Published in final edited form as:

Infect Control Hosp Epidemiol. 2008 May ; 29(5): 446–449. doi:10.1086/586721.

Inter-observer Variability of Clinical Criteria in Nursing Home Residents with Suspected UTI

Manisha Juthani-Mehta, MD^{*}, Mary Tinetti, MD^{*}, Eleanor Perrelli, MSN, Virginia Towle, MPhil^{*}, Peter H. Van Ness, PhD, MPH^{*}, and Vincent Quagliarello, MD^{*}

^{*} Department of Internal Medicine, Yale University School of Medicine, New Haven, CT 06520

Abstract

We determined the inter-observer variability of clinical criteria for urinary tract infection (UTI) in nursing home residents. Pairs of nursing home staff caring for thirty residents were interviewed at times of suspected UTI. At least one measure from each clinical criteria category was reliably observed by nursing home staff members.

Keywords

urinary tract infection; nursing homes; reliability

Diagnosing urinary tract infection (UTI) in nursing home residents is problematic. Given the high incidence of asymptomatic bacteriuria and pyuria, a positive urine culture and pyuria on urinalysis are non-diagnostic.¹ Asymptomatic bacteriuria has not been associated with significant adverse outcomes;² therefore, practitioners must distinguish symptomatic UTI from asymptomatic bacteriuria in making therapeutic decisions. Practitioners utilize clinical criteria to differentiate symptomatic UTI from asymptomatic bacteriuria, but existing clinical criteria were developed by expert consensus.

The McGeer consensus criteria for UTI are widely accepted as surveillance and treatment standards.³ For residents without an indwelling catheter, three of the following criteria must be met to identify UTI: 1) fever $\geq 100.4^{\circ}\text{F}$; 2) new or increased burning on urination, frequency, or urgency; 3) new flank or suprapubic pain or tenderness; 4) change in character of urine; and 5) worsening of mental or functional status.⁴ The Loeb consensus criteria for UTI are minimum criteria necessary for empiric antibiotic therapy. For residents without an indwelling catheter, criteria include acute dysuria alone *or* fever ($>37.9^{\circ}\text{C}$ [100°F] or 1.5°C [2.4°F] increase above baseline temperature) *plus* at least one of the following: new or worsening urgency, frequency, suprapubic pain, gross hematuria, costovertebral angle tenderness, or urinary incontinence.⁵ The reliability, specifically inter-observer variability, for elements of these consensus criteria has not been determined.

Primary care givers in nursing homes are nursing home staff (i.e., registered nurses and aides). Nursing home practitioners (i.e., physicians, physician assistants, nurse practitioners) are called for clinical decision making when residents have a change in status, and often practitioners do

Corresponding author: Manisha Juthani-Mehta, M.D., Yale University School of Medicine, Department of Internal Medicine, Section of Infectious Disease, LMP 5040A, P.O. Box 208022, New Haven, CT 06520, Phone: (203) 737-1856; Fax: (203) 737-2531, Email: E-mail: manisha.juthanimehta@yale.edu.

Conflict of Interest: V. Q. has served as a consultant to Cubist Pharmaceuticals once in the past year for expertise unrelated to this study. No other authors receive financial support for consultantships or speakers forums. No authors have company holdings or patents.

not examine residents before clinical decisions are made. Therefore, nursing home practitioners rely on the observations of nursing home staff communicated by phone. The reliability of clinical criteria among different nursing home staff members is crucial to determining which criteria should be utilized for clinical decision making. In this study, we determined the inter-observer variability of clinical criteria at the time of suspected UTI.

METHODS

Setting and Participants

942 residents were screened from 5 New Haven area nursing homes. Residents were excluded if they: 1) were <65 years old; 2) had indwelling catheters; 3) were on chronic suppressive antibiotic or anti-infective therapy for recurrent UTI; 4) were short-term rehabilitation residents; 5) were terminal; 6) were on dialysis. Through medical record review, 700 residents were eligible. All eligible residents (or surrogates) were approached for verbal consent; 575 consented (82% participation rate). Twenty-one residents died or became ineligible at, or prior to, a baseline interview and 3 refused the baseline interview, resulting in 551 enrolled participants. Thirty participants with suspected UTI, defined as the nursing home resident's physician or nurse clinically suspecting a UTI, and for whom two nursing home staff members could be interviewed, participated in inter-observer reliability testing of clinical criteria. Interviews were conducted from February 2006-January 2007 in order to capture many different nursing home staff members.

