Nexus Junior Scientists Journal | Vol. 1, Issue 1 (2026) | Paper ID: 2605

Blood-Based Biomarkers for the Early Diagnosis of Pediatric Acute Lymphoblastic Leukemia: A Literature Review

Zayedan Ahmed Tayyib
Affiliation: Aitchison College, Lahore, Pakistan
Corresponding Author: Zayedan Ahmed Tayyib (zayedan.tayyib@gmail.com)

ABSTRACT

As the most common childhood malignancy, pediatric acute lymphoblastic leukemia (ALL) demands timely and accurate diagnosis to achieve favorable clinical outcomes. Because current diagnostic methods rely heavily on invasive bone marrow aspiration, reliable and minimally invasive alternatives are urgently needed. This review aims to synthesize and critically evaluate the current evidence on blood-based biomarkers, including circulating RNAs, proteins, metabolites, autoantibodies, and artificial intelligence-assisted diagnostic models, for the early diagnosis of pediatric ALL. Circulating RNA biomarkers and multi-marker diagnostic approaches appear to hold the greatest potential, although the existing evidence remains limited by small, single-center studies, methodological heterogeneity, and insufficient external validation. Several individual markers, including microRNA-146a and a multi-index artificial intelligence model built on routine hematological data, achieved area-under-the-curve values exceeding 0.95, underscoring the diagnostic promise of both molecular and computational approaches. Overall, blood-based biomarkers represent a promising complement to current diagnostic strategies, but standardized multicenter validation and robust multi-marker approaches are required before routine clinical implementation. Continued research in this field could improve early detection while reducing the need for invasive diagnostic procedures in children.

Keywords: acute lymphoblastic leukemia; blood-based biomarkers; pediatric oncology; early diagnosis; liquid biopsy; microRNA

1. INTRODUCTION

Acute Lymphoblastic Leukemia (ALL), a heterogeneous hematological malignancy marked by the uncontrolled proliferation of primitive lymphoid progenitor cells within the bone marrow, is the most prevalent form of pediatric cancer, accounting for approximately 25% to 30% of all childhood malignancies (Kowal et al., 2025; Papadopoulou et al., 2024; Zhang et al., 2019). Advancements in multidisciplinary treatment protocols have raised five-year survival rates to roughly 90% in developed nations, yet ALL remains a leading cause of disease-related mortality among children and adolescents (Papadopoulou et al., 2024; Shahid et al., 2021; Yu et al., 2019). This ongoing clinical challenge is compounded by the heterogeneous nature of the disease and the late clinical presentation of high-risk subtypes.

Primary prevention strategies, such as reducing exposure to environmental carcinogens, are largely ineffective against childhood cancers, which makes secondary prevention through early diagnosis

essential to maintaining high cure rates (Cavalcante et al., 2016). Timely detection matters not only for initiating rapid therapeutic intervention but also for managing immediate, life-threatening complications such as hyperleukocytosis and tumor lysis syndrome, which can occur at disease onset (Kowal et al., 2025). Early-stage ALL, though, is often difficult to identify: its initial symptoms are frequently non-specific and may mimic benign pediatric infections or normal physiological development (Cavalcante et al., 2016; Kowal et al., 2025).

The current "gold standard" for diagnosis integrates Morphology, Immunology, Cytogenetics, and Molecular biology (MICM) classification (Cheng et al., 2024; Yu et al., 2019; Zhang et al., 2019). These methods achieve high diagnostic accuracy, but they rely on invasive bone marrow aspiration and biopsy. These procedures are technically demanding, expensive, and profoundly traumatic for pediatric patients, often necessitating sedation or general anesthesia (Kowal et al., 2025). Standard morphological evaluation of peripheral blood may also fail to detect the disease in its earliest stages, when leukemic lymphoblasts have not yet emerged in the systemic circulation in significant numbers (Yu et al., 2019).

