(6), 1725–1729. https://doi.org/10.21037/jtd.2017.05.34

Ejaz, H., McGrath, H., Wong, B. L. H., Guise, A., Vercauteren, T., & Shapey, J. (2022). Artificial intelligence and medical education: A global mixed-methods study of medical students' perspectives. *Digital Health, 8*, Article 20552076221089099. https://doi.org/10.1177/20552076221089099

Hunt, V. M., Sprehe, L. K., de Lomba, W. C., Gameiro, R. R., Woite, N. L., Wade, C. G., Chen, J. H., Rougas, S. C., Fraser, H. S. F., Sullivan, A. M., & Fregni, F. (2026). What the AI era doctor should know: A scoping review of proposed artificial intelligence competencies for medical education. *npj Digital Medicine*. Advance online publication. https://doi.org/10.1038/s41746-026-02761-9

Kwak, S. K., & Kim, J. H. (2017). Statistical data preparation: Management of missing values and outliers. *Korean Journal of Anesthesiology, 70*(4), 407–411. https://doi.org/10.4097/kjae.2017.70.4.407

Lanfranchi, G., Gembillo, G., Santoro, D., & Villari, M. (2026). Artificial intelligence adoption in healthcare: A systematic review of implementation challenges and health services implications. *BMC Health Services Research, 26*, Article 1187. https://doi.org/10.1186/s12913-026-14960-x

Lee, J., Wu, A. S., Li, D., & Kulasegaram, K. M. (2021). Artificial intelligence in undergraduate medical education: A scoping review. *Academic Medicine, 96*(11S), S62–S70. https://doi.org/10.1097/ACM.0000000000004291

Mousavi Baigi, S. F., Sarbaz, M., Ghaddaripouri, K., Ghaddaripouri, M., Mousavi, A. S., & Kimiafar, K. (2023). Attitudes, knowledge, and skills towards artificial intelligence among healthcare students: A systematic review. *Health Science Reports, 6*(3), Article e1138. https://doi.org/10.1002/hsr2.1138

Nahm, F. S. (2016). Nonparametric statistical tests for the continuous data: The basic concept and the practical use. *Korean Journal of Anesthesiology, 69*(1), 8–14. https://doi.org/10.4097/kjae.2016.69.1.8

Tolentino, R., Baradaran, A., Gore, G., Pluye, P., & Abbasgholizadeh-Rahimi, S. (2024). Curriculum frameworks and educational programs in AI for medical students, residents, and practicing physicians: Scoping review. *JMIR Medical Education, 10*, e54793. https://doi.org/10.2196/54793

World Health Organization Regional Office for Europe. (2025a). *Artificial intelligence for health in the WHO European Region* [Data set]. European Health Information Gateway. https://gateway.euro.who.int/en/datasets/aira/

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*.

World Health Organization. (2024). *Artificial intelligence for health: Supporting countries to deploy responsible AI technologies to accelerate equitable health for all*.

World Health Organization Regional Office for Europe. (2025b). *Artificial intelligence is reshaping health systems: State of readiness across the WHO European Region*.