Expanding the Scope of an Automated Radiology Recommendation-Tracking Engine: Initial Experiences and Lessons Learned

Mindy Y. Licurse¹ · Darco Lalevic¹ · Hanna M. Zafar^{1,2} · Mitchell D. Schnall^{1,2} · Tessa S. Cook¹

Published online: 10 November 2016
© Society for Imaging Informatics in Medicine 2016

Abstract An automated radiology recommendation-tracking engine for incidental focal masses in the liver, pancreas, kidneys, and adrenal glands was launched within our institution in July 2013. For 2 years, the majority of CT, MR, and US examination reports generated within our health system were mined by the engine. However, the need to expand the system beyond the initial four organs was soon identified. In July 2015, the second phase of the system was implemented and expanded to include additional anatomic structures in the abdomen and pelvis, as well as to provide non-radiology and non-imaging options for follow-up. The most frequent organs with incidental findings, outside of the original four, included the ovaries and the endometrium, which also correlated to the most frequently ordered imaging follow-up study of pelvic ultrasound and non-imaging follow-up study of endometrial biopsies, respectively. The second phase expansion has demonstrated new venues for augmenting and improving radiologist roles in optimal communication and management of incidental findings.

Keywords Communication · Data mining · Imaging informatics · Cancer detection

Background

Radiology practices are increasingly interested in tracking and appropriately communicating unexpected incidental findings on imaging studies, particularly those which may represent malignancy, so that appropriate therapy can be delivered in a timely fashion. Prior studies have demonstrated such indeterminate or suspicious lesions in up to 18 % of patients with no history of malignancy [1] and up to 31 % of patients with known malignancy [2]. A recent meta-analysis demonstrated that up to 37 % of patients undergoing CT colonography examinations had extracolonic findings that may have required additional work-up or had potential clinical significance [3]. Therefore, in an effort to improve and streamline the identification, communication, and follow-up of incidental indeterminate or suspicious findings on imaging, our institution implemented the first version of an automated radiology recommendation-tracking engine in July 2013 [4].

The initial scope of the internally designed and implemented engine was the monitoring of follow-up recommendations for focal masses in a subset of solid abdominal organs—the liver, pancreas, kidneys, and adrenal glands. These four organs were chosen because of the frequency and clinical significance of masses found in these organs [5], as well as in light of the initial white paper by the ACR Incidental Findings Committee in 2010 [6]. The engine depends on structured reporting used in combination with Code Abdomen (Table 1), a new lexicon and categorization system developed to be analogous to the Breast Imaging Reporting and Data System (BI-RADS). Code Abdomen enables the monitoring of patients with lesions indeterminate (category 0 or 3) or suspicious (category 4 or 5) for malignancy identified on computed tomography (CT), magnetic resonance imaging (MR), or ultrasound (US) examinations [4]. Category 99 was introduced a few months after the system was originally deployed,

✉ Mindy Y. Licurse
mindylicurse@gmail.com

¹ Department of Radiology, Hospital of the University of Pennsylvania, 3400 Spruce Street, 1 Silverstein Building, Philadelphia, PA 19146, USA

² Leonard Davis Institute of Health Economics, University of Pennsylvania, Philadelphia, PA, USA

Table 1 Code abdomen categorization scheme utilized by a radiologist to categorize focal lesions found in the abdominal organs in the report

Category	Description
0	Incompletely evaluated. If indicated within the patient's clinical context, follow-up INSERT MODALITY is advised.
1	No mass.
2	Benign. No further evaluation needed.
3	Indeterminate. Further imaging follow-up may be needed. If indicated within the patient's clinical context, follow-up INSERT MODALITY is advised within INSERT TIME FRAME.
4	Suspicious. May represent malignancy.
5	Highly suspicious. Clear imaging evidence of malignancy.
6	Known cancer.
7	Completely treated cancer.
99	Technically inadequate for evaluation of masses.

