Review

Tokenizing DNA-Encoded Chemical Libraries with Non-Fungible Tokens (NFTs): A Scalable Framework for Registration, Provenance, and Transfer of Ultra-Large Small-Molecule Asset Collections

Stanley Cho , Bomi Woo , Sung-Ung Kang *

Posted Date: 27 February 2026

doi: 10.20944/preprints202602.1981.v1

Keywords: DNA-encoded chemical library; DEL; DECL; non-fungible token; NFT; blockchain; intellectual property; provenance; technology transfer

 doi:10.20944/preprints202602.1981.v1

Review

Tokenizing DNA-Encoded Chemical Libraries with Non-Fungible Tokens (NFTs): A Scalable Framework for Registration, Provenance, and Transfer of Ultra-Large Small-Molecule Asset Collections

Stanley Cho [1], Bomi Woo [2] and Sung-Ung Kang [2,3,*]

[1] Department of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853, USA
[2] Neuroregeneration and Stem Cell Programs, Institute for Cell Engineering, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
[3] Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
* Correspondence: skang34@jhmi.edu

Abstract

DNA-encoded chemical libraries (DECLs/DELs) enable the pooled synthesis and selection of millions to billions of DNA-barcoded small molecules, providing an efficient route to discover binders and early leads against diverse biological targets. As DEL-derived programs advance toward identifying clinical candidates, the asset surface of a DEL platform expands from a small set of optimized hits to include library designs, building-block combinations, DNA tags, selection data, and physical library stocks, thus creating new challenges in registration, traceability, and scalable ownership in transfer practices. Non-fungible tokens (NFTs) are unique blockchain-native tokens that can represent digital assets that can be coupled to smart contracts to enable traceable transactions and programmable rights management, which inspire proposals to tokenize intellectual-property (IP) assets such as patents. Here, we review (i) the scientific and commercial value of DEL in modern drug discovery, (ii) NFT/blockchain concepts, specifically in reported biomedical-IP and supply-chain use cases, and (iii) a conceptual architecture for NFT-enabled registration and controlled transfer of DEL libraries or sublibraries using on-chain identifiers with off-chain encrypted metadata and legal agreements.

Keywords: DNA-encoded chemical library; DEL; DECL; non-fungible token; NFT; blockchain; intellectual property; provenance; technology transfer

1. Introduction

DNA-encoded chemical libraries (DELs) couple small molecules to DNA barcodes, enabling pooled selection against targets and sequencing-based identification of enriched ligands at a scale that far exceeds conventional plate-based screening logistics. Because DEL campaigns can interrogate enormous chemical spaces quickly and with low material consumption per member, DEL has become a mainstream modality for ligand discovery and hit generation across diverse target classes. However, DEL's scale also creates a rights-management bottleneck, such as infeasibility to patent "everything that could be made" in a DEL. Still, it is equally risky to leave valuable library-derived chemical matter and selection datasets without robust evidence of provenance, timing, and ownership/control.

Blockchain and NFT technology have been proposed as mechanisms for immutable timestamping, tamper-evident provenance, and programmable ownership/rights management for cryptographically unique digital assets, including frameworks that explicitly discuss representing patents as NFTs. This review explores whether an NFT-enabled registration and transfer layer could

 doi:10.20944/preprints202602.1981.v1

complement existing IP and internal governance practices for DEL platforms, especially when the "asset unit" is defined at the library or sublibrary level, rather than per enumerated compounds.

2. DNA-Encoded Chemical Libraries in Drug Discovery: Value, Scale, and Translation

DELs are collections of small molecules covalently linked to DNA tags that serve as amplifiable barcodes. These barcodes enable pooled handling and selection against protein targets followed by sequencing-based identification of enriched binders. Modern DEL workflows integrate DNA-compatible chemistry, split-and-pool synthesis logic, affinity selection, next-generation sequencing, and computational deconvolution, allowing libraries that can reach millions to billions of members in a practical experimental footprint. As an example, a concrete translational signal is the existence of DEL-enabled discovery programs that progressed to clinical candidates, including well-documented RIPK1 inhibitor series where DEL played a role in hit identification and subsequent optimization toward clinical candidates. The growing diversity of DEL chemistries and selection modalities have led to systematic off-DNA resynthesis and structure-activity relationship (SAR) development. [1–3] The large scale of DEL changes the composition of "what must be managed," since library design files, barcode maps, building-block inventories, QC artifacts, and selection datasets become as important as the isolated compounds in the value chain. [2,4,5] Accordingly, DEL programs can be viewed as data-and-material hybrid assets, where the physical library stock and the digital representation of its selection history must remain synchronized to preserve reproducibility and downstream utility. [4–7]

