

Imaging for Appendicitis: Should Radiation-induced Cancer Risks Affect Modality Selection?¹

Sorapop Kiatpongsan, MD, PhD
Lesley Meng, MPH
Jonathan D. Eisenberg, BA
Maurice Herring, BA
Laura L. Avery, MD
Chung Yin Kong, PhD
Pari V. Pandharipande, MD, MPH

Purpose:

To compare life expectancy (LE) losses attributable to three imaging strategies for appendicitis in adults—computed tomography (CT), ultrasonography (US) followed by CT for negative or indeterminate US results, and magnetic resonance (MR) imaging—by using a decision-analytic model.

Materials and Methods:

In this model, for each imaging strategy, LE losses for 20-, 40-, and 65-year-old men and women were computed as a function of five key variables: baseline cohort LE, test performance, surgical mortality, risk of death from delayed diagnosis (missed appendicitis), and LE loss attributable to radiation-induced cancer death. Appendicitis prevalence, test performance, mortality rates from surgery and missed appendicitis, and radiation doses from CT were elicited from the published literature and institutional data. LE loss attributable to radiation exposure was projected by using a separate organ-specific model that accounted for anatomic coverage during a typical abdominopelvic CT examination. One- and two-way sensitivity analyses were performed to evaluate effects of model input variability on results.

Results:

Outcomes across imaging strategies differed minimally—for example, for 20-year-old men, corresponding LE losses were 5.8 days (MR imaging), 6.8 days (combined US and CT), and 8.2 days (CT). This order was sensitive to differences in test performance but was insensitive to variation in radiation-induced cancer deaths. For example, in the same cohort, MR imaging sensitivity had to be 91% at minimum (if specificity were 100%), and MR imaging specificity had to be 62% at minimum (if sensitivity were 100%) to incur the least LE loss. Conversely, LE loss attributable to radiation exposure would need to decrease by 74-fold for combined US and CT, instead of MR imaging, to incur the least LE loss.

Conclusion:

The specific imaging strategy used to diagnose appendicitis minimally affects outcomes. Paradigm shifts to MR imaging owing to concerns over radiation should be considered only if MR imaging test performance is very high.

©RSNA, 2014

¹From the Massachusetts General Hospital Institute for Technology Assessment, 101 Merrimac St, 10th Floor, Boston, MA 02114 (S.K., L.M., J.D.E., M.H., C.Y.K., P.V.P.); Department of Obstetrics and Gynecology, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand (S.K.); Department of Radiology, Massachusetts General Hospital, Boston, Mass (L.L.A., C.Y.K., P.V.P.); and Harvard Medical School, Boston, Mass (C.Y.K., P.V.P.). Received November 13, 2013; revision requested December 19; revision received April 1, 2014; accepted April 12; final version accepted April 18. P.V.P. supported by the Medical Imaging and Technology Alliance. Address correspondence to P.V.P. (e-mail: pari@mgh-ita.org).

The research content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

©RSNA, 2014

While appendicitis has long been the leading indication for emergent abdominal surgery, controversy remains regarding the best way to make this diagnosis (1–16). In prior decades, clinical evaluation alone informed the decision to perform surgery. With the advent and evolution of computed tomographic (CT) technology, CT quickly became the diagnostic mainstay owing to its high accuracy, reliability, and efficiency (4,6,12). More recently, concerns about radiation-induced cancer risks from CT have prompted practice drifts toward other imaging techniques that spare patients exposure to ionizing radiation (6). There are two primary alternatives: right lower quadrant ultrasonography

(US) and magnetic resonance (MR) imaging (2–5,7,11,15–23). At present, in most institutions, US is the first-line imaging test for children, and both US and MR imaging are preferred to CT in pregnant women (20,22–26).

The use of US and MR imaging to diagnose appendicitis in a general adult population is evolving. US is rarely used as a stand-alone modality; when US results are negative or indeterminate, follow-up CT is typically requested (6,11,24,27). The workflow demands of this two-part algorithm, and its dependence on operator and patient factors, challenge its reliable implementation in some emergency department settings (6). MR imaging is less frequently used than US, but is more accurate and provides greater anatomic coverage (2,3,11,15,16). In the largest, most methodologically rigorous evaluation of MR imaging to date, its reported sensitivity (97%) and specificity (93%) for the detection of appendicitis were comparable to—or exceeded—corresponding reported values for combined US and CT and CT alone (2–4,11,15,16). However, MR imaging is more expensive and is substantially more difficult to access in an emergent setting.

As institutions update emergency department imaging algorithms for the diagnosis of appendicitis, how should decision makers select the best imaging test, and to what extent should concerns over radiation exposure factor into this decision? We addressed this question by using decision-analytic methods to model the risks and benefits of imaging, including radiation-induced cancer risks. Our purpose was to compare life expectancy losses attributable to three imaging strategies for appendicitis in adults—CT, US followed

by CT for negative or indeterminate US results, and MR imaging—by using a decision-analytic model.

Materials and Methods

Model Overview

We developed a decision-analytic model to assess long-term health effects associated with three imaging strategies for patients suspected of having appendicitis: CT alone, combined US and CT, and MR imaging alone (Fig 1). Key model inputs—including the assumed diagnostic performance of each strategy—are included in Table 1. For selection of data sources informing test performance, priority was given to sources that synthesized test performance across multiple reported studies and to large prospective multicenter studies (4,11).

The life expectancy loss associated with each imaging strategy was our primary outcome measure. We incorporated expected days of life lost attributable to surgical mortality, missed appendicitis, radiation-induced cancers, and competing age- and sex-based mortality risks (4,11,13,29–32). Our base-case analysis was performed

Advances in Knowledge

- Long-term outcomes associated with the use of different imaging strategies for the work-up of appendicitis—including CT, combined US and CT, and MR imaging approaches—differ very minimally, even when accounting for radiation-induced cancer risks, with life expectancy (LE) losses for each strategy projected to be 8.2 days (CT alone), 6.8 days (combined US and CT), and 5.8 days (MR imaging).
- In our analysis, MR imaging was associated with the least LE loss, but required very high test performance characteristics to retain this standing—for example, sensitivity of 91% or greater and specificity of 62% or greater in 20-year-old men.
- Minimal LE differences across imaging strategies were more sensitive to changes in test performance than to varied assumptions about radiation-induced cancer risks and associated deaths—unless the LE loss attributable to radiation-induced cancers from CT decreased by approximately 74-fold relative to our initial assumption, the ranking of imaging strategies by LE loss would not change.

