
Article

Not peer-reviewed version

An Algorithm for the Integration of Data from Surgical Robots and Operation Room Management Systems

[Paola Picozzi](#) , [Umberto Nocco](#) , [Chiara Labate](#) , [Greta Puleo](#) , [Veronica Cimolin](#) *

Posted Date: 3 June 2025

doi: [10.20944/preprints202506.0171.v1](https://doi.org/10.20944/preprints202506.0171.v1)

Keywords: data Integration; automation in healthcare; clinical engineering; robotic surgery



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

An Algorithm for the Integration of Data from Surgical Robots and Operation Room Management Systems

Paola Picozzi ^{1,2}, Umberto Nocco ², Chiara Labate ², Greta Puleo ² and Veronica Cimolin ^{1,3,*}

¹ Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milano, Italy

² Clinical Engineering Department of ASST Grande Ospedale Metropolitano Niguarda, 20162 Milano, Italy

⁴ Istituto Auxologico Italiano, IRCCS, S. Giuseppe Hospital, Strada Cadorna 90, 28824 Piancavallo, Italy

* Correspondence: veronica.cimolin@polimi.it

Abstract: This study presents an algorithm developed by the Clinical Engineering department to automatically match surgical events recorded by robotic systems with corresponding entries in the hospital's OR management software. At ASST Grande Ospedale Metropolitano Niguarda, robotic procedures were previously identified manually by surgical staff within the operating room management system, leading to frequent inconsistencies and data quality issues. Two heterogeneous datasets—robot logs and hospital procedure records—were aligned based on common features such as date, duration, and operating room, despite lacking a unique identifier. The matching algorithm enables the accurate identification of robotic procedures within the hospital system and facilitates integration of clinical and technical data into a unified framework. Results show a substantial increase in correctly matched records, significantly improving the reliability of robotic procedure tracking. This integrated approach supports more effective data utilization for clinical engineering activities, operational monitoring, and Health Technology Assessment (HTA) analyses. The work provides a practical solution to a real-world data integration problem and lays the foundation for future extensions, including the application of machine learning to enhance matching precision.

Keywords: data Integration; automation in healthcare; clinical engineering; robotic surgery

1. Introduction

In recent years, data collection and analysis have become increasingly central in all sectors to support strategic decision-making processes [1,2]. The concept of data-driven decision-making has become increasingly popular in all sectors, including healthcare [3–5]. As a result, the concept of data-driven healthcare is emerging as a crucial approach to improve patient care and optimize resource allocation, relying on data and evidence to guide clinical and organizational choices [6,7].

To exploit the full potential of health data, it is essential that they are properly structured and integrated [8–10]. Fragmented or poorly organized information risks being underutilized, thus limiting its strategic value. Today, clinical data is not only generated by medical records or manually entered administrative systems, but an increasing amount of data comes directly from medical devices and equipment [11,12]. However, to be useful, this data needs to be carefully linked to patient information and hospital information systems.

This study presents a project carried out by the Clinical Engineering Department of the ASST Grande Ospedale Metropolitano Niguarda, aimed at integrating data from robotic surgical systems (Da Vinci®) with data from the operating room management system. In these two datasets, there is no unique identifier linking them together. For this reason, the main objective was to automatically identify robotic surgical procedures in the operating room (OR) management system that, until recently, could only be tracked through manually entered fields, such as the designation of the type of procedure (e.g. endoscopic, robotic, ...) or the use of the code 00.39 to indicate robotic assistance.

As this information was entered manually by the operating room staff, it was subject to a significant margin of error and inconsistency.

The proposed integration approach uses objective data extracted from the robotic systems, such as procedure dates, console start times and total duration of console use, to match and accurately identify robotic procedures among all procedures recorded in the hospital's OR management system. This allows not only a more reliable identification, but also an enrichment of clinical and operational data through the fusion of technical parameters and administrative records.

In addition, robotic data, being automatically generated by the system itself, provide objective metrics - unaffected by human error - such as the precise duration of console use. In contrast, durations recorded in the hospital system may be less accurate due to manual input. This integration is therefore crucial to improve data quality and enable more accurate monitoring and evaluation of robotic surgical activity.

