의료질평가지원금 평가지표의 변별력 검토

Discrimination Ability of Indicators for the Healthcare Quality Evaluation Grant Initiative in Korea National Health Insurance Program

강희정, 하솔잎

한국보건사회연구원

Hee-Chung Kang, Solleep Ha

Korea Institute for Health and Social Affairs(KIHASA)

Correspondence to: Hee-Chung Kang

370, Sicheong-daero, Sejong-si, Korea

Tel: 044-287-8106

E-mail: hee-chungkang@kihasa.re.kr

Acknowledgements

이 연구는 2016년도 보건복지부 보건의료정책과의 지원을 받아 수행되었다.

Received: April 9, 2018

Revised: August 16, 2018

Accepted after revision: October 11, 2018

© Published by Korean Society for Public Health and Medicine; all rights reserved

Abstract

Objective: We examined how well each quality indicator discriminate quality differences among hospitals in the Healthcare Quality Evaluation Grant initiative of Korea National Health Insurance Program.

Methods: We analyzed hospital-level quality indicators which were used for the decision of 2016 Healthcare Quality Evaluation Grant initiative. We examined each indicator's discrimination ability using the criteria (median $\geq$ 95 percentile & inter-quartile range percent $<$ 4.5%)

Results: Current Healthcare Quality Evaluation Grant initiative is mainly composed of structure or process indicators. Most structural quality indicators of tertiary general hospitals have no discrimination ability unlike general hospitals.

Conclusion: This study shows that indicators of the Quality Evaluation Grant initiative are systematically reflecting structural differences between tertiary general hospitals and general hospitals. We need to restructure quality indicators from structure-based to outcome-based and to examine the introduction of absolute rating approaches for overcoming limitations of current relative ratings and improve its effects as a quality based P4P program.

Keywords: quality of care, hospital, quality assessment, quality evaluation

서론

국제연합(United Nations, UN) 지속가능발전목표(Sustainable Development Goal: SDG)의 세 번째 목표인 “생애 전 단계에서 모든 사람의 건강과 웰빙(well-being) 보장”을 달성하기 위해서는 의료의 질(quality of care) 향상을 전제로 인구(population), 서비스(service), 비용(cost) 측면의 보장성 확대가 중요하다[1, 2]. 의료의 질 향상 정책은 국민이 질 높은 서비스에 접근하는 공평한 기회를 보장하고 보험자가 보장성 확대에 따른 지출 압박의 상황 속에서 효율성을 유지할 수 있는 정책적 도구로서 유용하다. 아울러, 모든 공급자가 보다 투명한 경쟁을 통해 공평하게 수입을 증가시키는 기회를 얻을 수 있다. 의료의 질은 의료 효과성(effectiveness) 중심의 전통적 개념에서 환자안전(patient safety), 환자중심성(patient centeredness), 낭비를 줄이는 효율성(efficiency), 의료제공자 간 의료 연계(care coordination), 질의 격차(disparities) 해소를 포함하는 포괄적 개념으로 확대되어 왔고 접근성(access), 비용(cost)과 함께 보건의료체계의 핵심 성과 중 하나로 평가되고 있다[3-6].

정부는 건강보험 보장성 강화의 일환으로 전액 환자가 부담하는 선택 진료 제도를 2014년부터 2017년까지 단계적으로 축소하여 폐지하였고, 의료기관의 손실을 보상하기 위하여 2015년 9월부터 ‘의료질평가지원금’제도를 도입했다. 이 제도는 종합병원 이상 의료기관에 대해 의료의 질을 평가하고 등급에 따라 차등 수가를 보상한다. 공급자 지불보상이 주로 행위별수가제에 기초하고 있는 우리나라 건강보험제도에서 의료질평가지원금 제도의 도입은 양(volume) 중심에서 질(quality) 중심으로 지불보상 방식을 변화시키는 계기로 의미가 크다[1]. 다만, 제도 설계 당시 사용가능한 평가 지표가 충분하지 않았고 선택진료제도 축소에 따른 의료기관의 손실보상 목적이 우선되었기 때문에 평가지표가 의료의 질 향상을 유도하는 관련성에 있어서 국내 의료현장의 실제 경험에 기초하지 못했던 한계가 있었다.

의료질평가지원금제도는 의료의 질 향상이 비용을 통제하는 핵심전략이라는 관점에서 질 향상 인센티브를 통해 제도 전반의 효율화를 기대하는 성과보상지불(pay for performance, P4P)모형에 이론적 근거를 두고 있다[7-8]. 올해는 제도 도입이후 3년차가 되며 선택진료제도가 완전히 폐지되는 해이다. 이제부터는 평가와 보상을 통해 의료기관의 자발적 질 향상 활동을 촉진하도록 동 제도의 평가지표와 질 향상의 관련성 제고에 더욱 집중해야 한다. 이를 위해서는 단계적으로 성과보상지불 모형의 성공 요소를 확대하는 제도 교정이 필요한데, 평가지표의 개편은 무엇보다 우선되어야 하는 핵심 요소이다[8-10].

