Objectives

- How to critically appraise a paper reporting the results of a diagnostic test.
- Inspired to explore and/or teach EBM
- *‘interactive’ and ‘hands on’ and ‘fun’*

Clinical scenario

- Mr PF: 61 years, retired accountant from Cairns.
- Left foot pain 2 months
  - mild warmth over mid-foot, no antecedent illnesses, small ulcer 1st metatarsal head - probe to bone negative, systemically well, ESR 63.
- Type 2 diabetes for 11 years, retinopathy, peripheral neuropathy, on insulin
- X-ray: destructive process involving midfoot bones and joints
- He “wants to avoid surgery if possible!”

The diagnosis of osteomyelitis of the foot in diabetes: microbiological examination vs. magnetic resonance imaging and labelled leucocyte scanning

Diabetic Medicine 2006;23:649-653

- Background
  - Foot infections are a frequent and severe complication of diabetes often leading to amputations. Clinical assessment is unreliable. The rate of unnecessary biopsies might be reduced with non-invasive screening tests. The comparative diagnostic value of microbiology, MRI, and labelled leucocyte scanning is not established.

Critical Appraisal of Studies

Question: diagnosis

1. Is the evidence about the accuracy of the diagnostic test valid?

Measurement: was there a blinded comparison with a reference ("gold") standard of diagnosis?
Representative: was the diagnostic test evaluated in an appropriate spectrum of patients?
Ascertainment: was the reference standard applied regardless of the diagnostic test result?

2. Using this valid evidence, does the test accurately distinguish patients who do and do not have a specific disorder?

Sensitivity, **Specificity**, and **Likelihood ratios**.

3. How can I apply this valid, accurate diagnostic test to a specific patient?

Is the test **available, affordable**, **accurate**, and **precise** in our setting? Can we generate a clinically sensible estimate of our patient's **pre-test probability**? Will the resulting post-test probabilities affect our management and help our patient?

1. Is the evidence about the accuracy of the diagnostic test valid?

Measurement: was there a blinded comparison with a reference ("gold") standard of diagnosis?

Ascertainment: was the reference standard applied regardless of the diagnostic test result?

****Read the **first paragraph of Methods** (p 650), discuss with partner, then answer the 2 questions above****

1. Is the evidence about the accuracy of the diagnostic test valid?

- Was there a blinded comparison with a reference ("gold") standard of diagnosis?
  - Reference standard - YES - bone specimens.
  - Blind – not stated; potentially important given 'subjective' nature of histopathology (degree of leucocyte infiltration).
  - Histopathology as reference standard may be questioned – sampling errors, expertise of pathologist.

1. Is the evidence about the accuracy of the diagnostic test valid?

- Was the reference standard applied regardless of the diagnostic test result?
  - YES, all patients presented had bone specimens taken.

1. Is the evidence about the accuracy of the diagnostic test valid?

Representative: was the diagnostic test evaluated in an appropriate spectrum of patients (like those in whom it would be used in practice)?

****Read the **first 3 paragraphs of Results** (p 651), discuss with partner, then answer the questions above****

1. Is the evidence about the accuracy of the diagnostic test valid?

Was the diagnostic test evaluated in an appropriate spectrum of patients (like those in whom it would be used in practice)?

- Turkish hospital clinic patients
- n=31 - small sample
- Wagner classification 3 (36%), 4 (48%), 5 (16%) – severe disease means higher pre-test probability of disease than usual patient group
- MRSA 24%, pseudomonas 33% - hospital based patients

Wagner’s classification for diabetic foot disease

Grade-0	High risk foot and no ulceration.
Grade -1	Superficial Ulcer.
Grade -2	Deep Ulcer (cellulitis)
Grade -3	Osteomyelitis with ulceration or abscess.
Grade -4	Gangrenous Patches. Partial foot gangrene.
Grade -5	Gangrene of entire foot.

