



OPEN ACCESS

EDITED BY

Farah Kidwai-Khan,
Yale University, United States

REVIEWED BY

Harishchander Anandaram,
Amrita Vishwa Vidyapeetham University, India

*CORRESPONDENCE

Esteban Zavaleta-Monestel
✉ ezavaleta@clinicabiblica.com

RECEIVED 03 October 2025

ACCEPTED 31 October 2025

PUBLISHED 13 November 2025

CITATION

Zavaleta-Monestel E, Arguedas-Chacón S and Cordero-Bermúdez K (2025) Artificial intelligence for surgical data management and decision support: lessons from wound care. *Front. Artif. Intell.* 8:1718436. doi: 10.3389/frai.2025.1718436

COPYRIGHT

© 2025 Zavaleta-Monestel, Arguedas-Chacón and Cordero-Bermúdez. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Artificial intelligence for surgical data management and decision support: lessons from wound care

Esteban Zavaleta-Monestel^{1*}, Sebastián Arguedas-Chacón¹ and Katherine Cordero-Bermúdez^{2,3}

¹Health Research Department, Hospital Clínica Bíblica, San José, Costa Rica, ²Department of General Surgery, Rafael Ángel Calderón Guardia Hospital, Costa Rican Social Security Fund (CCSS), San José, Costa Rica, ³Hernia and Abdominal Wall Rehabilitation Clinic, Clínica Bíblica Hospital, San José, Costa Rica

KEYWORDS

natural language processing, artificial intelligence, clinical decision-making, wound healing, outcome assessment, health care

Introduction

The surge in surgical data has made converting raw documentation, operative notes, pathology reports, laboratory results, and outcomes registries into actionable clinical insights a considerable challenge for clinical care and research. Though these sources hold valuable information, manual chart review is slow, inconsistent, and unsustainable in high-volume environments (Murff et al., 2011; Mellia et al., 2021). Surgeons now dedicate a significant portion of their workday to documentation instead of direct patient care. A time-and-motion study observed that nearly half of physicians' working hours are consumed by electronic health record (EHR) tasks (Sinsky et al., 2016). These findings underscore the need for innovative solutions that enhance documentation without compromising care quality (Joukes et al., 2018).

Intelligent data-processing methods are beginning to bridge this gap. For example, natural language processing (NLP), which allows computers to interpret and generate human language, has been used in medicine for over a decade. Early research showed stronger detection of postoperative complications compared to billing-code reviews (Murff et al., 2011). More recently, large language models (LLMs), such as ChatGPT, have demonstrated the ability to rapidly and accurately analyze narrative clinical text (Chen et al., 2022; Dagli et al., 2024; Friedman et al., 2025).

AI applications go beyond text interpretation. Similarly, advances in image recognition, intraoperative navigation, and robotic assistance are evolving, positioning structured data management as pivotal in surgery's digital transformation (Beyaz et al., 2025). Collectively, these breakthroughs foreshadow a future where automated information extraction aids perioperative decision-making and redesigns surgical workflows (Han et al., 2025).

Evidence from chart review

For over a decade, NLP has proven valuable in surgical research. Automated text analysis has outperformed billing-code reviews in detecting postoperative complications, achieving higher sensitivity without loss of specificity (Murff et al., 2011). A meta-analysis confirmed that NLP detects complications more accurately than manual methods while maintaining comparable specificity (Mellia et al., 2021).

Recent work has evolved from feasibility to refinement. For example, a data-extraction pipeline developed for breast cancer reports identified 48 outcome variables with near-human accuracy, achieving *F*-scores above 0.90 for most measures. *F*-score is a statistic that combines sensitivity and precision to measure accuracy (Chen et al., 2022). In another instance, a framework that combined NLP and LLM integration for spinal surgery attained near-perfect sensitivity for key operative variables, reduced review time by more than 3,000-fold, and dramatically lowered costs (Dagli et al., 2024).

Taken together, these studies indicate that intelligent data-processing tools are evolving from experimental demonstrations to realistic clinical applications. However, routine implementation across hospitals remains uncommon and will require further validation in real-world settings.

Wound care as a case study

Plastic surgery and wound care in particular present unique documentation challenges due to their reliance on narrative detail. In a recent study, employing an LLM reduced the average chart review time from 7.56 to 1.03 min per case, all while maintaining an overall accuracy of 95.7% (ranging from 74.7% to 98.6% across variables). Furthermore, the model generated wound summaries that closely echoed clinician notes (Friedman et al., 2025).

