

# Ductal Lavage for Detection of Cellular Atypia in Women at High Risk for Breast Cancer

William C. Dooley, Britt-Marie Ljung, Umberto Veronesi, Massimiliano Cazzaniga, Richard M. Elledge, Joyce A. O'Shaughnessy, Henry M. Kuerer, David T. Hung, Seema A. Khan, Ragsbert F. Phillips, Patricia A. Ganz, David M. Euhus, Laura J. Esserman, Bruce G. Haffty, Bonnie L. King, Mark C. Kelley, Maxine M. Anderson, Paul J. Schmit, Ramona R. Clark, Frederic C. Kass, Benjamin O. Anderson, Susan L. Troyan, Raquel D. Arias, John N. Quiring, Susan M. Love, David L. Page, Eileen B. King

**Background:** Breast cancer originates in breast epithelium and is associated with progressive molecular and morphologic changes. Women with atypical breast ductal epithelial cells have an increased relative risk of breast cancer. In this study, ductal lavage, a new procedure for collecting ductal cells with a microcatheter, was compared with nipple aspiration with regard to safety, tolerability, and the ability to detect abnormal breast epithelial cells. **Methods:** Women at high risk for breast cancer who had nonsuspicious mammograms and clinical breast examinations underwent nipple aspiration followed by lavage of fluid-yielding ducts. All statistical tests were two-sided. **Results:** The 507 women enrolled included 291 (57%) with a history of breast cancer and 199 (39%) with a 5-year Gail risk for breast cancer of 1.7% or more. Nipple aspirate fluid (NAF) samples were evaluated cytologically for 417 women, and ductal lavage samples were evaluated for 383 women. Adequate samples for diagnosis were collected from 111 (27%) and 299 (78%) women, respectively. A median of 13 500 epithelial cells per duct (range, 43–492 000 cells) was collected by ductal lavage compared with a median of 120 epithelial cells per breast (range, 10–74 300) collected by nipple aspiration. For ductal lavage, 92 (24%) subjects had abnormal cells that were mildly (17%) or markedly (6%) atypical or malignant (<1%). For NAF, corresponding percentages were 6%, 3%, and fewer than 1%. Ductal lavage detected abnormal intra-ductal breast cells 3.2 times more often than nipple aspiration (79 versus 25 breasts; McNemar's test,  $P < .001$ ). No serious procedure-related adverse events were reported. **Conclusions:** Large numbers of ductal cells can be collected by ductal lavage to detect atypical cellular changes within the breast. Ductal lavage is a safe and well-tolerated procedure and is a more sensitive method of detecting cellular atypia than nipple aspiration. [J Natl Cancer Inst 2001;93:1624–32]

Breast cancer is the most common malignancy diagnosed among women in the United States and is the second leading cause of cancer death in women (1). Tamoxifen and prophylactic mastectomy decrease the incidence of breast cancer in high-risk women (2–4). Because the accurate determination of an individual woman's risk of breast cancer remains difficult, a technique is needed that can reliably identify women who have biologic markers associated with an increased (or decreased) risk of breast cancer.

The vast majority of breast cancers begin in the epithelium lining the ductal system of the breast (5,6). Invasive breast cancer originating in ductal epithelial cells is believed to result from progressive molecular and morphologic changes, including, in the early phases, the phenotypic appearance of cellular atypia (7,8).

Two prospective studies (9,10) with long-term follow-up have shown that women with cellular atypia detected by the cytologic examination of breast specimens have a higher relative risk of developing breast cancer than women without cellular atypia. The specimens in these studies were collected by nipple aspiration (9) or by random periareolar fine-needle aspiration (10).

Ductal lavage is a procedure developed to enhance the ease and efficiency of collecting breast epithelial cells for cytologic analysis. The procedure involves the use of a microcatheter to cannulate ductal orifices on the nipple identified by fluid droplets elicited by nipple aspiration. Each fluid-yielding duct is

*Affiliations of authors:* W. C. Dooley, Institute for Breast Health, University of Oklahoma Health Sciences Center, Oklahoma City; B.-M. Ljung, E. B. King (Department of Pathology), L. J. Esserman (Department of Surgery), University of California, San Francisco; U. Veronesi, M. Cazzaniga, Istituto Europeo di Oncologia, Milan, Italy; R. M. Elledge, Breast Care Center, Baylor College of Medicine, Houston, TX; J. A. O'Shaughnessy, Baylor-Sammons Cancer Center, US Oncology, Dallas, TX; H. M. Kuerer, Department of Surgery, The University of Texas M. D. Anderson Cancer Center, Houston; D. T. Hung, Pro•Duct Health, Inc., Menlo Park, CA; S. A. Khan, Department of Surgery, Northwestern Memorial Hospital, Chicago, IL; R. F. Phillips, Metro Surgical Associates, Decatur, GA; P. A. Ganz, Division of Cancer Prevention and Control Research, Jonsson Comprehensive Cancer Center, University of California, Los Angeles; D. M. Euhus, Division of Surgical Oncology, The University of Texas Southwestern Medical Center, Dallas; B. G. Haffty, B. L. King, Department of Therapeutic Radiology, Yale University School of Medicine, New Haven, CT; M. C. Kelley, Division of Surgical Oncology, Vanderbilt University Medical Center, Nashville, TN; M. M. Anderson, Department of Surgery, Martin Luther King/Charles Drew Medical Center, Los Angeles; P. J. Schmit, Department of Surgery, Olive View—University of California at Los Angeles Medical Center, Sylmar; R. R. Clark, The Breast Care Center, Santa Barbara, CA; F. C. Kass, Cancer Center of Santa Barbara; B. O. Anderson, Department of Surgery, University of Washington, Seattle; S. L. Troyan, Department of Surgery, Beth Israel Deaconess Medical Center, Boston, MA; R. D. Arias, Department of Obstetrics and Gynecology, University of Southern California, Los Angeles; J. N. Quiring, QST Consultations, Ltd., Allendale, MI; S. M. Love, University of California, Los Angeles; D. L. Page, Department of Pathology, Vanderbilt University.

*Correspondence to:* William C. Dooley, M.D., Institute for Breast Health, University of Oklahoma Health Sciences Center, 920 Stanton Young Blvd., Rm. WP-1140, Oklahoma City, OK 73104 (e-mail: william-dooley@ouhsc.edu).

See "Notes" following "References."

© Oxford University Press

cannulated to a maximum depth of 1.5 cm, and the ductal system is infused with normal saline. Ductal effluent collected through the microcatheter is then analyzed cytologically. These results, which are independent of the risk assessed by the Gail model (11), should provide additional information to a woman about her risk of developing breast cancer.