Implication for Patient Care

- In the clinical scenario of suspected appendicitis, paradigm shifts toward MR imaging owing to concerns about radiation exposure from CT should be pursued with caution and considered only if high MR imaging test performance is reliably achievable.

Published online before print

10.1148/radiol.14132629 Content codes: **GI HP**

Radiology 2014; 273:472–482

Abbreviations:

BEIR = Biologic Effects of Ionizing Radiation
 CI = confidence interval
 EAR = excess absolute risk
 ERR = excess relative risk

Author contributions:

Guarantors of integrity of entire study, S.K., L.L.A., P.V.P.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; manuscript final version approval, all authors; literature research, S.K., L.M., L.L.A., C.Y.K., P.V.P.; clinical studies, M.H., L.L.A.; experimental studies, S.K., L.L.A.; statistical analysis, S.K., L.M., L.L.A., C.Y.K., P.V.P.; and manuscript editing, S.K., L.M., M.H., L.L.A., C.Y.K., P.V.P.

Funding:

This research was supported by the National Institutes of Health (grants K07CA133097 and K25CA133141).

Conflicts of interest are listed at the end of this article.

Figure 1

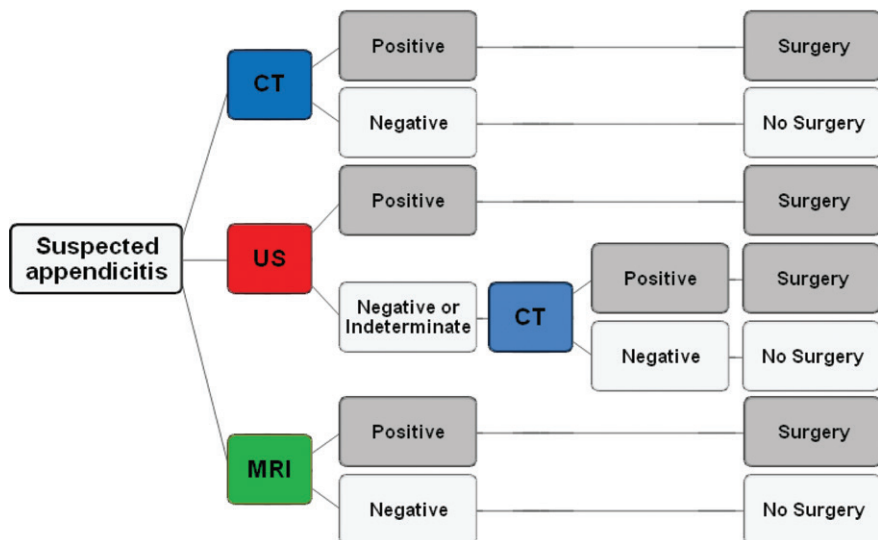


Figure 1: Flow diagram shows three imaging strategies for suspected acute appendicitis: (a) CT alone, (b) combined US and CT, and (c) MR imaging alone. Patients who underwent CT alone were triaged to surgery for appendectomy if they had positive results; patients with negative results did not undergo surgery. For the combined US and CT approach, patients underwent US, with triage to CT in circumstances of negative or indeterminate US results. Patients who underwent MR imaging alone were triaged to surgery for appendectomy if they had positive results; patients with negative results did not undergo surgery.

Table 1

Parameter Estimates for Base-Case and Sensitivity Analyses

Model Parameter and Source	Base Case	Sensitivity Range
Prevalence (%) (4,11,13,28)	45.5	10.0–90.0
Sensitivity of CT examination (%) (4)	94.0	50.0–100.0
Specificity of CT examination (%) (4)	94.0	50.0–100.0
Sensitivity of first-line US (%) (11)	77.0	50.0–100.0
Specificity of first-line US (%) (11)	94.0	50.0–100.0
Sensitivity of post-US CT (%) (11)	89.0	50.0–100.0
Specificity of post-US CT (%) (11)	97.0	50.0–100.0
Sensitivity of MR imaging (%) (11)	97.0	50.0–100.0
Specificity of MR imaging (%) (11)	93.0	50.0–100.0
Mortality from nondelayed laparotomy and appendectomy (%) (29)	0.05	0.0–0.1
Mortality from delayed laparotomy and appendectomy (%) (29)	0.31	0.1–1.0
Mortality from appendectomy in patients without appendicitis (%)	0.05	0.0–0.1
Life expectancy loss from one CT examination (d)	1.7	0.0–6.0
Life expectancy of men at age 20 (y) (30)	56.4	...
Life expectancy of women at age 20 (y) (30)	61.2	...
Life expectancy of men at age 40 (y) (30)	37.8	...
Life expectancy of women at age 40 (y) (30)	41.9	...
Life expectancy of men at age 65 (y) (30)	17.2	...
Life expectancy of women at age 65 (y) (30)	19.9	...

in a population of 20-year-old men; however, additional male and female cohorts (20-year-old women and 40- and 65-year-old men and women) were

further evaluated in sensitivity analysis. We did not incorporate differences in test performance for men and women, because adequate data are not available

to inform sex-specific differences for the performance of each imaging strategy being considered in the diagnosis of appendicitis. Additionally, adequately granular published data were not available to project the long-term mortality implications of detecting (or missing) conditions other than appendicitis on an age- and sex-specific basis. In women, the performance of each strategy for the detection of alternative causes of pelvic pain (eg, adnexal etiologies) may differ more than in men. Therefore, for parsimony, men were used in the base-case analysis. We elected to use a younger population because in this population radiation-induced cancer risks are higher, and therefore of greater concern.

We used commercially available software (TreeAge Pro 2009; TreeAge Software, Williamstown, Mass), to develop the primary decision-analytic model in which life expectancy losses were computed for each imaging strategy for appendicitis. We used a previously developed Markov model, programmed in C++, to project life expectancy losses associated with radiation-induced cancer risks from CT (32). These were then incorporated into the primary decision-analytic model. Below, we outline our model assumptions and approach to sensitivity analysis.

Imaging Strategies and Test Performance Characteristics

CT alone.—Patients who underwent CT alone for the diagnosis of appendicitis were triaged to surgery for appendectomy if they had positive results; patients with negative results did not undergo surgery. By definition, all patients in this strategy were exposed to ionizing radiation. Test performance characteristics for CT (94% sensitivity, 94% specificity) were elicited from a meta-analysis by Doria and colleagues (4) that represented pooled results from 21 studies (Table 1). To our knowledge, this source represents the most recent and comprehensive synthesis of related evidence to date.