Furthermore, these findings are consistent with results from other surgical areas, including breast cancer and spine surgery, indicating that wound care exemplifies a broader trend toward data-driven surgical documentation (Chen et al., 2022; Dagli et al., 2024). Wound care is particularly illustrative because it integrates diverse data types, involves multidisciplinary teams, and addresses both functional and aesthetic outcomes.

This complexity matches challenges in other plastic surgery subspecialties, such as breast reconstruction. Here, clinical data are unstructured, and outcomes include quality of life, cosmetic satisfaction, and survival (Spoer et al., 2022; Rugină et al., 2025). Recent reviews show wound care is an ideal area for machine learning and language-based systems. Applications range from wound assessment and prognosis to treatment personalization and outcome prediction (Ganesan et al., 2024).

As a result, wound care offers a practical model for evaluating the implementation of intelligent data tools prior to broader adoption. Automation in this context can improve efficiency without compromising accuracy, even in domains characterized by free-text narratives (Friedman et al., 2025).

Toward decision support

Improving efficiency is only the first step; the real promise of intelligent systems lies in their ability to support informed decision-making in surgery. Beyond documentation, several experimental tools are now being tested for intraoperative use, including platforms that interpret real-time imaging and assist robotic procedures (Byrd IV and Tignanelli, 2024; Matheny et al., 2020). These applications remain in early

development and require validation and regulatory review before being used in routine clinical settings (Liang et al., 2025).

For instance, one NLP-based system accurately predicted unplanned intensive care admissions in elective neurosurgical patients by analyzing free-text clinical notes. Another machine learning approach, applied to ventral hernia repairs, identified recurrence, surgical site occurrence, and 30-day readmission rates using preoperative EMR data (Hassan et al., 2022).

Together, these advances indicate a transition from retrospective data review to real-time risk assessment. This shift demonstrates that digital and learning-based systems can evolve from record-keeping to clinical decision support, enabling surgeons to anticipate complications rather than only documenting them (Ive et al., 2025).

Challenges and governance

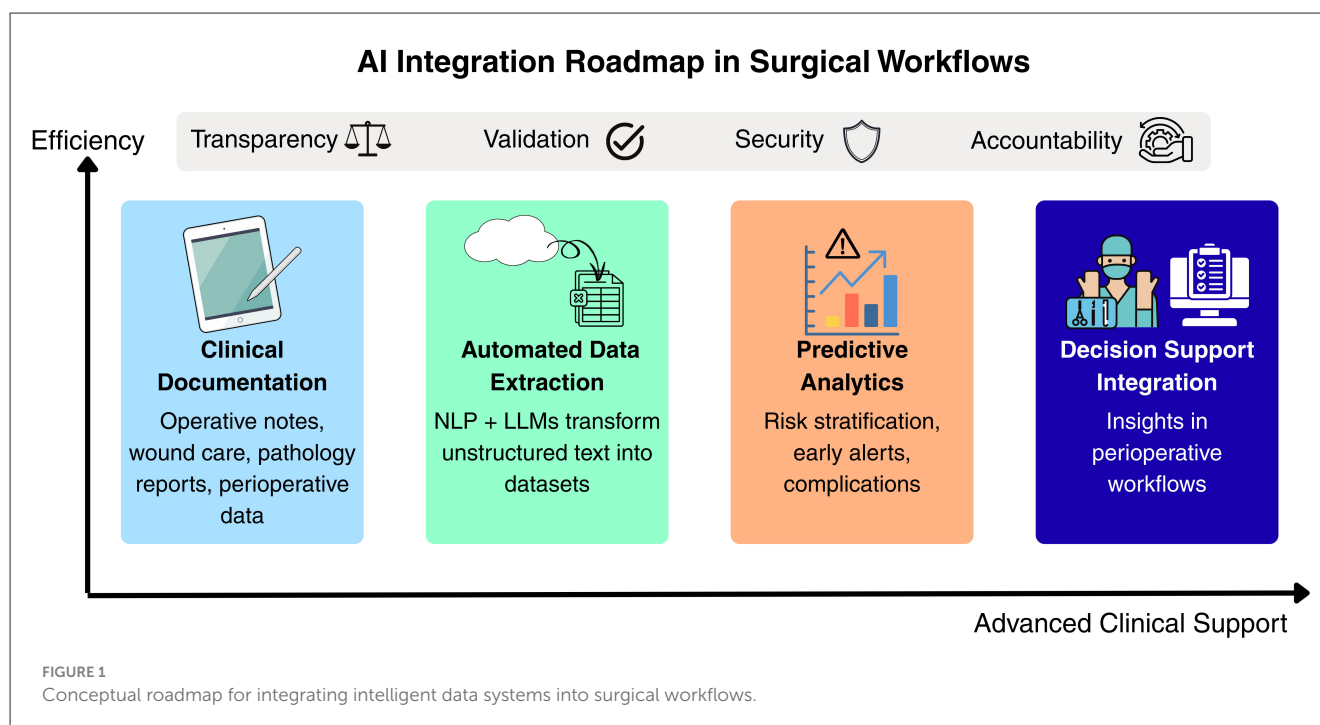
Despite these advances, substantial limitations persist. Most current studies are retrospective, single-center, and focus on feasibility; thus, their relevance to other institutions using diverse EHR systems and documentation styles is constrained. Moreover, model performance fluctuates across different data types, and occasional errors, such as hallucinations or misclassifications, underscore the continued necessity for human supervision (Mellia et al., 2021; Liang et al., 2025).

