CME/SAM

The Influence of Rapid Onsite Evaluation on the Adequacy Rate of Fine-Needle Aspiration Cytology

A Systematic Review and Meta-Analysis

Robert L. Schmidt, MD, PhD, MBA, Benjamin L. Witt, MD, Leslie E. Lopez-Calderon, MD, and Lester J. Layfield, MD

Key Words: Fine-needle aspiration (FNA); Rapid onsite evaluation (ROSE); Modeling; Adequacy; Meta-analysis

DOI: 10.1309/AJCPEGZMJKC42VUP

Upon completion of this activity you will be able to:
- predict situations where rapid onsite evaluation (ROSE) is likely to be effective.
- define adequacy and diagnostic yield.
- identify factors that cause variation in fine-needle aspiration cytology adequacy rates.
- describe the deficiencies of single-cohort studies of ROSE performance.

The ASCP is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians. The ASCP designates this journal-based CME activity for a maximum of 1 *AMA PRA Category 1 Credit*™ per article. Physicians should claim only the credit commensurate with the extent of their participation in the activity. This activity qualifies as an American Board of Pathology Maintenance of Certification Part II Self-Assessment Module.

The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose.

Questions appear on p 396. Exam is located at www.ascp.org/ajcpcme.

Abstract

Rapid onsite evaluation (ROSE) has the potential to improve the adequacy rates of fine-needle aspiration (FNA) cytology. Studies have obtained variable results on the influence of ROSE. We conducted a systematic review and meta-analysis of studies on the influence of ROSE on FNA adequacy. We synthesized evidence across all anatomic locations. We only included studies that contained a control arm and compared cohorts with ROSE against cohorts without ROSE at a single location. We screened 2,179 studies and identified 25 studies that met our inclusion criteria. On average, ROSE improves the adequacy rate by 12%, but there was considerable variability across studies. The adequacy rate with ROSE depends on the non-ROSE adequacy rate. Sixty-five percent of the variability in the adequacy rate with ROSE was found to occur because of differences in the adequacy rate without ROSE. Studies with high non-ROSE adequacy rates showed low improvement after ROSE was implemented. Studies must account for the effect of the non-ROSE adequacy rate to determine the effect of ROSE on FNA adequacy rates.

Fine-needle aspiration (FNA) is a well-established procedure that is commonly used for investigating lesions at many anatomic locations. It is regarded as safe and accurate and has a low complication rate. The adequacy rate is a key aspect of FNA performance. Adequacy is defined in 2 ways: on a per-pass basis and a per-case basis. Adequacy is generally reported on a per-case basis. Sampling for a case is considered adequate if at least 1 adequate sample is obtained.

It is important to distinguish adequacy from diagnostic yield and accuracy. *Diagnostic yield* refers to the rate at which a diagnosis is made (per slide or per case) and is distinct from adequacy. *Adequacy* measures whether a sample provides sufficient material for a diagnosis. *Accuracy* refers to the correspondence between cases for which a diagnosis was rendered (nondiagnostic cases are excluded) and a gold standard (histopathology or clinical follow-up). Adequacy is necessary but not sufficient for diagnosis. Diagnostic yield and accuracy are more directly related to patient outcomes than adequacy; however, these concepts are less directly related to sampling performance because they depend on the performance of both sampling (adequacy rate) and interpretation (rate of inconclusive samples, accuracy). Thus, adequacy is a more direct measure of sampling performance than diagnostic yield or accuracy.

Studies have shown wide variation in adequacy rates across different study sites. Variation in adequacy most likely occurs because of the fact that FNA is a complex multistep process, and many factors have the potential to affect the overall diagnostic yield.[1] Adequacy rates can be affected by the number of needle passes,[2-4] the needle type and size,[5-8]

aspirator experience,[9,10] and the use of rapid onsite evaluation (ROSE) on aspirate specimens.[11,12]

ROSE has significant potential to improve adequacy rates; however, ROSE is significantly costly, and many sites do not have access to a cytopathologist to implement ROSE. Thus, it is important to quantify the influence of ROSE and to determine the circumstances under which ROSE is likely to increase adequacy. Studies on the influence of ROSE are complicated by the site-to-site variation in adequacy that can mask the effect of ROSE. Numerous studies have examined the effect of ROSE on FNA specimen adequacy; however, most studies report on the performance of a single cohort at a single study site. As we will show, it is difficult to distinguish the influence of ROSE from other factors in a single cohort design. In contrast, studies that compare the performance of 2 cohorts, with and without ROSE, at a single study site are much more likely to isolate the effect of ROSE from other factors that ordinarily vary from site to site.

To our knowledge, the literature on the influence of ROSE on FNA adequacy has never been reviewed. We therefore conducted a systematic review and meta-analysis based on high-quality studies with head-to-head comparisons of 2 cohorts (with and without ROSE). We aimed to determine the influence of ROSE on the adequacy and diagnostic yield of FNA from lesions in various anatomic sites and to identify factors that influence ROSE.

Materials and Methods

Literature Search

We followed the guidelines for systematic reviews of diagnostic accuracy studies.[13,14] We searched the MEDLINE and EMBASE databases on November 6, 2011, using the following search string: "needle biopsy" AND "assessment or onsite OR on-site or immediate or rapid"/title or abstract. We used no restrictions on language or period of study. We included studies from all anatomic locations.

Only studies comparing either adequacy or diagnostic yield between 2 cohorts (with ROSE vs without ROSE) at a single site were eligible for inclusion to increase the overall rigor of our study. Our objective was to synthesize evidence from high-quality studies. Although single-cohort studies are more frequently published, 2-cohort studies provide much higher-quality evidence on the influence of ROSE because they minimize the effect of site-to-site differences in adequacy and isolate the incremental influence of ROSE. Aside from the requirement for 2 cohorts at a single site, no restriction on study design was used.

