

Running Head: TOWARD NEW CRITERIA

Toward New Criteria for the Laboratory, Clinical, and Presumptive Diagnosis of UTI

Emily LaMond SN and Nancy Ryan-Wenger PhD, RN

The Ohio State University

College of Nursing

August 3, 2005

Toward New Criteria for the Laboratory, Clinical, and Presumptive Diagnosis of UTI

Chapter 1

Introduction

Urinary tract infections (UTIs) are the most common bacterial infection, accounting for 25% of all infections. UTIs occur in all populations and ages, however, infection is most common in women, especially sexually active women. One half of all women will experience a UTI in their lifetime, and one in three women will receive antimicrobial therapy for a UTI. In addition, the financial impact is enormous with costs exceeding \$1.6 billion for community acquired UTI (Foxman, 2003).

Escherichia coli and *Enterobacteriaceae* are the most common causative agents of UTIs. Unfortunately, there is an increasing antibiotic resistance to *E.coli*. Unnecessary treatment of false positive diagnoses contributes to this problem. Some clinicians consider symptom analysis and a positive urine culture to be accurate diagnostic tools, however most outpatient-setting diagnoses are made without a culture. Other diagnostic techniques include microscopy and urine dipstick. Optimal specimen collection, processing, and interpretation should provide a clinician with a clear-cut answer, however, no method is exact. Many discrepancies between diagnostic criteria exist, making accurate diagnosis difficult. Therefore, the reliability and validity of diagnostic techniques needs further evaluation (Graham & Galloway, 2001).

Pathophysiology of UTI

Urinary tract infection (UTI) is defined as the presence of microbial pathogens within the urinary tract. Bacterial contamination of normally sterile urine occurs by the retrograde movement of microorganisms, from the perineum through the urethra. Females have a shorter urethra compared to males, explaining the 14 times higher incidence of UTI in women. UTIs occur anywhere along the urinary tract, including the kidneys, prostate, bladder, ureter, and urethra. Classification of UTIs depends on the infection site. Cystitis, inflammation of the

bladder, is the most common. Infection of the renal pelvis and interstitium is referred to as pyelonephritis. Pyelonephritis is spread by ascending microorganisms along the ureters or the bloodstream. Bacteriuria simply indicates the presence of infection somewhere in the urinary tract (Huether, McCane, 2000).

An uncomplicated UTI is an infection of a normal genitourinary tract with no prior instrumentation, whereas a complicated UTI occurs in structurally or functionally abnormal tracts, such as those with indwelling catheters. Isolation of bacteria in significant quantities consistent with infection, but without symptoms, is referred to as asymptomatic bacteria. Uncomplicated UTIs can be symptomatic or asymptomatic. Complicated UTIs are frequently asymptomatic (Foxman, 2003).

An explanation of the pathophysiologic process of UTI reveals the significance of the indicators that are used in the diagnosis of UTI. Indicators of UTI include symptoms, bactericidal, nitrites, WBC, leukocyte esterase, and red blood cells.

1. Symptoms

Symptoms analysis is often used in diagnosis and treatment of UTI. The most common symptoms are frequency, urgency, painful urination, sensation of having to urinate after urination, suprapubic pain, and low back pain. Inflammation of the bladder due to bacteria decreases the bladder's capacity. Even small amounts of urine cause discomfort, which leads to frequency and urgency. Frequency is the need to urinate more often and urgency is a sudden, compelling urge to urinate. Frequency and urgency coincide with one another. Pain, discomfort, and a burning sensation, referred to as painful urination, are caused by infection somewhere in the urinary tract (Saber, 2004). Suprapubic pain, describe as pressure or discomfort in the abdomen midline just above the pubic bone, may be caused by muscle spasms in this region. Back pain can be indicative of a more serious upper UTI involving the kidneys (Saber, 2004)

2. Bacteruria

The most common microorganism, contributing to 90% of UTIs, is *Escherichia coli*. Less

frequent causative microorganisms include *Klebsiella*, *Proteus*, *Pseudomonas*, and *Staphylococcus*. *E.coli* heads a family of enteric bacteria which are facultative anaerobic Gram-negative rods living in the intestinal tracts of humans and animals, regardless of health or illness. This bacterium is versatile and well adapted; it can grow in glucose media and in the presence or absence of oxygen. It responds remarkably well to changes in the environment, such as changes in chemicals, pH, temperature, and osmolarity. Uropathogenic *E. coli* first colonizes in feces. Its presence in the region increases the risk for it to proceed up the urinary tract and into the bladder during sexual intercourse or by other means (Todar, 2002).

Certain characteristics of *E. coli* enhance its uropathogenic virulence (Todar, 2002). Uropathogenic *E.coli* possesses the adhesin pyelonephritis associated pili [PAP] pili, called P.fimbria. P.fimbria binds to the P blood group antigen of red blood cells and on a specific galactose disaccharide found on the surfaces of uroepithelial cells. Type I fimbriae play a similar role in causing *E.coli* to adhere to uroendothelial cells. Siderophores produced by *E.coli* cause hemolysis of the red blood cells to release iron needed for survival. They also lyse lymphocytes and inhibit phagocytosis of neutrophils (Todar, 2002). The presence of bacteria in the urine is detected by microscopy, urinalysis, and urine cultures.

3. Nitrates

Normal urine contains nitrates. Gram-negative bacteria present in the urinary tract produce an enzyme, reductase, which reduces nitrate to nitrite. The presence of nitrites on urinalysis may be indicative of a UTI. The conversion of nitrates to nitrites takes about four hours; therefore a morning void provides the best sample. Gram-positive organisms and *Pseudomonas* species do not reduce urinary nitrate. This contributes to its low sensitivity. False negatives can be due to inadequate numbers of bacteria and diluted or acidic urine. If urine pH is low, nitrites may be present, but will not show up on dipstick urinalysis (Anisman, 2002).

3. White Blood Cells (WBC), Leukocyte esterase (LE)

Elevations in WBC are the result of the inflammatory response of urogenital mucosa to

colonizing bacteria. Leukocyte esterase is an enzyme made by neutrophils in response to the presence of bacteria (Anisman, 2002).

4. Red blood cells

Elevation of red cells indicates inflammation and complications anywhere in the urinary system, from the kidneys to the urethra. The reporting of the absence or presence of red cells in microscopy is usually sufficient in most clinical practices (Graham & Galloway, 2001).

Chapter 2

Literature Review

Diagnostic Indicators of UTI

The accuracy of diagnostic methods is evaluated in terms of sensitivity and specificity with respect to a gold standard. Test accuracy is dependent on two important factors, reliability and validity. Reliability refers to the repeatability of a test. A test is reliable if repeated tests yield equivalent findings and if it is precise. Precision is the smallest unit of analysis that can be measured, how much variation is present in the test itself, as well as, variation introduced by the examiner (Valanis, 1999). For instance, the counting of colony forming units on a culture and color evaluation on the dipstick is subject to the examiner's decision and bias. Observations are influenced by systematic error, distorting the true nature of events, and are therefore misleading (Valanis, 1999).

Validity is the degree to which a test accurately represents reality. Measures of validity are specificity and sensitivity. Table 1.1a describes these measures of validity. Sensitivity is the frequency in which people with a true infection test positive. Specificity is the frequency in which non-cases test negative for infection. In other words, sensitivity is a measure of true positives, whereas specificity is a measure of true negatives (Valanis, 1999). Table 1.1b displays the formulas used to calculate sensitivity and specificity (Valanis, 1999).