Bias and equity also present significant concerns. Numerous algorithms are built on data from high-income countries and may not represent patient populations in low- and middle-income regions. Deploying these systems globally without adaptation and tailoring increases the risk of reinforcing health disparities. Additionally, differences in wound healing, influenced by comorbidities, nutrition, or access to follow-up care, highlight the importance of population-specific validation (Ganesan et al., 2024).

Robust governance frameworks are crucial. Intelligent data systems must offer transparency, undergo validation across institutions, and adhere to international standards. Privacy and data security are vital. Clinicians may be reluctant to adopt platforms that rely on external processing, emphasizing the need for secure, locally hosted solutions (Liang et al., 2025; European Parliament and Council, 2024).

Regulators are starting to act. In the European Union, the AI Act establishes risk-based requirements for medical software, including transparency, documentation, and post-market monitoring (European Parliament and Council, 2024). In the United States, the Food and Drug Administration (FDA) has issued guidance for adaptive software that uses machine learning. This signals that surgical data tools will face more regulatory scrutiny (US Food Drug Administration, 2025).

Finally, patient trust remains central. A recent mixed-methods study found that while most patients view AI-assisted decision support positively, concerns persist about safety, accountability, and clinician oversight. Unless these issues are directly addressed, clinical adoption is likely to remain cautious and uneven (Ben Hmido et al., 2025).



Discussion

Current evidence shows that intelligent data systems can strengthen surgical data management by improving both efficiency and accuracy. Systematic reviews and meta-analyses confirm that NLP consistently outperforms traditional chart review in identifying perioperative outcomes (Mellia et al., 2021). More recent studies demonstrate that advanced NLP pipelines and LLM-based frameworks can reliably extract complex operative and pathology data across diverse specialties, including oncology and spine surgery (Chen et al., 2022; Dagli et al., 2024). These results indicate that digital tools in surgery are supported not only by conceptual promise but also by practical demonstrations of value.

Wound care exemplifies this progress clearly. Friedman et al. found that using a language model reduced chart review time by more than 80% while maintaining high accuracy, even in documentation that relies heavily on narrative text (Friedman et al., 2025). When viewed alongside findings from breast cancer and spine surgery, this evidence highlights how automated text analysis can transform unstructured information into organized data that supports both patient care and research (Chen et al., 2022; Dagli et al., 2024).

There are still significant challenges. Most studies remain retrospective, single-institution, and rarely explore the realities of real-time workflow integration. Occasional errors, such as hallucination or misclassification, require human oversight and external validation. The lack of standardized reporting frameworks also limits comparison across studies and slows clinical translation (Liang et al., 2025). Ethical, privacy, and governance issues are equally important. Without transparent validation, clear accountability, and equitable implementation, digital systems

risk eroding rather than building trust among clinicians and patients (European Parliament and Council, 2024; US Food Drug Administration, 2025; Ben Hmido et al., 2025).