Before this project, verifying the correct identification of the operations performed with the robot required a manual comparison between the data from the equipment and the information entered in the operating theatre management system. This process depended on human intervention, making it time-consuming and prone to errors. Therefore, the development of an automatic system that allows autonomous identification of the operations performed with the surgical robots and the integration of the data from the two databases makes it possible to considerably reduce the working time by making the entire process more efficient.

Ultimately, this project aims to facilitate the comprehensive integration of technical and clinical data to improve the management and evaluation of robotic surgery. Given the higher costs typically associated with robotic procedures compared to traditional surgical methods, access to detailed and accurate data can support more informed decisions, optimize resource utilization and potentially contribute to cost containment through increased operational efficiency.

2. Materials and Methods

This project was conducted by the Clinical Engineering Department of the ASST Grande Ospedale Metropolitano Niguarda in Milan, with the aim of solving a practical and recurring problem: the difficulty of reliably identifying robot-assisted surgical procedures within the hospital's operating room information system.

To address this issue, an algorithm capable of automatically linking robotic procedures recorded by da Vinci® surgical systems with the corresponding entries in the hospital's operating room management system (Ormaweb [13]) was developed. The aim was to achieve an accurate and reproducible match, enabling the integration of clinical and technical data between the various systems and ultimately supporting analyses of procedure types, resource utilisation and performance of the robotic platforms.

2.1. Data Source

Research manuscripts reporting large datasets that are deposited in a publicly available database should specify where the data have been deposited and provide the relevant accession numbers. If the accession numbers have not yet been obtained at the time of submission, please state that they will be provided during review. They must be provided prior to publication.

Two main data sources were used:

1. Operating room database (Ormaweb): This database contains structured records of all surgical procedures performed in the hospital. Each row represents one surgical procedure. This dataset contains various data such as:
 - The surgery identification number and the patient identification number allow this dataset to be linked to other hospital information systems.
 - Information about the operating theatre and operating block where the operation is performed.

- The surgical speciality and type of surgery together with the DRG (Diagnosis Related Group) reimbursement code
- Information about the operating theatre staff.
- Information about the date of surgery and all operating times (entry into the operating room, entry into the operating room, start of anaesthetic preparation, patient ready, start of surgery, end of surgery, exit from the operating room, exit from the operating room).

2. Records of robotic procedures: These were exported from the da Vinci® robotic surgical systems and include the serial number of the system, the date and local time of each procedure, and the total duration of the procedure in minutes.

Thus, the two data sources do not have a common key to match them easily.

The data extracted from the operating room management system and the surgical robot are from 01/01/2023 to 30/06/2024.

2.2. Data Pre-Processing

Both datasets were imported and processed using the Python programming language. For the robotic dataset, a new column was created combining the date and procedure start time to generate a single reference timestamp for each case (Local Procedure Start DateTime). In Ormaweb, the first step was to convert the data type of the date and time fields to “datetime”, since they were strings in the original dataset.

As a second aspect, the algorithm took into account cases in which some fields had been left blank because the operating room staff had not filled them in during the procedure. Therefore, auxiliary columns were constructed to define the most reliable available time window during which a procedure could have taken place. These were filled in by prioritising the available timestamps in a hierarchical order (e.g. entry to the block, entry to the room, start time of the procedure).

2.2. Matching Algorithm

The core of the methodology consists of a deterministic algorithm (Figure 1) developed in Python. For each robotic procedure represented by a row in the robotic dataset, the algorithm attempted to identify a matching intervention within Ormaweb by applying a stepwise filtering process:

1. Same-day filter: Only surgical procedures that occurred on the same date as the robotic registry were considered.
2. Time window filter: The start time of the robotic procedure on the console had to fall within the estimated entry and exit times recorded in the clinical system.
3. Duration filter: The duration of the robotic procedure had to be shorter than the candidate's total time window.

If more than one potential match was found after these steps, the algorithm applied additional filters based on known associations between robotic systems' serial number and surgical rooms:

- System SK5054 was associated with OR 08.
- System SK7255 was associated with OR 10.
- System SK5389 was associated with OR 12.

In the rare cases where an ambiguity remained after all filters had been applied, fallback logic was used to refine the selection. This involved comparing robotic data with narrower intra-operative time windows, rather than with wider entry/exit times.

Only if a single unique match was identified and the associated operating theatre was consistent with expectations was the match accepted.

