換気血流シンチとヘリカルCTの肺塞栓診断特性のメタ分析比較に関する研究

(医療技術評価総合研究事業) 分担研究報告書

分担研究者　野口善令　　所属　名古屋第二赤十字病院
研究協力者　林野泰明　　所属　京都大学医学部大学院
研究協力者　福井次矢　　所属　聖路加国際病院

研究要旨：
Purpose
To perform a meta-analysis of the literatures on helical computed tomography (hCT) and ventilation-perfusion (V/Q) scan for the detection of acute pulmonary embolism (PE) using the methodological tool of summary receiver operating characteristic (SROC) curve analysis.
Materials and Methods
Studies in MEDLINE and EMBASE database published from January 1985 for V/Q scan and January 1990 for hCT to March 2003 were included if (a) these tests were performed for the evaluation of acute PE (b) conventional angiography was the reference standards, and (c) absolute numbers of true positive, false-negative, true-negative, false-positive results were available. Two investigators independently abstracted data sets. Sensitivity analysis was conducted by excluding articles published before 1995.
Results
Data from 12 articles were included. With a random effects model, pooled sensitivity for hCT was 86.0% [95% CI: 80.2%, 92.1%] with specificity of 93.7% [95% CI: 91.1%, 96.3%]. V/Q scan yielded low sensitivity of 39.0% [95% CI: 37.3%, 40.8%] but high specificity of 97.1% [95% CI: 96.0%, 98.3%] with high probability threshold, while it yielded the high of 98.3% [95% CI: 97.2%, 99.5%] and low specificity of 4.8% [95% CI: 4.7%, 4.9%] with normal/near-normal threshold. Regression coefficient for hCT versus V/Q scan with high probability threshold was 0.588 [95%CI: -1.55, 2.74], and 4.14 [95%CI: -0.002 to 8.28] with normal/near-normal threshold.
Conclusion
To exclude PE with low pretest probability, hCT was superior to V/Q scan using high probability threshold; we can use hCT and V/Q scan using normal/near-normal threshold comparably to confirm PE with moderate-to-high pretest probability.

A. 研究目的

The diagnosis of pulmonary embolism (PE) remains elusive and difficult, because the symptoms are non-specific and all available tests have substantial limitations in clinical settings (1, 2). Selective pulmonary angiography remains the gold-standard diagnostic test, but this procedure is invasive, causing morbidity and mortality rates of 4% and 0.2%, respectively (3, (1). Moreover, it is also costly and time-consuming. Several minimally invasive modalities have been used to detect PE. The ventilation-perfusion (V/Q) scan is one of the most frequently used procedures for confirming or excluding diagnosis of PE because of its general availability. The Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED) study assessed the usefulness of V/Q lung scans in acute pulmonary embolism (5). Despite accuracy as high as 96% in some instances, the V/Q scan alone is inadequate because up to 75% of patients fall in the inconclusive category, therefore other modalities, such as serial compression ultrasonography and D-dimer testing, are useful adjuncts (6, 7).

In recent years, contrast agent-enhanced helical computed tomography (hCT) of the pulmonary arteries has been proposed, and data are accumulating (8-17). The choice between V/Q scan and hCT should be determined by comparing their diagnostic accuracy using the available data. Interpretation of these data as a whole is difficult because of the wide variation in the background of the patients. In prior review articles comparing the test performance of hCT and V/Q (18,

19), the difference in test performance has not been systematically compared by statistical method frequently used in meta-analysis on diagnostic tests.

Summary receiver operating characteristic (SROC) analysis is a method that enables quantitative combination of the multiple studies with heterogeneous results; each point on the summary ROC curve represents a combination of sensitivity and specificity that could result from each study (20-22). Reasons for heterogeneity among different studies is due to differences between the way how clinicians define a "positive" test or wide variation in terms of the patients' background, and summary ROC curve analysis could resolve those problems by stringent inclusion criteria and meta-regression analysis. It is suitable to compare different diagnostic tests' performance, and reports using this analysis has recently been accumulating (23, 24). Thus, the purpose of our study was to perform a meta-analysis of the literatures on hCT and V/Q scan for the detection of acute PE using the methodological tool of SROC analysis.

B. 研究方法

Identification of trials

A computerized search was performed to identify relevant English language articles published in MEDLINE. In searching for articles on hCT, we combined "computed tomography" in text words, "image interpretation, computer-assisted" in medical subject headings, and "pulmonary embolism" in medical subject headings or text words. Similarly, we searched for articles on perfusion scan using "radionuclide imaging" in medical subject headings, and "perfusion scan" or "ventilation-perfusion lung scan" as text words. The hCT search was limited to papers between January 1, 1990 and May 30, 2003, since earlier CT equipment was substantially different in technology. We included papers on perfusion scan from January 1st, 1985 since this technique began to be used for the diagnosis of pulmonary embolism from the mid 1980s. We also scanned the references in retrieved articles and contacted knowledgeable individuals in radiology to see if they knew of any other relevant reports. A database search using a similar strategy was conducted in the EMBASE database from 1985 to 2003. We did not search for unpublished reports.

Inclusion criteria

We included a study provided that 1) it used hCT or perfusion scan as a diagnostic tool for acute pulmonary embolism; 2) it gave absolute numbers of true-positive (TP), false positive (FP), true-negative (TN), and false negative (FN) cases or their equivalent; 3) it used pulmonary angiography as the reference standard for diagnosing pulmonary embolism; and 4) the time interval between the findings obtained from the test and reference standard was 48 hours or less taking into account that pulmonary emboli might disappear during the interval of two tests. We determined this time interval by discussing with radiologists and reviewing literatures of diagnostic tests on pulmonary embolism (2).

A study was excluded if 1) it used pulmonary angiography, in combination with any other reference standard, as the reference standard; 2) hCT was not performed for acute pulmonary embolism (e.g., chronic pulmonary embolism or septic embolism); 3) non-comparable CT methods (e.g., electron-beam CT) were used; 4) hCT was performed after anticoagulant therapy or surgery for pulmonary embolism; or 5) the published information was incomplete. We tried to examine if different studies from the same institutions used the same group of patients, because more than one studies could be published by one author,

Data collection

Two investigators (Y.H., M.G.) independently abstracted the data from all articles included in our analysis. The information abstracted included descriptive data (authors, title, journal citation, year of publication), study group characteristics (sample size,

mean age, proportion of women, prevalence of pulmonary embolism), study design characteristics (involving the criteria used to define a positive result and protocol information), extent of blinding between readers of hCT, perfusion scan and reference standard, the information about the extent of the disease, and any evidence of verification bias and test interpretation bias. For each study the results were classified as TP, FP, TN, and FN. For the V/Q scan, PIOPED criteria were generally used according to the probability of PE: high probability, intermediate probability, low probability, near-normal/normal (3). We specified three thresholds for calculating sensitivity and specificity (Table 1): high probability perfusion scan as a positive and others (intermediate probability, low probability, near-normal/normal) negative (threshold 1, high probability threshold); high and intermediate perfusion scan probability as a positive and low probability and near-normal/normal scan as a negative (threshold 2, high/intermediate threshold); near-normal/normal scan as a negative and others (high probability, intermediate probability, low probability) positive (threshold 3, normal/near-normal threshold). One study developed and employed original criteria consisting of five categories, which were reduced to four categories to match PIOPED criteria (4). As for hCT, the presence or absence of pulmonary emboli during the test defined as an intraluminal filling defect or complete nonfilling of a pulmonary artery was used as a positive or negative criterion.

