Saudi Journal of Medicine and Public Health

https://saudijmph.com/index.php/pub

https://doi.org/10.64483/202412605

# Unlocking Insights in Thyroid Care: Applying Natural Language Processing to Electronic Health Records for Improved Patient Stratification and Clinical Decision Support

**Shama Rshaid Bin Zwayed (1), Rowaf Zghir A Alrowaili (2), Mohammed Abdullah Ail Al Nosyan (3), Ali Mohammed Kleibi (4), Dalia Nawash Alanzi (5), Nawal Ibrahim Yaqub Qadah (6), Hanan Abraham Al Gezani (7), Albandary Awadh Almutairi (8), Majed Mislat Eid Albaqami (9), Zahra Ahmed Bosilly (10), Mofarhe Yahya Mashike (7), Ghaliah Musallam Alhawit (5)**

*(1) The First Health Cluster In Riyadh,Ministry of Health, Saudi Arabia,*
*(2) Turaif General Hospital,Ministry of Health, Saudi Arabia,*
*(3) National Guard Health Affairs,Ministry of Health, Saudi Arabia,*
*(4) Jazan Health Cluster,Ministry of Health, Saudi Arabia,*
*(5) Maternity And Children's Hospital In Tabuk,Ministry of Health, Saudi Arabia,*
*(6) Sabya General Hospital,Ministry of Health, Saudi Arabia,*
*(7) Bish General Hospital, Ministry of Health, Saudi Arabia,*
*(8) Riyadh Specialized Dental Center,Ministry of Health, Saudi Arabia,*
*(9) King Salman Hospital,Ministry of Health, Saudi Arabia,*
*(10) Ahad Al-Masrah General Hospital,Ministry of Health, Saudi Arabia*

**Abstract**

**Background:** The clinical management of thyroid disorders, ranging from autoimmune conditions to cancer, generates vast amounts of unstructured data within Electronic Health Records (EHRs). This data, embedded in clinical notes, radiology reports, and pathology summaries, contains rich phenotypic details that are largely inaccessible to traditional analytical methods, creating a significant information gap for research and precision care. **Aim:** This narrative review aims to synthesize the current landscape, methodologies, challenges, and future directions of applying Natural Language Processing (NLP) to mine EHRs for thyroidology. It evaluates how NLP can transform unstructured text into structured data to enhance patient stratification, support clinical decisions, and advance epidemiological research. **Methods:** A comprehensive literature search was conducted across PubMed, IEEE Xplore, and ACL Anthology for studies published between 2010 and 2024, using keywords related to NLP, EHRs, and thyroid disorders. Relevant studies were selected and thematically analyzed. **Results:** The review identifies key NLP architectures—from rule-based systems to deep learning models—successfully applied to extract thyroid-specific concepts, automate TI-RADS scoring, predict outcomes, and identify adverse events. However, significant challenges persist, including data heterogeneity, clinical nuance, and ethical concerns regarding bias and generalizability. **Conclusion:** NLP is a powerful, transformative tool for thyroid care, poised to unlock latent insights from EHRs. Realizing its full potential requires interdisciplinary collaboration, robust validation, and the development of standardized, ethically aware frameworks to integrate these technologies into clinical workflows and research infrastructures.

**Keywords:** Natural Language Processing, Electronic Health Records, Thyroid Disorders, Clinical Decision Support, Artificial Intelligence.

## Introduction

The rising global prevalence of thyroid disorders, including hypothyroidism, hyperthyroidism, thyroid nodules, and thyroid cancer, presents a formidable challenge to healthcare systems (Taylor et al., 2018). The management of these conditions is inherently information-intensive, relying on the synthesis of symptoms, laboratory values, imaging descriptors, cytopathological findings, and longitudinal treatment responses. The widespread adoption of Electronic Health Records (EHRs) has digitized this clinical narrative, creating vast repositories of patient data. However, a critical paradox exists: while data abundance is greater than ever, actionable knowledge remains constrained. An estimated 80% of clinically relevant information in EHRs is stored as unstructured free text—radiologists' impressions, operative notes, endocrinologists'



*Corresponding author e-mail: oudramdan@gmail.com (Shama Rshaid Bin Zwayed).

