

STARD를 이용한 진단검사법 성능평가 연구의 질 평가와 메타분석을 통한 3세대 HCV 효소면역 항체검사의 통합민감도와 특이도 분석

김솔잎¹ · 오홍범¹ · 차충환¹ · 최성은¹ · 안홍엽² · 이관제²

울산대학교 의과대학 서울아산병원 진단검사의학과, 동국대학교 통계학과²

Quality Evaluation of the Performance Study of Diagnostic Tests Using STARD Checklist and Meta-Analysis for the Pooled Sensitivity and Specificity of Third Generation Anti-HCV EIA Tests

Sollip Kim, M.D.¹, Heung-Bum Oh, M.D.¹, Chung-Hwan Cha, M.D.¹, Sung-Eun Choi, M.A.¹, Hong-yup An, Ph.D.²,
and Kwan Jeh Lee, Ph.D.²

Department of Laboratory Medicine, Asan Medical Center and University of Ulsan College of Medicine¹, Seoul;
Department of Statistics, Dongguk University², Seoul, Korea

Background : The third generation anti-hepatitis C virus (HCV) enzyme immunoassay (EIA) is now in use for screening HCV infection. The aim of this study was to pool the data on the sensitivity and specificity of third generation anti-HCV EIA tests after evaluating the quality of the studies using Standards for Reporting of Diagnostic Accuracy studies (STARD) checklist.

Methods : We searched MEDLINE and PubMed databases using keywords about the accuracy of diagnostic tests for HCV infections. Methodological quality was assessed by two persons with a modified STARD checklist. A heterogeneity test was performed, and in case heterogeneity was present, a sub-group analysis was done. Fixed-effects model was used to obtain pool sensitivity and specificity with 95% confidence intervals (CI).

Results : A total of 41 studies from 16 papers were selected. The quality score ranged from 6 to 13 (median 10.5); Inter-observer agreement was 93.62% ($k=0.69$); and 41 studies revealed heterogeneity. We performed a sub-group analysis with only 28 studies from 13 papers that were evaluated to be of high quality. A subgroup using polymerase chain reaction as the reference test revealed homogeneity and was calculated the pooled sensitivity and specificity of 99.92% (CI 99.77-100.07%) and 99.66% (CI 99.45-99.86%) respectively. Studies on test kits with an increased reactivity to the core region also showed homogeneity in sensitivity and the pooled sensitivity was 99.78% (CI 99.53-100.03%).

Conclusions : For the first time in Korea, the diagnostic accuracy of test kits was evaluated by meta-analysis using STARD checklist. The methodology shown in this study should help extending laboratory medicine to an evidence-based medicine. (*Korean J Lab Med* 2006;26:307-15)

Key Words : STARD, Meta-analysis, Third generation anti-HCV EIA

서론

접 수 : 2006년 4월 18일 접수번호 : KJLM1943
수정본접수 : 2006년 5월 15일
게재승인일 : 2006년 5월 17일
교신저자 : 오 홍 범
우 138-736 서울시 송파구 풍납2동 388-1
서울아산병원 진단검사의학과
전화 : 02-3010-4505, Fax : 02-478-0884
E-mail : hboh@amc.seoul.kr

*본 연구는 아산생명과학연구소 연구비(2005-219) 지원으로 이루어졌음.

C형 간염 바이러스(hepatitis C virus, HCV)는 1989년 처음 발견된 이후 만성간질환의 주요 원인으로 알려져 있다. HCV 감염의 유병률은 전 세계적으로 약 2%이며 1억 2천만명 가량이 이환되어 있다[1]. 한국에서의 유병률은 약 1%이며, 한국에서 B형 간염 예방접종이 성공적으로 정착된 이후에는 C형 간염이 만성간

질환의 중요한 원인으로 대두되고 있다[2]. 한국에서는 매년 5%의 C형 간염환자에서 간암이 발생되고 있고, 간암의 원인 중 12%가 C형 간염인 것으로 보고된 바 있다[2]. C형 간염은 예방접종이 없고, HCV에 노출된 후의 치료법도 알려지지 않아 안전한 혈액 공급, 보건위생 분야에서의 안전한 주사관습, 정맥주사 약물 남용자들을 줄이는 등의 예방이 중요하다[1]. 한국에서는 1991년부터 헌혈 혈액에 대하여 HCV 항체검사를 시행하고 있으며[2], 3세대 효소면역 항체검사를 이용한 선별검사에서도 헌혈자의 0.34%가 양성이었다[3].

미국 질병관리본부(US Centers for Disease Control and Prevention)와 국립건강협회(National Institute of Health) 등에서는 고위험군에서 선별검사를 추천하고 있다. 이는 새로운 항바이러스제 조합 치료법으로 45-80%의 환자에서 바이러스를 없앨 수 있고, 조직학적인 치유가 이루어지며, 상담을 통해 알코올 남용을 비롯한 위험한 행동을 막고, A형, B형 간염 예방접종 및 동반감염을 치료함으로써 병의 진행과 2차 전파를 막을 수 있기 때문이다. 또한 대부분의 환자에서 초기증상이 없어 병의 발견이 어려우므로 선별검사가 필수적이기 때문이다[4].

선별검사로 효소면역 항체검사법(enzyme immunoassay, EIA)이 많이 사용된다. 1세대 항-HCV EIA가 1990년에 소개되었으나 민감도와 특이도가 낮아 곧 2세대 항-HCV EIA가 개발되었고 민감도와 특이도가 상당히 높아졌다. 3세대 항-HCV EIA는 1993년부터 쓰이기 시작했고 core, NS3, NS4, NS5 구역을 검출할 수 있는 재조합 단백질이나 합성 펩타이드 등을 사용하여 민감도를 높이고 감염원에 노출된 후 6-8주에 항체를 검출할 수 있도록 성능이 향상되었다[4, 5]. Standards for reporting of diagnostic accuracy studies (STARD)는 진단검사법의 정확도에 대한 연구 보고의 질을 높이기 위해 2003년 1월 처음 제안된 지침으로서 25개 항목의 체크리스트를 제시하였다[6]. 본 연구에서는 문헌검색을 통해 얻은 3세대 항-HCV EIA에 대한 자료에 대해 STARD 체크리스트를 이용하여 연구의 질을 평가하고, 질이 우수한 연구논문에 대해 메타분석을 시행하여 C형 간염 진단키트의 통합 민감도와 통합특이도를 분석하고자 하였다.