From the articles in which investigators tabulated the results for different observers, we extracted the data of the first observer, unless one of them was emphasized in the literature. When authors emphasized one result of different observers, e.g. level of experience, we extracted the most emphasized results. Any inconsistencies or controversies encountered in abstracted data were resolved through discussion and consensus.

Analysis and statistics

The overall suitability of the pooled and summary ROC curve analysis was evaluated using the Spearman correlation coefficient (5). We then checked heterogeneity separately for sensitivity and specificity. Since sensitivities for hCT and specificities for V/Q scan threshold 3 were not homogeneous ($p=0.006$ and $p<0.0001$, respectively), pooled sensitivity (TP/TP+FN) and specificity (TN/FP+TN) estimates were calculated using a random-effects model that weighted each report by its sample size (6).

To estimate the summary ROC curve for hCT and perfusion scan, we used a previously described method of variance-weighted least square regression (20, 22, 23, 28). Based on the 2×2 table constructed from each report, we made a logit transformation of the TP rate (sensitivity) and FP rate (1-specificity). Differences in the logit transformations (measures of the observed discriminatory power of hCT and perfusion scan) were then regressed on the sums of the logit transformations (measure of the positivity threshold used for the determination of a positive hCT and perfusion scan result). Summary ROC curves for hCT and perfusion scan were constructed by back transformation of the fitted line from the regression model. We weighted each study in the regression model by its variance (1/(TP+0.5)+ 1/(FP+0.5)+ 1/(FN+0.5)+ 1/(TN+0.5)) (7), and restricted the final summary ROC curves to the range of observed TP and FP rates.

Adjustment for clinical variables was accomplished by including them in the regression model. Inclusion of a dummy variable in the regression analysis for the type of diagnostic examination performed (1 for hCT, 0 for perfusion scan) then allows comparison of tests. The regression coefficient of this dummy variable is a measure of the difference in discriminatory power between the examinations. A positive regression coefficient implies increased discriminatory power for hCT compared to perfusion scan, and a negative regression coefficient implies reduced discriminatory power. To avoid undefined values for positivity criteria and their variance that arise from zeroes of the TP, FN,

TN, or FP values, 0.5 was added to that value (7).

We assessed the effect of publication year, mean age (55 years or younger vs. older than 55 years), prevalence of pulmonary embolism, duration of tests (<24 vs. <48 hours), study design (prospective vs. retrospective), presence of interpretation bias, and presence of verification bias (presence of verification bias vs. no available information) in combined model of hCT and V/Q scan. We could not consider the effect of the extent of the disease in the model comparing hCT and V/Q scan, since the information about the extent of the disease was not available in the literatures mainly dealing with V/Q scan (3, 4, 8). We dichotomized some variables (age, %women, and duration between tests) at median. Due to availability of data (e.g. data on collimation were only available for hCT) or missing data, the following variables were analyzed separately in each separate model; % of woman (=<25% vs. >25%), collimation (3mm or thinner vs. thicker than 3mm), size of pulmonary emboli (segmental vs. sub-segmental) for hCT model; type of ventilation scan radionuclide (99mTc-DTPA vs. other) for V/Q scan model. In the combined model, univariate analyses were performed to assess the effect of each clinical covariate. We incorporated the factors that had p<.20 in univariate analysis into a multivariate regression model, and then used backward elimination to remove variables with p-value equal to or greater than 0.05. For the main aim of this study, the dummy variable for type of diagnostic test (hCT=1) were always kept in this process. In both separate model and combined model, data on V/Q scan were treated separately in different thresholds (e.g., hCT vs. V/Q threshold1, hCT vs. V/Q threshold2, hCT vs. V/Q threshold3). Finally, we re-analyzed the final model with a random effects regression analysis (STATA technical bulletin no.42, College Station, Tex; Stata Statistical Software), which took inter-study variability as well as intra-study variability into account.

After sensitivities and specificities were pooled, we assessed the post-test probability of PE with different pre-test probabilities (low=0.03, moderate=0.27, high=0.78). The arbitrary pretest probability of 0.03, 0.27, and 0.78 were based on the Well's report in which pretest probability was determined using clinical signs and symptoms (9). First, pretest odds were converted into post-test odds, by means of the following formula: post-test odds = pretest odds * likelihood ratio. Likelihood ratio is defined as the probability of the test result in people with the disease divided by the probability of the result in people without disease. Post-test odds were converted back to posttest probabilities (10).

Since hCT methodology for detecting PE has undergone a remarkably rapid advance over the past decade, the test performance might be different between early 1990's and 2003. Therefore, we performed a sensitivity analysis by excluding articles published before 1995 on hCT, and compared the results with those in base-case analysis. All analyses were done using commercially available software (Intercooled STATA 7.0; STATA Corporation, TX, USA).

C. 研究結果

Summary of the literature review and data extraction

Our initial data search yielded a total of 1385 titles for studies using hCT or V/Q scan. We then excluded 1306 articles by reviewing titles and abstracts selected possible 79 articles for our analysis; the reasons for exclusion were summarized in Fig. 1. Of these, we identified 12 articles which met all the inclusion criteria (5, 8-11, 13-17, 27, 29). Two of these articles also reported test performance of both hCT and V/Q scan (10, 14), thus we had 9 studies for hCT and 5 studies for V/Q scan. Two articles were reported by the same lead author (8, 11), but these two studies were judged not to be overlapped because of the different study periods and different amount of contrast medium used. Overall, there were discrepancies in 19 out of 140 (13.6%) extracted data between the two authors, ranging from 0% to 35.7% depending on the sort of items extracted. All discrepancies were resolved by consensus.

hCT and ventilation-perfusion scan

The technique of helical CT was used to diagnose pulmonary embolism for the first time in 1992 (11). Of 12 articles we identified, 9 studies collectively reported on 520 subjects (range 10-151 per study) on hCT (12-19), while 5 studies reported on 1269 patients (range 20-731) on V/Q scan (3, 4, 8, 13, 16).

Variations in study protocols included the thickness of scan section, from 2.5mm to 5mm, and the amount of contrast medium varied from 70ml to 150ml. As for detector system, single-detector hCT was used in 8 reports and dual-detector hCT was used in only one report (19). The reported sensitivity of hCT ranged from 53% (15) to 100% (12), and the specificity from 75% (19) to 100% (12).

Ventilation studies used as nuclear isotope either Xenon-133 gas, Tc-99m-PYP or Tc-99m-DTPA and perfusion scanning used TC-99m-MAA. The PIOPED criteria was used in 4 studies, and original criteria were formulated in one study (4). The reported specificity of V/Q scan using threshold 1 ranged from 96.0% (3) to 100% (13, 16), with sensitivity were 54.5% (13) to 100% (16) or 98% (3) to 100% (4, 8) based on thresholds 2 and 3, respectively. Details of the articles included are summarized in Table 2.

