

ORIGINAL ARTICLES

Diagnostic tests for bacterial infection from birth to 90 days—a systematic review

P W Fowlie, B Schmidt

Abstract

Aim—To determine the clinical value of common diagnostic tests for bacterial infection in early life.

Methods—A Medline search (1966–95) was undertaken to identify studies that reported the assessment of a diagnostic “test,” predicting the presence or absence of bacterial infection in infants up to 90 days of age. The quality of each selected study was assessed using defined criteria. Data were extracted twice to minimise errors.

Results—Six hundred and seventy articles were identified. Two independent investigators agreed that 194 studies met the inclusion criteria ($\kappa = 0.85$), 52 of which met primary quality criteria; 23 studies reported data on (a) haematological indices, (b) C reactive protein evaluation, and (c) surface swab assessment. For haematological indices, the likelihood ratios for individual tests ranged from 20.4 (95% confidence interval 7.3 to 56.8) for a white cell count $< 7000/\text{mm}^3$ to 0.12 (0.04 to 0.37) for an immature:total (I:T) white cell ratio < 0.2 . For C reactive protein evaluation, the likelihood ratios ranged from 12.56 (0.79 to 199.10) for a value of $> 6 \text{ mg/l}$ to 0.22 (0.08 to 0.65) for a negative value. For surface swab assessment, the likelihood ratios ranged from 33.6 (2.1 to 519.8) for a positive gastric aspirate culture to 0.08 (0.006 to 1.12) for microscopy of ear swab material that did not show any neutrophils. Likelihood ratios for combinations of these individual tests ranged from 10.17 (3.64 to 28.41) to 0.47 (0.22 to 1.00).

Conclusions—The methodological quality of studies assessing the accuracy of diagnostic tests is generally poor. Even in rigorous studies, the reported accuracy of the tests varies enormously and they are of limited value in the diagnosis of infection in this population.

(Arch Dis Child Fetal Neonatal Ed 1998;78:F92-F98)

Keywords: sepsis; bacterial infection; diagnostic tests

these infants unclear.²⁻⁴ To be of practical use, any diagnostic test must fulfil certain criteria: it should accurately predict the presence or absence of infection and be reliable; it should be simple to perform; results should be available quickly; and it should be cost effective. If a test is not sufficiently accurate, then regardless of its other attributes it will be of limited value in clinical practice.

The Evidence Based Medicine Working Group suggest four criteria that will improve the validity of any results from a study assessing the accuracy of a diagnostic test: (1) there should be an independent, blind comparison with a reference standard; (2) the patient sample should include an appropriate spectrum of patients to whom the test will be applied in clinical practice; (3) the results of the test being evaluated should not influence the decision to perform the reference standard; and (4) the test should be described in sufficient detail to permit replication.^{5,6} In addition, it is suggested that in order to be clinically useful, the data should be presented in such a manner that likelihood ratios can be calculated.⁷

We carried out a systematic review to determine the methodological quality of clinical research into the accuracy of diagnostic tests for bacterial infection in the first three months of life, and to review the results from the studies most likely to provide valid data.

Methods

An extensive search for articles (English only) on the diagnosis of infection in the newborn period and during infancy published between 1966 and April 1995 was made in Medline at the National Library of Medicine, Bethesda, USA, using a strategy (available from the authors) designed to be highly sensitive at identifying articles on diagnosis.⁸ We identified 670 citations. Letters, editorials, commentaries, and reviews were excluded, leaving 572 articles and abstracts for possible inclusion.

To be included in the review, articles had to report the assessment of a diagnostic “test” (including signs and symptoms) predicting the presence or absence of bacterial infection, and also include extractable data relating to infants up to 90 days of age. Tests for chlamydial infection were included, although chlamydia is not strictly a bacterium. Diagnostic tests for viral infection were excluded as were tests using

Department of Child Health, University of Dundee, Scotland
P W Fowlie

Department of Paediatrics, McMaster University, Hamilton, Ontario, Canada
B Schmidt

Correspondence to:
Dr P W Fowlie, Department of Child Health, Ninewells Hospital, Dundee DD1 9SY.

Accepted 20 October 1997

Infection in early life is a major cause of mortality and morbidity.¹ The symptoms and signs are often non-specific, making diagnosis difficult and the optimal strategy for managing