This article reports the results of a prospective multicenter study designed to compare ductal lavage and nipple aspiration with regard to safety, tolerability, and ability to detect abnormal breast epithelial cells in women at high risk of breast cancer who have nonsuspicious mammograms and clinical breast examinations.

## SUBJECTS AND METHODS

### Eligibility

Written informed consent was obtained from all subjects before enrollment in the study. The protocol was approved by the institutional review boards/ethics committees at all 19 participating sites. Women enrolled in the study were at high risk for breast cancer as determined by a 5-year risk of invasive breast cancer development of at least 1.7%, according to the model by Gail et al. (11); a personal history of invasive breast cancer, ductal carcinoma *in situ*, or lobular carcinoma *in situ*; or a documented genetic mutation in the BRCA1 gene or the BRCA2 gene (12). In the calculation of risk, the Gail model includes the variables of current age, number of first-degree relatives with breast cancer, age at menarche, age at first live birth, number of breast biopsies, whether or not the woman has a personal history of atypical hyperplasia, and race (13). Participants were required to be at least 18 years old; there was no upper age limit.

All women were required to have had a mammogram and a clinical breast examination interpreted as not suspicious for breast cancer within 12 months before entry in the study. Women who had undergone lumpectomy without radiation therapy for a prior breast cancer were eligible, and both breasts were studied. Women who had undergone lumpectomy plus radiation therapy or mastectomy for a prior breast cancer were also eligible, but only the contralateral breast was studied. Breasts that had undergone surgery within 2 cm of the nipple were excluded because of probable proximal disruption of the ductal system. Women undergoing chemotherapy within 6 months of enrollment or any treatment with tamoxifen or raloxifene were not eligible.

### Anesthesia

Seventy-two percent of the subjects underwent the study procedures under local anesthesia in the investigator's office or outpatient facility. The method of local anesthesia was at the discretion of the investigator and the subject. Early in the study, subcutaneous periareolar injections, using a 30-gauge needle, of 1% lidocaine without epinephrine or marcaine were used in 52% (150 of 291) of the subjects undergoing lavage; this practice was subsequently abandoned because of subject discomfort. For most subjects who opted for anesthesia, EMLA cream (i.e., 2.5% lidocaine–2.5% prilocaine; Astra USA, Westborough, MA) was applied topically and then covered with an occlusive dressing for approximately 1 hour before the procedure. Sometimes, a 2% lidocaine jelly was used on the catheter tip. Approximately 1–3 mL of 1% lidocaine without epinephrine was infused into the duct after cannulation. Twenty-eight percent of the subjects underwent the study procedures in an operating room while under general anesthesia or sedation for a planned surgical procedure on the contralateral breast, such as an excisional biopsy or mastectomy.

### Nipple Aspiration

For nipple aspiration, nonsedated subjects were seated and instructed in breast self-massage, which they then performed. Subjects who underwent nipple aspiration in the operating room were supine; breast massage was performed by the physician. The nipple was dekeratinized with a mild abrasive gel (Omni Prep Skin Prep; D.O. Weaver & Co., Aurora, CO). After an approximately 1-minute massage, nipple aspiration was performed by placing a suction cup (Pro•Duct Health, Inc., Menlo Park, CA) fitted with a 20-mL syringe over the nipple and applying a small amount of suction (7–15 mL). If no nipple aspiration fluid (NAF) appeared on the nipple, the lactiferous sinus was manually compressed. Repeated efforts at breast massage and suction were attempted until fluid was

elicited or until the investigator determined that the breast would not yield fluid. Subjects whose breasts did not yield fluid on the first attempt were invited to return for up to three repeat attempts before being discontinued from the study (n = 80). If NAF was elicited, it was pooled and collected into capillary tubes and deposited into CytoLyt cell preservative (Cytoc Corp., Boxborough, MA) for fixation.

### Ductal Lavage

Ductal lavage was attempted immediately after nipple aspiration on all ducts that yielded NAF. The patient was placed in the supine position, the skin in the nipple area was cleaned with 70% alcohol, and a fenestrated sterile drape was placed over the nipple. Ductal orifices were sometimes enlarged with dilators to facilitate cannulation. A separate microcatheter (Pro•Duct Health, Inc.) was used for each duct cannulation to prevent cellular cross-contamination between different individual ductal systems. Sometimes, a 2% lidocaine jelly was used on the catheter tip. After the microcatheter was inserted to a maximum depth of 1.5 cm, a total of 1–3 mL of 1% lidocaine without epinephrine was infused intraductally in most subjects. Approximately 2–6 mL of normal saline was then infused, and the breast was compressed to facilitate recovery of ductal fluid into the collection chamber. This lavage procedure (infusion, compression, and effluent collection) was repeated multiple times, instilling a total volume of approximately 10 mL of normal saline per duct and recovering approximately 5 mL of ductal effluent per duct.

The location of each fluid-yielding duct and of each cannulated duct was carefully marked on a 64-square nipple grid. The recovered ductal effluent was placed into tubes half filled with CytoLyt solution and individually labeled for each cannulated duct.

Two similar versions of the microcatheter were used in the study. The DucWash microcatheter (Pro•Duct Health, Inc.) was used for the first two thirds of the subjects enrolled, and the Pro•Duct microcatheter was used for the remaining subjects. The DucWash catheter was often inserted after duct dilation with one to three external dilators. The Pro•Duct catheter contained an internal tapered dilator and, therefore, generally did not require other dilators to aid insertion.

