Uncovering Surgical Trajectories in Pediatric Lower Extremity Care Using NLP on Clinical Notes

Nazgol Tavabi, Nishita Kalepalli, Timothy Williamson, Ata M. Kiapour

Department of Orthopedic Surgery and Sports Medicine, Boston Children's Hospital and Harvard Medical School, Boston, MA
Nazgol.Tavabi@childrens.harvard.edu

Disclosures: N. Tavabi: 4; BonePixel. N. Kalepalli: None. T. Williamson: None. Kiapour: 3B; MIACH Orthopaedics, 4; BonePixel, 8; BMC Musculoskeletal Disorders and American Journal of Sports Medicine.

INTRODUCTION: Lower extremity (LE) procedures constitute the majority of pediatric orthopedic surgeries, yet the long-term surgical trajectories of these patients across anatomical regions remain poorly quantified. Understanding repeat surgical burden, age- and sex-specific risk, and cross-specialty progression is critical for surgical planning and family counseling. However, much of the rich clinical information needed to study these patterns is captured only in unstructured operative notes, making it difficult to extract and analyze at scale. Even advanced EMR tools (e.g., EPIC SlicerDicer) cannot adequately process these data in busy, high-volume pediatric orthopedic settings. To address this gap, we applied a robust NLP-based pipeline to convert operative notes into structured, searchable datasets, enabling systematic characterization of surgical trajectories.

METHODS: We retrospectively analyzed all operative notes from a large pediatric hospital (2000–2025). Natural language processing identified LE orthopedic surgeries, which were further grouped by anatomical site (hip, knee, thigh, lower leg, ankle, foot). Patients were tracked longitudinally to capture subsequent orthopedic procedures across subspecialties (upper extremity, spine, oncology, other). Group differences were evaluated using chi-squared tests for categorical variables and t-tests for continuous variables.

RESULTS: We identified 100,438 LE surgeries from 56,606 patients (mean age 13.7 ± 7.3 years; 51% female). 22,028 patients (39%) underwent ≥1 subsequent orthopedic procedure, accounting for 59,094 follow-up surgeries. Patients with repeat surgeries were younger at index procedure (12.4 vs 14.6 years, $p<0.001$) but showed no sex difference in reoperation likelihood (F 40% vs M 38%, $p=0.07$). The most common index sites were knee (31%), hip (21%), and foot (9%). Among those with follow-up, hip (26%) and knee (23%) remained the leading sites, while thigh (8%) and ankle (4%) were proportionally enriched compared to single-surgery patients ($p<0.001$ for distribution).Across the cohort, 55.8% of repeat surgeries remained LE-focused, while 14.5% involved other regions: spine (3,201 cases), upper extremity (2,318), and orthopedic oncology (1,879). Spine surgeries were disproportionately common among repeat patients compared with the overall cohort ($p<0.001$). Follow-up analysis showed a median time to second surgery of 8.6 months. Surgical burden was highly skewed: 56% had only 1 additional procedure, while 6% had ≥5, and the maximum observed was 61 surgeries in a single patient. The distribution of high-burden patients differed by index anatomy, with hip and thigh surgeries showing the highest cumulative counts ($p<0.001$).

DISCUSSION: Nearly two in five pediatric patients undergoing LE surgery required additional orthopedic procedures, often involving other anatomical regions. Younger age at index surgery strongly predicted repeat intervention. While most follow-up operations remained LE-focused, the enrichment of spine and oncology procedures suggests syndromic and systemic conditions drive many trajectories. The small subgroup of "high-burden" patients represent a disproportionate surgical load and merits focused perioperative planning and multidisciplinary care.

SIGNIFICANCE: This 25-year, hospital-wide analysis establishes benchmarks for repeat surgical risk following pediatric lower extremity procedures. By leveraging a robust NLP-based pipeline to extract structured data from unstructured operative notes, we systematically capture detailed patient trajectories across anatomy, age, sex, and cross-specialty involvement. These analyses provide actionable insights for prognosis, surgical planning, family counseling, and care coordination, while demonstrating the potential of NLP to unlock large-scale clinical datasets that are otherwise difficult to analyze in high-volume pediatric orthopedic practice.

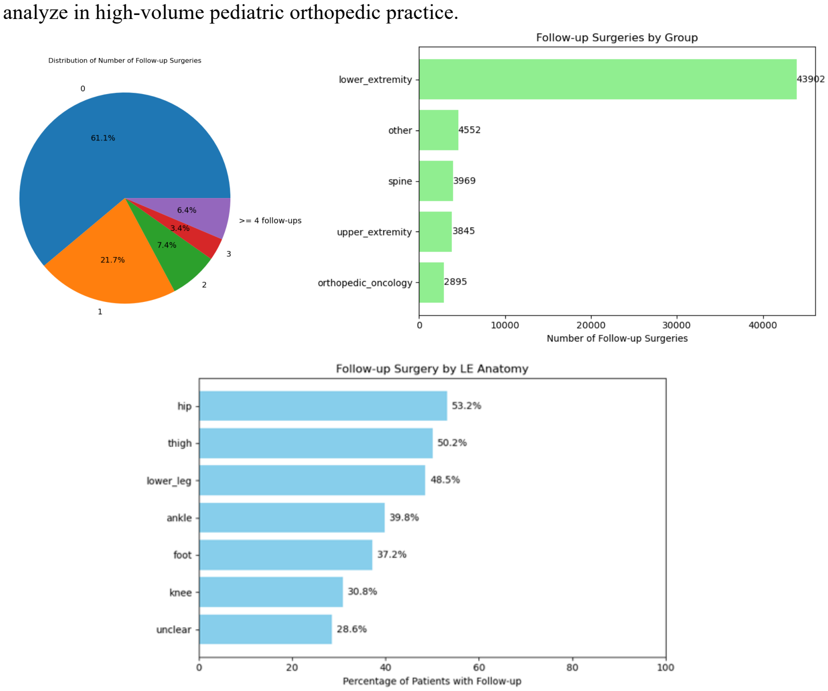

Figure 1: (A) Distribution of follow-up surgeries per patient. **(B)** count of type of follow-up surgeries, **(C)** percentage of follow-ups for LE anatomy