

발 간 등 록 번 호

G000DY2-2022-143

# 평가지표 개발 가이드라인 마련 연구

---

## 연구진

연구책임자 이후연 교수

공동연구자 조경희 교수

윤난희 교수

이미선 교수

연구보조원 김지은

박소연

김동준

김성주

권남희

보조원 임지환

---



건강보험심사평가원

HEALTH INSURANCE REVIEW & ASSESSMENT SERVICE





---

## 제1장 서론 1

---

1. 연구 배경 및 목적 .....	3
가. 영양급여 적정성 평가 .....	3
나. 연구의 배경과 필요성 .....	7
다. 연구 목적 .....	7
2. 연구 내용 및 방법 .....	8
가. 연구 내용 .....	8
나. 연구 방법 .....	9

---

## 제2장 국내외 평가지표 개발 현황 11

---

1. 미국의 Center for Medicare & Medicaid(CMS) .....	13
2. 미국의 Agency for Healthcare Research and Quality's(AHRQ) .....	38
3. 미국의 National Quality Forum(NQF) .....	47
4. 영국의 Shared commitment to quality .....	55
5. 영국의 National Institute for Health and Care Excellence(NICE) .....	58
6. 한국 국가건강검진 검진항목 선정 .....	63

---

## 제3장 평가지표 개발의 문제점 및 개선사항 67

---

1. 의견 수렴 과정 .....	69
2. 주요 내용 .....	70
3. 개선 방안 .....	70

## 제4장 평가지표 개발 가이드라인

73

1. 의료의 질을 이해하기 위한 기본 개념 .....	75
가. 바람직한 보건의료의 특성 .....	75
나. 의료의 질 정의와 개념 .....	77
다. 의료의 질 향상 접근방안 .....	78
라. 의료환경의 변화와 의료의 질 평가 .....	81
2. 평가지표 개발의 기본원칙 .....	81
가. 평가지표의 개념과 유형 .....	81
나. 개발한 평가지표에 대한 평가 기준 .....	83
3. 평가지표 개발의 단계별 지침 .....	84
가. 평가지표 개발 계획 수립 .....	85
나. 후보지표 개발 .....	95
다. 후보지표 검증 .....	107
라. 후보지표 평가 .....	113
마. 최종 평가지표 선정 .....	119

## 제5장 제언

125

1. 국내외 평가지표 개발 현황을 통한 시사점 .....	127
가. 평가지표 개발 우선순위 설정을 위한 시사점 .....	127
나. 평가지표 개발 가이드라인 개발을 위한 시사점 .....	128
2. 국가 차원의 질 향상 정책목표와 우선순위 영역 도출 .....	129
3. 적정성 평가 영역의 주요 우선순위 도출 .....	134
가. 국가수준의 의료의 질 격차를 확인하고 평가영역의 우선순위 결정 .....	134
나. 평가 영역별 평가지표의 균형적 분포 .....	135
■ 참고 문헌 .....	139



## 표 목차

〈표 1〉 요양급여 적정성 평가 분야별 평가항목 .....	3
〈표 2〉 미국 NQF 2020년 보고서 주요 결과 .....	25
〈표 3〉 CMS의 질 평가 유형 .....	32
〈표 4〉 CMS의 질 평가 활동 .....	36
〈표 5〉 평가의 관점과 평가지표 분류 .....	40
〈표 6〉 NQF 평가지표 검토 기준 .....	42
〈표 7〉 AHRQ 평가지표 개발 주기 단계 및 소요시간 .....	45
〈표 8〉 타당성 평가에 사용되는 기준 .....	62
〈표 9〉 국가건강검진 항목 선정의 기본원칙, 평가내용, 근거자료 .....	65
〈표 10〉 평가지표 개발 가이드라인 마련 연구 실무자 인터뷰 대상 명단 및 일정 .....	69
〈표 11〉 의료의 질 구성요소 .....	75
〈표 12〉 평가지표 유형 ① .....	80
〈표 13〉 평가지표 유형 ② .....	82
〈표 14〉 평가지표 유형 ③ .....	86
〈표 15〉 문제 확인 예시 .....	87
〈표 16〉 평가 대상 선정 예시 .....	91
〈표 17〉 결과지표의 예시 .....	97
〈표 18〉 전문가 구성(안)의 장·단점 .....	98
〈표 19〉 평가자료원 예시 .....	99
〈표 20〉 핵심 결과지표 개발 예시 .....	100
〈표 21〉 권고의 등급화 기준 .....	102
〈표 22〉 근거수준의 등급 .....	103
〈표 23〉 SIGN의 권고등급체계 (2015년 이전) .....	104
〈표 24〉 SIGN의 권고등급체계 추가 설명표 .....	104
〈표 25〉 구조와 과정 후보지표 개발 예시 .....	105
〈표 26〉 후보지표로 사용 가능한 지표의 유형 .....	106
〈표 27〉 1차 검증과 2차 검증의 특징 비교 .....	109
〈표 28〉 1차 검증계획서 예시 .....	109
〈표 29〉 2차 검증계획서 예시 .....	110
〈표 30〉 필요성 평가 기준 .....	116
〈표 31〉 동의여부 판정 기준 .....	116
〈표 32〉 패널 규모에 따른 동의여부 판정 기준 .....	117
〈표 33〉 적합여부 평가 기준 .....	117

〈표 34〉 후보지표에 대한 전문가 합의과정 예시 .....	117
〈표 35〉 평가지표 개발 check list .....	119
〈표 36〉 표준화된 양식에 따라 선정된 지표 요약정리 .....	122
〈표 37〉 요양급여 적정성평가의 정책 목표 .....	129
〈표 38〉 국내 다양한 평가제도의 내용과 주관기관 .....	132



## 그림 목차

[그림 1] 적정성 평가 평가체계도 .....	4
[그림 2] 적정성 평가지표 개발 및 운영과정 .....	5
[그림 3] 주요 연구 내용 .....	8
[그림 4] 연구 방법 .....	10
[그림 5] CMS Quality Strategy Aims .....	14
[그림 6] CMS Quality Strategy Goals .....	16
[그림 7] 2022년 MDP의 주요 내용 요약 .....	20
[그림 8] 환경 평가 및 격차 분석 과정 .....	22
[그림 9] 환경 분석 과정과 결과 .....	24
[그림 10] 의미 있는 평가 지표 2.0 구성 (Meaningful Measures 2.0 Framework) .....	28
[그림 11] CMS 목표 달성을 위한 측정지표 개발 .....	29
[그림 12] 환자안전을 위한 측정 단계 (Cascade of Measures) 활용의 예 .....	30
[그림 13] CMS 목표 달성을 위한 측정지표 개발 .....	31
[그림 14] 평가지표의 가치 분석을 위한 5가지 항목 기준 .....	34
[그림 15] CMS 평가 측정의 생애주기 .....	34
[그림 16] 평가 지표 개발 단계에서 이해관계자의 참여와 역할 ① .....	37
[그림 17] 평가 지표 개발 단계에서 이해관계자의 참여와 역할 ② .....	38
[그림 18] AHRQ 질 평가의 생애주기 .....	41
[그림 19] NQF의 평가지표 검토 주기 일정 .....	49
[그림 20] NQF의 평가지표 검토 주제 영역 .....	50
[그림 21] NQF의 제안된 합의 개발 프로세스 .....	51
[그림 22] NQF의 평가지표 유형별 업무흐름도 .....	53
[그림 23] 의료의 질에 대한 관점의 공유 (A single shared view of quality) .....	56
[그림 24] 의료의 질 향상을 위한 7가지 단계 .....	57
[그림 25] 지표 검증 방법에 대한 예시 .....	61
[그림 26] 국가건강검진 검진(검사)항목 선정 및 평가 담당 관련 조직 .....	64
[그림 27] 평가지표 개발 프로세스 .....	84
[그림 28] 평가대상선정에 대한 모식도 .....	90
[그림 29] 평가대상자 선정 흐름도 예시 .....	94
[그림 30] 인공관절치환술 30일 이내 재입원의 위험도 보정에 사용된 변수 예시 .....	101
[그림 31] 제1차 국민건강보험 종합계획(2019-2023) 중 의료 질 평가 및 보상 내실화 .....	131
[그림 32] 평가발전위원회의 평가체계 혁신방향 및 7개 권고(안) .....	133
[그림 33] The Measure Applications Partnership (MAP) 구성 .....	136