Combined US and CT (US with triage to CT for negative or indeterminate results).—To model this strategy,

we needed to know US performance—and the performance of CT after US—to incorporate the conditional dependence of these sequential tests into our model. The largest study capable of informing these parameters was a 2013 multicenter study by Leeuwenburgh and colleagues (11). The authors performed a multicenter evaluation of 230 patients who underwent a protocol identical to our combined US and CT strategy, as well as unenhanced MR imaging, as further discussed below. On the basis of the data provided, we were able to extract (a) combined US and CT test performance (sensitivity = 97%, specificity = 91%) and (b) individual test performance metrics for US (sensitivity = 77%, specificity = 94%) and CT (sensitivity = 89%, specificity = 97%) when conditional on a negative or indeterminate US result (Table 1). For the base-case analysis only, combined US and CT test performance was required, but further breakdown of each test's contribution allowed us to individually explore different performance levels for each in sensitivity analysis.

MR imaging alone.—The multicenter study by Leeuwenburgh and colleagues (11) was also used to inform MR imaging test performance metrics (sensitivity = 97%, specificity = 93%) (Table 1). Notably, this study utilized a noncontrast (unenhanced) MR imaging protocol. Patients who underwent MR imaging alone were triaged to surgery for appendectomy if they had positive results; patients with negative results did not undergo surgery or CT (Fig 1).

Pretest probability (appendicitis prevalence).—The baseline prevalence of acute appendicitis in the target population was designated to be 45.5%, the midpoint of the range of reported prevalence values from four large studies (40%–51%) (4,11,13,28). Recognizing that disease prevalence assumptions are highly variable, this value was varied extensively in sensitivity analysis (Table 1).

Projecting Life Expectancy Loss in the Model

Life expectancy loss from the treatment of patients suspected of having appendicitis.—Mortality risks associated with

appendectomy, laparotomy, and missed appendicitis were informed by a recent large retrospective inpatient study of appendectomy outcomes (29). Within this study, the outcomes of open appendectomy informed our analysis (Table 1) (29). Mortality rates were reported for patients undergoing open appendectomy for nonperforated appendicitis ($n = 131\,172$) and for perforated appendicitis ($n = 68\,344$) (29). In our model, correctly diagnosed, true-positive cases were assumed to have a 0.05% attributable risk of death; the reported mortality rate of nonperforated appendicitis was used as a proxy for this value (29). Patients with missed appendicitis (false-negative cases) were assumed to have a higher attributable risk of death owing to complications resulting from delayed diagnosis, perforation, and treatment. Here, we used mortality from perforated appendicitis as a proxy (0.31%) (29). Patients suspected of having appendicitis at imaging but who did not have appendicitis at surgery (false-positive cases) were assumed to have the same mortality rate as true-positive cases (0.05%) in this base-case analysis; however, this assumption was tested further in sensitivity analysis. Patients who were given a correct diagnosis of not having appendicitis at imaging (true-negative cases) incurred no life expectancy loss unless they underwent CT, as detailed below.

Age- and sex-specific baseline life expectancies for all cohorts were drawn from publicly available life tables published by the U.S. Social Security Administration (Table 1) (30). For each imaging strategy, life expectancy loss attributable to the management of suspected appendicitis was computed as follows. On the basis of a combination of imaging test performance and disease prevalence (eg, likelihood of appendicitis), patients were placed into one of four result categories: true-positive, true-negative, false-positive, and false-negative. Life expectancy losses associated with each result category were computed as a product of the risk of death related to the management of suspected appendicitis and the cohort's life expectancy. For example, for

20-year-old men, false-negative results incurred a life expectancy loss of 0.0031 (0.31%) \cdot 56.4 years = 63.8 days. Life expectancy losses for each imaging strategy were calculated as a weighted average of losses associated with each result category. Weighting was based on the proportions of patients distributed across the four result categories and therefore reflected the diagnostic accuracy of each strategy.

Life expectancy loss from radiation-induced cancers.—Cohorts who underwent abdominopelvic CT incurred additional life expectancy loss because of radiation-induced cancers. On the basis of published data, we assumed that abdominopelvic CT had an effective dose of 8.3 mSv (32). To project life expectancy loss associated with typical CT coverage (lung bases to pubis) at this effective dose level, we adapted a previously described organ-specific Markov model, designed to project radiation-induced cancer deaths and life expectancy loss according to patient age, sex, and the specific anatomy imaged (32).

Briefly, this radiation-risk model (32) is based on key assumptions of the Biologic Effects of Ionizing Radiation (BEIR) VII report (31), including a linear no-threshold risk-exposure relationship for all solid cancers, a linear-quadratic relationship for leukemia, and suggested methods of cancer risk transport from Japanese atomic bomb survivors to a current U.S. population (31). The model accounts for exposure to 16 solid organs during abdominopelvic CT (lung, esophagus, stomach, pancreas, liver, colon, rectum, kidney, bladder, prostate, uterus, ovary, breast, central nervous system, thyroid, and oral cavity) and for leukemia, using organ-specific parameters from the BEIR VII report (31), from the report of Little et al (33), and from the report of Preston and colleagues (34). For most cancers, organ-specific risks of cancer death were estimated as a geometric mean of excess relative risk (ERR) and excess absolute risk (EAR), in keeping with BEIR VII methods (31). Exceptions included breast cancer, for which methods from Preston et al (34) were used and for which only an EAR

Table 2

Mean Life Expectancy Losses and Lifetime Risks of Radiation-induced Cancer Death Projected for a Single Abdominopelvic CT Study

Patient Age (y)	Life Expectancy Loss for Men (d)	Lifetime Risk of Radiation-induced Cancer Death per 100 000 for Men	Life Expectancy Loss for Women (d)	Lifetime Risk of Radiation-induced Cancer Death per 100 000 for Women
20	1.7 (0.9, 2.5)	33 (18, 49)	2.0 (1.3, 2.7)	37 (24, 50)
40	1.1 (0.6, 1.6)	24 (13, 35)	1.3 (0.9, 1.7)	26 (17, 35)
65	0.5 (0.3, 0.7)	18 (10, 26)	0.6 (0.4, 0.8)	19 (12, 26)

Note.—Data in parentheses are 95% CIs.

model was used, and thyroid and central nervous system cancers, for which only an ERR model was used (31–33).