Data Collection

Two nursing home staff members, usually one nurses' aide and one nurse, were interviewed regarding 30 participants at the time of suspected UTI by one study nurse. Two different nursing home staff members that cared for the resident when suspected of having UTI (e.g., 3–11pm shift) were interviewed within 3 days of suspected UTI. If two members from the same shift could not be identified, a second member from the next shift (e.g., 11–7am shift) was interviewed. All interviewees were asked the open ended question, "What triggered suspicion of UTI?" and up to three responses were recorded. Interview questions regarding mental status, behavior, and functional status were based upon the Minimum Data Set, a federally mandated questionnaire for nursing home residents.⁶ The clinical criteria assessed at the time of suspected UTI include 5 measures of behavior, 12 measures of functional status, 6 measures of mental status, change in urinary frequency, decreased urinary output, urinary retention, change in incontinence, increased diaper changes, urinary urgency, flank pain, abdominal pain, dysuria, gross hematuria, change in color of urine, change in odor of urine, new weakness, new fatigue, and new malaise.

Analysis

Simple kappa statistics are reported when the data were nominal. Weighted kappa statistics using Fleiss-Cohen quadratic weighting are reported when the data were ordinal; 90% confidence intervals (CI) for the kappa statistics were calculated. The PABAK (prevalence and bias adjusted kappa) is a descriptive reliability statistic for two-by-two agreement tables and is reported for binary data. The PABAK is calculated as two times the overall percentage of agreement, minus 1, and takes into account bias and prevalence factors which can result in a low kappa even when overall agreement is high.^{7, 8} It is a useful supplement to the kappa statistic when highly imbalanced marginal rating distributions threaten the accuracy of kappa results.^{9,10} Prevalence of the binary clinical criteria were calculated as the number of positive responses over the total number of responses (n=60). This study was approved by the Yale School of Medicine Human Investigation Committee.

RESULTS

Demographics of the 30 participants were as follows: mean age 87 years; 100% women; 83% white, 17% black; 7% Hispanic. Triggers causing suspicion of UTI are listed in Table 1. Clinical criteria were grouped into categories of behavioral symptoms, functional status, mental status, change in voiding pattern, urinary tract specific symptoms, change in character of urine, and non-specific symptoms. If the lower bound of the 90% CI for a kappa or weighted kappa estimate was ≥ 0.4 , then the clinical criteria was inferred to be moderately reliable. PABAK results greater than 0.4 suggested that the measure is moderately reliable (see Table 2).¹¹ Prevalence of the binary clinical criteria are listed in Table 2. Of the 30 participants with suspected UTI, 11 participants had definite UTI ($>100,000$ CFU on urine culture plus >10 WBC on urinalysis), 9 participants had possible UTI ($>100,000$ CFU and <10 WBC or urine culture between 10,000 and 100,000 CFU with any amount of pyuria), 7 had no UTI (urine culture with no growth or $<10,000$ CFU with any amount of pyuria), and 3 participants did not have a urine culture performed. Twenty of 30 participants were treated for UTI.

DISCUSSION

This study had two main findings. First, measures of behavior and mental status were less reliably observed by nursing home staff than measures of functional status. Only one of five measures of behavior (i.e., resists care) and three of six measures of mental status (i.e., periods of altered perception, disorganized speech, lethargy) were reliable. Eleven of twelve measures of functional status were reliable except for dressing. Second, change in voiding pattern and urinary tract specific symptoms were observed infrequently. Nevertheless, when adjusting for prevalence and bias factors, these clinical criteria categories appeared reliable according to PABAK results. Hence, these data suggest that at least one measure from each of the clinical criteria categories assessed for suspected UTI is reliably assessed by nursing home staff.

