The American Journal of Engineering and Technology
ISSN 2689-0984 Volume 07 - 2025

AI-Powered Regulatory Surveillance for Mitigating Pharmaceutical Manufacturer Product Hopping under the IRA

[1]Pinaki Bose
[1]Independent Researcher, USA

Received: 11th Sep 2025 | Received Revised Version: 19th Oct 2025 | Accepted: 23th Nov 2025 | Published: 30th NOV 2025

Volume 07 Issue 11 2025 | Crossref DOI: 10.37547/tajet/v7i11-308

Abstract

The Inflation Reduction Act (IRA) of 2022 aims to curb soaring pharmaceutical costs by mandating price negotiation for Qualifying Single Source Drugs (QSSDs) that have been on the market for 7 (small molecules) or 11 (biologics) years and lack generic competition. However, this time-based metric introduces a critical systemic vulnerability: Product Hopping (PH). PH is an established anticompetitive tactic wherein manufacturers introduce minor, non-therapeutic reformulations (a New Drug Application (NDA) or Biologics License Application (BLA)) solely to reset the negotiation clock, thereby extending their monopoly. Current reactive regulatory frameworks are insufficient to counteract these complex, data-driven manipulative strategies. This theoretical paper proposes an expert-level conceptual framework for an Artificial Intelligence (AI)-powered Regulatory Surveillance Architecture (RSA) within the Centers for Medicare & Medicaid Services (CMS). This RSA leverages predictive analytics, Natural Language Processing (NLP), and anomaly detection across multi-modal data streams—including patent filings, clinical trial documents, and market data—to quantify the economic and therapeutic rationale underlying reformulation, yielding a Probabilistic Intent Score (PIS). Central to the framework is the mandatory implementation of Explainable AI (XAI) to ensure that regulatory interventions, particularly those triggering high-stakes negotiation, are transparent, auditable, and legally defensible, meeting rigorous standards of administrative due process and governance.

Keywords: AI Governance, Product Hopping, Inflation Reduction Act (IRA), CMS, Predictive Analytics, Explainable AI (XAI), Regulatory Informatics, Pharmaceutical Manipulation.

Cite This Article: Bose, P. (2025). AI-powered regulatory surveillance for mitigating pharmaceutical manufacturer product hopping under the IRA. The American Journal of Engineering and Technology, 7(11), 188–195. https://doi.org/10.37547/tajet/v7i11-308.

1. Introduction

1.1 The Integrity Challenge to Medicare Drug Price Negotiation Context and Mandate of the IRA

The IRA represents a landmark legislative effort to reduce the high cost of prescription drugs in the United States by empowering CMS to negotiate the Maximum Fair Price (MFP) for a select number of high-expenditure drugs. The eligibility for negotiation is strictly defined around the status of a Qualifying Single Source Drug (QSSD), which must meet several criteria: high overall Medicare spending, status as a brand drug or biologic without generic or biosimilar competition, and the fulfillment of a time-on-market requirement—at least 7 years post-FDA approval for small-molecule drugs or 11 years post-licensure for

biologics [1]. The primary goal of this intervention is to curb the costs driven by high-priced, patent-protected drugs after a reasonable period of market exclusivity has elapsed.

1.2 Defining Product Hopping as a Systemic Vulnerability

Product Hopping (PH), or formulation manipulation, is an established anticompetitive strategy used by brand-name manufacturers to circumvent the entry of cheaper generic alternatives. The mechanism involves introducing a marginally modified version of a drug—often changing only the dosage frequency (e.g., from two-a-day to one-a-day) or delivery system, with little to no demonstrable clinical improvement [2]. This minor change is sufficient to secure a new patent and, critically, receive a new NDA or BLA, which is leveraged to reset the statutory exclusivity period [3].

The complexity of PH arises because it exploits the boundaries between disparate regulatory structures: patent law, antitrust law, the Hatch-Waxman Act, and state drug product substitution laws. CMS has explicitly acknowledged this threat, recognizing that manufacturer modifications that are "modest or minor" are used to "effectively reset the 7- or 11-year periods" required for QSSD eligibility. Since manufacturers have historically proven adept at navigating these fragmented regulatory structures to maintain monopolies, the introduction of a new, time-based IRA clock creates a new, specific metric for strategic exploitation. A failure to counter these strategies will result in billions of dollars in forgone Medicare savings, signaling a validation of regulatory gaming over genuine pharmaceutical innovation.