TABLE 1.1a RESEARCH CONCEPT: MEASURE OF VALIDITY

TEST RESULTS	INFECTION PRESENT	INFECTION ABSENT
Positive	True positive A	False positive B
Negative	False negative C	True negative D

TABLE 1.1b Sensitivity and Specificity

MEASURE	CALCULATION
Sensitivity	$(A / A + C) * 100$ (percent of true positives, predictive value of positive test)
Specificity	$(D / D + B) * 100$ (percent of true negatives, predictive value of negative test)

Diagnostic methods can also be evaluated in terms of positive predictive value and negative predictive value. Predictive values represent the probability in which test results correctly identify a disease status. Positive predictive value is the probability a person with a positive test has the infection. The probability a person without the infection has a negative test is referred to as the negative predictive value (Valanis, 1999).

Unlike sensitivity and specificity, predictive values are influenced by the prevalence of an infection in a given population in which the test is applied. Predictive values are useful in deciding whether to use a particular test in a specific population. Screening high-risk groups improves the predictive value of a positive test. Table 1.2 displays the formulas used to calculate positive predictive value and negative predictive power (Valanis, 1999).

Table 1.2 Measure of Predictive Value

Positive predictive value	$\text{True positives} / (\text{true positives} + \text{false positives}) \times 100$
Negative predictive value	$\text{True negatives} / (\text{true negatives} + \text{false negatives}) \times 100$

Different diagnostic tests can be used in the diagnosis of UTI based on the presence of bacteria or indicators of infection. Examples of tests include dipstick urinalysis, urine culture, and analysis of reported symptoms. The midstream clean catch method is standard for urine sample collection (Flach, et al., 2003). Specimens are easily contaminated if the perineum is not adequately cleansed or if urine is collected at the beginning of the stream rather than midstream. Ideally, a sample via catheter or suprapubic aspiration minimizes contamination (Flach, et al. 2003).

Urine dipstick

Urine dipstick consists of chemically treated paper, which displays different colors indicating the presence of leukocyte esterase, nitrites, blood, and protein when dipped into urine sample (Flach, et al. 2003). Various specificities and sensitivities have been reported for these components. In addition, different cut-off values for components such as leukocyte esterase and nitrites are used in diagnosis.

In many cases, dipstick urinalysis is used only as a supplement to clinical assessment if the presence or absence of infection is questionable. A study was performed to determine whether dipstick urinalysis significantly alters the accuracy of clinical assessment in the diagnosis of UTI. Four hundred patients for who UTI was under consideration and who were seen consecutively in an adult emergency department participated in the study. Clinicians first noted signs and symptoms and then estimated the probability of UTI before conducting a dipstick analysis. Probabilities were compared to the results of urine culture and dipstick urinalysis.

Most urine specimens were mid-stream ($n=376$), and 24 were obtained via catheter. Ames Multiple Reagent Strips were used for the dipstick analysis. Positive dipstick was defined by the presence of nitrites or reaction of greater than or equal to a trace of leucocytes. Negative dipstick was defined by the absence of any reaction for leukocytes and nitrites. Positive specimens yielded a growth of pathogens greater than or equal to 10,000 colony forming units per millimeter (cfu/mL) of urine. Negative specimens were classified by growth less than 10,000 cfu/mL or those that grew normal flora (mixed diptheriods, lactobacilli, staphylococci, or streptococci). UTI was defined by the presence of a positive urine culture. To measure diagnostic accuracy, McNeman's test was used to compare sensitivity and specificity before and after dipstick urinalysis (Sultana et al., 2000).

The urine culture was negative in 78% of the patients and 22% were positive. Dipstick urinalysis was positive in 30% (120) of the patients. Table 2 compares the specificity and sensitivity of the dipstick urinalysis and its specific components after a positive urine culture indicated UTI. The initial probability of UTI assigned by clinicians changed in 26% (104/400) of the patients after dipstick urinalysis. After dipstick urinalysis, 10% (23/70) of patients were assigned a higher probability of UTI, prior to being assigned a low pretest probability. Negative dipstick urinalysis occurred in 33% of the patients placed in the high pre-test probability category. In terms of clinician diagnostic performance before and after dipstick urinalysis, there was a statistically significant difference between sensitivity and specificity for all probability groups (Sultana et al., 2000).

This study concluded that dipstick urinalysis provides information to clinicians that significantly improves diagnostic accuracy compared to assessment alone. While other studies have documented the accuracy of dipstick urinalysis alone in diagnosis, this study demonstrates that dipstick urinalysis significantly augments clinical assessment in symptomatic patients. Dipstick urinalysis often agreed with assessment in the low and high probability groups. Therefore, tests are advised when uncertainty exists about a patient's low or high probability

designation of UTI (Sultana et al., 2000).

TABLE 2. SENSITIVITY AND SPECIFICITY OF THE COMPONENTS OF DIPSTICK URINALYSIS

	SENSITIVITY	SPECIFICITY
Total dipstick urinalysis	94%	26%
Blood	86%	46%
Leukocytes or nitrites	82%	84%
Leukocytes	72%	86%
Protein	71%	53%
Nitrite	48%	96%

Lack of an adequate explanation for the heterogeneity of dipstick accuracy stimulates an ongoing debate. A meta-analysis of literature from 1990 to 1999 evaluated the diagnostic accuracy of the urine dipstick test and took into consideration different pre-defined sources of heterogeneity. Criteria for inclusion included the following: focus of publication should be the diagnosis of bacteruria or UTIs; they should investigate the use of dipstick tests for nitrites and/or leukocyte esterase, and should present empirical data. Seventy-two studies were included in which 17 studied nitrites only, two examined leukocyte-esterase only, and the remaining studied combinations (Dewille et al., 2004).

Sensitivity of the dipstick test for nitrites was low (45- 60%) but specificity was higher with an 85-98% range. Sensitivity of the urine dipstick test for leukocyte-esterase was slightly higher than the dipstick test for nitrites, however specificity was lower. Combination of nitrites and leukocyte esterase on the dipstick test increased sensitivity but had different effects on specificity. Table 4 discusses the cut off criteria and accuracy of combinations of one or both of these tests in dipstick urinalysis (Dewille et al., 2004).

TABLE 4 COMPARISON OF DIFFERENT CUT-OFF VALUES IN NITRITES AND LEUKOCYTES: **ONE OR BOTH** TESTS POSITIVE

CUT OFF CRITERION	N (35)	Sensitivity	Specificity
$>10^3$ mcu/ml	2	0.45	0.62
$>10^4$ mcu/ml	11	0.67	0.78
$>10^5$ mcu/ml	27	0.80	0.80

N= number of studies

COMPARISON OF DIFFERENT CUT-OFF VALUES IN TESTS FOR **BOTH** NITRITES AND LEUKOCYTES IN URINE DIPSTICKS

CUT OFF	N	Sensitivity	Specificity
$>10^3$ mcu/ml	1	0.43	0.62
$>10^4$ mcu/ml	3	0.31	0.96
$>10^5$ mcu/ml	12	0.44	0.96

This study concludes urine dipstick test alone appears to be useful in all populations to rule out the possibility of infection if both tests of nitrites and leukocyte-esterase are negative. The combination of positive test results is sensitive in family practice, however the usefulness of only dipstick test to confirm infection remains doubtful, even when pre-test probabilities are high because of its low specificity (Dewille et al., 2004).

Studies are often conducted under stringent protocols that are not practical to carry out in a busy primary care setting. The pressure of time and other factors often compromise the accuracy of urinalysis. To determine the accuracy of urinalysis in the detection of urinary tract infection in a primary care setting, a cross sectional study was undertaken on 100 patients with UTI symptoms at the Primary Care Clinic of University Malaya Medical Center in Malaysia (Othman, Chia, & NG, 2003).