in order to provide an option for reporting unenhanced examinations in addition to technically inadequate images. Focal lesions still felt to be indeterminate or suspicious on unenhanced studies were reported with the corresponding 0, 3, 4, or 5 category. Follow-up recommendations are extracted from the study reports using natural language processing, and monitoring is performed in conjunction with the scheduling system to determine if a patient has scheduled, completed, or missed a follow-up imaging exam. During the first 2 years after its launch, users identified the need to provide follow-up recommendations that could not be captured by the scope of the original system. This need was also supported by the literature, given that incidental findings in other anatomic organs—such as the adnexa—are relatively common [7]. As such, the decision was made to expand the system to include categorization of findings in additional anatomic structures throughout both the abdomen and pelvis, as well as to offer non-radiology and non-imaging options for follow-up evaluation. The second stage of the system was implemented and launched in July 2015 and remains in use operationally at our institution. Key outcomes and lessons obtained since the implementation are discussed in this work.

Methods

The first phase was carried out from July 2013 through July 2015. These first 2 years' worth of data collected by the recommendation-tracking engine were analyzed to determine radiologists' compliance with the new structured templates, the overall distribution of indeterminate or suspicious findings identified, and the organ-specific breakdown of these indeterminate or suspicious findings. Perceptions of referring

physicians and non-physician providers receiving reports containing Code Abdomen were gathered using an online survey. Their responses were used to improve the system for the second stage. Specifically, specialists, and in particular, oncologists, felt that since they closely followed their patients' clinical course, they had a better appreciation of the necessity of following up incidental imaging findings. In light of this feedback, a new RECIST-like structured reporting template was developed to be used instead of Code Abdomen for oncology patients.

The first 2 years of system use also made it clear that there was a need to expand beyond the original four organs, given the relative frequency of incidental findings identified elsewhere, particularly in the pelvis and more specifically, in the adnexa. Therefore, the second stage update consisted of extending the recommendation structured template to include a number of pelvic organs (e.g., the uterus, ovaries, prostate, bladder, etc.), additional solid abdominal organs (e.g., the spleen), as well as hollow viscera (e.g., the small bowel and colon). The lung bases were also provided as an option since they are typically included in the scan region of abdominal CT examinations. Radiologists were now required to use the lexicon and template for all abdominal imaging examinations and non-musculoskeletal pelvic imaging performed with the same three modalities (CT, MR, and US). Furthermore, the update allowed for recommendation of non-radiology imaging examinations (such as direct visualization by endoscopy) as well as non-imaging (including office consultation, laboratory evaluation, and tissue sampling) follow-up options, as feedback from both radiologists and referring physicians indicated that in many instances, a follow-up radiology examination was not the most appropriate next step in the patient's management. All radiologists interpreting abdominal and pelvic exams were required to use Code Abdomen, with the exception of those interpreting musculoskeletal and pediatric imaging. Participation was tied to the faculty incentive, and email reminders were again sent when reports were non-compliant with the reporting requirements. However, no radiologist has yet to be penalized for insufficient use of the system.

For the second stage, organs beyond the original four in the abdomen only required assignment of a category if they required follow-up (i.e., contained indeterminate or suspicious lesions). Radiologists categorizing findings from such exams were required to pick from a list of 18 possible additional organs (including the bladder, peritoneum/mesentery, ovaries, etc.); the list included a choice of "other" for any anatomic structure not included on the provided list. A list of possible follow-up tests was also provided and offered the option to recommend a specific imaging follow-up examination (e.g., enhanced MRI abdomen with Eovist, enhanced MRI of the abdomen, pelvic ultrasound, small bowel follow through, etc.) or non-imaging follow-up (e.g., cytology, endometrial biopsy, etc.). While completion of imaging follow-up can be automatically monitored by querying an institutional data repository,

completion of non-imaging pathology follow-up is still being implemented.