Weighted Pooled Results

Weighted pooled data are presented in Table 3. Pooled sensitivity for hCT was 86.0% [95% CI: 80.2%, 92.1%] with specificity of 93.7% [95% CI: 91.1%, 96.3%]. V/Q scan yielded low sensitivity, 39.0% [95%CI: 37.3%, 40.8%] with threshold 1, while it yielded high sensitivities, 86.0 % [95%CI: 83.3%, 88.8%] and 98.4% [95%CI: 97.2%, 99.5%], respectively, with thresholds 2 and 3. V/Q scan yielded high specificity, 97.1% [95%CI: 96.0%, 98.3%] with threshold 1; sensitivities were 39.0% [95%CI: 37.3%, 40.8%] and 4.8% [95%CI: 4.7%, 4.9%], respectively, with thresholds 2 and 3.

SROC Analysis

No significant predictors were found in the separate univariate analysis for hCT (% of women, collimation, size of pulmonary emboli) or V/Q scan (type of radionuclide for ventilation scan). The univariate analysis for the comparison of two tests revealed that only the following variables of those stated in the method section had a p-value smaller than .20; age (p=0.032) in the model including V/Q scan threshold 1; duration between two tests (0.027) and presence of verification bias (0.067) in the model including V/Q scan using threshold 2; duration between two tests (p=0.111) and presence of verification bias (p=0.146) in the model including V/Q scan threshold 3. Therefore, we included these variables in each multivariate model; however, as a result of backward elimination, no significant predictors were kept in the final model. n a model comparing performance of hCT to V/Q scan using thresholds 3, the estimate was large enough to suggest the superiority of hCT over V/Q scan thought it was not statistically significant (β-coefficient 4.14 [95% CI: -0.002, 8.28]). In the comparison between hCT and V/Q scan using thresholds 1, no significant difference was observed in terms of discriminatory power (β-coefficient 0.588 [95% CI: -1.55, 2.74]). In a model comparing hCT and perfusion scan threshold 2, hCT displayed superior discriminatory power (β-coefficient 3.73 [95% CI: 2.56, 4.9]). IThese results are shown in Table 4 and Fig. 2.

Post-test probability of PE

Post-test probabilities of PE according to positive or negative results of each test were summarized in Table 5. When pretest probabilities were moderate (0.27) or high (0.78), post-test probabilities on positive results of V/Q scan using threshold 1 and hCT were same and 84.1%and 98%, respectively. When pretest probability was low (0.03), post-test probabilities of negative hCT was 0.05, and that of negative V/Q scan using threshold 3 was 0.01; when pretest probabilities were moderate and high, post-test probabilities of negative hCT or V/Q scan with any threshold were not low enough to exclude PE.

Sensitivity Analysis

Results of sensitivity analysis were shown in Table 3 and Table 4. By excluding the articles before 1995, sensitivity (true positive ratio) decreased from 86.0% to 84.6% (95%CI, 78.3%-91.6%), while specificity did not change substantially (93.7% (95%CI, 91.1%-96.4%) in weighted pooled. SROC analysis yielded smaller β-coefficient of 3.61 with slightly increased p=value (0.118) in the comparison between hCT and V/Q scan using threshold 3.

D. 考察

The present study provides a meta-analysis of the literatures on hCT and perfusion scanning. Based on SROC analysis, hCT and perfusion scan had similar discriminatory power using high probability threshold (threshold 1). The data also suggest that hCT has a greater discriminatory power than the V/Q scan with high/intermediate probability threshold (threshold 2). The results of the model comparing hCT and V/Q scan using normal/near-normal threshold (threshold 3) was confusing, because the β-coefficient was large but it was borderline statistical significant (p=0.05). As discussed later, the small number of articles, included in this meta-analysis because of stringent inclusion criteria might, could have decreased the power to detect true difference of two diagnostic tests. To evaluate these issues, it is crucial to take a quantitative view of the data and their interpretation -- that is, it is essential to think in terms of estimation rather than testing (p-value) (20). Therefore, we would rather state that β-coefficient of 4.14 is high enough to suggest that hCT has higher discriminatory power than V/Q scan using normal/near-normal threshold,

Statistical Methods/Results

The SROC analysis allows us to compare different tests by summarizing the sensitivity and specificity results from several studies into a single ROC curve (21). In the past decade, many manuscripts have suggested that hCT might be more effective than V/Q scan in diagnosing and screening of PE (22), but formal comparison of the two diagnostic tests is yet to be done. Although former studies discuss about superiority of hCT because of its higher sensitivity or specificity, one cannot state that a test is more useful only for that reason. In general, a negative result essentially rules out pulmonary embolism in a test with very high sensitivity, and a positive result effectively confirms the diagnosis in when a test with very high specificity (10). In the real clinical setting, for example, high probability threshold of V/Q scan with high specificity has been used to confirm and normal/near-normal threshold with high sensitivity to exclude pulmonary embolism. It is therefore necessary to compare the sensitivity of hCT with that of V/Q scan using normal/near-normal threshold in order to exclude the possibility of pulmonary embolism, and compare specificity to diagnose pulmonary embolism by SROC analysis.

There has been criticism of the claimed very high sensitivities and specificities of earlier hCT studies (23). This sounds very curious because CT methodology for diagnosing PE has had a remarkably rapid advance in this last decade. Therefore, we performed sensitivity analysis, which revealed that regression coefficient in favor for CATP including literatures published in 1990-1994 was slightly higher than that excluding those literatures in the model comparing with V/Q scan using threshold 3. This suggested that the diagnostic test performance was slightly better or almost equivalent in 1990-1994 to those in later years. New diagnostic tests are generally often described in glowing terms when first introduced, only to be found wanting when more experience has been gained (24). An example is use of carcinoembryonic antigen (CEA), which was originally considered a very promising tool for diagnosing colon cancer, but subsequently was found to be increased in a wide variety of other conditions (e.g., smokers) (10). This arises not from any dishonesty on the part of early investigators, or unfair medical skepticism by the medical community later. Rather, it is related to limitations in the methods used to evaluate test

characteristics in the first place, e.g. severity of patients included or other biases.

Another criticism of high accuracy as pointed out in recently published review articles is that methodological problems are common among the studies evaluating hCT in the diagnosis of PE, e.g. several studies were missing key data on the methods used to select patients. However, It is unclear how these methodological problems have actually influenced the results in our analysis (25). The Second Prospective Investigation of Pulmonary Embolism Diagnosis (PIOPED 2), funded by the US National Institutes of Health, is currently evaluating the accuracy of helical CT in the diagnosis of pulmonary embolism in more than 1000 patients, and should bring us closer to real answer to the problem (26).

Clinical Implications/Results

Our work shows that hCT and the V/Q scan using high-probability threshold probability threshold have comparable discriminatory power, while hCT had greater discriminatory power than V?Q scan using normal/near-normal threshold. More, now, is the importance of considering other aspects of the tests. First, the main problem with the V/Q scan is that definitive diagnosis can be obtained in less than 30% of patients tested, and that the remaining patients need to undergo further testing (3). The use of hCT would therefore reduce the number of patients subject to further diagnostic tests than for the V/Q scan (22). In contrast, one should consider the presence of contraindications before conducting hCT. In one report, about 24% of patients suspected of being pulmonary embolism were excluded from hCT because of contraindications such as impaired renal function, or allergy to contrast agent (23). This is a substantial proportion, similar to that suffering inconclusive results of the V/Q scan.