평가지표의 개편은 지표의 선정과 관리의 두 가지 측면에서 접근될 수 있다. 지표 선정 원칙과 관리체계에 따라 지속적으로 진입, 유지, 갱신, 퇴출되는 순환이 필요하기 때문이다[10]. 의료질평가지원금제도 도입 시 관련 연구를 바탕으로 평가지표 선정기준을 영향(impact), 측정가능성(feasibility), 개선가능성(improvability), 수용성(acceptability)으로 설정한 바 있으나 명확히 규정화되어 있지는 않다[7]. 이 원칙은 의료질평가지원금제도가 의료기관 간 질적 차이를 바르게 측정하여 의료기관이 평가 지표와 결과를 신뢰하고 개선 노력을 함으로써 의료의 질 향상에 대한 제도의 영향력을 높이도록 관련 연구와 외국의 제도들에서 공통으로 제시하고 있는 선정 기준을 반영하고 있다[10].

지표 선정 원칙에 따른 평가지표의 적절성 평가는 통계적으로 지표의 변별력을 평가하는 것에서 출발한다[10]. 지표의 순환은 통계적 타당성과 임상적 전문성의 결합을 바탕으로 주기적인 검토와 논의를 통해 이루어 질 수 있다. 그러나 질 평가에 기초한 지불보상제도에서 평가지표 선정 과정 상 체계적(systematic), 참여적(participatory), 근거 기반 방식(evidence based manner)의 부재로 문제가 제기될 수 있다[11]. 의료질평가지원금제도 또한 이러한 문제를 안고 있다.

보건복지부 고시인 「의료질평가지원금 산정을 위한 기준」이 평가지표를 포함하여 의료 질 평가 관련 사항을 심의하도록 의료질평가심의위원회의 운영을 규정하고 있지만, 이 위원회는 이해단체 대표로 구성된 심의기구이고 평가지표 순환을 위한 전문적 역할을 기대하기 어렵다[12].

2018년 의료질평가지원금제도의 평가지표는 ‘의료의 질과 환자안전’, ‘공공성’, ‘의료전달체계’, ‘교육수련’, ‘연구개발’의 다섯 영역을 평가하는 59개 지표로 구성되어 있다[13]. 제도 도입 당시 일천억 원에서 출발하여 올해 칠천억 원으로 규모가 확대됨에 따라 평가지표의 선정기준과 관리체계에 대한 관심이 확대될 것이다.

이 연구의 목적은 현재 평가지표가 의료기관의 질적 차이를 통계적으로 구분해내는 변별력을 검토하고자 하는 것이다. 연구결과는 향후 평가지표 개편과 관리체계 구축을 목적으로 임상적 전문성을 반영하는 참여적 논의의 기초정보를 제공할 것이다.

연구방법

1. 연구 대상 및 자료

2016년 의료질평가지원금 평가등급 결정을 위해 수집된 종합병원과 상급종합병원의 평가지표 정보를 바탕으로 기관별 연 단위 데이터를 구축하였다. 예를 들어, 간호차등제 등급과 같이 분기별로 연 4회 조사되는 경우는 가장 최근 값인 12월 값으로 데이터를 구성했다. 분석에 포함된 의료기관은 모두 322개 기관으로 상급종합병원 43개(13.4%), 종합병원 279개(86.6%)였다(Table 1).

2. 분석방법

322개 의료기관 단위로 의료 질과 환자안전 영역의 16개 지표, 공공성 영역의 5개 지표, 의료전달체계 영역의 6개 지표가 분석에 사용되었다(Table 2~4). 분석 결과는 상급종합병원과 종합병원, 16개 시도별로 구분하여 비교되었다. 모든 지표를 0~1사이 범위를 갖도록 표준화한 후 지표별로 중앙값이 95 백분위수(percentile) 이상이고 지표 값들의 중간 50%가 흩어진 정도(Q3-Q1)를 의미하는 사분범위(IQR, the interquartile range)가 4.5%미만이면 변별력이 없는 지표로 구분하였다(Reeves et al., 2010; 신영석 외 2015). 이 기준은 평가대상 병원들의 대부분이 높은 수준의 지표 값을 달성하여 해당 지표가 의료기관간 질적 차이를 구분하는 변별력이 없다고 정의하는 것이다.

연구결과

1. 분석 대상

43개 종합병원은 모든 영역에서 평가등급이 발생되었으나 종합병원은 의료질과 환자안전, 공공성, 의료전달체계 영역에서 78.5%(219/279), 교육수련 영역에서 55.6%(155/279), 연구개발 영역에서 29.7%(83/279)만 평가등급이 발

생하였다(Table 1). 평가지원금 대상에서 제외되는 등급제외 기관 발생률을 지역별로 비교해보면, 광주, 전남, 울산지역에서 높았다(Table 1, Figure 1).