Participants submitted freshly voided urine samples after given clear instructions on how to collect a midstream urine specimen. The sample was collected in a clean urine bottle as well as a urine culture bottle. Urine dipstick tests were performed by the researchers and outcomes for

leukocyte esterase, nitrites, or red blood cells was documented. The residual urine from the container was sent to laboratory for microscopy and the urine culture bottle was sent to the microbiology lab. Significant bacteriuria was regarded as 10^5 per mL of growth of a single organism. The urine culture was the gold standard test (Othman, Chia, & NG, 2003).

In this study, the nitrites test was the most specific test in the urine dipstick (82%). These results parallel the results of many other studies. Red blood cells had the highest specificity (75%). Similar to other studies, this study also found a combination of any two or three positive tests will increase the sensitivity but decrease the specificity (Othman, Chia, & NG, 2003).

In urine microscopy, the presence of both leukocytes and bacteria give higher sensitivity levels compared to red blood cells. However, RBC has the highest specificity. In general, urine dipstick is comparable to urine microscopy. In conclusion, when in a primary care setting, urine dipstick to diagnose simple, uncomplicated UTI is appropriate (Othman, Chia, & NG, 2003).

Culture

To obtain a urine culture, urine is placed on an agar dish, incubated, and colony-forming units (cfu) are counted. In the past, a quantitative urine culture yielding greater than 100,000 colony-forming units (cfu) of bacteria per millimeter of urine was termed “significant bacteriuria”. Even in asymptomatic persons, this value has a high specificity for true infection. For that reason, it was chosen as the standard. However, many studies have established that one third of asymptomatic women have cfu counts below this level. Additionally, bacterial counts of 100 cfu per mL of urine in symptomatic women have a high positive predictive value for cystitis. Some laboratories only report counts greater than 10,000 cfu per mL, resulting in the under diagnosis of low-coliform-count infections (Orenstein & Wong, 1999). Even though urine culture is considered the “gold standard” it can result in false positives due to specimen contamination as well (Graham & Galloway, 2001).

Urine cultures are expensive and the stringent cut-off value of greater than or equal to 100,000 cfu/ml is not conclusive. An alternative method to culture is the demonstration of

significant pyuria, typically defined as 10 or more leucocytes per microlitre (10 WBC/ul) in freshly voided urine. A study was conducted with the goal of correlating the leukocyte esterase results of rapid urinalysis assay with direct microscopy for pyuria. Samples from 206 volunteer healthcare professionals were subjected to urinalysis by microscopy and rapid urinalysis assay (Moore et al., 2001).

Of the 206 samples, 74 were positive for leucocytes with rapid urinalysis assay. Of those, 39 were negative for significant pyuria on direct microscopy. Significant pyuria was defined as >10 leucocytes/ul on direct microscopy. Fifteen leucocytes/ul is referred to as “trace” on the visual scale whereas >25 leucocytes/ul is “positive” on the visual scale. When the leukocyte esterase results were correlated with the direct urine microscopy results, an assay reading of 15 leucocytes/ul had a sensitivity of 91%, specificity of 79%, and positive predictive value of 53% and a negative predictive value of 97%. This was compared to an readings of >25 leucocytes/ul which had a sensitivity of 63%, specificity of 95%, positive predictive value of 75% and a negative predictive value of 91%.. Table 5 compares the statistical significance of rapid urinalysis to that of significant pyuria on direct microscopy. This study concluded that a rapid urinalysis assay result of 25 leucocytes/ul or greater in non-pregnant, pre-menopausal females predicts significant pyuria on urine microscopy and reduces the need for urine cultures. (Moore et al., 2001).

In addition, this study found negative leukocyte esterase has a strong correlation (97% negative predictive value) with the absence of pyuria on microscopy. A negative urine dipstick is a good indicator the patient does not have pyuria. Findings also suggest that symptomatic patients with negative urine dip should receive a mid-stream urinalysis culture. There are many reasons for a false negative test of leukocyte esterase, which include high specific gravity, elevated glucose, early infection, antibiotics, or high concentration of oxalic acid. In addition, false positives may result from vaginal discharge, tuberculosis of the bladder, or vesicle calculi (Moore et al., 2001).

TABLE 5. Rapid urinalysis compared to significant pyuria on direct microscopy

Urinalysis result	Sensitivity	Specificity	FNR	FPR	PPV	NPV	Overall
15+ UL/ul	62%	79%	9%	22%	53%	97%	81
25+	75%	95%	37%	6%	75%	91%	88

FNR- false negative rate

Overall- true positives + true negatives/ total sample

FPR- false positive rate

PPV- positive predictive value

NPV- negative predictive value

Signs and Symptoms

Signs and symptoms are often used alone in the diagnosis of recurrent, uncomplicated UTIs. To determine the probability of correctly diagnosing UTIs from signs, symptoms, and examination, a study compared sensitivity and specificity probabilities derived from the results of reactive strip tests to clinical history and signs (Medina- Bombardo et al., 2003).

In this epidemiological analysis, the cut-off value for the reactive strip was > 70 leukocytes/ml, and for agar culture the cut-off was greater than or equal to 100,000 cfu/ul. Cultures were positive in 38.9% of the cases and negative in 42.4%, giving a 48% pre-test probability of having a UTI among women with incident urinary symptoms. The common symptoms of frequency, burning, tenesmus, urgency, and painful voiding had a low specificity for UTI, which was always less than 50%. In addition, no sign or symptom identified increases the likelihood of UTI. Table 3a illustrates the relative frequency, sensitivity and specificity of urinary symptoms and table 3b displays the results from the reactive strip test. In conclusion, overall findings that are strongly related to positive UTI are a positive test for pyuria and nitriuria. Coexistence of these factors increases the probability of UTI more than seven times. Evaluation of clinical signs and symptoms together with the reactive strip can help clinicians make accurate decisions (Medina- Bombardo et al., 2003).

TABLE 3a Relative frequency, sensitivity, and specificity of common symptoms described by patients during clinical interview

Symptom	Relative Frequency	Sensitivity	Specificity
Frequency	78.9%	0.90	0.23
Burning	70.5%	0.76	0.30
Urgency	65.0%	0.74	0.43
Painful voiding	57.6%	0.68	0.48
Tenesmus	67.4%	0.75	0.35
Difficulty	41.5%	0.44	0.54
Diurnal incontinence	19.7%	0.18	0.79
Nocturnal incontinence	16.1%	0.16	0.82

TABLE 3b Relative frequency, sensitivity, and specificity of reactive strip compared to urine culture

Reactive strip	Relative Frequency	Sensitivity	Specificity
Pyuria	58.0%	0.72	0.57
Nitrite test positive	22.0%	0.41	0.92

Presumptive and Self-diagnosis

In addition to practical diagnostic techniques in busy clinic settings, it should also be taken into consideration that many women, up to 30%, experience recurrent UTIs. Thus, it is important to effectively manage recurrent UTI in a convenient, safe, and cost-effective manner. Patient self-diagnosis and self-treatment of recurrent UTIs may improve patient convenience and decrease antimicrobial use. A study was conducted on a large, generalized sample of 172 women with recurrent infection were seen at a university-based health care clinic to assess the accuracy of self-diagnosis and the cure rates seen with the self-treatment of UTIs (Gupta, Hooton, Roberts, & Stamm, 2001).

Women, at least 18 years old with at least two episodes of UTI in the previous 12 months,

participated in the study. The eligible 172 women were given instruction on how to obtain a clean catch urine sample. Cultures were performed and significant bacteriuria was considered present when uropathogens were present in quantities of at least 10^2 colony forming units/ mL. Participants were given six 200-mg offloading tablets to take twice daily for three days if UTI symptoms developed. During the study period, they were instructed to use past symptoms of UTI as a guide for self-diagnosis. When symptoms developed, the participants collected a clean catch urine specimen and initiated treatment. The women were followed for up to 12 months. The only interactions between the study staff and participants occurred over monthly telephone reviews and during follow-up visits after self-diagnosis and self-treatment events (Gupta, et al 2001).