Second, as for the V/Q scan, large differences (25-30%) of interpretation among expert readers have been reported, especially in the classification of low or intermediate-probability scan. In contrast, hCT has better inter-observer agreement than the V/Q scan (κ=0.85 vs. 0.61) (27). Third, there are inconsistent results concerning relative cost-effectiveness, with the controversy continuing (28). These advantages or disadvantages of either test are crucial in application to the patients when PE is suspected. What is really important is to judge the advantages and disadvantages of each test, and select the most appropriate procedure for favorable outcome. Based on the results of our analyses and considering above, we recommend the following strategies; 1) to confirm PE with moderate-to-high pretest probability, use either hCT or V/Q scan with high-probability threshold; 2) to exclude PE with low pretest probability, hCT is a better test than V/Q scan. If hCT is not available, V/Q scan using normal/near-normal threshold could be a next alternative; 3) if pre-test probability is moderate-to-high, PE can not be excluded even when hCT or V/Q scan with any threshold shows negative result.

Limitations

Our review has several limitations. First, by the nature of meta-analysis, the result is subject to publication bias. Only published studies were examined, and studies with poor results are less likely to be written, submitted, and accepted. Our results may therefore be biased towards the favorable direction. This tendency should affect hCT and V/Q scan equally and not alter our qualitative conclusion. Second, as in all meta-analyses of diagnostic testing, verification bias could be present since about half the studies included in our meta-analysis did not control or mention this. Verification bias occurs when the result of the test influences the decision as to which patients receive the verification test. This can have dramatic results on the sensitivity and specificity of a test (29). We are unable to correct for this bias because the original studies do not provide the necessary information on the entire population tested. However, in our study, covariate analysis did not show a significant difference between studies that did and did not control for verification bias. Third, the large degree of variation

between observers in reading V/Q scans could limit the interpretation of our results when combining studies.

Numerous studies were excluded from analysis and smaller number of studies were finally included in meta-analysis compared with previous meta-analysis on hCT (30). One reason is that some studies used combination reference standards for comparing hCT and V/Q scan (e.g. normal results on V/Q lung scan were accepted as an alternative reference standard for the absence of pulmonary embolism) (22). The small number of studies included in our current analysis might have decreased the power to detect the true difference; however, this could not have been avoided, because it is usually supposed that the test is being compared to a sole reference standard when conducting meta-analysis of diagnostic tests. We therefore chose the optimal strategy that pulmonary angiography should be the sole reference standard. For the same reason above, only one literature on multi-detector hCT was included in our analysis. This might have underestimated the test performance of hCT, because the recent introduction of multi-detector row spiral CT is expected to offer further increase in performance, in particular the ability to scan larger anatomic volumes with high spatial resolution (31). The stringent inclusion criteria used in the current study, however, should have assured the quality of our results.

Independently pooled estimates of sensitivity and specificity could easily calculated; however, these frequently used methods have come under strong criticism, because they do not take into account the fact that different studies may have used test thresholds (32). In spite of that, the reason why we pooled estimates of sensitivity and specificity separately is that the results of β-coefficients are not always easy for readers to intuitively understand, and in the real setting, articles of SROC analysis including those cited below present the results of pooled sensitivities and specificities are presented even lately in the articles using summary ROC analysis (33). Other reason is that authorities in the field of decision sciences recommend to pool sensitivities and specificities of diagnostic tests and use these data for cost-effectiveness analysis (34). These pooled estimates, however, should carefully be interpreted when utilized to compare these two diagnostic tests directly.

E. 結論

In conclusion, hCT has the greater discriminatory power to exclude PE than V/Q scan using normal/near-normal threshold, while hCT and V/Q scan using high probability threshold had similar discriminatory power to diagnose PE.

References

1. Wells PS, Ginsberg JS, Anderson DR, et al. Use of a clinical model for safe management of patients with suspected pulmonary embolism. Ann Intern Med. 1998;129:997-1005.
2. Hyers TM. Venous thromboembolism. Am J Respir Crit Care Med. 1999;159:1-14.
3. Mills SR, Jackson DC, Older RA, Heaston DK, Moore AV. The incidence, etiologies, and avoidance of complications of pulmonary angiography in a large series. Radiology. 1980;136:295-9.
4. Stein PD, Athanasoulis C, Alavi A, et al. Complications and validity of pulmonary angiography in acute pulmonary embolism. Circulation. 1992;85:462-8.
5. The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). JAMA. 1990;263:2753-9.
6. Stein PD, Hull RD, Raskob GE. Withholding treatment in patients with acute pulmonary embolism who have a high risk of bleeding and negative serial noninvasive leg tests. Am J Med. 2000;109:301-6.
7. Righini M, Goehring C, Bounameaux H, Perrier A. Effects of age on the performance of common diagnostic tests for pulmonary embolism. Am J Med. 2000;109:357-61.

8. Remy-Jardin M, Remy J, Wattinne L, Giraud F. Central pulmonary thromboembolism: diagnosis with spiral volumetric CT with the single-breath-hold technique--comparison with pulmonary angiography. Radiology. 1992;185:381-7.

9. Blum AG, Delfau F, Grignon B, et al. Spiral-computed tomography versus pulmonary angiography in the diagnosis of acute massive pulmonary embolism. Am J Cardiol. 1994;74:96-8.

10. Goodman LR, Curtin JJ, Mewissen MW, et al. Detection of pulmonary embolism in patients with unresolved clinical and scintigraphic diagnosis: helical CT versus angiography. AJR Am J Roentgenol.1995;164:1369-74.

11. Remy-Jardin M, Remy J, Deschildre F, et al. Diagnosis of pulmonary embolism with spiral CT: comparison with pulmonary angiography and scintigraphy. Radiology. 1996;200:699-706.

12. Mayo JR, Remy J, M. , Muller NL, et al. Pulmonary embolism: prospective comparison of spiral CT with ventilation-perfusion scintigraphy. Radiology. 1997;Nov,205:447-52.

13. Drucker EA, Rivitz SM, Shepard JA, et al. Acute pulmonary embolism: assessment of helical CT for diagnosis. Radiology. 1998;209:235-41.

14. Garg K, Welsh CH, Feyerabend AJ, et al. Pulmonary embolism: diagnosis with spiral CT and ventilation-perfusion scanning--correlation with pulmonary angiographic results or clinical outcome. Radiology. 1998;208:201-8.

15. Qanadli SD, Hajjam ME, Mesurolle B, et al. Pulmonary embolism detection: prospective evaluation of dual-section helical CT versus selective pulmonary arteriography in 157 patients. Radiology. 2000;217:447-55.

16. Ruiz Y, Caballero P, Caniego JL, et al. Prospective comparison of helical CT with angiography in pulmonary embolism: global and selective vascular territory analysis. Interobserver agreement. Eur Radiol. 2003;13:823-9.

17. Nilsson T, Soderberg M, Lundqvist G, et al. A comparison of spiral computed tomography and latex agglutination D-dimer assay in acute pulmonary embolism using pulmonary arteriography as gold standard. Scand Cardiovasc J. 2002;36:373-7.

18. van Beek EJ, Brouwers EM, Song B, Bongaerts AH, Oudkerk M. Lung scintigraphy and helical computed tomography for the diagnosis of pulmonary embolism: a meta-analysis. Clin Appl Thromb Hemost. 2001;7:87-92.

19. Safriel Y, Zinn H. CT pulmonary angiography in the detection of pulmonary emboli: a meta-analysis of sensitivities and specificities. Clin Imaging. 2002;26:101-5.

20. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Med Decis Making. 1993;13:313-21.

21. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. Stat Med. 1993;12:1293-316.

22. Irwig L, Tosteson AN, Gatsonis C, et al. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med. 1994;120:667-76.

23. Fleischmann KE, Hunink MG, Kuntz KM, Douglas PS. Exercise echocardiography or exercise SPECT imaging? A meta-analysis of diagnostic test performance. JAMA. 1998;280:913-20.

24. Owens DK, Holodniy M, McDonald TW, Scott J, Sonnad S. A meta-analytic evaluation of the polymerase chain reaction for the diagnosis of HIV infection in infants. JAMA. 1996;275:1342-8.

25. Harvey RT, Gefter WB, Hrung JM, Langlotz CP. Accuracy of CT angiography versus pulmonary angiography in the diagnosis of acute pulmonary embolism: evaluation of the literature with summary ROC curve analysis. Acad Radiol. 2000;7:786-97.

26. Rathbun SW, Raskob GE, Whitsett TL. Sensitivity and specificity of helical computed tomography in the diagnosis of pulmonary embolism: a

systematic review. Ann Intern Med. 2000;132:227-32.
27. Trujillo NP, Pratt JP, Talusani S, Quaife RA, Kumpe D, Lear JL. DTPA aerosol in ventilation/perfusion scintigraphy for diagnosing pulmonary embolism. J Nucl Med. 1997;38:1781-3.
28. Midgette AS, Stukel TA, Littenberg B. A meta-analytic method for summarizing diagnostic test performances:
receiver-operating-characteristic-summary point estimates. Med Decis Making. 1993;13:253-7.
29. Woods ER, Iles S, Jackson S. Comparison of scintigraphic diagnostic criteria in suspected pulmonary embolism. Can Assoc Radiol J. 1989;40:194-7.
30. Fletcher RH, Fletcher SW, Wagner EH. Chapter 3 - Diagnosis. In: Satterfield TS, Taylor C, editors. Clinical Epidemiology. 3rd ed. Philadelphia: Lippincott Williams & Wilkins 1996: 43-66.
31. Rothman KJ. Chapter 6 - Random Error and Role of Statistics. Epidemiology. An introduction. New York: Oxford University Press; 2002:113-29.
32. van Erkel AR, van Rossum AB, Bloem JL, Kievit J, Pattynama PM. Spiral CT angiography for suspected pulmonary embolism: a cost-effectiveness analysis. Radiology. 1996;201:29-36.
33. van Rossum AB, Pattynama PM, Mallens WM, Hermans J, Heijerman HG. Can helical CT replace scintigraphy in the diagnostic process in suspected pulmonary embolism? A retrolective-prolective cohort study focusing on total diagnostic yield. Eur Radiol. 1998;8:90-6.
34. Paterson DI, Schwartzman K. Strategies incorporating spiral CT for the diagnosis of acute pulmonary embolism: a cost-effectiveness analysis. Chest. 2001;119:1791-800.
35. Perrier A, Howarth N, Didier D, et al. Performance of helical computed tomography in unselected outpatients with suspected pulmonary embolism. Ann Intern Med. 2001;135:88-97.
36. Rozanski A, Diamond GA, Berman D, Forrester JS, Morris D, Swan HJ. The declining specificity of exercise radionuclide ventriculography. N Engl J Med. 1983;309:518-22.
37. Mullins MD, Becker DM, Hagspiel KD, Philbrick JT. The role of spiral volumetric computed tomography in the diagnosis of pulmonary embolism. Arch Intern Med. 2000;160:293-8.
38. Kearon C. Excluding pulmonary embolism with helical (spiral) computed tomography: Evidence is catching up with enthusiasm. CMAJ. 2003;168:1430-1.
39. Roger VL, Pellikka PA, Bell MR, Chow CW, Bailey KR, Seward JB. Sex and test verification bias. Impact on the diagnostic value of exercise echocardiography. Circulation. 1997;95:405-10.
40. Ghaye B, Szapiro D, Mastora I, et al. Peripheral pulmonary arteries: how far in the lung does multi-detector row spiral CT allow analysis? Radiology. 2001;219:629-36.
41. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol. 1995;48:119-30.
42. Oei EH, Nikken JJ, Verstijnen AC, Ginai AZ, Myriam Hunink MG. MR imaging of the menisci and cruciate ligaments: a systematic review. Radiology. 2003;226:837-48.
43. Visser K, Hunink MG. Peripheral arterial disease: gadolinium-enhanced MR angiography versus color-guided duplex US--a meta-analysis. Radiology. 2000;216:67-77.
44. Hunink M, Glasziou P, Siegel J, et al. Chapter 8 - Finding and summarizing the evidence. Decision making in health and medicine: Integrating evidence and values. Cambridge: Cambridge University Press; 2001:214-244.

Table 1 Comparison between PIOPED criteria and our thresholds defined in the current study.

PIOPED criteria	Thresholds in our study		
	Threshold 1	Threshold 2	Threshold 3
High probability	positive	positive	positive
Intermediate probability	negative	positive	positive
Low probability	negative	negative	positive
Near-normal/normal probability	negative	negative	negative

Positive indicates the test result of V/Q scan is positive in each threshold; negative indicates the test r esult of V/Q scan is negative in each threshold. For example, a patient, classified as intermediate prob ability in PIOPED criteria in original study, is classified into negative using our threshold 1 and into po sitive using thresholds 2 and 3.

Table 2. Clinical Characteristics of the Studies Included in the Meta-analysis

Helical CT

Author	Age, mean, year	Women (%)	Duration between the Tests	Study Design	Enrollment of Patients	Collimation/ pitch	Size of Pulmonary Emboli	Verification Bias	Interpretation Bias	TP	FN	FP	TN	No. of Patients
Remy-Jardin (8)	34	29	<24	Prospective	Consecutive	5mm/1	Segmental-Main PA	No	No	18	0	1	23	42
Blum (9)	43	60	<12	Prospective	N/A	5mm/1	Segmental-Main PA	No	No	7	0	0	3	10
Remy-Jardin (11)	59	57	<24	Prospective	Consecutive	3-5mm/1.7	Segmental-Main PA	No	No	39	4	0	32	75
Goodman (10)	53	40	<24	Prospective	N/A	5mm/1	Subsegmental-Main PA	No	No	7	4	1	8	20
Garg (14)	63	4	<24	Prospective	N/A	3mm/2	Segmental-Main PA	Likely	No	4	2	0	18	24
Drucker (13)	57	53	<12	Prospective	N/A	5mm/1	Segmental-Main PA	No	No	8	7	1	31	47
Quandli(16)	58	54	<12	Prospective	Consecutive	2*2.5mm/1	Subsegmental-Main PA	No	No	56	3	3	89	151
Nilsson(15)	53	53	<12	Prospective	N/A	3-5mm/1.3-1.7	Subsegmental-Main PA	No	No	30	3	2	55	90
Ruiz(16)	N/A	44	<24	Prospective	Consecutive	3mm/1.6	Segmental-Main PA	No	No	21	2	7	31	61

(Table 2-Continued)