〈Table 1〉 Subject hospitals by 5 dimensions of Healthcare Quality Evaluation Grant Initiative (as of 2016)

unit: N(%)

Dimension /Region	① Healthcare Quality & Patient Safety, ② Publicness, ③ Healthcare Delivery System			④ Education & Training			⑤ Research & Development		
	Excluding*	Ratings*	Subtotal**	Excluding*	Ratings*	Subtotal**	Excluding*	Ratings*	Subtotal**
Tertiary H.		43(100.0)	43(100.0)		43(100.0)	43(100.0)		43(100.0)	43(100.0)
Kangwon		1(100.0)	1(2.3)		1(100.0)	1(2.3)		1(100.0)	1(2.3)
Gyeonggi		5(100.0)	5(11.6)		5(100.0)	5(11.6)		5(100.0)	5(11.6)
Gyeongnam		2(100.0)	2(4.7)		2(100.0)	2(4.7)		2(100.0)	2(4.7)
Gyeongbuk									
Gwangju		2(100.0)	2(4.7)		2(100.0)	2(4.7)		2(100.0)	2(4.7)
Daegu		4(100.0)	4(9.3)		4(100.0)	4(9.3)		4(100.0)	4(9.3)
Daejeon		1(100.0)	1(2.3)		1(100.0)	1(2.3)		1(100.0)	1(2.3)
Busan		4(100.0)	4(9.3)		4(100.0)	4(9.3)		4(100.0)	4(9.3)
Seoul		14(100.0)	14(32.6)		14(100.0)	14(32.6)		14(100.0)	14(32.6)
Ulsan		1(100.0)	1(2.3)		1(100.0)	1(2.3)		1(100.0)	1(2.3)
Incheon		3(100.0)	3(7.0)		3(100.0)	3(7.0)		3(100.0)	3(7.0)
Jeonnam		1(100.0)	1(2.3)		1(100.0)	1(2.3)		1(100.0)	1(2.3)
Jeonbuk		2(100.0)	2(4.7)		2(100.0)	2(4.7)		2(100.0)	2(4.7)
Jeju									
Chungnam		2(100.0)	2(4.7)		2(100.0)	2(4.7)		2(100.0)	2(4.7)
Chungbuk		1(100.0)	1(2.3)		1(100.0)	1(2.3)		1(100.0)	1(2.3)
General H.	60(21.5)	219(78.5)	279(100.0)	124(44.4)	155(55.6)	279(100.0)	196(70.3)	83(29.7)	279(100.0)
Kangwon	4(28.6)	10(71.4)	14(5.0)	4(28.6)	10(71.4)	14(5.0)	11(78.6)	3(21.4)	14(5.0)
Gyeonggi	10(19.6)	41(80.4)	51(18.3)	24(47.1)	27(52.9)	51(18.3)	32(62.7)	19(37.3)	51(18.3)
Gyeongnam	9(42.9)	12(57.1)	21(7.5)	16(76.2)	5(23.8)	21(7.5)	19(90.5)	2(9.5)	21(7.5)
Gyeongbuk	2(10.5)	17(89.5)	19(6.8)	8(42.1)	11(57.9)	19(6.8)	16(84.2)	3(15.8)	19(6.8)
Gwangju	7(38.9)	11(61.1)	18(6.5)	16(88.9)	2(11.1)	18(6.5)	16(88.9)	2(11.1)	18(6.5)
Daegu	1(14.3)	6(85.7)	7(2.5)	1(14.3)	6(85.7)	7(2.5)	4(57.1)	3(42.9)	7(2.5)
Daejeon	1(12.5)	7(87.5)	8(2.9)	2(25.0)	6(75.0)	8(2.9)	4(50.0)	4(50.0)	8(2.9)
Busan	1(4.3)	22(95.7)	23(8.2)	2(8.7)	21(91.3)	23(8.2)	12(52.2)	11(47.8)	23(8.2)
Seoul	4(9.5)	38(90.5)	42(15.1)	14(33.3)	28(66.7)	42(15.1)	16(38.1)	26(61.9)	42(15.1)
Ulsan	2(50.0)	2(50.0)	4(1.4)	3(75.0)	1(25.0)	4(1.4)	3(75.0)	1(25.0)	4(1.4)
Incheon	3(21.4)	11(78.6)	14(5.0)	7(50.0)	7(50.0)	14(5.0)	11(78.6)	3(21.4)	14(5.0)
Jeonnam	11(50.0)	11(50.0)	22(7.9)	15(68.2)	7(31.8)	22(7.9)	21(95.5)	1(4.5)	22(7.9)
Jeonbuk	3(30.0)	7(70.0)	10(3.6)	6(60.0)	4(40.0)	10(3.6)	9(90.0)	1(10.0)	10(3.6)
Jeju	0(0.0)	6(100.0)	6(2.2)	0(0.0)	6(100.0)	6(2.2)	4(66.7)	2(33.3)	6(2.2)
Chungnam	1(10.0)	9(90.0)	10(3.6)	4(40.0)	6(60.0)	10(3.6)	10(100.0)	0(0.0)	10(3.6)
Chungbuk	1(10.0)	9(90.0)	10(3.6)	2(20.0)	8(80.0)	10(3.6)	8(80.0)	2(20.0)	10(3.6)
Total	60(18.6)	262(81.4)	322(0.0)	124(38.5)	198(61.5)	322(0.0)	196(60.9)	126(39.1)	322(0.0)