1.3 Thesis and Scope

Current regulatory surveillance relies heavily on retrospective auditing and manual analysis, which is too slow and fragmented to identify nascent manipulation strategies. This paper asserts that to secure the fiscal integrity of the IRA, CMS must implement a proactive, AI-powered capability designed specifically to model manufacturers with an intent to manipulate the negotiation cycle. This necessitates a shift in regulatory paradigm, utilizing advanced data science to integrate and monitor signals across the traditionally siloed domains of intellectual property, clinical efficacy, and market dynamics. The proposed conceptual framework outlines the architecture and governance required for this expert-level regulatory informatics capability.

2. The Product Hopping Mechanism and IRA Avoidance Calculus

2.1 The Antitrust Analogy: Quantifying Exclusionary Intent

In antitrust litigation, the legality of product hopping often hinges on the "no-economic-sense test". The challenge for CMS is translating this subjective legal standard of "intent" into objective, quantifiable metrics suitable for automated regulatory intervention. The AI system must provide empirical evidence proving that the primary purpose of the reformulation was strategic delay, rather than genuine innovation. This determination is critical because IRA negotiation constitutes a high-stakes, adjudicative regulatory decision that requires a strong data-driven foundation to withstand inevitable legal challenges [7].

2.2 Time-to-Market Exploitation

The IRA’s time requirement provides a precise window for exploitation. Manufacturers typically initiate the development and marketing shift approximately 2 to 3 years before the QSSD clock expires (Years 5-6 for small molecules, Years 9-10 for biologics). This strategic timing ensures that the new product achieves critical market saturation and prescriber adoption before the original drug becomes eligible for generic substitution. This maneuver dramatically delays the expected price reduction for Medicare beneficiaries, undermining the entire fiscal mechanism of the IRA.

The full cycle of this manipulation, which the AI system must predict and prevent, can be visualized as a systematic regulatory bypass:

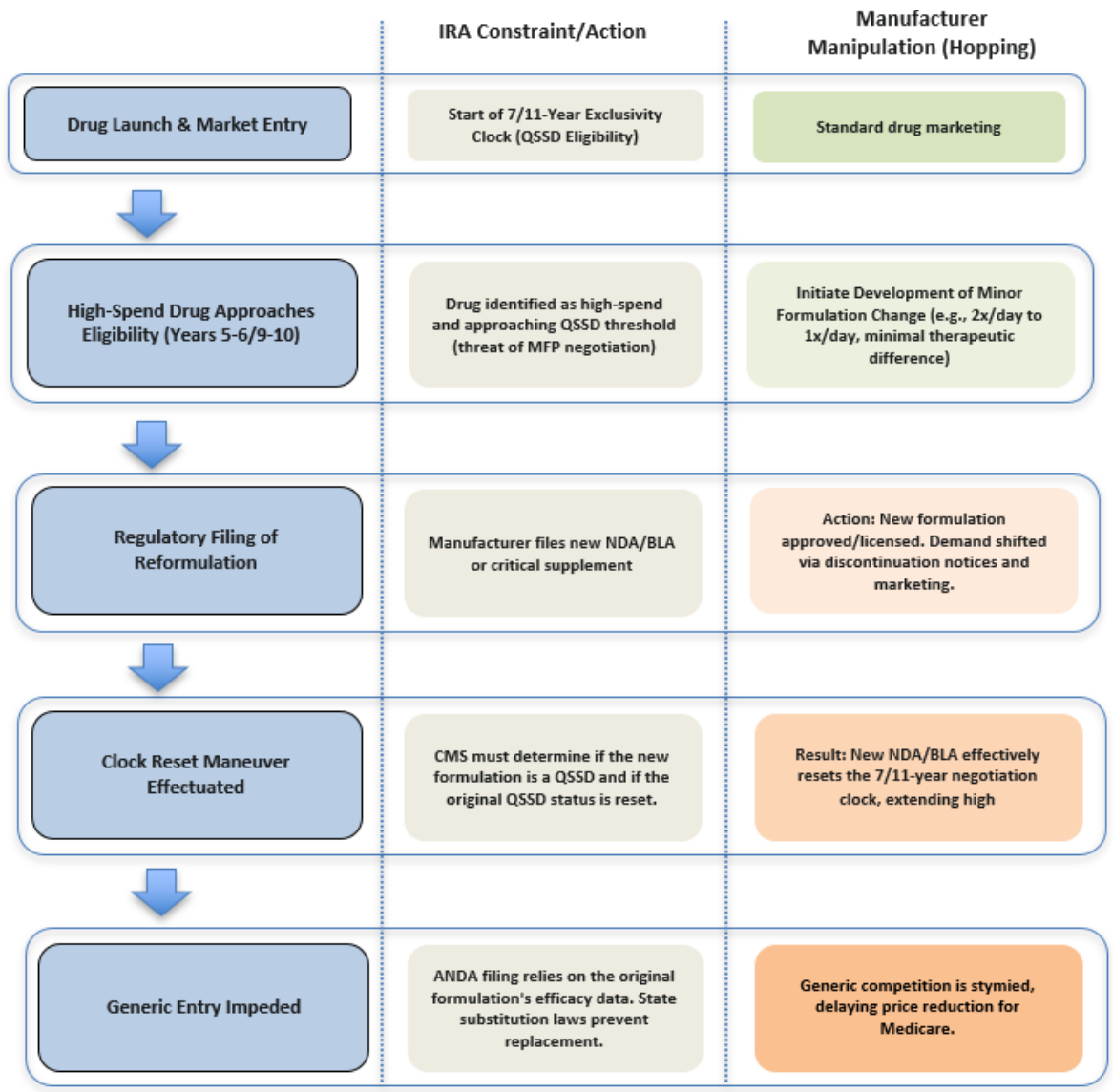

Fig 1.: The Product Hopping Loophole and Exclusivity Clock Reset

3. AI-Driven Predictive Intelligence: Modeling Manufacturer Intent

3.1 Paradigm Shift: From Retrospective Audit to Algorithmic Foresight

The complexity and fragmentation of the data required to prove anticompetitive intent—spanning proprietary cost data, unstructured regulatory filings, and dynamic market metrics—necessitates an integrated AI solution. The AI framework enables CMS to transition from reactive analysis ("what happened?") to predictive intelligence ("what is likely to happen, and why?"). The core output of this transition is the generation of a **Probabilistic Intent Score (PIS)**, which forecasts the likelihood and mechanism of a manufacturer successfully executing a product hop before the QSSD eligibility period expires.

3.2 Multi-Modal Data Acquisition and Integration

The CMS RSA requires harmonizing heterogeneous data streams, which are often siloed across different regulatory and commercial entities:

- **Regulatory and Clinical Data:** Raw text from FDA NDA/BLA submission summaries, clinical trial efficacy data, adverse event reports, and structured regulatory approval dates [8].

- **Intellectual Property (IP) Data:** Patent thicket architecture, specific claims for new patents (often secondary features like manufacturing methods [10]), Orange Book listings, and records of Hatch-Waxman litigation [6].
- **Market and Pricing Data:** Detailed Medicare Part D and Part B gross covered costs, Non-Federal Average Manufacturer Price (NFAMP) submissions, and commercial sales volume data used to calculate predicted price erosion curves [8].

3.3 Application of Natural Language Processing (NLP) for Therapeutic Assessment

A crucial element in determining manipulative intent is objectively quantifying the clinical benefit of the reformulated drug. NLP, specifically transformer-based models tailored for clinical text, should be employed to analyze the unstructured clinical trial summaries of the new formulation against the original drug's profile. This analysis generates a **Therapeutic Equivalence Score (TES)**.

A TES approaching zero indicates negligible therapeutic novelty, suggesting a non-substantive change. When a manufacturer submits detailed R&D cost data (as required by the IRA) claiming significant expense for a reformulation that the AI system scores with a low TES, **this differential provides strong empirical evidence supporting the determination of a strategic maneuver rather than genuine innovation**. The utilization of an Agentic AI framework, akin to systems used in drug discovery, allows the CMS system to dynamically execute complex, multi-step regulatory intelligence workflows—for instance, querying patent filings, extracting corresponding clinical text, and synthesizing a comprehensive TES—autonomously.