Ventilation-perfusion Scan																
Author	Age, mean, year	Women (%)	Duration between the Tests	Study Design	Enrollment of Patients	Radionuclide Used		Size of Pulmonary Emboli	Verification Bias	Interpretation Bias	Thresholds Used for Data Extraction	TP	FN	FP	TN	No. of Patients
						V	Q									
PIOPED(5)	56	55	<24	Prospective	N/A	^{133}Xe	^{99m}Tc-MAA	N/A	No	N/A	Threshold 1	102	149	14	466	731
											Threshold 2	207	44	231	249	
											Threshold 3	246	5	430	50	
Woods(29)	NA	NA	<48	Retrospective	N/A	^{99m}Tc-DTPA	^{99m}Tc-MAA	N/A	Likely	No	Threshold 1	6	7	1	24	38
											Threshold 2	10	3	13	12	
											Threshold 3	13	0	23	2	
Trujillo (27)	NA	NA	<48	Prospective	N/A	^{99m}Tc-DTPA	^{99m}Tc-MAA	N/A	Likely	N/A	Threshold 1	72	100	6	277	455
											Threshold 2	164	8	177	106	
											Threshold 3	172	0	280	3	
Garg (14)	63	NA	<48	Prospective	N/A	^{99m}Tc-DTPA	^{99m}Tc-MAA	Segmental-Main PA	Likely	No	Threshold 1	0	7	0	18	25
											Threshold 2	7	0	14	4	
											Threshold 3	7	0	18	0	
Goodman (10)	53	8	<48	Prospective	N/A	^{99m}Tc-PYP	^{99m}Tc-MAA	Subsegmental-Main PA	Likely	No	Threshold 1	1	10	0	9	20
											Threshold 2	6	5	4	5	
											Threshold 3	11	0	9	0	

V indicates ventilation; Q, perfusion; Xe, xenon; Tc, technetium; MAA. macroaggregated albumin; DTPA, diethylenetriamine pentacetic acid;PYP, pyro phosphate; PA, pulmonary artery; N/A, not available

Table 3. [Y. H. 1]**Pooled sensitivity and specificity of hCT and ventilation-perfusion scan using different thresholds***

	Year Searched		Sensitivity	95%CI	Specificity	95%CI
V/Q Scan	1985-2003	Threshold 1	39.0%	37.3% - 40.8%	97.1%	96.0% - 98.3%
		Threshold 2	86.0%	83.3% - 88.8%	45.5%	43.9% - 47.1%
		Threshold 3	98.4%	97.2% - 99.5%	4.8%	4.7% - 4.9%
hCT	1990-2003		86.0%	80.2% - 92.1%	93.7%	91.1% - 96.2%
	1995-2003		84.6%	78.3% - 91.6%	93.7%	91.1% - 96.4%

*CI indicates confidence interval; hCT, helical computed tomography

Table 4[Y. H. 2]**. Comparison of SROC Analyses between hCT and V/Q Scan***

Comparison	β-Coefficient	(95% CI)	P
hCT (1990-2003) vs.			
V/Q scan (Threshold 1)	0.588	(-1.55 to 2.74)	0.457
V/Q scan (Threshold 2)	3.73	(2.56 to 4.9)	<0.001
V/Q scan (Threshold 3)	4.14	(-0.002 to 8.28)	0.05
hCT (1995-2003) vs.			
V/Q scan (Threshold 1)	0.72	(-2.45 to 2.9)	0.515
V/Q scan (Threshold 2)	3.63	(2.4 to 4.86)	<0.001
V/Q scan (Threshold 3)	3.61	(-9.22 to 8.14)	0.118

* SROC indicates summary receiver operating characteristics: CI, confidence interval

Table 5 [Y. H. 3]**Post-test probabilities of PE on different pretest probabilities and test results**

Tests	Test Results	Pretest Probability=0.03		Pretest Probability=0.27		Pretest Probability=0.78	
		Post-test probability	(95%CI)	Post-test probability	(95%CI)	Post-test probability	(95%CI)
hCT	Positive	0.296	(0.243-0.399)	0.841	(0.802-0.893)	0.980	(0.974-0.987)
	Negative	0.005	(0.003-0.006)	0.055	(0.032-0.074)	0.347	(0.235-0.422)
V/Q scan							
Threshold 1	Positive	0.296	(0.225-0.420)	0.841	(0.785-0.901)	0.980	(0.971-0.988)
	Negative	0.019	(0.018-0.020)	0.196	(0.190-0.203)	0.522	(0.457-0.573)
Threshold 2	Positive	0.047	(0.044-0.049)	0.380	(0.366-0.394)	0.848	(0.841-0.856)
	Negative	0.009	(0.007-0.011)	0.107	(0.085-0.129)	0.522	(0.457-0.574)
Threshold 3	Positive	0.031	(0.0306-0.0313)	0.287	(0.284-0.289)	0.786	(0.783-0.788)
	Negative	0.010	(0.003-0.017)	0.116	(0.036-0.185)	0.546	(0.257-0.675)

hCT, helical computed tomography; PE, pulmonary embolism; V/Q scan, ventilation-perfusion scan

Fig. 1. Results of Literature Search*

*PA indicates pulmonary angiography; hCT, helical computed tomography; PE, pulmonary embolism