These limitations provide a strong rationale for investigating minimally invasive blood-based biomarkers that reflect the molecular and biochemical changes associated with pediatric ALL. Numerous biomarkers have demonstrated promising diagnostic potential, yet no single biomarker has consistently shown sufficient accuracy or clinical applicability to replace current diagnostic methods. Instead, the available evidence increasingly supports integrated multi-marker approaches that combine complementary biological information to improve diagnostic performance. This review critically evaluates the current evidence on blood-based biomarkers for the early diagnosis of pediatric ALL, synthesizing findings across biomarker categories while examining their diagnostic performance, methodological strengths, clinical limitations, and future translational potential.

METHODS: APPROACH TO THE REVIEW

This is a narrative literature review rather than a formal systematic review. It synthesizes findings from 14 peer-reviewed primary research studies, cited throughout the text and listed in the References section, on blood-based biomarkers for the early diagnosis of pediatric acute lymphoblastic leukemia (ALL). Relevant studies were identified through searches of PubMed and Google Scholar using combinations of keywords related to pediatric acute lymphoblastic leukemia, blood-based biomarkers, liquid biopsy, early diagnosis, microRNAs, proteins, long non-coding RNAs, metabolomics, and artificial intelligence. Studies were selected for their relevance to blood-based biomarkers for the early diagnosis of pediatric ALL; those focused primarily on prognosis, treatment response, or bone marrow biomarkers were excluded. Studies were organized thematically by biomarker class—including protein and glycoprotein markers, microRNAs, long non-coding RNAs, metabolomic and lipidomic signatures, artificial intelligence-assisted diagnostic models, and autoantibodies—to facilitate comparison of diagnostic performance, methodological strengths, and limitations across categories. For each study, reported findings, sample characteristics, and diagnostic metrics, including sensitivity, specificity, and area under the curve, were compared to identify recurring biomarkers, methodological gaps, and directions for future research. No predefined systematic search protocol, database restrictions, publication date range, or formal quality assessment framework was applied.

2. BLOOD-BASED BIOMARKERS FOR EARLY DIAGNOSIS

2.1. Protein and Glycoprotein Biomarkers

Proteomic profiling of serum has identified several protein signatures that differentiate ALL patients from healthy controls. Using Surface-Enhanced Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (SELDI-TOF-MS), one study developed a 7-protein peak model that achieved high sensitivity (91.8%) and specificity (90.0%) in distinguishing pediatric ALL from Acute Myeloid Leukemia (AML) and healthy children (Shi et al., 2009). In that model, fragments of Complement Component 3a (C3a) were upregulated, while Platelet Factor 4 (PF4) and Connective Tissue Activating Peptide III (CTAP-III) were downregulated (Shi et al., 2009).

Other investigations using iTRAQ-based proteomics (a technique that uses chemical tags to simultaneously measure and compare protein levels across multiple samples) have corroborated the upregulation of Leucine-rich alpha-2-glycoprotein 1 (LRG1) and identified S100A8 as a strong diagnostic candidate (Cavalcante et al., 2016; Yu et al., 2019). Soluble L-selectin (sL-selectin), meanwhile, appears to have lineage-specific utility, as it is significantly higher in T-ALL than in B-ALL (Yu et al., 2019).

Critical Summary: Protein panels show high sensitivity, making them promising tools for initial triage. Yet many of these proteins, such as C3 and LRG1, are also markers of general inflammation, which may limit their specificity when used individually to diagnose ALL in children with intercurrent infections. The collective evidence suggests that protein panels outperform single markers but still require validation in cohorts with non-leukemic inflammatory diseases. These limitations have encouraged researchers to pursue more disease-specific molecular biomarkers, particularly circulating non-coding RNAs, which may offer greater biological specificity while remaining detectable in peripheral blood.

2.2. MicroRNAs (miRNAs)

MicroRNAs are stable, small non-coding RNAs that regulate gene expression and show significant dysregulation in leukemic environments. Circulating miR-146a has emerged as a potentially robust biomarker: one Pakistani study reported nearly perfect discrimination between ALL cases and healthy controls (Shahid et al., 2021). miR-31 and miR-493-3p are likewise significantly downregulated in the peripheral blood of newly diagnosed pediatric ALL patients, and these low expression levels often correlate with higher white blood cell counts and increased disease risk (Yang et al., 2025; Zhang et al., 2019).

