

## Paper RE04

### Future-Focused Real-World Data Utilization Strategies: Resilience and Innovation from Chugai

Nauta Yamanaka, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

Wataru Ohtsuka, Chugai Pharmaceutical Co., Ltd., Tokyo, Japan

#### ABSTRACT

Under “Chugai Digital Vision 2030”, Chugai Pharmaceutical promotes Digital Transformation for Drug Discovery and Development, positioning Real-World Data (RWD) as a key driver for regulatory, clinical, and operational decision-making. Innovation is further supported by cross-functional data-science collaboration and strengthened internal and external networking.

Within this strategic context, this paper provides an overview of Chugai’s enterprise-wide initiatives to advance the utilization of RWD and data science. Particular attention is given to efforts involving unstructured clinical information, including a case example assessing the feasibility of automated identification of brain metastasis from head MRI narrative reports across multiple institutions.

These activities illustrate an evolving framework for enhancing analytical capability, improving evidence generation processes, and supporting data-driven practices within pharmaceutical development.

#### INTRODUCTION

The use of real-world data (RWD) in pharmaceutical R&D has been growing, providing opportunities to complement clinical trials, broaden clinical understanding, and support decision-making across development stages. In Japan, the potential for RWD utilization has expanded further with nationwide health coverage and increasing availability of large-scale datasets.

In 2019, Chugai established a dedicated DX Promotion Department to orchestrate company-wide digital initiatives, laying the foundation for “CHUGAI DIGITAL VISION 2030” and the initiative “Digital Transformation for Drug Discovery and Development”. This paper describes Chugai’s approach to developing and applying practical RWD capabilities that align with the company’s digital vision. It also introduces an example initiative involving the extraction of clinical outcome information—specifically, the presence or absence of brain metastasis—from unstructured radiology text collected across multiple institutions. This paper also summarizes initiatives that support effective RWD utilization through data-science activities at Chugai, including cross-functional efforts such as an internal RWD portal and company-wide data-science events.

#### WHY CHUGAI IS COMMITTED TO RWD UTILIZATION

Pharmaceutical R&D has faced a long-term decline in productivity worldwide, increasing the importance of applying digital technologies to enhance efficiency and evidence generation throughout drug development. Chugai Pharmaceutical, as a research-driven company with strengths in unique technologies and science, has positioned digital transformation as one of its core strategic pillars. Among the company’s defining

characteristics—its role as one of Japan’s leading prescription pharmaceutical companies, a unique business model, distinctive technological capabilities, and industry leadership in digital transformation—the advancement of RWD utilization plays a central role.

Under “CHUGAI DIGITAL VISION 2030” (Figure 1), the company promotes enterprise-wide digital transformation with the aim of continuously delivering innovative medicines while improving productivity across the value chain. The vision explicitly highlights the active use of RWD/RWE as essential components for creating innovative new drugs through digital technologies.

RWD is expected to contribute across multiple stages of drug development - from early consideration of target product profiles and clinical development planning to post-marketing activities. By enabling more comprehensive understanding of patient populations, treatment patterns, and clinical outcomes, RWD offers opportunities to support timely and evidence-based decision-making.