Future progress will depend not only on validation but also on robust infrastructure. Multi-institutional collaborations can generate diverse datasets that capture variations in surgical practice, while standardized benchmarks can help clinicians compare model performance across institutions (Li et al., 2025). Like clinical trial registries, shared databases for algorithm performance could promote transparency, fairness, and reproducibility in real-world applications.

International coordination will also be essential. Without harmonized regulatory frameworks, global deployment of surgical data systems could become fragmented, slowing innovation and disadvantaging regions with fewer resources (Mennella et al., 2024; Rosenthal et al., 2025).

Figure 1 summarizes this staged trajectory, showing a roadmap for integrating intelligent tools into surgical workflows, from initial documentation support toward advanced perioperative decision assistance guided by transparency, validation, security, and accountability.

Looking ahead, collaboration between institutions and clear performance benchmarks will be key to ensuring safe adoption. The evolution of these systems should follow a structured path: beginning with documentation efficiency, advancing to predictive analytics for complication detection, and eventually integrating into perioperative decision support. Achieving this will require multidisciplinary cooperation, regulatory guidance, and a firm commitment to preserving clinical judgment while embracing the benefits of automation. The surgical community must take an active role in defining how these technologies are validated and applied so that innovation translates into safer and more

efficient patient care (Maleki Varnosfaderani and Forouzanfar, 2024).

Conclusions

Intelligent data systems emerge as practical tools in surgery, improving both efficiency and accuracy in data extraction across multiple specialties. However, most current studies remain retrospective and single-center, highlighting the need for broader validation, standardized benchmarks, and strong governance. A gradual pathway, from documentation support to risk prediction and perioperative decision assistance, will be key to responsible implementation. With collaboration across disciplines and active regulatory guidance, these technologies can help transform innovation into safer and more efficient surgical care.

Author contributions

EZ-M: Conceptualization, Funding acquisition, Investigation, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SA-C: Project administration, Validation, Writing – original draft, Writing – review & editing. KC-B: Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

References

- Ben Hmido, S., Abder Rahim, H., Ploem, C., Haitjema, S., Damman, O., Kazemier, G., et al. (2025). Patient perspectives on AI-based decision support in surgery. *BMJ Surg Interv Health Technol.* 7, e000365. doi: 10.1136/bmjst-2024-000365
- Beyaz, S., Özgören, A. L., Turgut, N., and Ülke, H. C. (2025). Artificial intelligence and robotic surgery in clinical medicine: progress, challenges, and future directions. *Future Sci OA.* 11, 2540742. doi: 10.1080/20565623.2025.2540742
- Byrd IV, T. F., and Tignanelli, C. J. (2024). Artificial intelligence in surgery—a narrative review. *Journal of Medical Artificial Intelligence* [cited 2025 Sept 19]; 7(0). Available from: <https://jmai.amegroups.org/article/view/9200> doi: 10.21037/jmai-24-111
- Chen, Y., Hao, L., Zou, V. Z., Hollander, Z., Ng, R. T., Isaac, K. V., et al. (2022). Automated medical chart review for breast cancer outcomes research: a novel natural language processing extraction system. *BMC Medical Research Methodology.* 22, 136. doi: 10.1186/s12874-022-01583-z
- Dagli, M. M., Ghenbot, Y., Ahmad, H. S., Chauhan, D., Turlip, R., Wang, P., et al. (2024). Development and validation of a novel AI framework using NLP with LLM integration for relevant clinical data extraction through automated chart review. *Sci Rep.* 14, 26783. doi: 10.1038/s41598-024-77535-y
- European Parliament and Council. (2024) [cited 2025 Sept 19]. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (AI Act). Available from: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
- Friedman, R., Lisk, R., Cordero-Bermudez, K., Singh, S., Ghani, S., Gillette, B. M., et al. (2025). Optimizing Chart Review Efficiency in Pressure Injury Evaluation Using ChatGPT. *Annals of Plastic Surgery.* 94(4S):S327. doi: 10.1097/SAP.0000000000004275
- Ganesan, O., Morris, M. X., Guo, L., and Orgill, D. (2024). A review of artificial intelligence in wound care. *ais.* 4, 364–75. doi: 10.20517/ais.2024.68
- Han, F., Huang, X., Wang, X., Chen, Y. F., Lu C, Li S, et al. (2025). Artificial Intelligence in Orthopedic Surgery: Current Applications, Challenges, and Future Directions. *MedComm.* 6, e70260. doi: 10.1002/mco2.70260
- Hassan, A. M., Lu, S. C., Asaad, M., Liu, J., Offodile, A. C., Sidey-Gibbons, C., et al. (2022). Novel Machine Learning Approach for the Prediction of Hernia Recurrence, Surgical Complication, and 30-Day Readmission after Abdominal Wall Reconstruction. *J Am Coll Surg.* 234, 918–27. doi: 10.1097/XCS.00000000000000141
- Ive, J., Olukoya, O., Funnell, J. P., Booker, J., Lam, S. H. M., Reddy, U., et al. (2025). AI assisted prediction of unplanned intensive care admissions using natural language processing in elective neurosurgery. *npj Digit Med.* 8, 549. doi: 10.1038/s41746-025-01952-0
- Joukes, E., Abu-Hanna, A., Cornet, R., and de Keizer, N. F. (2018). Time Spent on Dedicated Patient Care and Documentation Tasks Before and After the Introduction of a Structured and Standardized Electronic Health Record. *Appl Clin Inform.* 9, 46–53. doi: 10.1055/s-0037-1615747
- Li, M., Xu, P., Hu, J., Tang, Z., and Yang, G. (2025). From challenges and pitfalls to recommendations and opportunities: Implementing federated learning in healthcare. *Medical Image Analysis.* 101, 103497. doi: 10.1016/j.media.2025.103497
- Liang, E. N., Pei, S., Staibano, P., and van der Woerd, B. (2025). Clinical applications of large language models in medicine and surgery: A scoping review. *J Int Med Res.* 4, 53.(7):03000605251347556. doi: 10.1177/03000605251347556