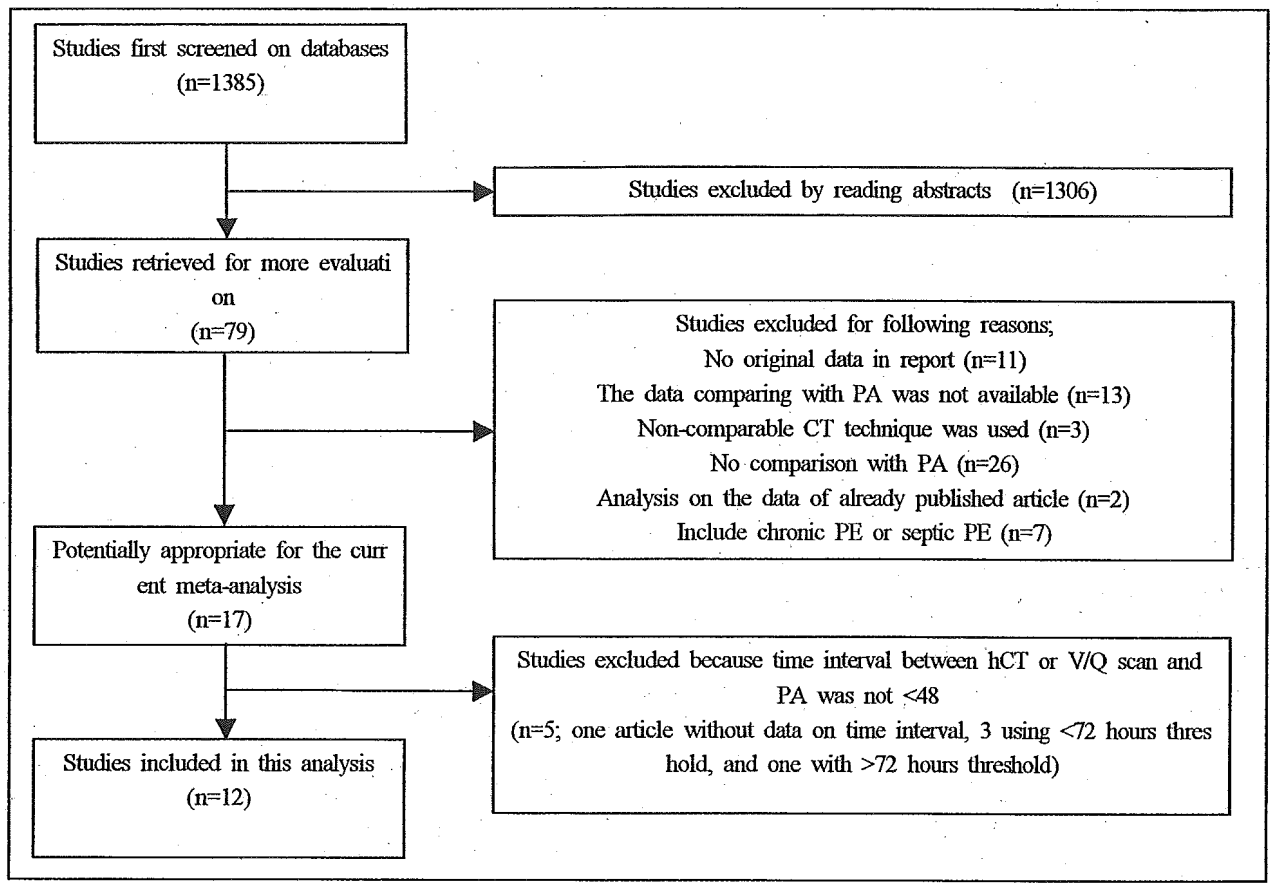
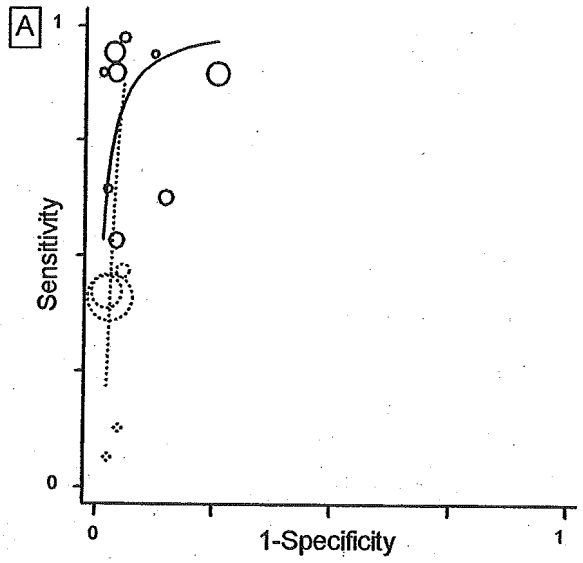
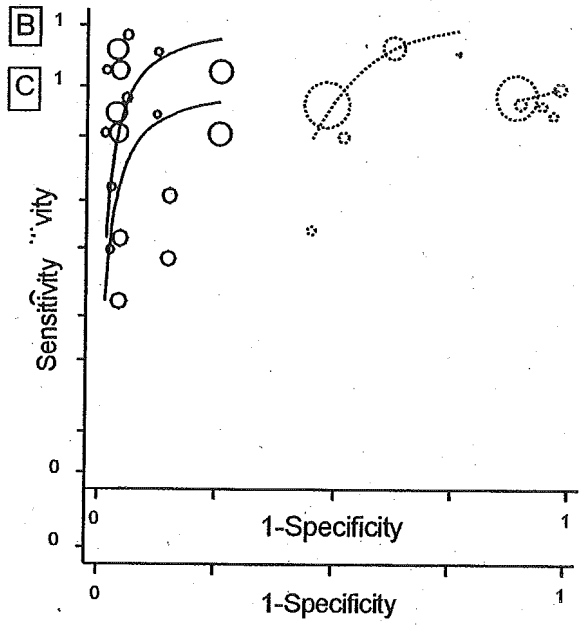
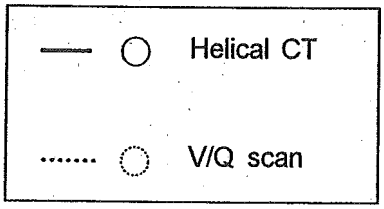


Fig.2 Panel A shows summary receiver operating characteristics (SROC) curves for helical computed tomography (hCT) based on a model comparing ventilation-perfusion (V/Q) scan using scan threshold 1. Panel B shows SROC curves for hCT based on a model comparing ventilation-perfusion (V/Q) scan using scan threshold 2. Panel C shows SROC curves for hCT based on a model comparing ventilation-perfusion (V/Q) scan using scan threshold 3. In all panels, the horizontal axis represents the false-positive ratio (1-specificity) and the vertical axis the true-positive ratio (sensitivity). The size of plotting symbol is inversely proportional to the variance of an observation. hCT has a similar discriminatory power to that V/Q scan with thresholds 1, while overall discriminatory power of hCT is better than that of V/Q scan with threshold 2 and threshold 3.[Y. H. 4]







F. 研究発表

学会発表

平成17年7月20日　消化器外科にランダム化比較試験は必要か？
ランダム化比較試験と観察研究の結果比較（シンポジウム）
第60回　日本消化器外科学会総会（東京）

論文発表

1. Shikata S, Yamagishi H, Taji Y, Shimada T, Noguchi Y. Single- versus two- layer intestinal anastomosis: a meta-analysis of randomized controlled trials. BMC Surg. Jan 27 2006;6(1):2.

2. Hayashino Y, Goto M, Noguchi Y, Fukui T. Ventilation-Perfusion Scanning and Helical CT in Suspected Pulmonary Embolism: Meta-Analysis of Diagnostic Performance. Radiology. Mar 2005;234(3):740-748.

3. Hayashino Y, Noguchi Y, Fukui T. Systematic evaluation and comparison of statistical tests for publication bias. J Epidemiol. Nov 2005;15(6):235-243.

4. Noguchi Y, Nagata-Kobayashi S, Stahl JE, Wong JB. A meta-analytic comparison of echocardiographic stressors. Int J Cardiovasc Imaging. Apr 2005;21(2-3):189-207.

5. Shikata S, Noguchi Y, Fukui T. Early Versus Delayed Cholecystectomy for Acute Cholecystitis: A Meta-analysis of Randomized Controlled Trials. Surg Today. 2005;35(7):553-560.

6. 四方　哲、野口善令、福井次矢：急性胆嚢炎に対する手術時期、ランダム化比較試験のメタアナリシス．自治医大医学部紀要 27:1-8, 2004（総説）

その他

投稿中

1. Satoru Shikata, Takeo Nakayama, Yoshinori Noguchi, Yoshinori Taji and Hisakazu Yamagishi.: Comparison of effects in randomized controlled trials with observational studies in digestive surgery

2. SETA Takeshi, SHIMADA Toshihiko, NOGUCHI Yoshinori: Protease inhibitors for preventing complications associated with ERCP: a meta- analysis