Upon subsequent implementation of the second stage in July 2015, a subset of data between the time frame of July and November 2015 was further analyzed to determine what impact, if any, the additional categorization outside of the original four organs and follow-up recommendations had in terms of radiologist compliance (with use of the recently updated templates), number of indeterminate or suspicious findings identified, and imaging and non-imaging follow-up recommended. The subset of cases was taken from July 2015–November 2015 to account for all cases performed after the implementation of the second stage of the system up until the time considered reasonable for the system to be adapted proficiently by users. These cases are considered representative for the purpose of initial statistical analysis. Follow-up time interval recommendations were also reviewed to determine overarching trends and consistency among radiologists. Reports from 29 abdominal imagers were included in the analyzed dataset.

Results

In over 2 years of experience, we have determined that approximately 16 % of all abdominal imaging examinations (including inpatients, outpatients, and emergency department patients) receive follow-up recommendations for indeterminate or suspicious findings, with 47 % of these in the liver, 14 % in the pancreas, 28 % in the kidneys, and 12 % in the adrenal glands. Figure 1 demonstrates the overall distribution of exams categorized by this system over the 2-year period,

excluding those categorized as “1” or “2” with the corresponding categorization scheme explained in Table 1.

Between July and November 2015, indeterminate or suspicious findings in organs other than the four original abdominal structures were identified in 1049 of 16,725 examinations (6.5 %). Twenty-nine examinations had at least two or more indeterminate/suspicious findings. Figure 2 demonstrates a steep increase in the number of indeterminate and suspicious lesions categorized between July and August 2015. This increase was likely related to radiologists’ increasing familiarity with the updated system (both faculty and new trainees who started in July 2015) and adaptation to the availability of new organs and new follow-up options. The suggestion of decrease in the number of such lesions coded between October and November 2015 reflects a variability in the case mix from month-to-month rather than a decrease in usage of the system. Between July and November 2015, the most frequently categorized non-abdominal organs were the ovaries (316/1049 exams, 30 %), endometrium (196/1049 exams, 19 %), uterus (94/1049 exams, 9 %), lymph nodes (74/1049 exams, 7 %), and bones (64/1049 exams, 6 %) (Fig. 3).

Additionally, in July 2015, options for non-imaging follow-up evaluation were added in the form of a pick list, ranging from choices such as endometrial biopsy to microbiology. Between July and November 2015, 141 of 1049 abdominopelvic imaging examinations (13.4 %) categorized for indeterminate or suspicious lesions (outside of the initial four abdominal organs) recommended non-imaging follow-up, including cystoscopy, cytology, endoscopy, endoscopic retrograde cholangiopancreatography (ERCP), hysteroscopy, or laboratory testing such as urinalysis, office consultation, and tissue biopsy. The most frequently recommended non-imaging follow-up included endometrial biopsy (56/141

Fig. 1 Distribution of categories and organs from July 2013 to June 2015

Fig. 2 Number of exams coded for organs outside of the original four abdominal organs, between July and November 2015, after implementation of the updated system in July 2015

exams, 40 %), hysteroscopy (35/141 exams, 25 %), and office consultation (16/141 exams, 11 %) (Fig. 4).

Of all the indeterminate lesions (outside of the initial four major organs) coded during this time period, 26 % specified time intervals for follow-up in the following frequency: 7 % recommended 0–1-month follow-up, 44 % recommended >1–3-month follow-up, 24 % recommended >3–6-month follow-up, 22 % recommended >6–12-month follow-up, and 2 % recommended >12-month follow-up. However, the majority of indeterminate lesions had no specified time intervals.

Tracking compliance has also been a key feature of this system. After the original implementation in July 2013, first-pass compliance, defined as a report that was compliant immediately upon finalization by a faculty member, rose quickly from 48 % in July 2013 to 72 % in August 2013, stabilizing by mid-2014 between 93 and 99 %. Radiologists' compliance

with structured templates after the second stage implementation in July 2015 was high, with first-pass compliance (i.e., after the first time the report was finalized by a faculty member) as of November 2015 ranging between 95–98 % (Fig. 5). After initial implementation, it was found that radiologists required <1 min to incorporate the coding scheme into a single report [4]. This was essentially unchanged with the second version of the system.