Table 1 Characteristics of studies included in the review

Study	Year	Reference	Population	Criterion standard	"Test"
Scanlon	1971	50	"Newborn infants"	Positive blood culture	External ear canal swab (M and C)
Scanlon	1972	49	"Newborn infants"	Positive blood culture	External ear canal swab (M and C); gastric aspirate (M and C)
Boyle	1978	12	Infants admitted with respiratory distress (25–40 weeks gestation)	Positive blood culture	Total WCC; neutrophil count; gastric aspirate
Squire	1979	55	Infants who died within 72 h of admission (term and preterm)	Extensive inflammation at necropsy plus positive culture of blood, CSF, lung, or other body fluid	Total WCC; neutrophil count; platelet count
Voorra	1982	58	Term newborn infants admitted to NICU because of fever	Positive blood culture	Total WCC
El-Radhi	1983	22	Infants <22 days	Positive culture of blood, urine, or CSF, or radiological diagnosis of pneumonia	Gastric aspirate (M and C)
Adhikari	1986	9	LBW infants in NICU <1 week old	Positive blood culture	CRP
King	1987	33	Infants <8 weeks old admitted with fever	Positive blood culture or CSF	Total WCC, I:T ratio
Leibovich	1987	36	Newborn infants	Positive blood culture	Gastric aspirate (M)
Evans	1988	23	Infants admitted to NICU (mean gestation 35 weeks)	Positive blood or CSF culture or other body fluid "as appropriate"	Ear canal swabs; axilla swab; ET aspirate; gastric aspirate; NP swab; rectal swab; skin swab; umbilical swab (all C)
Koenig	1988	34	Infants <3 months referred from community	Positive blood or urine culture	WCC >15 000/mm ³ plus I:T ratio >0.2
Misra	1989	38	Infants <7 days	Positive blood culture	Total WCC; I:T ratio; platelet count
Rigal	1990	45	Intubated newborn infants (28–44 weeks gestation)	Open lung biopsy: positive culture and pathological diagnosis	ET aspirate ©
Seibert	1990	52	Infants ≤31 weeks	Positive blood culture	CRP; total WCC; I:T ratio; combination
Bonadio	1991	11	Infants ≤8 weeks old	Positive blood, CSF, urine culture	Neutrophil count
Thomson	1992	56	Infants <33 weeks gestation admitted to NICU	Positive blood or CSF culture	Gastric aspirate (M and C); ear swab (C); nasal swab (C); umbilical swab (C)
Peakman	1992	39	Infants admitted to NICU (25–40 weeks gestation)	Positive blood/CSF culture; heavy growth in urine or >1 surface swab	Total WCC; neutrophil count; platelet count; CRP
Hiew	1992	31	Infants ≤3 months old	Positive bacterial culture (blood, CSF, urine, pustule, sputum, umbilicus)	Total WCC; neutrophil count; I:T ratio; platelet count; CRP; combination
Krediet	1992	35	Infants admitted to NICU (24–44 weeks gestation)	Positive blood or CSF culture or pneumonia diagnosed on CXR and positive ET aspirate culture	CRP; I:T ratio; combination
Pourcyrous	1993	42	Infants admitted to NICU (23–44 weeks gestation)	Positive blood or CSF culture	CRP
Rodwell	1993	47	Neutropenic newborn infants (25 weeks to term)	Positive blood/CSF culture; necropsy evidence of infection	Haematology score†
Edgar	1994	21	Infants admitted to NICU (23–39 weeks gestation)	Decision based on clinical course, blood culture results, and haematology indices	CRP
Philip	1994	40	Infants admitted to NICU	Positive blood or CSF culture or pneumonia diagnosed on CXR and positive ET aspirate culture	CRP; I:T ratio; combination
Wagle	1994	59	Infants <30 weeks	Positive culture of blood, CSF, or urine	CRP
DaSilva	1994	18	Neonates after day 4 admitted to NICU (mean gestation 28 weeks)	Positive blood or CSF culture	Total WCC; I:T ratio; band cell count; combination

†Based on total white cell count, immature:total white cell ratio, total neutrophil count, immature:mature neutrophil ratio, regenerative changes in neutrophils, and platelet count.⁵¹ C = culture; CRP = C reactive protein; CSF = cerebrospinal fluid; CXR = chest x ray; ET = endotracheal; I:T ratio = immature:total white cell ratio; LBW = low birth weight; M = microscopy; NICU = neonatal intensive care unit; NP = nasopharyngeal; WCC = white cell count.

amniotic fluid or cord blood. The abstract, and if necessary, the complete manuscript of each of the articles identified by the search were assessed independently by the authors. One hundred and ninety four articles met the inclusion criteria. Agreement as to which articles to include or exclude was excellent (Cohen's κ statistic = 0.85), and any disagreement was resolved by discussion.

The quality (validity) of each article included in the review was assessed independently by the two authors. Three primary criteria were used: (1) Was there an independent blind comparison with a reference (gold) standard? For the purposes of the review, an acceptable diagnostic gold standard was regarded as a diagnosis of infection based on pure growth of an organism from blood, cerebrospinal fluid, urine, or deep tissue culture, or chest x ray changes supported by bacteriological growth from endotracheal tube aspirate; (2) Did the population studied include an appropriate spectrum of babies to whom the test would be applied in clinical practice? (3) Were the results reported in such a manner that they could be expressed as likelihood ratios? Two other secondary criteria were also assessed: was the test described in

sufficient detail to allow duplication, and was there any reference to the reliability of the test?

Data from the studies retained after this process were extracted twice by PWF. Where possible, 2×2 tables, or equivalent, were created. For tests with a positive or negative result only, the accuracy of each "test" was then reported as sensitivity and specificity, positive and negative predictive values, and the likelihood ratios associated with positive and negative results. In the case of multilevel tests, the results were reported as likelihood ratios associated with each level of test result. Confidence intervals are limited to the likelihood ratio in order to avoid reporting too many data. The precise definitions of these measures and how to use the likelihood ratio are described in the appendix.

Results

Of the 194 studies accepted for inclusion, the authors agreed that 73 (38%) reported an independent, blind comparison with an acceptable reference standard (agreement between authors, $\kappa = 0.30$), 148 (76%) studied a population that included an appropriate spectrum of babies to whom the test could be applied in clinical practice ($\kappa = 0.82$), and 58 (30%)