There is no evidence that non-specific symptoms (e.g. change in mental status) are indicative of UTI.¹⁰ Nevertheless, the McGeer criteria, which are the accepted standard for use in the nursing home setting,⁴ list “worsening of mental or functional status” as a criterion to identify UTI. Hence, it is important to determine the inter-observer reliability of both urinary tract specific and non-specific symptoms. Although not all measures of behavior or mental status were reliable, one measure of behavior and three measures of mental status were still reliably assessed. The three measures of mental status that were reliable (i.e., periods of altered perception, disorganized speech, and lethargy) are consistent with validated measures of delirium.¹¹ Urinary tract specific symptoms, specifically flank pain, appeared reliable, but they had a low prevalence in this subset of the population.

The main strength of this study is that nursing home staff members were interviewed for data collection. A limitation is that residents were not directly examined or interviewed. However, reliability assessments by nursing home staff were more clinically meaningful than assessments by a study nurse who did not know the participant’s baseline status.

Identifying symptomatic presentations of urinary infection in older institutionalized populations is an infectious disease research priority. Consensus based criteria are the current standards and may not be optimal for clinical decision making because of poor sensitivity and moderate predictive accuracy (50–60%) for bacteriuria plus pyuria.¹² However, without evidence-based criteria to guide clinical management, they have remained the basis for clinical practice. Based on this study, at least one measure from each of the clinical criteria categories appears to be moderately reliable. These findings suggest that at least one measure from each of these clinical criteria categories can be reliably measured and should be tested in future studies designed to identify clinical criteria associated with UTI. The next step will be to

determine whether reliable evidence-based criteria for the diagnosis of UTI in nursing home residents can be developed.

Acknowledgements

We are very grateful to the nursing home staff from the four nursing homes that participated in this study.

Grant Support: Dr. Juthani-Mehta was supported by a training grant from the NIA (T32-AG019134) and the ASP-Infectious Diseases Society of America (IDSA) Education and Research Foundation-National Foundation for Infectious Diseases (NFID) Young Investigator Award in Geriatrics. This study was supported, in part by the NIH/NIA Claude D. Pepper Older Americans Independence Center, Atlantic Philanthropies, IDSA/NFID, John A. Hartford Foundation, and ASP.

References

1. Nicolle LE. Urinary tract infection in long-term-care facility residents. *Clin Infect Dis* 2000;31(3): 757–61. [PubMed: 11017826]
2. Eberle CM, Winsemius D, Garibaldi RA. Risk factors and consequences of bacteriuria in non-catheterized nursing home residents. *J Gerontol* 1993;48(6):M266–71. [PubMed: 8227997]
3. Centers for Medicare and Medicaid (CMS) Manual System, State Operations Manual. Appendix. 2005. p. 183–4.
4. McGeer A, Campbell B, Emori TG, et al. Definitions of infection for surveillance in long-term care facilities. *Am J Infect Control* 1991;19(1):1–7. [PubMed: 1902352]
5. Loeb M, Bentley DW, Bradley S, et al. Development of minimum criteria for the initiation of antibiotics in residents of long-term-care facilities: results of a consensus conference. *Infect Control Hosp Epidemiol* 2001;22(2):120–4. [PubMed: 11232875]
6. Hawes C, Morris JN, Phillips CD, Mor V, Fries BE, Nonemaker S. Reliability estimates for the Minimum Data Set for nursing home resident assessment and care screening (MDS). *Gerontologist* 1995;35(2):172–8. [PubMed: 7750773]
7. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993;46(5):423–9. [PubMed: 8501467]
8. Kopp BJ, Erstad BL, Allen ME, Theodorou AA, Priestley G. Medication errors and adverse drug events in an intensive care unit: direct observation approach for detection. *Crit Care Med* 2006;34(2):415–25. [PubMed: 16424723]
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74. [PubMed: 843571]
10. Boscia JA, Kobasa WD, Abrutyn E, Levison ME, Kaplan AM, Kaye D. Lack of association between bacteriuria and symptoms in the elderly. *Am J Med* 1986;81(6):979–82. [PubMed: 3799658]
11. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med* 1990;113(12):941–8. [PubMed: 2240918]
12. Juthani-Mehta M, Tinetti M, Perrelli E, Towle V, Van Ness PH, Quagliarello V. Diagnostic accuracy of criteria for urinary tract infection in a cohort of nursing home residents. *J Am Geriatr Soc* 2007;55(7):1072–7. [PubMed: 17608881]

