MEDINFO 2023 — The Future Is Accessible
J. Bichel-Findlay et al. (Eds.)

doi:10.3233/SHTI230918

Design of HL7 FHIR Profiles for Pathology Reports Integrated with Pathology Images

Chung-Yueh Lien[a,1], Tzu-Yun Ting[a], Li-Chun Kuo[a], Pau-Choo Chung[b], Yuan-Chia Chu[c], Chen-Tsung Kuo[c],
[a] *Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan*
[b] *Electrical Engineering and Computer Science, National Cheng Kung University, Tainan, Taiwan*
[c] *Department of Information Management, Taipei Veterans General Hospital, Taipei, Taiwan*

ORCiD ID: Chung-Yueh Lien https://orcid.org/0000-0002-7230-6303, Pau-Choo Chung https://orcid.org/0000-0002-8660-570X, Yuan-Chia Chu https://orcid.org/0000-0002-0997-5168, Chen-Tsung Kuo https://orcid.org/0000-0002-7139-4663

Abstract. This paper describes the development of Health Level Seven Fast Healthcare Interoperability Resource (FHIR) profiles for pathology reports integrated with whole slide images and clinical data to create a pathology research database. A report template was designed to collect structured reports, enabling pathologists to select structured terms based on a checklist, allowing for the standardization of terms used to describe tumor features. We gathered and analyzed 190 non–small-cell lung cancer pathology reports in free text format, which were then structured by mapping the itemized vocabulary to FHIR observation resources, using international standard terminologies, such as the International Classification of Diseases, LOINC, and SNOMED CT. The resulting FHIR profiles were published as an implementation guide, which includes 25 profiles for essential data elements, value sets, and structured definitions for integrating clinical data and pathology images associated with the pathology report. These profiles enable the exchange of structured data between systems and facilitate the integration of pathology data into electronic health records, which can improve the quality of care for patients with cancer.

Keywords: pathology report, FHIR, Implementation Guide, structured report, interoperability

1. Introduction

Pathology reports contain comprehensive information about cancer diagnosis, grading, margin status, and histopathological characteristics. The pathologist's observation of a sample of tumor tissue plays a crucial role in both cancer diagnosis and clinical treatment.

[1] Corresponding Author: Chung-Yueh Lien, Address: No. 365, Ming-te Road, Peitou District, Taipei City, Taiwan, e-mail: chungyueh@ntunhs.edu.tw

To facilitate data collection and analysis, it is important to have an efficient method for gathering cancer information, and statistical analysis can provide a powerful tool for diagnosis and risk assessment. However, data on cancer are often stored in electronic format, which are unstructured and heterogeneous across different Electronic Health Record systems. The lack of interoperability can pose challenges to data collection and analysis, especially when medical data must be exchanged across multiple institutions or systems [1]. To address this issue, data collection and sharing must establish clear data standards in order to achieve interoperability [2]. The Fast Healthcare Interoperability Resource (FHIR) for Health Level Seven (HL7) is expected to solve this interoperability problem and improve the usability of health data.

To enhance the interoperability of clinical and genetic data for precision oncology, the American Society of Clinical Oncology created the minimal common oncology data elements (mCODE) profile. Osterman et al. outlined the development of mCODE, which is divided into six thematic groups for instructional purposes. The mCODE profile, which consists of 23 profiles and 90 data elements, serves as a framework for incorporating structured data from patients with cancer into data exchange [3]. Guérin et al. also proposed a common data model to share standardized and scalable data concepts, including a minimum set of clinical and genomic data that is compatible with the HL7 FHIR format for accelerating data sharing in oncology [4]. Additionally, there are ongoing efforts to align clinical cancer data elements and concepts with FHIR standards, such as Zong et al.'s extension of an existing cancer profile by extracting common data elements from four colorectal cancer clinical trial casereport forms and adapting the FHIR-based data model as needed [5].