재료 및 방법

1. 3세대 항 HCV EIA의 성능평가에 대한 연구문헌 선택과 자료 추출

1989년부터 2006년 동안 MEDLINE (1989-2005년 12월), PubMed (1989-2006년 3월) 데이터베이스에서 "hepatitis C", "mass screening", "hepatitis C antibodies", "sensitivity and specificity", "diagnostic accuracy", "HCV EIA" 등의 용어를 사용하여 문헌을 검색하였다. MEDLINE 검색에서는 모든 언어를 포함하였고 PubMed에서는 영어로 된 문헌만을 검색하였다. 검색된 연

구문헌과 함께 검색한 연구문헌의 참고문헌 목록과 관련문헌도 검토하여 선택에 포함시켰다.

선택된 문헌 중 3세대 EIA 검사키트의 이름, 대상 집단, 검체 수, 진양성 수, 진음성 수, 위양성 수, 위음성 수, 참고검사법 등을 추출하였다. 참고검사법으로는 혈청학적 검사법과 중합효소연쇄반응(polymerase chain reaction, PCR), 추적조사를 받고 있는 환자 또는 이들의 조합으로 하였다. 추출한 자료를 바탕으로 2×2 분할표를 작성하였다. 동물에 관한 연구, 연구결과가 제시되지 않은 중설(review article), 1세대 또는 2세대 EIA에 관한 연구 등은 제외하였다.

2. 방법론적 질의 평가(Methodological quality)

2명의 검토자가 독립적으로 각 문헌의 방법론적 질을 평가하였다. 평가기준으로는 STARD[6]를 본 연구의 특성에 맞추어 변형시킨 점검표를 사용하였다(Table 1). 해당 문헌에 평가항목이 잘 기술되어 있으면 1점, 잘 기술되어 있지 않으면 0점으로 하여 점수화 하였으며 그의 평균을 구하였다. 2명의 검토자간의 의견차이가 있었던 항목에 대해서는 토론을 통하여 의견일치를 이루도록 노력하였고, 검토자간의 일치도와 kappa 통계량을 계산하였다.

3. 자료 분석

각 문헌에 대하여 민감도, 특이도와 95% 신뢰구간(confidence interval, CI)을 계산하여 도식화하였다. STARD 평가에서 우수한 문헌만을 선택하여 분석하였다. 고정효과 모형(fixed effect model)으로 cochrane's Q test를 시행하여 선택된 문헌들간 동질성이 만족되는지 알아보았고, 문헌들간 이질성을 보이는 경우 하위집단 분석(sub-group analysis)을 시행하였다. 모든 통계분석은 Excel (version 2003, Microsoft, Redmond, WA, USA)과 SAS (version 9.1, Cary, NC, USA)를 사용하였다. 동질성을 만족하는 연구들을 기반으로 통합민감도(pooled sensitivity)와 통합특이도(pooled specificity)를 다음과 같은 식으로 구했으며 95% 신뢰구간을 제시하였다.

$$\text{Pooled sensitivity} = \frac{\sum_{i=1}^n TP_i}{\sum_{i=1}^n (TP_i + FN_i)}$$

$$\text{Pooled specificity} = \frac{\sum_{i=1}^n TN_i}{\sum_{i=1}^n (TN_i + FP_i)}$$

i = 통합하려는 연구의 수, TP = 진양성, FN = 위음성, TN = 진음성, FP = 위양성

결 과

1. 검색된 문헌과 자료 추출

MEDLINE (1989-2005.12)에서 "hepatitis C virus"와 "mass

Table 1. 진단검사법의 정확도에 대한 연구의 보고를 개선하기 위한 변형된 STARD 점검표

구분과 주제	항목 #	설명	쪽 #
제목/초록/중심단어	1	진단검사법 정확도에 대한 논문이라는 점을 명시(MeSH 제목으로는 '민감도와 특이도'가 추천됨)	
서론	2	연구의 논점이나 목적을 명시(예, 진단검사법의 정확도를 측정 또는 실험간과 참여자 그룹간의 정확도를 비교)	
방법	3-5	대상 집단에 대한 설명(포함기준, 제외기준, 자료수집의 시기와 장소, 참여자 모집과 자료 추출과정 등)	
참조표준검사	7	참조표준검사에 대한 설명(참조표준검사명과 이론적 배경)	
검사방법	8	검사방법에 대한 설명(방법을 상세히 설명하거나 검사방법에 대한 참고문헌이나 제조사의 설명지 제시)	
	9	단위, cutoff 값, 결과해석에 대한 설명	
	10	연구를 수행하고 결과를 해석하는데 전문가가 몇 명 참여하였는지 설명	
	11	맹검(blindness) 여부에 대한 설명	
통계학적 모형	12	정확도를 계산하고 비교하는 방법과 불확실도를 측정하기 위해 쓰인 통계법 설명(예, 95% confidence interval)	
	13	검사의 재현성을 계산한 방법에 대한 설명	
결과	14	언제 연구가 완성되었는지에 대한 설명(모집을 시작한 날과 끝난 날짜를 포함)	
참여자	15	참여자의 임상적, 인구학적 특성(예, 나이, 성별, 증상의 범위, 동반질환, 현재의 치료법, 모집기관 등)	
	16	포함기준을 만족시킨 참여자 중 몇 명이 index test와 참조표준검사를 받거나 받지 못했는지와 그 이유에 대한 설명 (순서도 사용을 추천)	
검사결과	17	Index test와 참조표준검사간의 시간 차이와 그 사이의 치료에 대한 설명	
	18	표적조건하에서 병의 중증도 분포에 대한 설명(표적조건이 없는 경우는 다른 진단명을 기술)	
	19	결과를 표로 작성하였는가	
평가	20	서로 다른 결과, 애매한 결과, 없어진 결과, outlier 등을 어떻게 처리하였는가에 대한 설명	
	21	연구과정 동안 생긴 adverse effect	
	22	진단검사법의 정확도와 통계적 불확실도를 측정(예, 95% confidence interval)	
	23	참여자의 하위집단, 해석자, 모집기관 사이의 정확도의 변이정도를 계산	
	24	검사의 재현성을 측정	
고찰	25	연구결과의 임상적 적용정도에 대한 토의	