3. 四方　哲、野口善令　腹腔鏡下胆嚢摘出術の手術時期、メタアナリシスによるランダム化比較試験と観察研究の比較

循環器疾患の危険因子・予防因子に関するメタ・アナリシス
(医療技術評価総合研究事業) 分担研究報告書

分担研究者　横山徹爾　国立保健医療科学院技術評価部

研究要旨：
観察的な疫学研究の結果から、野菜・果物またはこれらの食物に多く含まれる栄養素の摂取量が多い者では、循環器疾患リスクが低いことが示唆されている。これら栄養素のバイオマーカーとして血中ビタミンC濃度を測定し、前向きに追跡して循環器疾患リスクを評価したコホート研究が近年複数報告されている。本研究では、そのメタ・アナリシスを行う。MedLine(PubMed)による検索の結果、該当する原著論文は脳卒中4編、虚血性心疾患6編あった。各論文に示されたカテゴリー毎の相対危険と信頼区間に基づき、量反応関係を仮定して Greenland and Longnecker の方法により血中ビタミンC濃度 10 μmol/L 増加あたりの相対危険を推定して統合したところ、脳卒中では 0.90(95%信頼区間 0.86-0.95)、虚血性心疾患では 0.99(0.96-1.02)であった。また、Greenland and Longnecker の方法を用いて解析を行うための簡便なソフトウェアを開発した。

A. 研究目的

これまでの国内外のコホート研究によって、野菜・果物またはこれらの食物に多く含まれる栄養素の摂取量が多い者では、循環器疾患リスクが低いことが示唆され[1]、また、最近10年ほどの間に、血中ビタミンC濃度と循環器疾患(脳卒中、虚血性心疾患)リスクが逆相関を示すというコホート研究の報告が相次いでいる[2-9]。このようなことから、現在では循環器疾患発症をエンドポイントに含む複数の大規模な無作為化比較試験が米国(WACS, 8000人; Physicians' Health Study II, 15000人)とフランス(SUVIMAX, 14000人)で進行中であるが[10]、その結果の解釈のために観察研究の知見を整理しておくことは重要である。コホート研究等の観察的な研究では多くの場合、危険因子をカテゴリー分けして相対危険を推定しているが、そのカテゴリーわけの方法が研究によって統一されておらず、そのためメタ・アナリシスによって結果を統合することは必ずしも容易ではない。

本研究では、コホート研究のメタ・アナリシスによって血中ビタミンC濃度と循環器疾患リスクとの関連の強さを、Greenland and Longnecker の方法(G-L 法)[11]を用いて定量的に評価し、その危険因子・予防因子としての意義の大きさを明らかにする。また、多くの研究者が G-L 法をメタ・アナリシスに応用できるように、解析ソフトウェアを提供する。

B. 研究方法

血中ビタミンC濃度と循環器疾患リスクとの関連について報告されている原著論文についての系統的な情報収集を行った。英文は MedLine(PubMed)を用い、検索式として、脳卒中をエンドポイントにしたものは、cohort studies[MeSH] AND (ascorbic acid[MeSH] OR ascorbic acid[tw] OR ascorbate[tw] OR vitamin c[tw]) AND (cerebrovascular disease[MeSH] OR stroke[tw])、虚血性心疾患をエンドポイントとしたものは cohort studies[MeSH] AND (ascorbic acid[MeSH] OR ascorbic acid[tw] OR ascorbate[tw] OR vitamin c[tw]) AND (cardiovascular disease[MeSH] OR heart disease[tw])とした。

該当する論文に記載されている情報を整理し、相対危険を統合するために、血中ビタミンC濃度と循環器疾患罹患・死亡リスクの(対数の)間に直線的な関係があると仮定し、カテゴリー

化されて示された論文中の相対危険と信頼区間のデータからG-L法を用いて、血中ビタミンCの一定濃度増加あたりの相対危険とその分散を各研究ごとに推定した。血中ビタミンC濃度はほぼ正規分布に近いことが経験的に知られているので[4]、論文中に示された各カテゴリーの代表値としては、そのカテゴリーの中央値（例：四分位の場合は12.5, 37.5, 62.5, 87.5パーセンタイル値）を、正規分布を仮定して推定して用いた。

研究間の相対危険の均質性を検討し、均質であれば母数モデル、異質であれば変量モデルを採用して相対危険の統合を行った。

C. 研究結果

脳卒中をエンドポイントとしたコホート研究は2004年3月3日に検索を行った結果、MedLineで35件が抽出された。これらのうち、血中ビタミンC濃度と脳卒中リスクを評価したコホート研究で原著論文は、4編だけであった[2-5]。

虚血性心疾患をエンドポイントとしたものは2005年10月3日に検索を行った結果、151件が抽出された。これらのうち、血中ビタミンC濃度と虚血性心疾患リスクを評価したコホート研究で原著論文は、6編だけであった[2, 3, 6-9]。2006年2月8日にも再検索をしたが、いずれの疾患もこれ以上増えなかった。

表1に脳卒中リスクの要約表を示す。国別内訳はフィンランド2編[2, 5]、日本1編[4]、英国1編で、脳卒中を病型に分けたもの2編、全脳卒中をまとめたもの2編であった。血中ビタミンC濃度は、4分位で分けたものが2編、3分位が1編、第1四分位(<22.7 μmol/L)対それ以外としたものが1編であった。いずれも最低値群（または最高値群）を基準とした性年齢調整ハザード比と多変量調整（血圧と血清総コレステロールは共通、他は研究によって異なる）ハザード比と95%信頼区間が示されていた。低値対それ以外とした1編では、血中ビタミンC濃度単独での脳卒中との関連は有意でなく、血中βカロチンと同時に低値の場合のみ有意なリスク上昇があった。表中にはβカロテン濃度に関係なくビタミンC濃度のみでの比較となるように統合し

表1. 血中ビタミンC濃度と脳卒中リスクに関するコホート研究の結果の要約

著者 (発表年)[ref.#]	集団	追跡年数	脳卒中のイベント数	血中ビタミンC濃度 (μmol/L)		多変量調整ハザード比		
				カテゴリー	値	点推定	95% 信頼区間	調整変数
Gey et al. (1993)[2]	**Finland-1,** 男性2974人	12年	31	Q1	<22.7	1.53	(1.12-2.09)	Age, smoking, BP, TC
				Q2-Q4	≥22.7	1	referent	
Gale et al. (1995)[3]	**UK,** 男性359人 女性307人	20年	124	T1	≤11.91	1	referent	Sex, age, BP, TC
				T2	11.92-27.82	1.1	(0.7-1.8)	
				T3	>27.82	0.7	(0.4-1.1)	
Yokoyama et al. (2000)[4]	**Japan,** 男性880人 女性1241人	20年	196	Q1	≤40	1	referent	Sex, age, BP, TC, BMI, atrial fibrillation, antihypertensive medication, personal history of IHD, physical activity, smoking, alcohol drinking.
				Q2	41-51	0.89	(0.60–1.32)	
				Q3	52-63	0.84	(0.55–1.29)	
				Q4	≥64	0.71	(0.45–1.14)	
Kurl et al. (2002)[5]	**Finland-2,** 男性2419	10年	120	Q1	<28.4	2.10	(1.17-3.80)	Age, BP, TC, BMI, smoking, alcohol drinking, seasons, diabetes, ischemic change in ECG
				Q2	28.4–47.56	1.76	(0.96-3.20)	
				Q3	48.14–64.96	1.47	(0.79-2.75)	
				Q4	>64.96	1	referent	

Q1-Q4, 1st-4th quartiles; T1-T3, 1st-3rd tertiles.

BP, blood pressure; TC, serum total cholesterol; IHD, ischemic heart disease.

Gey et al.[2] の研究はhigh-とlow-caroteneグループを分けて分析しているが、ここではプールした推定値のみ示した。