# 서론



## 제1장 서론

### 1. 연구 배경 및 목적

#### 가. 요양급여 적정성 평가

- 요양급여 적정성 평가는 건강보험으로 제공된 진찰·시술·투약·검사 등의 의료서비스가 의약학적 측면과 비용 효과적 측면에서 적정하게 제공되었는지를 평가함
- 요양급여 적정성 평가의 목적은 다음과 같음
  - 의료제공자에게 평가 결과를 통보하여, 의료기관의 자율적인 질 향상 및 진료행태 개선을 유도하고 국민이 질 높은 의료서비스를 받을 수 있도록 하고자 함
  - 평가 결과 공개를 통해 국민의 의료선택권을 보장하고, 의료서비스 질 향상과 비용 부담 적정화를 도모하고자 함
- 2021년 6월 기준 8개 분야 35개 항목의 323개 평가지표가 있음 (〈표 1〉 참고)

〈표 1〉 요양급여 적정성 평가 분야별 평가항목

구분	분야(항목 수)	세부 항목(52개)
1	환자 중심(1)	• 환자경험
2	급성질환(5)	• 관상동맥우회술 • 폐렴 • 경피적관상동맥중재술 • 급성기뇌졸중 • 급성심근경색증
3	만성질환(4)	• 고혈압 • 천식 • 당뇨병 • 만성폐쇄성폐질환
4	암 질환(5)	• 대장암 • 폐암 • 간암 진료결과 • 유방암 • 위암
5	감염질환(1)	• 결핵
6	정신건강영역(4)	• 의료급여정신과 • 우울증 외래 • 정신건강 입원영역 • 치매

7	진료행위 및 약제(9)	<ul style="list-style-type: none"> <li>• 항생제처방률</li> <li>• 처방건당 약품목수</li> <li>• 수술의 예방적 항생제 사용(18개 수술)</li> <li>• 마취</li> <li>• 수혈</li> </ul>	<ul style="list-style-type: none"> <li>• 주사제 처방률</li> <li>• 투약일당 약품비</li> <li>• 혈액투석</li> <li>• 치과 근관치료</li> </ul>
8	기관단위(6)	<ul style="list-style-type: none"> <li>• 요양병원</li> <li>• 신생아중환자실</li> <li>• 위험도표준화재입원비</li> </ul>	<ul style="list-style-type: none"> <li>• 중환자실</li> <li>• 병원표준화사망비</li> <li>• 중소병원</li> </ul>

\* '19년 '질병군 포괄수가(7개 질병군)' 평가는 종료, '유소아중이염 항생제' 평가는 약제 평가로 통합됨

○ 적정성 평가 과정은 [그림 1]과 같이 기획, 수행, 활용 단계로 구분됨

- 기획 단계 : 평가대상 후보 항목 선정, 예비 평가, 평가계획 수립 등
- 수행 단계 : 평가자료 구축, 평가자료 분석, 평가결과 산출 등
- 활용 단계 : 국민홍보, 질 향상 지원, 정부 정책에 활용 등



[그림 1] 적정성 평가 평가체계도

○ 기획 단계의 구체적인 평가지표 개발 과정은 다음과 같음 ([그림 2] 참고)



[그림 2] 적정성 평가지표 개발 및 운영과정

## 1) 평가대상 후보항목 선정

- 상시 평가 제안시스템을 통해 임상학회, 의약계단체, 소비자단체, 학계, 복지부 등 정부 부처 및 심사평가원 등에서 신규 평가 항목을 제안
- 신규 및 기존 평가 항목을 비교 검토하고, 우선순위 선정 기준(빈도 및 비중, 의약학적 중요성, 사회적 관심, 개선 효과, 평가 용이성)에 대해 전문가의 의견수렴
- 의료평가조정위원회 심의를 거쳐 최종 신규 평가항목 우선순위 선정

## 2) 지표개발

- 예비평가 지표 개발을 전문 학회 등에 위탁하여 외부 전문가 중심으로 개발
- 해당 분야 외부 전문가단체의 평가지표 개발을 통해 평가의 객관성 및 수용성을 제고하고자 함

### 3) 예비평가

- 평가지표의 타당성 및 평가 실행 가능성을 검토하는 것이 목적임
- 연구에서 개발한 평가지표, 외국의 평가지표, 타 평가제도의 유사 지표 등을 비교 검토
  - 평가지표 구성(구조·과정·결과)의 적절성 확보
  - 평가지표의 타당성 검토 후 평가지표 통합·삭제·추가
  - 평가지표별 세부기준 마련
  - 청구자료, 의무기록 조사 자료 등을 이용한 모의 분석

### 4) 연간평가계획 수립

- 적정성 평가 관련 연간 사업은 매년도 말에 다음 해에 추진할 평가 계획을 수립하여 보건복지부 장관의 승인절차를 거침
- 평가대상 항목과 목적, 대상기관 · 기간, 평가기준, 평가결과의 활용 등 포함

### 5) 세부시행계획 수립

- 평가 항목별로 대상기관 · 기간, 평가 지표, 평가 방법, 자료조사 방법 등을 포함하여 구체적으로 수립
- 평가 계획, 평가 기준, 평가 등급 등에 관한 사항은 분과위원회에서 전문가 의견을 수렴한 후 의료평가조정위원회 심의
- 수립된 세부추진계획을 평가 실시 2개월 전에 요양기관 등에 사전 공개

## 나. 연구의 배경과 필요성

- 2013년부터 외부 전문가단체 위탁 연구를 통해 평가지표 개발이 이루어지고 있음
  - 이해당사자인 의료전문가의 참여를 통해 평가의 투명성 및 수용성을 높일 수 있고, 평가지표 개발 초기부터 임상 전문가의 참여를 통해 향후 지표 적용 시 발생할 수 있는 장애요인과 성공요인을 심층적으로 파악할 수 있는 장점이 있음
  - 한편, 구조, 과정, 결과 영역별로 균형 있는 평가지표를 개발하는데 어려움이 있음. 결과지표 확대의 필요성에도 불구하고 자료수집 및 중증도 보정의 어려움과 결과 지표 평가에 대한 부담으로 결과지표 개발에 소극적인 경우가 있음
  - 또한 개발된 평가지표가 학회의 관심 분야에 집중되는 경우가 있고, 표준화된 지침 부재로 개발의 어려움과 개별 학회 간 개발 절차가 일관되지 않음
- 최근에는 신규 평가 제안시스템을 통한 평가지표 개발은 후보 항목의 제안부터 도입 까지 걸리는 기간이 약 3~5년 정도로 장시간이 소요됨에 따라, 적시성 있는 신규 항목 도입을 위한 개선방안이 요구됨
- 따라서 우선순위가 높은 영역의 핵심적인 지표가 균형적으로 개발될 수 있도록 하면서, 항목별 특수성이 반영되는 체계적 개발 절차를 마련하는 것이 필요함