The titles and abstracts of the resulting set of studies were independently screened by 2 authors (B.L.W. and R.L.S.), and discrepancies were resolved by a third author (L.J.L.). Screening was performed in 2 stages. In the first stage, we included studies that reported any outcome (eg, accuracy, adequacy, cost, and complication rate) associated with ROSE. We screened studies a second time to exclude those that did not involve a comparison of 2 separate arms (with ROSE vs without ROSE) at a single institution. A citation search ("forward search") and reference search ("backward search") was conducted using Scopus on February 6, 2012 (updated May 22, 2012). Duplicates were removed, and the titles and abstracts of these additional studies were screened for potentially relevant studies. Full-text articles were then obtained for all potentially relevant studies. Studies from this set were included if they contained data comparing the 2 cohorts, with and without ROSE, at a single study site.

Data Extraction

Each of the included studies was independently assessed by 2 authors (R.L.S. and L.J.L.) using a standardized data extraction form. Discrepancies were resolved by discussion. Where possible, we formed homogeneous cohorts from studies that obtained data from multiple anatomic locations or that used different methods.

Statistical Analysis

Statistical calculations were performed using Stata 12 software (Stata, College Station, TX). Meta-analysis of adequacy was completed using a random effects model as implemented in the *metan* routine defined in Stata. Tests for heterogeneity were conducted using the inconsistency statistic.[15] Meta–regression analysis was performed using *metareg* in Stata. Statistical tests were conducted at the 5% significance level. Changes in adequacy were expressed in terms of the risk difference (RD), which is the difference between the adequacy rate with ROSE and the adequacy without ROSE.

Analysis of Risk Difference

We used metaregression to analyze the effect of the initial adequacy rate:

Equation 1

$$RD_{ij} = \kappa + \alpha_i + \beta X_j$$

where RD_{ij} = the predicted RD for tissue type i, for i = 2 … 9, and study, j; κ = a constant corresponding to the baseline RD (breast); α_i = the effect of each tissue type, i, for i = 2 … 9, on the RD relative to breast (i = 1); X_j = the non-ROSE adequacy rate for study, j; and β = the coefficient of the non-ROSE adequacy rate.

This model accounts for the possible effect of tissue type and the initial adequacy rate on the RD. Statistical tests were performed to determine whether the non-ROSE adequacy rate had an effect on the RD (ie, H_0: $\beta = 0$). We also tested whether the RD varied by tissue type (H_0: $\alpha_i = \alpha_j$) .

DOI: 10.1309/AJCPEGZMJKC42VUP

Results

Literature Search

We obtained 2,179 unique studies from the initial search of MEDLINE and EMBASE. Screening of abstracts and titles provided a set of 73 potentially relevant studies **Figure 1**. Citations and references of the set of potentially relevant studies provided an additional 1,031 studies, which were screened for relevancy and provided 2 additional studies. A secondary screen of the 73 potentially relevant studies yielded 25 studies with a total of 12,407 cases that met our inclusion criteria.[3,11,16-40]

Several studies reported subgroups from different anatomic locations or that had been sampled using different methods. These were separated into distinct data sets. For example, pancreas data were extracted from the studies by Klapman et al[11] and Saleh and Khatib.[33] The Ghofrani et al[20] data were separated into data sets corresponding to ultrasound-guided and palpation-guided FNA. This produced a total of 31 data sets.

Characteristics of Included Studies

All included studies compared 2 cohorts, with and without ROSE, at a single site. Nineteen of the 25 studies were conducted in the United States **Table 1**. The included studies used 3 main types of designs: prospective randomized trials (n = 5), retrospective "before and after" studies in which ROSE was allocated by period (n = 5), and retrospective studies in which the allocation of ROSE was not specified (n = 15). Eight of the 21 studies specified the number of pathologists involved in the study and whether the pathologists in the ROSE and non-ROSE cohorts were the same. None of the studies specifically mentioned whether the pathologist making the final diagnosis was blinded to the initial assessment of adequacy by ROSE. Fifteen of the 25 studies used pathologists or cytopathology fellows as ROSE assessors, 3 studies used cytotechnicians, and 5 studies did not specify the type of assessor. Nine of the 25 studies specified criteria for adequacy.

Effect of ROSE on Adequacy

The results of the meta-analysis are presented in **Table 2**. On average, ROSE was associated with a 12% improvement (95% confidence interval [CI] 0.08-0.16) in adequacy rate ($P < .001$). Lung (RD = 0.18, $P < .001$), soft tissue (RD = 0.14, $P = .02$), head and neck (RD = 0.20, $P < .001$), thyroid (RD = 0.10, $P = .02$), and lymph node (RD = 0.12, $P = .007$) all showed statistically significant improvement in adequacy