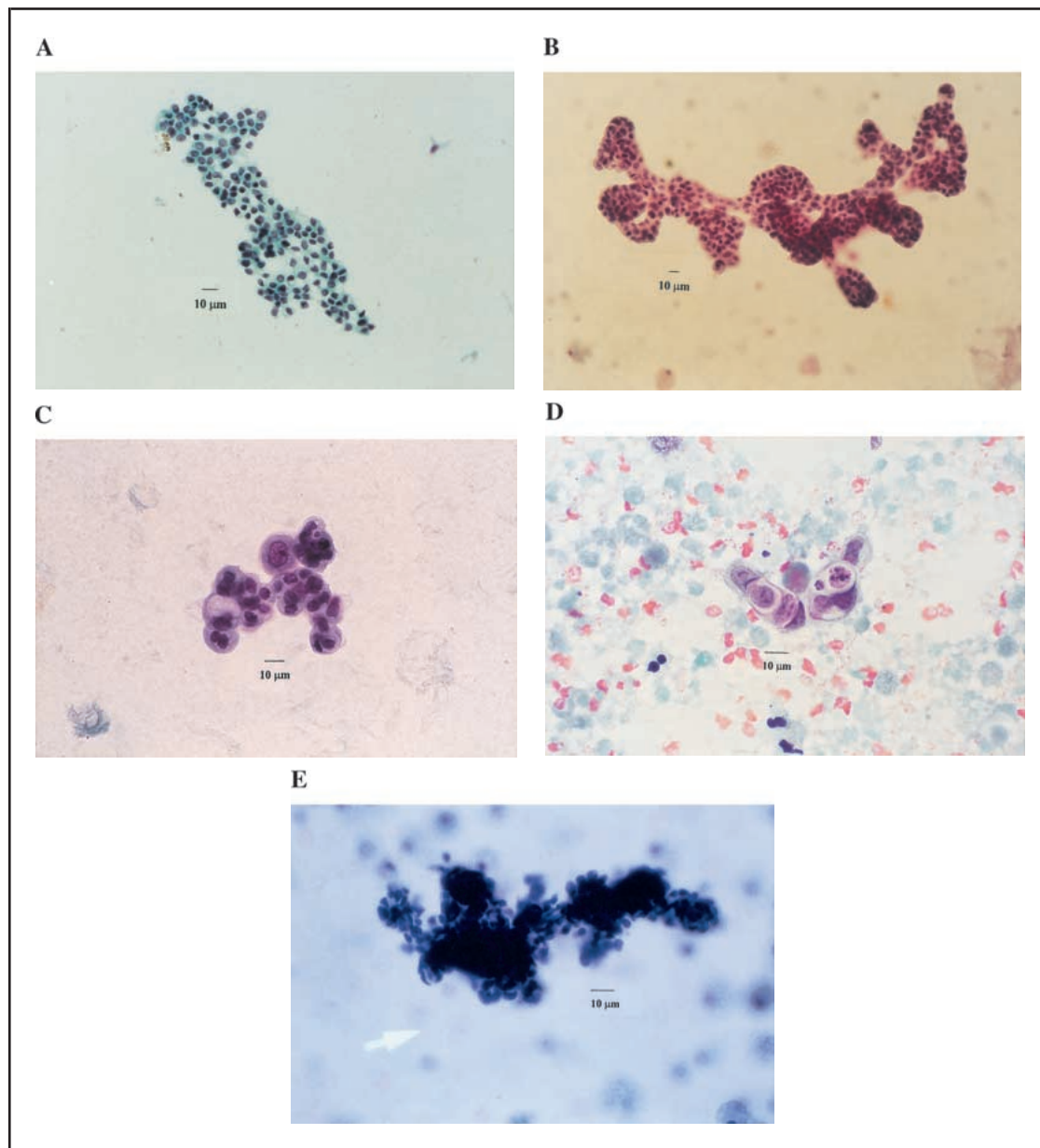
Surgeons or surgical oncologists performed most of the ductal cannulations in this trial. However, medical oncologists, an obstetrician–gynecologist, a radiologist, a radiation oncologist, a nurse practitioner, and a physician's assistant successfully performed ductal lavage in the trial.

### Cytologic Processing and Examination

All samples were shipped overnight to the University of California, San Francisco, where the samples were prepared by use of the Millipore filtration technique (Millipore Corp., Bedford, MA) and standard Pap staining (14,15). Samples from Istituto Europeo di Oncologia (Milan, Italy) were prepared at that site's cytopathology laboratory by the same methods before shipment to the University of California, San Francisco, for evaluation.

The cytologic diagnostic categories were very similar to the 1997 consensus criteria for breast fine-needle aspiration biopsy samples published by the National Cancer Institute (Bethesda, MD) (16). There were five diagnostic categories: inadequate cellular material for diagnosis (samples with <10 epithelial cells per sample or unacceptable technical quality), benign, mild atypia, marked atypia, and malignant. Markedly atypical cells have features that raise serious concerns about a possible malignant process but do not have all criteria for a definitive diagnosis of malignancy. Representative examples of the cellular categories are shown in Fig. 1.

Each NAF sample and ductal lavage sample were reviewed independently by two cytopathologists, B.-M. Ljung and E. B. King. When the diagnoses were discordant, the slides were reread jointly to obtain a unified diagnosis. Cell numbers were quantified by directly counting epithelial cell clusters (groups containing  $\geq 10$  cells), single cells, and cells in small groups (groups containing nine or fewer cells). All cells within 10 representative epithelial cell clusters were counted and averaged. If more than 10 clusters were present, the cells within 10 randomly selected clusters were counted and averaged. The average number of cells was then multiplied by the total number of clusters present. Single cells and cells in small groups were counted by selecting a representative field from each quadrant of the sample. The types and numbers of single cells and cells in small groups were determined by randomly selecting up to 200 cells and by counting and typing all observed cells. The percentage of epithelial cells present alone and in small groups was determined by this method. The total number of epithelial cells was determined by adding the number of epithelial cells in clusters and the number of those present alone and in small groups.



**Fig. 1.** Cellular cytologic diagnostic categories. **A)** Benign. **B)** Mild atypia. **C)** Marked atypia. **D)** Malignant. **E)** Cells from a terminal ductal lobular unit. Scale bars = 10  $\mu\text{m}$ .

## Data Collection and Monitoring

Subjects, investigators, and study coordinators filled out case report forms recording relevant information about the subjects' medical history, eligibility, study procedures, adverse events, and follow-up. Subjects were contacted 1 day and 2 weeks after the study day to collect information about adverse events. All data were confirmed by source verification by the study sponsor, Pro•Duct Health, Inc., and all data were double entered to ensure accuracy.

## Statistical Design

Using McNemar's test (17), we determined that a sample size of 300 breasts would be required to permit detection of a threefold difference between cytologic diagnoses from ductal lavage and nipple aspiration, with a power in excess of 95% and an  $\alpha$  value of .05. We assumed that NAF would be obtained from 75% of the subjects. Therefore, approximately 500 women would need to be enrolled to ensure 300 breasts for the within-breast comparison analysis between ductal lavage and nipple aspiration.