Critical Summary: miRNAs are among the most statistically impressive individual biomarkers identified. The literature is not entirely consistent, though: miR-146a is upregulated in plasma, whereas older studies on bone marrow samples reported its downregulation, indicating that the choice of biofluid significantly affects biomarker interpretation (Shahid et al., 2021). Near-perfect diagnostic results from single-center studies should also be viewed with caution until they are replicated in larger, multicenter trials.

2.3. Long Non-Coding RNAs (lncRNAs)

LncRNAs are essential modulators of oncogenesis and exhibit lineage-specific expression profiles. Three of them—LINC01221, RP11-472G21.2, and CRNDE—are significantly overexpressed in pediatric T-ALL compared with B-ALL and healthy controls (Sharma et al., 2024). In B-ALL, circulating HOTAIR has been proposed as a diagnostic biomarker; notably, its levels in peripheral blood plasma closely match those in bone marrow, supporting the feasibility of using blood as a non-invasive surrogate (Rahimi Dashti et al., 2025).

Critical Summary: LncRNAs offer better subtype specificity (e.g., T-lineage vs. B-lineage) than many protein markers. Markers like HOTAIR, though, currently show lower diagnostic sensitivity (~48%) than miRNAs do, suggesting they may be more useful for subtype differentiation once a primary diagnosis is already suspected (Rahimi Dashti et al., 2025).

2.4. Metabolomics and Lipidomics

Metabolic reprogramming in leukemic cells creates a unique systemic "fingerprint." 1H NMR-based metabolomics (which uses proton nuclear magnetic resonance spectroscopy to measure small molecules, or metabolites, in biological fluids), for instance, has identified a "triad" of choline, tyrosine, and unsaturated lipids that distinguishes ALL patients from healthy children with high accuracy (Serrafi et al., 2026). Patients with ALL typically show elevated formate and citrate, linked to the Warburg effect, alongside decreased glutamine and tyrosine (Serrafi et al., 2026).

Lipidomic analysis (the comprehensive study of all lipids, or fat-based molecules, in a biological sample) has further identified significant elevations in serum sphingolipids—specifically Cer 18:0 and Cer 20:0—that correlate strongly with white blood cell counts (Yan et al., 2025). These shifts may be influenced by gut microbial dysbiosis, which would provide a broader biological context for leukemogenesis (Yan et al., 2025).

Critical Summary: Metabolomics offers a functional "signature" of the disease, and the discovery of the choline-tyrosine triad is a significant strength. The primary 1H NMR pilot study, however, was hindered by a substantial age gap between patients and controls, a known confounding factor in pediatric metabolic research (Serrafi et al., 2026).

2.5. AI-Enhanced Hematological and Spectroscopy Models

Advanced machine learning models leverage routine blood data to improve early screening. One such model, a 10-index Random Forest classifier incorporating albumin, LDH, hemoglobin, and white blood cell parameters, achieved a high AUC of 0.950 for ALL in a large Chinese cohort of 1,281 children (Cheng et al., 2024).

Fourier-transform infrared (FTIR) spectroscopy offers a similar advantage, detecting biochemical shifts in serum proteins and lipids within minutes. When combined with machine learning, this low-cost approach achieved moderate accuracy in distinguishing ALL from controls, including children with non-leukemic cytopenias (Kowal et al., 2025).

Critical Summary: The AI model demonstrates the strongest statistical robustness among the reviewed studies, largely because of its substantially larger validation cohort and its reliance on routinely collected hematological indices. Unlike many molecular biomarkers that remain in the discovery phase, AI-based approaches could integrate more readily into existing clinical workflows. External validation across diverse populations is still required, though, before these models can be considered for widespread clinical implementation. These findings point to a broader shift from evaluating individual biomarkers toward integrating multiple diagnostic indicators—a theme explored further in the comparative analysis that follows.

2.6. Autoantibodies

Children with B-ALL produce autoantibodies against tumor-associated antigens (TAAs) such as α-enolase and VDAC1, and these antibodies appear early in the oncogenic process, often before the antigens themselves become detectable (Yu et al., 2022).