If the algorithm could not find any matches using the filters listed above, it would insert a note indicating that the match did not take place. The operator must then manually match the remaining transactions. This may happen because in some cases the data entered manually in the room management system may be incorrect and therefore in no way match the data recorded by the robot.

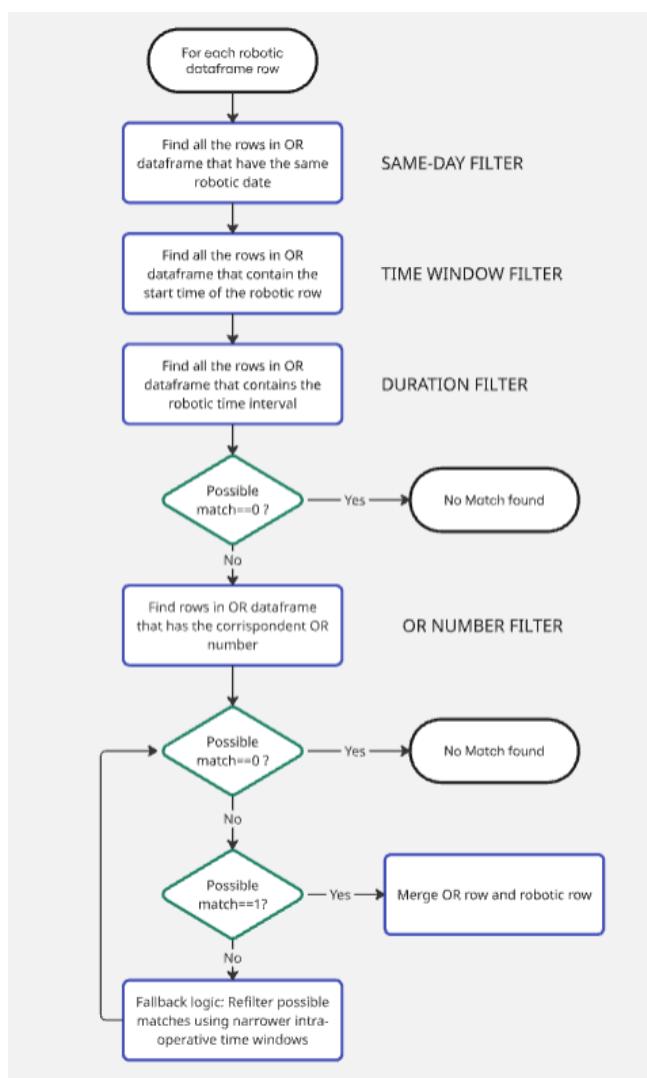


Figure 1. Matching algorithm flowchart.

3.2. Outcome Metrics

To evaluate the performance of the algorithm, two main metrics were considered:

- The total number of robotic procedures not matched to an entry in the Ormaweb database.
- The accuracy of matching between the two datasets, determined by checking the match between surgeries matched by the algorithm and those manually matched by two people.

3. Results

The developed algorithm was applied to a dataset consisting of 14,500 surgical procedures recorded in the operating room management system (Ormaweb) from 1 January 2023 to 30 June 2024 and 1,372 procedure logs extracted from the Da Vinci® robotic systems. The primary objective was to accurately identify robot-assisted surgeries through the integration of these two independent data sources. The matching algorithm successfully matched 1,362 robotic procedures with corresponding entries in the Ormaweb database, achieving a 99.27% match rate.

To validate the accuracy of the algorithm, the matched cases were manually reviewed by two independent operators. All 1,362 matched procedures were confirmed as correct, achieving a 100% accuracy rate compared to manual validation. This result demonstrates the robustness and reliability of the matching strategy, which is mainly based on the synchronization of timestamps and the use of common metadata, such as operating theatre identifiers and duration of surgery.

Only 10 robotic diaries (0.73%) remained unpaired. In most of these cases, the inability to establish a match was attributable to missing or inconsistent data, such as incomplete timestamps or procedural details in the operating room management system. These mismatched cases are flagged for manual review and resolution by clinical engineering staff, ensuring that no robotic procedure is excluded from subsequent analysis. This approach balances automation with a final manual validation step to ensure complete data integration.