2. Using this valid evidence, does the test accurately distinguish patients who do and do not have a specific disorder?

		Target disorder (osteomyelitis)		Totals
		Present	Absent	
Diagnostic test result	Positive	a	b	a+b
	Negative	c	d	c+d
Totals		a+c	b+d	a+b+c+d

$$\text{Sens.} = \frac{a}{a+c}$$

$$\text{Spec.} = \frac{d}{b+d}$$

$$\text{LR+} = \frac{\text{Sens.}}{1-\text{Spec.}}$$

$$\text{LR-} = \frac{1-\text{Sens.}}{\text{Spec.}}$$

****Read Table 2 (p 651), discuss with partner, then using the '2x2' table calculate the sensitivity, specificity, and likelihood ratios (LR+ and LR-) for each of labelled leucocyte scan, MRI, and microbiology****

2. Using this valid evidence, does the test accurately distinguish patients who do and do not have a specific disorder?

		Target disorder (osteomyelitis)		Totals
		Present	Absent	
Labelled leucocyte scan	Positive	21	1	22
	Negative	2	2	4
Totals		23	3	26

$$\text{Sens.} \quad \frac{21}{23} = 0.91$$

$$\text{Spec.} \quad \frac{2}{3} = 0.67$$

$$\text{LR+} = \frac{\text{Sens.}}{\text{1-Spec.}} = 2.7$$

$$\text{LR-} = \frac{\text{1-Sens.}}{\text{Spec.}} = 0.13$$

2. Using this valid evidence, does the test accurately distinguish patients who do and do not have a specific disorder?

		Target disorder (osteomyelitis)		Totals
		Present	Absent	
MRI	Positive	18	2	20
	Negative	5	3	8
Totals		23	5	28

$$\text{Sens.} \quad \frac{18}{23} = 0.78$$

$$\text{Spec.} \quad \frac{3}{5} = 0.60$$

$$\text{LR+} = \frac{\text{Sens.}}{\text{1-Spec.}} = 2.0$$

$$\text{LR-} = \frac{\text{1-Sens.}}{\text{Spec.}} = 0.53$$

2. Using this valid evidence, does the test accurately distinguish patients who do and do not have a specific disorder?

		Target disorder (osteomyelitis)		Totals
		Present	Absent	
Micro-biology	Positive	24	2	26
	Negative	2	3	5
Totals		26	5	31

$$\text{Sens.} \quad \frac{24}{26} = 0.92$$

$$\text{Spec.} \quad \frac{3}{5} = 0.60$$

$$\text{LR+} = \frac{\text{Sens.}}{\text{1-Spec.}} = 2.3$$

$$\text{LR-} = \frac{\text{1-Sens.}}{\text{Spec.}} = 0.13$$

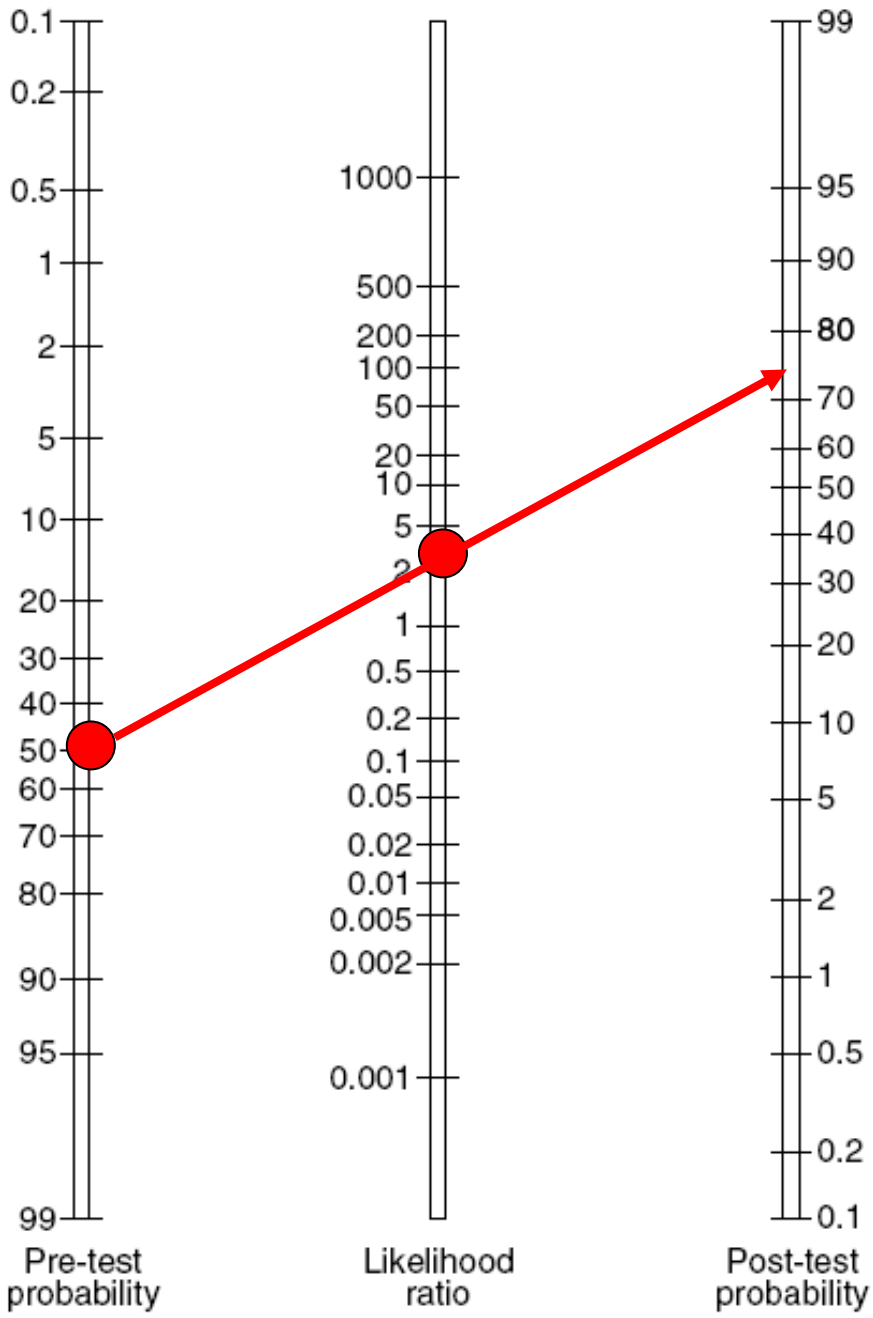
Why Likelihood Ratios?