## Statistical Analyses

Ductal lavage and nipple aspiration results were evaluated independently, and an analysis comparing ductal lavage results with nipple aspiration results was also performed. McNemar's test (17) was used to compare cytologic results from

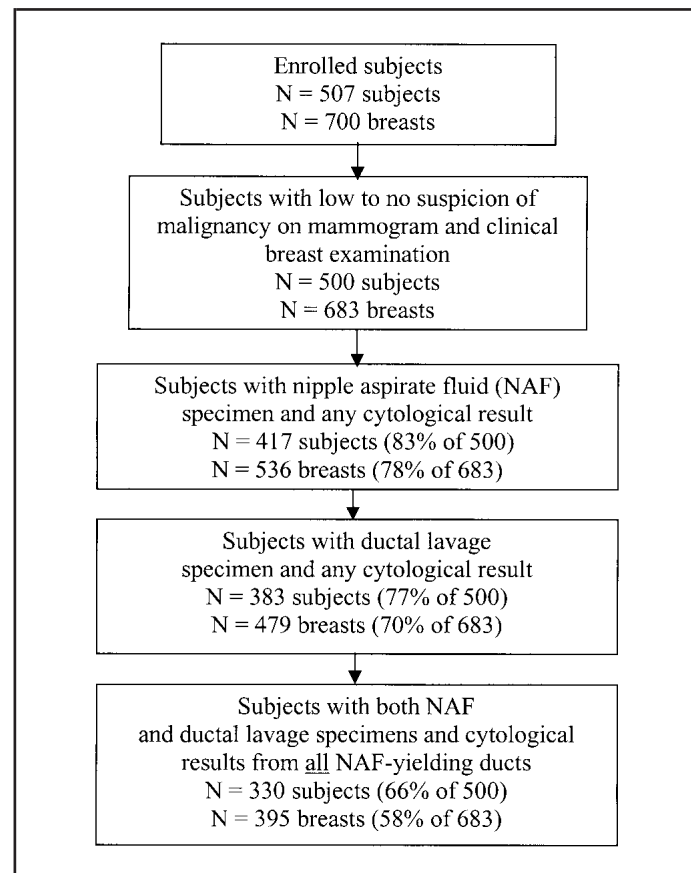
paired NAF and ductal lavage specimens. A flow chart showing the number of subjects included in each analysis is presented in Fig. 2. All statistical tests were two-sided.

## RESULTS

### Enrollment and Demographics

A total of 507 women were enrolled, and 700 of their breasts were studied. Table 1 shows the distribution of patients according to their high-risk eligibility criteria. Fifty-seven percent (291 of 507) of enrolled women had a history of breast cancer, either invasive or ductal carcinoma *in situ*, and 39% (199 of 507) had a 5-year Gail risk of breast cancer of 1.7% or more. Sixty-two percent of the subjects in the study had one breast studied, and 38% had two breasts studied. Subject demographics are listed in Table 2. The mean age was 51.9 years, and the mean 5-year risk of breast cancer development according to the Gail model (among subjects with a risk of at least 1.7%) was 3.3%. The majority of subjects enrolled in the study were white. The distribution of subjects by race was consistent with the U. S. Bureau





**Fig. 2.** Enrollment and number of subjects in the clinical trial and the number of subjects included in the major analyses.

**Table 1.** Distribution of study subjects according to factors associated with a high risk of breast cancer\*

Risk type	Total No. (%) of subjects	No. of subjects with one breast studied	No. of subjects with two breasts studied
Prior breast cancer	291 (57)	282	9
Invasive	202	200	2
DCIS	83	76	7
Unknown	6	6	0
LCIS	10 (2)	4	6
5-y Gail risk of $\geq 1.7\%$	199 (39)	28	171
BRCA1 or BRCA2 mutation	3 (<1)	0	3
Not high risk	4 (<1)	0	4
All risk categories (%)	507 (100)	314 (62)	193 (38)

\*DCIS = ductal carcinoma *in situ*; LCIS = lobular carcinoma *in situ*.

of the Census population estimates for women at the time the study was conducted (18). On data verification, four enrolled subjects were found not to have a high risk of breast cancer, and seven were found to have a suspicious or outdated mammogram or clinical breast examination.

### Procedure Tolerability and Adverse Events

Subjects who underwent nipple aspiration and ductal lavage in the office setting completed a 100-mm visual analog scale, with 0 mm representing “no pain” and 100 mm representing

**Table 2.** Patient characteristics\*

Characteristics	
Age, y	51.9 (26–81)
Age at menarche, y	12.6 (7–25)
Age at first live birth, y	25.3 (14–42)
No. of first-degree relatives with breast cancer	0.5 (0–5)
No. of previous breast biopsies	1.7 (0–18)
Previous diagnosis of atypical ductal hyperplasia, No. (%)	
No	375 (74)
Yes	96 (19)
Unknown	36 (7)
Racial heritage, No. (%)	
White	425 (84)
African-American	63 (12)
Other	11 (2)
Asian	8 (2)
5-y Gail risk, % (n = 199)	3.3 (1.7–11.0)

\*Data are either the mean (range) or number (%).

“most severe pain.” Subjects were asked to complete the scale immediately after each procedure. The median rating was 8 mm for nipple aspiration and 24 mm for ductal lavage. When asked to compare the comfort of ductal lavage with the comfort of mammography, 49% (127 of 261) of the subjects responding to the question reported that ductal lavage was at least as comfortable as mammography, including 29% reporting ductal lavage as more comfortable and 20% reporting comfort to be about the same. Fifty-one percent reported ductal lavage as less comfortable than mammography.

No serious procedure-related adverse events were reported during the clinical trial. Two subjects reported possible infections that were treated with oral antibiotics. The most common adverse events among subjects who underwent both nipple aspiration and attempted duct cannulation included breast pain (44%), ecchymoses (17% overall, but only 4% in subjects who did not receive periareolar injections of anesthetic), and breast engorgement (5%). Most adverse events were mild and of short duration.

### Nipple Aspiration

All enrolled subjects underwent nipple aspiration. Fluid-yielding ducts were identified in 84% (427 of 507) of all subjects. Only 13 of the initial 85 subjects who did not yield fluid on the first attempt underwent repeat attempts at nipple aspiration. Subsequent attempts were successful in five of the 13 subjects. The percentage of NAF-yielding breasts was 96% (135 of 141) among women who underwent nipple aspiration during general anesthesia in the operating room.

NAF-yielding ducts were identified in 83% of parous women and in 87% of nulliparous women. Nipple aspiration identified NAF-yielding ducts in 90% of premenopausal and perimenopausal women and in 78% of postmenopausal women. The average number of NAF-yielding ducts per breast was 1.5. Women with both breasts studied had, on average, a total of 2.6 NAF-yielding ducts.

Of the 427 women with fluid-yielding ducts, six did not meet the entry criteria of low to no suspicion of malignancy on mammography and clinical breast examination, and four did not have their NAF samples sent for analysis. Thus, NAF results are

reported for 417 subjects. The results of NAF cytology are presented in Table 3. The majority of subjects (73%) had NAF samples with inadequate cellular material for diagnosis, defined as fewer than 10 epithelial cells in the sample, or samples with unacceptable technical quality.