The entire matching process was performed in less than 3 seconds on a standard personal computer (Intel Core i7, 16 GB RAM), highlighting the scalability and practical applicability of the method. This level of performance suggests that the algorithm can be integrated into real-time or near real-time data pipelines for continuous monitoring of robotic surgery. Furthermore, the low computational cost makes it suitable for regular use by hospital staff, without the need for specialized hardware or technical expertise.

4. Discussion

The integration of clinical and technical data represents one of the most significant challenges in the digital transformation of healthcare. The project described in this study addressed a concrete and recurring problem, namely the difficulty of reliably identifying robot-assisted surgical procedures in OR management systems. This problem is common to many clinical settings where data from medical devices are not directly interoperable with hospital administrative and management systems [14,15].

The results obtained demonstrate the effectiveness of a deterministic approach, based on the synchronization of timestamps and metadata, for automatically linking logs generated by Da Vinci® robotic systems with records in the Ormaweb platform. With a success rate of 99.27% and 100% accuracy compared to manual validation, the method proved to be highly robust. These results are in line with the literature, which highlights that time synchronization, and the use of rules based on contextual metadata are effective strategies for record linkage in healthcare [16,17].

Another relevant aspect that emerged from the study concerns data quality and reliability. Data generated automatically by surgical robots provide objective metrics, such as the precise duration of console use, which are not subject to human error, unlike manual records in hospital management systems. This allows for a more accurate measurement of surgical activities and provides a stronger basis for operational analysis, performance evaluations and strategic decisions.

The adoption of an automated algorithm for matching procedures also allows for a significant reduction in work time for clinical and technical staff. In the past, such verification required time-consuming and error-prone manual checks, whereas now it can be completed in seconds with simple computational processing. This approach aligns with the principles of “digital health” and “learning health systems”, where automation and intelligent use of data enable improved decision-making in healthcare [18,19].

However, the study also highlights some limitations. The failure of matching in about 0.73% of the cases was attributed to incomplete or inconsistent data, particularly in the manually entered records in the Ormaweb system. This highlights the need to improve data quality at source and underlines the importance of supplementing the automatic process with a manual review phase in ambiguous cases, as also recommended by previous studies on data linkage.

A more conservative approach was adopted, avoiding automatic matching of interventions without an exact correspondence, in order to prevent incorrect associations that could result from relaxing the matching criteria. In such cases, the task of manually linking the two datasets is left to the operator. A further limitation of this study is that integration does not take place directly between the two systems, but it is necessary to download the excel files containing all the data and then integrate them using the code developed in this study. In the future, there should be a direct integration, i.e. the robot can send data directly to the operating theatre management system.

In addition, code should also be developed that integrates data from other types of surgical robots, not just those from the Da Vinci®.

Finally, the possibility of integrating this algorithm into near real-time monitoring pipelines opens interesting perspectives for the continuous analysis of robotic activity, cost evaluation and optimization of resource utilization. Considering the high cost associated with robotic surgery, access to accurate and timely data is crucial to ensure sustainable and efficient management of healthcare services.