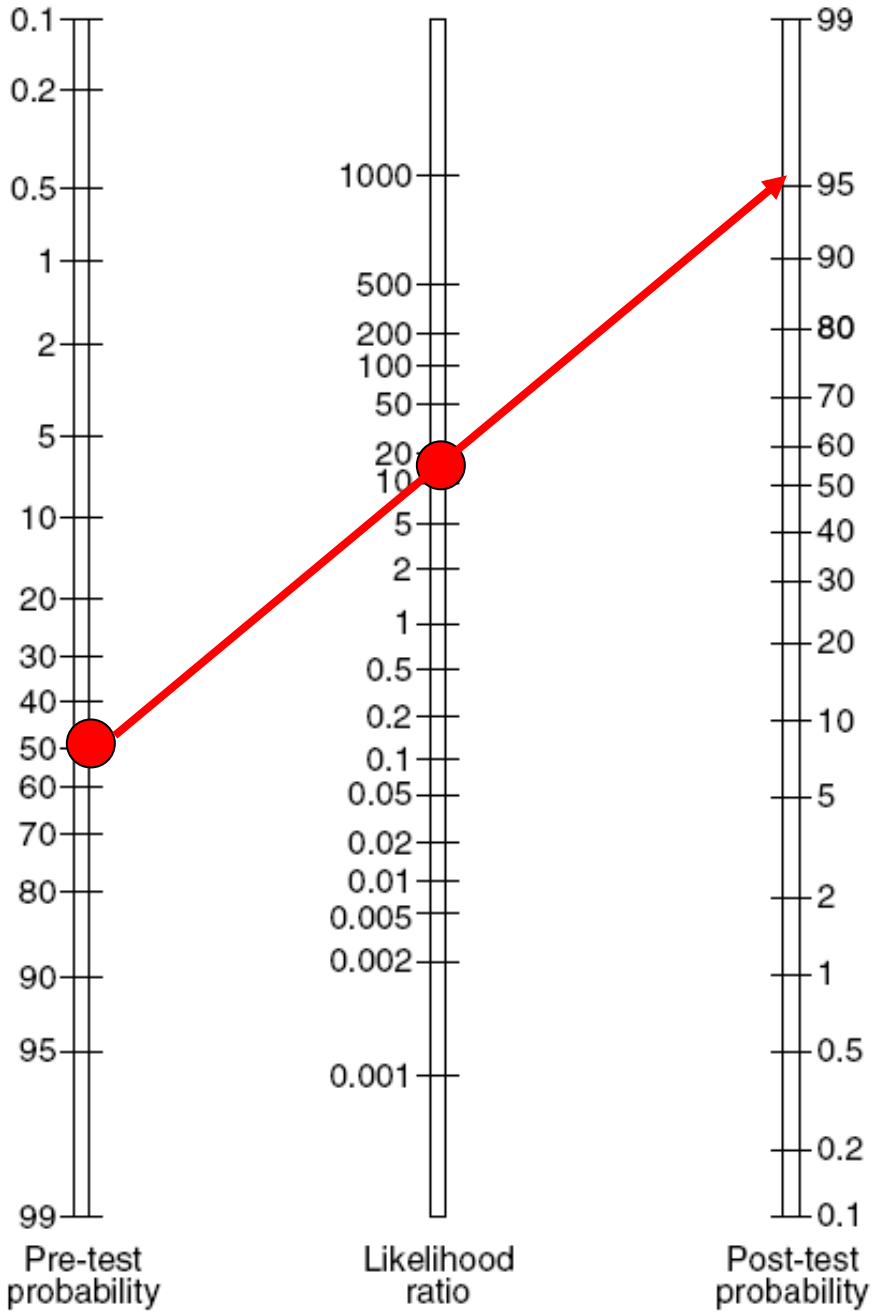
LR+

- LR > 10
  - usually conclusive change
- LR 5 –10
  - moderate shift
- LR 2 – 5
  - small, sometimes important
- LR 1 – 2
  - small, rarely important shift

LR-

- LR < 0.1
  - usually conclusive change
- LR 0.1-0.2
  - moderate shift
- LR 0.2-0.5
  - small, sometimes important
- LR 0.5-1
  - small, rarely important shift


0.1
0.2
0.5
1
2
5
10
20
30
40
50
60
70
80
90
95
99
Pre-test probability
1000
500
200
100
50
20
10
5
2
1
0.5
0.2
0.1
0.05
0.02
0.01
0.005
0.002
0.001
Likelihood ratio
99
95
90
80
70
60
50
40
30
20
10
5
2
1
0.5
0.2
0.1
Post-test probability


0.1
0.2
0.5
1
2
5
10
20
30
40
50
60
70
80
90
95
99
Pre-test probability
1000
500
200
100
50
20
10
5
2
1
0.5
0.2
0.1
0.05
0.02
0.01
0.005
0.002
0.001
Likelihood ratio
99
95
90
80
70
60
50
40
30
20
10
5
2
1
0.5
0.2
0.1
Post-test probability

Which tests are useful in making a positive diagnosis?

		+LR
A.	RF for RA	3
B.	CCP for RA	14
C.	Uric acid for gout	1.1
D.	ANA for SLE	1.5
E.	cANCA+PR3 for Wegeners	60
F.	Ro+La for Sjogrens	4
G.	ACE for sarcoid	1.5

3. How can I apply this valid, accurate diagnostic test to a specific patient?

- Is the test **available, affordable**, **accurate**, and **precise** in our setting?
- Can we generate a clinically sensible estimate of our patient's **pre-test probability**?
- Will the resulting post-test probabilities affect our management and help our patient?

3. How can I apply this valid, accurate diagnostic test to a specific patient?

In our setting is the test?:

- **available**
  - generally regional and metropolitan settings
- **affordable**
  - for patient (patchy medicare funding)
  - for health system – discriminating use
- **accurate & precise**
  - likely to be radiologist dependent, no published data on this

3. How can I apply this valid, accurate diagnostic test to a specific patient?

Can we generate a clinically sensible estimate of our patient's **pre-test probability**? maybe…

- Sources:
  - Clinical experience….memory distortions (recency, emotional response, limited experience)
  - Prevalance statistics (useful, if accurate diagnostic data)
  - Local databases (useful, if they exist)
  - From this research study (was it representative?)
  - From research study specifically designed to measure pre-test probabilities (none published)