Table 2 Accuracy of haematological variables

Study	Year	Reference	Positive test definition	Incidence of sepsis	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Positive LR (95% CI)	Negative LR (95% CI)
Boyle	1978	12	Absolute neutrophil count $\leq 5000/\text{mm}^3$	9/114	89	81	29	99	4.67 (2.96 to 7.37)	0.14 (0.02 to 0.87)
Boyle	1978	12	WCC $<7000/\text{mm}^3$	9/114	78	96	64	98	20.42 (7.34 to 56.77)	0.23 (0.07 to 0.79)
Squire	1979	55	Abnormal band count (as referenced)	23/46	13	83	43	49	0.75 (0.19 to 2.98)	1.05 (0.82 to 1.35)
Squire	1979	55	Abnormal neutrophil count (as referenced)	23/46	61	83	78	68	3.50 (1.36 to 9.04)	0.47 (0.28 to 0.82)
Squire	1979	55	Platelet count $<150\ 000/\text{mm}^3$	23/46	61	78	74	67	2.80 (1.21 to 6.50)	0.50 (0.29 to 0.87)
Squire	1979	55	WCC <5000 or $>24\ 000/\text{mm}^3$	23/46	44	96	91	63	10.00 (1.39 to 71.90)	0.59 (0.41 to 0.86)
Voorra	1982	58	WCC $<5000/\text{mm}^3$	9/59	22	98	67	88	11.11 (1.12 to 110.72)	0.79 (0.56 to 1.13)
King	1987	33	I:T ratio ≥ 0.2	16/321	69	75	13	98	2.76 (1.88 to 4.05)	0.42 (0.20 to 0.86)
King	1987	33	WCC $\leq 5000/\text{mm}^3$	16/340	44	96	37	97	11.81 (5.39 to 25.91)	0.58 (0.38 to 0.90)
Misra	1989	38	Band cell count	33/78					<10 , LR = 0.36 (0.13 to 1.00) $10-20$, LR = 0.89 (0.52 to 1.51) >20 , LR = 2.18 (1.14 to 4.18)	
Misra	1989	38	I:T ratio ≥ 2	33/78	90	73	71	92	3.41 (2.08 to 5.60)	0.12 (0.04 to 0.37)
Misra	1989	38	Platelet count (per mm^3)	33/78					<1 , LR = 4.08 (0.17 to 96.95) $1-1.5$, LR = 1.48 (0.78 to 2.81) >1.5 , LR = 0.79 (0.56 to 1.10)	
Misra	1989	38	WCC (per mm^3)	33/78					<5000 , LR = 2.96 (1.25 to 6.96) $5-15\ 000$, LR = 0.64 (0.45 to 0.92) $>15\ 000$, LR = 1.36 (0.29 to 6.34)	
Seibert	1990	52	WCC <5000 or $>20\ 000/\text{mm}^3$ or I:T ratio ≥ 0.2 (early infection)	Not given†	63	68	11	96	2.0†	0.54†
Seibert	1990	52	WCC <5000 or $>20\ 000/\text{mm}^3$ or I:T ratio ≥ 0.2 (late infection)	Not given†	33	61	26	80	0.9†	1.10†
Bonadio	1991	11	Neutropenia $<1500/\text{mm}^3$	80/1000	1	93	1	92	0.17 (0.02 to 1.18)	1.07 (1.04 to 1.10)
Krediet	1992	35	I:T ratio ≥ 0.2 (early infection)	14/49	86	51	41	90	1.77 (1.18 to 2.64)	0.28 (0.07 to 1.04)
Krediet	1992	35	I:T ratio ≥ 0.2 (late infection)	25/48	68	65	68	65	1.96 (1.05 to 3.64)	0.49 (0.26 to 0.94)
Peakman	1992	39	Neutrophil count <2000 or $>7500/\text{mm}^3$	17/42	29	64	36	57	0.82 (0.33 to 2.02)	1.10 (0.72 to 1.69)
Peakman	1992	39	Platelet count $<150\ 000/\text{mm}^3$	17/42	6	84	20	57	0.39 (0.05 to 3.01)	1.12 (0.91 to 1.38)
Peakman	1992	39	WCC <5000 or $>20\ 000/\text{mm}^3$	17/42	18	76	33	58	0.74 (0.21 to 2.55)	1.08 (0.79 to 1.48)
Hiew	1992	31	I:T ratio >0.2	30/70	13	84	27	69	0.85 (0.29 to 2.45)	1.03 (0.87 to 1.22)
Hiew	1992	31	Neutrophil count ≤ 1000 or $\geq 5400-14\ 400/\text{mm}^3$ (age dependent)	30/70	53	44	29	69	0.96 (0.65 to 1.42)	1.05 (0.66 to 1.68)
Hiew	1992	31	Platelet count $<150\ 000/\text{mm}^3$	30/70	3	94	20	70	0.58 (0.07 to 5.00)	1.03 (0.94 to 1.12)
Hiew	1992	31	WCC <5000 or $>21\ 000, 30\ 000, 150\ 000/\text{mm}^3$ (age dependent)	30/70	27	81	38	72	1.44 (0.67 to 3.10)	0.90 (0.71 to 1.15)
DaSilva	1994	18	Band cell count $>1000/\text{mm}^3$	28/147	50	91	56	89	5.41 (2.76 to 10.61)	0.55 (0.38 to 0.80)
DaSilva	1994	18	I:T ratio ≥ 0.2	28/147	18	96	50	83	4.25 (1.32 to 13.68)	0.86 (0.72 to 1.02)
DaSilva	1994	18	WCC $\geq 20\ 000$ or $\leq 5000/\text{mm}^3$	28/147	36	80	29	84	1.77 (0.96 to 3.26)	0.81 (0.60 to 1.08)
Philip	1994	40	I:T ratio >0.2	32/311	53	82	25	94	2.91 (1.93 to 4.38)	0.57 (0.40 to 0.83)

†The data presented in this report did not allow 2×2 tables to be created. It is not therefore possible to calculate the incidence of infection or confidence intervals around the likelihood ratios. CI = confidence interval; I:T ratio = immature/total white cell count; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value; WCC = white cell count.

Table 3 Accuracy of C reactive protein assay as a diagnostic test for bacterial infection