When NAF samples did have cellular material adequate for diagnosis, the samples contained a median of 120 epithelial cells per breast (range, 10–74 300 epithelial cells per breast). Epithelial cells represented a median of 13% of single cells or cells in small groups, whereas foam cells represented 77%. Only 18% of all of the NAF samples evaluated (96 of 536) contained clusters of epithelial cells (groups of 10 or more epithelial cells). Among the NAF samples with clusters, 81% (78 of 96) had five or fewer epithelial cell clusters.

Abnormal cells were detected in 10% of the subjects (41 of 417) undergoing nipple aspiration. The cytology results of NAF specimens from all subjects undergoing nipple aspiration are presented in Table 3. NAF samples from 6% (27 of 417) of the subjects had mildly atypical cells detected, and NAF samples from 3% (12 of 417) had markedly atypical cells detected. NAF samples from two subjects contained malignant cells. The overall concordance between the initial independent cytologic diagnoses of the two cytopathologists in this study for both NAF and ductal lavage specimens was 89%.

### Ductal Lavage

Ductal lavage was attempted in 426 subjects on a total of 740 NAF-yielding ducts. Successful cannulation, defined as successful catheter insertion into the ductal orifice with ductal lavage fluid sent for cytology, was achieved in 82% of ducts (610 of 740). Ninety-two percent (392 of 426) of the subjects had successful cannulation of at least one duct. Unsuccessful attempts were primarily due to an inability to cannulate the duct or to fully seat the catheter.

Surgeons attempted the most duct cannulations (477 of 740); surgeons had successful cannulation rates of 87% overall and 95% when subjects were under general anesthesia. The cannulation success rate was 82% for all investigators and was similar for both versions of the microcatheter used in the study. Parous women were cannulated successfully more often than nulliparous subjects (95% versus 84%;  $P < .001$ ).

The mean volume of normal saline infused during ductal lavage was 14 mL. The mean effluent volume collected was 5 mL. Cytologic samples were obtained from 610 ducts in 392 subjects. Samples from three subjects were lost, and six did not fulfill the mammography and/or clinical breast examination eli-

gibility criteria, leaving 383 subjects and 591 ducts available for analysis.

In contrast to the NAF results, the majority (78%) of subjects had ductal lavage samples with adequate cellular material for diagnosis (Table 3). The median cell count for the ducts in which lavage was performed with the first version of the microcatheter (DucWash) was 4000 epithelial cells per duct (range, 24–143 000). The median cell count for the ducts in which lavage was performed with the second version of the microcatheter (Pro•Duct Health, Inc.) was 13 500 epithelial cells per duct (range, 43–492 000). In contrast to NAF samples, 67% of the single cells and cells in small groups collected per duct were epithelial cells, whereas only 17% of the cells collected were foam cells. Of the 591 ductal lavage specimens evaluated, 397 (67%) contained epithelial cell clusters. In samples containing cell clusters, 68% contained more than 10 clusters per duct, and 20% contained more than 100 clusters per duct.

Abnormal cells were detected in 24% (92 of 383) of all subjects who underwent ductal lavage (Table 3). Of those subjects, 17% (66 of 383) had cells classified as mild atypia, and 6% (24 of 383) had cells classified as marked atypia. Two subjects had ductal lavage samples with malignant cells. These data include only the highest grade (most severe) cytologic diagnosis per subject. The identical dataset presented by duct, rather than by subject, is also presented in Table 3.

Of the 114 subjects with any abnormal result on NAF and/or ductal lavage, nine (8%) had abnormal results in both breasts. One of these subjects had two abnormalities in the same breast: one duct with markedly atypical cells and one with mildly atypical cells. Eight additional subjects had multiple ducts with abnormal results within a single breast. Two women had markedly atypical cells in two ducts in the same breast.

The median number of epithelial cells per duct for each of the cytologic diagnoses increased with increasing cytologic grade (Kruskal–Wallis test,  $P < .001$ ). The median number of epithelial cells collected per duct was as follows: benign = 4000; mild atypia = 13 400; marked atypia = 40 000; and malignant = 83 000.

Subjects with abnormal findings on evaluation of ductal lavage specimens were not statistically significantly different from those with benign or inadequate cellular material for diagnosis with respect to history of breast cancer, number of first-degree relatives with breast cancer, age at menarche, biopsy history, or use of hormone replacement therapy.

### Comparison of Nipple Aspiration and Ductal Lavage

Twenty-seven percent of the subjects had at least one NAF sample with adequate material for cellular diagnosis, whereas 78% of the subjects had at least one ductal lavage sample with adequate cellular material for diagnosis (McNemar's test,  $P < .001$ ). In samples adequate for diagnosis, the median number of epithelial cells collected per breast by NAF was 120, and the median number of epithelial cells collected per duct by the second version of the ductal lavage microcatheter was 13 500 (Kruskal–Wallis test,  $P < .001$ ). Eighteen percent of NAF samples contained epithelial cell clusters compared with 67% of ductal lavage samples (Cochran–Mantel–Haenszel test,  $P = .001$ ), and when clusters were present, lavage samples contained statistically significantly more clusters than did NAF samples (Cochran–Mantel–Haenszel test,  $P = .001$ ).

To directly compare nipple aspiration and ductal lavage with

**Table 3.** Overall cytologic diagnoses\*

Cytologic diagnosis	Frequency (%)		
	NAF by subject (n = 417)	Ductal lavage by subject (n = 383)	Ductal lavage by duct (n = 591)
ICMD	306 (73)	84 (22)	173 (29)
Benign	70 (17)	207 (54)	313 (53)
Mild atypia	27 (6)	66 (17)	77 (13)
Marked atypia	12 (3)	24 (6)	26 (4)
Malignant	2 (<1)	2 (<1)	2 (<1)

\*ICMD = inadequate cellular material for diagnosis; NAF = nipple aspirate fluid.

respect to their ability to collect cells and to detect cellular atypia, cytologic diagnoses were compared in paired NAF and ductal lavage samples. Only breasts in which all fluid-yielding ducts were successfully cannulated were included in this comparative analysis. Consequently, this analysis included 395 breasts from 330 subjects (Table 4). For the ductal lavage results, the highest grade cytologic diagnosis per breast was used. Among these paired samples, adequate NAF samples were obtained from 82 breasts, and adequate ductal lavage samples were obtained from 284 breasts. Thus, ductal lavage was 3.5 times more likely than nipple aspiration (McNemar's test,  $P < .001$ ) to result in a cytologic diagnosis (Table 4).

When the comparison included breasts in which either nipple aspiration or ductal lavage resulted in a diagnosis of mildly atypical, markedly atypical, or malignant cells, ductal lavage was 4.7 times more likely than nipple aspiration to result in a higher grade abnormal diagnosis (66 versus 14 breasts; McNemar's test,  $P < .001$ ). The complete concordance data for breasts with an NAF sample and ductal lavage of all fluid-yielding ducts are presented in Table 4.