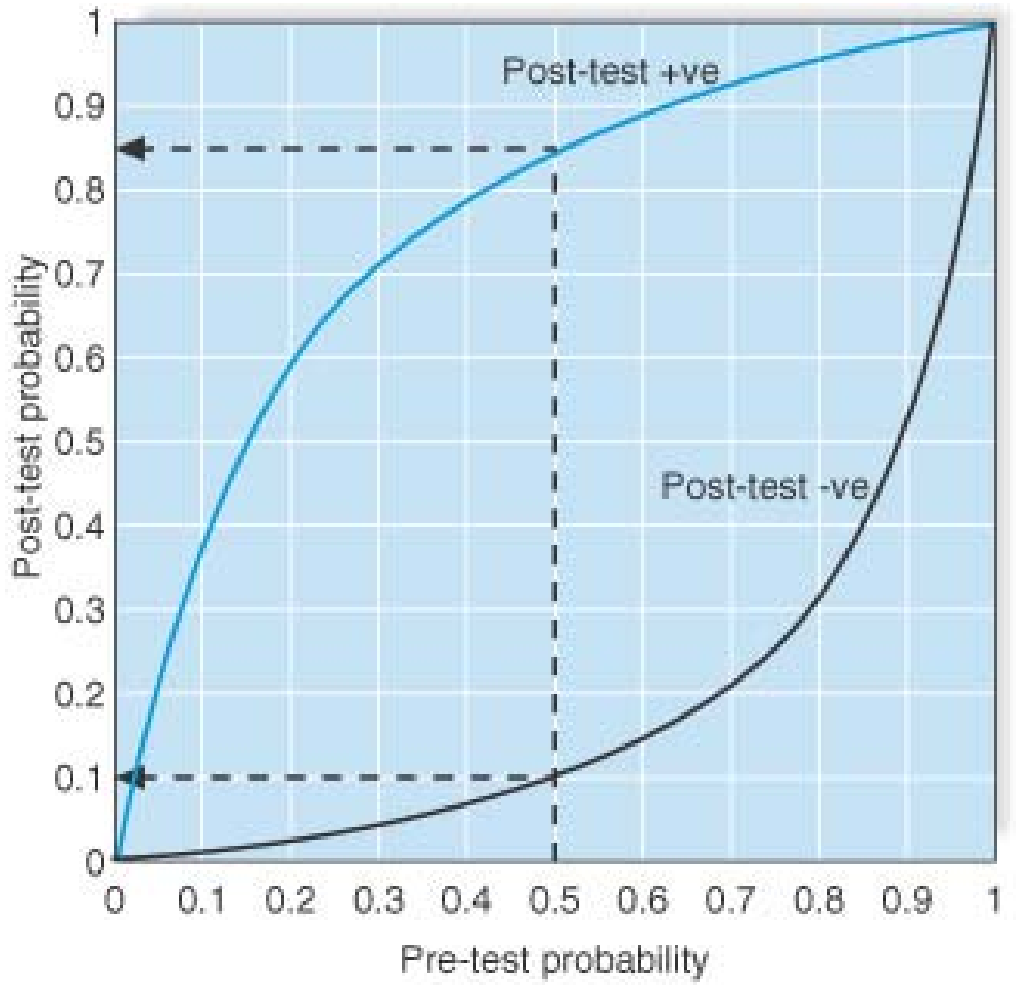
3. How can I apply this valid, accurate diagnostic test to a specific patient?

- Will the resulting **post-test probabilities** affect our management and help our patient?
  - **Test-treatment threshold** likely to be crossed?
  - Is our patient informed and willing participant?

Treat

Test

Wait



3. How can I apply this valid, accurate diagnostic test to a specific patient?

- Will the resulting **post-test probabilities** affect our management and help our patient?
  - **Test-treatment threshold** likely to be crossed?
  - Is our patient informed and willing participant?

