

Visualizing Imaging Pathways after Anatomy-Specific Follow-Up Imaging Recommendations

Thusitha Mabotuwana, Christopher S. Hall

Abstract—Radiologists routinely make follow-up imaging recommendations, usually based on established clinical practice guidelines, such as the Fleischner Society guidelines for managing lung nodules. In order to ensure optimal care, it is important to make guideline-compliant recommendations, and also for patients to follow-up on these imaging recommendations in a timely manner. However, determining such compliance rates after a specific finding has been observed usually requires many time-consuming manual steps. To address some of these limitations with current approaches, in this paper we discuss a methodology to automatically detect finding-specific follow-up recommendations from radiology reports and create a visualization for relevant subsequent exams showing the modality transitions. Nearly 5% of patients who had a lung related follow-up recommendation continued to have at least eight subsequent outpatient CT exams during a seven year period following the recommendation. Radiologist and section chiefs can use the proposed tool to better understand how a specific patient population is being managed, identify possible deviations from established guideline recommendations and have a patient-specific graphical representation of the imaging pathways for an abstract view of the overall treatment path thus far.

Keywords—Follow-up recommendations, care pathways, imaging pathway visualization, follow-up tracking.

I. INTRODUCTION

RADIOLOGY reports often contain follow-up imaging recommendations to monitor stability of potentially malignant findings, to ensure resolution of potentially serious disease, and for further diagnostic characterization [1]. Radiologists often make such recommendations based on established clinical practice guidelines. For instance, the Fleischner Society, an international medical society for thoracic radiology dedicated to diagnosis and treatment of chest related diseases, recommends that solitary pulmonary nodules that are 6-8 mm in diameter in patients at high-risk for primary lung cancer be followed-up with an initial follow-up CT at 6-12 months and again at 18-24 months if no change [2].

Radiologists' recommendation adherence to Fleischner Society guidelines for managing pulmonary nodules incidentally detected on CT exams has been reported to be high (82.8% in one study [3]), although patient (and/or referring physician) compliance to follow-up imaging recommendations has been reported to be low, with over 35% of follow-up imaging recommendations usually not followed-up [4]. Not all

recommendations need to be followed-up, but failure to comply with certain imaging follow-up recommendations in a timely manner can lead to 'upstaging' of malignancy, delayed treatment, poor patient outcomes, unnecessary testing, lost revenue, and legal liability [1], [5], [6]. Therefore, it is important to follow-up as recommended, and also for radiology administrators, insurers, and population health managers to ensure that specific patient populations are managed according to established guidelines. However, tools that can be used to visualize common imaging pathways are limited. In current practice, if a radiologist wants to determine what type of exams a patient has had following the detection of a specific finding (e.g., a lung nodule), the effort required is largely manual. To accomplish this, a radiologist will need to follow a series of manual steps, such as (1) review individual radiology reports (and/or patient medical history) to determine when the lung nodule was first mentioned, (2) determine related follow-up exams, and (3) determine if the imaging pathway the patient has followed is reasonable (e.g., determine if follow-up examinations were performed every 6-months per recommendations). This is a time-consuming effort, and not scalable. As such, the primary goal of this study is to present a methodology that can be used to visualize the imaging transitions at a population level as well as a patient level, following a specific finding of interest.

II. METHODS

A. Dataset

We extracted data for radiology exams performed between January 1 2010 and February 28 2017 from an academic radiology information system. The dataset contained 2,972,164 exams performed across multiple network hospitals. For each exam, the dataset contained the report text as well as several metadata fields, including exam code, exam date, radiology subspecialty, patient setting (inpatient, outpatient or emergency), patient ZIP Code, and modality. The human subjects division at the University approved this project as minimal risk. All unique patient identifiers (such as medical record numbers and accession numbers) were de-identified, while all date fields were shifted by a multiple of 7-days (to support any seasonality related analysis).