3.4 Predictive Modeling and Anomaly Detection

To forecast and flag potential manipulation, the RIA employs several advanced machine learning techniques:

1. **Anomaly Detection:** Techniques such as Isolation Forest and Local Outlier Factor (LOF) are utilized to identify statistically abnormal events preceding the QSSD deadline. These anomalies include unusual spikes in patent filings for secondary features (patent evergreening) [5], unexpected changes in Medicare utilization patterns, or sudden inventory shifts and discontinuation notices for the original product [9].
2. **Regression and Time-Series Forecasting:** Predictive models analyze market elasticity and forecast the expected rate of generic price erosion had the product hop not occurred. If the introduction of the new formulation causes a statistically significant impairment of the *predicted* generic market success, the resulting financial loss provides a quantifiable measure of the exclusionary impact.
3. **Economic Feasibility Index (EFI):** The models synthesize the TES, Patent Thicket Density, and R&D cost data to calculate an EFI. A strongly negative EFI—where the new formulation is economically unjustifiable unless generic competition is explicitly excluded—serves as objective proof aligning with the antitrust standard of "no economic sense."

The required fusion of these signals into a coherent, actionable regulatory prediction is synthesized in Table I.

Table I: Critical Data Signals and AI Techniques for Proactive Product Hopping Detection

Data Signal Cluster	Data Source	AI Technique	Primary Analytical Output	Regulatory Flag/Insight
Regulatory & Clinical Efficacy	FDA NDA/BLA Efficacy Summaries, Phase III Results, Structured Data Points	Clinical NLP, Feature Engineering	Therapeutic Equivalence Score (TES) vs. Original; Justification extraction	Low TES + New Patent/New NDA implies formulation manipulation.
Intellectual Property Strategy	Orange Book Listings, Patent Filing Velocity, Hatch-Waxman Litigation Dockets	Graph Neural Networks (GNN), Anomaly Detection	Patent Thicket Density; Non-Substantive Patent Index; Litigation Delay Metrics	Evidence of "Evergreening" and intentional delay exceeding R&D value.
Market & Behavioral Coercion	Medicare Part D/B Spending, Inventory/Discontinuation Notices, Marketing Scrapes	Change Point Detection, Sentiment Analysis, Time-Series Forecasting	Sudden drop in original drug claims/inventory; Demand shift velocity; Deviation from forecasted generic erosion	Early warning of coercive market withdrawal aimed at destroying the generic base.
Manufacturer Economics	NFAMP Submissions, R&D Cost Submissions (IRA Mandate)	Regression Analysis, Cost-Benefit Modeling	Economic Feasibility Index (EFI)	EFI is significantly negative *unless* generic competition is excluded = strong PH intent.

4. Conceptualizing the CMS Regulatory Surveillance Architecture

4.1 Alignment with CMS Technical Reference Architecture (TRA)

Any advanced technological deployment within CMS must strictly comply with the agency's Technical Reference Architecture (TRA). The TRA establishes authoritative technical standards for security, infrastructure, data management, and interoperability. Given the necessity of handling proprietary manufacturer data and Medicare claims data, the RSA must incorporate robust security controls compliant with federal standards (e.g., FISMA) and manage data integrity throughout the model's lifecycle [11]. Compliance with the TRA ensures a standardized operating environment and facilitates the reuse of shared infrastructure across CMS systems.

4.2 The Data Lakehouse as the Foundational Platform

To accommodate the diverse, multi-modal data required for intent modeling, Lakehouse Data architecture is proposed as the foundational platform. This structure is necessary because traditional data warehouses struggle with the high volume of unstructured text data (patent filings, clinical notes) required for NLP feature extraction. The data Lakehouse combines the flexibility of storing raw, unstructured data (the data lake) with the transactional performance and governance capabilities typical of a data warehouse [12]. This architecture supports the continuous ingestion and harmonized preparation of data necessary for iterative model training and real-time regulatory surveillance.