Critical Summary: Autoantibodies represent a promising "early warning" system, but they currently suffer from notably low sensitivity (~23–27%), meaning they cannot serve as standalone screening tools (Yu et al., 2022).

Taken together, these findings demonstrate that no single biomarker category consistently fulfills all the characteristics of an ideal diagnostic test. Each biomarker class offers distinct biological and clinical advantages, but their reported diagnostic performance varies according to study design, patient population, and analytical methodology. Evaluating these biomarkers in isolation therefore provides only a partial picture of their clinical potential. The following section synthesizes the available evidence to compare their relative strengths, limitations, and readiness for clinical application.

3. COMPARISON AND SYNTHESIS

Table 1 summarizes the diagnostic performance, strengths, and limitations of each biomarker class discussed in this review.

Table 1. Comparison of blood-based biomarker classes for the early diagnosis of pediatric acute lymphoblastic leukemia.

Biomarker Class	Representative Markers	Best Reported Performance	Key Strength	Key Limitation
Proteins / glycoproteins	C3a fragments, PF4, CTAP-III, LRG1, S100A8, sL-selectin	91.8% sensitivity, 90.0% specificity (7-protein SELDI-TOF-MS model) (Shi et al., 2009)	High sensitivity in multi-marker panels (Cavalcante et al., 2016; Shi et al., 2009; Yu et al., 2019)	Overlap with general inflammatory markers limits specificity (Cavalcante et al., 2016; Yu et al., 2019)
MicroRNAs	miR-146a, miR-31, miR-493-3p	Up to 100% sensitivity, single-center cohort (Shahid et al., 2021)	Strongest single-marker accuracy reported (Shahid et al., 2021; Yang et al., 2025; Zhang et al., 2019)	Findings vary by biofluid and population (Shahid et al., 2021)
Long non-coding RNAs	LINC01221, RP11-472G21.2, CRNDE, HOTAIR	~48% sensitivity (HOTAIR) (Rahimi Dashti et al., 2025)	Lineage-specific expression distinguishes T-ALL from B-ALL (Rahimi Dashti et al., 2025; Sharma et al., 2024)	Lower sensitivity than miRNAs (Rahimi Dashti et al., 2025)
Metabolomics / lipidomics	Choline-tyrosine-lipid triad, Cer 18:0, Cer 20:0	AUC > 0.95 (choline-tyrosine-lipid triad) (Serrafi et al., 2026)	High-accuracy composite metabolic signature (Serrafi et al., 2026; Yan et al., 2025)	Small pilot cohorts; age mismatch between groups (Serrafi et al., 2026)
AI-assisted models	10-index Random Forest classifier; FTIR + machine learning	AUC 0.950, n = 1,281 (Cheng et al., 2024)	Largest validation cohort; uses routine, low-cost data (Cheng et al., 2024; Kowal et al., 2025)	Requires external validation across diverse populations (Cheng et al., 2024)

Biomarker Class	Representative Markers	Best Reported Performance	Key Strength	Key Limitation
Autoantibodies	Anti-α-enolase, anti-VDAC1	~23–27% sensitivity (Yu et al., 2022)	Near-perfect specificity; detectable early in oncogenesis (Yu et al., 2022)	Low sensitivity precludes standalone use (Yu et al., 2022)

A comparative analysis of the identified biomarkers reveals that although multiple modalities show statistical promise, they vary considerably in diagnostic performance, biological specificity, and readiness for clinical implementation. When the strongest evidence for early diagnosis is weighed, a clear hierarchy emerges based on the statistical robustness of the findings and the size of the validation cohorts. Circulating microRNAs and metabolomic "triads" currently report the highest individual diagnostic accuracies: studies on miR-146a and the 1H NMR-derived choline-tyrosine-lipid triad both yielded Area Under the Curve (AUC) values exceeding 0.95 (Serrafi et al., 2026; Shahid et al., 2021). These statistically impressive results, though, must be weighed against the relatively small, single-center populations from which they were derived. The Random Forest model developed by Cheng et al. used a substantial cohort of 1,281 children to achieve an AUC of 0.950 (Cheng et al., 2024). This suggests that patterns within existing, high-volume clinical data may translate more readily into clinical practice, as they rely on routinely collected laboratory data rather than specialized molecular assays.