B. Previous Work: Follow-Up Detection

In previous research, we developed natural language processing based algorithms that can be used to detect follow-up recommendation sentences in radiology reports [7]. A report can contain one or more follow-up recommendations. While recommendations in radiology reports can be categorized into several classes (e.g., imaging recommendations, clinical or therapy follow-up, tissue sampling/biopsy and so on), and the algorithm distinguishes between nine different types of follow-up recommendations, the focus of the current work was on follow-up imaging recommendations only.

The follow-up detection algorithm first parses the radiology report to extract standardized sections (such as, “Clinical Indications”, “Findings” and “Impression”), paragraph headers within each section (e.g., “Abdomen” and “Pelvis” may appear under the “Findings” section) and sentences within each paragraph. The algorithm then evaluates each sentence within the “Findings” and “Impression” sections to determine if a sentence contains a follow-up imaging recommendation (e.g., “Follow-up CT chest recommended in 3 months”). Follow-up detection is performed using key-word searches and other heuristics. Output of this first step is a list of follow-up recommendation sentences (along with metadata, such as whether it is a negated sentence – e.g., “no further follow-up is necessary”). We evaluated the detection algorithm using 532 reports annotated by three board certified radiologists as the ground truth, to have 93.2% PPV (95% CI: 89.8-94.5%), 99.5 NPV (95% CI: 98.4-99.9%) and 97.9% accuracy (95% CI: 96.2-98.5%) [8]. The algorithm was subsequently improved based on detection errors identified in the test set so that 100% accuracy was achieved on the annotated dataset.

C. Anatomy Extraction

To identify the anatomy associated with a follow-up recommendation, an ontology based natural language processing engine we previously developed was used [9] along

with the publicly available NCBO annotation service [10]. We constrained queries to both services to extract anatomies as specified in SNOMED-CT ontology. Results were then merged and unique values selected. We used this approach to optimize the capabilities of the two systems, for instance, if the text contains “right lower lobe”, the anatomy engine would detect “Structure of right lower lobe of lung” corresponding to SNOMED ID 266005 whereas NCBO would not find a mapping. Conversely, from the sentence “hypervascular liver lesion, MRI follow-up is suggested”, NCBO detected “Liver Structure”, corresponding to SNOMED ID 10200004 whereas the internal engine identified “Lesion of liver” (SNOMED ID 300331000), which is a finding. Since our focus is on identifying anatomy, in this instance, the engine did not find any relevant anatomy since a longer phrase was already matched. The follow-up anatomy detection module first attempts to extract anatomy from the follow-up sentence – for instance, concept “Thoracic Structure” corresponding to ID 51185008 will be extracted from “Follow-up CT chest is recommended”. If no anatomy is identified in the follow-up sentence, the algorithm steps backwards from the follow-up sentence, processing one sentence at a time, until at least one anatomy is identified in a sentence. Search was restricted to the section in which the follow-up sentence occurred (which is usually ‘Findings’ and/or ‘Impression’ sections). Once identified, the ‘anatomy context’ becomes the text from the beginning of matched sentence to end of follow-up sentence. This process was repeated for all follow-up sentences when a report contained multiple recommendations. Table I shows four examples of extracted anatomy as reported in prior work [7]. For each follow-up recommendation, we also keep track of the previous two sentences, which is referred to as the ‘search context’. This search context can then be queried using regular expressions to detect the type of follow-up (e.g., whether follow-up recommendation is for a pulmonary nodule).

TABLE I
EXTRACTED ANATOMY FOR SEVERAL FOLLOW-UP RECOMMENDATION SENTENCES. (DETECTED FOLLOW-UP SENTENCE IS ITALICIZED) [7]

Anatomy Context	Extracted Anatomy	SNOMED-CT Description(s) and ID(s)
<i>These can be reassessed on CT <u>chest</u> for <u>lung</u> nodule follow-up</i>	chest lung	Thoracic Structure, 51185008 Entire lung, 181216001
<i>There is a <u>right adrenal</u> nodule which is likely benign and could be further evaluated by CT at the time of <u>lung</u> nodule follow-up.</i>	right adrenal lung	Entire right adrenal gland, 281625001 Entire lung, 181216001
<i>1 cm hypoechoic focal lesion in the mid portion of the <u>left kidney</u>. Although it is possible that it may represent a simple cyst, it is not adequately characterized on this study. <u>Recommend follow up US in 6 months to establish stability.</u></i>	left kidney	Left kidney structure, 18639004
<i>Nodular opacities in the <u>right lung</u> may represent infection versus aspiration. <u>Dedicated CT may be helpful.</u></i>	right lung	Right lung structure, 3341006