Author Contributions: Conceptualization, P.P., U.N., C.L., G.P. and V.C.; methodology, P.P. and V.C.; software, P.P.; validation, P.P.; data curation, P.P.; writing—original draft preparation, P.P. and V.C.; writing—review and editing, P.P., U.N., C.L., G.P. and V.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. S. Hjelle, P. Mikalef, N. Altwaijry, and V. Parida, "Organizational decision making and analytics: An experimental study on dashboard visualizations," *Information and Management*, vol. 61, no. 6, Sep. 2024, doi: 10.1016/j.im.2024.104011.
2. S. Fanelli, L. Pratici, F. P. Salvatore, C. C. Donelli, and A. Zangrandi, "Big data analysis for decision-making processes: challenges and opportunities for the management of health-care organizations," *Management Research Review*, vol. 46, no. 3, pp. 369–389, Feb. 2023, doi: 10.1108/MRR-09-2021-0648.
3. A. T. Freitas, "Data-Driven Approaches in Healthcare: Challenges and Emerging Trends," *Law, Governance and Technology Series*, vol. 58, pp. 65–80, 2024, doi: 10.1007/978-3-031-41264-6_4.
4. L. J. Basile, N. Carbonara, R. Pellegrino, and U. Panniello, "Business intelligence in the healthcare industry: The utilization of a data-driven approach to support clinical decision making," *Technovation*, vol. 120, p. 102482, Feb. 2023, doi: 10.1016/J.TECHNOVATION.2022.102482.
5. Ibrahim Adedeji Adeniran, Christianah Pelumi Efunniyi, Olajide Soji Osundare, and Angela Omozele Abhulimen, "Data-driven decision-making in healthcare: Improving patient outcomes through predictive modeling," *International Journal of Scholarly Research in Multidisciplinary Studies*, vol. 5, no. 1, pp. 059–067, Aug. 2024, doi: 10.56781/ijsrms.2024.5.1.0040.
6. F. Cascini, F. Santaroni, R. Lanzetti, G. Failla, A. Gentili, and W. Ricciardi, "Developing a Data-Driven Approach in Order to Improve the Safety and Quality of Patient Care," *Front Public Health*, vol. 9, p. 667819, May 2021, doi: 10.3389/FPUBH.2021.667819.
7. L. J. Basile, N. Carbonara, U. Panniello, and R. Pellegrino, "The exploitation of data to support decision-making in healthcare: a systematic literature review and future research directions," *Management Review Quarterly* 2024, pp. 1–33, Apr. 2025, doi: 10.1007/S11301-024-00482-5.
8. K. Chao, M. N. I. Sarker, I. Ali, R. B. R. Firdaus, A. Azman, and M. M. Shaed, "Big data-driven public health policy making: Potential for the healthcare industry," *Helijon*, vol. 9, no. 9, p. e19681, Sep. 2023, doi: 10.1016/J.HELJON.2023.E19681.
9. M. Javaid, A. Haleem, and R. P. Singh, "Health informatics to enhance the healthcare industry's culture: An extensive analysis of its features, contributions, applications and limitations," *Informatics and Health*, vol. 1, no. 2, pp. 123–148, Sep. 2024, doi: 10.1016/J.INFOH.2024.05.001.
10. M. Zarour et al., "Ensuring data integrity of healthcare information in the era of digital health," *Healthc Technol Lett*, vol. 8, no. 3, p. 66, Jun. 2021, doi: 10.1049/HTL2.12008.
11. E. Martens et al., "Smart hospital: achieving interoperability and raw data collection from medical devices in clinical routine," *Front Digit Health*, vol. 6, p. 1341475, Mar. 2024, doi: 10.3389/FDGTH.2024.1341475/BIBTEX.
12. J. B. Withall, J. M. Schwartz, J. Usseglio, and K. D. Cato, "A Scoping Review of Integrated Medical Devices and Clinical Decision Support in the Acute Care Setting," *Appl Clin Inform*, vol. 13, no. 5, p. 1223, Oct. 2022, doi: 10.1055/S-0042-1759513.

13. "Ormaweb" web page: <https://www.dedalus.com/italy/it/la-nostra-offerta/prodotti/o4c/> (Accessed on 03/02/2025)
14. P. B. Jensen, L. J. Jensen, and S. Brunak, "Mining electronic health records: towards better research applications and clinical care," *Nature Reviews Genetics* 2012 13:6, vol. 13, no. 6, pp. 395–405, May 2012, doi: 10.1038/nrg3208.
15. S. M. Meystre, C. Lovis, T. Bürkle, G. Tognola, A. Budrionis, and C. U. Lehmann, "Clinical Data Reuse or Secondary Use: Current Status and Potential Future Progress," *Yearb Med Inform*, vol. 26, no. 1, pp. 38–52, Aug. 2017, doi: 10.15265/IY-2017-007.,
16. K. Harron et al., "Challenges in administrative data linkage for research," *Big Data Soc*, vol. 4, no. 2, Dec. 2017, doi: 10.1177/2053951717745678.
17. P. Christen, "Data matching: Concepts and techniques for record linkage, entity resolution, and duplicate detection," *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*, pp. 1–270, Jan. 2012, doi: 10.1007/978-3-642-31164-2/COVER.
18. C. P. Friedman, J. C. Rubin, and K. J. Sullivan, "Toward an Information Infrastructure for Global Health Improvement," *Yearb Med Inform*, vol. 26, no. 1, pp. 16–23, Aug. 2017, doi: 10.15265/IY-2017-004.,
19. W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," *Health Inf Sci Syst*, vol. 2, no. 1, p. 3, Feb. 2014, doi: 10.1186/2047-2501-2-3.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.