Ductal lavage detected abnormal intraductal breast cells better than NAF (79 versus 25 breasts, respectively; McNemar's test,  $P < .001$ ). This difference was calculated by adding the number of breasts with concordant abnormal diagnoses to the number of breasts with abnormal diagnoses detected by ductal lavage alone and comparing the sum with the number of breasts with

concordant abnormal diagnoses plus the breasts with abnormal diagnoses detected by NAF alone.

When the comparison of nipple aspiration and ductal lavage was broadened to include paired breasts in which cannulation of any fluid-yielding duct was successful and unpaired breasts that were missing either nipple aspiration or ductal lavage results, the findings were very similar to those described above. When all NAF results are compared with all ductal lavage results (Table 3), NAF detected mildly atypical, markedly atypical, or malignant (abnormal) cytologic findings in 10% (41 of 417) of women, whereas ductal lavage detected abnormal cytologic findings in 24% (92 of 383) of women (Fisher's exact test,  $P < .001$ ). Twenty of the 41 women with abnormal cytologic findings from NAF had normal findings from ductal lavage, whereas 71 of 92 women with abnormal cytologic findings from ductal lavage had normal findings from NAF. More important, in this analysis, 27% of the subjects had an NAF sample adequate for diagnosis, whereas 78% of the subjects had a ductal lavage sample adequate for diagnosis.

## DISCUSSION

Our findings demonstrate that ductal lavage is a safe and well-tolerated procedure that can be used to collect and detect atypical and malignant breast ductal epithelial cells in women at high risk of developing breast cancer. A cytologic diagnosis of atypical epithelial cells provides additional information to help a woman assess her risk for developing breast cancer and to assist her in determining if she is a candidate for risk reduction therapy or closer surveillance.

The percentage of women in whom NAF was elicited in this study (84%) was higher than the percentage of women yielding NAF (50%) in the study by Wrensch et al. (9), the largest previously reported series in which NAF was examined. The current study included only women at high risk for the development of breast cancer, whereas women in the study by Wrensch et al. were not selected on the basis of breast cancer risk. In the study by Wrensch et al., women with NAF were found to have a higher relative risk of developing breast cancer than women without NAF. The higher prevalence of NAF in the current study is consistent with previous observations that women with NAF have an increased risk of breast cancer and with the fact that subjects in the current study were selected by their high-risk status.

It is important to stress that only ducts that yielded NAF were targeted for ductal lavage in this study. The rationale for focusing ductal lavage on NAF-yielding ducts is based on the data demonstrating that women with NAF have a higher relative risk of developing breast cancer than those without NAF (9). In addition, fluid-yielding ducts are easier to locate for cannulation than are non-fluid-yielding ducts. Although it is hypothesized that proliferating abnormal ductal epithelial cells will produce more fluid than normal intact epithelium, the rate of atypia has not been studied in non-fluid-yielding ducts.

The much higher number of subjects with samples adequate for diagnosis and the higher grade diagnoses obtained by ductal lavage demonstrate that ductal lavage is superior to nipple aspiration in detecting intraductal cellular abnormalities. Furthermore, the marked increase in ductal epithelial cell recovery by ductal lavage compared with nipple aspiration is a diagnostic

**Table 4.** Paired nipple aspirate fluid (NAF) and ductal lavage results from 395 breasts\*

<i>Any diagnosis versus inadequate cellular material for diagnosis (ICMD)†</i>				
Diagnosis by examination of NAF specimen	Diagnosis by examination of ductal lavage specimen, No.			Total
	ICMD	Any diagnosis		
ICMD	97	216		313
Any diagnosis	14	68		82
Total	111	284		395

<i>Paired NAF versus ductal lavage cytologic diagnosis concordance (by breast)‡</i>					
Diagnosis by examination of NAF specimen	Diagnosis by examination of ductal lavage specimen, No.				
	ICMD	Benign	Mild atypia	Marked atypia	Malignant
ICMD	97	165	39	12	0
Benign	11	32	12	2	0
Mild atypia	3	7	5	1	0
Marked atypia	0	1	3	3	0
Malignant	0	0	0	0	2

\*Concordance tables for 395 breasts distributed by NAF cytologic diagnoses (rows) versus ductal lavage cytologic diagnoses (columns). This paired analysis is performed by breast rather than by subject and includes only breasts that had an NAF sample and ductal lavage samples from all fluid-yielding ducts.

†Ductal lavage resulted in any diagnosis in 72% [(216 + 68)/395] of breasts, whereas NAF resulted in any diagnosis in 21% [(14 + 68)/395] of breasts. Ductal lavage was 3.5 times more likely than nipple aspiration (284 versus 82 breasts; McNemar's test,  $P < .001$ ) to result in a cytologic diagnosis. ICMD = inadequate cellular material for diagnosis.

‡Concordant diagnoses for paired NAF and ductal lavage samples are in **boldface type**. Ductal lavage resulted in higher grade diagnoses than NAF for the breasts whose values are above the diagonal, and NAF resulted in higher grade diagnoses for breasts whose values are below the diagonal.



advantage of ductal lavage. Table 4 shows a total of 66 women who had a higher grade diagnosis by ductal lavage than by NAF, including 14 with markedly atypical cytology by ductal lavage, but either inadequate or benign samples by NAF.

Ductal lavage collected statistically significantly more breast epithelial cells than did nipple aspiration. When NAF samples are abnormal, the atypical cells available for interpretation are typically very few, often fewer than 10 cells, and generally represent fewer than 5% of the total cell population. Lavage samples may also contain only 5% abnormal cells, but because so many more cells are obtained, statistically significantly more abnormal cells are available for study than that in NAF samples, which allows for more confident cytologic diagnoses. These highly cellular ductal lavage samples should aid the assessment of molecular markers (19) and enable the progression of cellular changes associated with increased or decreased breast cancer risk to be investigated over time. Of interest, samples with abnormal cells contained the highest total number of epithelial cells, consistent with the presumption that the more abnormal the epithelium, the greater the number of exfoliated cells available to lavage.

The epithelial cells seen in NAF samples and in ductal lavage samples are morphologically indistinguishable. However, ductal lavage samples often contain many more clusters of cells than NAF samples, including clusters with architectural features of terminal ductal lobular units. Similar clusters are not seen in NAF specimens. If we had used nipple aspiration alone in this study to screen women at high risk of breast cancer for cellular atypia, only 27% of the women would have had samples adequate for diagnosis. Conversely, ductal lavage alone resulted in at least one sample adequate for diagnosis in 78% of the women. More important, this adequate sample rate reflects the inclusion of all training cases when investigators were less proficient with the ductal lavage technique.