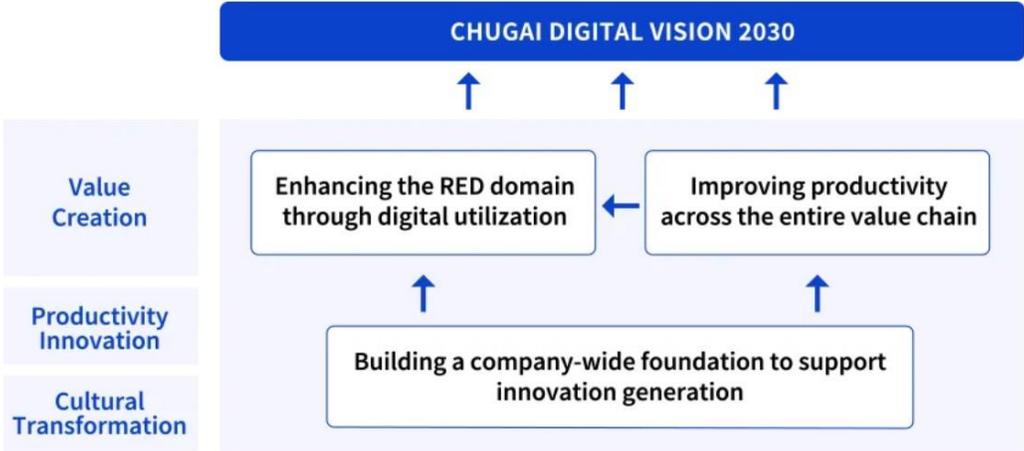


Figure 1: CHUGAI DIGITAL VISION 2030

## UTILIZATION OF RWD SPECIFIC TO THE JAPANESE POPULATION

### Health Insurance Structure in Japan

Japan's universal health insurance system (Figure 2) enables individuals to receive medical care at any clinic or hospital by presenting a health insurance card, which has been integrated into the Individual Number Card from December 2025. Under this reimbursement structure, patients typically pay 30% of medical costs, with the remaining portion covered by insurance; the copayment rate is lower for children and older adults, with adjustments based on age, income, or local subsidies. The system also allows free choice of medical provider.

These characteristics contribute to the generation of comprehensive administrative and clinical records, forming an important foundation for RWD utilization, while simultaneously requiring rigorous privacy protection and ethical oversight.

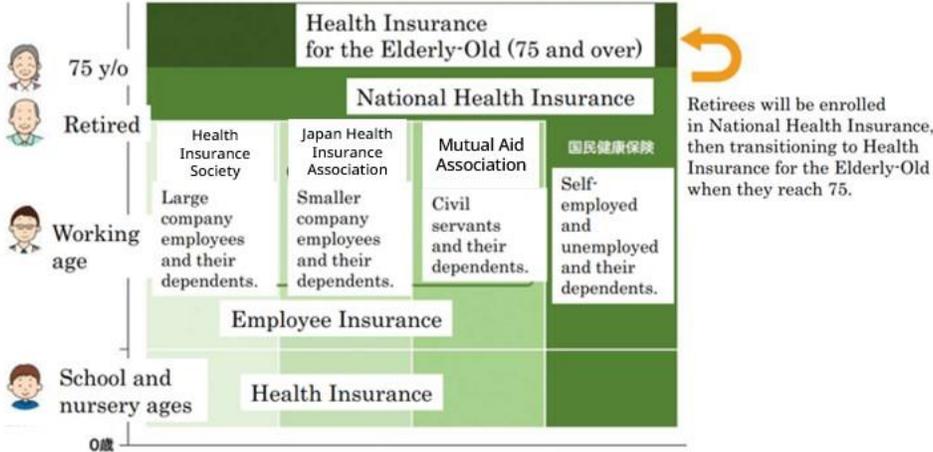


Figure 2: Japan Medical Association

**Data Sources Available in Japan**

Major real-world datasets available in Japan include claims data, health checkup records, electronic medical records, and patient registries. Each dataset has its own strengths and limitations, reflecting differences in structure, level of clinical detail, and update frequency.

Key challenges associated with the use of these sources include data standardization, privacy protection, and ethical considerations, particularly when combining multiple datasets to enhance analytical completeness.

**Policy Momentum: Next Generation Medical Infrastructure Act**

The 2024 revision of the Next Generation Medical Infrastructure Act (Figure 3) expanded the permissible use of pseudonymized medical data, enabling access to diagnostic categories and test values—including those related to rare conditions—for research and regulatory applications.

The revision also allows linkage between medical datasets and national public databases, such as the National Database (NDB) and long-term care datasets, thereby creating opportunities for broader multi-source analyses while maintaining necessary oversight through appropriate approvals and safeguards.