Magnetic Resonance Imaging for Diagnosing Foot Osteomyelitis

A Meta-analysis

Arch Intern Med. 2007;167:125-132

Alok Kapoor, MD; Stephanie Page, MD; Michael LaValley, PhD; Daniel R. Gale, MD; David T. Felson, MD, MPH

Table 3. Diagnostic Performance of 4 Technologies in Studies That Compared MRI With Another Imaging Test

	MRI		Technetium Tc 99m Bone Scan		Plain Radiography		WBC Scan	
Source	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
Croll et al,[15] 1996	88.9	100.0	50.0	50.0	22.2	94.4	33.3	69.2
Enderle et al,[16] 1999	100.0	75.0	83.3	75.0	71.4	80.0		
Horowitz et al,[18] 1993	100.0	100.0			43.8	100.0		
Kearney et al,[19] 1999	100.0	50.0	88.9	100.0	66.7	100.0		
Levine et al,[21] 1994	76.9	100.0	100.0	25.0	60.0	81.3	80.0	28.6
Lipman et al,[22] 1998	77.3	40.0			73.3	40.0		
Nigro et al,[25] 1992	100.0	95.2	90.5	33.3	69.6	33.3		
Remedios et al,[26] 1998	100.0	81.8	100.0	0	38.5	100.0	90.9	84.6
Yuh et al,[30] 1989	100.0	89.5	94.4	18.2	75.0	60.0		

Objectives

- How to critically appraise a paper reporting the results of a diagnostic test.
- Inspired to explore and/or teach EBM
- *‘interactive’ and ‘hands on’ and ‘fun’*