3. Why Conventional IP and Asset-Management Workflows Strain at DEL Scale

Patent and trade-secret strategies for small molecules have historically been optimized for a limited set of lead series and discrete optimized structures, whereas DEL platforms generate and curate a far larger "pre-lead" universe where many members can never be individually resynthesized. [8,9] Though patenting is pursued only for prioritized DEL series, organizations still need internal mechanisms to establish provenance, document library creation, and control access to library designs and physical stocks across teams, partners, and time. [4] Moreover, DEL data analysis itself has recognized complexity sources, such as count-affinity relationships, frequent hitters, inhomogeneous composition, DNA damage, and linker effects, that motivate careful recordkeeping of library synthesis and selection context to interpret results reliably.[6] Thus, a "library registration" problem emerges: stakeholders may want verifiable timestamps and audit trails for when a library existed, what it contained, and who controlled access at key decision points. Traditional enterprise tools (ELN/LIMS, internal repositories, and contract documents) can store such information but cannot provide a single tamper-evident, cross-organization provenance layer that remains consistent when assets are transferred between entities or when collaborations dissolve and restart years later. [4,10–13]

In a biomedical context, the blockchain has been repeatedly proposed as a mechanism to provide immutable logging and shared auditability across parties that do not fully trust each other, suggesting an analogy for cross-institution DEL collaborations and licensing workflows. [10–14] A blockchain-based registry provides a single append-only audit trail that can be verified by any authorized party, thus reducing DEL reliance on third-party notarization. However, healthcare blockchain literature also stresses that native on-chain storage is impractical for large datasets due to privacy, governance, and regulatory constraints that often necessitate hybrid on-chain/off-chain architectures, which is relevant for sensitive DEL designs and selection data.[10–13,15–17]

Therefore, the question is not whether NFTs can replace patents on DELs, but whether NFTs can serve as a scalable registry and transaction primitive that complements legal instruments while preserving confidentiality and enabling controlled transfer of rights to use a library or its derivatives.[14,18]

 doi:10.20944/preprints202602.1981.v1

4. NFTs and Blockchain for Syntheric Chemistry: Concepts and Reported Biomedical-Industrial Patterns

A blockchain is a distributed ledger technology that records transactions in a tamper-evident manner across a network of participants, and it is frequently discussed in healthcare as a way to improve integrity, auditability, and multi-party coordination where trust is fragmented.[10–13] Smart contracts are programmable logic deployed on certain blockchain platforms that automate conditional actions and logging, thus motivating frameworks for automated governance in data sharing, monitoring, and supply-chain contexts.[13,19,20] In healthcare prototypes, smart contracts have been implemented for event logging and access actions, illustrating the feasibility of coupling real-world processes to on-chain records while preserving an auditable trace.[19]

An NFT is commonly defined as a unique, non-interchangeable token recorded on a blockchain that can represent a digital asset and can be transferred while retaining a transparent transaction history.[18,21–24] Recent biomedical discussions of NFTs have focused on potential use cases such as data provenance, credentialing, patient controlled data access, and incentive mechanisms, while simultaneously noting unresolved issues in governance, interoperability, privacy, and speculative misuse.[21–24]A key technical design pattern across healthcare blockchain reviews is to store only minimal identifiers, hashes, and permissions on-chain, while keeping sensitive or large payloads off-chain in controlled repositories, which reduces privacy leakage and addresses scalability constraints.[10–13,15,17,25]

For pharmaceutical supply chains, blockchain has been investigated as a means to improve traceability and detect falsified or substandard drugs, providing an existence proof that multi-stakeholder chemical-product ecosystems potentially benefit from shared provenance layers.[26–28] These patterns imply that an NFT-enabled DEL registry should treat a DEL as a permissioned information object whose on-chain footprint is a cryptographic commitment and pointer, rather than an on-chain disclosure of uniquely synthesized chemical structures that could undermine confidentiality or patent strategy.[14,15,17,18,25]