Abnormal cytologic diagnoses, including mild atypia, marked atypia, or malignancy, were detected in 24% of all high-risk women who underwent ductal lavage compared with only 10% of all women who underwent nipple aspiration (Fisher's exact test,  $P < .001$ ). In a previous study of random periareolar fine-needle aspirates in 480 women at high risk of breast cancer (10), hyperplasia with atypia was described in 21% of the subjects, although no malignant cytologic findings were reported.

Only two subjects in the current study had markedly atypical cells detected in more than one duct. When marked atypia was present, it was not uncommon to also observe mild atypia in other ducts, implying more widespread proliferative breast disease in some of these high-risk women.

Nipple aspiration (9,20-23), filling breast ducts with fluid to facilitate the collection of ductal epithelial cells (20,24), and breast cytology (9,10,14,20-26) are not new concepts. However, clinicians have previously been unable to routinely collect large numbers of cells for confident cytologic diagnosis. The current study demonstrates that ductal lavage is a safe and efficient method for collecting breast ductal epithelial cells. While ductal perforation may occur during lavage, it has been reported previously during galactography and has no known adverse consequence (27).

Published studies (9,10) with long-term follow-up clearly demonstrate that women with atypical breast ductal epithelial cells are at increased risk for developing breast cancer.

Wrensch et al. (9) followed 2300 women for an average of 12.7 years to examine the association between atypical findings on analysis of NAF and subsequent development of breast cancer. In their study, the finding of cellular atypia in NAF specimens was associated with a 4.9-fold increase in the relative risk of subsequently developing breast cancer. The combination of cellular atypia and a family history of breast cancer in that study raised the relative risk of developing breast cancer to 18-fold (9).

The current study was designed to replicate as much as possible the cytologic diagnostic criteria for the analysis of fine-needle aspirates and NAF samples. Specifically, the criteria used to evaluate ductal lavage specimens in this study were very similar to the consensus criteria of the National Cancer Institute established in 1997 (16) for the analysis of fine-needle aspirates and the criteria used for nipple aspirate cytology described by King et al. (28,29) and used by Wrensch et al. (9).

Cell preparation and staining methods used in the study by Wrensch et al. (Millipore filtration and Pap staining) were also used in this study. Finally, one of the cytopathologists (E. B. King) who interpreted the specimens in this ductal lavage study also interpreted all of the specimens in the study by Wrensch et al. (9). The population of cells defined as "atypical" in the ductal lavage specimens reported in this study is likely to represent a population of cells very similar to that referred to as "atypical hyperplasia" in references (9,28,29) above and are likely to have a similar prognostic value.

Fabian et al. (10) have reported recently that the relative risk of developing breast cancer in a cohort of high-risk women with epithelial hyperplasia with atypia detected cytologically in random breast fine-needle aspirates was fivefold higher than that of women without such a diagnosis. In that study, 15% of high-risk women with a 10-year Gail risk of greater than 4% and hyperplasia with atypia developed breast cancer within 3 years. More important, an elevated Gail risk and the finding of hyperplasia with atypia on fine-needle aspiration were independent predictors of breast cancer risk (10).

The similarity in the increased relative risk of developing breast cancer conferred by the diagnosis of cellular atypia in the study by Wrensch et al. (9) (a 4.9-fold increase) and in the study by Fabian et al. (10) (a 5.0-fold increase) is notable. Also of interest is the similarity between the increased relative risks conferred by atypical cytology in these studies and the risk of developing breast cancer conferred by the pathologic diagnosis of atypical hyperplasia on biopsies documented in independent studies (30-32). These studies report that the pathologic diagnosis of atypical hyperplasia confers a 3.7- to 5.3-fold increase in the relative risk of developing breast cancer.

When coupled with a family history of breast cancer, atypical hyperplasia confers an additional 11- to 22-fold increase in the relative risk of developing breast cancer (30,32). This increase again parallels the increased relative risk observed for women with atypical ductal cytology detected in NAF coupled with a family history of breast cancer (18-fold increase) in the study by Wrensch et al. (9). The consistency and similarity of the relative risks of developing breast cancer conferred by atypical ductal epithelial cell cytology and pathologic atypical hyperplasia in these independent studies are striking. These results strongly suggest that ductal epithelial cell atypia is associated with an increased risk of breast cancer. Thus, the identification of this

atypia may be used to further stratify women at elevated risk of breast cancer.

The identification of cellular atypia by ductal lavage and the increased relative risk it confers take on further potential importance in light of the findings of the National Surgical Adjuvant Breast and Bowel Project P-1 Study (2). That study of 13 388 women who had a 5-year Gail risk of 1.7% or more demonstrated that administration of tamoxifen could reduce the risk of invasive breast cancer by 49% ( $P < .001$ ). More important, in women with a history of atypical hyperplasia, tamoxifen reduced the risk of breast cancer by 86%.

However, how to balance the benefits and risks of tamoxifen is still a matter of debate, particularly for women 50 or more years of age. Ductal lavage can provide a woman with some information about whether she has evidence of cellular changes associated with increased breast cancer risk at the time she is weighing the risks and benefits of tamoxifen therapy. However, it must be emphasized that the relationship between a benign ductal lavage cytologic result and the likelihood that a woman will not develop breast cancer is unknown and warrants further study.

Ductal lavage also offers the potential opportunity to follow a specific ductal system over time and to identify it should surgical therapy be indicated. Indeed, in the current clinical trial, the feasibility of utilizing ductal lavage to detect and direct the surgical resection of several ductal carcinomas *in situ* that were occult on mammography and clinical breast examination has been demonstrated. To date, 11 subjects with abnormal ductal lavage findings have had additional imaging and breast surgery to further evaluate their cytologic findings. Four of the subjects have had a pathologically confirmed ductal carcinoma *in situ* located in the same region as the ductal system with the abnormal ductal lavage findings. It is important to note, however, that the sensitivity and specificity of ductal lavage for cancer detection have not yet been determined. Follow-up data on all available subjects from this trial will be collected under institutional review board-approved research and published in the future.

In summary, ductal lavage is a safe and well-tolerated method of accessing specific milk ducts to collect and detect atypical and malignant cells within the breast. Ductal lavage is statistically significantly more sensitive than nipple aspiration for the detection of cellular atypia. Detection of intraductal cellular abnormalities can provide women at elevated risk for breast cancer and their physicians additional information to aid their decision about risk-reduction therapy and ongoing surveillance.