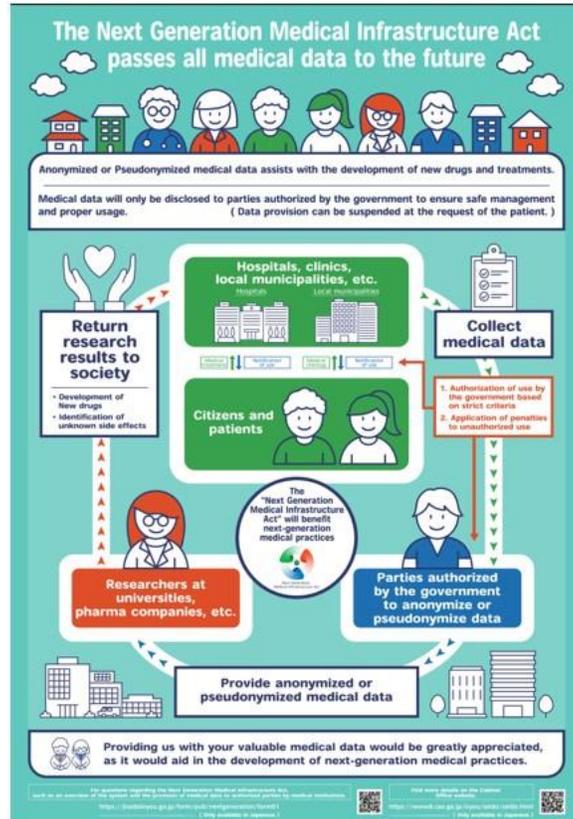


Figure 3: Next Generation Medical Infrastructure Act

## Beyond Structured Data

Structured datasets often do not capture key clinical details such as treatment effects, adverse events, the rationale behind clinical decisions, reasons for treatment changes, or outcome-related information. In routine practice, these elements are frequently documented in unstructured text sources, including clinical notes, radiology reports, and pathology reports, which contain richer contextual information than structured fields.

Natural language processing (NLP) offers a potential means to extract such information in a systematic manner. In this paper, an exploratory initiative is introduced to assess whether essential clinical details—specifically, the presence or absence of brain metastasis—can be derived from unstructured radiology texts collected across multiple institutions.

## CASE STUDIES OF RWD UTILIZATION AT CHUGAI

### Background

Brain metastasis is a clinically significant complication in cancer patients, adversely impacting QoL and survival outcomes; therefore, understanding its epidemiology is important to treatment strategies. While there are some international reports on brain metastases in breast cancer patients, comprehensive clinical

data on brain metastases from breast cancer remain limited in Japan. Clinical information regarding brain metastases is usually documented in an unstructured manner, such as narrative radiology reports, this situation poses significant challenges for automated data extraction and analysis.

### Objective

To evaluate the feasibility of developing an automated detection system for extracting brain metastasis information from unstructured MRI reports across multiple institutions within the J-CONNECT consortium framework.

### Methods

Patients who underwent chemotherapy and were registered in the CONNECT2 study – one of the J-CONNECT projects, the "Multicenter Collaborative Study on the Comprehensive Collection and Utilization of RWD in Cancer Care" (UMIN Study ID: UMIN000051605) – were identified from institutional cancer registries between 2018 and 2021. From these, 258 patients with advanced or recurrent breast cancer were selected across three J-CONNECT participating institutions based on detailed tumor-site and histopathological information.

To assess the feasibility of automated identification of brain metastasis, a rule-based determination approach using dependency parsing was developed. The method examined metastasis-related expressions within the "Diagnosis" and "Findings" sections of each MRI report (Figure 4) and applied predefined lexical and syntactic rules to classify the presence or absence of brain metastasis (Figure 5).

The system was initially developed at one institution and subsequently refined at additional sites. Performance was evaluated using F1-score and error rate.