Author	Year	Reference	Positive test	Incidence of infection	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Positive LR (95% CI)	Negative LR (95% CI)
Adhikari	1986	9	"Positive" v "negative"	19/43	84	71	70	85	2.89 (1.50 to 5.55)	0.22 (0.08 to 0.65)
Seibert	1990	52	>10 mg/l (early infection)	Not given†	63	70	13	96	2.1†	0.53†
Seibert	1990	52	>10 mg/l (late infection)	Not given†	57	61	30	82	1.5†	0.70†
Krediet	1992	35	>7 mg/l (early infection)	14/49	43	71	38	76	1.50 (0.67 to 3.34)	0.80 (0.49 to 1.32)
Krediet	1992	35	>7 mg/l (late infection)	25/48	84	48	64	73	1.61 (1.04 to 2.47)	0.34 (0.12 to 0.90)
Peakman	1992	39	"Positive" v "negative"	17/42	29	64	36	89	0.82 (0.33 to 2.02)	1.11 (0.72 to 1.69)
Hiew	1992	31	>10 mg/l	30/70	83	41	38	85	1.42 (1.10 to 1.83)	0.40 (0.17 to 0.94)
Pourcyrous	1993	42	>9 mg/l (day 1)	242/689	57	88	72	79	4.69 (3.57 to 6.16)	0.49 (0.43 to 0.57)
Pourcyrous	1993	42	>9 mg/l (day 1, 2, or 3)	242/689	77	76	63	85	3.17 (2.66 to 3.78)	0.30 (0.24 to 0.38)
Edgar	1994	21	>6 mg/l	43/60	35	100	100	38	12.56 (0.79 to 199.10)	0.67 (0.53 to 0.85)
Philip	1994	40	>10 mg/l	32/311	53	92	43	95	6.44 (3.87 to 10.72)	0.51 (0.35 to 0.74)
Wagle	1994	59	>10 mg/l (day 1)	51/309	63	87	49	92	4.91 (3.35 to 7.19)	0.43 (0.30 to 0.61)
Wagle	1994	59	>10 mg/l (day 1 and/or day 2)	51/309	90	81	48	98	4.65 (3.57 to 6.07)	0.12 (0.53 to 0.28)

†The data presented in this report did not allow 2×2 tables to be created. It is not therefore possible to calculate the incidence of infection or confidence intervals around the likelihood ratios. CI = confidence interval; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value.

reported results that could be expressed as a likelihood ratio ($\kappa = 0.56$). Initial agreement by the two authors on the secondary criteria suggested that in only 93 studies (51%) was the actual test described in sufficient detail such that it could be repeated; and in only six studies (3%) was there any mention of the reliability of the tests.

In order to minimise the inclusion of studies reporting potentially biased results, it was decided to examine further only those articles that met all three of the primary "quality" criteria. Initially it was agreed that 57 articles

appeared to do this; however, during data extraction it became apparent that a further five of these articles did not, in fact, meet all three methodological criteria. After discussion it was decided to drop these five articles, leaving 52 papers for further assessment.⁹⁻⁶⁰

We assessed 155 individual tests in these 52 papers. We present the results of the review of individual tests based on (1) haematological indices, (2) C reactive protein evaluation, and (3) surface swab assessment. These particular data are taken from 23 studies^{9 11 12 18 21-23 31 33 35 36 38-40 42 45 49 50 52 55 56 58 59}

Table 4 Accuracy of surface swabbing (including gastric aspiration) for diagnosing bacterial infection

Study	Year	Reference	Test definition	Incidence of infection	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Positive LR (95% CI)	Negative LR
Scanlon	1972	49	Ear swab culture	8/67	100	40	18	100	1.56 (1.19 to 2.03)	0.15 (0.01 to 2.23)
Scanlon	1972	49	Ear swab microscopy: ≥ 1 PMN	8/67	88	95	70	98	17.21 (5.54 to 53.44)	0.13 (0.02 to 0.83)
Scanlon	1972	49	Gastric aspirate culture	8/67	88	61	23	97	2.25 (1.49 to 3.39)	0.21 (0.03 to 1.30)
Scanlon	1972	49	Gastric aspirate microscopy: >5 PMN/high power field	8/67	88	85	44	98	5.74 (2.98 to 11.05)	0.15 (0.02 to 0.93)
Scanlon	1971	50	Ear swab culture	6/28	100	47	43	100	1.79 (1.05 to 3.04)	0.16 (0.01 to 2.40)
Scanlon	1971	50	Ear swab microscopy: any PMN	6/21	100	93	86	100	14.32 (2.12 to 96.49)	0.08 (0.006 to 1.12)
Scanlon	1971	50	Ear swab microscopy: any organisms seen	6/21	83	67	50	91	2.50 (1.12 to 5.57)	0.25 (0.04 to 1.55)
El-Radhi	1983	22	Gastric aspirate culture (4–28 days old)	44/67	71	100	100	64	33.65 (2.13 to 519.75)	0.30 (0.19 to 0.47)
El-Radhi	1983	22	Gastric aspirate culture (<4 days old)	2/21	50	90	33	94	4.75 (0.71 to 32.00)	0.56 (0.14 to 2.25)
El-Radhi	1983	22	Gastric aspirate microscopy: >3 PMN/HPF (4–28 days)	44/67	82	87	92	71	6.27 (2.16 to 18.19)	0.21 (0.11 to 0.40)
El-Radhi	1983	22	Gastric aspirate microscopy: >3 PMN/HPF (<4 days old)	2/21	100	90	50	100	7.80 (1.83 to 33.32)	0.22 (0.02 to 2.67)
Boyle	1978	12	Gastric aspirate microscopy: any bacteria	7/90	71	78	28	97	3.29 (1.77 to 6.13)	0.37 (0.11 to 1.18)
Boyle	1978	12	Gastric aspirate smear: any PMN	7/90	71	72	18	97	2.58 (1.44 to 4.62)	0.40 (0.12 to 1.29)
Rigal	1990	45	Tracheal aspirate culture	2/11	50	22	13	67	0.64 (0.15 to 2.68)	2.25 (0.36 to 14.28)
Evans†	1988	23	Axilla swab culture	Not given†	53	86	9	3.7†	0.55†	
Evans†	1988	23	Ear canal swab culture	Not given†	50	87	9	3.7†	0.58†	
Evans†	1988	23	ET tube aspirate and culture	Not given†	59	74	7	2.3†	0.56†	
Evans†	1988	23	Gastric aspirate culture	Not given†	48	77	5	2.1†	0.68†	
Evans†	1988	23	Nasopharyngeal swab culture	Not given†	59	87	12	4.6†	0.47†	
Evans†	1988	23	Rectal swab culture	Not given†	69	31	5	1.0†	1.01†	
Evans†	1988	23	Skin swab culture	Not given†	0	84	0	0.0†	1.19†	
Evans†	1988	23	Umbilical swab culture	Not given†	80	77	4	3.5†	0.26†	
Leibovich	1987	36	Gastric aspirate microscopy: ≥ 5 WC/HPF or any bacteria	8/140	75	68	13	98	2.36 (1.47 to 3.78)	0.37 (0.11 to 1.23)
Thomson	1992	56	Ear swab culture	9/134	78	90	35	98	7.48 (4.02 to 13.93)	0.2 (0.07 to 0.84)
Thomson	1992	56	Gastric aspirate culture	9/127	67	85	25	97	4.37 (2.33 to 8.19)	0.39 (0.15 to 0.99)
Thomson	1992	56	Gastric aspirate microscopy (pus cells)	9/127	89	49	12	98	1.75 (1.31 to 2.34)	0.2 (0.04 to 1.45)
Thomson	1992	56	Gastric aspirate microscopy (pus cells and organisms)	9/127	89	80	24	99	4.37 (2.86 to 6.69)	0.14 (0.02 to 0.89)
Thomson	1992	56	Nasal swab culture	9/130	56	95	46	97	11.20 (4.23 to 26.69)	0.4 (0.23 to 0.97)
Thomson	1992	56	Umbilical swab culture	9/136	78	92	41	98	9.89 (4.96 to 19.69)	0.2 (0.07 to 0.82)