* Number of hospitals(row %), ** Number of hospitals(column %)

Fig. 1 General hospitals excluded from the initiative by dimensions and regions

* Participation requirement for the initiative: Hospitals must submit at least half of the indicators in Healthcare Quality & Patient Safety dimension.

2. 평가 영역별 평가지표의 변별력 검토

의료 질과 환자 안전 영역에서, 상급종합병원은 16개 지표 중 8개 지표가, 종합병원은 1개 지표가 변별력이 없는 것으로 분석되었다.

〈Table 2〉 Discriminating power of the indicators for Healthcare Quality & Patient Safety(as of 2016)

Measurement Indicators	Type	Tertiary Hospitals(n=43)								
		mean	(s.d.)	(1) median	Q1	Q3	(2) IQR	95pctl	(1)≥ 95pctl	(2)< 0.045
Hospital Accreditation	Structure	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Number of Specialists per Inpatient Day	Structure	0.33	0.07	0.31	0.45	0.27	0.38	0.11		
Number of Nurses per Inpatient Day	Structure	0.52	0.07	0.51	0.62	0.48	0.56	0.08		
Number of Dedicated Specialists per ICU Bed	Structure	0.46	0.22	0.38	0.88	0.25	0.63	0.38		
Number of Dedicated Nurses per ICU Bed	Structure	0.74	0.12	0.73	0.96	0.70	0.79	0.09		
Dedicated Personnel for Patient Safety	Structure	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Outpatient Visits per Physician per Day	Process	0.27	0.07	0.27	0.36	0.21	0.32	0.11		
Operation of Patient Safety Reporting System	Structure	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Operation of POA Coding System	Structure	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Participation in National Healthcare-associated Infections Surveillance System(KONIS)	Structure	0.98	0.15	1.00	1.00	1.00	1.00	0.00	*	+
Participation in Korean Antimicrobial Resistance Monitoring System(KARMS)	Structure	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Operation of Hospital Patient Visitor Policy	Structure	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Airborne Infection Isolation Rooms	Structure	0.81	0.39	1.00	1.00	1.00	1.00	0.00	*	+
Number of Critically Ill Children Inpatients	Process	0.19	0.24	0.11	0.83	0.07	0.16	0.09		
Percent of Inpatient with Rare Diseases	Process	0.36	0.14	0.35	0.60	0.29	0.39	0.11		
Percent of High-risk Pregnant Women Inpatients	Process	0.31	0.11	0.30	0.45	0.26	0.36	0.10		
Indicators with lower discrimination(N)									8	

Measurement Indicators	Type	General Hospitals(n=279)								
		mean	(s.d.)	(1) median	Q1	Q3	(2) IQR	95pctl	(1)≥ 95pctl	(2)< 0.045
Hospital Accreditation	Structure	0.40	0.49	0.00	1.00	0.00	1.00	1.00		
Number of Specialists per Inpatient Day	Structure	0.19	0.10	0.17	0.35	0.13	0.23	0.10		
Number of Nurses per Inpatient Day	Structure	0.36	0.17	0.39	0.57	0.28	0.45	0.17		
Number of Dedicated Specialists per ICU Bed	Structure	0.06	0.16	0.00	0.50	0.00	0.00	0.00		+
Number of Dedicated Nurses per ICU Bed	Structure	0.37	0.23	0.39	0.72	0.21	0.53	0.32		
Dedicated Personnel for Patient Safety	Structure	0.66	0.48	1.00	1.00	0.00	1.00	1.00	*	
Outpatient Visits per Physician per Day	Process	0.35	0.12	0.34	0.54	0.27	0.41	0.13		
Operation of Patient Safety Reporting System	Structure	0.65	0.48	1.00	1.00	0.00	1.00	1.00	*	
Operation of POA Coding System	Structure	0.99	0.10	1.00	1.00	1.00	1.00	0.00	*	+
Participation in National Healthcare-associated Infections Surveillance System(KONIS)	Structure	0.35	0.48	0.00	1.00	0.00	1.00	1.00		
Participation in Korean Antimicrobial Resistance Monitoring System(KARMS)	Structure	0.24	0.42	0.00	1.00	0.00	0.00	0.00		+
Operation of Hospital Patient Visitor Policy	Structure	0.66	0.47	1.00	1.00	0.00	1.00	1.00	*	
Airborne Infection Isolation Rooms	Structure	0.27	0.44	0.00	1.00	0.00	1.00	1.00		
Number of Critically Ill Children Inpatients	Process	0.01	0.02	0.00	0.06	0.00	0.00	0.00		+
Percent of Inpatient with Rare Diseases	Process	0.16	0.12	0.13	0.37	0.08	0.21	0.14		
Percent of High-risk Pregnant Women Inpatients	Process	0.08	0.16	0.00	0.38	0.00	0.09	0.09		
Indicators with lower discrimination(N)									1	