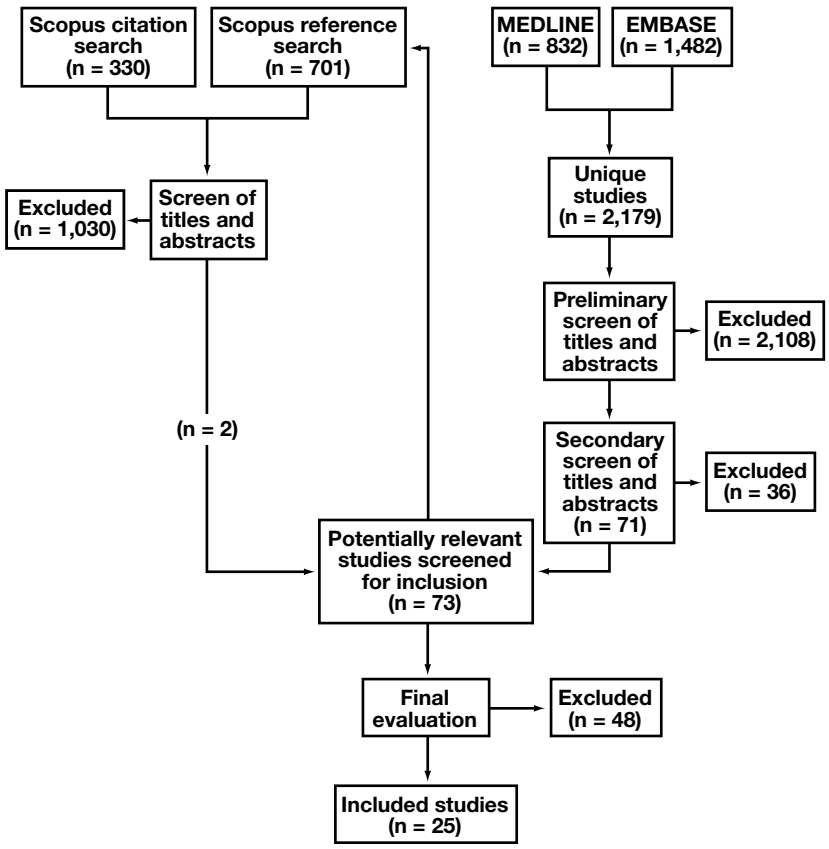

Figure 1 Flow diagram for literature search.

DOI: 10.1309/AJCPEGZMJKC42VUP

after implementation of ROSE. ROSE was not associated with improvement in studies on the breast (RD = 0.06, P = .28), pancreas (RD = 0.08, P = .14), and mediastinum (RD = .04, P = .57) or in studies reporting results aggregated from several different anatomic locations (RD = 0.11, P = .11). ROSE was associated with a decrease in adequacy rate in 2 of 31 data sets and, in each of these cases, the decreases were small (.01 and .04) and not statistically significant.

Heterogeneity was noted to be statistically significant in the overall average RD (I^2 = 93.2%; P <.001). With the exception of head and neck studies, heterogeneity in each of the subgroups was statistically significant. We investigated the adequacy rate without ROSE as a potential source of heterogeneity. We plotted the RD against the non-ROSE adequacy rate **Figure 2**. Meta–regression analysis **Table 3** showed that the non-ROSE adequacy rate was negatively correlated with the RD (t = –9.2; P < .001). Meta–regression analysis showed that the adequacy rate without ROSE accounted for 65% of the between-study heterogeneity. Our model shows a good fit between the actual and predicted RD **Figure 3**.

The influence of ROSE varied by anatomic location (α in Table 3). ROSE had very little effect on adequacy in breast, mediastinum, and soft tissue, but the effect on other tissues was significant. We designated these as low and high ROSE "impact" groups (Table 3). On average, the improvement in adequacy was 14 percentage points greater in the "high-impact" group relative to the "low-impact" group (95% CI = 9.6%-19.1%, P < .001).

Variation in the Initial (Non-ROSE) Adequacy Rate

The non-ROSE adequacy rate varied from 44% to 100% **Figure 4**. The non-ROSE adequacy rate showed significant heterogeneity in each tissue group except for head and neck **Table 4**. Meta–regression analysis showed that the variation among tissue groups was not significant (P = .56).

Discussion

To our knowledge, this is the first study to synthesize the evidence for the effect of ROSE on FNA adequacy. We found that, on average, ROSE improves the per-case adequacy rate by about 12%. After adjusting for the non-ROSE adequacy rate, ROSE had a statistically significant effect on adequacy in all 9 types of tissue (Table 3). These results were obtained from high-quality studies (ie, those having a comparison arm) and used data from 25 different studies and 9 different anatomic locations. Thus, our results have broad applicability.

Table 1
Characteristics of Included Studies

			Study Design			Assessors			
Study	No.	Location	Prosp	Rand	ROSE Allocation	No.	Same[a]	Type[b]	Adequacy Defined?
Akalin et al[39]	160	USA	0	0	BA	2	Yes	NS	No
Azabdaftari et al[16]	144	USA	0	0	NS	NS	NS	0, 2	No
Cleveland et al[3]	487	USA	0	0	NS	4	Yes	0	No
Davenport[17]	207	USA	0	0	NS	NS	NS	1	Yes
Diette et al[18]	204	USA	1	0	NS	NS	NS	NS	No
Dray et al[19]	1213	NZ	0	0	BA	2	Yes	1	Yes
Eisele et al[40]	884	USA	0	0	NS	NS	NS	0, 1	Yes
Ghofrani et al[20]	1502	USA	0	0	NS	NS	NS	NS	No
Hamill et al[21]	720	NZ	0	0	NS	NS	NS	1	No
Iglesias-Garcia et al[12]	182	Spain	0	0	NS	NS	NS	1	No
Jing et al[24]	1588	USA	0	0	NS	7	Yes	1	Yes
Klapman et al[11]	243	USA	0	0	NS	1	Yes	1	No
Kucuk et al[25]	143	Turkey	0	1	Random	1	Yes	1	No
Lachman et al[26]	331	USA	0	0	BA	NS	NS	0	No
Moberly et al[27]	274	USA	0	0	NS	3	Yes	1	No
O'Malley et al[28]	121	USA	0	0	NS	NS	NS	1	No
Padhani et al[29]	80	USA	1	1	Random	NS	NS	1, 2	No
Raab et al[31]	1176	USA	0	0	BA	NS	NS	NS	Yes
Redman et al[32]	693	USA	0	0	NS	NS	NS	0,1	Yes
Saleh and Khatib[33]	396	USA	0	0	BA	NS	NS	1	No
Santambrogio et al[34]	220	Italy	1	1	Random	NS	NS	1	No
Trisolini et al[35]	189	Italy	1	1	Random	NS	NS	1	Yes
Virayavanich et al[36]	299	USA	0	0	NS	NS	NS	NS	Yes
Yarmus et al[37]	68	USA	1	1	Random	1	Yes	1	No
Zhu and Michael[38]	883	USA	0	0	NS	NS	NS	1, 2	Yes