4.3 Components of the Regulatory Intelligence Agent (RIA)

The RIA operates as the central processing engine within the RSA:

1. **Data Ingestion Layer:** High-throughput pipelines provide real-time updates from federal data partners (e.g., FDA regulatory submission registries) and commercial data providers (e.g., pricing compendia). This layer must enforce standardized data exchange protocols to ensure interoperability.
2. **Feature Engineering and Model Training Layer:** This layer executes the complex calculations (TES, Patent Thicket Density) and trains the predictive models (PIS). Given the high propensity for manufacturers to adapt their tactics, the models are highly susceptible to "model drift". Consequently, this layer must mandate a continuous life-cycle approach to model maintenance, requiring frequent retraining and rigorous validation against new Product Hopping case studies to ensure that the AI remains effective and compliant with regulatory expectations.
3. **Inference Engine (PIS Generator):** This core machine learning engine synthesizes the results from all specialized models (NLP, GNN, Anomaly Detection) to produce the final, weighted Probabilistic Intent Score (PIS).
4. **XAI Justification Module:** This mandatory component, detailed further below, translates the PIS and its driving factors into a human-readable explanation package.

For the system to acquire the necessary comprehensive view, policy measures must mandate clear data submission standards and ensure standardized, real-time data exchange agreements with external regulatory entities, particularly the FDA and commercial pricing compendia.

5. Governance and Explainable AI for Due Process in Regulatory Intervention

5.1 The Imperative of Explainable AI (XAI)

In public sector decision-making, particularly where high-stakes financial penalties or market interventions are involved, reliance on opaque "black box" algorithms is unacceptable. Explainable AI (XAI) is essential to satisfy the legal and ethical demands for transparency, due process, and accountability [4]. If CMS bases its decision to trigger negotiation—effectively challenging a manufacturer's claim of innovation—on an unintelligible score, the decision will invariably face litigation challenging the algorithmic transparency and methodology [7]. XAI ensures that the reasoning behind the PIS is auditable, building trust among clinicians, regulated entities, and the public.

The utilization of XAI shifts the legal focus away from challenging the AI's internal, complex logic, and redirects the debate toward challenging the underlying data inputs or the policy weights assigned to the model outputs. This substantially strengthens the legal defensibility of CMS actions and places the burden of proof back on manufacturers to justify their data submissions and strategic intent.

5.2 XAI Methods for Regulatory Justification

The XAI module must generate clear, actionable rationales, which can be accomplished through:

1. **Rationale Generation:** Generating narrative justifications that explicitly link the final PIS to specific inputs, model weighting, and causal factors. For example, the explanation might state: "Intervention justified because the Probabilistic Intent Score (PIS) exceeded 0.90, primarily driven by a Therapeutic Equivalence Score (TES) of 0.03 (indicating marginal therapeutic novelty) and an Anomaly Detection flag showing a 4x increase in secondary patent claims filed in the preceding 24 months, suggesting strategic patent thicketing".
2. **Interpretable Optimization:** Utilizing methods such as surrogate modeling or near-optimal solution analysis, the XAI can simplify complex predictive results, allowing CMS experts to understand the trade-offs and alternative decision pathways the model considered. This capability is crucial for justifying why a regulatory action was taken against one drug but not a seemingly similar one, ensuring consistent and equitable enforcement.

5.3 Establishing an AI Governance Framework within CMS

Effective deployment of the RSA requires an explicit governance structure applied across People, Process, Technology, and Operations (PPTO).

- **Oversight and Process:** CMS must establish a centralized AI Governance Committee (AGC) with ultimate decision-making authority over model deployment and interpretation. This committee oversees continuous validation, performance monitoring, and adaptation to new manufacturer strategies.
- **Algorithmic Fairness:** The framework must proactively address potential implicit algorithmic biases that could unjustly target specific therapeutic classes or disproportionately impact access for vulnerable populations, thereby compromising health equity. Rigorous bias assessment must be integrated into the continuous validation cycle.

5.4 Human-in-the-Loop Adjudication

The AI system is designed strictly as a decision *support* tool, not a decision *maker*. The XAI output—the comprehensive evidence package—serves as the primary documentation presented to CMS economists, legal experts, and regulatory professionals. These human experts retain the final authority to vet the algorithmic findings against established antitrust and IRA legal standards before any formal regulatory action (e.g., triggering negotiation) is implemented. This closed-loop process ensures that the system is both efficient and ethically accountable.