Discussion

Overall, the second iteration of the automated tracking engine was well received. The near-100 % compliance rate suggests a high acceptance and overall usability of the system by radiologists at our institution. The 1049 examinations in which

Fig. 3 Most frequently categorized additional organs with intermediate or suspicious lesions

Fig. 4 Most frequently recommended non-imaging follow-up

indeterminate or suspicious findings were found validate the originally perceived need to expand the system beyond the liver, pancreas, kidneys, and adrenal glands.

Not surprisingly, the most frequently recommended additional imaging, for the ovaries and endometrium, corresponded with the most frequently recommended imaging follow-up exam—pelvic ultrasound—and the most frequently recommended non-imaging follow-up—endometrial biopsy and hysteroscopy. This is a prime example of how, in conjunction with evidence-based guidelines such as the Society of Radiologists in Ultrasound (SRU) recommendations for ovarian cysts published in 2010, radiologists are able to complement the care provided by their referring physician colleagues and provide clinically appropriate direction for the next steps in a patient's care. Since the initial white paper in 2010 on

incidental findings in the abdomen [6], subsequent additional white papers have been produced by the ACR on management of incidental findings within the adnexa, vasculature, spleen, lymph nodes, gallbladder, and biliary system as found by CT or MRI of the abdomen and pelvis. While these are not specified as medicolegal guidelines, they nonetheless give appropriate directional guidance for both radiologists and referring physicians [8–11]. Furthermore, tracking this data allows for continued analysis and fine-tuning of guiding principles with improved sensitivity and specificity.

Of note, many follow-up time intervals were not specified. While some of these may be due to categorization of exam findings with the underlying rationale that patients can be imaged by a different modality relatively quickly to make a more definitive diagnosis, this may also suggest possible lack

Fig. 5 First pass compliance rates (%) from July to November 2015 after implementation of the updated system in July 2015

of evidence-based guidelines for certain clinical scenarios, lack of awareness among radiologists regarding guidelines which may exist in practice, or lack of consensus among radiologists regarding appropriate follow-up time intervals. Whichever the dominant factor, this brings to attention the need for appropriate education or discussion for consensus to improve provided clinical guidance. On the other hand, it is also important to consider that a universally specified time interval for follow-up may not be easily determined for certain lesions and that without detailed knowledge of the clinical context of the patient, it may be difficult to choose an appropriate follow-up interval.

One limitation of the system is the inability to identify follow-up completion in a fully automated fashion. While this is feasible for imaging follow-up, it remains under development for non-imaging laboratory and pathology follow-up and will be tackled in a subsequent version for non-test-based clinical follow-up. Another limitation of the system is the current inability to quantify the degree of suspicion for malignancy associated with the indeterminate and suspicious categories. For example, when a radiologist determines that a breast lesion falls into BI-RADS category 3, he or she is conveying a less than 2 % chance of suspicion that the lesion is malignant. Associating the Code Abdomen categories with this type of quantitative data will require additional data collection, which is ongoing and mimics the development process also used for BI-RADS. Furthermore, while the system is a home-grown institutional system with no current plans to commercialize, reproduction at a multi-institutional level may be indicated in the future and is technically possible although potential obstacles may arise in adapting to variable EHR and RIS systems. While there are similar systems in existence such as the Alert Notification of Critical Results (ANCR), such systems have generally been designed with the primary objective focused on communication of critical imaging results rather than non-critical incidental findings [12]. Nonetheless, recent studies have demonstrated success of non-critical result communication [13].

As the automated recommendation-tracking engine continues to evolve, there is clear utility of the modified elements added during the second stage. The option of categorizing findings for organs outside of the original four abdominal organs as well as the options for non-imaging follow-up has been frequently utilized by radiologists in our health system and demonstrates that for a fully robust follow-up notification system to be in place, there is a necessary evolution of the radiologist's responsibility beyond solely reporting and monitoring imaging results. Radiologists do not practice in a vacuum and therefore, there must be appropriate tools in place for proper integrated clinical care. Additionally, optimal management of incidental findings can lead to a curbing of inappropriate utilization of imaging as well as ensure proper and timely follow-up for patients.