BA, Before and after; NS, not specified; NZ, New Zealand; Prosp, prospective; Rand, randomized; ROSE, rapid onsite evaluation.
[a] Whether the same assessors made the final diagnoses in both arms of the study (ROSE and non-ROSE).
[b] Assessor types: 0 = technologist, 1 = pathologist, 2 = fellow.

DOI: 10.1309/AJCPEGZMJKC42VUP

Table 2
Meta-Analysis of Adequacy With and Without ROSE

			95% CI for RD		Adequacy Rate		
Tissue	**Study**	**RD**	**LCL**	**UCL**	**With ROSE**	**Without ROSE**	**Heterogeneity**
Lung	Davenport[17]	0.38	0.26	0.50	0.82	0.44	I^2 = 73.8%; P = .001
	Diette et al[18]	0.31	0.19	0.43	0.82	0.50	
	Saleh and Khatib[33]	0.16	0.03	0.30	0.82	0.66	
	Padhani et al[29]	0.09	–0.11	0.29	0.79	0.70	
	Kucuk et al[25]	0.15	0.04	0.25	1.00	0.85	
	Santambrogio et al[34]	0.12	0.06	0.18	1.00	0.88	
	Yarmus et al[37]	0.06	–0.08	0.19	0.94	0.88	
	Subgroup average	0.18	0.10	0.27	P < .001		
Multiple	Saleh and Khatib[33]	0.26	0.13	0.39	0.73	0.48	I^2 = 83.8%; P < .001
	Azabdaftari et al16	0.10	–0.06	0.26	0.84	0.74	
	Klapman et al[11]	0.10	–0.05	0.24	0.87	0.77	
	Lachman et al[26]	0.00	–0.01	0.01	1.00	1.00	
	Subgroup average	0.11	–0.02	0.24	P = .11		
Breast	Dray et al[19]	0.14	0.09	0.20	0.77	0.63	I^2 = 80.8%; P =.006
	Hamill et al[21]	0.00	–0.07	0.07	0.70	0.69	
	Akalin et al[39]	0.01	–0.10	0.13	0.84	0.83	
	Subgroup average	0.06	–0.05	0.16	P = .28		
Soft tissue	Virayavanich et al[36]	0.14	0.02	0.25	0.77	0.63	NA
	Subgroup average	0.14	0.02	0.25	0.84	0.83	
Head and neck	Moberly et al[27]	0.24	0.12	0.37	0.86	0.81	I^2 = 0.0%; P =.500
	Eisele et al[40]	0.20	0.15	0.25	0.91	0.71	
	Subgroup average	0.20	0.16	0.25	P < .001		
Pancreas	Saleh and Khatib[33]	0.01	–0.32	0.35	0.67	0.65	I^2 = 75.1%; P = .007
	Klapman et al[11]	0.20	0.06	0.34	0.93	0.73	
	Iglesias-Garcia et al[12]	0.12	0.04	0.19	0.99	0.87	
	Cleveland et al[3]	0.01	–0.07	0.05	0.99	1.00	
	Subgroup average	0.08	–0.02	0.19	P < .14		
Thyroid	Jing et al[24]	0.28	0.25	0.32	0.94	0.66	I^2 = 96.3 %; P < .001
	Zhu and Michael[38]	0.26	0.21	0.31	0.94	0.68	
	O'Malley et al[28]	0.04	–0.19	0.11	0.76	0.80	
	Moberly et al[27]	0.06	–0.12	0.23	0.86	0.81	
	Ghofrani et al[20]	0.10	0.04	0.16	0.93	0.83	
	Raab et al[31]	0.04	0.01	0.07	0.96	0.92	
	Ghofrani et al[20]	0.03	–0.01	0.06	0.95	0.93	
	Redman et al[32]	0.02	–0.03	0.07	0.96	0.94	
	Subgroup average	0.10	0.01	0.19	P = .02		
Mediastinum	Trisolini et al[35]	0.04	–0.09	0.16	0.79	0.75	NA
	Subgroup average	0.04	–0.09	0.16	P = .57		
Lymph node	Cleveland et al[3]	0.12	0.03	0.20	0.96	0.84	NA
	Subgroup average	0.12	0.03	0.20	P = .007		
Overall	Total	0.12	0.08	0.16	P < .001		I^2 = 93.2%; P < .001

CI, confidence interval; I^2, inconsistency statistic (percentage of total variation that can be attributed to between-study variation); LCL, lower confidence limit; NA, not available; RD, risk difference (adequacy with ROSE – adequacy without ROSE); ROSE, rapid onsite evaluation; UCL, upper confidence limit.