Table 2. Studies including only the sensitivity of third generation anti-HCV EIA tests

Studies	Reference	Sample size	Spectrum of patients	Reference standard	Test assay	TP	FN	FP	TN	Sensitivity	Specificity
Ré (1)	[7]	22	HCV patients confirmed	PCR	Detect-HCV 3.0	20	2	-	-	90.9	-
Ré (2)			by PCR		Wiener anti-HCV	22	0	-	-	100.0	-
Ré (3)					Saronno Equipar anti-HCV	21	1	-	-	95.5	-
Ré (4)					Murex anti-HCV 4.0	22	0	-	-	100.0	-
Ré (5)					Fujirebio Serodia-anti-HCV	22	0	-	-	100.0	-
Berger (1)	[8]	45	Positive or borderline	PCR	Abbott HCV EIA 3.0	24	0	-	-	100.0	-
Berger (2)			samples by Abbott		Murex anti-HCV	24	0	-	-	100.0	-
Berger (3)			HCV EIA 2.0		Innotest anti-HCV III	24	0	-	-	100.0	-
Berger (4)					Roche Cobas Core anti-HCV EIA	24	0	-	-	100.0	-
Courouce (1)	[9]	309	HCV patients confirmed	Chiron RIBA	Abbott HCV EIA 3.0	307	2	-	-	99.4	-
Courouce (2)			by Chiron RIBA 3.0 SIA	3.0 SIA and	Innotest anti-HCV III	296	13	-	-	95.8	-
Courouce (3)			and PCR	PCR	Murex anti-HCV ver III	308	1	-	-	99.7	-
Courouce (4)					Sanofi Monolisa new Ag	309	0	-	-	100.0	-
Courouce (5)					Ortho HCV 3.0	308	1	-	-	99.7	-
Courouce (6)					Sorin ETI-Ab-HCV	301	8	-	-	97.4	-
Courouce (7)					UBI anti-HCV 4.0	301	8	-	-	97.4	-

screening” 조합으로 검색한 경우 233건, “hepatitis C antibodies”와 “sensitivity and specificity” 조합으로 검색한 경우 351건, “hepatitis C antibodies”와 “diagnostic accuracy” 조합으로 검색한 경우 14건, PubMed (1989-2006.3)에서 “HCV EIA”로 검색한 경우 447건이었다. 이들 검색된 문헌들과 각각의 참고문헌 목록과 관련문헌들을 검토한 결과 최종적으로 16개의 문헌이 선택되었다. 이 중 민감도만을 측정한 문헌이 3개[7-9], 특이도만 측정한 문헌이 4개[10-13], 민감도와 특이도 모두를 측정한 문헌이 9개였다[14-22](Table 2-4).

총 18종류의 검사키트가 포함되었다: Ortho HCV 3.0® (Ortho Diagnostics, Amersham, UK) 8회; Abbott HCV EIA 3.0® (Abbott Diagnostic, Abbott Park, IL, USA) 6회; Abbott ARCHITECT Anti-HCV® (Abbott Diagnostic), Cobas Core Anti-HCV EIA® (Roche Diagnostics, Basel, Switzerland), Murex Anti-HCV® (Murex Diagnostics, Dartford, UK) 4회; Abbott AxSYM HCV version 3.0® (Abbott Diagnostic), Innotest Anti-HCV III® (Innogenetics, Zwijnaarde, Belgium) 2회; Abbott IMX HCV version 3.0® (Abbott Diagnostic), Mo-

Table 3. Studies including only the specificity of third generation anti-HCV EIA tests

Studies	Reference	Sample size	Spectrum of patients	Reference standard	Test assay	TP	FN	FP	TN	Sensitivity	Specificity
Hennig (1)	[10]	4383	Low risk blood donors	Abbott Matrix HCV 2.0	Abbott AxSYM HCV ver 3.0	-	-	7	4374	-	99.8
Hennig (2)					Abbott IMX HCV ver 3.0	-	-	1	4380	-	100.0
Hennig (3)					Abbott HCV EIA 3.0	-	-	9	4372	-	99.8
Jonus (1)	[11]	3811	Low risk blood donors	Chiron RIBA 3.0 SIA	Abbott Modified Architect anti-HCV	-	-	3	3808	-	99.9
Jonus (2)					Abbott Current Architect anti-HCV	-	-	9	3802	-	99.8
Jonus (3)		1984	Hospitalized or diagnostic patients (random selection)		Abbott Current Architect anti-HCV	-	-	15	1918	-	99.2
Jonus (4)					Abbott Modified Architect anti-HCV	-	-	11	1922	-	99.4
Zachary (1)	[12]	2020	Routine samples to virology laboratory	Deciscan, RIBA, INNO-LIA	Biorad Monolisa anti-HCV Plus on the Evolis automate	-	-	7	1931	-	99.6
Zachary (2)					Abbott AxSYM HCV ver 3.0	-	-	16	1922	-	99.2
Lee (3)	[13]	9936	Commercial blood donors	Chiron RIBA 3.0 SIA and PCR	Ortho HCV 3.0	-	-	16	9846	-	99.8

