**Effective application of artificial intelligence to a workflow in global health: a case study of categorization of ASTMH abstracts**

*Published at: American Society of Tropical Medicine and Hygiene annual meeting, 2026*

Toufiq Musah[1]*, Olivia Zahn[2]*, Sourabh Kulhare[2], Ishan Shah[2], Davidson H. Hamer[3,4], Kristy O. Murray[5], Judy DeAcetis[6], Charles B. Delahunt[2, 7]

* Equal contribution. [1] Kwame Nkrumah University of Science and Technology, Kumasi, Ghana; [2] Formerly Global Health Labs, Bellevue WA (ceased operations in 2025); [3] Department of Global Health, Boston University School of Public Health; [4] Section of Infectious Diseases, Department of Medicine, Boston University Chobanian and Avedisian School of Medicine; [5] Department of Pediatrics, Emory University School of Medicine and Children's Healthcare of Atlanta; [6] American Society of Tropical Medicine and Hygiene, Arlington, VA, USA; [7] University of Washington, Department of Electrical and Computer Engineering, Seattle WA.

---

**Abstract:**

Each year the ASTMH annual meeting receives over 2000 abstracts in 54 scientific categories. For peer review, the abstracts are distributed to category-specific subcommittees. But many categories overlap (e.g. malaria and mosquitoes), and if an abstract is submitted with an incorrect category and then distributed to the wrong subcommittee, it is a challenge to detect and reassign. Therefore, we built and deployed (in 2024 - 2026, retraining each year) an artificial intelligence (AI) model to automatically vet abstracts' submitted categories, flag potentially incorrect cases, and suggest alternative, more appropriate categories for the abstract peer review process. We focused on 25 priority (particularly difficult) categories.

The 2026 classifier was trained on ~6,400 abstracts from 2023–2025 along with their finalized category assignments (including corrections based on AI models in 2024 - 2025). Abstracts were converted into numerical representations using ModernPubMedBERT, a large language model (LLM) pre-trained on biomedical literature that captures semantic meaning beyond simply matching keywords. Abstracts were kept private by the LLM utility. These representations trained a Neural Network classifier that outputs ranked category recommendations with confidence scores, and generates a report comparing author-selected categories with model recommendations. This enables the Scientific Program Chair and Associate Chair (SCP Chairs) to identify and review potentially misclassified

submissions before distribution to the scientific program subcommittees. The code (not data) is organized in a version-controlled repository for future reuse.

We give results of the 2026 model for 567 abstracts submitted to 25 priority categories, for which we have Author's category, AI's top-1 choice, AI's top-2 choices, and final true category based on SP Chairs' reviews: AI top-1 disagreed with the Author in 45% of cases. Within this disagreement subset: Author matched Final - 33%; AI top-1 matched Final - 45%; AI top-2 matched Final - 74%. When AI top-1 score was high (over 90), AI top-1 matched Final – 80%, Author matched Final – 9%.

AI was a valuable triage tool, highly effective at identifying the best candidates for expert manual review. This allowed the SPC Chairs to efficiently recategorize abstracts before distribution to subcommittees, mitigating an historical headache in the ASTMH abstract review process.