## 다. 연구 목적

- 이 연구는 신규 평가지표의 신뢰도와 수용성을 높이고, 개발 과정의 효율성을 높이고자 함. 구체적인 목적은 다음과 같음
  - 적정성 평가지표 개발 가이드라인 개발을 통해 평가지표 개발 절차를 표준화함
  - 평가지표 개발 초기부터 다양한 전문가와 이해관계자의 의견을 적극적으로 수렴하여, 평가지표의 신뢰도와 타당도를 향상시킴
  - 구조, 과정, 결과지표의 균형적인 개발을 촉진할 수 있는 가이드라인을 개발함

## 2. 연구 내용 및 방법

### 가. 연구 내용

- 이 연구의 주요 목적은 크게 세 가지로 ‘요양급여 적정성 평가지표 개발 절차의 표준화’, ‘요양급여 적정성 평가지표 개발 투명성 강화’, ‘요양급여 적정성 평가지표 개발의 시의성 확보’임. 이를 위한 세부적인 연구 내용은 다음 [그림 3]과 같음
- 첫째, 국내외 국가 수준의 질 지표 개발·관리 체계 검토를 통해 요양급여 적정성 평가 지표 개발에서 활용할 수 있는 시사점을 도출함
- 둘째, 평가지표 개발 과정의 표준화를 위한 가이드라인을 개발하고, 환자를 포함한 다양한 이해관계자와 의료전문가의 참여를 통해 지표 개발의 전문성·투명성·활용성을 높이는 방안을 마련함
- 셋째, 현재 구조와 과정 중심 평가지표 개발에서 결과지표를 중심으로 지표의 개발이 이루어지도록 가이드라인을 개발함. 이를 통해 향후 구조, 과정, 결과 평가지표가 균형 있게 개발될 수 있도록 함

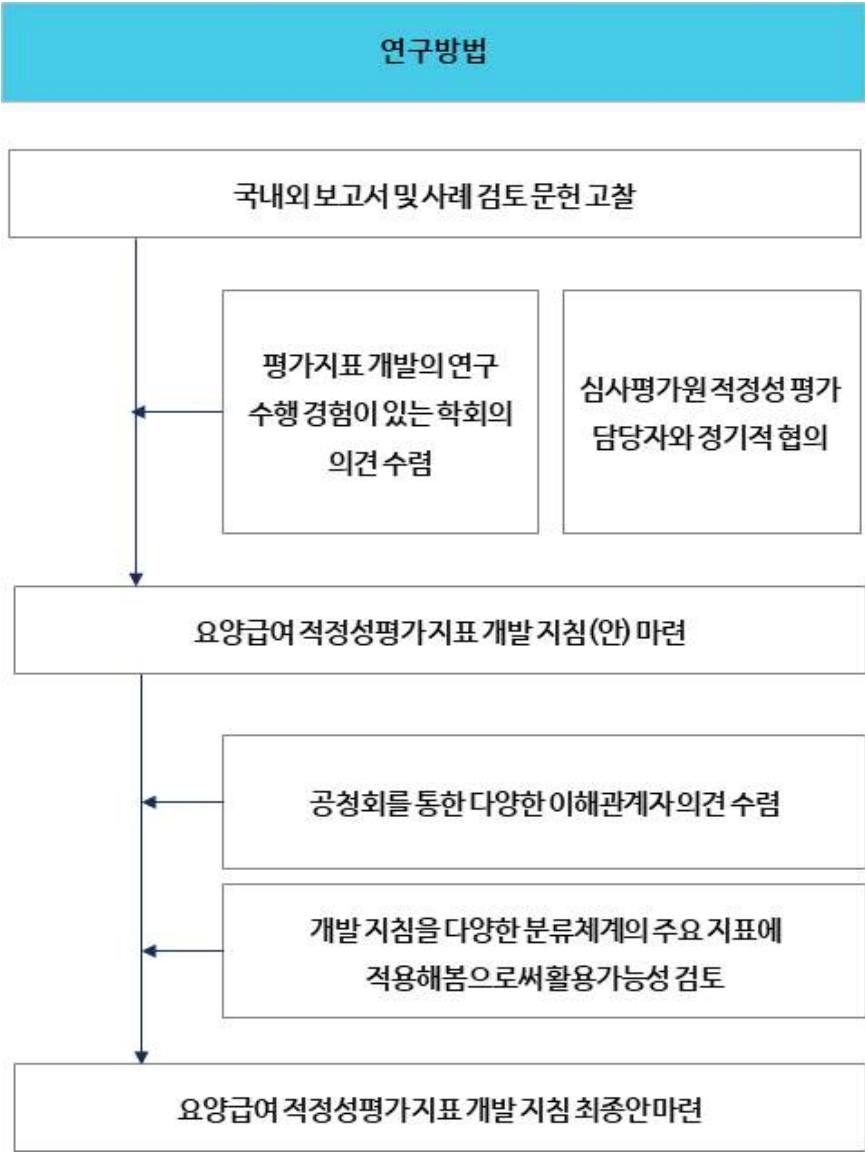
평가 지표 개발 가이드라인 개발	이해관계자와 전문가의 참여 확대	구조,과정,결과 지표의 균형적 개발
<ul style="list-style-type: none"> <li>• 국내외 평가지표 개발 절차 검토</li> <li>• 평가지표 개발 경험에 있는 임상, 평가 실무자를 대상으로 문제점과 개선방안에 대한 의견 수렴</li> <li>• 단계별 개발 내용, 주체, 기간을 명시</li> </ul>	<ul style="list-style-type: none"> <li>• 다양한 이해관계자의 참여를 체계화</li> <li>• 평가지표 개발 초기 단계부터 참여</li> <li>• 결과지표 산출의 타당도와 신뢰도를 높이기 위한 외부 또는 내부 전문가를 적극 활용</li> </ul>	<ul style="list-style-type: none"> <li>• 결과(최종,중간)지표 필수(예외 여부는 위원회에서 논의) =&gt; 결과지표</li> <li>• 결과와 관련이 있는 근거가 있는 경우에 한해 구조와 과정 포함 =&gt; 구조와 과정 지표</li> <li>• 중증도 보정과 통계분석에 대한 지원(교육, 내외부 전문가 참여)을 통해 결과지표의 신뢰도와 타당도 향상</li> </ul>
요양급여 적정성 평가지표 개발 지침을 주요 지표에 적용해보고 개발 지침의 적용가능성 검토		

[그림 3] 주요 연구 내용



## 나. 연구 방법

- 국내외 평가지표 개발 방법 및 가이드라인 등 현황 파악 및 고찰
  - 요양급여 적정성 평가 등 국내 평가제도들의 지표 개발 방법
  - 주요 국가의 평가지표 개발 방법
- 전문가와 업무 담당자의 의견 수렴
  - 자문회의를 통해 평가지표 개발의 연구 수행 경험이 있는 임상전문가와 보건의료 평가 전문가의 의견 수렴
  - 건강보험심사평가원의 평가개발 및 수행 실무자와 인터뷰를 통해 현재 평가지표 개발 과정의 개선사항 발굴
  - 공청회를 통해 다양한 이해관계자와 전문가의 의견을 수렴
- 평가지표 분류체계의 지표를 대상으로 개발 지침에 적용해보고 개발 모형의 타당성과 활용 가능성을 검토함