Table 4. Studies including both the sensitivity and specificity of third generation anti-HCV EIA tests

Studies	Reference	Sample size	Spectrum of patients	Reference standard	Test assay	TP	FN	FP	TN	Sensitivity	Specificity
Vrielink (1)	[14]	1923	403 Blood donors,	PCR	Abbott HCV EIA 3.0	398	0	3	1051	100.0	99.7
Vrielink (2)			212 non-A, non-B hepatitis patients,		Murex anti HCV VK 47	397	1	7	1047	99.7	99.3
Vrielink (3)			253 multi-transfused patients and 1055 blood donors		Ortho HCV 3.0	398	0	1	1053	100.0	99.9
Abdel-Hamid	[15]	1134	Community-based longitudinal sample	Chiron RIBA 3.0 SIA	Abbott HCV EIA 3.0	97	1	2	1023	99.0	99.8
Ismail	[16]	495	Suspicious hepatitis patients	Chiron RIBA 3.0 SIA and chart review	Ortho/ECi	175	0	6	311	100.0	98.1
Huber	[17]	1090	Suspicious hepatitis patients	PCR	Roche Cobas Core anti-HCV EIA	107	7	30	946	93.9	96.9
Prince (1)	[18]	301	Blood donor (ALT >100 IU/L)	Chiron RIBA 2 or Abbott Matrix HCV 2.0 or PCR	Abbott HCV EIA 3.0	54	0	2	245	100.0	99.2
Prince (2)					Ortho HCV 3.0	54	0	8	239	100.0	96.8
Kodama	[19]	600	298 HCV patients and 302 normal people	RIBA, PCR and chart review	Ortho HCV 3.0	284	0	1	315	100.0	99.7
Stuyver	[20]	68	Hemodialysed patients	PCR	Ortho HCV 3.0	23	0	0	45	100.0	100.0
Lavanchy (1)	[21]	2545	400 HCV patients, 2000 routine samples and 134 tricky panels	INNO-LIPA HCV Ab III	Roche Cobas Core anti-HCV EIA	506	2	6	2028	99.6	99.7
Lavanchy (2)		2500	400 HCV patients, 2000 blood donors, 100 tricky panels	Abbott Matrix HCV	Roche Cobas Core anti-HCV EIA	401	0	0	2099	100.0	100.0
Lavanchy (3)		2545		INNO-LIPA HCV Ab III	Ortho HCV 3.0	506	6	6	2022	98.8	99.7
Lavanchy (4)		2500			Murex anti-HCV	499	9	1	2033	98.2	100.0
Marcellin	[22]	45	Blood donors (increased ALT)	Ortho RIBA 3 and PCR	Ortho HCV 3.0	12	0	0	33	100.0	100.0

nolisa Anti-HCV Plus version 2[®] (Bio-Rad, Marnes-La-Coquette, France), Ortho Vitros ECi Anti-HCV[®] (Ortho Diagnostic Systems, NJ, USA), Murex Anti-HCV VK 47[®] (Murex Diagnostics), Detect-HCV 3.0[®] (Biochem Immuno Systems Inc., Montreal, Quebec, Canada), Serodia-Anti-HCV[®] (Fujirebio Inc., Tokyo, Japan), Sanofi Monolisa New Ag[®] (Sanofi

diagnostics Pasteur, Marnes la Coquette, France), Equipar Anti-HCV[®] (Saronno (Va), Italy), ETI-Ab-HCV[®] (Sorin Biomedica, Saluggia, Italy), UBI Anti-HCV 4.0[®] (United Biomedical, Hauppauge, NY, USA), Wiener Anti-HCV[®] (Wiener Lab., 2000 Rosario, Argentina)이 1회씩 연구되었다.

2. 선택된 문헌의 방법론적 질 평가

2명의 검토자가 16개의 선택된 문헌을 대상으로 방법론적 질을 평가하였고 각 항목에 대한 두 검토자간의 일치도와 kappa 통계량 값은 Table 5와 같다. 일치를 이루지 못한 항목들은 대부분 해석상의 약간의 차이 때문이었으며, 이는 토론을 통해 대부분 해결되었으나 9개 항목에서는 일치를 이루지 못하였다. 각 문헌의 질 평가점수의 중앙값은 10.5 (범위, 6-13)였다. 메타분석에서 문헌의 방법론적 질을 평가할 때 질이 낮고 높음을 나눌 수 있는 기준점수가 없어서 본 연구에서는 히스토그램상 곡선이 양분되는 기준점인 7점 이하를 질이 낮은 문헌으로 간주하였다.

3. 선택된 각 문헌의 민감도, 특이도

각 문헌에 대하여 민감도, 특이도와 95% 신뢰구간은 Fig. 1과 같았다. 민감도는 90.9-100%, 특이도는 96.8-100%로 대상 환자군의 특성, 검체 수, 검사키트별 방법의 차이, 참조표준검사의 방법 등에 따라 차이가 있었다.

4. 연구의 동질성 평가

전체 연구에 대하여 동질성 평가를 한 결과 민감도는 Q통계량 131.22, $P<0.05$, 특이도는 Q통계량 447.49, $P<0.05$ 로 선택된 연구들이 이질적임을 알 수 있었다. 전체 16개 문헌의 총 41개의 연

Table 5. Inter-observer agreement (%) and kappa statistics of modified STARD checklist

Item	Inter-observer agreement (%)	Kappa statistics
1	93.33	0.76
2	93.33	0.76
3-5	100	1
6	100	1
7	100	1
8	100	0.5
9	93.33	0.84
10	93.33	0
11	66.67	0.53
12	80	0.44
13	53.33	-0.30
14	100	1
15	100	1
16	100	1
17	100	0.5
18	100	0.5
19	86.67	0.59
20	93.33	0.81
21	100	0.5
22	100	1
23	100	0.5
24	100	1
25	100	1
Mean	93.62	0.69

구들 중 표준검사법으로 PCR을 이용한 경우가 14개, 혈청학적 검사를 이용한 경우가 14개, PCR과 혈청학적 검사, 임상적 판단 등을 종합하여 이용한 경우가 13개였다. 따라서 표준검사법의 종류에 따라 하위집단 분석을 실시하였으나 동질성을 보이지 않았다. 연구의 질이 낮은 것으로 평가된 3개의 문헌을 제외한 13개 문헌의 총 28개 연구들에 대한 동질성 평가에서도 민감도는 Q통계량 32.80, $P<0.05$, 특이도는 Q통계량 206.27, $P<0.05$ 로 이질성을 보였다.