We found significant heterogeneity in our results and showed that the non-ROSE adequacy rate is a significant source of heterogeneity. Thus, the non-ROSE adequacy rate is an important confounder. This result is not surprising because studies with high non-ROSE adequacy rates have little opportunity for improvement. Thus, the improvement in adequacy will be small when the non-ROSE adequacy rate is high. In contrast, sites with low non-ROSE adequacy rates have significant opportunity for improvement. We found that the non-ROSE adequacy rate accounted for 61% of the between-study variability. This is an important finding. Previous studies on the effect of ROSE have shown considerable variability in adequacy rates. Our study shows that the non-ROSE adequacy rate is an important source of variability, and the results of ROSE studies cannot be compared without knowledge of the non-ROSE adequacy rate. Our study is the first to identify this confounder and to adjust for this factor in our estimates of the effect of ROSE.

The studies included in our analysis compared 2 cohorts at a single site. The 2-cohort design is superior to a single-cohort design because it potentially controls for many site-specific factors and reduces variation; however, such designs can still suffer from bias. The method of allocation is a potential source of bias because certain types of cases might be allocated to ROSE or to certain ROSE assessors. The studies in this group used 3 different approaches for allocation of ROSE to cases: randomization, before and after, and nonspecified allocation as determined by retrospective case review.

DOI: 10.1309/AJCPEGZMJKC42VUP

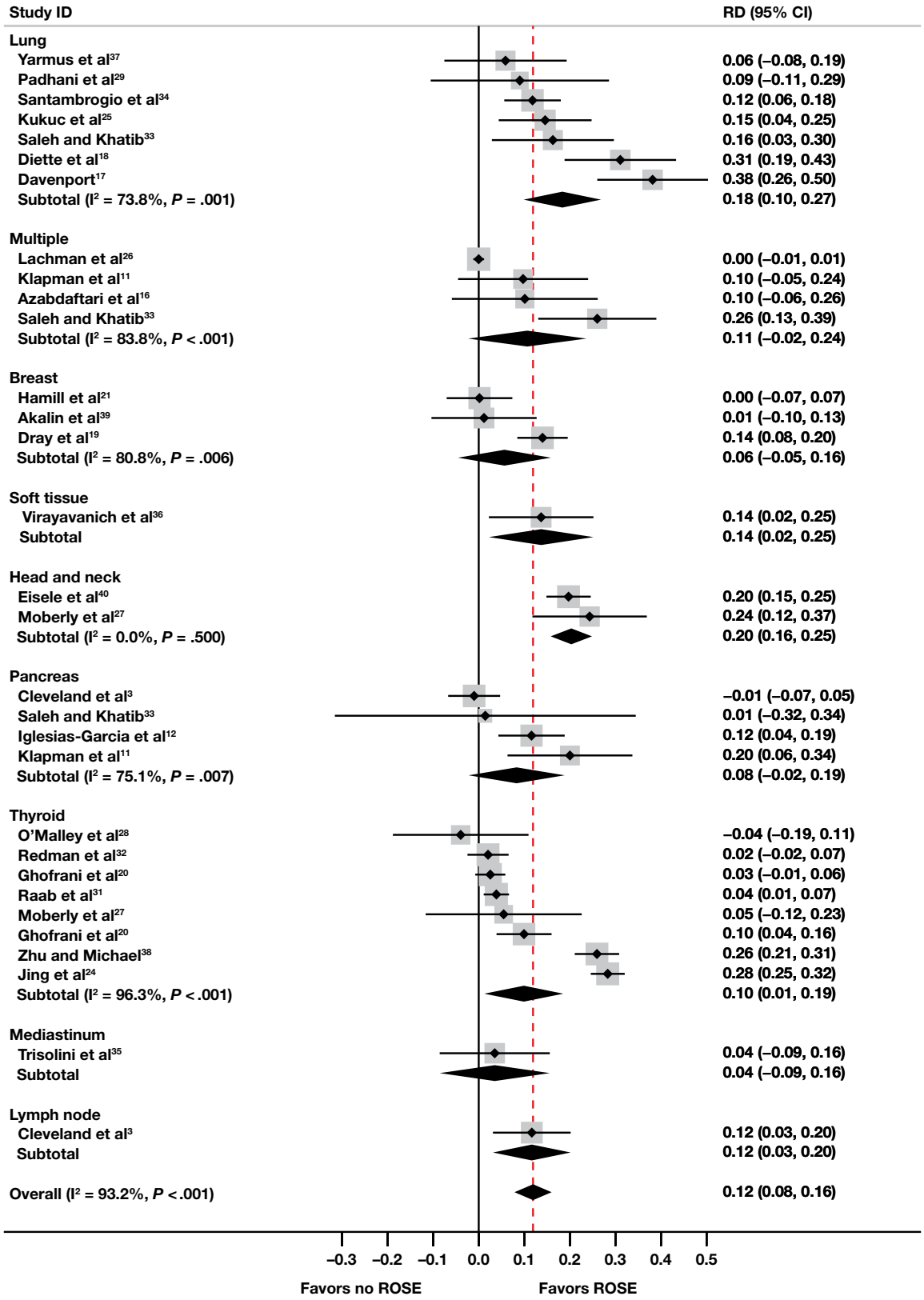

Figure 2 Forest plot of included studies. The figure shows the risk difference (RD; change in adequacy rate due to rapid onsite evaluation [ROSE]) for individual studies and subgroups of studies based on anatomic site. CI indicates confidence interval; squares, the RD estimate for each study; lines, 95% CIs; and diamonds, estimates for the subgroup and overall summary estimates. The diamonds are centered on the estimate and the width is equal to the 95% CI. I^2 is the inconsistency statistic, which measures the percentage of total variability that can be attributed to between-study variation. The analysis is unadjusted for initial adequacy rate.