†The sensitivities, specificities, and positive predictive values relating to these tests are given in the paper and it is therefore possible to calculate the likelihood ratios. However, insufficient data are provided to calculate the negative predictive value or the confidence limits around any of the likelihood ratios. CI = confidence interval; HPF = high power field; LR = likelihood ratio; NPV = negative predictive value; PMN = polymorphonuclear leucocyte; PPV = positive predictive value; WC = white cell

Table 5 Accuracy of combinations of tests

Study	Year	Reference	Positive test definition	Incidence of sepsis	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Positive LR (95% CI)	Negative LR (95% CI)
Koenig	1988	34	WCC >15 000/mm ³ and I:T ratio >0.2	10/100	60	86	32	95	4.15 (2.04 to 8.48)	0.47 (0.22 to 1.00)
Seibert	1990	52	CRP or abnormal haematology (day 7 or less)*	Not given†	75	50	9	95	1.5	0.50 (0.15 to 1.70)
Seibert	1990	52	CRP or abnormal haematology (day 8–60)*	Not given†	73	45	28	85	1.33	0.57 (0.28 to 1.19)
Krediet	1992	35	CRP + I:T ratio (day 8–60)*	25/48	60	74	71	63	2.30 (1.08 to 4.91)	0.54 (0.32 to 0.93)
Hiew	1992	31	CRP + neutrophil count*	30/100	47	64	36	74	1.31 (0.80 to 2.14)	0.83 (0.57 to 1.21)
Hiew	1992	31	CRP + WCC*	30/100	27	89	50	74	2.33 (0.97 to 5.64)	0.83 (0.66 to 1.04)
Hiew	1992	31	WCC + neutrophil count*	30/100	23	83	37	72	1.36 (0.60 to 3.12)	0.93 (0.74 to 1.16)
Rodwell	1993	47	Haematology score (as in table 1)	28/170	N/A	N/A	N/A	N/A	Score ≥ 5 : LR = 15.21 (1.64 to 141.02) Score = 4: LR = 6.20 (2.84 to 13.55) Score = 3: LR = 3.80 (2.03 to 7.14) Score = 2: LR = 0.32 (0.08 to 1.25) Score <2 : LR = 0.04 (0.00 to 0.67)	
DaSilva	1994	18	Band cell count or WBC or I:T ratio*	28/147	61	76	37	89	2.49 (1.61 to 3.85)	0.52 (0.32 to 0.83)
Philip	1994	40	I:T ratio + CRP*	32/311	22	98	54	92	10.17 (3.64 to 28.41)	0.80 (0.66 to 0.96)

*See tables 2 and 3 for individual CRP/haematology cut off values indicating a positive result. †The data presented in this report did not allow 2×2 tables to be created. It is therefore not possible to calculate the incidence of infection or confidence intervals around the likelihood ratios. CI = confidence interval; CRP = C reactive protein; I:T ratio = immature to total white cell ratio; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value; WBC = white blood cell; WCC = white cell count.

and were chosen for reporting here because they relate to tests that are commonly used in clinical practice. Data on diagnostic accuracy when combinations of these tests are used are also presented; these were reported in seven studies.^{18 31 34 35 40 47 52} Data from the remaining 27 papers assessing the other tests (tests specific for group B streptococcal infection, tests to diagnose neonatal conjunctivitis, tests using acute phase proteins other than C reactive protein, combinations of other diagnostic tests, clinical signs used to diagnose infection,

and a variety of miscellaneous tests) are available from the authors.

Individual study details are listed in table 1. The accuracy of each of the tests, or combination of tests, reviewed is shown in tables 2–5.

Discussion

The methodological difficulties in carrying out systematic reviews and meta-analyses evaluating diagnostic tests have been reported⁶² and we have addressed many of these in this study.

We used a search strategy designed to be as sensitive as possible⁸ and we believe that by using this strategy we will have found most of the appropriate studies. However, we did not search any database other than Medline, we did not specifically consult individual authors, and we did not attempt to identify possible unpublished studies. As with all systematic reviews, therefore, it is possible that some data have been missed.

The level of agreement between the investigators on which articles were relevant was exceptionally good. However, when the review process progressed to assessing the quality of individual studies, agreement fell, though it remained acceptable.⁶³ Both authors are experienced at assessing study design, and a significant factor contributing to the level of disagreement was that when reporting this type of study authors are not explicit enough in describing their methods.