## REFERENCES

- (1) Greenlee RT, Hill-Harmon MB, Murray T, Thun M. Cancer statistics, 2001. *CA Cancer J Clin* 2001;51:15–36.
- (2) Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst* 1998;90:1371–88.
- (3) Hartmann LC, Schaid DJ, Woods JE, Crotty TP, Myers JL, Arnold PG, et al. Efficacy of bilateral prophylactic mastectomy in women with a family history of breast cancer. *N Engl J Med* 1999;14:77–84.
- (4) Hartmann LC, Sellers TA, Schaid DJ, Nayfield S, Grant CS, Bjoraker JA, et al. Clinical options for women at high risk for breast cancer. *Surg Clin North Am* 1999;79:1189–206.
- (5) Wellings SR, Jensen HM, Marcum RG. An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J Natl Cancer Inst* 1975;55:231–73.
- (6) Wellings SR. A hypothesis of the origin of human breast cancer from the terminal ductal lobular unit. *Pathol Res Pract* 1980;166:515–35.
- (7) Waldman FM, DeVries S, Chew KL, Moore DH 2<sup>nd</sup>, Kerlikowske K, Ljung BM. Chromosomal alterations in ductal carcinomas *in situ* and their *in situ* recurrences. *J Natl Cancer Inst* 2000;92:313–20.
- (8) Lenington WJ, Jensen RA, Dalton LW, Page DL. Ductal carcinoma *in situ* of the breast. Heterogeneity of individual lesions. *Cancer* 1994;73:118–24.
- (9) Wrensch MR, Petrakis NL, King EB, Miike R, Mason L, Chew K, et al. Breast cancer incidence in women with abnormal cytology in nipple aspirates of breast fluid. *Am J Epidemiol* 1992;135:130–41.
- (10) Fabian CJ, Kimler BF, Zalles CM, Klemp JR, Kamel S, Zeiger S, et al. Short-term breast cancer prediction by random periareolar fine-needle aspiration cytology and the Gail risk model. *J Natl Cancer Inst* 2000;92:1217–27.
- (11) Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, et al. Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *J Natl Cancer Inst* 1989;81:1879–86.
- (12) Claus EB, Schildkraut JM, Thompson WD, Risch NJ. The genetic attributable risk of breast and ovarian cancer. *Cancer* 1996;77:2318–24.
- (13) Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, et al. Validation studies for models projecting the risk of invasive and total breast cancer incidence. *J Natl Cancer Inst* 1999;91:1541–8.
- (14) Papanicolaou GN, Holmquist DG, Bader GM, Falk EA. Exfoliative cytology of the human mammary gland and its value in the diagnosis of cancer and other diseases of the breast. *Cancer* 1958;11:377–409.
- (15) Barrett DL, King EB. Comparison of cellular recovery rates and morphologic detail obtained using membrane filter and cytocentrifuge techniques. *Acta Cytol* 1976;20:174–80.
- (16) The uniform approach to breast fine-needle aspiration biopsy. NIH Consensus Development Conference. *Am J Surg* 1997;174:371–85.
- (17) McNemar Q. Psychological statistics. 3<sup>rd</sup> ed. New York (NY): John Wiley & Sons, Inc.; 1962.
- (18) Population estimates program, Population Division, U.S. Bureau of the Censuses. Washington (DC): July 1, 1999 (<http://www.census.gov/population/estimates/nation/intfil3-1.txt>).
- (19) Evron E, Dooley WC, Umbricht CB, Rosenthal D, Sacchi N, Gabrielson E, et al. Detection of breast cancer cells in ductal lavage fluid by methylation-specific PCR. *Lancet* 2001;357:1335–6.
- (20) Sartorius OW, Smith HS, Morris P, Benedict D, Friesen L. Cytologic evaluation of breast fluid in the detection of breast disease. *J Natl Cancer Inst* 1977;59:1073–80.
- (21) Buehring GC. Screening for breast atypias using exfoliative cytology. *Cancer* 1979;43:1788–99.
- (22) King EB, Chew KL, Petrakis NL, Ernster VL. Nipple aspirate cytology for the study of breast cancer precursors. *J Natl Cancer Inst* 1983;71:1115–21.
- (23) Sauter ER, Ross E, Daly M, Klein-Szanto A, Engstrom PF, Sorling A, et al. Nipple aspirate fluid: a promising non-invasive method to identify cellular markers of breast cancer risk. *Br J Cancer* 1997;76:494–501.
- (24) Hou M, Tsai K, Lin H, Chai C, Liu C, Huang T. A simple intraductal aspiration method for cytodagnosis in nipple discharge. *Acta Cytol* 2000;44:1029–34.
- (25) Dunn JM, Lucrative ME, Wood SJ, Munford A, Webb AJ. Exfoliate cytology in the diagnosis of breast disease. *Br J Surg* 1995;82:789–91.
- (26) Nathan M. Early diagnosis of a breast neoplasm by histological examination of bloody nipple discharge. *Critique. Paris (France)*:1914;60:38–9.
- (27) Slawson SH, Johnson BA. Ductography: how to and what if? *Radiographics* 2001;21:133–50.
- (28) King EB, Barrett D, King MC, Petrakis NL. Cellular composition of the nipple aspirate specimen of breast fluid. I. The benign cells. *Am J Clin Pathol* 1975;64:728–38.
- (29) King EB, Barrett D, Petrakis NL. Cellular composition of the nipple aspirate specimen of breast fluid. II. Abnormal findings. *Am J Clin Pathol* 1975;64:739–48.
- (30) Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med* 1985;312:146–51.



- (31) London SJ, Connolly JL, Schnitt SJ, Colditz GA. A prospective study of benign breast disease and the risk of breast cancer. *JAMA* 1992;267:941-4.
- (32) Dupont WD, Parl FF, Hartmann WH, Brinton LA, Winfield AC, Worrell JA, et al. Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. *Cancer* 1993;71:1258-65.

## NOTES

W. C. Dooley and B.-M. Ljung contributed equally to this article.

*Editor's note:* Pro•Duct Health, Inc., sponsored this clinical study, and representatives of Pro•Duct Health, Inc., co-authored, edited, and reviewed all or

part of this article. D. T. Hung is an employee of and holds stock options in Pro•Duct Health, Inc., the maker of the InDuct™ breast aspirator and the InDuct™ breast microcatheter. L. J. Esserman is on the scientific advisory board of Pro•Duct Health, Inc., but has no financial interest in the company. S. M. Love is the Founder and Director of Pro•Duct Health, Inc. D. L. Page and B.-M. Ljung hold stock options in Pro•Duct Health, Inc., and are paid consultants for this company.

We thank Jane Thomas (Pro•Duct Health, Inc.) and Stephanie Demming (The University of Texas M. D. Anderson Cancer Center) for their assistance in the preparation and editing of this manuscript.

Manuscript received March 26, 2001; revised August 27, 2001; accepted September 6, 2001.