Paper PP05

Redefining the role of the Statistical Programmer in an AI-Driven World

James Shipley-Burgess, Coronado Research, London, England

Chris Wood, Coronado Research, London, England

Abstract

As AI transforms the life sciences, the role of the statistical programmer is evolving. Traditional coding tasks are becoming automated, prompting the question: what's next for programmers? This poster explores how a reimagined drug development model places programmers at the centre of innovation. With rising clinical trial costs and increasing pressure to streamline processes, programmers are uniquely positioned to make a strategic impact. Beyond coding, they are becoming cross-functional collaborators by guiding AI outputs, optimising data workflows, and uncovering insights hidden in trial data. Rather than being replaced, programmers are emerging as essential drivers in a human-in-the-loop model. This shift highlights a new era where programmers are not just technical contributors, but key enablers of smarter, more efficient development.

Introduction

For decades, the pharmaceutical industry operated under the "blockbuster drug" model: a single treatment addressing a widespread condition, generating enormous commercial returns. A landmark case is Lipitor, approved in 1997, which went on to generate more than $160 billion in sales for Pfizer by treating high cholesterol, a condition affecting much of the global population [1]. The blockbuster drug model is straightforward: identify a prevalent health issue, develop a drug to treat said issue, then distribute it on a global scale.

The gradual shift towards precision medicine has steadily disrupted this paradigm [2]. What began with early genomics and biomarker-guided therapies has grown into a model where interventions are increasingly tailored to genetic profiles, phenotypes, or other individual characteristics. Yet clinical development processes have largely evolved only incrementally, inherited from the blockbuster era. As a result, we face a growing mismatch: treatments are becoming more precise, but the systems supporting them remain blunt. Addressing this gap requires not another small adjustment, but a step-change in how we design and deliver clinical trials.

Precision medicine creates unique economic and operational pressures. Drug development costs remain high, while the populations eligible for treatment are smaller and more fragmented [3]. New technologies, AI/ML-driven analytics, data automation platforms and smart trial designs, are gaining traction. In particular, AI is being positioned as the industries "silver bullet", with efficiencies being promised across the entire drug development lifecycle. Some tech companies claim to have cracked end-to-end automation, bridging raw data and statistical analysis plans, completing the programming activities in between. What once took programmers weeks to program will soon be completed by AI within a matter of hours.

Understandably, this narrative has generated concern among statistical programmers, who fear obsolescence as routine tasks become automated. Yet such fears echo earlier societal transitions.

Parallels with the Industrial Revolution

The hysteria that surrounds todays AI revolution are reminiscent of the industrial revolution, where mechanisation spurred widespread fear of job displacement. Prior to industrialisation, approximately 80-90% of the global population was engaged in agriculture [4]. The combustion engine not only reduced the need for agricultural labour but also enabled new industries, ultimately expanding human creativity and innovation. History demonstrates that humans were not replaced by machines; rather, they were freed from arduous manual labour, unlocking entirely new domains of opportunity.

Implications for Statistical Programmers

In the same vein, AI stands to liberate programmers from repetitive, standardised work. For example, the automation of regulatory submission datasets and outputs reflects the kind of task most vulnerable to structured automation. However, higher-order tasks which require contextual understanding, cross-functional insight and domain knowledge will remain firmly human-driven. At Coronado Research, statistical programming teams have already shifted to integrative roles. By

engaging across functions and applying domain expertise across the development lifecycle, programmers are positioned not as redundant coders, but as critical connectors of data, insight and decision-making.

What Does the Future Proof Programmer Look Like?

If AI is reshaping the way we work, the role of the statistical programmer will evolve alongside it. But what does that role look like in practice? The model below illustrates what the future of statistical programmer may look like if the opportunities that present themselves today are seized.