STARD 점수가 높은 총 28개의 연구들을 표준검사법에 따라 하위집단 분석을 시행한 결과 표준검사법으로 PCR을 이용한 경우 민감도는 Q통계량 2.30, $P=1.0$, 특이도는 Q통계량 5.34, $P=1.0$ 로 동질성을 보였다. 반면 표준검사법으로 혈청학적 검사를 이용한 경우는 이질성을 보였다. 표준검사법으로 PCR과 혈청학적 검사, 임상정보를 종합하여 사용한 경우는 모든 연구의 민감도 및 특이도가 각각 100%이었다(Table 6). 28개의 연구들을 검사방법별로 나누어 하위집단 분석을 한 결과 응집법(particle agglutination), 미세입자 효소면역검사(microparticle enzyme immunoassay), 화학발광 미세입자 검사(chemiluminescent microparticle immunoassay) 등의 방법을 제외하고 순수 EIA 방법을 이용한 연구들의 경우에도 이질적인 것으로 평가되었다(민감도 Q통계량 30.49, $P<0.05$; 특이도 Q통계량 120.80, $P<0.05$). 다시 NS3 영역에 반응성을 높인 검사법[5, 12, 21](Ortho HCV 3.0®, Murex

Table 6. Pooled sensitivity and specificity of third generation anti-HCV EIA tests grouped according to the type of reference test

Reference test	Number of journals (studies)	Total number of patients	Homo-geneity	Estimates (95% CI)
PCR				
Sensitivity	3 (8)	2036	Yes ($P=1.0$)	99.92% (99.77-100.07%)
Specificity	2 (4)	1991	Yes ($P=1.0$)	99.66% (99.45-99.86%)

Abbreviation: CI, confidence interval.

Table 7. Pooled sensitivity and specificity of third generation anti-HCV EIA tests grouped according to the region of index test

Region of increased reactivity	Number of journals (studies)	Total number of patients	Homo-geneity	Estimates (95% CI)
NS3				
Sensitivity	8 (12)	9161	No ($P<0.05$)	99.33% (99.00-99.66%)
Specificity	10 (13)	25895	No ($P<0.05$)	99.77% (99.71-99.83%)
Core				
Sensitivity	3 (5)	7013	Yes ($P=0.79$)	99.78% (99.53-100.03%)
Specificity	2 (3)	6968	No ($P<0.05$)	99.75% (99.61-99.89%)

Abbreviation: CI, confidence interval.