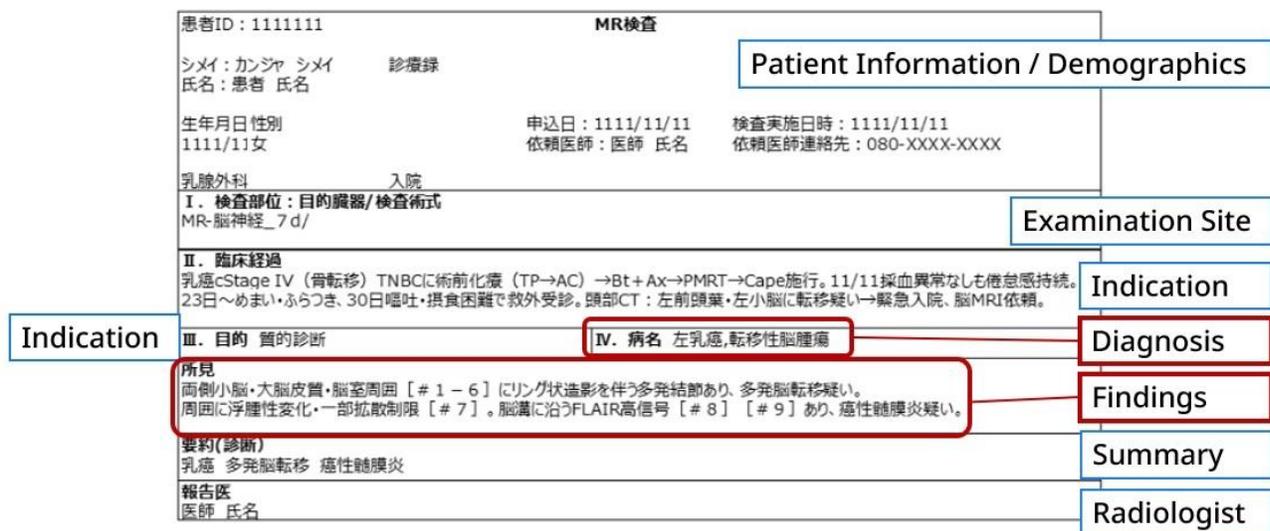


Figure 4: Example of a Head MRI Report

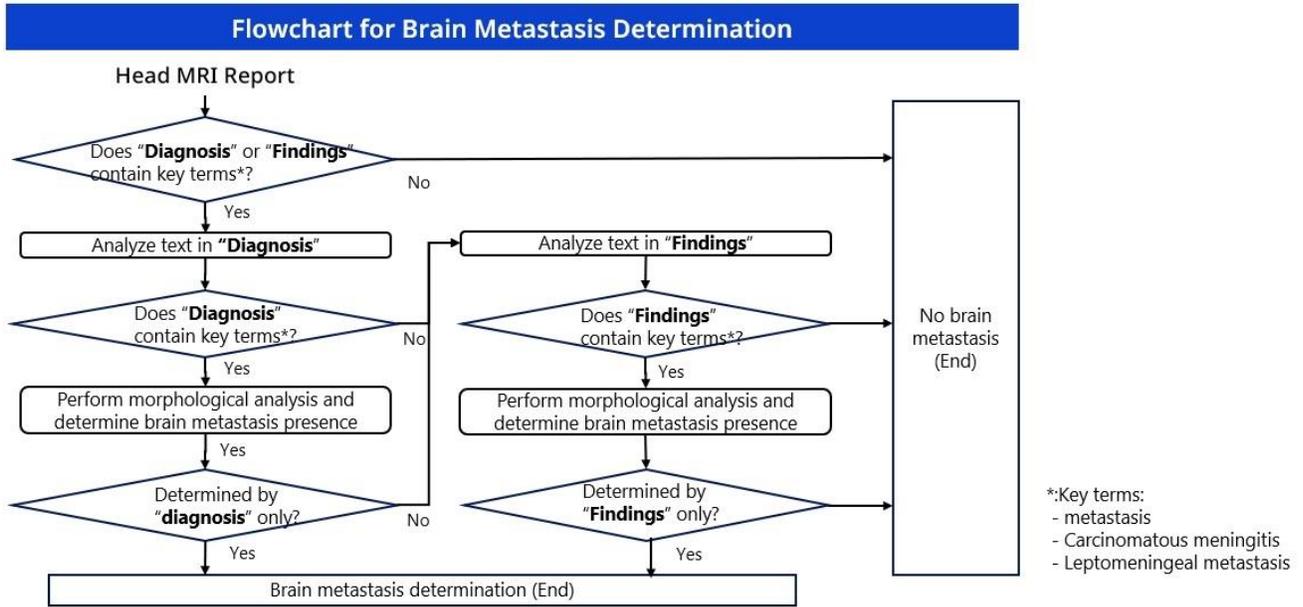


Figure 5. Flowchart of Brain Metastasis Determination Logic.

## Results

Of the 258 patients, 43 (16.7%) had brain metastases, while 215 (83.3%) did not. A total of 845 head MRI reports were reviewed for final judgment, comprising 228 from the first institution, 201 from the second institution, and 416 from the third institution.