[그림 4] 연구 방법

## 국내외 평가지표 개발 현황



## 제2장 국내외 평가지표 개발 현황

### 1. 미국의 Center for Medicare & Medicaid(CMS)

#### 가. 평가 영역 우선순위 결정

- 미국은 2015년 Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act(MACRA)에 따라 새로운 의료서비스 전달모형과 지불 제도를 도입함. MACRA(Section 102)에 따라 미국보건부(Health and Human Services, HHS)는 질 지표 개발 계획(Measure Development Plan Purpose, MDP)을 수립하고, 이에 따라 지표를 개발하며 지표 개발 전반적인 과정을 웹사이트에 공개하고 있음
- 여기서는 CMS가 MDP에 따라 지표 개발의 우선순위를 결정하는 과정과 성과를 살펴보고, 최근의 건강 환경과 격차를 분석한 내용을 소개하고자 함
- 또한 CMS에서는 최근 평가지표 우선순위를 결정하기 위한 목적으로 기존 지표가 균형적으로 배치되어 있는지와 새로운 지표의 개발이 필요한 영역은 어디인지를 파악하는데 도움을 줄 수 있는 도구(평가지표 개발 단계, Cascade of Measure)를 개발하였고, 최근까지는 다양한 전문가와 이해관계자의 의견을 수렴하는 과정에 있음. Cascade of Measure를 통해 개발이 필요한 평가 영역의 우선순위를 결정하는 과정을 알아보하고자 함

#### 1) CMS 평가 지표 개발 계획 (Quality Measure Development Plan, MDP)

- MACRA는 1) 더 좋은 의료(환자중심, 신뢰, 접근, 안전한 의료), 2) 더 현명한 지출(비용 절감), 3) 더 건강한 국민(행태, 사회, 환경 관련 건강결정 요인에 대한 개입을 통해) 이라는 세 가지 목적을 달성하기 위해 질 향상 전략을 수립하고, 이를 위해 CMS가 평가 지표를 개발하도록 함 ([그림 5] 참고)

- 이에 따라 CMS는 질 지표 개발 계획의 초안을 만들고, 이를 웹사이트에 공개하여 다양한 이해관계자와 국민의 의견을 수렴함. 약 210여 개인과 기관 등 이해관계자의 의견을 수렴하여, 2016년 계획을 최종적으로 확정하였음



[그림 5] CMS Quality Strategy Aims

#### 가) 질 지표 개발 계획 (Measure Development Plan Purpose, MDP)

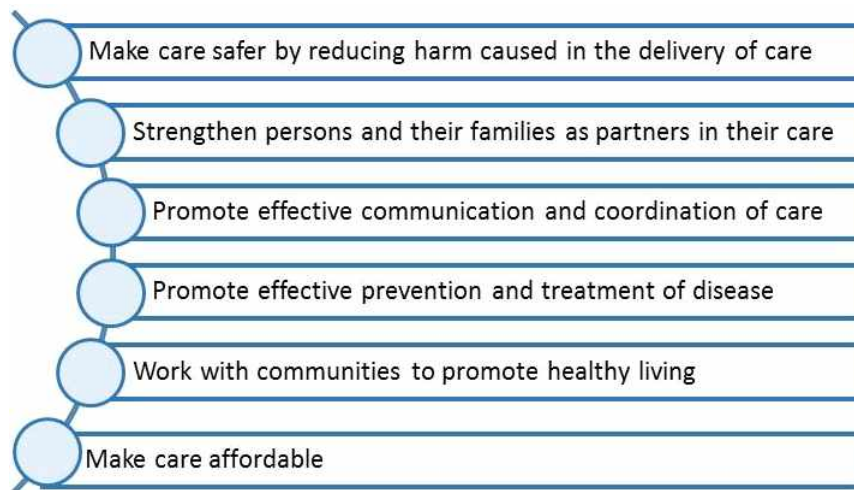
- MDP의 목적은 법에서 요구하는 사항을 충족하고 향후 평가지표 개발에 필요한 전략적 틀을 제공하고자 하는 것임
- MDP는 현재 평가지표의 성과를 비교 분석하고, 달성하고자 하는 목표와 격차가 큰 평가 영역에 높은 우선순위를 매김. 해당 영역에 대한 평가 지표를 개발하거나, 기존 지표를 보완하여 격차를 줄이기 위한 우선순위 선정과 접근 방식을 권장함
- 관련 정부 부처, 민간 기관, 전문가 등 광범위한 전문가의 경험을 공유하고 역할 조정에 대한 내용도 포함되어 있음

- 질 지표 개발 계획의 운영 요건(Operational Requirements of the Quality Measure Development Plan)은 다음과 같음
  - 다양한 지불자가 활용 가능(Multi-Payer Applicability)
  - 평가지표 개발자 사이의 조율과 공유(Coordination and Sharing Across Measure Developers)
  - 임상진료지침 활용(Clinical Practice Guidelines)
  - 미승인 지표는 근거 제시(Evidence Base for Non-Endorsed Measures)
  - 격차 분석(Gap Analysis)
  - 우선순위 영역 설정(Quality Domains and Priorities)
  - 다양한 의료기관에 적용(Applicability of Measures Across Health Care Settings)
  - 질 향상 활동과 연계(Clinical Practice Improvement Activities)
  - 전산 시스템 고려(Consideration for Electronic Specifications)
  
- 질 지표 개발 시 고려해야 하는 주요 사항은 다음과 같음
  - 환자, 보호자 및 지역사회와 협력
  - 임상 의사 및 전문가 단체와 협력
  - 다양한 지불자 간의 조정
  - 데이터 수집과 제출 부담 감소
  - 개발 기간을 단축
  - 예비 평가를 위한 데이터 수집 효율화
  - 의미 있는 결과 지표를 개발
  - 환자 보고 결과 측정(PROM) 지표 개발
  - 의료기관과 의료제공자의 공유된 책임감(shared accountability)을 촉진하는 지표 개발

## 나) 질 지표 개발 계획의 전략 (Strategic Vision of the Measure Development Plan)

○ CMS는 HHS의 National Quality Strategy (NQS)<sup>1)</sup>에 기반하여 질 향상을 위한 지표를 개발하는데, 2016년 기준 CMS의 6가지 질 향상 목표는 다음과 같음 ([그림 6] 참고)

- 의료서비스 전달 과정에서 위해를 감소하고, 안전을 향상
- 환자와 가족의 참여를 강화
- 효과적인 의사소통과 의료의 조정
- 효과적인 예방과 치료
- 건강증진을 위한 지역사회와 연계
- 적정 부담



[그림 6] CMS Quality Strategy Goals

1) NQS(National Quality Strategy)는 의료의 질 향상을 위한 국가 전략으로 미국 보건부(Department of Health and Human Services, HHS)를 대신하여 Agency for Healthcare Research and Quality(AHRQ)에서 주도하에, 2011년 3월 최초 발표됨



- 질 향상 목표를 달성하기 위한 기본원칙은 다음과 같음
  - 형평성 강화
  - 자료수집체계 강화
  - 지역사회역량 강화
  - 질 향상 프로그램을 통한 학습지원
  