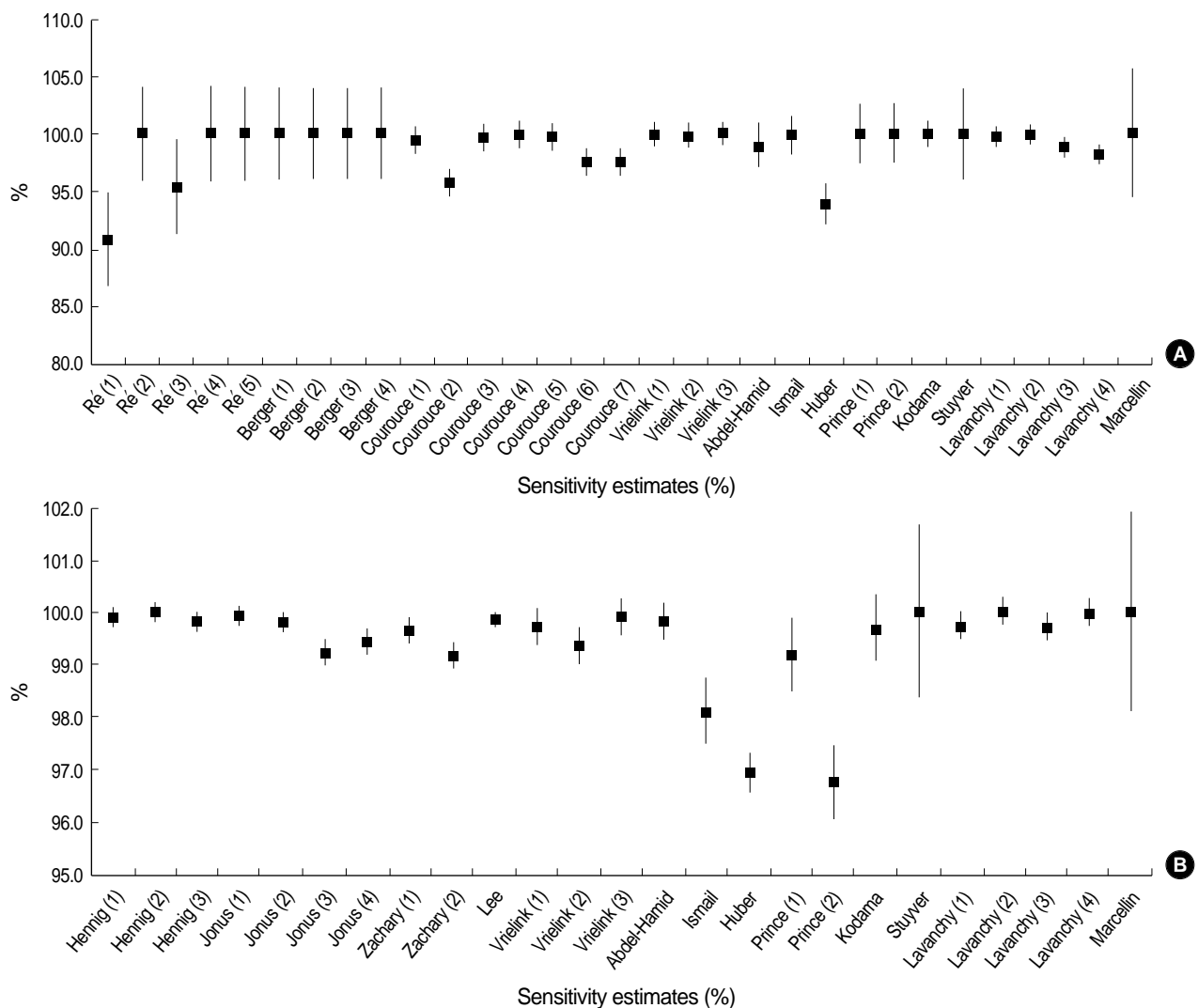


Fig. 1. Estimates from the studies of sensitivity and specificity of third generation anti-HCV EIA tests. Points indicate estimates of sensitivity (A) and specificity (B). Vertical lines are 95% confidence intervals for estimates.

Anti-HCV®, Abbott HCV EIA 3.0®, Monolisa Anti-HCV Plus on the Evolis Automate®)을 이용한 연구들을 모아서 동질성 평가를 한 결과 민감도는 Q통계량 19.96, $P < 0.05$, 특이도는 Q통계량 108.83, $P < 0.05$ 로 이질성을 보였다. 반면 core 영역에 반응성을 높인 검사법[5, 12, 21](Innotest Anti-HCV III®, Roche Cobas Core Anti-HCV EIA®, Murex Anti-HCV VK 47®)을 이용한 연구들을 모아서 동질성 평가를 한 결과 민감도는 Q통계량 1.70, $P = 0.79$ 로 동질성을 보였고, 특이도는 Q통계량 12.64, $P < 0.05$ 로 이질성을 보였다(Table 7).

5. 통합민감도 및 통합특이도

동질성을 보인 연구결과들에 대해 통합민감도와 통합특이도를 구하였다. STARD 점수가 높은 28개 연구결과 중 PCR 표준검사법을 사용한 경우는 통합민감도 99.92% (CI 99.77-100.07%), 통

합특이도는 99.66% (CI 99.45-99.86%)이었다(Table 6). 또한 검사방법에 따른 분류에서는 core 영역에 반응성을 높인 검사법에서 민감도 부분이 동질하였는데 이들 연구결과와 통합민감도는 99.78% (CI 99.53-100.03%)이었다(Table 7).

고 찰

본 연구에서는 문헌검색을 통해 얻은 3세대 항 HCV EIA에 대한 자료를 메타분석하여 C형 간염 진단을 위한 통합민감도와 통합특이도를 구하였다. STARD 점검표를 기준으로 각 연구의 질을 평가한 후 우수한 문헌을 선정하고 연구의 동질성 여부를 판정하는 과정을 거쳤다. STARD 점검표를 이용하여 진단검사법 평가논문의 질을 평가한 것은 국내에서 처음 시도된 일이다. 본 연구에서는 외국 문헌만을 대상으로 하였는데 선정된 문헌의 평가점수는 중

양값이 10.5 (범위, 6-13)이었다. 각 항목별로 검토자간 일치율과 kappa 통계량 값을 구하였는데, 일치율 93.62%, kappa 통계량 0.69로 비교적 일치율이 좋았다. 16개의 선택된 문헌의 41개 연구 결과에 대한 동질성 평가에서는 이질적인 것으로 평가되었다. 연구의 질이 낮은 것으로 평가된 3개의 문헌을 제외하고 표준검사법에 따라 나누어 하위집단 분석을 하였을 경우, PCR법에 의한 경우가 동질한 것으로 평가되었다. 본 연구에서는 동질한 문헌을 대상으로 통합민감도 및 통합특이도를 구하였는데 각각 99.92% (CI 99.77-100.07%), 99.66% (CI 99.45-99.86%)로 매우 우수함을 알 수 있었다. 검사방법의 차이에 따라 나누어 하위집단 분석을 한 경우에는 core 영역에 반응성을 높인 검사법을 이용한 연구들에서 동질하다는 결론을 얻을 수 있었고 통합민감도는 99.78% (CI 99.53-100.03%)이었다. 이러한 통합민감도 및 통합특이도는 새로운 키트가 개발되어 인허가 과정에서 생물학적 동등성 평가를 받고자 할 때 기존 검사법의 성능 수준으로 이용될 수 있을 것이다.

진단검사법의 정확도에 대한 연구들은 이질성을 보이는 경우가 많고 또한 방법론적으로도 제한된 정보만을 제공하기 때문에 완벽한 메타분석이 어려운 것으로 알려져 있다[23]. 이번 연구에서도 16개 문헌의 41개 연구 결과들을 모두 통합하였을 때 이질적으로 나와 하위집단 분석이 필요하였다. 표준검사법의 종류와 비교대상 검사법의 차이에 따라 나누었을 때 동질성을 보였다. 따라서 HCV 항체검사 키트의 정확도에 대한 연구를 시행할 때 표준검사법의 종류와 비교대상 검사의 방법을 명확히 명시해야 함을 알 수 있었다.