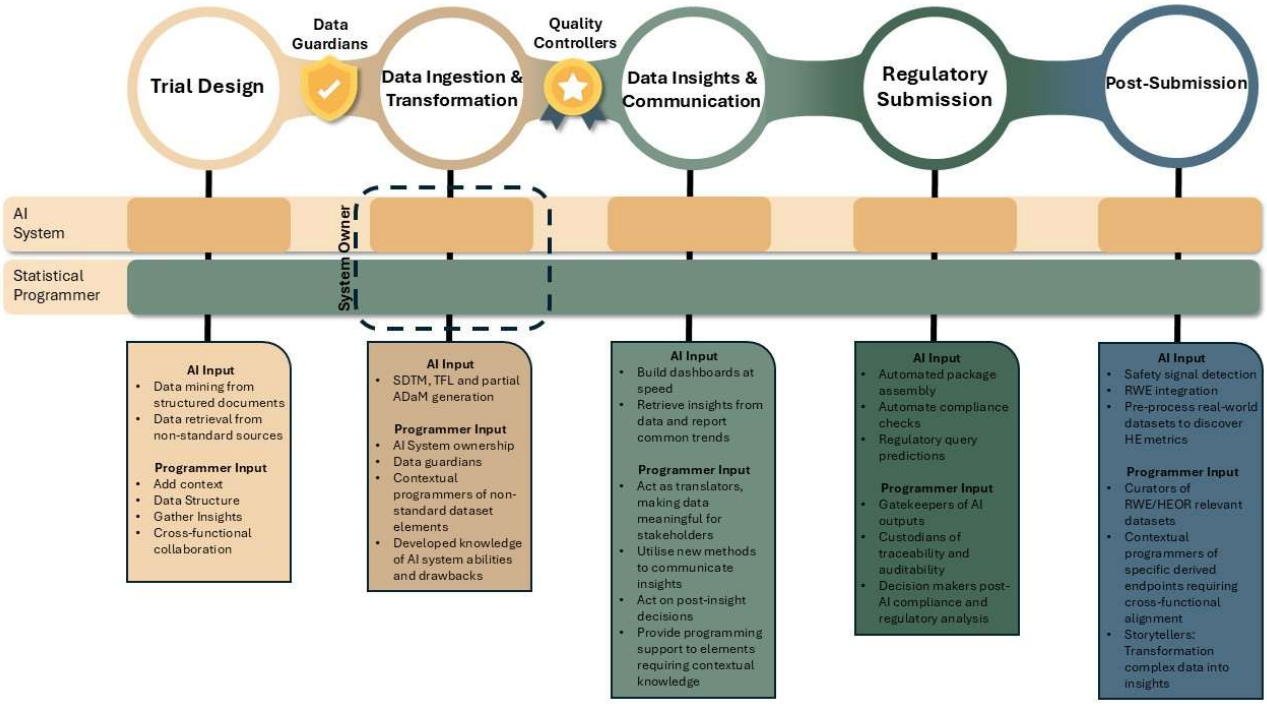

Figure 1. Future drug development lifecycle model. AI automation accelerates core data transformations (yellow lane), but statistical programmers remain the linchpins (green lane), acting as system owners, guardians, and translators across the lifecycle. Programmers work with Regulatory Affairs and Market Access from study design through submission and commercialisation, ensuring aligned endpoints, trustworthy transformations, and structured evidence for both regulators and payers.

In this future model, AI becomes the engine that automates core transformations, raw trial data into SDTM, ADaM, and standard outputs. But statistical programmers remain the linchpins, positioned along every stage of the lifecycle to ensure structure, compliance, and meaning. Before trials begin, they work with Regulatory Affairs and Market Access to structure mined evidence, aligning endpoints to both payer and regulator expectations. During data transformation, they act as guardians of the AI platform, programming complex derivations when context is required, validating AI outputs, and safeguarding standards. Beyond submission, they translate outputs for statisticians, clinicians, and regulators, and they extend their role into market access by contextualising real-world data for HEOR and reimbursement. This positions programmers not merely as coders, but as system owners and connectors; humans-in-the-loop who guarantee that AI driven development is not only faster, but also trustworthy, compliant, and strategically aligned.

This model illustrates what is possible, there are other ways this model may develop if action is not taken now to ensure statistical programmers have a seat at the table. Both data managers and clinical programmers have skills relating to data ingestion and transformation, placing them as effective system owners of future AI systems transforming raw data into submission ready datasets and outputs. Our future is not guaranteed, and it requires the statistical programming community to advocate for our ownership of these systems, to champion our knowledge, abilities and relevance as AI systems become more widespread. To remain futureproof, statistical programmers must begin developing skills and perspectives that extend far beyond writing datasets or tables. The following areas signal where focus and opportunities lie for programmers today.

1. Domain Expertise as a Differentiator
The programmers who thrive will be those who understand not only the code, but also the science behind the trial and the regulatory context in which data is used. This means:

- Deepening knowledge of clinical trial design, endpoints, and therapeutic areas.
- Understanding regulatory standards (such as CDISC and ICH guidance) not just as formats to implement, but as frameworks that shape evidence.
  With AI taking care of standardised outputs, programmers will increasingly be the interpreters and explainers of data, ensuring outputs are both valid and meaningful.

2. Becoming the "Critical Reviewer" of AI
AI may be able to generate code, harmonise datasets, or suggest analyses at speed, but it cannot guarantee accuracy, appropriateness, or context. Future-proof programmers will need to:

- Apply critical, logical thinking to review AI generated outputs.
- Identify nuances where automated logic may fail and provide necessary correction.
- Act as the safeguard for quality, transparency, and reproducibility.
  This oversight role is where the human-in-the-loop adds irreplaceable value.