Receive Date: 23 December 2024, Revise Date: 30 December 2024, Accept Date: 31 December 2024

progress notes, and pathology reports (Wang et al., 2018). This unstructured data encapsulates the nuance and clinical reasoning essential for personalized care but is inaccessible to conventional data analytics, which depend on structured fields like diagnosis codes and lab numbers.

This gap between data availability and insight extraction is where Natural Language Processing (NLP) emerges as a pivotal technology. NLP, a subfield of artificial intelligence (AI) and computational linguistics, provides the methodologies to computationally understand, interpret, and manipulate human language (Devlin et al., 2019). In healthcare, NLP applied to EHRs—often called clinical NLP—aims to transform unstructured text into structured, computable data. This capability is particularly salient for thyroidology. The specialty's dependence on descriptive diagnostics (Ultrasound TI-RADS lexicon, Bethesda System for cytology), nuanced treatment decisions, and long-term surveillance generates text-rich records perfect for NLP mining (Haugen et al., 2016).

This narrative review synthesizes the evolution, current state, and future trajectory of NLP applications in thyroid care. We will explore the fundamental architectures of NLP systems, from early rule-based approaches to contemporary deep learning models. A critical analysis of their application to specific tasks—such as phenotype identification, nodule risk stratification, outcome prediction, and adverse event detection—will be presented. The review will also confront the significant challenges of data heterogeneity, model bias, and clinical integration, while proposing a roadmap for translating NLP research into robust, equitable tools that enhance patient stratification and empower clinical decision support.

**Thyroid Care in the Era of the Electronic Health Record**

The modern EHR is a mosaic of data types, each contributing a piece to the thyroid patient's puzzle. Structured data includes laboratory results (TSH, Free T4, Thyroglobulin), medication lists (levothyroxine, methimazole), and coded diagnoses using the International Classification of Diseases (ICD). While crucial, these elements offer a skeletal view. The "clinical flesh" is found in the unstructured narratives. Radiology reports describe nodule echogenicity, margins, and composition using standardized lexicons like TI-RADS, but in prose form (Tessler et al., 2017). Pathology reports detail cytologic and histologic findings, referencing the Bethesda System and cancer staging criteria (Cibas & Ali, 2017). Progress notes capture symptom evolution, family history, and therapeutic rationale. Discharge summaries consolidate the entire care episode.

Mining this textual data manually for research or quality improvement is prohibitively time-consuming and prone to error. Furthermore, reliance on structured codes alone is deeply flawed for thyroid research (Xue et al., 2022). ICD codes for thyroid cancer, for example, lack specificity for histologic subtype, and codes for hypothyroidism do not capture etiology (e.g., Hashimoto's vs. post-RAI) (Brito et al., 2014). This leads to misclassification bias and hinders precise cohort identification for clinical trials or outcomes research. NLP promises to overcome these limitations by extracting and codifying the detailed information locked in text, enabling the creation of rich, patient-specific phenotypes that reflect true clinical complexity (Shi et al., 2022).

**Foundations of Clinical Natural Language Processing**

The evolution of clinical NLP mirrors advances in AI. Early systems were predominantly rule-based or dictionary-driven. These methods rely on lexicons of clinical terms and hand-crafted grammatical rules to identify concepts. For instance, a rule might specify that the phrase "hypoechoic nodule" within three words of "microcalcifications" indicates a high-risk feature. Systems like the Clinical Text Analysis and Knowledge Extraction System (cTAKES) were pioneers in this space, providing open-source frameworks for concept extraction (Savova et al., 2010). While transparent and effective for specific, well-defined tasks, rule-based systems are brittle. They struggle with linguistic variation, negation ("no suspicious features"), and contextual ambiguity, requiring extensive, domain-specific engineering.