Note 1) Each indicator was standardized to have a range of 0 to 1, then discriminating power of the value was measured.

2) *: (1) median≥95pctl

3) +: (2) IQR<0.045

공공성 영역에서는 상급종합병원만 5개 지표 중 2개 지표가 변별력이 없었다(Table 3). 의료전달체계 영역에서도 상급종합병원만 6개 지표 중 한 개 지표가 변별력이 없었다(Table 4).

〈Table 3〉 Discriminating power of the indicators for Publicness(as of 2016)

Measurement Indicators	Type	Tertiary Hospitals(n=43)								
		mean	(s.d.)	(1) median	Q1	Q3	(2) IQR	95pctl	(1)≥ 95pctl	(2)< 0.045
Percent of Medicaid Outpatients with Mild Diseases	Process	0.57	0.09	0.58	0.72	0.52	0.61	0.09		
Operation of Delivery Room	Structure	0.10	0.03	0.11	0.15	0.09	0.13	0.04		+
Number of NICU Inpatients	Process	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Disposition of NICU Neonatologist	Structure	0.33	0.23	0.30	0.75	0.16	0.47	0.31		
Number of Dedicated Nurses per ICU Bed	Structure	0.98	0.15	1.00	1.00	1.00	1.00	0.00	*	+
Indicators with lower discrimination(N)									2	

Measurement Indicators	Type	General Hospitals(n=279)								
		mean	(s.d.)	(1) median	Q1	Q3	(2) IQR	95pctl	(1)≥ 95pctl	(2)< 0.045
Percent of Medicaid Outpatients with Mild Diseases	Process	0.14	0.17	0.06	0.48	0.02	0.20	0.18		
Operation of Delivery Room	Structure	0.36	0.16	0.34	0.63	0.24	0.45	0.21		
Number of NICU Inpatients	Process	0.52	0.50	1.00	1.00	0.00	1.00	1.00	*	
Disposition of NICU Neonatologist	Structure	0.03	0.08	0.00	0.19	0.00	0.00	0.00		+
Number of Dedicated Nurses per ICU Bed	Structure	0.11	0.31	0.00	1.00	0.00	0.00	0.00		+
Indicators with lower discrimination(N)									-	

Note 1) Each indicator was standardized to have a range of 0 to 1, then discriminating power of the value was measured.

2) *: (1) median≥95pctl

3) +: (2) IQR<0.045

〈Table 4〉 Discriminating power of the indicators for Healthcare Delivery System(as of 2016)

Measurement Indicators	Type	Tertiary Hospitals(n=43)								
		mean	(s.d.)	(1) median	Q1	Q3	(2) IQR	95pctl	(1)≥ 95pctl	(2)< 0.045
Percent of Medicaid Outpatients with Mild Diseases	Process	0.57	0.09	0.58	0.72	0.52	0.61	0.09		
Operation of Delivery Room	Structure	0.10	0.03	0.11	0.15	0.09	0.13	0.04		+
Number of NICU Inpatients	Process	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Disposition of NICU Neonatologist	Structure	0.33	0.23	0.30	0.75	0.16	0.47	0.31		
Number of Dedicated Nurses per ICU Bed	Structure	0.98	0.15	1.00	1.00	1.00	1.00	0.00	*	+
Indicators with lower discrimination(N)									1	

Measurement Indicators	Type	Tertiary Hospitals(n=43)								
		mean	(s.d.)	(1) median	Q1	Q3	(2) IQR	95pctl	(1)≥ 95pctl	(2)< 0.045
Percent of Medicaid Outpatients with Mild Diseases	Process	0.57	0.09	0.58	0.72	0.52	0.61	0.09		
Operation of Delivery Room	Structure	0.10	0.03	0.11	0.15	0.09	0.13	0.04		+
Number of NICU Inpatients	Process	1.00	0.00	1.00	1.00	1.00	1.00	0.00	*	+
Disposition of NICU Neonatologist	Structure	0.33	0.23	0.30	0.75	0.16	0.47	0.31		
Number of Dedicated Nurses per ICU Bed	Structure	0.98	0.15	1.00	1.00	1.00	1.00	0.00	*	+
Indicators with lower discrimination(N)									-	

Note 1) Each indicator was standardized to have a range of 0 to 1, then discriminating power of the value was measured.