Life expectancy losses attributable to an abdominopelvic CT study for cohorts of different ages and sexes were projected as follows. We first identified organ-specific equivalent doses that corresponded to an abdominopelvic CT study with an effective dose of 8.3 mSv. This was accomplished by using commercially available software (ImPACT CT, London, England), simulation data from phantom studies, and published tissue-weighting factors (32,35,36). These equivalent dose values were then used to compute risks of radiation-induced cancer death, on an organ-specific basis, by using risk functions (for EAR and ERR) and parameters from the BEIR VII investigators (31), Preston et al (34), and Little and colleagues (33). Using Markov modeling methods, we developed a mathematical model in which these aggregate organ-specific risks of radiation-induced cancer death—from a single exposure—could be modeled over a lifetime, and in which consequent effects on life expectancy could be computed (32). Notably, the model accounted for the cohort's age and sex at exposure, minimum latency periods between exposure and cancer death (2 years for leukemia, 5 for solid cancers [31,37]), the timing of cancer deaths over a cohort's lifetime, and age-sex all-cause mortality risks (32). By using this model, life expectancy losses attributable to abdominopelvic CT studies for cohorts of varying ages and sexes could be computed.

Projections of uncertainty for each estimate of life expectancy loss

attributable to radiation-induced cancers (with 95% confidence intervals [CIs]) were computed by using the Markov chain Monte Carlo method, as previously described in detail (32).

Life expectancy loss due to radiation-induced cancers from CT was projected by using the radiation-risk model—specific to patient age and sex—and was then incorporated for all strategy outcomes for which a CT study was used (true-positive, true-negative, false-positive, and false-negative). Thus, such losses were incorporated in the CT-alone and combined-US-and-CT strategies (when US results were negative or indeterminate). For example, in the above-mentioned 20-year-old male cohort, if a false-negative result were obtained after a CT study, then life expectancy loss attributable to the false-negative result would be calculated as follows: $[0.0031 (0.31\%) \cdot 56.4 \text{ years}] + \text{life expectancy loss attributable to radiation-induced cancers from one abdominopelvic CT study}$. As described previously, this value would then be combined in a weighted average with corresponding life expectancy losses attributable to other possible results (eg, true-positive) to determine the life expectancy loss attributable to the strategy of interest.

Sensitivity Analysis to Evaluate the Uncertainty of Strategy Preference

One-way and two-way deterministic sensitivity analyses were performed to evaluate the effects of varied parameter estimates on our model results—in particular, to determine whether the “preferred” strategy (defined as that with the lowest associated life expectancy loss) changed. Sensitivity ranges are

included in Table 1 alongside parameter estimates. Test performance characteristics associated with all strategies; the prevalence of appendicitis; and mortality risks from surgery, missed appendicitis, and radiation-induced cancers were varied extensively. Notably, in a two-way sensitivity analysis of test performance, we determined threshold sensitivity and specificity values by assuming 100% sensitivity and computing the minimum required specificity, and by assuming 100% specificity and computing the minimum required sensitivity. This allowed us to define a parameter space of combined sensitivity and specificity values over which the preferred strategy would be unchanged.

Results

Life Expectancy Losses Projected for a Single Abdominopelvic CT Study

The life expectancy loss and corresponding uncertainty estimate associated with radiation exposure from one abdominopelvic CT study for the base-case hypothetical cohort of 20-year-old men was 1.7 days (95% CI: 0.9, 2.5 days), corresponding to a lifetime probability of radiation-induced cancer death of 33 per 100 000 (95% CI: 18, 49 per 100 000). For 20-year-old women, these values were 2.0 days (95% CI: 1.3, 2.7 days) and 37 per 100 000 (95% CI: 24, 50 per 100 000), respectively. For all other cohorts, corresponding values are shown in Table 2.

Base-Case Analysis

In a hypothetical cohort of 20-year-old men suspected of having acute

Table 3

Life Expectancy Loss Associated with Each Imaging Strategy

Imaging Strategy	Total Projected Life Expectancy Loss (d)	Projected Life Expectancy Loss Attributable to Test Performance and Consequent Management (d)	Projected Life Expectancy Loss Attributable to Radiation-induced Cancers (d)
CT	8.18	6.48	1.71
CT and US	6.84	5.79	1.05
MR imaging	5.81	5.81	0.00

appendicitis, differences in life expectancy loss across imaging strategies were minimal: 5.8 days for MR imaging, 6.8 days for combined US and CT, and 8.2 days for CT alone (Table 3). For the combined US and CT strategy, 64% of hypothetical patients underwent follow-up CT. Radiation-induced cancers accounted for 1.1 days of total life expectancy loss in the combined US and CT strategy, versus 1.7 days when CT was used as a stand-alone strategy. An example is provided of our calculation of life expectancy loss for the CT-alone strategy in the base case—for a cohort of 20-year-old men, in keeping with the model structure defined in Figure 1 and the model parameters defined in Table 1, the projected life expectancy loss was calculated as follows: $\{56.4 \text{ years} \cdot 365 \text{ days/year} \cdot [(0.455 \cdot 0.94 \cdot 0.0005) + (0.455 \cdot 0.06 \cdot 0.0031) + ((1 - 0.455) \cdot 0.06 \cdot 0.0005)]\} + 1.7 \text{ days} = 8.2 \text{ days}$ of life expectancy loss.

Sensitivity Analysis

Test performance characteristics.—When concurrently varying MR imaging sensitivity and specificity, we found that only at very high levels of test performance was MR imaging associated with the least life expectancy loss across strategies (Fig 2). MR imaging sensitivity needed to be 91% at minimum (with 100% specificity), and MR imaging specificity needed to be 62% at minimum (with 100% sensitivity). When we concurrently varied the sensitivity and specificity of CT (as a stand-alone strategy) across specified ranges in two-way sensitivity analysis, and when we similarly concurrently varied test performance characteristics for US and for conditional CT (following negative

or indeterminate US results), the MR imaging strategy continued to have the least associated life expectancy loss, consistent with the base-case result.

Disease prevalence.—When we varied the prevalence of appendicitis across a wide range (10%–90%), MR imaging continued to be associated with the least life expectancy loss by a small margin, as depicted in Figure 3; there were no resulting changes in strategy rankings.

Mortality risks from surgery and delayed diagnosis.—We evaluated the dependency of our results on mortality risks from routine surgery (surgery prompted by true- or false-positive results) in two ways. First, we performed one-way sensitivity analyses in which surgical mortality risks were varied across specified ranges (Table 1) in true-positive and false-positive cases, respectively; MR imaging remained preferred. Second, we performed a two-way sensitivity analysis, treating mortality risks from true- and false-positive results as independent variables and varying them concurrently each across the same specified range. This two-way analysis allowed us to determine if differential surgical mortality risks in patients with appendicitis (true-positive) versus patients without appendicitis (false-positive) could affect our results. Nevertheless, across all tested mortality risk combinations, MR imaging remained preferred.

When we varied the mortality risk associated with a delayed diagnosis of appendicitis (false-negative result), MR imaging remained preferred.