The advent of machine learning (ML) introduced statistical models that learn patterns from annotated examples. Models like Conditional Random Fields could learn to label sequences of words (token-level) to identify entities such as "Medication" or "Disease" (Li et al., 2022). The true paradigm shift, however, came with deep learning and the development of transformer-based language models. Models like BERT (Bidirectional Encoder Representations from Transformers) and its biomedical variants (BioBERT, ClinicalBERT) are pre-trained on massive corpora of text, learning deep contextual representations of language (Lee et al., 2020). They understand that "mass" in an oncology note differs from "mass" in physics. These models can be fine-tuned with relatively small sets of labeled clinical notes to perform state-of-the-art tasks like named entity recognition (NER), relation extraction, and document classification with remarkable accuracy, significantly reducing the need for manual feature engineering (Liu et al., 2020). Table 1 & Figure 1 depict a schematic overview of an NLP-driven analytics pipeline for thyroid care. Unstructured EHR text sources (clinical notes, radiology reports, pathology summaries) are processed through NLP architectures ranging from rule-based systems to transformer models.

**Saudi Journal of Medicine and Public Health**

**Table 1: Comparison of NLP Architectures for Clinical Text Mining**

https://doi.org/10.64483/202412605

| Architecture | Description | Pros | Cons | Example Thyroid Application |
|---|---|---|---|---|
| **Rule-Based** | Uses lexicons & hand-written grammatical/syntactic rules. | Transparent, interpretable, and effective for consistent phrasing. | Brittle, labor-intensive to create/maintain, poor with variation/negation. | Extracting "nodule size > 4cm" from ultrasound reports. |
| **Traditional Machine Learning (e.g., SVM, CRF)** | Uses statistical models trained on feature-annotated datasets. | More robust to variation than rules; well-established. | Requires heavy feature engineering; performance plateaus. | Classifying pathology reports as benign vs. malignant. |
| **Deep Learning (e.g., RNNs, CNNs)** | Uses neural networks to automatically learn feature representations. | Reduces need for feature engineering; good sequential modeling. | Requires large datasets; can be a "black box." | Temporal modeling of treatment response from notes. |
| **Transformer Models (e.g., BERT, BioBERT)** | Uses self-attention mechanisms pre-trained on vast text corpora. | State-of-the-art performance; understands deep context. | Computationally intensive; requires fine-tuning; "black box." | Extracting complex phenotypes (e.g., RAI-refractory disease). |

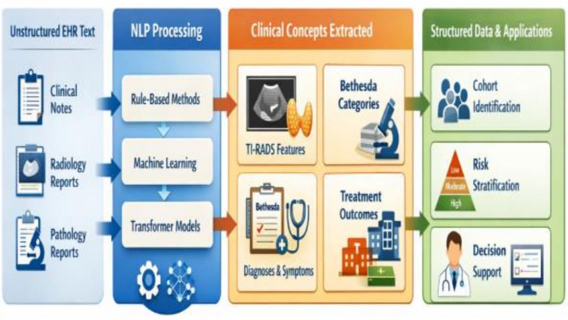


**Figure 1. NLP-Enabled Pipeline for Thyroid Patient Stratification from EHRs**

**NLP Applications in Thyroidology**

**Phenotyping and Cohort Identification**

A foundational application is the accurate identification of patient cohorts. NLP can move beyond ICD codes to find patients with specific phenotypes. For example, Shin et al. (2021) used an NLP algorithm on clinical notes to identify patients with thyroid eye disease (TED), accurately distinguishing active from inactive disease based on descriptive symptoms and clinical findings mentioned in ophthalmology and endocrinology notes. Similarly, NLP can parse notes to determine the etiology of hypothyroidism (e.g., identifying mentions of "Hashimoto's," "thyroiditis," or "post-thyroidectomy") or to find patients with rare conditions like resistance to thyroid hormone (Wang et al., 2020). This enables more precise recruitment for clinical trials and more accurate observational studies on disease natural history (Shcherbak et al., 2022).