The discrepancy between biomarker categories often stems from their underlying biological relevance and sensitivity to confounding factors. Protein biomarkers like Complement C3, Leucine-rich α-2-glycoprotein 1 (LRG1), and S100A8, for example, are frequently identified as significantly upregulated at diagnosis (Cavalcante et al., 2016; Yu et al., 2019), but these molecules are also deeply involved in general systemic inflammation and the acute phase response. This involvement limits their specificity relative to non-coding RNAs such as the lncRNAs LINC01221 and CRNDE, which exhibit lineage-specific expression capable of distinguishing T-ALL from B-ALL and healthy states (Sharma et al., 2024). Serum autoantibodies against α-enolase and VDAC1 present the opposite trade-off: they offer near-perfect specificity, but their clinical utility is severely hampered by low sensitivity, as they were detectable in fewer than 30% of B-ALL patients in early-stage trials (Yu et al., 2022).

Methodological differences across these studies also create challenges for direct synthesis. miR-146a, for instance, showed 100% sensitivity in a Pakistani cohort (Shahid et al., 2021), yet older studies on bone marrow reported its downregulation—a discrepancy that highlights how the choice of biofluid (plasma versus bone marrow) and ethnic population can shape findings. In metabolomics, the high accuracy of the choline-tyrosine triad is promising, but the primary study was limited by a significant age gap between patients and controls, a known confounding factor in pediatric metabolic profiles (Serrafi et al., 2026). These inconsistencies suggest that many current findings reflect population-specific or age-specific "fingerprints" rather than universal markers of leukemogenesis.

The most significant trend across the literature is the shift from single-analyte discovery to multi-marker diagnostic panels. Replicated findings, such as the consistent identification of LRG1 across separate proteomic and glycoprotein studies (Cavalcante et al., 2016; Yu et al., 2019), reinforce the value of established signatures. Multi-marker models—such as the 7-peak proteomic model or the 10-index AI model (Cheng et al., 2024; Shi et al., 2009)—consistently outperform individual markers by capturing the multifaceted systemic response to malignancy. These panels also provide the redundancy needed to maintain accuracy across clinically heterogeneous patient groups.

In summary, the current body of evidence provides strong preliminary support for the diagnostic potential of several blood-based biomarker signatures, but the field still lacks the large-scale, prospective validation needed before these biomarkers can supplement or replace bone marrow aspiration in routine clinical practice. The continued reliance on single-center cohorts, together with limited evaluation of biomarkers against clinically relevant "mimicker" diseases—such as benign infections that cause cytopenia—remains a major barrier to clinical translation. Overall, these limitations underscore the need for more rigorous validation and methodological standardization before blood-based biomarkers can be adopted as reliable diagnostic tools.

4. LIMITATIONS OF THE CURRENT LITERATURE

The findings summarized in this review are encouraging, but the current body of literature is primarily composed of discovery-phase studies with several systemic limitations that affect the reliability and clinical applicability of the evidence. A significant portion of the research, particularly in proteomics and metabolomics, relies on very small sample sizes, often involving cohorts of fewer than 40 participants (Cavalcante et al., 2016; Rahimi Dashti et al., 2025; Sharma et al., 2024). Such limited cohorts increase the risk of over-optimistic results and statistical overfitting, which reduces the reliability of the findings. As a result, many identified signatures—ranging from specific ceramide profiles (Yan et al., 2025) to lncRNA expressions (Rahimi Dashti et al., 2025; Sharma et al., 2024)—must be considered preliminary until they are confirmed in larger, multi-center prospective trials.

The reliance on single-center study designs within localized geographic regions—such as single institutions in China (Shi et al., 2009), Pakistan (Shahid et al., 2021), Poland (Kowal et al., 2025; Serrafi et al., 2026), or Greece (Papadopoulou et al., 2024)—also limits the generalizability of these biomarkers across diverse ethnic and socioeconomic populations.