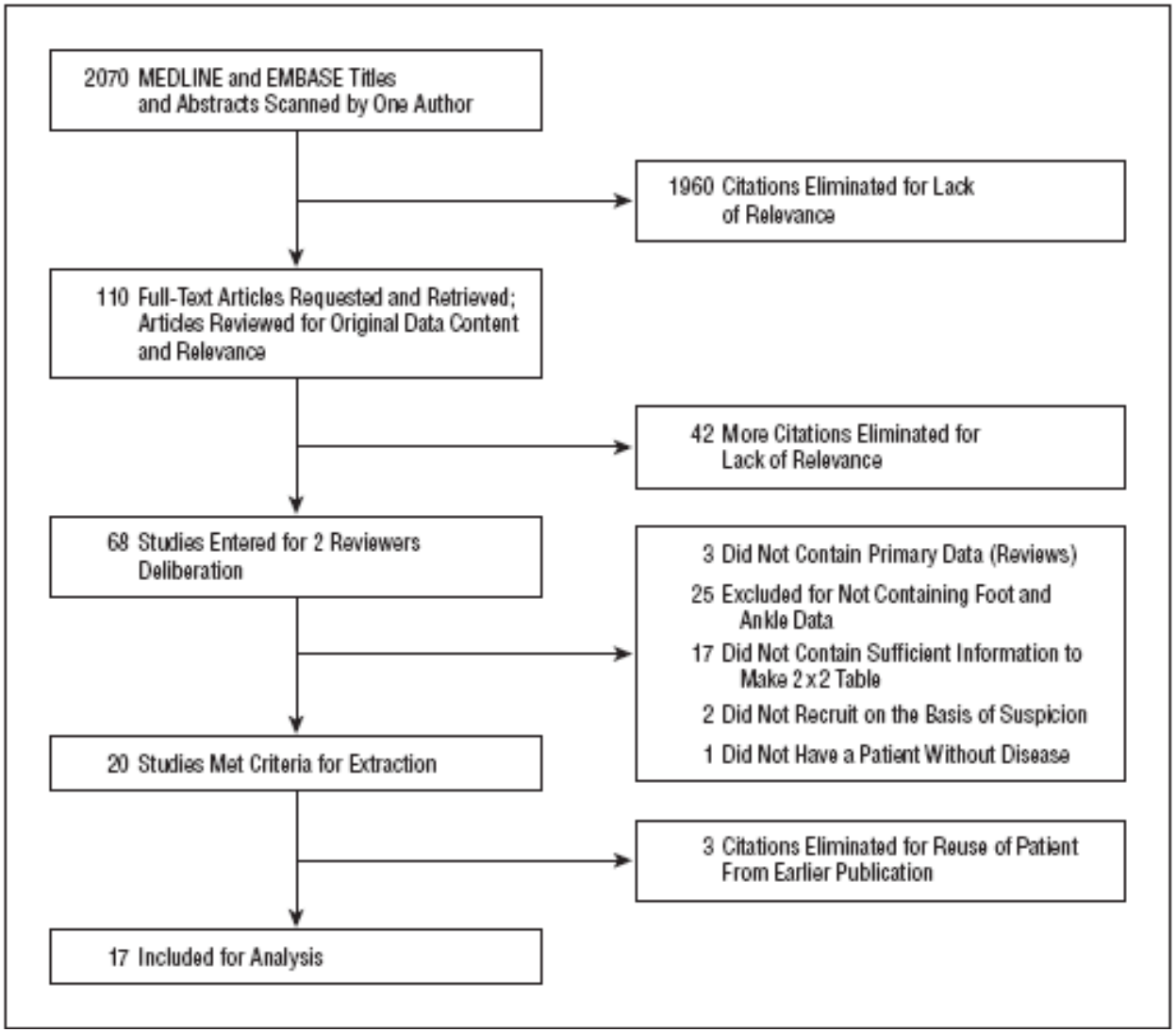

2070 MEDLINE and EMBASE Titles and Abstracts Scanned by One Author
1960 Citations Eliminated for Lack of Relevance
110 Full-Text Articles Requested and Retrieved; Articles Reviewed for Original Data Content and Relevance
42 More Citations Eliminated for Lack of Relevance
68 Studies Entered for 2 Reviewers Deliberation
3 Did Not Contain Primary Data (Reviews)
25 Excluded for Not Containing Foot and Ankle Data
17 Did Not Contain Sufficient Information to Make 2 x 2 Table
2 Did Not Recruit on the Basis of Suspicion
1 Did Not Have a Patient Without Disease
20 Studies Met Criteria for Extraction
3 Citations Eliminated for Reuse of Patient From Earlier Publication
17 Included for Analysis

Table 1. Characteristics of Included Studies

Source	Enrollment Criteria	No. of Sites in the Study (No. of Patients)*	Mean Age of Patients, y	Prevalence of Diabetes, %	Prospective Design	Consecutive Enrollment	MRI Assessors
Craig et al,[14] 1997	Diabetic patients scheduled for partial amputation	57 (13)	57.0	100.0	Yes	NS	NS
Croll et al,[15] 1996	Patients admitted with nongangrenous diabetic foot infections	27 (27)	66.0	100.0	Yes	NS	No
Enderle et al,[16] 1999	Diabetic patients suspected of having chronic osteomyelitis from random surgery and medicine clinics	19 (19)	60.7	100.0	Yes	Yes	Yes
Ertugrul et al,[17] 2006	Diabetic patients with ulcers at Wagner grade ≥3‡	31 (31)	62.0	100.0	Yes	NS	NS