Table 1

Triggers for Suspecting UTI

Trigger	Number of episodes	Percent of total number of interviews(n=60)
Change in mental status	20	33.3
Fever or chills	17	28.3
Dysuria	9	15.0
Change in character of urine	9	15.0
Change in behavior	9	15.0
Other infection	7	11.7
Change in voiding pattern	4	6.7
Change in gait or new fall	3	5.0
Flank pain	2	3.3
Other work-up	2	3.3

Table 2
Inter-observer reliability testing at suspected UTI

Clinical Criteria (n=30)	Prevalence	Kappa [†] (90% CI)	Weighted Kappa [†] (90% CI)	PABAK [†]
Change in voiding pattern				
1. Change in urinary frequency	11.7%	0.53 (0.15–0.91)		0.80
2. Decreased urinary output	1.7%	**		0.94
3. Urinary retention	1.7%	**		0.94
4. Change in incontinence	0%	**		1.00
5. Increased diaper changes	3.3%	**		0.86
6. Urinary urgency	8.3%	0.78 (0.44–1.12)		0.94
Any change in voiding pattern (#1–6)	13.3%	0.44 (0.09–0.79)		0.74
Urinary Tract Specific Symptoms				
1. Flank pain	3.3%	1.00 (1.00–1.00)		1.00
2. Abdominal pain	1.7%	**		0.94
3. Dysuria	16.7%	0.76 (0.49–1.03)		0.86
Any urinary tract specific symptom (#1–3)	20%	0.79 (0.56–1.03)		0.86
Change in Character of Urine				
1. Gross hematuria	1.7%	**		0.94
2. Change in color of urine	8.3%	0.35 (–0.13–0.83)		0.80
3. Change in odor of urine	25%	0.56 (0.28–0.84)		0.66
Any change in character of urine (#1–3)	26.6%	0.50 (0.22–0.78)		0.60
Non-specific Symptoms				
1. New weakness	16.7%	0.52 (0.17–0.87)		0.74
2. New fatigue	25%	0.73 (0.49–0.97)		0.80
3. New malaise	26.7%	0.66 (0.41–0.91)		0.74
Any non-specific symptom (#1–3)	48.3%	0.67 (0.45–0.89)		1.00
Mental Status Symptoms				
1. Periods of altered perception		0.61 (0.39–0.84)		
2. Disorganized speech		0.77 (0.57–0.96)		
3. Lethargy		0.65 (0.44–0.85)		
4. Easily distracted		0.41 (0.16–0.66)		
5. Restlessness		0.55 (0.31–0.80)		
6. Mental function varies during the day		0.31 (0.08–0.54)		
Behavioral Symptoms				
1. Resists care			0.62 (0.40–0.85)	
2. Wandering			0.45 (–0.13–1.03)	
3. Verbally abusive			0.35 (–0.04–0.73)	
4. Physically abusive			0.30 (–0.14–0.74)	
5. Socially inappropriate			0.52 (0.06–0.97)	
Functional Status				

Clinical Criteria (n=30)	Prevalence	Kappa [†] (90% CI)	Weighted Kappa [†] (90% CI)	PABAK [†]
1. Bed mobility			0.73 (0.55–0.92)	
2. Transfer			0.93 (0.87–0.98)	
3. Walk in room			0.75 (0.51–1.00)	
4. Walk in corridor			0.72 (0.49–0.94)	
5. Locomotion on unit			0.79 (0.61–0.98)	
6. Locomotion off unit			0.74 (0.42–1.07)	
7. Eating			0.82 (0.64–1.00)	
8. Toilet use			0.79 (0.62–0.95)	
9. Personal hygiene			0.76 (0.52–1.01)	
10. Bathing			0.80 (0.49–1.11)	
11. Bladder continence			0.65 (0.44–0.86)	
12. Dressing			0.66 (0.34–0.98)	

** Insufficient distribution for kappa, 100% 'No' response at either time 1 or time 2.

[†] Both kappa statistics and PABAK have been presented for change in voiding pattern, urinary tract specific symptoms, change in character of urine, and non-specific symptoms (binary variables with yes/no responses). For variables with more than 2 levels, a simple kappa is presented when the levels are not ordered, such as mental status (3 levels). A weighted kappa is presented when there are more than 2 levels and the levels indicate increasing disability, such as behavioral symptoms (4 levels) and functional status (6 levels). PABAK is not presented for variables having more than 2 levels.