D. Clinical Use Case, Relevant Exam Identification and Data Visualization

We decided to focus first on three commonly occurring findings for which published follow-up guidelines exist: lung, thyroid and adrenal nodules. We used keywords *lesion*, *lump*, *nodule*, *tumor* and *mass* for all three, along with a few other descriptors at a finding-specific level. We included ‘opacity’ for lung findings; ‘hypodensity’ and ‘fullness’ for adrenal findings; and ‘hypodensity’ and ‘opacity’ for thyroid findings. We

required one of these nodule-related words to be within a six-word proximity (after removing stop words) of where the anatomy was detected within the anatomy context to ensure the finding was actually related to the detected anatomy.

To validate the algorithm’s ability to correctly determine the anatomy associated with a follow-up imaging recommendation, we manually selected a total of 200 reports – 50 reports for each of the three follow-up finding types as well as 50 reports that contained a follow-up recommendation but were unrelated to the lung, adrenal or thyroid. This was performed by searching

for the specific finding types in the “Findings” and “Impression” sections of randomly selected reports and repeating the process until the required dataset of 200 reports was created. The algorithm performance was 98.7% sensitivity (95% CI: 96.5-98.7%), 100% specificity (95% CI: 93.6-100%) and 99% accuracy (95% CI: 95.8-99%). There were two false-negatives, one related to an adrenal nodule and the other related to a lung nodule. A false-negative was defined as an instance where follow-up detection *or* anatomy extraction failed. Overall accuracy of 99% was slightly better than 97.9% follow-up detection accuracy reported previously since detection errors observed in the training set were rectified prior to anatomy extraction.

Once the anatomy associated with a follow-up recommendation is determined, the next step is to determine the relevant subsequent exams. For this, we used filters on the exam description, patient setting and the subspecialty. For instance, all lung, thyroid and/or adrenal related findings are reported on exams that have “chest”, “lung”, “adrenal” or “thyroid” as part of the exam description or read by subspecialties “body” or “chest”. We also restricted subsequent exams to outpatient exams only for patients who lived within a 50-mile radius of the hospital in order to avoid incorrectly counting patients who likely followed up outside the university system as “not following up”.

For imaging pathway visualization, we used Graphviz, which is an open source graph visualization software, with an application program interface [11]. A control program synthesizes the data extracted from the previous processing steps to create a representation using plain text graph description language (commonly referred to as the “DOT language”) of population and patient level modality transitions that can then be interpreted and rendered by the Graphviz engine.

III. RESULTS

There were 318,186 (10.7%) reports that had at least one follow-up imaging recommendation sentence. Of these, 29,310 exams contained at least one of the three specific finding types of interest (21,829 were lung related, 6,204 were thyroid related and 1,983 were adrenal related). The population level modality transitions following a lung related follow-up imaging recommendation are shown in Fig. 1, where a given transition occurred at least 400 times (the threshold can be adjusted as needed). The underscore followed by a numerical number in the fig. represents the transition since the particular finding, for instance, Pre_diagnosis_state_0 to CT_1 indicates that 9,610 patients had a lung nodule related follow-up first recommended after a CT exam. Similarly, CT_2 to XR_3 indicates that 560 patients had a radiograph as the third exam following a CT (first exam in this path could have been a CT or an XR per Fig. 1).

The results observed in Fig. 1 are largely as expected, since most lung nodules are detected on CT and radiograph exams. However, how many patients continue to have follow-up exams was largely unknown, and it was interesting to observe that out of the 21,829 lung-related findings, 474 patients continued to have at least eight subsequent outpatient CT exams (CT_3

through to CT_10) during the seven year period following the lung-related follow-up imaging recommendation.