This study aimed to create and test an interoperability model for pathology reports with associated whole-slide images that adhere to the American Joint Committee on Cancer (AJCC) 8th checklist for diagnostic items in cancer pathological reports. The proposed model uses the FHIR methodology and tools to define the specifications of clinical data, which reduces time and labor costs associated with the data exchange. By providing an interoperability model that meets the AJCC 8th checklist, this study offers a standard approach for pathological reporting that can enhance the accuracy and consistency of cancer diagnosis and treatment.

2. Methods

Using the dataset of 190 prospective admissions to Taipei Veterans General Hospital, this study aimed to create FHIR-compliant pathology reports for non–small-cell lung cancer. Both digital pathology slides and pathology reports were included in the dataset, which were collected prospectively. To ensure the accuracy of the data, all pathology slides were reviewed by board-certified pathologists with expertise, and representative slides of primary tumors were chosen. Specifically, due to the lack of an appropriate FHIR profile for reporting cancer pathology, we adopted the mCODE concept in designing our FHIR profiles.

2.1. The design of FHIR profile

We looked at the structure of pathology reports and identified well-defined cancer-related vocabularies in the microscopic findings. To map these observations to a structured terminology, we used regular expressions for pattern matching. The pathology

reports are divided into three sections: pathological diagnosis, gross findings, and microscopic findings. The narrative descriptions of pathological diagnosis and gross findings were encoded as unstructured text, whereas microscopic findings were described using both free text and itemized forms. Cancer-related observations were labeled according to the AJCC 8th checklist and recorded as a series of paragraphs in the report's section on microscopic findings. The structure of the report template is shown in detail below.

- **Pathological diagnosis:** state the diagnostic results for tumor information.
- **Gross finding:** describing the specimen's status in the microscopy.
- **Microscopic findings:** describing the detailed histologic type, histologic pattern, total tumor size, cell type, size of invasive focus, margin, and non–tumorous-parenchyma.

We created multiple FHIR profiles for cancer observations, each with structured terms, and distributed them into an FHIR that includes FHIR resources covering critical data elements, value sets, extensions for supplementary elements, and a structured definition meant to incorporate other resources linked to the pathology report. The architecture of the proposed structured pathology report is depicted in Figure 1.

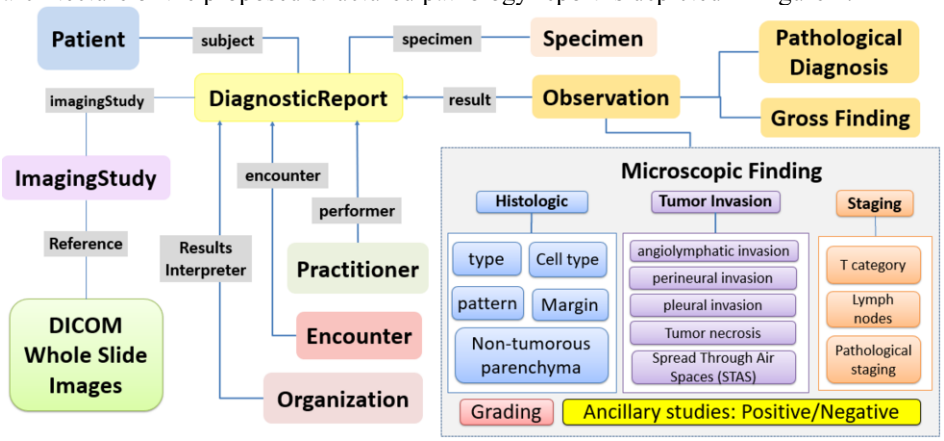

Figure. 1. The architecture of the FHIR profiles of the non–small-lung cancer pathology report.