5. Tokenizing Patents and IP

Peer-reviewed literatures, although there have been much more practical but unreviewed cases and actions, explicitly proposing “patents as NFTs” emphasize that NFTs provide traceable ownership records, facilitate market liquidity, and enable programmable royalty logic, while also stressing that legal enforceability and standardization remain open challenges.[14,18] In biomedical innovation contexts, blockchain-based frameworks have been proposed to secure IP transactions and reduce transaction costs in multi-stakeholder settings, reflecting a broader trend toward “infrastructure for collaboration” rather than purely speculative crypto-assets.[14] This point is relevant for pharmaceuticals because many valuable preclinical assets, including platform know-how, datasets, and early-stage chemical matter, are routinely shared via NDAs, MTAs, and licensing agreements that require auditable records, clear scope, and friction-reduced contracts.[14,18] From a design perspective, NFT-based IP frameworks typically decompose the system into identity and permission, asset metadata standardization, token minting and wallet custody, transaction execution, and dispute-governance layers, which can be mapped onto DEL assets with appropriate domain-specific metadata.[14,18]

The core claim that can be made is therefore not that NFT tokenization replaces formal patent filing, but that it can provide a tamper-evident registry and transaction log that complements IP operations and potentially enable steps from novel licensing to royalty automation if coupled to enforceable contracts.[14,18] Because DEL assets are often mixtures plus metadata rather than individually isolated compounds, the unit of tokenization should be carefully chosen (e.g., library, sublibrary, selection dataset, or a validated hit series) to align with what can be legally transferred and experimentally reconstituted.[2,4,5,14,18]

 doi:10.20944/preprints202602.1981.v1

6. Applying NFTs to DEL: Conceptual Architecture, Benefits, and Limitations

6.1. Asset Definition and Tokenized Selection

A practical NFT-for-DEL design begins by defining an "asset unit" that is meaningful scientifically and legally, such as a complete DEL, a sublibrary defined by scaffold building-block partitions, or a selection package that bundles selection conditions and enriched hit lists.[2,4,5,14] This asset definition should explicitly include both (i) physical material custody where applicable and (ii) the digital description needed to interpret or reproduce the asset.[4,5] Hierarchical tokenization should be used to avoid minting billions of NFTs for each small molecule in a DEL by issuing a master token for the library and derivative tokens for sublibraries or curated subsets, which mirrors how DEL is operationally managed in pools rather than as isolated members.[2,5,7]

6.2. Hybrid on-Chain/off-Chain Metadata with the Confidentiality by Design

Healthcare blockchain reviews frequently conclude that large or sensitive payloads should remain off-chain, and that on-chain data should be limited to hashes, pointers, and access-control events, which is directly applicable to DEL designs that would embody trade secrets or pre-patent information.(10-13, 15, 17, 25) Accordingly, the NFT metadata can store a cryptographic hash of an encrypted "DEL package" including design files, synthesis route summary, barcode map, QC summaries, and selection provenance, while the package itself is stored in controlled infrastructure with role-based access. (10-13, 15, 17, 25) A minimal, domain-appropriate metadata standard is essential for interoperability, which can be established by encoding a unique library identifier, asset unit type, creation timestamp, a cryptographic hash of the associated off-chain DEL, and ownership references. This approach supports tamper-evident integrity checking, as any change to the DEL package would break the on-chain hash match without publishing chemical structures to a public ledger.(10-13, 18)

6.3. Transaction and Transfer Workflows from Registry to Rights Movement

NFT transfers can provide a transparent transaction history, but peer-reviewed IP-tokenization work stresses that enforceable meaning requires coupling token transfers to off-chain legal agreements that define what rights move such as sublicense, reproduce, resynthesize, or access data, and under what constraints.[14,18] A workable workflow is to bind each NFT to (i) an NDA/MTA/license template, (ii) a controlled data-room access policy, and (iii) a physical chain-of-custody procedure for any transferred library stock, thereby aligning cryptographic ownership with operational control.(14, 18, 26-28)

Smart contracts can automate milestone logging, time-bounded access, and royalty-trigger events, but healthcare smart-contract literature also emphasizes the need for careful governance and security review to avoid brittle automation and unintended disclosures.[19,20]

6.4. Expected Positive Impacts for Pharma DEL Operations

A primary benefit for storing DELs as NFTs is providing tamper-evident timestamping and provenance, which healthcare blockchain literature frames as a way to reduce disputes and increase auditability in multi-party settings that lack a single trusted administrator.[10–14] For DEL, this can translate into standardized "library birth certificates" that record creation events, versioning of barcode maps, and custody transitions, improving continuity across personnel turnover and long program timelines.[4,5,14] A second benefit is establishing a scalable registry and enabling the discovery of assets, because tokenization frameworks for IP discuss how standardized metadata and transparent ownership logs can reduce friction in licensing and collaboration marketplaces.[14,18] A third benefit is promoting programmable incentives, where NFT-linked royalty logic is proposed as a way to enable downstream value sharing, which could be relevant for cross-company collaborations, CRO-mediated screening, or academic–industry DEL partnerships.[14,18] Finally,

 doi:10.20944/preprints202602.1981.v1

blockchain supply-chain work in pharma suggests that provenance layers can strengthen trust in distributed ecosystems, implying that a similar approach may help govern the movement of DEL materials and datasets across organizational boundaries.[26–28]