Conclusion

As radiology continues to move from a volume-based to value-based model [14], the necessity for an efficient, integrated, and complete follow-up recommendation monitoring system is vital. Systems such as our recommendation-tracking engine open up the possibilities to assist radiologists with increasing legal/legislative requirements for communication, decrease the number of patients who fail to get clinically indicated follow-up testing, and improve the diagnostic utility of radiology reports by including clinically relevant guidance, all of which will add value to patient care and improve quality metrics across the board. The implementation of categorizing findings outside of the four original major organs as well as non-imaging follow-up has augmented the role of radiologists at our institution as hands-on diagnosticians and will furthermore assist in setting a backdrop for useful quality performance metrics.

References

1. Kimberly JR, Phillips KC, Santago P, Perumpillichira J, Bechtold R, Pineau B, Vining D, Bloomfield RS: Extracolonic findings at virtual colonoscopy: an important consideration in asymptomatic colorectal cancer screening. *J Gen Intern Med* 24(1):69–73, 2009
2. Lardinois D, Weder W, Roudas M, von Schulthess GK, Tutic M, Moch H, Stahel RA, Steinert HC: Etiology of solitary extrapulmonary positron emission tomography and computed tomography findings in patients with lung cancer. *J Clin Oncol* 23(28):6846–53, 2005
3. Lin JS, Piper MA, Perdue LA, Rutter CM, Webber EM, O'Connor E, Smith N, Whitlock EP: Screening for colorectal cancer: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA* 315(23):2576–94, 2016
4. Zafar HM, Chadlavada SC, Kahn Jr, CE, Cook TS, Sloan CE, Lalevic D, Langlotz CP, Schnall MD: Code abdomen: an assessment coding scheme for abdominal imaging findings possibly representing cancer. *J Am Coll Radiol* 12(9):947–950, 2015
5. Berland LL: The American College of Radiology strategy for managing incidental findings on abdominal computed tomography. *Radiol Clin North Am* 49:237–243, 2011
6. Berland LL, Silverman SG, Gore RM, Mayo-Smith WW, Megibow AJ, Yee J, Brink JA, Baker ME, Federle MP, Foley WD, Francis IR, Herts BR, Israel GM, Krinsky G, Platt JF, Shuman WP, Taylor AJ: Managing Incidental Findings on Abdominal CT: White Paper of the ACR Incidental Findings Committee
7. Pickhardt PJ, Hanson ME: Incidental adnexal masses detected at low-dose unenhanced CT in asymptomatic women age 50 and older: implications for clinical management and ovarian cancer screening. *Radiology* 257(1):144–50, 2010
8. Managing Incidental Findings on Abdominal and Pelvic CT and MRI, Part 1: White Paper of the ACR Incidental Findings Committee II on Adnexal Findings
9. Managing Incidental Findings on Abdominal and Pelvic CT and MRI, Part 2: White Paper of the ACR Incidental Findings Committee II on Vascular Findings

10. Managing Incidental Findings on Abdominal and Pelvic CT and MRI, Part 3: White Paper of the ACR Incidental Findings Committee II on Splenic and Nodal Findings
11. Managing Incidental Findings on Abdominal and Pelvic CT and MRI, Part 4: White Paper of the ACR Incidental Findings Committee II on Gallbladder and Biliary Findings
12. Lacson R, O'Connor SD, Andriole KP, Prevedello LM, Ramin K: Automated critical test result notification system: architecture, design, and assessment of provider satisfaction. *AJR Am J Roentgenol* 203(5):W491–W496, 2014
13. O'Connor SD, Dalal AK, Sahni VA, Lacson R, Khorasani R: Does integrating nonurgent, clinically significant radiology alerts within the electronic health record impact closed-loop communication and follow-up? *J Am Med Inform Assoc* 23(2):333–8, 2016
14. Levin DC: The 2014 RSNA annual oration in diagnostic radiology: transitioning from volume-based to value-based practice-a meaningful goal for all radiologists or a meaningless platitude? *Radiology* 275(2):314–20, 2015