Clinical utility is further obscured by a frequent lack of comparison against relevant "mimicker" diseases. Most studies evaluate biomarkers by comparing ALL patients to healthy children, yet a real-world screening tool must be able to differentiate leukemia from benign pediatric conditions, such as severe infections or iron-deficiency anemia, that also cause cytopenia. Because many identified biomarkers—particularly inflammatory proteins like Complement C3 and LRG1 (Cavalcante et al., 2016; Yu et al., 2019)—are also involved in the general acute-phase response, omitting disease-control groups significantly reduces confidence in their leukemia-specific diagnostic performance.

Methodological inconsistencies and confounding variables further hinder the translation of these findings into routine clinical practice. Age is a particularly critical variable in pediatric research, and significant age disparities between study groups—as noted in the 1H NMR metabolic pilot studies (Serrafi et al., 2026)—can introduce biological bias that statistical adjustments cannot fully resolve. Variations in laboratory techniques, ranging from label-free FTIR spectroscopy to complex iTRAQ proteomics (Kowal et al., 2025; Yu et al., 2019), together with differences in biofluid selection (serum versus plasma), also make direct comparison and replication across clinical settings considerably more challenging. Taken together, the current evidence provides strong preliminary support for the diagnostic potential of several blood-based biomarker signatures. However, substantial methodological standardization and large-scale external validation are still required before these approaches can supplement or replace bone marrow aspiration in routine clinical practice.

Table 2 summarizes these limitations alongside corresponding priorities for future research.

Table 2. Key methodological limitations of the current literature and corresponding priorities for future research.

Limitation	Description	Future Research Priority
Small sample sizes	Many proteomic and metabolomic studies involved fewer than 40 participants (Cavalcante et al., 2016; Rahimi Dashti et al., 2025; Sharma et al., 2024)	Large-scale, multicenter prospective validation
Single-center, narrow geography	Studies drawn largely from single institutions in China, Pakistan, Poland, or Greece (Kowal et al., 2025; Papadopoulou et al., 2024; Serrafi et al., 2026; Shahid et al., 2021; Shi et al., 2009)	Multiethnic, multi-socioeconomic validation cohorts
Lack of disease-control groups	Most studies compare ALL patients only to healthy children rather than to mimicker conditions such as infections or iron-deficiency anemia	Inclusion of clinically relevant disease-control groups, as in the FTIR study (Kowal et al., 2025)
Age and biofluid confounds	Significant age disparities and inconsistent biofluid selection (serum vs. plasma) across studies (Kowal et al., 2025; Serrafi et al., 2026; Yu et al., 2019)	Standardized, age-matched cohorts and biofluid protocols
Single-analyte focus	Individual biomarkers often lack sufficient standalone sensitivity or specificity	Integration of multi-omic data into comprehensive diagnostic models (Cheng et al., 2024; Shi et al., 2009)

5. RESEARCH GAPS AND FUTURE DIRECTIONS

Numerous blood-based biomarkers have demonstrated encouraging diagnostic potential, yet the field remains largely in the discovery phase, leaving a substantial gap between biomarker identification and routine clinical implementation. The most critical research gap is the lack of large-scale, multicenter prospective validation in diverse ethnic and socioeconomic populations. Most existing literature relies on small, single-institution cohorts, which increases the risk of statistical overfitting and results in over-optimistic diagnostic performance that may not translate to broader populations. Standardizing laboratory protocols, biofluid selection (serum versus plasma), and measurement thresholds across clinical settings will be essential to ensure that molecular signatures—ranging from specific ceramide profiles (Yan et al., 2025) to miRNA expressions (Shahid et al., 2021; Yang et al., 2025; Zhang et al., 2019)—are universally reproducible.

A vital unanswered question concerns biomarker specificity in the context of common pediatric conditions that mimic leukemia symptoms. Many studies achieve near-perfect accuracy by comparing ALL patients to healthy controls, but a functional screening tool must be able to differentiate malignancy from benign conditions—such as severe infections or iron-deficiency anemia—that also cause cytopenia or elevated white cell counts. Future research should therefore prioritize the inclusion of clinically relevant disease-control groups, as seen in the FTIR spectroscopy study (Kowal et al., 2025), to verify that identified markers are truly leukemia-specific rather than generalized indicators of systemic inflammation or the acute-phase response.