3. Results

3.1. FHIR implementation guide

The FHIR implementation guide (IG) was created using FHIR Shorthand and includes an introduction, background description, logical model, FHIR profile, terminology system, search parameter, and capability statement. Based on previous stage results, the profile was created, including flags, cardinality, type, description, and constraints, as well as a completed differential table. The formulation of code terminology, which includes code systems, value sets, and concept maps between different coding systems, enabled the integration of commonly used medical vocabulary and code systems, such as Anatomic Pathology Lexicon, International Classification of Diseases (ICD), LOINC,

and SNOMED CT. The logical model outlines how to implement the profile, using the pathology report for non–small-cell lung cancer as an example. The resulting FHIR IG, which can be found at https://mitw.dicom.org.tw/IG/NSCLC/, includes 25 profiles for essential data elements, value sets, and a structured definition to facilitate the integration of clinical data and pathology images related to the pathology report. As shown in Figure 2, these profiles are included in an FHIR DiagnosticReport.

Code	Value	When For
Pathology report gross observation Narrative	The specimen received in formalin has 3 parts. Part (A) consists of a piece of lung, labeled RLL segment and measures 6.5 x 3.3 x1.8 cm. A gray white firm tumor, 1.2 x 0.5 x 0.5 cm, is noted 0.8 cm from the bronchial cut end, 0.1 cm from visceral pleura.Part (B) consists of groups of lymph nodes labeled as above.Part (C) is a piece of lung labeled 'RLL margin', measuring 3.8x 1.8 x 1 cm. Representative parts are taken for sections in 5 blocks. (A:bronchial and vascular cut end, B:stapled margin,C:tumor, D:specimen labeled 'RLL margin', E:interlobar LNs)(WKW)	2021-06-21
Pathology report microscopic observation Narrative Other stain	1. Histologic type: adenocarcinoma, micropapillary predominant 2. Histologic pattern: micropapillary (50%), papillary (25%), acinar (5%), lepidic (5%), solid (5%) 3. Cell type: non-mucinous tumor cells 4. Total tumor size: 4.5 x 2.5 x 1.8 5. Size of invasive focus: 4.4 cm 6. Tumor grading (WHO 2021): grade 3 (poorly differentiated) 7. Angiolymphatic invasion: present 8. Perineural invasion: absent 9. Spread Through Air Spaces (STAS): absent 10. Tumor necrosis: present 11. Pleural invasion: absent(PL0) (demonstrated with EVG) 12. Bronchus cut end: free of tumor involvement 13. Non-tumorous parenchyma: peribronchial chronic inflammation and anthracotic depositions 14. T category (based on the size of invasive focus): Tumor >4 cm but <= 5 cm in greatest dimension (pT2b) 15. Lymph nodes: all without metastatic tumor (N0) 16. Pathological staging: pT2bN0 (AJCC 8th edition). 17. Ancillary studies: - The tumor cells are positive for TTF-1; negative for p40. The multinucleated giant cells are negative for CK.	2021-06-21
Histologic type (observable entity)	adenocarcinoma,acinar predominant	2021-06-21
Pattern (attribute)	acinar(90%),lepidic(10%)	2021-06-21
Type of cell (attribute)	non-mucinous tumor cells	2021-06-21
Surgical margins (body structure)	free of tumor involvement	2021-06-21
Non-tumorous parenchyma	congestion	2021-06-21

Tumor size (observable entity)	0.6x04cm	2021-06-21
Depth of invasion by tumor (observable entity)	0.6 cm	2021-06-21
Status of venous (large vessel)/lymphatic (small vessel) invasion by tumor (observable entity)	absent	2021-06-21
Perineural spread (qualifier value)	absent	2021-06-21
Pleural invasion by tumor indeterminate (finding)	absent(PL0)	2021-06-21
Spread Through Air Spaces	absent	2021-06-21
Tumor necrosis (morphologic abnormality)	absent	2021-06-21
Primary tumor.pathology Cancer	tumor <=1 cm or less in greatest dimension(pT1a)	2021-06-21
Regional lymph nodes.pathology [Class] Cancer	all without metastatic tumor(pN0)	2021-06-21
Distant metastases.pathology [Class] Cancer	unknown	2021-06-21
Stage group.other Cancer	pT1aN0	2021-06-21
World Health Organization tumor classification (observable entity)	Grade group 2 (microscopic measurement)	2021-06-21
Ancillary studies	Tumor cell positive for TTF-1 and negative for p40;Multinucleate giant cell negative for CK	2021-06-21

Adenocarcinoma, acinar predominant, pT1aN0

Figure. 2. The result of DiagnosticReport for various profiles

The following is an example of a designed itemized text-based structured report defined in microscopic findings based on the AJCC 8th checklist. The report is processed through the defined structured report template, and the data elements are mapped to FHIR resources.