Life expectancy loss from radiation-induced cancers.—Unless the life expectancy loss attributable to

Figure 2

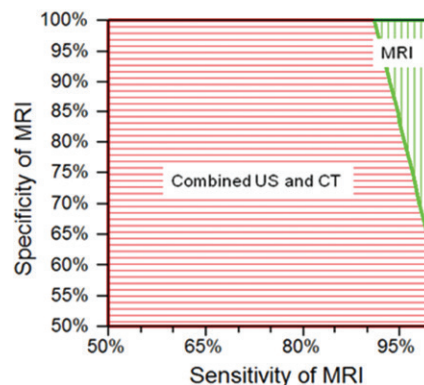


Figure 2: Graph shows results of two-way sensitivity analysis of MR imaging test performance characteristics. Two-way sensitivity analysis was performed by concurrently varying the sensitivity of MR imaging from 50% to 100% and the specificity from 50% to 100%. Vertically shaded area = combinations of MR imaging sensitivity and specificity for which MR imaging incurred the least life expectancy loss across strategies. Horizontally shaded area = combinations of MR imaging sensitivity and specificity for which MR imaging was no longer preferred and instead, combined US and CT yielded the least life expectancy loss. CT alone was not preferred in any of the circumstances tested in this sensitivity analysis. We found that MR imaging needed to have a sensitivity of at least 91% (with 100% specificity) and a specificity of at least 62% (with 100% sensitivity) to remain the preferred imaging strategy.

radiation-induced cancers from CT decreased to 0.023 days per CT study (a 74-fold decrease from the base-case value of 1.7 days), strategy rankings would not change (Fig 4). Were this to occur, the combined US and CT strategy would have the least associated life expectancy loss.

Age and sex.—Across all age and sex groups, MR imaging was the imaging strategy associated with the least life expectancy loss.

Discussion

Recently, reported test performance characteristics for the diagnosis of appendicitis at MR imaging have been highly favorable—comparable to, or even exceeding, those of CT-based strategies (2–4,11,15,16,21). However, even when accounting explicitly

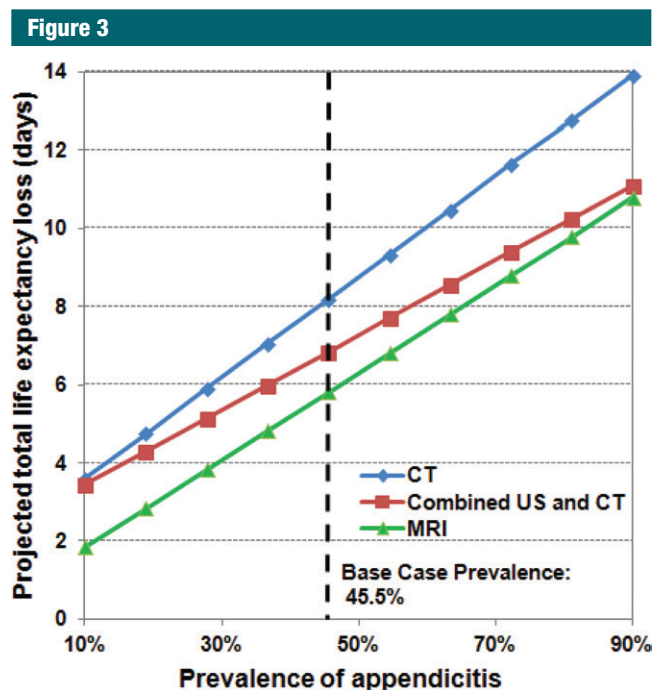


Figure 3: Graph shows results of one-way sensitivity analysis of the prevalence of acute appendicitis. In this analysis, we varied the prevalence of acute appendicitis from 10% to 90%. Across prevalence values in this range, MR imaging alone was always associated with the least life expectancy loss.

for radiation-induced cancer risks from CT, selection of MR imaging versus CT-based strategies affects long-term health outcomes only minimally. For example, in a cohort of 20-year-old men, projected life expectancy losses across all strategies (MR imaging alone, combined US and CT, and CT alone) varied by less than 3 days. In our model, on the basis of published data, we assumed that the sensitivity of MR imaging equaled or exceeded that of both CT strategies studied, whereas the specificity of MR imaging was lower than that of CT alone (4,11). False-negative results (driven by sensitivity) had relatively greater negative consequences than false-positive results (driven by specificity), which led to MR imaging's favorability over CT-based strategies from a life expectancy standpoint. Use of a combined US and CT strategy similarly resulted in small life expectancy gains over CT alone; its higher sensitivity, compared with CT alone,

similarly outweighed the negative consequences of its lower specificity. Notably, each imaging modality's test performance—and not their associated radiation-induced cancer risks—determined their relative ranking in terms of life expectancy loss.

What do the results mean for institutional policymakers—how does one make sense of these small, differential risks? In the setting of suspected appendicitis in adults, concerns over radiation exposure from CT should not directly translate into paradigm shifts toward MR imaging. For an institution to adopt MR imaging as a first-line strategy for the diagnosis of appendicitis, minimum test performance requirements are high. Institutions should consider their unique infrastructural challenges, including technologist and radiologist expertise as well as patient case mix, to ensure that MR imaging performance criteria can be reliably met to justify its use over CT-based strategies.

Some reports of MR imaging performance suggest that achievable sensitivity and specificity may not yet be high enough to warrant widespread paradigm shifts toward MR imaging (2,10,38). In a recent study by Chabanova and colleagues (2), none of three readers were able to achieve the concurrent sensitivity and specificity requirements delineated by our analysis that would be needed to justify the use of MR imaging over CT-based strategies in 20-year-old men (sensitivity range = 83%–93%, specificity range = 50%–83%). In a recent study of “nonexpert” readers evaluating appendicitis at MR imaging (38), reported performance values (sensitivity = 89%, specificity = 83%) similarly were not high enough to justify the use of MR imaging over CT-based strategies, according to our analysis results.

That being said, it is possible to invest in infrastructure and training to improve and sustain a high level of MR imaging performance. We selected a multicenter study by Leeuwenburgh and colleagues (Optimizing Imaging in Suspected Appendicitis, or OPTIMAP [39]) to inform test performance characteristics in our model on the basis of the high level of methodologic rigor and study power. In six centers in the Netherlands that participated, paradigm shifts away from CT as a stand-alone strategy were already in place—CT alone was not a comparator arm (39). The same group previously demonstrated the potential to improve MR imaging interpretation skills for appendicitis through training, achieving statistically significant sensitivity gains of 10% (82% vs 92% pre- and posttraining); analogous changes in specificity were not achieved (10). Notably, in that study, neither the pre- nor the posttraining performance metrics would have been sufficient to justify MR imaging over CT-based strategies for 20-year-old men (10).