The literature also suggests that the future of pediatric ALL diagnosis lies in moving from single-analyte discovery toward the validation of integrated multi-marker panels. Multi-index models, such as the 10-index AI model and the 7-peak proteomic signature (Cheng et al., 2024; Shi et al., 2009), consistently capture the multifaceted biological response to leukemia more effectively than individual biomarkers alone. The most promising direction for future research, then, appears to be the integration of multi-omic data—including genomic, transcriptomic, proteomic, and metabolomic information—to generate comprehensive diagnostic models. Such approaches may better reflect the biological complexity of pediatric ALL while improving diagnostic accuracy and resilience to biological variability, though their clinical value must still be confirmed through standardized multicenter prospective studies before they can be incorporated into routine clinical practice.

6. CONCLUSION

The current body of evidence suggests that blood-based biomarkers represent a promising complement to existing diagnostic methods for the early diagnosis of pediatric ALL. Numerous circulating biomarkers have demonstrated diagnostic potential, but the most consistent evidence supports circulating RNA biomarkers and integrated multi-marker diagnostic approaches. Individual biomarkers such as miR-146a and metabolomic signatures (Serrafi et al., 2026; Shahid et al., 2021) have shown strong diagnostic performance, yet AI-based models trained on routine hematological data may offer greater potential for clinical translation, largely because they have been evaluated in substantially larger patient cohorts (Cheng et al., 2024).

Despite these encouraging findings, important barriers remain before these biomarkers can be incorporated into routine clinical practice. Much of the existing evidence is derived from small, single-center studies with limited external validation and susceptibility to confounding factors such as patient age. Individual biomarkers also often lack sufficient sensitivity or specificity when used alone, reinforcing the need for integrated multi-marker panels.

The development of reliable blood-based biomarkers could facilitate earlier identification of children at risk for ALL while reducing the need for invasive diagnostic procedures. The evidence reviewed in this paper suggests that integrated multi-marker strategies are more likely than individual biomarkers alone to achieve clinically meaningful diagnostic performance. Widespread clinical implementation, however, will require standardized multicenter validation, evaluation in diverse populations and clinically relevant disease-control groups, and continued efforts to develop robust diagnostic models capable of distinguishing ALL from benign pediatric conditions.

DATA AVAILABILITY STATEMENT

This is a literature review; no new data were generated. All studies discussed are publicly available via the DOIs provided in the References section.

AUTHOR CONTRIBUTIONS

Zayedan Ahmed Tayyib conceived the review, conducted the literature search and study selection, analyzed and synthesized the included literature, and wrote and revised the manuscript.

ACKNOWLEDGMENTS

The author has no acknowledgments to declare.

CONFLICTS OF INTEREST

The author declares no competing interests.

REFERENCES

Cavalcante, M. S., Torres-Romero, J. C., Lobo, M. D. P., Moreno, F. B. M. B., Bezerra, L. P., Lima, D. S., Matos, J. C., Moreira, R. A., & Monteiro-Moreira, A. C. O. (2016). A panel of glycoproteins as candidate biomarkers for early diagnosis and treatment evaluation of B-cell acute lymphoblastic leukemia. Biomarker Research, 4, Article 1. https://doi.org/10.1186/s40364-016-0055-6

Cheng, Z. J., Li, H., Liu, M., Fu, X., Liu, L., Liang, Z., Gan, H., & Sun, B. (2024). Artificial intelligence reveals the predictions of hematological indexes in children with acute leukemia. BMC Cancer, 24, Article 993. https://doi.org/10.1186/s12885-024-12646-3

Kowal, A., Jakubczyk, P., Bal, W., Piasecka, Z., Szuler, K., Łach, K., Sopel, K., Cebulski, J., & Chaber, R. (2025). Serum Fourier-transform infrared spectroscopy with machine learning for screening of pediatric acute lymphoblastic leukemia: A proof-of-concept study. Cancers, 17(21), Article 3548. https://doi.org/10.3390/cancers17213548