6. Conclusion

6.1 Securing the Fiscal Integrity of the IRA Through Regulatory Informatics

The Inflation Reduction Act faces a critical, systemic challenge from sophisticated manufacturer strategies like Product Hopping, which seek to exploit the time-on-market requirements to maintain high prices. The inherent difficulty lies in accurately distinguishing genuine, incremental innovation from anticompetitive manipulation designed solely to reset the negotiation clock.

The conceptual framework presented here—an AI-powered Regulatory Surveillance Architecture anchored by the Probabilistic Intent Score and Therapeutic Equivalence Score—provides CMS with the proactive, data-driven methodology required to counter formulation manipulation effectively. By leveraging advanced NLP to quantify therapeutic novelty and anomaly detection to flag strategic behavior in the patent landscape, the system can provide the necessary empirical evidence to quantify manufacturer intent.

To realize the promised savings of the IRA and establish a robust regulatory regime, CMS must prioritize strategic investments in three critical areas:

1. **Data Infrastructure:** Development of a secure, multi-modal data Lakehouse capable of integrating high-volume, disparate data streams (clinical, IP, market).
2. **Algorithmic Development:** Resource allocation toward building the Regulatory Intelligence Agent (RIA) and its specialized predictive models (PIS generator), ensuring a continuous life-cycle approach to mitigate model drift.
3. **Governance and Legal Defensibility:** Mandatory integration of Explainable AI (XAI) across all decision-making layers. This governance approach ensures transparency and auditability, securing the CMS's legal position and maximizing the efficacy of the Medicare drug negotiation program.

References

1. FAQs about the Inflation Reduction Act's Medicare Drug Price Negotiation Program | KFF, https://www.kff.org/medicare/faqs-about-the-inflation-reduction-acts-medicare-drug-price-negotiation-program/
2. Medicare Drug Price Negotiation Program: Final Guidance, Implementation of Sections 1191 - CMS, https://www.cms.gov/files/document/ipay-2028-final-guidance.pdf
3. A Legal Test for the Pharmaceutical Company Practice of "Product Hopping" - AMA Journal of Ethics, https://journalofethics.ama-assn.org/article/legal-test-pharmaceutical-company-practice-product-hopping/2015-08
4. Explainable AI in Healthcare Risk Prediction - Censinet, https://www.censinet.com/perspectives/explainable-ai-in-healthcare-risk-prediction
5. PRODUCT HOPPING: A NEW FRAMEWORK - Michael A. Carrier* & Steve D. Shadowen - Congress.gov, https://www.congress.gov/116/meeting/house/109970/documents/HHRG-116-IF17-20190919-SD005.pdf
6. 1 Michael A. Carrier Distinguished Professor Rutgers Law School Having Our Cake and Eating it Too: Five Things Congress Can Do,, https://www.congress.gov/117/meeting/house/112

518/witnesses/HHRG-117-JU05-Wstate-CarrierM-20210429.pdf
7. Medicare Drug Price Negotiation Program: Revised Guidance, Implementation of Sections 1191 – 1198 of the Social Security Act f - CMS, https://www.cms.gov/files/document/revised-medicare-drug-price-negotiation-program-guidance-june-2023.pdf
8. Natural Language Processing Applied to Clinical Documentation in Post-acute Care Settings: A Scoping Review - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC10792659/
9. The Role of Patents and Regulatory Exclusivities in Drug Pricing | Congress.gov, https://www.congress.gov/crs-product/R46679
10. COST OF GENERIC DRUG DEVELOPMENT AND APPROVAL FINAL - https: // aspe . hhs . gov., https://aspe.hhs.gov/sites/default/files/documents/20e14b66420440b9e726c61d281cc5a5/cost-of-generic-drugs-erg.pdf
11. Interpretable Optimization: Why and How We Should Explain Optimization Models - MDPI, https://www.mdpi.com/2076-3417/15/10/5732
12. The application of artificial intelligence in health financing: a scoping review - PMC, https://pmc.ncbi.nlm.nih.gov/articles/PMC10626800/