During the study, 88 of the 172 women (51%) self-diagnosed at least one episode of UTI. Pre-therapy urinalysis and culture was performed to confirm these presumed UTI episodes. Uropathogen were present in 144 cases (84%), sterile pyuria in 19 cases (11%) and no pyuria or bacteriuria in 9 cases (5%) (Gupta, et al. 2001).

In conclusion, this study provides strong evidence supporting the feasibility and safety of a self management strategy for women with recurrent UTI. 94% of the suspected 172 UTIs met the criteria for definite or probable UTI. A strategy in which health care providers give a prescription for self-diagnosis and self-treatment of UTIs in young women with a history of uncomplicated recurrent UTIs seems safe and practical. This reduces the need for repeated office visits and antimicrobial prophylaxis. However, it should be emphasized this approach should only be used in adherent women with uncomplicated UTIs and can not be generalized to the population at large. (Gupta, et, al. 2001). Symptoms of UTI may occur in other conditions such as Chlamydia, arthritis, or vaginitis, making diagnosis more difficult (Flach, et al. 2003).

Researchers have reported UTIs can be managed safely and conveniently over the phone with low recurrence rates and gynecological complications. A team of researchers led by David Vinson, MD, of the Permanente Medical Group in Sacramento evaluated data on 4,177 women

who received treatment over seven weeks through a regional advice and appointment call center at Kaiser Permanente (Managed Care, 2003).

In this study, women were screened by telephone for UTI using a highly structured interview process designed to identify women who might have other conditions. The interview process was conducted by nurse practitioners. Women who received the diagnosis of UTI were required to have one of the following symptoms for at least ten days: burning, painful urination, frequency, urgency, pressure, and/or increased urination at night or blood in the urine.

Vinson concluded telephone management of UTI is more efficient for health care providers, more convenient for patient, and comparable to office-based care. Rates of recurrent cystitis and/or pyelonephritis within six weeks were low. In addition, incidence of gynecologic infections and noninfectious urologic conditions was very low (Managed Care, 2003).

Improvements can still be made in the diagnosis of UTI by performing studies with clear inclusion and exclusion criteria and with double blind studies. Reports on the distribution of microorganisms, the way urine was collected, the time between collection and analysis, whether the first void urine was used, and who was reading the test may improve future reviews of accuracy (Deville et al., 2004). It is difficult to compare the results of different studies because there is no definition of the clinical terminology employed. For example, there is no uniform definition of the symptom dysuria although this is the term most frequently associated with UTI (Medina-Bombardo et al., 2003). Some may consider urine culture the standard diagnostic method to indicate "significant bacteriuria", however there is no defined cut-off level indicating of significant bacteriuria. It is apparent discrepancies in diagnostic criteria and cut off values exist, fueling the need for further research towards a gold standard.

Clinicians are also challenged by new research and technology. To some, diagnosis is more of an art than science, in which diagnoses are made on the basis of a hunch, intuition, or experience, rather than the scientific method. Costs and patient convenience may influence clinicians' decisions as well. Various tests and assessments may aid in the diagnosis and

treatment decisions. It is important to have diagnostic criteria for urinary tract infections so optimum therapeutic outcomes can be achieved. The current challenge is to develop guiding principles for the judicious use of antibiotics for persons with a urinary tract infection (Flach, et al. 2003).

Chapter 3

Methods

Purpose

In current practice, the gold standard for diagnosis is urine culture, although a standard cut-off value does not exist. The primary purpose of this study is to determine what combination of symptoms and/or diagnostic tests best predict the incidence of UTI in non-pregnant females in terms of sensitivity, specificity, positive predictive value, and negative predictive value. The goal is to develop a gold standard for diagnosis.

Research Questions

- 1.) At which colony forming unit (cfu) level on the urine culture are the symptoms, signs, and dipstick results the most sensitive and specific?
- 2.) Which cfu level on the urine culture yields the highest positive predictive value and negative predictive value for each sign, symptom, and dipstick result?
- 3.) What sign, symptom, or combination yields the best clinical result?

Design

A study titled, "Genitourinary Infection Self-Diagnosis for Deployed Military Women" will provide the data for this secondary analysis. Data collection will be completed by January 31, 2006. Methodological standards for *diagnostic test* research include four criteria, which we have applied to this proposed study: 1) use of a prospective design; 2) sample derived from a consecutive series of patients from a relevant clinical population; 3) blind comparison of the diagnostic test with the reference standard; 4) same reference standard applied to both positive and negative diagnoses (Reid, Lachs, Feinstein, 1995). A prospective design will be used, in

that the women enter the study with vaginal and/or urinary symptoms. A relevant clinical population “is a group of patients covering the spectrum of disease that is likely to be encountered in the current or future use of the test” (Lijmer et al., 1999). Thus, subjects should not be asymptomatic, or have unrelated conditions, but should present with mild to severe symptoms associated with the disease (i.e. bacterial vaginitis//TV, Candida vaginitis, and/or UTI). The *diagnostic test* is the urine culture.

Sample

Military women are the target population for this study. The sample will include Army and Navy women because the ultimate goal of the parent research is to develop a self-diagnosis and treatment kit for genitourinary symptoms for deployed military women. There are approximately 163,414 women in the Army (72,747 active duty; 90,667 reserves and National Guard), and 87,642 women in the Navy (54,142 active duty; 33,500 reserves) (U.S. Army, 2001 & U.S. Navy, 2001). Both active duty and reserve women are likely to be deployed for military operations. Subjects are recruited from the following sites:

- 1 Troop Medical Clinic at Fort Hood, an Army base in Texas
- 2 32nd Street Clinic at Naval Station San Diego in California
- 3 Soldier and Family Medical Clinic at Fort Bliss, an Army base in El Paso, Texas
- 4 Troop Medical Clinic at Fort Carson, an Army base in Colorado Springs, Colorado

Sample size

A traditional power analysis to determine sample size is not appropriate for diagnostic test research. Rather, sample size is determined using a method for estimating a population proportion, because sensitivity and specificity are proportions. To determine the sample size for a population proportion, we used the formula $n = z^2 pq / d^2$ where z represents the standard score at 0.05 level of significance, p represents either sensitivity or specificity, q represents $1 - p$, and d represents precision of the estimator (Ryan-Wenger & Lowe, 2000). The resulting sample sizes were then adjusted for the prevalence of the condition using the estimates from

the pilot study ($N=n / \text{prevalence}$). The sample sizes must be adjusted for prevalence in order to ensure that an adequate number of diseased women will be tested to calculate a given sensitivity and an adequate number of non-diseased women will be tested to calculate a given specificity, with a desired level of significance and sampling error (precision). If the prevalence of disease is low, then the number of diseased women sampled will need to be increased to obtain adequate sensitivity. If the prevalence of disease is high, however, then the number of non-diseased women sampled will need to be increased to ensure specificity at a desired level. Sample size estimates for sensitivity and specificity were calculated for a minimum $\pm 3\%$ precision. The overall sample size needed for a particular diagnosis is driven by desire for a sensitivity and specificity of at least 95%. Sample size calculations were based on the prevalence of urine culture diagnoses of UTI of the first 500 subjects in the current parent study. For a sensitivity and specificity of 95% and a precision of 2.45% a minimum of 310 participants were needed. However, the secondary analysis included 468 participants.