- 평가지표 개발자는 CMS 전략과 원칙에 따라 개발해야 함
  - CMS는 환자중심의 의료를 제공하기 위해 의료제공자인 의료기관과 의사들 간의 공유된 책임감을 강조함
  - 기존의 질 평가 프로그램과의 연계성을 통해 임상진료진들의 부담을 줄이고, 평가 지표 개발 과정 전체에 환자와 소비자가 적극적으로 참여하고, 소비자가 의사결정에 활용할 수 있도록 결과를 공개하는 것을 원칙으로 함
  
- 지표개발 계획에 다음의 사항을 반영해야 함
  - 한 가지 이상의 만성질환을 가진 환자에게 제공되는 의료서비스의 연속성을 추적해야 함
  - 치료의 목적과 환자와 가족의 선택을 강조하고, 의사와 환자의 관계를 강조함
  - 만성 질환의 위험을 증가시키는 약물남용과 정신건강 문제를 가진 환자의 통합적인 건강관리 강조함
  - 결과와 관련이 큰 과정지표 또는 결과지표를 강조함
  - 환자 경험, 의료서비스의 조정, 적절한 의료서비스 제공(과다와 과소 이용)에 관심을 가짐
  - 다양한 수준에서 책무성을 향상(의사 수준, 의료기관 수준, 인구집단 수준)시켜야 함
  - 모든 전문과목과 세부 전문과목을 포괄할 수 있어야 함

- 의사, 의사 외 의료인, 환자를 대면하지 않는 의료진 모두를 대상으로 적용할 수 있어야 함
- 병원, 요양시설 등 다양한 형태의 의료기관에 적용할 수 있어야 함
- 현재의 진료환경을 고려하고, 가능한 EHRs에서 생성되는 자료를 활용해야 함
- 나이, 성, 인종 등 인구 통계학적 변수에 따른 결과를 산출하여, 형평성을 향상하는데 활용할 수 있어야 함
- CMS Physician Compare website에 공개함
- 다양한 형태의 지불 모형에 적용 가능해야 함
- 메디케이드를 포함하여 관련한 연방 기관 및 민간 부문을 포함한 다양한 평가 모형과 보고 체계에 적용 가능해야 함
- 진료 건수가 적은 농어촌 또는 개인 의원에도 적용 가능(e.g., rural providers, small and independently owned physician practices)해야 함

#### 다) 격차의 확인과 우선순위 결정 (Priorities and Gaps)

- CMS는 다양한 이해관계자의 의견, 최신의 출판물, 연방 보고서, 이슈(예: HHS의 약물 부작용 예방 계획, Core Quality Measures Collaborative<sup>2)</sup>), 의사의 질 보고 시스템(Physician Quality Reporting System, PQRS)<sup>3)</sup>을 통해 질 개선이 필요한 영역이나 격차를 확인하고, 우선순위를 결정함
- 2016년 우선순위가 높은 영역은 다음과 같음
  - 임상 진료

2) CQMC(Core Quality Measures Collaborative)는 다양한 지불자가 활용할 수 있는 핵심 평가지표개발을 위한 보건의료 전문가 연합

3) 의사의 질 보고 시스템(Physician Quality Reporting System, PQRS)은 의사가 환자에게 적절한 시간 안에 적절한 의료가 제공되었는지를 평가하고, 이를 CMS에 보고하는 프로그램임. CMS의 피드백을 통해 다른 의료제공자와 비교평가할 수 있음. 2017년부터 새로운 성과 기반 인센티브 프로그램(Merit-based Incentive Payment System, MIPS)으로 통합운영됨.

- 환자의 선호와 공유된 의사결정
  - 한 개 이상의 전문 과목을 포괄
  - 격차가 큰 영역에 초점
  - 결과 지표
- 안전
    - 정확한 진단
    - 안전한 약물 사용
- 진료의 연속성
    - 팀 기반 평가(e.g., 환자 정보 공유의 적시성)
    - 원격의료와 같은 새로운 의료기술의 효과
- 환자 경험
    - 환자 보고 결과(Patient-reported outcome measures, PROMs)
    - 환자와 가족에게 중요한 주제(자가 관리를 위한 지식, 기술, 자신감 등)
    - 인구집단의 건강과 예방
    - 건강 증진과 예방 서비스의 효과를 평가하기 위한 인구집단의 건강 결과(기대 여명, 삶의 질, 중독)
    - 만성질환의 발견과 예방
- 적정 의료
    - 검사나 시술의 과다이용

## 2) CMS 평가지표 개발 계획 (Quality Measure Development Plan, MDP) 성과 보고서<sup>4)</sup>

- CMS는 2016년 평가지표 개발 계획(Quality Measure Development Plan, MDP) 발표 이후 매년 평가지표 개발 과정과 운영 성과를 담은 보고서를 발표하고 있음
- 2022년 발표된 MDP 보고서에는 재정 운영 현황, 지표의 개발 과정, 의료 질 수준의 차이 확인 및 우선순위, 추진 일정, 지표목록 관리 과정, 가치기반 의료와 형평성 강화를 위한 CMS 정책에 관한 내용이 포함되어 있음 ([그림 7] 참고)
- 2021년에는 \$13.2 million 중 \$10.6을 지표 개발에 지출하였고, 개발된 지표 중 결과 또는 중간 결과지표는 12개이고, 과정지표는 5개, 환자 보고 결과지표는 8개였음
- 2023년에 지표개발이 필요한 높은 우선순위 영역은 건강행태, 당뇨, 에이즈, 고혈압, 신장질환, 여성 및 모성 관리 등 6개 영역임

2022 Quality Payment Program Measure Inventory			MACRA* Section 102 Funding	
<b>Merit-based Incentive Payment System (MIPS)</b>		<b>Qualified Clinical Data Registries</b>	<b>\$13.2 million</b> expended in fiscal year (FY) 2021	
200 total quality measures for all specialties (24% eCQMs)		CMS-approved: 49	\$10.6 million for measure development	
Outcome/intermediate outcome:	41	Applicable to specialties prioritized in Measure Development Plan and Annual Reports:	\$2.6 million for support activities	
Patient-reported outcome-based performance measures (PRO-PMs):	18	<b>Advanced Alternative Payment Models (APMs)</b>		
Efficiency:	6	CMS-approved: 10		
Process:	132			
Patient experience:	2			
Structural:	1			

Key Accomplishments for the Quality Payment Program		25 Measures Developed in FY 2021	
<ul style="list-style-type: none"> <li>Established 7 MIPS Value Pathways for the 2023 performance year and prioritized 6 clinical areas for developing MVP measure sets focused on outcomes: behavioral health, diabetes, HIV/AIDS, hypertension, kidney disease, and women's health/maternal care</li> <li>Added the APM Performance Pathway as a reporting option for eligible clinicians in MIPS APMs</li> <li>Expanded MIPS telehealth services and made 39 electronic clinical quality measures (eCQMs) applicable for clinician reporting</li> <li>Added 5 episode-based cost measures for a total of 23 such measures used in CY 2022, which cover an estimated 13% of Medicare Part A and B costs toward a statutory target of 50%<sup>1</sup></li> </ul>		Outcome/Intermediate:	12
		Process:	5
		PRO-PMs:	8
		<b>3 Measures Continuing Development in 2022</b>	
		PRO-PM:	1
		Process:	1
		Composite:	1

[그림 7] 2022년 MDP의 주요 내용 요약

4) <https://www.cms.gov/Medicare/Quality-Payment-Program/Measure-Development/Measure-development>