2) *: (1) median≥95pctl

3) +: (2) IQR<0.045

고찰

의료질평가지원금 평가지표는 구조와 과정지표 위주로 구성되어 있으며, 영역별로 변별력이 없는 구조 및 과정지표들이 확인되었다. 이 연구에서 검토한 지표의 변별력은 의료질평가지원금의 지표 선정 기준인 영향(impact), 측정가능성(feasibility), 개선가능성(improvability), 수용성(acceptability) 중 측정가능성과 직접 관련된 조건이지만, 개선가능성 등 기타 기준에도 영향을 미치는 핵심 요소이다. 특히, 상급종합병원에서 변별력이 없는 지표가 많고 해당 지표들은 100% 달성률을 보이고 있다. 구조지표 중심의 지표 구성은 제도적으로 규정된 상급종합병원과 종합병원의 구조적 차이를 그대로 반영하고 있는 것으로 보인다. 지표는 목표를 달성할 수 있는 여지를 남겨두어야 피 평가대상인 의료기관에게 의료의 질 향상에 대한 동기를 강화시킨다[7]. 지표의 변별력 평가는 극단적인 달성수준을 보여주는 지표들을 구분해 냄으로써 지표를 통한 의료질의 개선 가능성을 높이기 위함이다. 특히, 절대평가가 아닌 상대평가로 중층화된 등급 구조에서, 변별력 없는 지표의 사용은 제도의 신뢰도를 낮추고 의료기관의 반응성을 낮추게 된다. 이러한 이유로 지표의 변별력은 바람직한 질 지표가 갖추어야 할 공통의 기준이다[14].

의료서비스의 질을 측정하고 모니터링 함에 있어서 임상 질 지표는 통계적으로, 임상적으로 모두 의미가 있어야 한다. Mainz[14]가 제시한 임상 질 지표 개발의 6단계는 지표의 선택과 제외 결정에 있어서 이러한 의미를 반영하기 위한 구체적인 요소들을 포함하고 있다. 첫 단계에서 평가될 임상영역을 선정한다. 높은 이용량(high volume), 비용(cost), 변이(variation)를 고려하여 중요하고 임상적 개입의 여지가 높은 영역을 확인해야 한다. 두 번째 단계에서는 측정팀을 구성해야 한다. 측정을 위한 역할을 규정하고 작업 분담을 해야 한다. 셋째 단계는 기존 근거와 사례를 검토하는 것이다. 잠재적 지표에 대한 과학적 문헌 검토를 통해 지식을 구축하고 기존의 지식과 실천에 대해 합의를 도출해야 한다. 네 번째 단계는 임상 지표와 표준(standards)을 선택하는 것이다. 과정과 결과지표를 선택하고 위험보정을 위한 예후 요인들을 확인하여 합의와 평가 절차를 수립해야 한다. 결과지표를 평가할 때는 환자구성의 차이와 기타 공변량 요인들을 통제하는 수준에서 적정성이 중요하다. 한편, 임상지표는 과학적 근거를 바탕으로 해야 하지만 과학적 근거를 찾기 어려운 경우는 의료 전문가 합의를 기반으로 선택할 수 있다. 다섯 번째 단계는 측정 방식을 설계하는 것이다. 지표 및 표준의 정의, 표적 인구집단 설정, 포함 및 제외기준 규정, 위험보정 방식의 설정, 자료원 파악, 데이터 수집 절차의 기술, 분석 계획 수립이 포함된다. 마지막 단계는 시범 평가(pilot test)를 해보는 것이다. 이와 대조하면, 현재의 의료질평가지원금 지표는 의료의 질 개선을 유도하는 통계적, 임상적 절차에 근거하지 못한 한계를 드러낸다.

통계적 검토를 위한 기준은 인센티브의 경제성(economics)과 지표의 성과(performance)에 대해 모두 고려해야 한다. 경제적 분석은 건강 편익과 관련 지표의 비용을 계량화함으로써 인센티브의 편익이 비용을 초과하면 좋은 성과에 대한 보상을 지속할 경제적 근거를 얻는 것이다[15, 16]. 하지만 이러한 분석은 최종의 결과적 건강편익을 추정하기 어려운 과정지표에 대해서는 적용이 어렵다. 따라서 과정지표에 대한 통계적 검토는 주로 지표별 성취율의 수준과 변이, 보고 예외율의 수준과 변이를 고려해야 한다. 지표가 어느 한쪽으로 치우친 분포에서 성취도 평가는 중위수와 사분범위(IQR)가 평균과 표준편차보다 적절한 척도가 된다[15]. 지표의 성취도는 여러 가지 요인들의 영향을 받는다. 공급자가 통제할 수 없는 요인들에 의해 평가 등급이 결정될수록 의료기관 간 성취도의 변이는 적어진다. 반면, 의료기관 간 성취율에서 변이가 클수록 해당 지표의 평가는 더 많은 의료기관의 성과 향상을 가져올 것이다. 아울러, 좋은 지표는 예외 보고율의 수준과 변이가 모두 낮아야 한다[15].