Table 4. Sample size need for particular diagnosis based on a sensitivity and specificity of at least 95%

Dx	Prevalence	n	Sensitivity	Specificity	Precision
UTI	65.4%	310	95%	95%	$\pm 2.45\%$

Calculations:

n = sample size before corrected for prevalence

$z = 1.96$ (significance level)

$p = .95$ (expected sensitivity and specificity of 95%)

$q = 1 - p = .05$

d = precision = minimum of $\pm 3.0\%$

$N = n / \text{prevalence}$

$$n = \frac{z^2 pq}{d^2} = \frac{(1.96)^2 (.95) (.05)}{9} = \frac{3.8416 (.95) (.05)}{9} = \frac{1824.76}{9} = 202.75$$

$$N = n / \text{prevalence} \quad \text{UTI} = 202.75 / .654 = 310$$

Inclusion and Exclusion Criteria

The subjects for this study will be military women with genitourinary symptoms who have a scheduled appointment or attend sick call at the clinics for diagnosis and treatment. It is expected that the majority of subjects will be active duty Army or Navy, but women from any branch of the service, active or reservists on active duty, who present consecutively at the clinic, will be invited to participate. Signed consent will be required for participation in the study. Few exclusions will be used in this study, as all military women are assumed to be deployable (although temporary deployment deferments are made for pregnancy, illness or injury). Repeat visits to the clinic by a woman who has already participated in the study will exclude the woman from participating again. Because menstrual blood interferes with diagnostic tests, menstruating women will be excluded from the study. Although the study protocol involves little risk to participants, pregnant women will be excluded. Pregnant women are not deployable until six months after delivery. Age is not an exclusion criterion for this study. Women age 17 and younger are not likely to be subjects because the minimum age for entry into the military is 18. However, many military women are ages 18, 19 and 20. There are no exclusions due to racial or ethnic group.

Procedure

A Signs and Symptom form will be used to gather information from the military women. One of three trained advanced nurse practitioners (APN) will administer the questionnaire, asking the same questions in the same order. A clean catch urine sample will be obtained from the women. The urine sample will then be used for a dipstick test and urine culture.

Instruments

Urine specimen kit.

The clinics' standard equipment for collecting clean catch urine specimens will be used. All military women will have had experience with providing such specimens (e.g. pre-enlistment or

pre-commissioning physical, annual pap smear and gynecologic exams, quadrennial physicals, and periodic mandatory drug screenings).

Urine dipstick.

Urine dipstick consists of a strip of chemically treated paper that displays different colors signifying different levels of sugar, blood, and ketones. Generally, a positive urine dipstick is defined by presence of nitrites or reaction of greater than or equal to a trace of leukocyte esterase. Negative dipstick is typically defined by the absence of any reaction for leukocyte esterase and nitrites.

Symptoms and Signs.

Symptoms were self-reported on clinical information forms initiated by the research nurse. Symptoms include: Burning, urgency, frequency, continued urge to void, myalgia, headache, nausea & vomiting. The research nurse evaluated the signs on clinical examination. Signs include suprapubic tenderness, upper abdominal tenderness, CVAT, and temperature. This form has content validity because it is derived from theory and research about UITs. Each subject is asked the same questions, in the same order, by one of three trained advanced practice nurse.

Urine culture and sensitivity.

The research APN will refrigerate the urine specimens at 2-8°C no more than 4 hours before they are transported to the hospital laboratory. Specimens are routinely transported twice daily, under refrigerated conditions, from the outlying clinics to the hospital laboratory. The military hospital laboratories at Fort Hood and Navy Station San Diego are CLIA-certified and supervised by properly licensed laboratory officers. Standardized laboratory protocols will be followed for preparation, incubation and analysis of the cultures. Results will be recorded on the patients' computerized records and printed by the research nurse within 48 hours.

Analysis.

This secondary analysis will need a total of 310 subjects with UTI, bacterial vaginitis, or

Candida vaginitis. Participants with bacterial vaginitis and Candida vaginitis will be included to serve as the controls. The 310 subjects will be randomly selected from the total 966 subjects. A series of contingency tables (2 x 2) will be constructed to calculate the sensitivity, specificity, positive predictive value, and negative predictive power for each symptom and dipstick test using the three culture levels as the standard.

Human Subjects Issues

Active duty military members are not permitted to accept payment for participation in research, unless blood specimens are taken, therefore no monetary incentive will be offered. Because the military is a hierarchical organization in which “orders” are given by superior ranks, and must be followed without question by lower ranks, it is extremely important to avoid any hint of coercion. Participation in research must be fully voluntary, as confidential as possible, given the research methodology, and allow for withdrawal at any time without penalty.

Potential Risks.

Potential risks are minimal. There is no apparent psychological, social, or legal risk to participation in the study. The medications are not experimental, but all medications have the potential for side effects. The likelihood of physical risks related to side effects of the medications is minimal, and are the same for women who do *not* participate in the study, but receive the same medications. There may be some risk of breach of confidentiality related to of the women’s participation or non-participation in the study and their diagnoses by clinic personnel. Breach of confidentiality is not unheard of in military settings as evidenced by our own previous research findings (Daniel, 1995), and is often a concern of military women who seek health care in military settings. An alternative to having the clinic as the entry point to the study would be to conduct the study in a neutral place, with only research personnel in attendance. However, that alternative would be cost-prohibitive.

Risk Protection.

Women will receive verbal and printed information from the research APN about potential

side effects of the medications, recommendations to avoid alcohol for 24 hours if metronidazole is prescribed, and signs and symptoms that should be reported immediately to a health care provider. Military women have 24-hour access to military health facilities, including the military hospital emergency room after clinic hours. Prior to implementation, clinic staff will be briefed on the research study. The importance of confidentiality and non-coercion of potential subjects will be emphasized by the site coordinators, research APNs and PIs. Other than the research APN's signature on the medical record progress notes, there will be no evidence in the medical record that the women participated in the study. It is not unusual for civilian contract nurses to be part of hospital and clinic staff, so civilian status alone is not indicative of a research study. Study forms will have only ID numbers; names will not be written on the forms. A codebook linking ID numbers with women's names, addresses and telephone numbers will be kept in a locked cabinet. This information is essential for the follow-up phase of the study. Findings will be reported for the aggregate; individuals will not be identifiable by the reported data.

Benefits of Research Participation.

Several immediate benefits to the participants include the opportunity to be examined by a female APN with expertise in women's health; education about signs and symptoms that distinguish between bacterial and yeast vaginal infections, and signs and symptoms of UTI; the opportunity to have single dose oral medications to treat their conditions, versus vaginal preparations or multiple dose regimens; a sense of empowerment from the experience of conducting a self-diagnosis (assuming their diagnosis matches the research APN's diagnosis). A long-term benefit is the knowledge that their participation is critical to developing a self-care method that may be available to them during future deployments.