MICROSCOPIC FINDINGS:
1. Histologic type: Minimally invasive adenocarcinoma
2. Histologic pattern: lepidic (70%) and acinar (30%)
3. Cell type: non–mucinous-tumor cells
4. Total tumor size: 1.4 × 1.1 × 1.1 cm^3
5. Size of invasive focus: 0.4 cm
6. Tumor differentiation: well differentiated (G1)
7. Angiolymphatic invasion: absent
8. Perineural invasion: absent
9. Spread Through Air Spaces (STAS): absent
10. Tumor necrosis: absent
11. Pleural invasion: absent (PL0)
12. Resection margin: free of tumor involvement
13. Lymph nodes: all without metastatic tumor
14. Non–tumorous-parenchyma: congestion
15. Pathological staging: pT1miN0 (AJCC 8th edition).

4. Discussion

The terminology section is critical in ensuring that the medical vocabulary used in the pathology report covers all the value sets required for the design of the value sets in the system analysis stage. It is important to note that tumor, node, metastasis (TNM) staging varies for different types of cancers, and even within the same staging symbol, there may be slight differences in its meaning. Certain cancers may lack a staging system in some cases, necessitating the need to define the TNM staging of different cancers in the terminology system. Furthermore, determining whether the hospital uses its own coding book is critical, as this may affect whether subsequent value sets need to be customized. Additionally, each hospital may require additional code mapping during the FHIR conversion process. To improve the report's integration with its entire slide images, it is advisable to utilize the ImagingStudy resource to link the DICOM-encoded images. Finally, the proposed FHIR IG will be submitted for commercial implementation to the medical informatics Taiwan (MI-TW) Connectathon [6]. We encourage software developers to attend the MI-TW Connectathon in Taiwan in order to test their products and validate the IG's implementation and interoperability in commercially available systems.

5. Conclusions

We have created cancer-related profiles in the form of an FHIR IG, containing FHIR resources relevant to non–small-cell lung cancer pathology reports. Our method involved converting text-based pathology reports into structured formats and mapping the vocabulary items to FHIR observation resources. We are confident that our approach will enable the development of research repositories to advance precision medicine.

Acknowledgments

This work is primarily supported by Taiwan's National Science and Technology Council (NSTC 111-2634-F-006-012 and 111-2321-B-075-003).

References

[1] Patt, D., P. Stella, and L. Bosserman, Clinical Challenges and Opportunities with Current Electronic Health Records: Practicing Oncologists' Perspective. Journal of Oncology Practice, 2018. 14(10): p. 577-579.
[2] Saripalle, R., C. Runyan, and M. Russell, Using HL7 FHIR to achieve interoperability in patient health record. J Biomed Inform, 2019. 94: p. 103188.
[3] Osterman, T.J., M. Terry, and R.S. Miller, Improving Cancer Data Interoperability: The Promise of the Minimal Common Oncology Data Elements (mCODE) Initiative. JCO Clin Cancer Inform, 2020. 4: p. 993-1001.
[4] Guérin, J., et al., OSIRIS: A Minimum Data Set for Data Sharing and Interoperability in Oncology. JCO Clin Cancer Inform, 2021. 5: p. 256-265.
[5] Zong, N., et al., Modeling cancer clinical trials using HL7 FHIR to support downstream applications: A case study with colorectal cancer data. Int J Med Inform, 2021. 145: p. 104308.
[6] Medical Informatics Taiwan Connectathon. https://mitw.dicom.org.tw/