The quality of many of the published studies is poor, in keeping with the findings of another review on this subject.⁶⁴ This is well illustrated by the fact that only 52 of a total of 572 reports met all three of the basic criteria designed to minimise the possibility of bias and improve the usefulness of any results. There were various common methodological flaws. Studies not infrequently repeated the test on a number of occasions in individual babies and then reported each result as a unique "event." Thus if those babies in whom the test was repeated were prone to a particular test result this could bias the overall assessment, although the direction of the bias could not necessarily be determined. The test and the criterion standard were often not independent of one another, again potentially introducing bias if only a given test result led to the gold standard being determined. Many studies also either used obviously septic infants or included a "control group" of perfectly well babies—populations in which the test would be unlikely to be used in clinical practice. Investigators conducting this type of research must consider these methodological issues when designing their studies, and also take care in reporting the methodology used as accurately and explicitly as possible.

We chose to report various measures assessing the accuracy of these tests, including the likelihood ratio. The interpretation of sensitivity and specificity is not intuitive to all clinicians and although positive and negative predictive values are perhaps of more value, they are only applicable in similar populations—their values vary depending on the prevalence of the outcome under consideration. To overcome this problem, use of the likelihood ratio allows clinicians to calculate the post-test probability of the outcome as long as some idea of the prevalence of the condition (pretest probability) is known. Less than one third (58/198) of the studies initially included in the review reported data in a way that allowed likelihood ratios to be calculated, thus limiting the clinical value of the information presented.

We have not carried out a formal meta-analysis on any of these results and have not therefore provided any pooled estimates. After the initial systematic review was complete and

the individual studies were available for scrutiny, it was felt that there was too much heterogeneity to justify any meta-analysis. The populations studied all varied in age, gestation, and selection criteria; there were numerous different criterion standards used; and very few of the tests were sufficiently similar, with a variety of cut off values being used. These differences are partly reflected in the heterogeneity of the result of similar tests reported in different studies.

Our choice of gold standard has often been used by others⁶⁵⁻⁶⁷ but does have some theoretical problems. If the number of genuinely infected infants is underrepresented by the gold standard—that is, some infected infants are not identified—then the positive predictive value of the test will be lower than in truth. However, this does not explain the poor negative predictive values that we frequently identified. For clinicians, this feature of the tests makes it very difficult to suggest either not starting treatment or stopping it on the basis of a negative result. Indeed, if some true infection were not picked up by culture, the true negative predictive value of these tests would be even lower than we report. On the other hand, in a small number of cases, bacterial isolates will actually represent poor aseptic technique, not true infection, and under these circumstances the accuracy of the results will be biased in the opposite direction.

Will the use of any of these investigations allow clinicians to alter their management? It has been suggested that a likelihood ratio between 0.1 and 10 is of limited use for predicting the presence or absence of a disease, since it will not substantially alter the pretest probability.⁷ Apart from a few exceptions, the likelihood ratios calculated from the studies included in this review lie within this indeterminate range and so they appear to be of limited value, either as individual tests or in combination. In our experience, when an infant presents with possible serious bacterial infection, clinicians understandably tend to act conservatively by performing some form of criterion standard (blood culture, urine culture, lumbar puncture, chest x ray, or a combination of these) and often start the infant on antibiotics, at least until the results of the criterion standard are available, when the situation is reviewed. Used singly, the diagnostic tests reported here are unlikely to change the pretest probability of a given child either being infected or not being infected, and so will not be much use in deciding whether to start or stop treatment. Our assessment of combinations of tests—a common clinical practice at present—showed equally disappointing results, although others have suggested this approach may be more promising.^{4 64} It is important, however, to recognise that regardless of the characteristics of any diagnostic test, the impact of different management strategies on any particular outcome can only truly be assessed by conducting appropriate randomised trials.

The quality of existing studies examining the accuracy of tests used to diagnose infection in the first three months of life is often poor and

future studies must be more rigorous. Valid data from existing studies suggest that tests are of limited value in the diagnosis of infection in this population.

Appendix

STATISTICAL DEFINITIONS USED

	Criterion standard		
	Positive	Negative	
"Test" positive	a	b	a + b
"Test" negative	c	d	c + d
	a + c	b + d	

Sensitivity = $a / a + c$, that is, the proportion of cases who are infected and have a positive test.

Specificity = $d / d + b$, that is, the proportion of cases who are not infected and have a negative test.

Positive predictive value = $a / a + b$, that is, the proportion of cases with a positive test who are infected.

Negative predictive value = $d / d + c$, that is the proportion of cases with a negative test who are not infected.

Likelihood ratio for a positive test = $[a / a + c] / [b / b + d]$, that is, post-test odds of infection = pretest odds \times likelihood ratio (odds = probability / $1 -$ probability).*

Likelihood ratio for a negative test = $[d / d + c] / [c / a + c]$, that is, post-test odds of no infection = pretest odds \times likelihood ratio (odds = probability / $1 -$ probability).*

*The mathematics can be avoided by using a nomogram for applying likelihood ratios (Fagan TJ. Nomogram for Bayes's theorem. *N Engl J Med* 1975;293:257). A straight line is drawn through the estimated pretest probability that the baby will experience the outcome of interest and the likelihood ratio associated with the given test result. The probability that the baby will now experience the outcome, given that particular test result (post-test probability), can simply be read off the nomogram.