영국의 QOF(Quality and Outcomes Framework)는 일차의료의사에 대한 질 평가 및 인센티브 프로그램으로 2010년 당시 134개 지표로 총점 1000점에 이르는 구조를 설계했다. 2011년 8개의 임상지표가 제외되었는데, 해당 지표들이 모두 검사 시행에 대한 과정지표로서 달성 수준이 높았고(중위수>95percentile, 사분위 범위<4.5%), 3회 연속(2005-6년, 2006-7년, 2007-8년) 변동이 거의 없었다는 것이 근거였다[15]. 2015년 의료질평가지원금 지표 선정을 위한 연구에서도 이 기준을 적용하여 기존의 적정성 평가에서 사용하던 임상지표의 통계적 변별력을 검토한 바 있다[7].

평가지표의 변별력은 구조와 과정 중심의 구성에서 특히 취약할 수 있다. 구조와 과정지표는 상대적으로 손쉽게 충족될 수 있기 때문에 반복적으로 사용할 경우 변별력이 떨어지게 된다. 이러한 이유로 구조와 과정지표보다 결과지표 중심으로 구조를 개편하고 과정지표의 가중치를 상대적으로 낮추며 지속적으로 변별력을 높이기 위해 새로운 지표 도입을 검토해야 한다[8,10]. 또한 등급제외기관들의 지역별 편차는 제도의 효과에 있어서 지속적으로 불평등 문제를 초래할 수 있다. 궁극적으로 지역 간 질의 격차가 발생하지 않도록 이들 지역 의료기관들의 질 향상 노력을 강화시켜 참여를 확대시키는 평가방식의 개편을 검토해야 한다. 다른 한편으로, 광주, 전남, 울산지역의 높은 등급제외율에 대한 정밀한 원인 분석을 통해 결과적으로 해당 서비스에 대한 지역 주민의 미충족 의료가 발생하지 않도록 공공의료 확대 등 적절한 대응책 마련도 검토되어야 할 것이다.

물론, 제도 도입 당시의 제한점을 고려해야 한다. 도입 당시에는 선택진료비가 주로 발생하던 상급종합병원의 손실 보전을 위해 상급종합병원과 종합병원의 시스템적 차이를 반영하는 구조지표의 활용은 불가피했던 측면이 있다. 아울러, 상대적으로 새로운 지표보다 안정적으로 지정기준에서 합의된 구조지표의 사용이 용이했을 수 있다[7]. 그러나 향후 제도의 개선은 본래 목적에 충실하도록 지표 구성을 개편해야 한다.

한편, 보건복지부는 현재의 상대평가 방식을 절대평가 방식으로 전환하는 계획을 추진 중에 있다. 상대평가 방식에서는 상대적 순위에 영향을 미치는 지표의 변별력이 중요하지만 절대평가 방식에서는 통계적 변별력 보다 지표 자체의 임상적 의미에 대한 많은 관심을 더 갖게 될 것이다. 하지만 두 경우 모두 평가지표의 통계적 의미를 갖추기 위해 지표의 선정과 제외 결정 시 지표의 달성수준과 변이를 평가하는 것은 필요하다. 지표를 제외시키는 경우에도 제외 이후의 변화를 사전에 검토하여 결정해야 하고 제외 이후 성과 수준이 감소하지 않는지 반드시 모니터링이 필요하다[15, 16]. 영국의 QOF에서도 관련 결과지표가 있으면서 지표를 제외한 경우는 제외된 지표의 성과 감소가 없었지만, 관련 결과지표가 없었던 경우에는 제외 이후 해당 지표의 수준이 감소하는 결과를 보였다[15, 17]. 결과적으로 지표 달성의 수준과 변이에 대해서는 지속적으로 모니터링이 필요하다[11].

향후 지표 개선을 위해서는 변별력이 없는 구조 및 과정지표를 제외하고 결과지표 중심으로 지표 구성을 개편해야 한다. 이러한 변화를 촉진하기 위해서는 질 지표가 체계적(systematic)이고 참여적(participatory)이며 근거 기반 방식(evidence based manner)으로 개발되고 관리되어야 한다[11]. 첫째, 참여적 접근으로 공급자 및 수요자(환자), 건강보험공단 및 심사평가원, 학계 전문가 등 다양한 이해관계자가 참여하는 사회적 합의 기전을 통해 지표 선정 기준, 평가 지표 수, 지표의 진·출입 기준, 구조/과정/결과지표의 비율, 지표별 가중치 설정 기준 등 관련 원칙의 결정이 가능해야 한다. 둘째, 근거기반의 접근으로 전문가 및 학회와 동반 관계(partnership) 구축을 통해 지표가 개발-인증-갱신되는 환류체계를 마련해야 한다.