Use of a combined US and CT approach for the evaluation of appendicitis—over CT alone—has already disseminated into many practice settings,

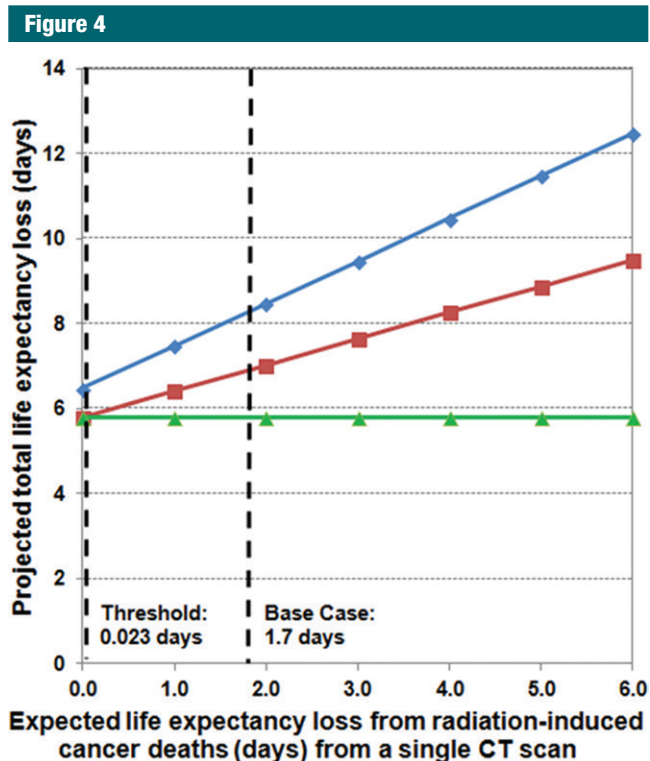


Figure 4: Graph shows results of sensitivity analysis of projected days of life expectancy loss from radiation-induced cancer deaths. We varied the projected days of life expectancy loss attributable to radiation-induced cancers from CT from 0 to 6 days. At the base-case value of 1.7 days, MR imaging alone was the preferred strategy. For life expectancy losses attributable to radiation-induced cancers less than a threshold value of 0.023 days, combined US and CT was the preferred strategy.

with a larger body of evidence available to support its use (4,5,11). Our analysis supports a combined US and CT approach, but similarly implies that if there are concerns over achievable test performance levels at a given institution, then a stand-alone CT strategy remains a reasonable option given minimal projected excess harms. It is important to recognize that while US may seem to be a “no-harm” modality when used in a combined approach with CT, this is not necessarily true; if achievable specificity is low, and false-positive rates are substantially higher than with CT alone, combined US and CT may incur more harm than benefit compared with CT alone. Therefore, prior to adopting a combined US and CT approach, close attention to optimizing test performance characteristics is similarly advised.

Our approach had specific limitations that merit consideration. First, by definition, decision analysis is a method that requires simplification of complex clinical and disease courses to construct a representative mathematical model. This limitation was addressed by ensuring that key model variables that may influence outcomes—and for which robust data were unavailable—were incorporated and varied in sensitivity analysis. For example, to inform the mortality risk of false-negative results and consequent delayed surgery, we used as a proxy mortality associated with open appendectomy for perforated appendicitis. This may represent an overestimation of risk, as not all patients with a delayed diagnosis will experience perforated appendicitis; however, varying

this estimate across a wide range did not meaningfully change our results, as verified in sensitivity analysis.

We deliberately excluded consideration of alternative diagnoses for abdominal pain in our analysis, a related limitation. In an analysis of this type, a top priority is to ensure that parallel advantages and disadvantages are incorporated across strategies, so as not to unfairly bias toward one or another. In the case of alternative diagnoses, US, CT, and MR imaging have different profiles of alternate diagnoses that may be readily detected. For example, US will be superior for detecting ectopic pregnancy, MR imaging will be superior for detecting biliary diseases, and CT will allow for more sensitive detection of urinary tract stones and acute bowel conditions. Because there is insufficient evidence to model how the differential detection of alternate conditions across imaging strategies may affect long-term patient outcomes, we did not incorporate alternative diagnoses into our analysis.

We also did not account for allergic reactions and contrast material-induced nephropathy from intravenously administered iodinated CT contrast materials in our analysis. On the basis of our study design, to account for these, we would have needed to incorporate deaths from each. Deaths from anaphylaxis are exceedingly rare (1:170000) and would not change our results meaningfully (40,41). While deaths attributable to contrast-induced nephropathy are also assumed to be extremely low, their accurate estimation is challenging given the paucity of related evidence and controversy about the magnitude of association between contrast material administration, acute kidney injury, and related death (40,42–44). Because we did not explicitly account for deaths in patients at risk for contrast-induced nephropathy, our analysis is best applied to populations who are not considered to be at risk for contrast-induced nephropathy. For patients who are at risk, a separate analysis would be appropriate, comparing noncontrast CT strategies with MR imaging.

Second, patient characteristics also factor into modality performance, particularly in the case of MR imaging. For example, more sequence and parameter customization may be needed for patients who are poor breath holders, who have difficulty tolerating the more demanding nature of MR imaging examinations, or who are obese. Our analysis does not account for these factors. That being said, with current MR imaging technologies and appropriately applied expertise, much can be done to optimize and shorten examinations and to overcome described barriers (3,15,39).

Third, we did not account for workflow considerations from the perspective of the emergency department. MR imaging studies take longer than CT studies, and most designated MR imaging slots—if available on an emergent basis—are used for head and spine imaging, indications for which the preference of MR imaging over other modalities is more compelling (eg, stroke, cord compression). An appendicitis case may not be prioritized in relation to neurologic MR imaging cases, which may result in longer wait times for patients with abdominal pain. Data that project the effects of queuing issues, to our knowledge, are not yet available. However, at the institutional level, workflow barriers that may compromise patient outcomes should be considered carefully prior to any paradigm shifts, particularly given the relatively small health benefits gained by the use of one strategy over another.

Fourth, radiation-induced cancer risks from imaging remain uncertain (31,33,45–47). While there is increasing empirical evidence to demonstrate an exposure-risk relationship, its precise nature remains unknown (48,49). We used a previously published model (32), which drew largely from assumptions and parameters put forth by the BEIR VII report (31); however, most empirical data from the BEIR VII report draws from the Japanese atomic bomb survivor experience, which represents a very different exposure circumstance (31). We addressed the uncertainty surrounding radiation risk projections

in sensitivity analysis, where we evaluated effects of radiation exposure throughout a range that included zero—and that exceeded life expectancy losses conceivable from the highest levels of exposure from a single CT study—computing effects of these different risk conditions on the relative favorability of different imaging strategies. While basic assumptions underlying our radiation-risk model are consistent with current standards (31,33,46,47), it should be noted that the type of risk information generated is appropriate for population-level, but not patient-level, decision making.