DOI: 10.1309/AJCPEGZMJKC42VUP

Table 3
Meta-Regression Results for Risk Difference[a]

Equation Factor	Coefficient		ROSE Impact	95% CI		t^b	*P*
	Symbol	Value		Lower	Upper		
Non-ROSE adequacy rate, X_j	β	–0.67		–0.82	–0.51	–8.91	<.001
Tissue effects							
Breast	α_1	0.00	Low	Reference			
Pancreas	α_2	0.14	High	0.05	0.23	3.26	.004
Lung	α_3	0.14	High	0.07	0.22	3.78	.001
Lymph node	α_4	0.15	High	0.02	0.28	2.42	.02
Thyroid	α_5	0.14	High	0.07	0.21	4.07	<.001
Multiple	α_6	0.12	High	0.03	0.20	2.81	.01
Mediastinum	α_7	0.01	Low	–0.15	0.16	0.12	.90
Soft tissue	α_8	0.03	Low	–0.12	0.18	0.44	.66
Head and neck	α_9	0.15	High	0.05	0.24	3.28	.004
Constant	κ	0.53		0.40	0.65	9.00	<.001

CI, confidence interval; ROSE, rapid onsite evaluation.
[a] The influence of each tissue type on the risk difference (RD) is shown. The coefficients correspond to Equation 1 in the text. The values of the coefficients indicate the relative effect of each factor on the RD (α) as determined by Equation 1. For example, assuming a non-ROSE adequacy rate of 0.76 (X = 0.76), the predicted RD for pancreas (α = 0.14) would be calculated using Equation 1: RD = κ + α + βX = 0.53 + 0.14 – 0.67(0.76) = 0.16. The coefficients for tissue effects (α_i) are expressed relative to breast, which is used as a reference.
[b] Student *t* statistic.

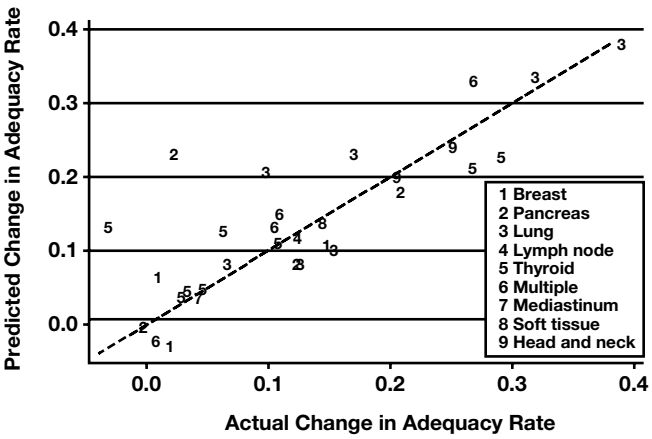

Figure 3 Change in adequacy rate. The figure shows the predicted vs actual change in adequacy rate. The change in adequacy rate or risk difference (RD) was obtained from the meta–regression analysis using the parameters in Table 3 (predicted RD = 0.53 – 0.67α + X, where α is the actual risk difference and X is an adjustment factor for the specific tissue). Each number corresponds to a specific study. The actual number corresponds to the tissue type in that study as indicated in the legend. The dashed line shows perfect correlation between the actual and predicted change in RD.

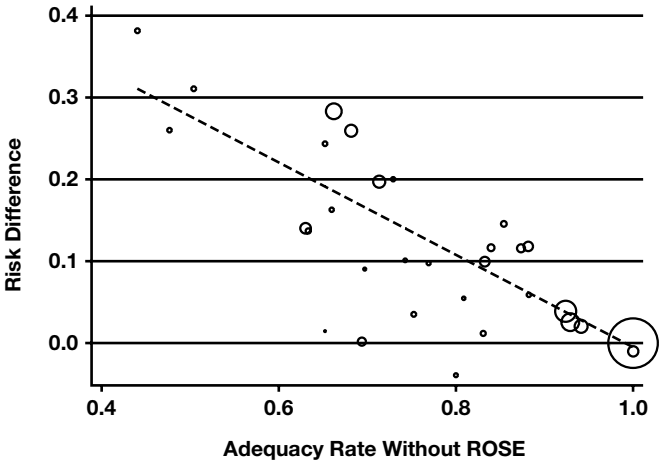

Figure 4 Correlation of adequacy rate with and without rapid onsite evaluation (ROSE). The figure shows the correlation between the risk difference and the adequacy rate without ROSE. Each circle represents a study. The size of the circle is proportional to the study size. The dashed line shows the best-fit line.

Random allocation of the intervention (ROSE) is the best design because the effects of case complexity, operator skill, and other unknown factors are distributed equally between study arms. Unfortunately, most studies did not provide a detailed description of the randomization procedure. For example, was both the intervention and the pathologist randomized or only the intervention?

The before-and-after design compares performance during 2 different periods. In this design, ROSE is used for all cases in the study period and not used at all in the control period. The validity of this approach rests on the assumption that no other significant changes occurred during that period (eg, changes in personnel, case mix, allocation, or technology). Changes in such confounders can lead to bias in a single study; however, it is likely that the direction and magnitude of this type of bias would vary across studies and not present a significant threat to bias in a meta-analysis.

