Moore and Tatonetti *BioData Mining* (2025) 18:46
https://doi.org/10.1186/s13040-025-00462-9

BioData Mining

EDITORIAL

Vibe coding: a new paradigm for biomedical software development

Jason H. Moore[1*] and Nicholas Tatonetti[1*]

*Correspondence:
Jason H. Moore
Jason.Moore@csmc.edu
Nicholas Tatonetti
Nicholas.Tatonetti@cshs.org
[1]Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

The rapid growth of biomedical and clinical data has created both an opportunity and an important challenge. While there is unprecedented potential for discovery, the complexity of transforming data into usable tools often requires extensive, expensive, and time-consuming algorithm and software development. Traditionally, this work has fallen to highly skilled computer programmers and data scientists, a bottleneck that slows research and limits the translation of basic findings into clinical applications which is a core objective of building learning health systems. We introduce here the emerging concept of *vibe coding*, in which artificial intelligence (AI) accelerates the coding process by converting natural language or abstract research intent (a vibe) into functioning software modules, dramatically shortening development cycles in biomedical research and clinical application. For instance, a biomedical researcher might describe a desired data analysis pipeline in plain English, like "load a sequencing dataset, remove low-quality reads, and run differential expression analysis," and the system generates working Python or R code within seconds. With a few more steps this code could be transformed into an app for easy deployment.

Vibe coding leverages recent advances in AI program synthesis, where models such as OpenAI Windsurf, Meta's Code Llama, Anysphere Cursor, Claude Code, and Google's AlphaCode interpret user prompts to generate syntactically valid, context-aware code. Unlike earlier AI code generators that frequently hallucinated or produced brittle scripts, today's large language models (LLMs) trained on vast corpora of open-source software have demonstrated expert-level performance on many programming tasks. Crucially, these models are now increasingly *agentic* in nature: they do not merely generate code snippets on demand but can autonomously engage in cycles of programming, debugging, optimization, evaluation, and even creation and connection of multiple components such as web apps and database backends. Such agentic AI systems operate as collaborators rather than tools, persistently refining outputs, testing functionality, and adapting solutions in real time as a human coder would. Tools like GitHub Copilot, Replit Ghostwriter, and LangChain agents exemplify this paradigm, enabling researchers to "converse" with AI assistants that understand the evolving requirements of complex

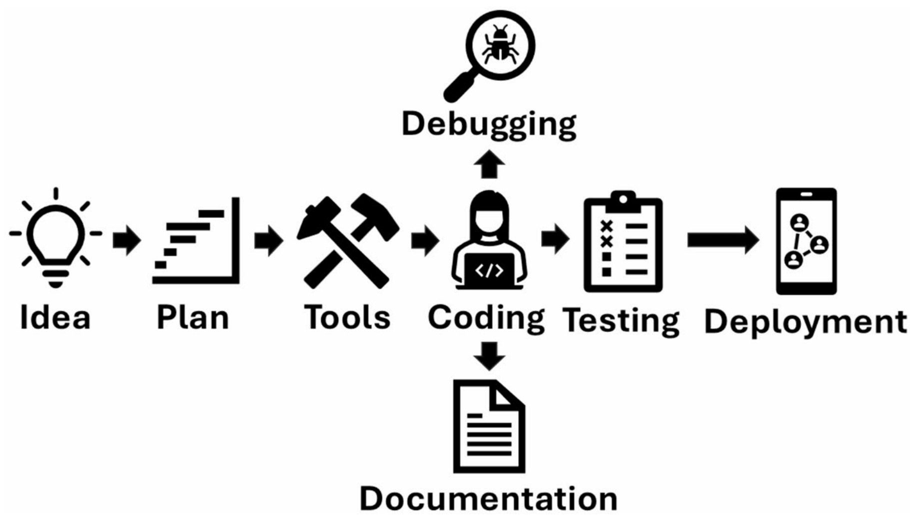

Fig. 1 Steps in the human-based software development pipeline that can all be replaced by AI agents

biomedical workflows. The AI learns the “vibe” of the task and builds meaningful, testable programs that reflect the user’s intent thus shifting the researcher’s role from coder to creative director.

For biomedical researchers and clinicians, many of whom lack formal training in computer and data science, vibe coding represents a democratizing force. With intuitive interfaces and increasingly reliable code generation, naïve users can construct data analysis pipelines, clinical decision support tools, or simulation environments without writing every line of code themselves. This significantly expands the reach of biomedical informatics to research teams and clinical programs that have historically relied on expensive technical staff or slow collaborative cycles to realize computational goals. Agentic AI further reduces barriers by handling tedious aspects like version control, parameter tuning, and unit testing, letting investigators focus on scientific questions rather than implementation details.

The implications for accelerating scientific discovery and clinical translation are transformative. Consider the time and cost involved in developing a genomic variant classifier or a radiomic biomarker extraction pipeline. These workflows traditionally require custom code, iterative debugging, and specialized programming knowledge (Fig. 1). With vibe coding, a researcher can describe their dataset and development goals in natural language, prompt the AI to generate an initial prototype, test the results, and iteratively refine the tool, all within hours instead of weeks or months. This paradigm enables rapid hypothesis testing, automated reproducibility, and on-demand tool generation during active studies or clinical deployments. As AI continues to integrate with biomedical and clinical infrastructure, vibe coding may become a foundational competency alongside pipetting, microscopy, and patient care.

Despite its promise, biomedical vibe coding must be approached with methodological rigor and domain-specific oversight. While modern LLMs have reduced hallucination rates, biomedical research operates in a high-stakes environment where correctness, reproducibility, and traceability are non-negotiable. The biomedical AI community must develop standardized validation pipelines for AI-generated code, comparable to unit testing in software engineering or QC metrics in genomics. These pipelines should include automated benchmarking against reference datasets and documentation of code provenance. Regulatory bodies and funding agencies should incentivize the development of audit-ready AI coding environments that align with HIPAA, FDA, or EMA guidelines, particularly in clinical decision support applications. Additionally, future iterations of agentic coding systems should be fine-tuned on biomedical corpora, integrate with electronic health record APIs, and support sandboxed execution environments to mitigate risk. By investing in these safeguards, we can ensure that vibe coding becomes both a faster route to software development and a trustworthy scalable method for delivering biomedical innovation to the front lines of care.

Published online: 01 July 2025

Publisher's note