Chapter 4

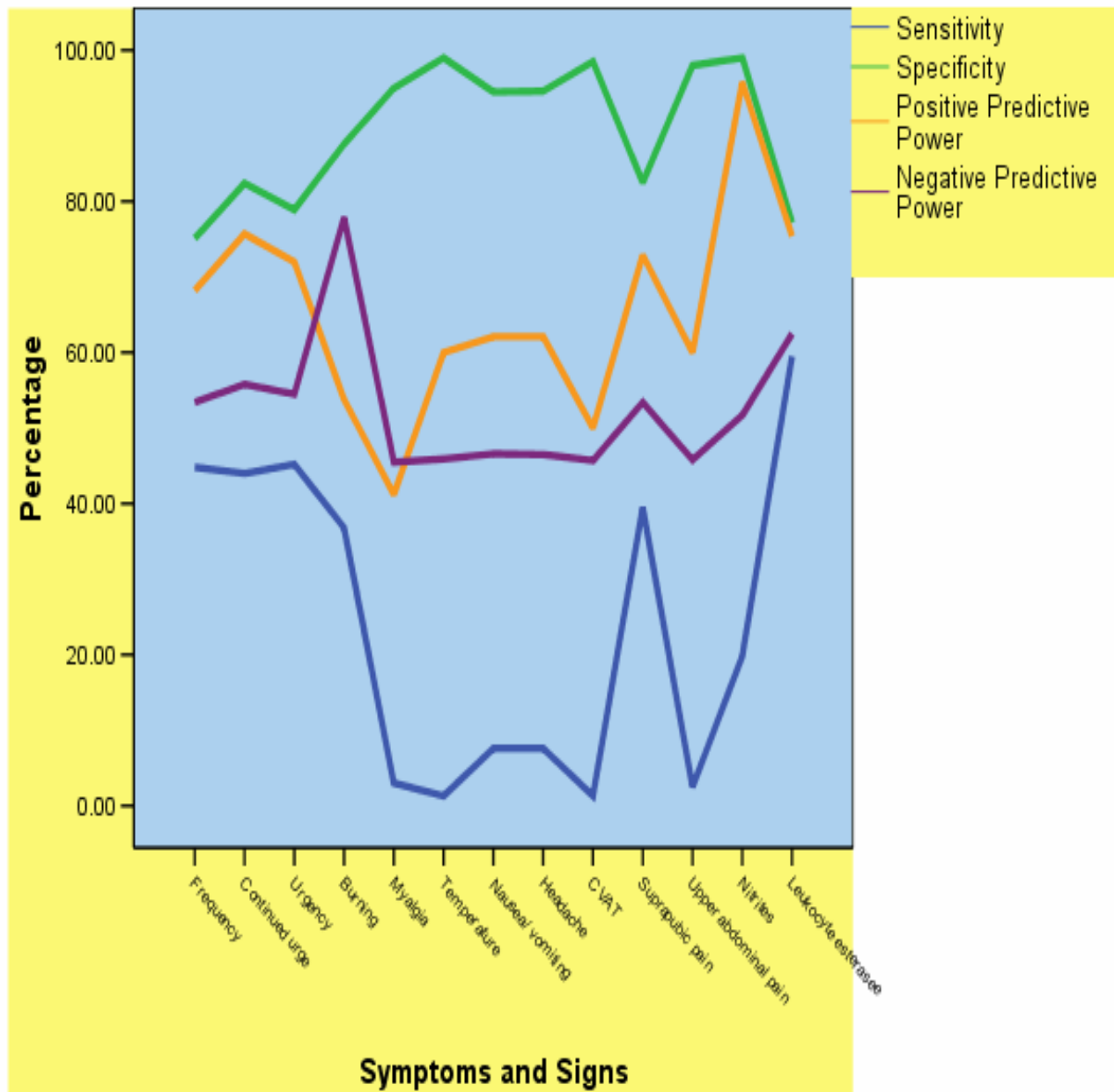
Results

The first research question investigated at which cfu level on the urine culture are the signs, symptoms, and dipstick results the most sensitive and specific. Symptoms examined included

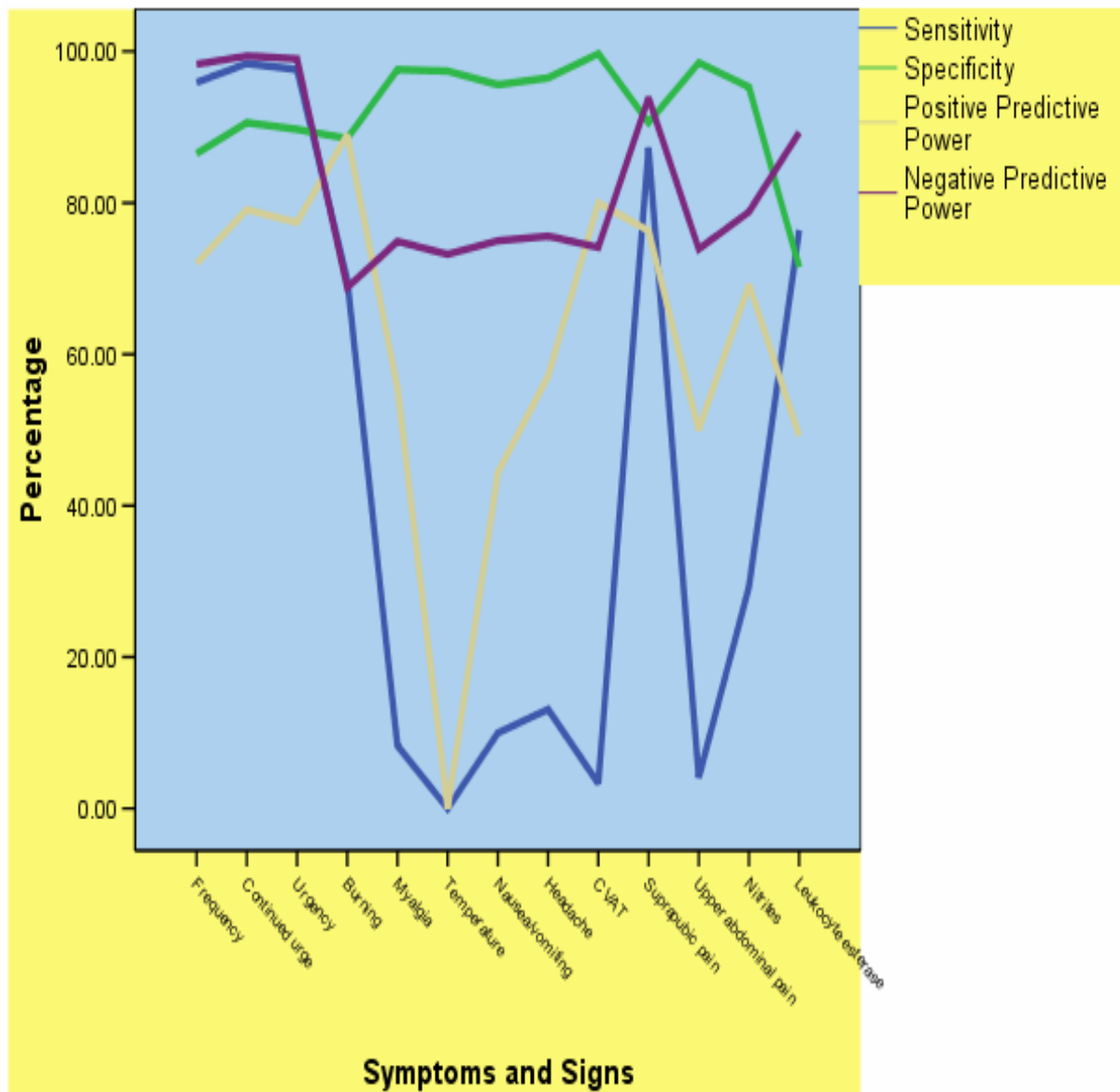
frequency, urgency, continued urge, burning, headache, nausea and vomiting, and myalgia. Signs include suprapubic pain, upper abdominal pain, CVAT, and temperature which were examined by the Nurse Practitioner. Nitrites and leukocytes were determined by dipstick.

First, the sensitivity, specificity, positive predictive value, and negative predictive value of the signs, symptoms, and dipstick results were examined at specific increments of cfu on the urine culture. CfU levels examined were ≤ 5000 cfu, $>5000 - \leq 10,000$ cfu, $>10,000 - \leq 50,000$ cfu, $>50,000 - \leq 100,000$ cfu and, $> 100,000$ cfu. No findings were significant at these intervals.