DOI: 10.1309/AJCPEGZMJKC42VUP

Table 4
Initial (Non-ROSE) Adequacy Rates by Tissue

	Non-ROSE Adequacy Rate			Heterogeneity	
Tissue	AVG	LCL	UCL	I^2	*P*
Breast	0.69	0.66	0.71	87.6	<.001
Pancreas	0.82	0.76	0.88	71.9	0.03
Lung	0.72	0.68	0.75	94.6	<.001
Lymph node	0.84	0.76	0.92	NA	NA
Thyroid	0.88	0.87	0.89	97.7	<.001
Multiple	0.64	0.57	0.70	89.7	<.001
Mediastinum	0.75	0.67	0.83	NA	NA
Soft tissue	0.63	0.53	0.73	NA	NA
Head and neck	0.71	0.67	0.74	2.7	.31
Overall	0.83	0.82	0.84	96.4	<.001

AVG, average; I^2, inconsistency statistic; LCL, lower confidence limit; NA, not calculable due to small sample size; ROSE, rapid onsite evaluation; UCL, upper confidence limit.

In many studies, the interventional allocation was determined by retrospective review. In this design, it is unclear how ROSE is allocated to cases. Although there is a "control arm," it is not clear whether the control arm is a true comparator. For example, ROSE may be assigned to difficult cases and not used in situations in which adequacy rates are typically high. Although this is the weakest design, it was the most frequently used design in our set of studies. This design presents a threat to bias; however, it is difficult to predict the direction and magnitude.

Our set of studies showed considerable heterogeneity in the non-ROSE adequacy rate (Figure 3). The reason for this heterogeneity is unclear. The studies showed much more within-tissue heterogeneity than between-tissue heterogeneity (Table 4). The reason for this heterogeneity is unclear but could be caused by different definitions of adequacy. Only a few of the included studies provided criteria for adequacy and, consequently, it was not possible to test whether definitions of adequacy contributed to heterogeneity.

In conclusion, our study showed that ROSE improves the adequacy rate across a wide range of tissue types and that the non-ROSE adequacy rate is an important confounder. Non-ROSE adequacy varies significantly among institutions. Studies need to account for the effect of the non-ROSE adequacy rate when reporting the influence of ROSE.

From the Department of Pathology, University of Utah School of Medicine and ARUP Laboratories, Salt Lake City, UT.

Dr Layfield is currently with the Department of Pathology, University of New Mexico, Albuquerque. Dr Lopez-Calderon is currently with the Department of Pathology and Anatomical Sciences, University of Missouri, Columbia.

Address reprint requests to Dr Schmidt: Dept of Pathology, University of Utah, 15 N Medical Dr E, Salt Lake City, UT 84112; Robert.l.schmidt@att.net.

References

1. Schmidt R, Factor RE, Affolter KE, et al. Methods specification for diagnostic test accuracy studies in fine-needle aspiration cytology: a survey of reporting practice. *Am J Clin Pathol*. 2012;137:132-141.
2. Moller K, Papanikolaou IS, Toermer T, et al. EUS-guided FNA of solid pancreatic masses: high yield of 2 passes with combined histologic-cytologic analysis. *Gastrointest Endosc*. 2009;70:60-69.
3. Cleveland P, Gill KRS, Coe SG, et al. An evaluation of risk factors for inadequate cytology in EUS-guided FNA of pancreatic tumors and lymph nodes. *Gastrointest Endosc*. 2010;71:1194-1199.
4. LeBlanc JK, Ciaccia D, Al-Assi MT, et al. Optimal number of EUS-guided fine needle passes needed to obtain a correct diagnosis. *Gastrointest Endosc*. 2004;59:475-481.
5. Sakamoto H, Kitano M, Komaki T, et al. Prospective comparative study of the EUS guided 25-gauge FNA needle with the 19-gauge Trucut needle and 22-gauge FNA needle in patients with solid pancreatic masses. *J Gastroenterol Hepatol*. 2009;24:384-390.
6. Iqbal S, Mir RN, Sohn W. Endoscopic ultrasound-guided fine-needle aspiration using 22- and 25-gauge needles alternately. *Endoscopy*. 2009;41(suppl 2):E87.
7. Imazu H, Uchiyama Y, Kakutani H, et al. A prospective comparison of EUS-guided FNA using 25-gauge and 22-gauge needles. *Gastroenterol Res Pract*. 2009;2009:546390.
8. Siddiqui AA, Lyles T, Avula H, et al. Endoscopic ultrasound-guided fine needle aspiration of pancreatic masses in a veteran population: comparison of results with 22- and 25-gauge needles. *Pancreas*. 2010;39:685-686.
9. Ljung BM, Drejet A, Chiampi N, et al. Diagnostic accuracy of fine-needle aspiration biopsy is determined by physician training in sampling technique. *Cancer*. 2001;93:263-268.
10. Choi SH, Han KH, Yoon JH, et al. Factors affecting inadequate sampling of ultrasound-guided fine-needle aspiration biopsy of thyroid nodules. *Clin Endocrinol*. 2011;74:776-782.
11. Klapman JB, Logrono R, Dye CE, et al. Clinical impact of on-site cytopathology interpretation on endoscopic ultrasound-guided fine needle aspiration. *Am J Gastroenterol*. 2003;98:1289-1294.