STARD는 2003년 1월 처음 제안된 이후 "Clinical Chemistry", "Radiology" 등 12개의 유수한 저널에 실려, 연구의 질 평가 기준으로 채택되고 있다. 전 세계적으로 무수히 많은 새로운 진단검사법이 빠른 속도로 개발되고 있으며 최근 그 속도는 더욱 빨라지고 있다. 진단검사법에 대한 연구가 잘못 설계, 보고된다면 이로 인해 잘못된 진단 및 치료를 하는 결과를 낳게 된다. 따라서 진단검사법의 정확도 연구에 있어 연구의 설계, 수행과정, 보고방법 등의 방법론적 부분이 강조되고 있다. STARD에서는 연구자가 방법론적으로 올바른 연구를 할 수 있도록 25개 항목의 점검표를 제시하고 있다[6]. 최근 STARD 점검표의 재현성에 대한 연구에서도 좋은 결과가 나와[24] 앞으로도 진단검사법의 정확도에 대한 연구에서 더욱 중요한 역할을 할 것으로 생각된다. 본 연구에서도 방법론적 질 평가에서 두 검토자간의 일치율이 93.62%, kappa 통계량 0.69로 비교적 높게 나타났다. 그러나 항목 11, 13에 대해서는 일치율이 각각 66.67%, 53.33%로 낮았는데(Table 5), 이는 본래의 문헌 자체에 각 항목에 대한 내용이 애매하게 기술되어 있었기 때문으로 여겨지며 이에 대한 보완이 필요할 것으로 사료된다. 또한 질이 낮고 높음을 나눌 수 있는 기준이 없어 이번 연구에서는 자료값들의 히스토그램을 그려 두 부류로 나누었다. 향후 연구의 질의 높고 낮음을 구분할 수 있는 기준점 설정에 대한 연구나, 질적 수준에 따라 각 연구의 비중을 어떻게 달리할 것인지에 대한 연구들이 진행되어야 할 것이다.

문헌으로 보고되는 연구결과들은 통계적으로 유의한 결과들이 대부분 선정되기 때문에 메타분석시 실제보다 더 긍정적인 결론을 낳는 오류를 범할 수 있다[25]. 이를 출판편견(publication bias)이라 하는데, 이런 출판편견은 randomized trial 원칙을 철저히 고수하는 임상연구에 비해 원칙을 철저히 준수하지 않는 진단검사법에 대한 연구에서 더 문제가 된다고 알려져 있다[26]. 따라서 본 연구와 같이 진단검사법의 정확도에 대한 메타분석을 실시할 때에는 출판편견이 있는지를 알아볼 필요가 있다. 그러나 출판편견을 구하기 위해서는 민감도와 특이도를 동시에 측정한 자료가 필요한데[25] 이번 분석에서는 민감도와 특이도를 동시에 측정한 연구의 수가 적어 출판편견을 계산하지 못하였다.

방대한 자료와 시간적 제약, 통계적 접근의 어려움으로 인해 지금까지 의학 분야에서 메타분석을 실시하기는 어려웠다. 그러나 각종 학술정보의 범람 속에서 체계적으로 압축된 지식이 필요하고, 새로운 검사법에 대해 상반된 결론이 야기될 때 메타분석법은 매우 유용하다. 이를 반영하듯 의학 분야에서도 계통적 문헌고찰을 통한 메타분석이 늘어나고 있다[23]. 진단검사의학이 근거중심의 학으로 진일보하는데 있어서도 메타분석은 매우 긴요한 방법론 중의 하나이므로 향후 이 분야에 대한 연구와 적용이 진단검사의학 분야에서도 활발히 이루어져야 할 것으로 생각한다.

결론적으로, 3세대 HCV 효소면역 항체검사의 정확도에 대한 연구들을 통합하여 메타분석 하였을 때 집단들은 매우 이질성을 보였으나 하위집단 분석에서는 일부에서 동질성을 보이는 것을 확인할 수 있었다. 동질성을 보인 연구들의 통합민감도는 99.92-99.78% (CI 99.53-100.07), 통합특이도는 99.66-99.75 (CI 99.45-99.86)로 매우 높게 나타났다. STARD를 이용하여 연구결과와 질을 평가하여 이를 토대로 메타분석을 시행한 사례는 국내에 본 연구가 처음이다. 본 연구에 제시된 방법론은 향후 진단검사의학이 근거중심의학으로 발전하는 데 긴요하게 이용될 것이다.

요 약

배경 : 3세대 효소면역 항체검사법(EIA)은 C형 간염의 선별검사로 많이 쓰이고 있다. 본 연구에서는 문헌검색을 통해 얻은 3세대 항 HCV EIA에 대한 자료에 대해 STARD 체크리스트를 이용하여 연구의 질을 평가하고, 질이 우수한 연구문헌을 대상으로 메타분석을 시행하여 C형 간염 진단키트의 통합민감도와 통합특이도를 분석하고자 하였다.

방법 : MEDLINE과 PubMed 데이터베이스에서 C형 간염 진단키트의 정확도와 관련된 용어로 검색하였다. 검색된 문헌을 2명의 검토자가 STARD 점검표를 이용하여 문헌의 방법론적 질을 평가하였다. 선택된 문헌들간 동질성이 만족되는지를 알아보고 고정효과모형을 사용하여 통합민감도, 통합특이도와 95% 신뢰구간(CI)을 구하였다. 문헌들간 이질성을 보이는 경우 하위집단 분석을 하였다.

결과 : 검색결과 최종적으로 16개 문헌에서 41개 연구결과가 선택되었다. 문헌의 질 평가점수의 중앙값은 10.5 (범위, 6-13)이었고, 일치율과 kappa 통계량은 각각 93.62%, 0.69로 비교적 좋았다. 41개 연구결과는 이질적인 것으로 평가되었다. 연구의 질이 높은 28개 연구를 표준검사법에 따라 나누어 하위집단 분석을 시행한 결과, PCR법에 의한 경우가 동질하였으며 통합민감도는 99.92% (CI 99.77-100.07%), 통합특이도는 99.66% (CI 99.45-99.86%)이었다. 또한 core 영역에 반응성을 높인 검사법에 대한 민감도 연구에서도 동질하였는데 통합민감도는 99.78% (CI 99.53-100.03%)이었다.