Fifth, we did not account for costs in our analysis. Given the minimal life expectancy gains achievable by MR imaging relative to CT-based strategies, MR imaging is not likely to be cost-effective in this setting from a societal standpoint. However, in a rigorous cost-effectiveness analysis, long-term health costs consequent to the initial imaging decision are considered, not just work-up costs. Hence, the relative cost-effectiveness of MR imaging cannot be accurately inferred from an analysis like ours. Future related efforts should consider the cost implications of MR imaging in this setting, ideally from a lifetime perspective.

While radiation exposure from CT has prompted consideration of alternate imaging modalities for appendicitis, life expectancy differences across imaging strategies are minimal and are driven more by test performance than by radiation-induced cancer risks. If an institution is capable of providing an MR imaging service that affords consistently high test performance, then MR imaging may be the best strategy. However, if institutional MR imaging performance is expected to be inconsistent or appreciably lower than that of combined US and CT or CT alone, then a CT-based strategy should still be used. In this scenario, the benefits gained by averting radiation exposure are unlikely to meaningfully outweigh the risks of poorer test performance, which is a larger driver of patients' long-term outcomes. While

our findings may seem to contradict current paradigm shifts toward lower-risk imaging strategies, to identify the lowest risk approach, risks of both pursuing a given approach under consideration—and of not pursuing that approach—must be weighed together. Viewing risks through this larger lens enables an objective understanding of the impact of radiation exposure from CT in the clinical setting of appendicitis, hopefully providing quantitative evidence to guide institutional policymakers who are considering practice changes.

Disclosures of Conflicts of Interest: **S.K.** No relevant conflicts of interest to disclose. **L.M.** No relevant conflicts of interest to disclose. **J.D.E.** No relevant conflicts of interest to disclose. **M.H.** No relevant conflicts of interest to disclose. **L.L.A.** No relevant conflicts of interest to disclose. **C.Y.K.** No relevant conflicts of interest to disclose. **P.V.P.** Financial activities related to the present article: none to disclose. Financial activities not related to the present article: has received research funding from the Medical Imaging and Technology Alliance. Other relationships: none to disclose.

References

1. Brenner DJ. Slowing the increase in the population dose resulting from CT scans. *Radiat Res* 2010;174(6):809–815.
2. Chabanova E, Balslev I, Achiam M, et al. Unenhanced MR imaging in adults with clinically suspected acute appendicitis. *Eur J Radiol* 2011;79(2):206–210.
3. Cobben L, Groot I, Kingma L, Coerkamp E, Puylaert J, Blickman J. A simple MRI protocol in patients with clinically suspected appendicitis: results in 138 patients and effect on outcome of appendectomy. *Eur Radiol* 2009;19(5):1175–1183.
4. Doria AS, Moineddin R, Kellenberger CJ, et al. US or CT for diagnosis of appendicitis in children and adults? a meta-analysis. *Radiology* 2006;241(1):83–94.
5. Gaitini D, Beck-Razi N, Mor-Yosef D, et al. Diagnosing acute appendicitis in adults: accuracy of color Doppler sonography and MDCT compared with surgery and clinical follow-up. *AJR Am J Roentgenol* 2008;190(5):1300–1306.
6. Hernanz-Schulman M. CT and US in the diagnosis of appendicitis: an argument for CT. *Radiology* 2010;255(1):3–7.
7. Keyzer C, Zalman M, De Maertelaer V, et al. Comparison of US and unenhanced multi-