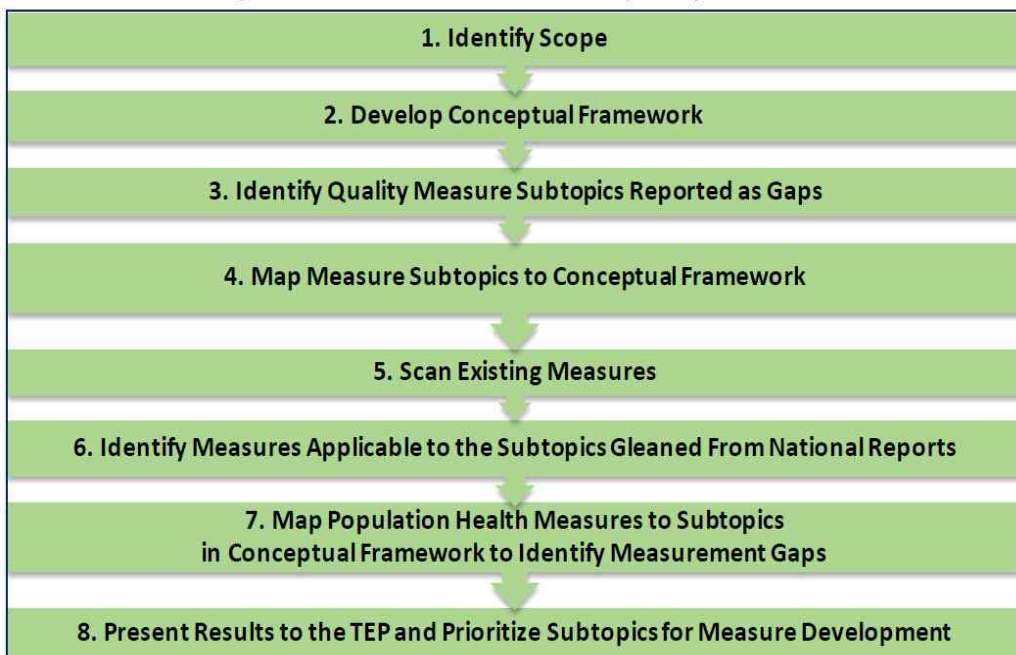
- 평가지표 개발 관련해서는 지표의 이름, 우선순위, 지표가 속한 영역, 개발자, 지표 유형, 승인 여부, 자료원에 대한 자세한 내용이 기술됨
- CMS는 매년 pre-rulemaking process<sup>5)</sup>를 거쳐 평가지표 개발의 우선순위를 결정함. 2023년 의료의 질 격차가 커서 우선순위가 높은 영역으로 선정된 내용은 다음과 같음
  - 전문과목(specialties) : 치과, 언어병리, 건강행태, 모성건강
  - 임상 상태(Clinical conditions) : 당뇨병환자의 다리 절단 예방
  - 형평성(Health equity) : 건강의 형평성과 건강의 사회적 결정요인 평가
  - 기타(Other topics) : 공유 의사결정, 환자 보고 건강결과
- 위의 영역 외에도 평가지표 개발의 우선순위가 높은 임상 영역은 건강행태, 당뇨, 에이즈, 고혈압, 신장질환, 여성 및 모성 관리 등 6개 영역임. MDP 프로젝트에서는 해당 임상영역의 환경을 분석하고, 질 수준의 격차를 분석하고, 전문가(MDP Technical Expert Panel)의 피드백을 통해 결과지표 개발을 위한 우선순위를 결정하게 됨

### 3) 2020 건강 환경과 격차 분석 (Population Health Environmental Scan and Gap Analysis Report)

- CMS Quality Measure Development Plan를 위해 의료 환경을 분석하고, 의료의 질 격차를 확인함. 이를 통해 2020년에는 인구집단의 건강을 향상시키기 위해 필요한 의료의 질 평가 주제를 선정함

5) 2010년 시행된 Affordable Care Act(ACA)에 따라 HHS는 의료의 질과 효율성 평가 지표를 선택하기 위한 사전 규칙 제정 프로세스(pre-rulemaking process)를 마련함. 구체적인 단계는 1) 매년 CMS에서 평가지표 개발자들에게 후보지표를 제안 받음 2) 매년 12월 1일 이전에 검토 중인 지표 목록(the Measures Under Consideration(MUC) List)을 공개함 3) 매년 2월 1일 이전까지 HHS의 검토 중인 지표에 대한 다양한 이해관계자의 의견 수렴을 마무리함 4) HHS에서는 다양한 전문가와 이해관계자의 의견을 취합함 5) HHS에서 선정한 평가지표는 연방정부에서 공개하고, 최종 확정 이전에 추가적인 의견 수렴 과정을 거침 6) NQF에 의해 승인 결정이 완료되지 않은 지표는 합리적인 근거를 제시하도록 함

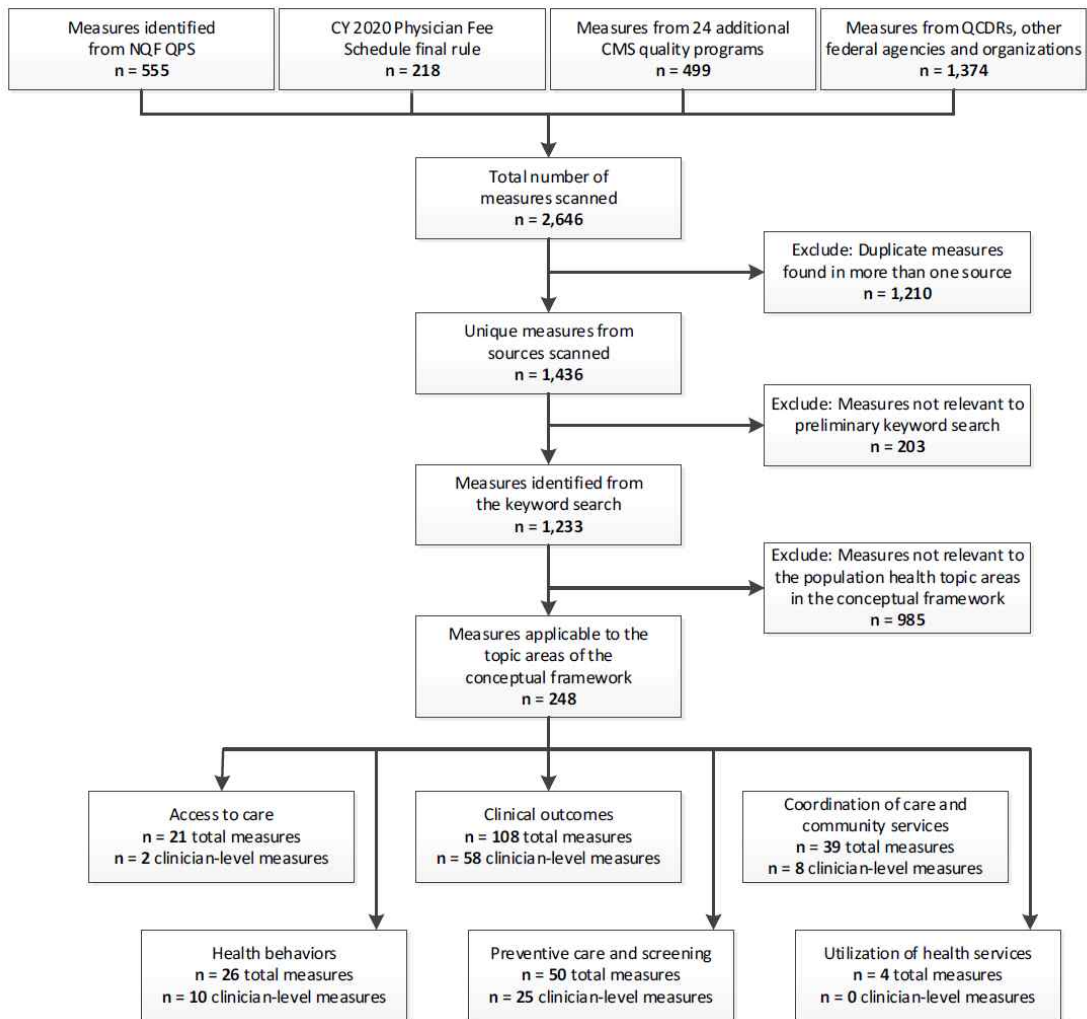
- 아래의 4가지 활동을 통해 임상 질 지표를 분석하고, 의료 환경을 파악하고 질의 격차를 분석함으로써 평가지표 개발을 위한 예산 배정의 우선순위를 결정함 ([그림 8] 참고)
- 개념 틀 구성(Conceptual framework) : 6개의 주제(접근성, 임상 결과, 의료서비스와 지역사회 서비스의 조정, 건강행태, 예방과 검진, 의료자원 활용)로 구성
- 의료의 질 격차 확인(Health care quality gaps) : 연방정부 자료, 정부 보고서, 논문 등을 검토하여 중요한 격차를 확인
- 현재의 평가지표 검토(Existing measures) : 다양한 기관의 평가 프로그램과 자료원을 활용
- 기존 지표의 채택과 보완 검토(Opportunities for adoption or adaptation of existing measures) : 다양한 프로그램의 평가지표의 도입 또는 보완