결론 : STARD를 이용하여 진단검사법 성능연구의 질을 평가하고 이를 토대로 검사키트의 메타분석을 시행한 사례는 국내에서 본 연구가 처음이다. 본 연구에 제시된 방법론은 향후 진단검사의 학이 근거중심의학으로 발전하는 데 진요하게 이용될 것이다.

참고문헌

- Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis* 2005;5:558-67.
- Suh DJ and Jeong SH. Current status of hepatitis C virus infection in Korea. *Intervirology* 2006;49:70-5.
- Oh HB, Hwang YS, Cho YJ, Kim DS, Kim SI. Experience of anti-HCV antibody immunoblot test in Korean blood donors. *Korean J Blood Transfus* 1997;8:1-8. (오흥범, 황유성, 조연정, 김두성, 김상인. 국내 헌혈 자에서의 항-HCV항체 면역블롯검사 경험. *대한수혈학회지* 1997;8:1-8.)
- Wong T and Lee SS. Hepatitis C: a review for primary care physicians. *CMAJ* 2006;174:649-59.
- Colin C, Lanoir D, Touzet S, Meyaud-Kraemer L, Bailly F, Trepo C. Sensitivity and specificity of third-generation hepatitis C virus antibody detection assays: an analysis of the literature. *J Viral Hepat* 2001; 8:87-95.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Standards for Reporting of Diagnostic Accuracy*. *Clin Chem* 2003;49:1-6.
- Re V, Gallego S, Trevino E, Barbas G, Dominguez C, Elbarcha O, et al. Evaluation of five screening tests licensed in Argentina for detection of hepatitis C virus antibodies. *Mem Inst Oswaldo Cruz* 2005;100: 303-7.
- Berger A, Doerr HW, Preiser W, Weber B. Lack of correlation between different hepatitis C virus screening and confirmatory assays. *J Virol Methods* 1996;59:141-6.
- Courouge AM, Barin F, Botte C, Lunel F, Maisonneuve P, Maniez M, et al. A comparative evaluation of the sensitivity of seven anti-hepatitis C virus screening tests. *Vox Sang* 1995;69:213-6.
- Hennig H, Schlenke P, Kirchner H, Bauer I, Schulte-Kellinghaus B, Bludau H. Evaluation of newly developed microparticle enzyme immunoassays for the detection of HCV antibodies. *J Virol Methods* 2000;84:181-90.
- Jonas G, Pelzer C, Beckert C, Hausmann M, Kapprell HP. Performance characteristics of the ARCHITECT anti-HCV assay. *J Clin Virol* 2005;34:97-103.
- Zachary P, Ullmann M, Djeddi S, Wendling MJ, Schvoerer E, Stoll-Keller F, et al. Evaluation of two commercial enzyme immunoassays for diagnosis of hepatitis C in the conditions of a virology laboratory. *Pathol Biol (Paris)* 2004;52:511-6.
- Lee SR, Wood CL, Lane MJ, Francis B, Gust C, Higgs CM, et al. Increased detection of hepatitis C virus infection in commercial plasma donors by a third-generation screening assay. *Transfusion* 1995; 35:845-9.
- Vrieling H, Zaaijer HL, Reesink HW, van der Poel CL, Cuypers HT, Lelie PN. Sensitivity and specificity of three third-generation anti-hepatitis C virus ELISAs. *Vox Sang* 1995;69:14-7.
- Abdel-Hamid M, El-Daly M, El-Kafrawy S, Mikhail N, Strickland GT, Fix AD. Comparison of second- and third-generation enzyme immunoassays for detecting antibodies to hepatitis C virus. *J Clin Microbiol* 2002;40:1656-9.
- Ismail N, Fish GE, Smith MB. Laboratory evaluation of a fully automated chemiluminescence immunoassay for rapid detection of HBsAg, antibodies to HBsAg, and antibodies to hepatitis C virus. *J Clin Microbiol* 2004;42:610-7.
- Huber KR, Sebesta C, Bauer K. Detection of common hepatitis C virus subtypes with a third-generation enzyme immunoassay. *Hepatology* 1996;24:471-3.
- Prince AM, Scheffel JW, Moore B. A search for hepatitis C virus polymerase chain reaction-positive but seronegative subjects among blood donors with elevated alanine aminotransferase. *Transfusion* 1997; 37:211-4.
- Kodama T, Ichiyama S, Sato K, Nada T, Nakashima N. Evaluation of a membrane filter assay system, Ortho HCV Ab Quick Pack, for detection of anti-hepatitis C virus antibody. *J Clin Microbiol* 1998; 36:1439-40.
- Stuyver L, Claeys H, Wyseur A, Van Arnhem W, De Beenhouwer H, Uytendaele S, et al. Hepatitis C virus in a hemodialysis unit: molecular evidence for nosocomial transmission. *Kidney Int* 1996;49:889-95.
- Lavanchy D, Steinmann J, Moritz A, Frei PC. Evaluation of a new automated third-generation anti-HCV enzyme immunoassay. *J Clin Lab Anal* 1996;10:269-76.
- Marcellin P, Martinot-Peignoux M, Gabriel F, Branger M, Degott C, Elias A, et al. Chronic non-B, non-C hepatitis among blood donors

- assessed with HCV third generation tests and polymerase chain reaction. *J Hepatol* 1993;19:167-70.
23. Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002;2:9.
24. Smidt N, Rutjes AW, van der Windt DA, Ostelo RW, Bossuyt PM, Reitsma JB, et al. Reproducibility of the STARD checklist: an instrument to assess the quality of reporting of diagnostic accuracy studies. *BMC Med Res Methodol* 2006;6:12.
25. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. *JAMA* 2006;295:676-80.
26. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 1995;48:119-30.