- De Louvois J, Harvey D, eds. *Infection in the newborn*. Chichester: John Wiley, 1990.
- Young PC. The management of febrile infants by primary-care pediatricians in Utah: comparison with published practice guidelines. *Pediatrics* 1995;95:623-7.
- Baraff LJ, Bass JW, Fleisher GR, et al. Practice guidelines for the management of infants and children 0 to 36 months of age with fever without source. *Pediatrics* 1993;92:1-11.
- Lieu TA, Baskin MN, Schwartz JS, et al. Clinical and cost-effectiveness of outpatient strategies for management of febrile infants. *Pediatrics* 1992;89:1135-44.
- Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A: Are the results of the study valid? *JAMA* 1994;271:389-92.
- Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B: What are the results and will they help me in caring for my patients? *JAMA* 1994;271:703-7.
- Sackett DL, Haynes RB, Guyatt GH, et al, eds. *Clinical epidemiology. A basic science for clinical medicine*. Boston: Little, Brown, 1991.
- Haynes RB, Wilczynski N, McKibbon A, et al. Developing optimal search strategies for detecting clinically sound studies in Medline. *J Am Med Informatics Assoc* 1994;1:447-58.
- Adhikari M, Coovadia HM, Coovadia YM, et al. Predictive value of C-reactive protein in neonatal septicaemia. *Ann Trop Paediatr* 1986;6:37-40.
- Bonadio WA, Hennes H, Smith D, et al. Reliability of observations in distinguishing infectious outcome of febrile young infants. *Pediatr Infect Dis J* 1993;12:111-4.
- Bonadio WA, Smith DS, Mathews S, et al. Clinical significance of newly documented neutropenia in febrile young infants evaluated for sepsis. *Pediatr Infect Dis J* 1991;10:407-8.

- Boyle RJ, Chandler BD, Stonestreet BS, et al. Early identification of sepsis in infants with respiratory distress. *Pediatrics* 1978;62:744-50.
- Braude H, Forfar JO, Gould JC, et al. Diagnosis of urinary tract infection in childhood based on examination of paired non-catheter and catheter specimens of urine. *BMJ* 1967;iv:702-5.
- Buck C, Gallati H, Pohlandt F, et al. Increased levels of tumour necrosis factor α (TNF- α) and interleukin 1 β (IL-1 β) in tracheal aspirates of newborns with pneumonia. *Infection* 1994;22:238-41.
- Caspe WB, Chamudes O, Louie B. The evaluation and treatment of the febrile infant. *Pediatr Infect Dis* 1983;2:131-5.
- Crain EF, Gershel JC. Urinary tract infections in febrile infants younger than 8 weeks of age. *Pediatrics* 1990;86:363-7.
- Dagan R, Sofer S, Phillip M, et al. Ambulatory care of febrile infants younger than 2 months of age classified as being at low risk of having serious bacterial infections. *J Pediatr* 1988;112:355-60.
- DaSilva O, Hammerberg O. Diagnostic value of leukocyte indices in late neonatal sepsis. *Pediatr Infect Dis J* 1994;13:409-10.
- de Bont ESJM, Martens A, van Raan J, et al. Diagnostic value of plasma levels of tumour necrosis factor α (TNF- α) and interleukin-6 (IL-6) in newborns with sepsis. *Acta Paediatr* 1994;83:696-9.
- Dumornay W, Roblin PM, Gelling M, et al. Comparison of a chemiluminometric immunoassay with culture for the diagnosis of chlamydial infections in infants. *J Clin Microbiol* 1992;30:1867-9.
- Edgar JDM, Wilson DC, McMillan SA, et al. Predictive value of soluble immunological mediators in neonatal infection. *Clin Sci* 1994;87:165-71.
- El-Radhi AS, Jawad M, Mansor N, et al. Sepsis and hypothermia in the newborn infant: value of gastric aspirate examination. *J Pediatr* 1983;103:300-2.
- Evans ME, Schaffner W, Federspiel CF, et al. Sensitivity, specificity, and predictive value of body surface cultures in a neonatal intensive care unit. *JAMA* 1988;259:248-52.
- Fransen L, Nsanze H, Klauss V, et al. Ophthalmia neonatorum in Nairobi, Kenya: the roles of *Neisseria gonorrhoeae* and *Chlamydia trachomatis*. *J Infect Dis* 1986;153:862-9.
- Girardin EP, Berner ME, Grau GE, et al. Serum tumour necrosis factor in newborns at risk for infections. *Eur J Pediatr* 1990;149:645-7.
- Hammerschlag MR, Gelling M, Roblin PM, et al. Comparison of Kodak Surecell Chlamydia Test Kit with culture for the diagnosis of chlamydial conjunctivitis in infants. *J Clin Microbiol* 1990;28:1441-2.
- Hammerschlag MR, Roblin PM, Cummings C, et al. Comparison of enzyme immunoassay and culture for diagnosis of chlamydial conjunctivitis and respiratory infections in infancy. *J Clin Microbiol* 1987;25:2306-8.
- Hamoudi AC, Marcon MJ, Cannon HJ, et al. Comparison of three major antigen detection methods for the diagnosis of group B streptococcal sepsis in neonates. *Pediatr Infect Dis* 1983;2:432-5.
- Hardy JD, Furnell PM, Brumfit W. Comparison of sterile bag, clean catch and suprapubic aspiration in the diagnosis of urinary infection in early childhood. *Br J Urol* 1976;48:279-83.
- Harris H, Wirtschaffner D, Cassady G. Endotracheal intubation and its relation to bacterial colonization and systemic infection of newborn infants. *Pediatrics* 1976;56:816-23.
- Hiew TM, Tan AM, Cheng HK. Clinical features and haematological indices of bacterial infections in young infants. *Singapore Med J* 1992;33:125-30.
- Ingram DL, Pendergrass EL, Bromberger PI, et al. Group B streptococcal disease. Its diagnosis with the use of antigen detection, Gram's stain, and the presence of apnoea, hypotension. *Am J Dis Child* 1980;134:754-8.
- King JC, Berman ED, Wright PF. Evaluation of fever in infants less than 8 weeks old. *South Med J* 1987;80:948-52.
- Koenig JM, Patterson LER, Rench MA, et al. Role of fibronectin in diagnosing bacterial infection in infancy. *Am J Dis Child* 1988;142:884-7.
- Krediet T, Gerards L, Fleeer A, et al. The predictive value of C reactive protein and I/T-ratio in neonatal infection. *J Perinat Med* 1992;20:479-85.
- Leibovich M, Gale R, Slater PE. Usefulness of gastric aspirate examination in the diagnosis of neonatal infection. *Trop Geogr Med* 1987;39:15-17.
- McIntosh EDG, Jeffery HE. Clinical application of urine antigen detection in early onset group B streptococcal disease. *Arch Dis Child* 1992;67:1198-200.
- Misra PK, Kumar R, Malik GK, et al. Simple hematological tests for diagnosis of neonatal sepsis. *Indian Pediatr* 1989;26:156-60.
- Peakman M, Senaldi G, Liossis G, et al. Complement activation in neonatal infection. *Arch Dis Child* 1992;67:802-7.
- Philip AGS, Tito AM, Gefeller O, et al. Neutrophil elastase in the diagnosis of neonatal infection. *Pediatr Infect Dis J* 1994;13:323-6.
- Philip AGS. Detection of neonatal sepsis of late onset. *JAMA* 1982;247:489-92.
- Pourcyrus M, Bada H, Korones SB, et al. Significance of serial C-reactive protein responses in neonatal infection and other disorders. *Pediatrics* 1993;92:431-5.
- Rabalais GP, Bronfin DR, Daum RS. Evaluation of a commercially available latex agglutination test for rapid