DOI: 10.1309/AJCPEGZMJKC42VUP

12. Iglesias-Garcia J, Dominguez-Munoz JE, Abdulkader I, et al. Influence of on-site cytopathology evaluation on the diagnostic accuracy of endoscopic ultrasound-guided fine needle aspiration (EUS-FNA) of solid pancreatic masses. *Am J Gastroenterol.* 2011;106:1705-1710.
13. Deeks JJ, Bossuyt PM, Gatsonis C, eds. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, Version 1.0.0.* Oxford, England: Cochrane Collaboration; 2009.
14. Leeflang MM, Deeks JJ, Gatsonis C, et al; Cochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Int Med.* 2008;149:889-897.
15. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002;21:1539-1558.
16. Azabdaftari G, Goldberg SN, Wang HH. Efficacy of on-site specimen adequacy evaluation of image-guided fine and core needle biopsies. *Acta Cytol.* 2010;54:132-137.
17. Davenport RD. Rapid on-site evaluation of transbronchial aspirates. *Chest.* 1990;98:59-61.
18. Diette GB, White P Jr, Terry P, et al. Utility of on-site cytopathology assessment for bronchoscopic evaluation of lung masses and adenopathy. *Chest.* 2000;117:1186-1190.
19. Dray M, Mayall F, Darlington A. Improved fine needle aspiration (FNA) cytology results with a near patient diagnosis service for breast lesions. *Cytopathology.* 2000;11:32-37.
20. Ghofrani M, Beckman D, Rimm DL. The value of onsite adequacy assessment of thyroid fine-needle aspirations is a function of operator experience. *Cancer.* 2006;108:110-113.
21. Hamill J, Campbell ID, Mayall F, et al. Improved breast cytology results with near patient FNA diagnosis. *Acta Cytol.* 2002;46:19-24.
22. Iglesias García J, Lariño Noia J, Domínguez Muñoz JE. Endoscopic ultrasound in the diagnosis and staging of pancreatic cancer. *Rev Esp Enferm Dig.* 2009;101:631-638.
23. Jamil LH, Jones A, Khoor A, et al. Onsite cytopathology improves yield in endobronchial ultrasound (EBUS) guided fine needle aspiration of mediastinal lymph nodes: a prospective blinded study. *Gastrointest Endosc.* 2009;69:AB332.
24. Jing X, Michael CW, Pu RT. The clinical and diagnostic impact of using standard criteria of adequacy assessment and diagnostic terminology on thyroid nodule fine needle aspiration. *Diagn Cytopathol.* 2008;36:161-166.
25. Kucuk CU, Yilmaz A, Akkaya E. Computed tomography-guided transthoracic fine-needle aspiration in diagnosis of lung cancer: a comparison of single-pass needle and multiple-pass coaxial needle systems and the value of immediate cytological assessment. *Respirology.* 2004;9:392-396.
26. Lachman MF, Cellura K, Schofield K, et al. On-site adequacy assessments for image-directed fine needle aspirations: a study of 341 cases. *Conn Med.* 1995;59:657-660.
27. Moberly AC, Vural E, Nahas B, et al. Ultrasound-guided needle aspiration: impact of immediate cytologic review. *Laryngoscope.* 2010;120:1979-1984.
28. O'Malley ME, Weir MM, Hahn PF, et al. US-guided fine-needle aspiration biopsy of thyroid nodules: adequacy of cytologic material and procedure time with and without immediate cytologic analysis. *Radiology.* 2002;222:383-387.
29. Padhani AR, Scott Jr WW, Chehma M, et al. The value of immediate cytologic evaluation for needle aspiration lung biopsy. *Invest Radiol.* 1997;32:453-458.
30. Parikh A, Paul BB, Hartz D, et al. Effectiveness of ultrasound-guided thyroid FNA with on-site pathology. *Thyroid.* 2009;19:S41-S42.
31. Raab SS, Grzybicki DM, Sudilovsky D, et al. Effectiveness of Toyota process redesign in reducing thyroid gland fine-needle aspiration error. *Am J Clin Pathol.* 2006;126:585-592.
32. Redman R, Zalaznick H, Mazzaferri EL, et al. The impact of assessing specimen adequacy and number of needle passes for fine-needle aspiration biopsy of thyroid nodules. *Thyroid.* 2006;16:55-60.
33. Saleh HA, Khatib G. Positive economic and diagnostic accuracy impacts of on-site evaluation of fine needle aspiration biopsies by pathologists. *Acta Cytol.* 1996;40:1227-1230.
34. Santambrogio L, Nosotti M, Bellaviti N, et al. CT-guided fine-needle aspiration cytology of solitary pulmonary nodules: a prospective, randomized study of immediate cytologic evaluation. *Chest.* 1997;112:423-425.
35. Trisolini R, Cancellieri A, Tinelli C, et al. Rapid on-site evaluation of transbronchial aspirates in the diagnosis of hilar and mediastinal adenopathy: a randomized trial. *Chest.* 2011;139:395-401.
36. Virayavanich W, Ringler MD, Chin CT, et al. CT-guided biopsy of bone and soft-tissue lesions: role of on-site immediate cytologic evaluation. *J Vasc Interv Radiol.* 2011;22:1024-1030.
37. Yarmus L, Van Der Kloot T, Lechtzin N, et al. A randomized prospective trial of the utility of rapid on-site evaluation of transbronchial needle aspirate specimens. *J Bronchol Interv Pulmonol.* 2011;18:121-127.
38. Zhu W, Michael CW. How important is on-site adequacy assessment for thyroid FNA? An evaluation of 883 cases. *Diagn Cytopathol.* 2007;35:183-186.
39. Akalin A, Lu D, Woda B, et al. Rapid cell blocks improve accuracy of breast FNAs beyond that provided by conventional cell blocks regardless of immediate adequacy evaluation. *Diagn Cytopathol.* 2008;36:523-529.
40. Eisele DW, Sherman ME, Koch WM, et al. Utility of immediate on-site cytopathological procurement and evaluation in fine needle aspiration biopsy of head and neck masses. *Laryngoscope.* 1992;102:1328-1330.

DOI: 10.1309/AJCPEGZMJKC42VUP