6.5. Risks, Drawbacks, and Unresolved Questions

Legal enforceability remains the primary limitation, as peer-reviewed "patent-as-NFT" work repeatedly notes that token records must be anchored to recognized legal instruments and that jurisdictions may not treat token possession as equivalent to IP title.[14,18] Confidentiality risk is also substantial, since metadata leakage (e.g., target name, scaffold hints, or timestamps) could compromise competitive positioning, motivating privacy preserving designs, and careful selection of what is placed on-chain.[14,15,17,18,26] Scalability and operational overhead are nontrivial, as healthcare blockchain systematic reviews repeatedly identify interoperability, governance, identity management, and integration with legacy systems as major adoption barriers.[10–13,17] Smart contract risk is not theoretical, because healthcare smart-contract discussions emphasize that correctness, security audits, and clear dispute-resolution mechanisms are prerequisites for safe automation in high-stakes contexts.[19,20] Finally, scientific interpretability constraints in DEL, such as data artifacts and context dependence in selection readouts, mean that ownership transfer must include sufficient method and QC provenance to avoid transferring an asset that is legally "owned" but scientifically ambiguous.[5]

6.6. Implementation Roadmap

A practical first pilot is library-level tokenization for a limited number of internal DELs, using a permissioned blockchain and off-chain encrypted storage to test integrity, access logging, and custody transfer without exposing competitive chemistry.(10-15, 17, 26) A pharmaceutical organization could mint an NFT to store the design files and barcode maps of a newly synthesized DEL with compatibility to upload subsequent datasets as derivative tokens referencing the parent library while keeping an uninterrupted license agreement. A second phase can extend to external collaborations by binding NFT transfers to standardized NDAs/MTAs and controlled data rooms, thereby testing whether token-based registries reduce negotiation friction while maintaining confidentiality.[14,18] A third phase can explore revenue-sharing or licensing automation for curated hit series or validated tool compounds, where the "asset unit" is experimentally grounded and legal scope is easier to define than for an entire enumerated library.[2,4,14,18] (Figure 1) Throughout, governance should be treated as a first-class design variable, consistent with healthcare blockchain lessons that technical immutability does not eliminate the need for clear policies, identity controls, and dispute pathways.(10-13, 17, 20).