- detector row CT in patients suspected of having acute appendicitis. *Radiology* 2005;236(2):527-534.
8. Kim K, Kim YH, Kim SY, et al. Low-dose abdominal CT for evaluating suspected appendicitis. *N Engl J Med* 2012;366(17):1596-1605.
 9. Lai V, Chan WC, Lau HY, Yeung TW, Wong YC, Yuen MK. Diagnostic power of various computed tomography signs in diagnosing acute appendicitis. *Clin Imaging* 2012;36(1):29-34.
 10. Leeuwenburgh MM, Wiarda BM, Bipat S, et al. Acute appendicitis on abdominal MR images: training readers to improve diagnostic accuracy. *Radiology* 2012;264(2):455-463.
 11. Leeuwenburgh MM, Wiarda BM, Wiezer MJ, et al. Comparison of imaging strategies with conditional contrast-enhanced CT and unenhanced MR imaging in patients suspected of having appendicitis: a multicenter diagnostic performance study. *Radiology* 2013;268(1):135-143.
 12. Raja AS, Wright C, Sodickson AD, et al. Negative appendectomy rate in the era of CT: an 18-year perspective. *Radiology* 2010;256(2):460-465.
 13. van Randen A, Bipat S, Zwinderman AH, Ubbink DT, Stoker J, Boermeester MA. Acute appendicitis: meta-analysis of diagnostic performance of CT and graded compression US related to prevalence of disease. *Radiology* 2008;249(1):97-106.
 14. Wan MJ, Krahn M, Ungar WJ, et al. Acute appendicitis in young children: cost-effectiveness of US versus CT in diagnosis—a Markov decision analytic model. *Radiology* 2009;250(2):378-386.
 15. Zhu B, Zhang B, Li M, Xi S, Yu D, Ding Y. An evaluation of a superfast MRI sequence in the diagnosis of suspected acute appendicitis. *Quant Imaging Med Surg* 2012;2(4):280-287.
 16. Inci E, Hocaoglu E, Aydin S, et al. Efficiency of unenhanced MRI in the diagnosis of acute appendicitis: comparison with Alvarado scoring system and histopathological results. *Eur J Radiol* 2011;80(2):253-258.
 17. Heverhagen JT, Pfestroff K, Heverhagen AE, Kloese KJ, Kessler K, Sitter H. Diagnostic accuracy of magnetic resonance imaging: a prospective evaluation of patients with suspected appendicitis (diamond). *J Magn Reson Imaging* 2012;35(3):617-623.
 18. Lee JH, Jeong YK, Park KB, Park JK, Jeong AK, Hwang JC. Operator-dependent techniques for graded compression sonography to detect the appendix and diagnose acute appendicitis. *AJR Am J Roentgenol* 2005;184(1):91-97.
 19. Leeuwenburgh MM, Jensch S, Gratama JW, et al. MRI features associated with acute appendicitis. *Eur Radiol* 2014;24(1):214-222.
 20. Long SS, Long C, Lai H, Macura KJ. Imaging strategies for right lower quadrant pain in pregnancy. *AJR Am J Roentgenol* 2011;196(1):4-12.
 21. Moore MM, Gustas CN, Choudhary AK, et al. MRI for clinically suspected pediatric appendicitis: an implemented program. *Pediatr Radiol* 2012;42(9):1056-1063.
 22. Pedrosa I, Lafornera M, Pandharipande PV, Goldsmith JD, Rofsky NM. Pregnant patients suspected of having acute appendicitis: effect of MR imaging on negative laparotomy rate and appendiceal perforation rate. *Radiology* 2009;250(3):749-757.
 23. Pedrosa I, Levine D, Eyvazzadeh AD, Siewert B, Ngo L, Rofsky NM. MR imaging evaluation of acute appendicitis in pregnancy. *Radiology* 2006;238(3):891-899.
 24. Krishnamoorthi R, Ramarajan N, Wang NE, et al. Effectiveness of a staged US and CT protocol for the diagnosis of pediatric appendicitis: reducing radiation exposure in the age of ALARA. *Radiology* 2011;259(1):231-239.
 25. Spalluto LB, Woodfield CA, DeBenedictis CM, Lazarus E. MR imaging evaluation of abdominal pain during pregnancy: appendicitis and other nonobstetric causes. *RadioGraphics* 2012;32(2):317-334.
 26. Cogley JR, O'Connor SC, Houshyar R, Al Dulaimy K. Emergent pediatric US: what every radiologist should know. *RadioGraphics* 2012;32(3):651-665.
 27. Poletti PA, Platon A, De Perrot T, et al. Acute appendicitis: prospective evaluation of a diagnostic algorithm integrating ultrasound and low-dose CT to reduce the need of standard CT. *Eur Radiol* 2011;21(12):2558-2566.
 28. Terasawa T, Blackmore CC, Bent S, Kohlwes RJ. Systematic review: computed tomography and ultrasonography to detect acute appendicitis in adults and adolescents. *Ann Intern Med* 2004;141(7):537-546.
 29. Masoomi H, Mills S, Dolich MO, et al. Comparison of outcomes of laparoscopic versus open appendectomy in adults: data from the Nationwide Inpatient Sample (NIS), 2006-2008. *J Gastrointest Surg* 2011;15(12):2226-2231.
 30. Period life table. Social Security Administration. <http://www.ssa.gov/oact/STATS/table4c6.html>. Published 2007. Accessed March 25, 2012.
 31. National Research Council (U.S.). Advisory Committee on the Biological Effects of Ionizing Radiation. Health risks from exposure to low levels of ionizing radiation, BEIR VII Phase 2. Washington, DC: National Academy of Sciences, 2006.
 32. Pandharipande PV, Eisenberg JD, Lee RJ, et al. Patients with testicular cancer undergoing CT surveillance demonstrate a pitfall of radiation-induced cancer risk estimates: the timing paradox. *Radiology* 2013;266(3):896-904.
 33. Little MP, Wakeford R, Tawn EJ, Bouffler SD, Berrington de Gonzalez A. Risks associated with low doses and low dose rates of ionizing radiation: why linearity may be (almost) the best we can do. *Radiology* 2009;251(1):6-12.
 34. Preston DL, Mattsson A, Holmberg E, Shore R, Hildreth NG, Boice JD Jr. Radiation effects on breast cancer risk: a pooled analysis of eight cohorts. *Radiat Res* 2002;158(2):220-235.
 35. International Commission on Radiological Protection. 1990 recommendations of the International Commission on Radiological Protection: user's edition. Oxford, England: Pergamon, 1992.
 36. Jones DG, Shrimpton PC. NRPB-SR250: Normalized organ doses for x-ray computed tomography calculated using Monte Carlo techniques. Health Protection Agency, UK. <http://www.hpa.org.uk/Publications/Radiation/NPRBArchive/NRPBSoftware>. Accessed October 27, 2012.
 37. Office of Radiation and Indoor Air. Modifying EPA Radiation Risk Models Based on BEIR VII, Draft White Paper. Washington, DC: U.S. Environmental Protection Agency, 2006.
 38. Leeuwenburgh MM, Wiarda BM, Jensch S, et al. Accuracy and interobserver agreement between MR-non-expert radiologists and MR-experts in reading MRI for suspected appendicitis. *Eur J Radiol* 2014;83(1):103-110.
 39. Leeuwenburgh MM, Laméris W, van Randen A, et al. Optimizing imaging in suspected appendicitis (OPTIMAP-study): a multicenter diagnostic accuracy study of MRI in patients with suspected acute appendicitis. *BMC Emerg Med* 2010;10:19.
 40. Halvorsen RA. Which study when? iodinated contrast-enhanced CT versus gado-

- linium-enhanced MR imaging. *Radiology* 2008;249(1):9-15.
41. Morcos SK, Thomsen HS. Adverse reactions to iodinated contrast media. *Eur Radiol* 2001;11(7):1267-1275.
42. McDonald JS, McDonald RJ, Comin J, et al. Frequency of acute kidney injury following intravenous contrast medium administration: a systematic review and meta-analysis. *Radiology* 2013;267(1):119-128.
43. McDonald RJ, McDonald JS, Bida JP, et al. Intravenous contrast material-induced nephropathy: causal or coincident phenomenon? *Radiology* 2013;267(1):106-118.
44. McDonald JS, McDonald RJ, Carter RE, Katzberg RW, Kallmes DF, Williamson EE. Risk of intravenous contrast material-mediated acute kidney injury: a propensity score-matched study stratified by baseline-estimated glomerular filtration rate. *Radiology* 2014;271(1):65-73.
45. Little MP, Heidenreich WF, Moolgavkar SH, Schöllnberger H, Thomas DC. Systems biological and mechanistic modelling of radiation-induced cancer. *Radiat Environ Biophys* 2008;47(1):39-47.
46. Preston DL, Pierce DA, Shimizu Y, Ron E, Mabuchi K. Dose response and temporal patterns of radiation-associated solid cancer risks. *Health Phys* 2003;85(1):43-46.
47. Preston RJ. Update on linear non-threshold dose-response model and implications for diagnostic radiology procedures. *Health Phys* 2008;95(5):541-546.
48. Mathews JD, Forsythe AV, Brady Z, et al. Cancer risk in 680,000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. *BMJ* 2013;346:f2360.
49. Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet* 2012;380(9840):499-505.