Horowitz et al,[18] 1993	Patients admitted with diabetic foot infections	47 (41)	54.4	100.0	Yes	NS	No
Kearney et al,[19] 1999	Diabetic outpatients suspected of having osteomyelitis	13 (13)	59.0	100.0	Yes	NS	NS
Ledermann et al,[20] 2002	Diabetic and nondiabetic patients suspected of osteomyelitis	84 (72)	NS	NS	NS	NS	Yes
Levine et al,[21] 1994	Diabetic patients with suspected osteomyelitis complicating soft tissue infection	29 (27)	51.6	100.0	No	NS	No
Lipman et al,[22] 1998	Patients with peripheral neuropathy and high clinical suspicion	20 (20)	46.0	85.0	Yes	Yes	Yes
Maas et al,[23] 2002	Patients with neuropathy and inflammation in addition to leprosy	18 (12)	63.0	0	No	Yes	Yes
Morrison et al,[24] 1998	Patients suspected of having osteomyelitis of the foot	73 (62)	56.0	84.9	No	NS	Yes
Nigro et al,[25] 1992	Patients with foot inflammation and possible osteomyelitis	47 (44)	55.0	70.5	NS	NS	NS
Remedios et al,[26] 1998	Diabetic patients with peripheral neuropathy, chronic foot ulcers, and signs of osteomyelitis	9 (9)	57.0	100.0	Yes	NS	NS
Seabold et al,[27] 1990	Patients highly suspected of having osteomyelitis in and around Charcot joint	12 (11)	50.6	91.7	No	No	No
Vesco et al,[28] 1999	Diabetic patients with foot ulcers	24 (24)	59.0	100.0	Yes	Yes	NS
Weinstein et al,[29] 1993	Diabetic patients admitted with suggestion of osteomyelitis, nonhealing ulcer, or soft tissue infection	75 (47)	49.6	100.0	Yes	Yes	Yes
Yuh et al,[30] 1989	Patients suspected of having osteomyelitis or nonhealing ulcer	44 (24)	58.2	100.0	NS	Yes	Yes