**Imaging and Pathology Report Mining**

Thyroid nodule management is guided by risk stratification systems. NLP excels at automating this. Algorithms can parse ultrasound reports to extract TI-RADS features (composition, echogenicity, shape, margin, echogenic foci) and calculate a risk score automatically (Zhao et al., 2021). This not only standardizes reporting but also enables large-scale audits of real-world nodule management against established guidelines. In pathology, NLP can classify fine-needle aspiration (FNA) reports according to the Bethesda System, identify non-diagnostic samples, and even extract key prognostic features from surgical pathology reports for thyroid cancer, such as lymphovascular invasion or extrathyroidal extension (Wong et al., 2023). This automates registry data entry, facilitating rapid case identification for tumor boards (Zhang et al., 2023).

**Outcome Prediction and Prognostication**

By integrating structured data with rich phenotypes extracted via NLP, predictive models gain power. For instance, combining extracted features from operative notes ("parathyroid autotransplantation") and post-operative notes (symptoms of "tingling" or "cramps") with lab values


*Corresponding author e-mail: oudramdan@gmail.com (Shama Rshaid Bin Zwayed).
Receive Date: 23 December 2024, Revise Date: 30 December 2024, Accept Date: 31 December 2024

can improve the prediction of hypoparathyroidism risk after total thyroidectomy (Al-Dhahri et al., 2014). In thyroid cancer, mining oncology notes for terms related to treatment response ("rising thyroglobulin," "structural progression") and imaging reports can help predict recurrence risk or identify patients with RAI-refractory disease earlier than structured data alone would allow (Lamartina et al., 2018).

**Adverse Event and Comorbidity Surveillance**

NLP is a powerful tool for pharmacovigilance and outcomes monitoring. It can scan clinical notes for undocumented side effects of anti-thyroid drugs (agranulocytosis, hepatotoxicity) or chronic conditions associated with thyroid dysfunction, such as atrial fibrillation in thyrotoxicosis or depression in hypothyroidism (Chaker et al., 2016). This supports real-world safety studies and helps map the complex comorbidity networks in thyroid patients (Table 2). Figure 2 illustrates key applications of natural language processing in thyroid care.

**Table 2: Exemplar Studies Applying NLP to Thyroid Care (2015-2024)**

| Study (Author, Year) | Primary NLP Task | Data Source | Key Finding/Application |
|---|---|---|---|
| **Wong et al., 2023** | Bethesda System classification from cytology reports. | Pathology reports from a multi-hospital network. | NLP model achieved >95% accuracy in classifying FNA reports, enabling automated cancer registry updates. |
| **Zhao et al., 2021** | Automated TI-RADS scoring from ultrasound reports. | Retrospective ultrasound reports. | Model extracted features with high F1-scores (>0.88), demonstrating feasibility for clinical decision support. |
| **Shin et al., 2021** | Phenotyping Thyroid Eye Disease (TED) activity. | Ophthalmology & endocrinology clinical notes. | NLP algorithm accurately identified active vs. inactive TED, outperforming billing code-based methods. |
| **Al-Dhahri et al., 2014** | Predicting post-thyroidectomy hypoparathyroidism. | Operative & post-operative notes + labs. | Model incorporating NLP-extracted surgical details outperformed models using labs alone. |
| **Wang et al., 2020** | Identifying rare thyroid disorder cases. | Longitudinal EHR notes across specialties. | NLP-enabled case-finding facilitated recruitment for a study on genetic thyroid disorders. |