Similarly, we created individual patient modality transition diagrams for each patient. Shown in Fig. 2 is the modality transition diagram for a patient who had the text “Small irregular-shaped nodule in the left upper lung which is new and cannot exclude a malignancy. Would consider CT imaging for further clarification” in the radiology report associated with the X-ray exam on 2/12/2010. As a result, this patient subsequently had a CT exam performed 14 days later where hyperinflated lungs with a nodule that could reflect nodular scarring were observed. As a result, a three-month follow-up exam was recommended to document stability. As shown along the arcs in Fig. 2, this CT exam was performed 100 days later (or 114 days later since diagnosis). Table II shows the corresponding exams for this patient. Date difference from previous exam as well as days since diagnosis are also shown which correspond to the values on the arcs in Fig. 2.

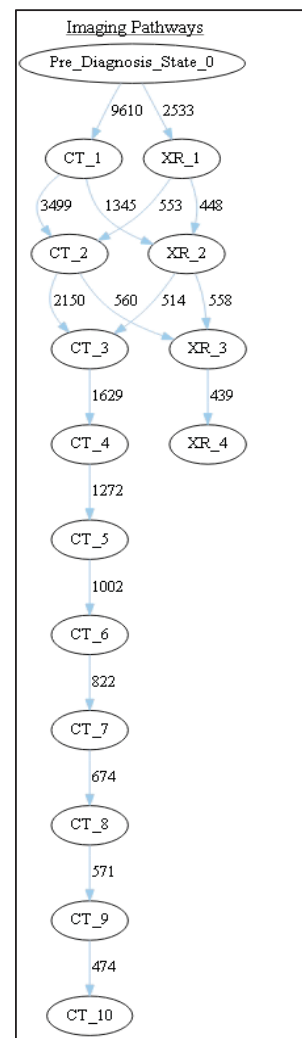


Fig. 1 Common imaging pathways after a lung nodule related follow-up recommendation. (The number on the arc represents the number of patients who had that transition)

TABLE II
SUBSEQUENT EXAMS AFTER A LUNG NODULE RECOMMENDATION

Modality	Exam Description	Shifted Exam Date	Date Difference	Days Since Finding
CR	XR CHEST 2VIEWS	2/12/2010		
CT	CT CHEST W CONTRAST	2/26/2010	14	14
CT	CT CHEST W CONTRAST	6/6/2010	100	114
CR	XR CHEST 2VIEWS	1/25/2011	233	347
CT	CT CHEST WO CONTRAST	2/3/2011	9	356
CR	XR CHEST 2VIEWS	4/5/2011	61	417
CR	XR CHEST 2VIEWS	9/20/2011	168	585
CT	CT CHEST WO CONTRAST	2/14/2012	147	732

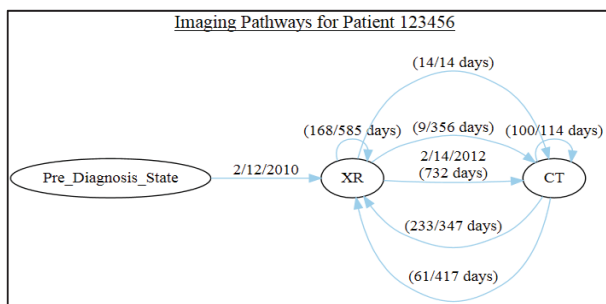


Fig. 2 Modality transitions after a lung nodule recommendation

Similarly, Table III shows an individual patient modality transition diagrams for a patient who had an ultrasound exam with a report on 3/25/2011 stating: “The left thyroid gland demonstrates two tiny cystic lesions. One of the lesions is complex and recommend continued surveillance in one year as these do not meet FNA criteria”. The patient had two subsequent lung nodule related CT exams in 2011, and none during 2012, as can be seen by the 709-day gap between the CT

on 12/14/2011 and radiograph on 11/22/2013. However, in 2013 this patient presents with cough for five days, pneumonia and “palpable inguinal adenopathy, in process of work up for lung nodules of uncertain origin, fevers and night sweat”. In a subsequent CT exam, it was reported that “Compared to 12/13/2013, there is a new right hilar mass, worrisome for lung cancer”. Being a more critical concern, lung related monitoring takes precedence, and although the patient gets cleared from the lung cancer concerns since the nodules have not increased in size, or resolved, the thyroid-related continued surveillance in one year recommendation is not carried out. This is an example of using the modality transition diagrams to quickly visualize the different states the patient has been in to understand the overall care pathway, and also identify cases where the follow-up exam was not performed in a timely manner.