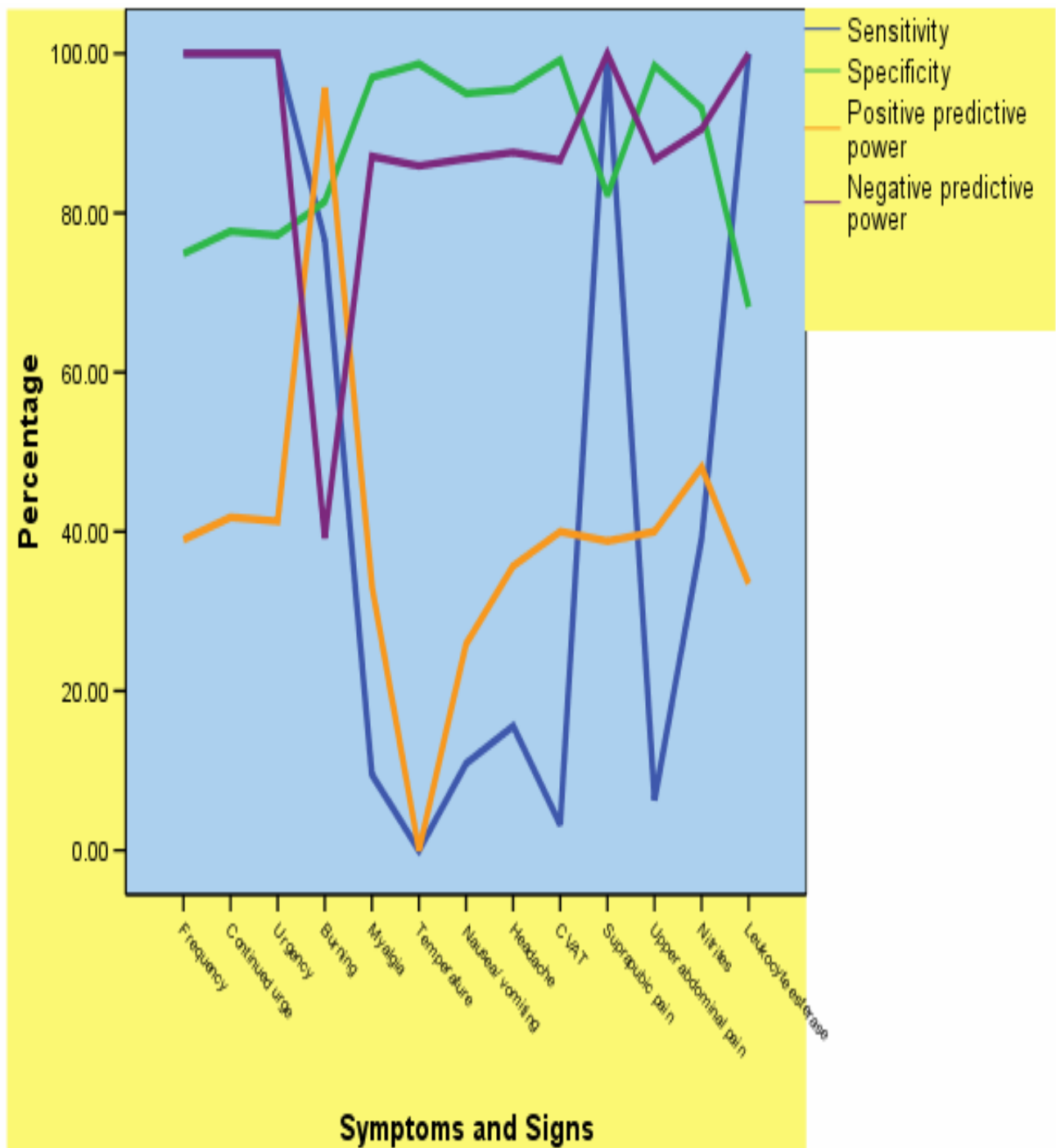
Data was then analyzed at the levels of $\geq 10,000$ cfu, $\geq 50,000$ cfu and $\geq 100,000$. At this cfu level, CVAT, myalgia, nausea and vomiting, and headache were the most specific with specificities greater than 94%. Leukocytes, urgency, and frequency were the most sensitive, however none yielded a sensitivity greater than 60%, making the sensitivity at this level inconclusive for diagnosis. In addition, no sign, symptom, or dipstick result provided a significant positive predictive value and negative predictive value. Figure 1 illustrates the sensitivity, specificity, positive predictive power, and negative predictive power of each sign, symptom, and dipstick result at $>10,000$ cfu on urine culture.

Figure 1 $\geq 10,000$ cfu on Urine Culture

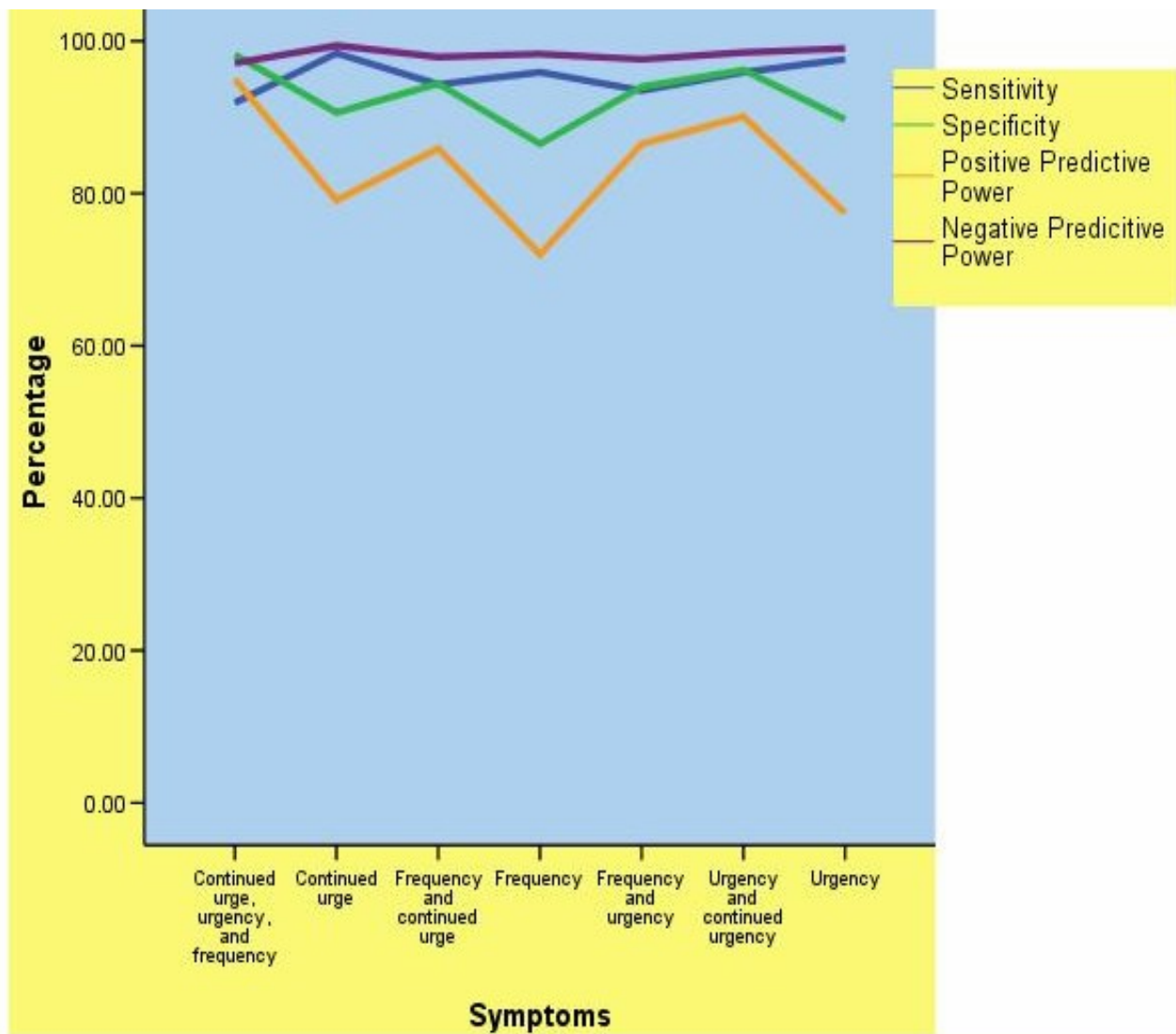
At $\geq 50,000$ cfu, frequency, urgency, and continued urge had sensitivities greater than 95%. In addition, these symptoms also yielded specificities greater than 90%. Although these symptoms were not as specific as other signs (i.e. temperature, CVAT, and upper abdominal pain), frequency, urgency, and continued urge yield the highest combination of both sensitivity and specificity. In addition, these symptoms produce the highest positive predictive values and negative predictive values. These symptoms are the best predictors of correctly identifying true cases of infection and dismissing negative cases of infection. Figure 2 illustrates the sensitivity, specificity, positive predictive value, and negative predictive value of each sign, symptom, and dipstick result at $>50,000$ cfu on urine culture.