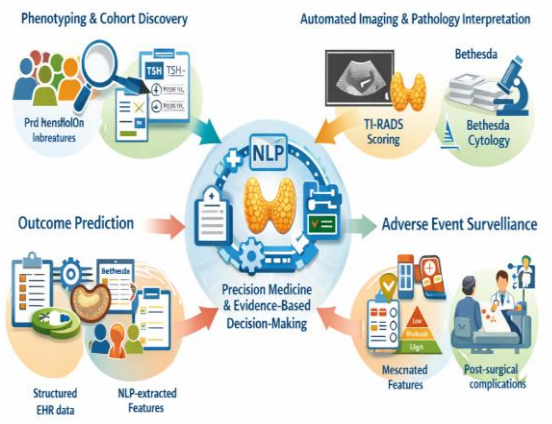


**Figure 2. Clinical Applications of NLP Across the Thyroid Care Continuum**

**Critical Challenges and Ethical Considerations**

Despite its promise, the path to integrating NLP into routine thyroid care is fraught with challenges.

**Data and Technical Hurdles**

EHR data is notoriously heterogeneous, with variations in documentation styles, templates, and abbreviations across institutions and even individual clinicians (Baclic et al., 2020). An NLP model trained on data from one academic hospital may fail in a community setting. Clinical language is dense with implicit knowledge, hedging ("suspicious for"), and complex coreferences (e.g., "the nodule" mentioned in paragraph three). Robust NLP systems must handle negation, temporality, and experiencer (e.g., family history vs. patient history) accurately (Uzuner et al., 2011). Furthermore, the "gold standard" for training and validation—manual chart review by clinicians—is expensive and time-consuming to produce at scale (Aversano et al., 2021).

**Bias, Fairness, and Generalizability**

AI models can perpetuate and amplify biases present in their training data. If EHR data from a particular demographic group (a specific racial, ethnic, or socioeconomic group) is underrepresented or documented differently, an NLP model may perform poorly for that group (Obermeyer et al., 2019). This could lead to disparities in automated risk stratification or clinical alerting. Ensuring fairness and auditing for bias are non-negotiable prerequisites for clinical deployment (Cary Jr et al., 2023).

**Clinical Integration and Evaluation**

The ultimate metric for any clinical NLP tool is whether it improves patient outcomes or clinician workflow. Moving from a research prototype to a validated clinical decision support system requires rigorous prospective evaluation in real-world settings. Integration into the clinician's EHR workflow must be seamless and non-disruptive, providing actionable information at the point of care (Sutton et al., 2020). Questions of liability, interpretability (why did the model make this recommendation?), and clinician trust remain significant barriers.

**Future Directions and Conclusion**

The future of NLP in thyroid care is integrative and translational. Next-generation systems will move beyond information extraction to true language understanding, capable of generating clinical summaries or drafting follow-up plans based on a patient's record (Kung et al., 2023). Multimodal AI models that jointly analyze text, medical images (ultrasound, CT scans), and genomic data will provide a holistic view of the patient (Huang et al., 2020). Federated learning approaches, where models are trained across multiple institutions without sharing raw patient data, offer a path to developing more robust and generalizable tools while preserving privacy (Rieke et al., 2020).

In conclusion, NLP stands as a key to unlocking the immense, untapped potential of unstructured data in thyroidology. It offers a pathway from descriptive, reactive documentation to predictive, proactive, and personalized care. By automating the extraction of detailed phenotypes, risk scores, and outcomes, NLP can power higher-quality research, refine clinical guidelines, and provide clinicians with intelligent tools for stratification and decision support. Realizing this vision demands a sustained, interdisciplinary collaboration among endocrinologists, surgeons, radiologists, pathologists, computer scientists, and ethicists. Together, they must build standardized, equitable, and clinically validated NLP systems that seamlessly integrate into the care continuum, ultimately enhancing the precision and quality of life for patients with thyroid disorders.