3. Data Storytelling and Cross Functional Collaboration
As drug development continues to demand faster decisions, programmers will increasingly act as "data translators". This involves:

- Presenting data clearly to clinicians, statisticians, regulators, and other functional teams.
- Collaborating with statisticians to communicate not just what the numbers show, but what they *mean* in the wider context of trial objectives and patient outcomes.
- Partnering across regulatory, clinical, and commercial functions to ensure evidence is aligned from trial through to market access.

4. Embracing New Tools and Technologies
Curiosity may have killed the cat, but it will save the statistical programmer. A future-proof programmer is unafraid of change and curious about new methods. Opportunities in most organisations already exist to strengthen skills in:

- Cloud computing and modern data pipelines.
- Data visualisation and interactive reporting tools.
- AI and machine learning, not necessarily as pure developers, but as informed users who can apply, evaluate, and challenge outputs.
- Statistical programmers have the opportunity to become AI system owners, becoming experts in the AI data transformation platform, just as they are today experts in SAS. This doesn't mean they know how to "build" the system, instead they will have a deep understanding of its internal mechanisms, abilities, drawbacks and ultimately be the guardians of the systems input, ensuring high quality output.

5. Championing a Critical and Logical Mindset
Perhaps most importantly, organisations of the future will recruit not just for programming ability, but for individuals who can think critically, solve problems logically, and adapt to new ways of working. Programmers should seek out opportunities within their organisations that place them at decision making tables, in study design discussions, data strategy workshops, or cross functional analytics initiatives. These are the spaces where their influence will grow.

Challenges to Consider

One of the key challenges statistical programming faces in the age of AI is ensuring that the next generation of programmers acquires the necessary domain knowledge. For many of us, this expertise was developed "on the job," by

tackling foundational programming tasks and learning through doing. Yet as automation and AI increasingly handle routine activities, those traditional learning opportunities are shrinking.

Take SDTM programming as an example. The growing use of automation and AI driven code generation means that the role of the statistical programmer will increasingly focus on critical review, verifying the logic of the code, ensuring accurate CDISC mapping, and safeguarding overall data integrity. In this future, the human contribution is less about writing every line of code, and more about applying domain knowledge and critical thinking to validate what the machine produces.

This raises an important question: how do we train the next generation to apply the same rigour, scepticism, and contextual understanding when they may no longer acquire this experience through foundational tasks? As the programming landscape evolves, entry level roles will demand a deeper initial knowledge of both programming and the broader clinical trial context. The tasks that once served as stepping stones for new programmers will increasingly be managed by AI. If we do not address this shift, we risk creating a skills gap that undermines the very expertise the industry needs most.

Conclusion

The rise of AI in drug development may initially appear to threaten the role of statistical programmers, but the reality is quite different. Throughout this paper we have shown how programmers' value is not diminished but redefined: as AI takes on repetitive, standardised tasks, programmers are freed to contribute where their domain knowledge, critical thinking, and contextual understanding are indispensable.

In this new model, statistical programmers move from the periphery to the centre of clinical development, acting as linchpins who connect data, insight, and decision-making across functions. By collaborating with AI rather than competing against it, they safeguard quality, bring coherence to complex datasets, and unlock opportunities for innovation.

At Coronado Research, we are already demonstrating what this future looks like. By embedding programmers across the development lifecycle and combining their expertise with AI-driven tools, we are not speculating about what might be, but showing what is possible today. This approach enables us to help shorten timelines, reduce costs, and accelerate the delivery of precision medicines to patients, without sacrificing scientific rigour.

The message is clear: statistical programmers are not becoming obsolete. They are becoming more essential than ever. By championing their skillset, embracing collaboration with AI, and positioning themselves as strategic partners in drug development, programmers can help unlock the full potential of precision medicine. But this future will not build itself. Together, we must champion our skillset, preserve the hard-won lessons of today, and ensure the next generation has the foundation to thrive in the world of AI. In doing so, we not only secure our place in the age of AI, but we also strengthen the foundation of the profession for generations to come.

References

[1] King S. The best-selling drugs of all time: Humira joins the elite [Internet]. *Forbes*. 2013 Jan 28 [cited 2025 Sep 22].

[2] DrugPatentWatch. As blockbuster drugs fizzle, biotech looks warily to the next big thing [Internet]. DrugPatentWatch Blog. 2020 Jun 22 [cited 2025 Sep 22]

[3] Moretti P. Economic Implications of Personalized Medicine: A Review of Cost-Effectiveness Studies. *J Health Econ Outcome Res*. 2024

[4] Encyclopedia.com. Industrial Revolution [Internet]. Encyclopedia.com; [cited 2025 Sep 22].