Figure 2 $\geq 50,000$ cfu on Urine Culture

At $\geq 100,000$ cfu, the sensitivities of frequency, urgency, and continued urge increased to 100%, however specificity decreased. Positive predictive value and negative predictive value also decreased significantly at this level. Therefore, $\geq 50,000$ cfu on the urine culture yields the best sensitivity, specificity, positive predictive value, and negative predictive value for the symptoms of frequency, urgency, and continued urge. Figure 3 illustrates the sensitivity, specificity, positive predictive value, and negative predictive value of each sign, symptom, and dipstick result at $\geq 100,000$ cfu on urine culture.

Figure 3 $\geq 100,000$ cfu on Urine Culture

The next research question asked which combination of signs, symptoms, and dipstick result yields the best clinical result. Frequency, urgency, and continued urge were each examined individually, in combination, and all together. The combination of all three symptoms yields the highest sensitivity, specificity, positive predictive value, and negative predictive value. Figure 4 illustrates the sensitivity, specificity, positive predictive value, and negative predictive value of the individual symptoms and various combinations.

Figure 4 Combination of Symptoms at > 50,000 c.f.u. on Urine Culture

Chapter 5

Discussion

In conclusion, $\geq 50,000$ cfu on urine culture yields the best results in terms of sensitivity, specificity, positive predictive value, and negative predictive value for the symptoms of frequency, urgency, and continued urge to void. Some clinicians use the criteria of $\geq 100,000$ to diagnosis UTI, however this practice is not recommended. When using the symptom cluster at this cfu level, 55 women with all three symptoms would not have been treated (53.8% negative predictive value). Dipstick analysis of leukocytes and nitrites does not provide a clear answer for diagnosis. A symptom cluster of frequency, urgency, and continued urge to void is better than dipstick and physical examination. In this study of 468 women, diagnosis of UTI based on the symptom cluster at $\geq 50,000$ cfu on urine culture would yield six medication errors, i.e. medicated for UTI when UTI is not present, and ten omissions of medication when UTI is present.

These findings support continued use of telephone triage, diagnosis, and treatment of UTI because of the high sensitivity, specificity, positive predictive value, and negative predictive value. A urine culture should not be necessary to validate diagnosis if all symptoms are present. The error that would result from this practice would not have been avoided even if physical examination, urine dipstick, and urine culture were used in diagnostic assessment. Physical examination and dipstick analysis are more valuable to rule out UTI then to confirm diagnosis UTI.

The delay of 24- 48 hours to receive urine culture results to validate a diagnosis or begin treatment can be avoided. The cost of unnecessary diagnosis testing can be avoided by using the symptom cluster of frequency, urgency, and continued urge to void as a diagnostic tool.

A strength of this secondary analysis is that a specific protocol of physical assessment, history taking, and lab assessment was followed for each subject by a research Nurse

Practitioner. Therefore, very few missing data existed in this study. In addition, sample size was large, including 468 subjects. Future research should evaluate the relief of symptoms within 48 hours experienced by women who are treated or not treated on the basis of the presence or absence of the symptom cluster.

References

- Daniel, W. W. (1995). Biostatistics: a foundation for analysis in the health sciences, (6th ed.). New York: John Wiley & Sons.
- Deville, w., Yzermans, J., Duijn, N., Bezemer, P., Windt, D., & Bouter, L. (2004). The urine dipstick test useful to rule out infections. A meta-analysis of the accuracy, *BMC Urology*, 4:4.
- Flach, S., Longenecker, C., Tape, T., Bryan, T., Parenti, C., & Wigton, R. (2003). The relationship between treatment objectives and practice patterns. *Medical Decision Making*, 23, 131-139.
- Foxman, B. (2003). Epidemiology of urinary tract infections: Incidence, morbidity, economic costs. *Disease a Month*, 49, 53-70.
- Heuther, S.E., & McCance, K.L. (2000). *Understanding Pathophysiology*. St. Louis, MO: Mosby.
- Gupta, K., Hooton, T., Roberts, P., & Stamm, W., (2001). Pateint initiated treatment of uncomplicated recurrent urinary tract infections in young women. *Annals of Internal Medicine*, 135, 9-16.
- Managed Care (October, 2003). Study: treatment of UT infection just a phone call away. Retrieved May 17, 2005 from, http://www.managedcaremag.com/archives/0310/0310.news_ut_infection.html
- Medina-Bombardo, M., Sequi-Diaz, C., Fusalba, R., & Llober, J. (2003) What is the predictive value of urinary symptoms for diagnosis of UTI in women. *Family Practice*, 20, 103-107.
- Moore, K., Murrays,L., & Wagg, A. (2001). Rapid urinalysis assays for the diagnosis of UTI. *British Journal of Nursing*, 10, 998-101.
- Orenstein, R., & Wong, E. (1999). Urinary tract infections in adults. *American Academy of Family Physicians*, Retrieved Feburary 13, 2005 from <http://www.aafp.org/afp/990301ap/1225.html>.
- Othman, S., Chia, Y., & NG, C. (2003). Accuracy of urinalysis in the detection of urinary tract

- infection in a primary care setting. *Asia Pacific Family Medicine*, 2, 206-212.
- Sultana, R., Zalstein, S., Cameron, P., & Campbell, D. (2001). Dipstick urinalysis and the accuracy of clinical diagnosis of UTI. *Journal of Emergency Medicine*, 20, 13-19.
- Todar, K. (2005). Urinary tract infections. *Todars Online Textbook of Bacteriology*, Retrieved February 28, 2005 from <http://textbookofbacteriology.net/e.coli.html>.
- Ryan-Wenger, N. A., & Lowe, N. K. (2000). Military women's perspectives on gynecologic health care during deployment. *Women's Health Issues*, 10, 333-343.
- Saberi, P. (2004). Urinary Tract Infection. University of Maryland Medical Center. Retrieved April 5, 2005, from www.umm.edu/ency/article/000505.htm.
- U.S. Army (January, 2001). The Soldiers Almanac, Situation Report. Retrieved on February 1, 2001 from the worldwide web: www.dtic.mil/soldiers/jan2001.
- U.S. Navy (February, 2001). Navy Public Affairs Library. Retrieved on February 1, 2001 from the worldwide web: www.chinfo.navy.mil/navpalib/people/women.
- Valanis, B. (1999). *Epidemiology in Health Care*. Stamford, CT: Appleton & Lange