[그림 8] 환경 평가 및 격차 분석 과정

## 가) 현재 평가 지표 분석

- 2020년 환경 분석을 위해 다양한 기관의 평가 프로그램과 자료원을 활용함
  - NQF Quality Positioning System (QPS)/NQF website : 임상 지표 관련 다양한 프로젝트의 지표를 검토하고, NQF 승인 상태 확인
  - CMS Measures Inventory Tool : 다양한 평가, 결과 공개, 지불 프로그램에서 사용하는 CMS의 지표 목록을 확인
  - CMS quality programs : CMS의 25개 질 향상 프로그램 검토
  
- 이 외에도 학술 단체, 의학계 단체, 공공 및 민간 의료기관의 다양한 정보 검토한 결과, 인구집단에 적용 가능한 248개의 지표를 확인하였고, 접근성(21개), 임상 결과(108개), 의료서비스와 지역사회 서비스의 조정(39개), 건강행태(26개), 예방 및 검진(50개), 의료자원 활용(4개)으로 구분함. 환경 분석 과정과 결과는 다음 [그림 9]와 같음



[그림 9] 환경 분석 과정과 결과

## 나) 격차 분석

- 연방정부 질 프로그램을 포함하여 인구집단의 건강을 평가할 수 있는 지표를 5가지 수준(임상의사(clinician/clinician group), 의료기관(facility), 보험(health care plan), 통합전달체계(integrated delivery system), 인구집단(population))으로 분석함. 인구집단의 건강 향상에 초점을 맞춘 2020년 보고서의 주요 결과는 다음과 같음



〈표 2〉 미국 NQF 2020년 보고서 주요 결과

**1) 평가지표 채택 (Opportunities for Population Health Measure Adoption)**

- 248개의 건강 평가지표 중 임상 의사 대상 103개의 지표: 환경과 격차 분석을 위해 6개 주제에 따라 배치
  - 1개의 지표 : 정신질환 입원 이후 추적관찰(MIPS<sup>6)</sup> #391/NQF #0576) 지표는 행정 자료를 활용하여 MVPs<sup>7)</sup>에서 활용하기 위한 후보 지표로 채택
  - 나머지 지표는 행정자료로 분석이 불가능하고 추가 자료 필요

**2) 평가지표 보완 (Opportunities for Population Health Measure Adaptation)**

- 35개의 임상 지표 중 14개는 현재의 지표를 보완하면 격차 축소 가능
  - 접근성 : 외국어 통역 서비스, 어린이 보험 가입율 향상, 원격의료 등을 통해 접근성 향상 가능
  - 임상 결과 : 암 사망, 모성 사망, 출산 결과
  - 의료 서비스와 지역사회 서비스의 조정 : 주거, 지역사회 서비스 제공, 마약 중독자 지원
  - 건강행태 : 흡연
  - 예방 및 검진 : 전립선 암 검진
  - 의료자원 활용 : 부적절한 응급실

**3) 새로운 평가지표 개발 (Opportunities for de Novo Population Health Measure Development)**

- 35개 지표 중 남아있는 21개의 평가지표는 격차를 줄이기 위해 새로운 평가지표의 개발이 필요
  - 접근성 : 도시가 아닌 지역의 노인 영양 상태 개선을 위한 지원

- 임상 결과 : 임신 간격, 마약 관련 이환율과 사망률, 미숙아 사망, 산욕기 합병증
- 의료 서비스와 지역사회 서비스의 조정 : 모유 수유지원, 취업, 모성건강/약물남용/신체 건강 통합, 노인대상 사회적 지지, 약물 사용 장애 의료서비스의 적시 이동(transition)
- 건강행태 : 사고 예방(머리 손상, 안전벨트, 부주의 운전), 건강정보 활용능력, 약품의 안전한 폐기(disposal)
- 예방 및 검진 : 학대 및 방치, 흉부 암 검진, 간병인의 건강 위험도 평가
- 의료자원 활용 : 격차가 발견된 주제 없음

#### 4) 의미 있는 평가지표 2.0 (Meaningful Measures 2.0 Framework)

- CMS는 질 평가 및 향상의 우선순위가 높은 영역을 선정하고, 선택과 집중을 통해 가치를 높이하고자 2017년 ‘Meaningful Measures Initiative’을 도입하였음. 이를 통해 많은 기존의 지표 중 의미가 있는 지표를 우선적으로 선별하여 활용하고, 의료진의 보고 부담을 줄여줄 수 있는 방안을 찾고자 함
- ‘Meaningful Measures’를 통해 임상적으로 관련이 없거나, 중복이 되는 지표를 삭제함(전체 지표 중 15% 감소). 또한 과정지표 보다 환자 보고를 포함하여 결과 지표의 우선순위를 높게 설정하고, 보고 해야 하는 지표의 수를 감소시켜 임상 의사들의 보고 부담을 줄이는 성과가 있었음
- 2020년에는 환자의 의사결정 역량을 강화하고, 전자 데이터 활용으로 의료진과 병원의 평가 부담을 감소 중점을 둔 ‘Meaningful Measures Framework 2.0’을 발표하였음
- ‘Meaningful Measures Framework 2.0’의 목표는 다음과 같음
  - 첫째, 환자중심의 평가와 정보 공개를 통해 소비자 선택의 역량 강화

6) Merit-based Incentive Payment System(MIPS) : The Medicare Access and Children’s Health Insurance Program (CHIP) Reauthorization Act of 2015(MACRA)에 따라 도입된 환자 의견, 환자 결과 및 비용 효과를 강조하는 새로운 지불 및 전달 모델

7) MIPS Value Pathways(MVPs)