Table 2. MRI Diagnostic Criteria, Frequency of Biopsy Use, and MRI Performance

Source	Signs on MRI Used to Determine Positive Result*	Biopsy Reference Standard, %		Prevalence of Osteomyelitis, %	MRI Performance	
		Positive†	Negative‡		Sensitivity, %	Specificity, %
Craig et al,[14] 1997	Primary, half of patients getting T1 with gadolinium, and soft tissue mass or ulcer	100.0	100.0	36.8	90.5	65.0
Croll et al,[15] 1996	NS	100.0	55.6	33.3	88.9	100.0
Enderle et al,[16] 1999	Increased uptake on STIR with ulcer or soft tissue mass and T1 with gadolinium§	100.0	100.0	73.7	100.0	75.0
Ertugrul et al,[17] 2006	Decreased T1, turboinversion recovery magnitude, and T1 with gadolinium	100.0	100.0	74.2	78.2	60.0
Horowitz et al,[18] 1993	Increased TR/TE or increased T2 and some gadolinium, with or without cortical disruption	100.0	NS	31.9	100.0	100.0
Kearney et al,[19] 1999	NS	NS	0	69.2	100.0	50.0
Ledermann et al,[20] 2002	Primary, T1 with gadolinium, and many secondary signs	100.0	100.0	63.1	90.6	83.9
Levine et al,[21] 1994	Standard primary	100.0	31.3	44.8	76.9	100.0
Lipman et al,[22] 1998	Primary and many secondary	100.0	20.0	75.0	77.3	40.0
Maas et al,[23] 2002	Standard primary, T1 with gadolinium, and many secondary‖	50.0	0	88.9	100.0	50.0
Morrison et al,[24] 1998	Standard primary, T1 with gadolinium for most patients, and many secondary signs‖	100.0	0	58.9	91.3	83.2
Nigro et al,[25] 1992	NS	92.3	NS	55.3	100.0	95.2
Remedios et al,[26] 1998	Standard primary	100.0	20.0	44.4	100.0	80.0
Seabold et al,[27] 1990	Standard primary	100.0	0	36.4	100.0	0¶
Vesco et al,[28] 1999	Standard primary, T1 with gadolinium, and ulcer, sinus tract, or soft tissue mass	0	0	54.2	100.0	81.8
Weinstein et al,[29] 1993	Standard primary	100.0	55.2	61.3	100.0	79.3
Yuh et al,[30] 1989	Standard primary	100.0	21.1	61.4	100.0	89.5

Table 3. Diagnostic Performance of 4 Technologies in Studies That Compared MRI With Another Imaging Test

	MRI		Technetium Tc 99m Bone Scan		Plain Radiography		WBC Scan	
Source	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %	Sensitivity, %	Specificity, %
Croll et al,[15] 1996	88.9	100.0	50.0	50.0	22.2	94.4	33.3	69.2
Enderle et al,[16] 1999	100.0	75.0	83.3	75.0	71.4	80.0		
Horowitz et al,[18] 1993	100.0	100.0			43.8	100.0		
Kearney et al,[19] 1999	100.0	50.0	88.9	100.0	66.7	100.0		
Levine et al,[21] 1994	76.9	100.0	100.0	25.0	60.0	81.3	80.0	28.6
Lipman et al,[22] 1998	77.3	40.0			73.3	40.0		
Nigro et al,[25] 1992	100.0	95.2	90.5	33.3	69.6	33.3		
Remedios et al,[26] 1998	100.0	81.8	100.0	0	38.5	100.0	90.9	84.6
Yuh et al,[30] 1989	100.0	89.5	94.4	18.2	75.0	60.0		

Table 4. Posttest Probability Stratified by Imaging Test Result Across a Spectrum of Pretest Probabilities

Pretest Probability, %	Posttest Probability, %			
	MRI Positive*	MRI Negative	Technetium Tc 99m Positive†	Technetium Tc 99m Negative
10.0	36.2	1.3	11.2	3.4
25.0	63.0	3.8	27.4	9.6
50.0	83.6	10.7	53.1	24.2
75.0	93.9	26.5	77.2	49.0
90.0	97.9	51.9	91.1	74.2