**References**


1. Al-Dhahri, S. F., Mubasher, M., Al-Muhawas, F., Alessa, M., Terkawi, R. S., & Terkawi, A. S. (2014). Early prediction of oral calcium and vitamin D requirements in post-thyroidectomy hypocalcaemia. *Otolaryngology--Head and Neck Surgery*, *151*(3), 407-414. https://doi.org/10.1177/0194599814536848
2. Aversano, L., Bernardi, M. L., Cimitile, M., Iammarino, M., Macchia, P. E., Nettore, I. C., & Verdone, C. (2021). Thyroid disease treatment prediction with machine learning approaches. *Procedia Computer Science*, *192*, 1031-1040. https://doi.org/10.1016/j.procs.2021.08.106
3. Baclic, O., Tunis, M., Young, K., Doan, C., Swerdfeger, H., & Schonfeld, J. (2020). Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, *46*(6), 161. https://doi.org/10.14745/ccdr.v46i06a02
4. Brito, J. P., Gionfriddo, M. R., Al Nofal, A., Boehmer, K. R., Leppin, A. L., Reading, C., ... & Montori, V. M. (2014). The accuracy of thyroid nodule ultrasound to predict thyroid cancer: systematic review and meta-analysis. *The Journal of Clinical Endocrinology & Metabolism*, *99*(4), 1253-1263. https://doi.org/10.1210/jc.2013-2928
5. Cary Jr, M. P., Zink, A., Wei, S., Olson, A., Yan, M., Senior, R., ... & Pencina, M. J. (2023). Mitigating racial and ethnic bias and advancing health equity in clinical algorithms: a scoping review: scoping review examines racial and ethnic bias in clinical algorithms. *Health Affairs*, *42*(10), 1359-1368. https://doi.org/10.1377/hlthaff.2023.00553
6. Chaker, L., van den Berg, M. E., Niemeijer, M. N., Franco, O. H., Dehghan, A., Hofman, A., ... & Peeters, R. P. (2016). Thyroid function and sudden cardiac death: a prospective population-based cohort study. *Circulation*, *134*(10), 713-722. https://doi.org/10.1161/CIRCULATIONAHA.115.020789
7. Cibas, E. S., & Ali, S. Z. (2017). The 2017 Bethesda system for reporting thyroid cytopathology. *Thyroid*, *27*(11), 1341-1346. https://doi.org/10.1089/thy.2017.0500
8. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)* (pp. 4171-4186). Doi: 10.18653/v1/N19-1423
9. Haugen, B. R., Alexander, E. K., Bible, K. C., Doherty, G. M., Mandel, S. J., Nikiforov, Y. E., ... & Wartofsky, L. (2016). 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid*, *26*(1), 1-133. https://doi.org/10.1089/thy.2015.0020
10. Huang, S. C., Pareek, A., Seyyedi, S., Banerjee, I., & Lungren, M. P. (2020). Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines. *NPJ digital medicine*, *3*(1), 136. https://doi.org/10.1038/s41746-020-00341-z

11. Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., ... & Tseng, V. (2023). Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS digital health*, *2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198
12. Lamartina, L., Grani, G., Durante, C., Borget, I., Filetti, S., & Schlumberger, M. (2018). Follow-up of differentiated thyroid cancer–what should (and what should not) be done. *Nature Reviews Endocrinology*, *14*(9), 538-551. https://doi.org/10.1038/s41574-018-0068-3
13. Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, *36*(4), 1234-1240. https://doi.org/10.1093/bioinformatics/btz682
14. Li, Y., Li, C., Li, X., Wang, K., Rahaman, M. M., Sun, C., ... & Wang, Q. (2022). A comprehensive review of Markov random field and conditional random field approaches in pathology image analysis. *Archives of Computational Methods in Engineering*, *29*(1), 609-639. https://doi.org/10.1007/s11831-021-09591-w
15. Liu, Y., Che, W., Qin, B., & Liu, T. (2020). Exploring segment representations for neural semi-Markov conditional random fields. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *28*, 813-824. https://doi.org/10.1109/TASLP.2020.2964960
16. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447-453. https://doi.org/10.1126/science.aax2342
17. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ digital medicine*, *3*(1), 119. https://doi.org/10.1038/s41746-020-00323-1
18. Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., & Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*, *17*(5), 507-513. https://doi.org/10.1136/jamia.2009.001560
19. Shcherbak, S. G., Changalidi, A. I., Barbitoff, Y. A., Anisenkova, A. Y., Mosenko, S. V., Asaulenko, Z. P., ... & Glotov, O. S. (2022). Identification of genetic risk factors of severe COVID-19 using extensive phenotypic data: A proof-of-concept study in a cohort of Russian patients. *Genes*, *13*(3), 534. https://doi.org/10.3390/genes13030534
20. Shi, M., Nong, D., Xin, M., & Lin, L. (2022). Accuracy of ultrasound diagnosis of benign and malignant thyroid nodules: a systematic review and meta-analysis. *International Journal of Clinical Practice*, *2022*(1), 5056082. https://doi.org/10.1155/2022/5056082
21. Shin, D., Kam, H. J., Jeon, M. S., & Kim, H. Y. (2021). Automatic classification of thyroid findings using static and contextualized ensemble natural language processing systems: development study. *JMIR Medical Informatics*, *9*(9), e30223. https://doi.org/10.2196/30223
22. Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ digital medicine*, *3*(1), 17. https://doi.org/10.1038/s41746-020-0221-y
23. Taylor, P. N., Albrecht, D., Scholz, A., Gutierrez-Buey, G., Lazarus, J. H., Dayan, C. M., & Okosieme, O. E. (2018). Global epidemiology of hyperthyroidism and hypothyroidism. *Nature Reviews Endocrinology*, *14*(5), 301-316. http://dx.doi.org/10.1038/nrendo.2018.18
24. Tessler, F. N., Middleton, W. D., Grant, E. G., Hoang, J. K., Berland, L. L., Teefey, S. A., ... & Stavros, A. T. (2017). ACR thyroid imaging, reporting and data system (TI-RADS): white paper of the ACR TI-RADS committee. *Journal of the American college of radiology*, *14*(5), 587-595. https://doi.org/10.1016/j.jacr.2017.01.046
25. Uzuner, Ö., South, B. R., Shen, S., & DuVall, S. L. (2011). 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association*, *18*(5), 552-556. https://doi.org/10.1136/amiajnl-2011-000203
26. Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., Afzal, N., ... & Liu, H. (2018). Clinical information extraction applications: a literature review. *Journal of biomedical informatics*, *77*, 34-49. https://doi.org/10.1016/j.jbi.2017.11.011

27. Wang, Y., Zhao, Y., Therneau, T. M., Atkinson, E. J., Tafti, A. P., Zhang, N., ... & Liu, H. (2020). Unsupervised machine learning for the discovery of latent disease clusters and patient subgroups using electronic health records. *Journal of biomedical informatics*, *102*, 103364. https://doi.org/10.1016/j.jbi.2019.103364
28. Wong, C. M., Kezlarian, B. E., & Lin, O. (2023). Current status of machine learning in thyroid cytopathology. *Journal of pathology informatics*, *14*, 100309. https://doi.org/10.1016/j.jpi.2023.100309
29. Xue, Y., Zhou, Y., Wang, T., Chen, H., Wu, L., Ling, H., ... & Wang, B. (2022). Accuracy of ultrasound diagnosis of thyroid nodules based on artificial intelligence-assisted diagnostic technology: a systematic review and meta-analysis. *International Journal of Endocrinology*, *2022*(1), 9492056. https://doi.org/10.1155/2022/9492056
30. Zhang, J., Li, J., Zhu, Y., Fu, Y., & Chen, L. (2023). Thyroidkeeper: a healthcare management system for patients with thyroid diseases. *Health Information Science and Systems*, *11*(1), 49. https://doi.org/10.1007/s13755-023-00251-w
31. Zhao, Z., Yang, C., Wang, Q., Zhang, H., Shi, L., & Zhang, Z. (2021). A deep learning-based method for detecting and classifying the ultrasound images of suspicious thyroid nodules. *Medical Physics*, *48*(12), 7959-7970. https://doi.org/10.1002/mp.15319 .