Papadopoulou, M. T., Panagopoulou, P., Paramera, E., Pechlivanis, A., Virgiliou, C., Papakonstantinou, E., Palabougiouki, M., Ioannidou, M., Vasileiou, E., Tragiannidis, A., Papakonstantinou, E., Theodoridis, G., Hatzipantelis, E., & Evangeliou, A. (2024). Metabolic fingerprint in childhood acute lymphoblastic leukemia. Diagnostics, 14(7), Article 682. https://doi.org/10.3390/diagnostics14070682

Rahimi Dashti, N., Fadavi, D., Rezaei, R., Rahgozar, S., & Moafi, A. (2025). Circulating lncRNA HOTAIR is a biomarker for pediatric acute lymphoblastic leukemia and mediator of miR-326 exosomal export. Scientific Reports, 15, Article 4901. https://doi.org/10.1038/s41598-025-87857-0

Serrafi, A., Pupek, M., Lewandowski, Ł., Janicka-Kłos, A., Wasilewski, A., Kasprzak, A., Matera-Witkiewicz, A., Zatoński, T., Połtyn-Zaradna, K., ściskalska, M., Brutkowski, T., & Kazanowska, B. (2026). 1H NMR-based metabolomics in pediatric acute lymphoblastic leukemia: A pilot study of plasma and cerebrospinal fluid profiles. Metabolites, 16(3), Article 160. https://doi.org/10.3390/metabo16030160

Shahid, S., Shahid, W., Shaheen, J., Akhtar, M. W., & Sadaf, S. (2021). Circulating miR-146a expression as a non-invasive predictive biomarker for acute lymphoblastic leukemia. Scientific Reports, 11, Article 22783. https://doi.org/10.1038/s41598-021-02257-4

Sharma, P., Kaur, P., Bhatia, P., Trehan, A., Sreedharanunni, S., & Singh, M. (2024). Novel lncRNAs LINC01221, RP11-472G21.2 and CRNDE are markers of differential expression in pediatric patients with T cell acute lymphoblastic leukemia. Cancer Cell International, 24, Article 65. https://doi.org/10.1186/s12935-024-03255-y

Shi, L., Zhang, J., Wu, P., Feng, K., Li, J., Xie, Z., Xue, P., Cai, T., Cui, Z., Chen, X., Hou, J., Zhang, J., & Yang, F. (2009). Discovery and identification of potential biomarkers of pediatric acute lymphoblastic leukemia. Proteome Science, 7, Article 7. https://doi.org/10.1186/1477-5956-7-7

Yan, F., Wang, S., Wang, Y., Sun, Y., Yang, J., Sun, L., Zaytseva, Y. Y., Deng, P., & Wang, L. (2025). LC-MS analysis of serum lipidomic and metabolomic signatures in pediatric patients with acute lymphoblastic leukemia. Italian Journal of Pediatrics, 51, Article 74. https://doi.org/10.1186/s13052-025-01921-z

Yang, H., Xu, Y., Zhang, X., Wang, Y., & Li, Z. (2025). Serum miR-493-3p as a diagnostic biomarker and epigenetic regulator targeting DPY30 in pediatric acute lymphoblastic leukemia. Annals of Hematology, 104, 6285–6296. https://doi.org/10.1007/s00277-025-06681-8

Yu, R., Zhang, J., Zang, Y., Zeng, L., Zuo, W., Bai, Y., Liu, Y., Sun, K., & Liu, Y. (2019). iTRAQ-based quantitative protein expression profiling of biomarkers in childhood B-cell and T-cell acute lymphoblastic leukemia. Cancer Management and Research, 11, 7047–7063. https://doi.org/10.2147/CMAR.S210093

Yu, R., Yang, S., Liu, Y., & Zhu, Z. (2022). Identification and validation of serum autoantibodies in children with B-cell acute lymphoblastic leukemia by serological proteome analysis. Proteome Science, 20, Article 3. https://doi.org/10.1186/s12953-021-00184-w

Zhang, Y., Li, X., Bai, L., Li, L., Li, D., Ding, X., Wang, B., & Li, C. (2019). MicroRNA-31 is a potential biomarker for screening B-lymphoblastic leukemia in children. Oncology Letters, 18(5), 4930–4935. https://doi.org/10.3892/ol.2019.10843