 doi:10.20944/preprints202602.1981.v1

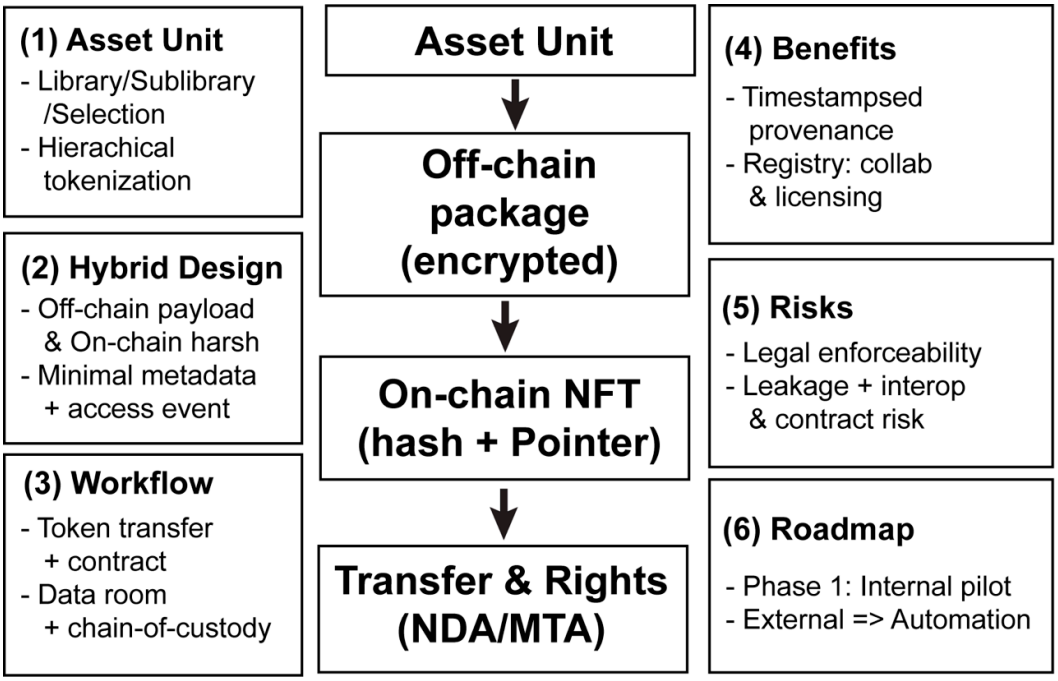

Figure 1. Technical schematic for NFT-enabled DEL registration and transfer. Schematic for an NFT-enabled hybrid on-chain/off-chain framework to register and transfer DNA-encoded chemical library (DEL) assets. A scientifically and legally meaningful DEL "asset unit" (library/sublibrary/selection package) is bundled into an encrypted off-chain DEL package whose integrity is committed on-chain via a cryptographic hash and a pointer URI. A permissioned NFT registry records minimal metadata and smart-contract records (mint/transfer/access-grant/milestone/royalty logs) while enforceable rights are defined by linked NDA/MTA/license agreements and, where applicable, physical chain-of-custody with QC provenance. Callouts summarize expected impacts, risks, and a phased implementation roadmap.

7. Conclusions

DEL technology has become a mainstream platform for small-molecule discovery and has demonstrated credible translational outputs, while simultaneously creating a large and complex surface of library-scale assets that strain traditional registration and transfer workflows. NFT and blockchain concepts, especially the peer-reviewed proposals for tokenizing patents and securing IP transactions, suggest a plausible direction for building a scalable, tamper-evident registry layer that can complement legal agreements rather than replace them. For DEL specifically, the most defensible near-term design is a hybrid architecture that stores minimal commitments on-chain and keeps sensitive chemical metadata off-chain under access control, coupled to enforceable NDAs/MTAs/licenses and physical chain-of-custody procedures. The key research agenda is therefore technical information including standardizing DEL token metadata, validating governance models, quantifying transaction-to-cost reductions, and ensuring confidentiality and compliance in real world pharma collaborations. If these gaps are addressed, NFT-enabled DEL registries could evolve from "speculative crypto tooling" into a practical infrastructure for scalable, auditable, and market compatible management of ultra large chemical asset collections.

Author Contributions: S.C. and S.U.K.: Conceptualization, S.C., B.W., and S.U.K.: DEL, NFT and IP literature review, S.C., B.W., and S.U.K.: writing - original draft and editing, S.U.K.: Supervision, S.C. and S.U.K.: Visualization.

Funding: no funding was used for this review article.

Institutional Review Board Statement: Not applicable.

 doi:10.20944/preprints202602.1981.v1

Informed Consent Statement: All authors have agreed to the publication of this manuscript.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: Authors declare that there is no conflict of interest.