TABLE III
SUBSEQUENT EXAMS AFTER A THYROID NODULE RECOMMENDATION

Modality	Exam Description	Shifted Exam Date	Date Difference	Days Since Finding
US	US THYROID	3/25/2011		
CT	CT CHEST WO CONT	6/8/2011	75	75
CT	CT CHEST WO CONT	12/14/2011	189	264
CR	XR CHEST 2 VIEWS	11/22/2013	709	973
CT	CT CHEST W CONT	12/13/2013	21	994
CR	XR CHEST 2 VIEWS	1/22/2014	40	1034
CT	CT CHEST WO CONT	5/12/2014	110	1144
CR	XR CHEST 2 VIEWS	1/7/2016	605	1749
CR	XR CHEST 2 VIEWS	2/14/2017	404	2153

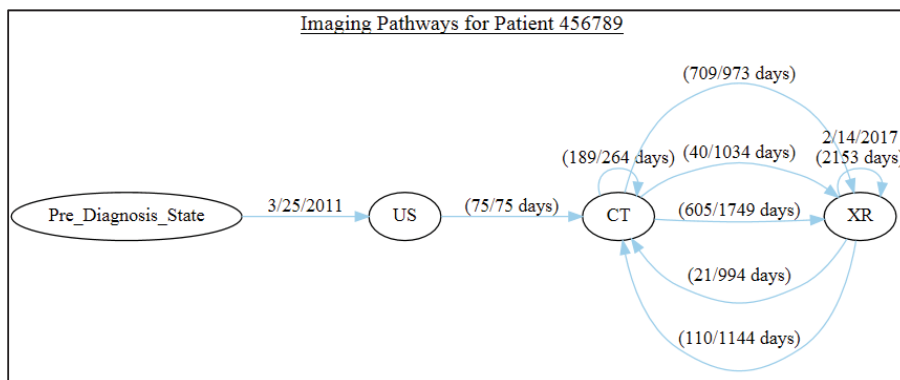


Fig. 3 Modality transitions after a thyroid nodule recommendation

IV. DISCUSSION

In this manuscript, we outlined a generic methodology to determine the modality transitions for a patient following a specific finding of interest. Using lung, adrenal and thyroid related findings, we demonstrated how the modality transitions can be visualized at a population level, as well as at a patient

level to understand how the population and/or individual patients are being managed. Accordingly, at a population levels such tools could possibly be used to ensure trends in the rate of guideline adherence, and to identify if there are any highly traversed secondary paths that are not guideline-compliant. We demonstrated two exemplary modality transition diagrams for

lung (Fig. 2) and thyroid (Fig. 3), and the same process can be used for other *events* as well, where the event could be some type of a finding. Using production data, we have also demonstrated a generic methodology that can be used to extract the finding/anatomy that is associated with a follow-up imaging recommendation.

As illustrated in Fig. 1, our results suggest that the treatment pathways at this particular institute were largely as expected, since most lung nodules are detected on CT and radiographs. However, if the threshold in Fig. 1 is reduced to 100 from 400, we observed that there were 257 instances where lung nodule follow-up was first recommended on a PET exam, and 178 instances where the recommendation was on an MRI exam. Out of the 257 patients who had a PET exam, 133 had a subsequent CT_2 exam (per Fig. 1). There were also 129 instances where a follow-up ultrasound exam was performed following CT_2. These pathways are still guideline compliant, but reducing the threshold further, to five for instance, showed that there were instances where a CT and an X-ray were performed on the same day starting from second exam, continuing through to 10th exam. Similarly, there were five instances where a CT and MR were performed on the same day starting from second exam through to 15th exam. These could be valid pathways and appropriate treatment for the individual patients, especially with complex patients, but whether these are strictly guideline compliant may need to be assessed on an individual basis.