이 연구는 몇 가지 제한점을 가지고 있다. 첫째, 2016년 단년도의 평가결과 자료만을 활용하여 분석하였다. 지표별 달성률 수준과 변이가 지속되는지 확인하기 위해서는 다년도 자료 분석이 필요하지만 축적된 자료의 부족으로 시행이 어려웠다. 그러나 구조와 과정 지표 중심의 지표 구성 문제를 확인하는 의미 있는 근거 정보를 제공할 것이다. 둘째, 의료

기관의 특성 정보 부족으로 세부적인 유형별 특성 분석을 시행하지 못했다. 향후 추가적인 분석이 필요하며, 등급제외율이 높은 지역에 대한 세부 분석도 향후 과제로 제안한다.

참고문헌

1. 강희정. 2017년 건강보험정책방향: 혁신을 통한 형평과 효율의 조화. 보건복지포럼 2017; 1:18-28.
2. Bull World Health Organ 2016;94:160-160A. doi http://dx.doi.org/10.2471/BLT.16.170605.
3. Arah OA, Westert GP, Hurst J, Klazinga NS. A Conceptual Framework for the OECD Health Care Quality Indicators Project. International Journal for Quality in Health Care. 2006;18(Suppl.1):5-13.
4. Brownwood I. Hospital performance-stage 1 report(DELSA/HEA/HQC(2015)1). Paris:OECD, 2015.
5. 박춘선, 최효정, 황수희, 임지혜, 김경훈, 김선민. OECD 병원 성과 프로젝트의 동향과 국내 시사점. 한국의료질향상학회지. 2016;22(1):12-27.
6. 황수희, 오동관, 김묘정, 박춘선. 우리나라 공공병원과 민간병원의 의료의 질. 대한공공의학회지 2017;1:145-156.
7. 신영석, 강희정, 신현웅, 김윤, 조민우, 윤영호, 신애선, 김은아, 강아람, 이채은, 옥민수, 우형택, 심진아. 선택진료료 개편에 따른 의료질향상 분담금 제도 시행방안 개발연구, 건강보험심사평가원, 한국보건사회연구원 2015.
8. Doran T, Maurer KA, Ryan AM. Impact of provider incentives on quality and value of health care. Annu Rev Public Health. 2017 Mar 20;38:449-465. doi: 10.1146/annurev-publhealth-032315-021457. Epub 2016 Dec 15.
9. Damberg CL, Sorbero ME, Lovejoy SL, Martsoff G, Raaen L, Mandel D. Measuring Success in Health Care Value-Based Purchasing Programs. RAND corporation 2014. Available at www.rand.org.
10. 강희정, 신영석, 이광수, 조민우, 하솔원 외. 의료질평가지원금 평가 중장기 모형 개발: 의료전달체계 및 공공성 영역 중심. 2017. 건강보험심사평가원, 한국보건사회연구원.
11. Khampang R, Teerawattananon Y, Tanivess S, Cluzeau F, Foskett-Tharby R, Gill P. Developing and testing quality indicators for the Thai Quality and Outcomes Framework. Safety in Health 2017;3(14). DOI 10.1186/s40886-017-0065-6.
12. 보건복지부고시 제2018-69호. 의료질평가지원금 산정을 위한 기준 일부개정. 2018.3.30.
13. 보건복지부 공고 제2018-217호. 2018 의료질평가계획. 2018.3.30.
14. Mainz J. Developing evidence-based clinical indicators: a state of the art methods primer. International Journal of Quality in Health Care, 2003; 15(Suppl 1):i5-i11.
15. Reeves D, Doran T, Valderas JM, Kontopantelis E, Trueman P, Sutton M, Campbell S, Lester H. How to identify when a performance indicator has run its course. BMJ. 2010 Apr 6;340:c1717. doi: 10.1136/bmj.c1717.
16. 김윤, 도영경, 조민우, 정혜민, 노준수 외. 요양급여 적정성 평가의 정책적 활용 극대화를 위한 이론적 근거와 개념 틀 개발 연구. 건강보험심사평가원, 서울대학교 산학협력단. 2014.
17. Kontopantelis E, Springate D, Reeves D, Ashcroft DM, Valderas JM, Doran T. Withdrawing performance indicators: retrospective analysis of general practice performance under UK Quality and Outcomes Framework. BMJ. 2014 Jan 27;348:g330. doi: 10.1136/bmj.g330.