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. The author(s) verify and take full responsibility for the use of generative AI in the preparation of this manuscript. Generative AI was used to support language refinement, grammar improvement, and structural editing of the text. No generative AI tools were used to create, analyze, or interpret data, figures, or results. The intellectual content, arguments, and conclusions presented in this article are entirely the responsibility of the author(s).

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Maleki Varnosfaderani, S., and Forouzanfar, M. (2024). The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century. *Bioengineering (Basel)*. 11, 337. doi: 10.3390/bioengineering11040337
- Matheny, M. E., Whicher, D., and Thadaney Israni, S. (2020). Artificial Intelligence in Health Care: A Report From the National Academy of Medicine. *JAMA*. 323, 509–10. doi: 10.1001/jama.2019.21579
- Mellia, J. A., Basta, M. N., Toyoda, Y., Othman, S., Elfanagely, O., Morris, M. P., et al. (2021). Natural Language Processing in Surgery: A Systematic Review and Meta-analysis. *Annals of Surgery*. 273, 900. doi: 10.1097/SLA.00000000000004419
- Mennella, C., Maniscalco, U., De Pietro, G., and Esposito, M. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*. 10, e26297. doi: 10.1016/j.heliyon.2024.e26297
- Murff, H. J., FitzHenry, F., Matheny, M. E., Gentry, N., Kotter, K. L., Crimin, K., et al. (2011). Automated Identification of Postoperative Complications Within an Electronic Medical Record Using Natural Language Processing. *JAMA*. 306, 848–55. doi: 10.1001/jama.2011.1204
- Rosenthal, J. T., Beecy, A., and Sabuncu, M. R. (2025). Rethinking clinical trials for medical AI with dynamic deployments of adaptive systems. *NPJ Digit Med*. 8, 252. doi: 10.1038/s41746-025-01674-3
- Rugină, A. I., Ungureanu, A., Giuglea, C., and Marinescu, S. A. (2025). Artificial Intelligence in Breast Reconstruction: A Narrative Review. *Medicina*. 61, 440. doi: 10.3390/medicina61030440
- Sinsky, C., Colligan, L., Li, L., Prgomet, M., Reynolds, S., Goeders, L., et al. (2016). Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. *Ann Intern Med*. 165, 753–60. doi: 10.7326/M16-0961
- Spoer, D. L., Kiene, J. M., Dekker, P. K., Huffman, S. S., Kim, K. G., Abadeer, A. I., et al. (2022). A Systematic Review of Artificial Intelligence Applications in Plastic Surgery: Looking to the Future. *Plast Reconstr Surg Glob Open*. 10, e4608. doi: 10.1097/GOX.00000000000004608
- US Food and Drug Administration. (2025). Artificial Intelligence in Software as a Medical Device. FDA [cited 2025 Sept 19]; Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-software-medical-device>