Despite using a production dataset from three institutions, the current study has several limitations. First, the anatomy extraction may not always perform perfectly (algorithm failed to identify the correct anatomy in two out of the 200 reports we examined), and further validation may help improve reliability of the module. Complementing the ontology-based approach with domain-specific dictionaries (e.g., using a text-to-anatomy dictionary where ontology based approach fails) could be one option. Second, we only had access to exams that were performed within the network hospitals and therefore some exams may have been omitted from the modality transition diagrams, thereby potentially providing an inaccurate pathway for certain patients. Restricting the patient cohort to patients whose last known address was within a 50-mile radius from the hospital would have largely accounted for this limitation, but this would be an inherent limitation with almost any US-based clinical system where a patient record is not maintained at a national level or where there is sufficient “patient leakage” from one healthcare system to another. Third, we did not have access to mortality data which may be a surrogate measure of the patients who are not following up, especially toward the more distant nodes in Fig. 1.

The visualization we have presented is currently a static representation and may not necessarily be the most optimal representation of the data. Therefore, further validation is needed to ensure the utility of the tool in routine practice. To this end, we are currently exploring how such a tool can be integrated into workflow to aid radiologists and/or radiology administrators get better insights into how patients are managed following a finding of interest. Making the visualizations more interactive could be a possible extension of this work, for

instance, showing the radiology report text associated with each transition in the individual modality transition diagrams (Figs. 2 and 3) may be useful to understand the clinical context.

Given the number of time-consuming manual steps required to understand the modality transitions associated with patient management following a specific finding, new techniques that can aid radiologists and administrators with this process are needed. Having access to such tools that are also integrated into routine workflow can help identify guideline compliance, as well as systematic deviations, if any, ultimately contributing towards establishing more standardized patient care.

REFERENCES

- [1] Harvey, H. B., et al., *Correlation of the Strength of Recommendations for Additional Imaging to Adherence Rate and Diagnostic Yield*. J Am Coll Radiol, 2015. 12(10): p. 1016-22.
- [2] *Fleischner Society – Society for Thoracic Imaging and Diagnosis*. (cited 2017 Aug 10); Available from: <https://fleischnersociety.org/>.
- [3] Eisenberg, R. L. and S. Fleischner, *Ways to improve radiologists' adherence to Fleischner Society guidelines for management of pulmonary nodules*. J Am Coll Radiol, 2013. 10(6): p. 439-41.
- [4] Callen, J. L., et al., *Failure to follow-up test results for ambulatory patients: a systematic review*. J Gen Intern Med, 2012. 27(10): p. 1334-48.
- [5] Dutta, S., et al., *Automated detection using natural language processing of radiologists recommendations for additional imaging of incidental findings*. Ann Emerg Med, 2013. 62(2): p. 162-9.
- [6] Sloan, C. E., et al., *Assessment of follow-up completeness and notification preferences for imaging findings of possible cancer: what happens after radiologists submit their reports?* Acad Radiol, 2014. 21(12): p. 1579-86.
- [7] Mabotuwana, T., et al. *Extracting Follow-Up Recommendations and Associated Anatomy from Radiology Reports*. in *16th World Congress on Medical and Health Informatics (MedInfo2017)*. Aug 2017. Hangzhou.
- [8] Mabotuwana, T., et al., *Determining Adherence to Follow-up Imaging Recommendations*. J Am Coll Radiol, 2018. 15(3 Pt A): p. 422-428.
- [9] Prinsen, P., et al., *A Novel Approach for Improving the Recall of Concept Detection in Medical Documents Using Extended Ontologies*, in *Society for Imaging Informatics in Medicine Conference*. 2017: Pittsburgh.
- [10] Computing, N. C. f. B. *NCBO Annotator*. (cited 2017 Aug 11); Available from: <https://biportal.bioontology.org/annotator>.
- [11] *Graphviz - Graph Visualization Software*. (cited 2017 Aug 10); Available from: <http://www.graphviz.org/>.