- diagnosis of group B streptococcal infection. *Pediatr Infect Dis J* 1987;6:177-81.
- 44 Rench MA, Metzger TG, Baker CJ. Detection of group B streptococcal antigen in body fluids by a latex-coupled monoclonal antibody assay. *J Clin Microbiol* 1984;20:852-4.
 - 45 Rigal E, Roze JC, Villers D, et al. Prospective evaluation of the protected specimen brush for the diagnosis of pulmonary infections in ventilated newborns. *Pediatr Pulmonol* 1990;8:268-72.
 - 46 Roberts KB, Borzy MS. Fever in the first eight weeks of life. *Johns Hopkins Med J* 1977;141:9-13.
 - 47 Rodwell RL, Taylor KMCD, Tudehope DI, et al. Hematological scoring systems in early diagnosis of sepsis in neutropenic newborns. *Pediatr Infect Dis J* 1993;12:372-6.
 - 48 Rodwell RL, Leslie AL, Tudehope DI. Evaluation of direct and buffy coat films of peripheral blood for the early detection of bacteraemia. *Aust Paediatr J* 1989;25:83-5.
 - 49 Scanlon J. Early recognition of neonatal sepsis. The value of bacteriological studies of gastric aspirate and external ear canal material. *Clin Pediatr* 1972;11:258-60.
 - 50 Scanlon J. The early detection of neonatal sepsis by the examination of liquid obtained from the external ear canal. *J Pediatr* 1971;79:247-9.
 - 51 Scheifele DW, Melton P, Whitchelo V. Evaluation of Limulus test for endotoxemia in neonates with suspected sepsis. *J Pediatr* 1981;98:899-903.
 - 52 Seibert K, Yu VYH, Doery JCG, et al. The value of C-reactive protein measurement in the diagnosis of neonatal infection. *J Paediatr Child Health* 1990;26:267-70.
 - 53 Sherman MP, Chance KH, Goetzman BW. Gram's stains of tracheal secretions predict neonatal bacteraemia. *Am J Dis Child* 1984;138:848-50.
 - 54 Singhi S, Dhawan A, Kataria S, et al. Clinical signs of pneumonia in infants under 2 months. *Arch Dis Child* 1994;70:413-7.
 - 55 Squire E, Favara B, Todd J. Diagnosis of neonatal bacterial infection: hematological and pathological findings in fatal and nonfatal cases. *Pediatrics* 1979;64:60-4.
 - 56 Thompson PJ, Greenhough A, Gamsu HR, et al. Congenital bacterial sepsis in very preterm infants. *J Med Microbiol* 1992;36:117-20.
 - 57 Thore M, Faxelius G, Hedin G, et al. The role of a commercial latex agglutination test in the diagnosis of group B streptococcal infection in neonates. *Acta Paediatr Scand* 1991;80:167-72.
 - 58 Voora S, Srinivasan G, Lilien LD, et al. Fever in full term newborns in the first four days of life. *Pediatrics* 1982;69:40-4.
 - 59 Wagle S, Grauaug A, Kohen R, et al. C-reactive protein as a diagnostic tool of sepsis in very immature babies. *J Paediatr Child Health* 1994;30:40-4.
 - 60 Yuan S, Chikao S, Jianghuai W, et al. Role of tumour necrosis factor in neonatal sepsis. *Chin Med Sci J* 1994;9:45-8.
 - 61 Rodwell RL, Leslie AL, Tudehope DI. Early diagnosis of neonatal sepsis using a hematologic scoring system. *J Pediatr* 1988;112:761-7.
 - 62 Irwig L, Tosteson ANA, Gatsonis C, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;120:667-76.
 - 63 Sackett DL, Haynes RB, Guyatt GH, et al, eds. *Clinical epidemiology. A basic science for clinical medicine*. Boston: Little, Brown, 1991:30-1.
 - 64 Klassen TP, Rowe PC. Selecting diagnostic tests to identify febrile infants less than 3 months of age as being at low risk for serious bacterial infection: a scientific overview. *J Pediatr* 1992;121:671-6.
 - 65 Stoll BJ, Gordon T, Korones SB, et al. Early-onset sepsis in very low birth weight neonates: a report from the National Institute of Child Health and Human Development Neonatal Research Network. *J Pediatr* 1996;129:72-80.
 - 66 Isaacs D, Barfield CP, Grimwood K, et al. Systemic bacterial and fungal infections in infants in Australian neonatal units. Australian Study Group for Neonatal Infections. *Med J Aust* 1995;162:198-201.
 - 67 Jensen HB, Pollock BH. Meta-analyses of the effectiveness on intravenous immune globulin for prevention and treatment of neonatal sepsis. *Pediatrics* 1997;99(2).