References

1. Gironda-Martinez A, Donckele EJ, Samain F, Neri D. DNA-Encoded Chemical Libraries: A Comprehensive Review with Succesful Stories and Future Challenges. ACS Pharmacol Transl Sci. 2021;4(4):1265-79.
2. Peterson AA, Liu DR. Small-molecule discovery through DNA-encoded libraries. Nat Rev Drug Discov. 2023;22(9):699-722.
3. Yuen LH, Franzini RM. Achievements, Challenges, and Opportunities in DNA-Encoded Library Research: An Academic Point of View. Chembiochem. 2017;18(9):829-36.
4. Dockerill M, Winssinger N. DNA-Encoded Libraries: Towards Harnessing their Full Power with Darwinian Evolution. Angew Chem Int Ed Engl. 2023;62(9):e202215542.
5. Wichert M, Guasch L, Franzini RM. Challenges and Prospects of DNA-Encoded Library Data Interpretation. Chem Rev. 2024;124(22):12551-72.
6. Liszczak G, Muir TW. Nucleic Acid-Barcoding Technologies: Converting DNA Sequencing into a Broad-Spectrum Molecular Counter. Angew Chem Int Ed Engl. 2019;58(13):4144-62.
7. Pang D, Ashkan K. Deep brain stimulation for phantom limb pain. Eur J Paediatr Neurol. 2022;39:96-102.
8. Bieck P, Antonin KH, Jedrychowski M. Monoamine oxidase inhibition in healthy volunteers by CGP 11305 A, a new specific inhibitor of MAO-A. Mod Probl Pharmacopsychiatry. 1983;19:53-62.
9. Belyanskaya SL, Ding Y, Callahan JF, Lazaar AL, Israel DI. Discovering Drugs with DNA-Encoded Library Technology: From Concept to Clinic with an Inhibitor of Soluble Epoxide Hydrolase. Chembiochem. 2017;18(9):837-42.
10. Agbo CC, Mahmoud QH, Eklund JM. Blockchain Technology in Healthcare: A Systematic Review. Healthcare (Basel). 2019;7(2).
11. Benchoufi M, Ravaud P. Blockchain technology for improving clinical research quality. Trials. 2017;18(1):335.
12. Drosatos G, Kaldoudi E. Blockchain Applications in the Biomedical Domain: A Scoping Review. Comput Struct Biotechnol J. 2019;17:229-40.
13. Vazirani AA, O'Donoghue O, Brindley D, Meinert E. Implementing Blockchains for Efficient Health Care: Systematic Review. J Med Internet Res. 2019;21(2):e12439.
14. Liang HW, Chu YC, Han TH. Fortifying Health Care Intellectual Property Transactions With Blockchain. J Med Internet Res. 2023;25:e44578.
15. Godyn M, Kedziora M, Ren Y, Liu Y, Song HH. Analysis of solutions for a blockchain compliance with GDPR. Sci Rep. 2022;12(1):15021.
16. Harris PA, Berger SB, Jeong JU, Nagilla R, Bandyopadhyay D, Campobasso N, et al. Discovery of a First-in-Class Receptor Interacting Protein 1 (RIP1) Kinase Specific Clinical Candidate (GSK2982772) for the Treatment of Inflammatory Diseases. J Med Chem. 2017;60(4):1247-61.
17. Kakarlapudi PV, Mahmoud QH. A Systematic Review of Blockchain for Consent Management. Healthcare (Basel). 2021;9(2).
18. Bamakan SMH, Nezhadsistani N, Bodaghi O, Qu Q. Patents and intellectual property assets as non-fungible tokens; key technologies and challenges. Sci Rep. 2022;12(1):2178.
19. Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T. Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring. J Med Syst. 2018;42(7):130.
20. Marino CA, Diaz Paz C. Smart Contracts and Shared Platforms in Sustainable Health Care: Systematic Review. JMIR Med Inform. 2025;13:e58575.
21. Abhari S, Morita P, Miranda P, Garavand A, Hanjahanja-Phiri T, Chumachenko D. Non-fungible tokens in healthcare: a scoping review. Front Public Health. 2023;11:1266385.

 doi:10.20944/preprints202602.1981.v1

22. Nunes T, da Cunha PR, de Abreu JM, Duarte J, Corte-Real A. Non-Fungible Tokens (NFTs) in Healthcare: A Systematic Review. Int J Environ Res Public Health. 2024;21(8).
23. Sibanda K, Ndayizigamiye P, Twinomurinzi H. Non-fungible tokens (NFTs) in healthcare: a thematic analysis and research agenda. Front Digit Health. 2024;6:1377531.
24. Yaghy A, Alberto NRI, Alberto IRI, Bermea RS, Ristovska L, Yaghy M, et al. The potential use of non-fungible tokens (NFTs) in healthcare and medical research. PLOS Digit Health. 2023;2(7):e0000312.
25. Hasselgren A, Kralevska K, Gligoroski D, Faxvaag A. GDPR Compliant Blockchain and Distributed Ledger Technologies in the Health Sector. Stud Health Technol Inform. 2020;270:1293-4.
26. Akram W, Joshi R, Haider T, Sharma P, Jain V, Garud N, et al. Blockchain technology: A potential tool for the management of pharma supply chain. Res Social Adm Pharm. 2024;20(6):156-64.
27. Sylim P, Liu F, Marcelo A, Fontelo P. Blockchain Technology for Detecting Falsified and Substandard Drugs in Distribution: Pharmaceutical Supply Chain Intervention. JMIR Res Protoc. 2018;7(9):e10163.
28. Zakari N, Al-Razgan M, Alsaadi A, Alshareef H, Al Saigh H, Alashaikh L, et al. Blockchain technology in the pharmaceutical industry: a systematic review. PeerJ Comput Sci. 2022;8:e840.