Model performance varied across institutions: perfect classification at Institution 1; F1-score 0.97 (error 0.04) improving to 0.98 (0.02) at Institution 2; and 0.68 (0.07) improving to 0.93 (0.01) at Institution 3 following refinement (Figure 6).

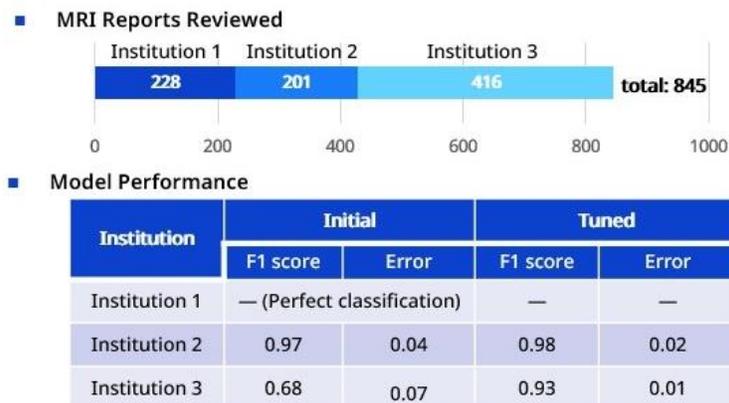


Figure 6: MRI Reports Reviewed and Reference Values

## Conclusion

The study demonstrated that dependency-parsing approaches applied to unstructured brain MRI reports can identify brain metastasis in breast cancer patients and may offer a scalable method for enhancing RWD-based clinical outcome assessment within multi-institutional frameworks.

## Note

This work is partially based on a study previously presented at the 45th Joint Conference on Medical Informatics (JcMI): “Extraction of Clinical Outcome Information from Unstructured Data - Detection of Brain Metastases from Head MRI Reports in Patients with Advanced and Recurrent Breast Cancer -”

## DATA SCIENCE ACTIVITIES AT CHUGAI

Advancing RWD utilization at scale requires an operating model that integrates infrastructure, data accessibility, talent development, and mechanisms that facilitate cross-functional collaboration. At Chugai, these activities are supported by a consultation framework that includes local data-science contact points within departments and dedicated channels for quick inquiries. Human-resource development is promoted through programs such as the CHUGAI DIGITAL ACADEMY (CDA), which provides structured training in digital and data-science skills.

Awareness and engagement are further strengthened through company-wide events such as Data Science Day, where case studies are shared to encourage collaboration and stimulate innovation. For day-to-day activities, departmental representatives share updates and gather needs across the organization via centralized RWD Portal and associated Teams channels.

Together, these elements contribute to an internal ecosystem for RWD utilization centered on data-science practices, enabling consistent, practical, and collaborative use of RWD across multiple functions.

## CONCLUSION

Progress in RWD utilization in Japan continues despite persistent challenges, supported by developments in data infrastructure, changes in regulatory frameworks, and the accumulation of practical use cases. At the same time, data-access frameworks remain under development, and further advancement may be facilitated through collaboration among industry, academia, and other relevant stakeholders.

Looking ahead, RWD is anticipated to play an increasingly recognized role in clinical trial design, regulatory submissions, and the demonstration of therapeutic value.

## REFERENCES

Figure 1: Chugai Pharmaceutical Co. Ltd., *CHUGAI DIGITAL VISION 2030*.

Available at: <https://www.chugai-pharm.co.jp/english/innovation/digital/vision.html> (accessed 17 December 2025).

Figure 2: Japan Medical Association. *Overview of Japan's Health Insurance System*.  
Available at: <https://www.med.or.jp/people/info/kaifo/system/> (accessed 17 December 2025).

Figure 3: Cabinet Office, Government of Japan. *Next-Generation Medical Infrastructure: Overview and Implementation*.

Available at: [https://www8.cao.go.jp/iryoinstitutions/pdf/jisedai\\_poster\\_e.pdf](https://www8.cao.go.jp/iryoinstitutions/pdf/jisedai_poster_e.pdf) (accessed 17 December 2025).

## **CONTACT INFORMATION**

Author Name: Nauta Yamanaka

Company: Chugai Pharmaceutical Co., Ltd.

Email: [yamanaka.nauta05@chugai-pharm.co.jp](mailto:yamanaka.nauta05@chugai-pharm.co.jp)