- 둘째, 건강 형평성 제고를 위한 평가지표 활용
  - 셋째, 평가지표 개발과 활용의 효율성 개선
  - 넷째, 정보 공개 및 지불보상 제도와 연계를 통해 결과 향상 촉진
  - 다섯째, 디지털 정보의 활용과 분석방법의 정교화를 통해 평가의 효율성 개선
- 기존 평가지표 중 향후 개선할 수 있는 가능성 없거나, 지표의 근거가 부족하거나, 다른 지표와 중복되거나, 환자 또는 임상적 관점에서 가치가 낮은 지표를 퇴출함
- 평가지표의 우선순위를 기획하는 단계는 다음과 같음. 첫째 환경 분석, 둘째 격차 분석, 셋째 주요 이해관계자 참여임. 해당 과정에서 필요한 평가지표에 대한 전문가 의견을 수렴하고, 평가지표 개발의 우선순위를 결정하고, 현재의 평가지표를 평가함. 이를 통해 피평가자(기관)의 평가로 인한 비용과 부담을 감소하고 환자중심성과 가치 기반 지불제도 연계를 강화함
- CMS는 매년 National Impact Assessment라는 보고서를 통해 질 평가의 성과를 발표하는데, 2021년 발표된 성과는 다음과 같음
- 첫째, CMS 질 평가 프로그램에서 사용하는 평가지표의 수가 24% 감소
  - 둘째, 2015년 이후로 결과지표가 39%에서 46%로 증가
  - 셋째, 2020년 평가지표의 약 80%가 디지털 자료 활용
  - 넷째, 15개 주요 지표에서 약 296~519억 달러 평가비용 감소
- ‘Meaningful Measures Initiative 2.0’에서는 평가지표 개발 단계(Cascade of Measures) 도구를 통해 기존의 평가지표를 평가의 목적, 목표, 국가 수준의 우선순위와 연결을 시킴. Cascade of Measures의 활용을 통해 기존 지표가 주요 영역별로 균형적으로 배치되어 있는지와 새로운 지표의 개발이 필요한 영역은 어디인지를 파악하는데 도움을 얻을 수 있음
- 평가지표 개발 단계(Cascade of Measure)는 2021년 가을까지 다양한 전문가와 이해관계자의 의견을 수렴하였고, 일부 내용을 보완하는 과정이지만, 기존의 CMS 평가지표 맵핑은 완료됨. Cascade of Measures의 활용을 통해 향후 질 향상 정책의 목적과 목표 달성에 필요한 결과 중심의 평가로의 전환을 촉진하고자 함

○ 평가지표 개발 단계(Cascade of Measures)

- Cascade of Measures는 기존 지표가 너무 많은 경우에는 지표수를 줄이고, 새로운 평가 지표 개발이 필요한 영역을 확인하면서, 지표의 우선순위를 설정하기 위한 도움임. 구체적인 최종 목표는 다음의 2가지 임. 1) 사용 중이거나 개발 중인 평가 지표의 우선순위 선정, 2) 목표와 목적 달성에 초점을 맞춘 결과 지표와 복합지표 개발에 초점

- ‘Meaningful Measures Initiative 2.0’에서 선정된 주요 영역은 환자 중심 케어, 안전, 만성 질환, 진료의 연속성, 형평성, 경제성 및 효율성, 웰니스 및 예방, 건강행태로 총 8개 영역임 ([그림 10] 참고)

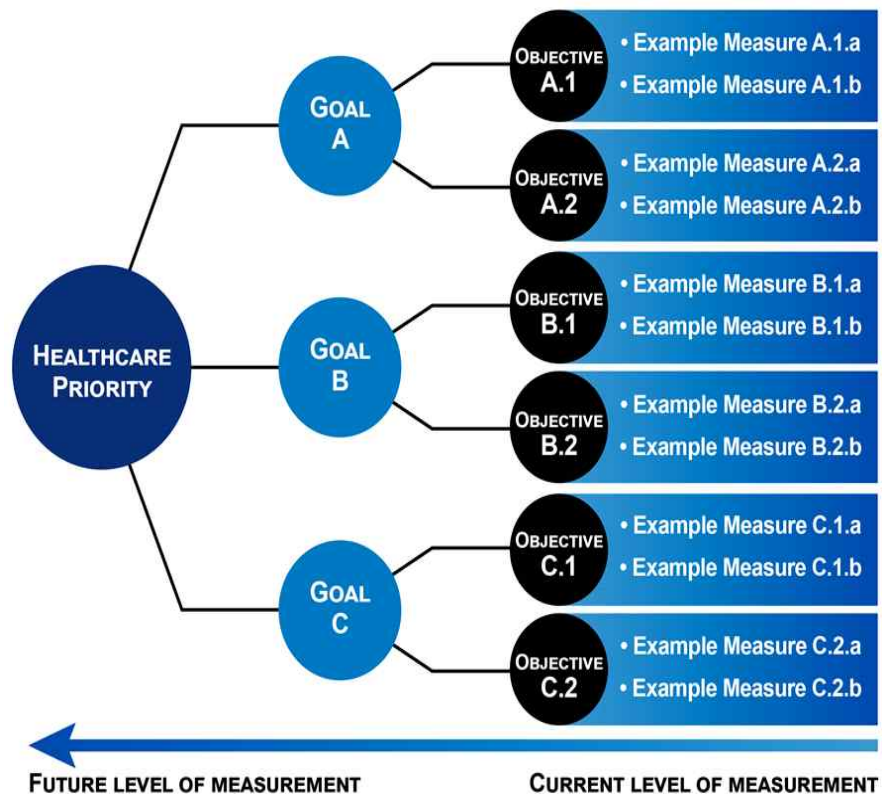


[그림 10] 의미 있는 평가 지표 2.0 구성  
(Meaningful Measures 2.0 Framework)

- ‘Meaningful Measures Initiative 2.0’의 8개 우선 순위에 따라 현재의 평가지표를 배치하는 것이 Cascade 도구의 활용의 첫 단계임
- 다음 [그림 11]과 같이 왼쪽의 주요 우선순위 영역에 따라 현재의 평가지표를 배열하는 과정을 통해 우선순위 영역별로 목표, 목적, 평가지표 묶음, 개별 평가지표가

연결됨 ([그림 11] 참고)

- CMS의 평가영역의 우선순위를 반영하여 평가지표를 개발함으로써, 거시적인 관점의 목적, 목표, 우선순위를 정량화해서 제시할 수 있음
  - 목표(Goals) : 우선순위 영역의 주요 목표
  - 목적(Objectives) : 목표 달성 여부를 평가할 수 있는 측정 가능한 지표
  - 지표 묶음(Measure Families) : 목적 달성 여부를 평가하는데 필요한 자원과 과정, 상태(conditions) 등의 지표 묶음
  - 지표의 예(Example Measures) : 목적 달성을 위한 지표 묶음의 개념을 보다 구체화할 수 있음



[그림 11] CMS 목표 달성을 위한 측정지표 개발

○ 다음 [그림 12]는 환자안전 영역의 평가지표 개발 단계를 보여줌. 각 목표는 다수의 평가지표로 구성된 여러 개의 목적으로 구성됨

- 예를 들어, 환자안전의 경우 심각한 환자안전 사고의 감소라는 목표를 정하고, 이 목표 달성을 위해 병원감염 감소, 합병증 감소, 진단오류 감소, 투약오류 감소 등의 목적을 설정할 수 있음. 여러 개의 목적 중 병원 감염 감소라는 목적을 달성하기 위해서는 중심 정맥관 감염이나 도뇨관 감염을 평가하는 개별 지표를 개발하고 이를 활용할 수 있음 ([그림 12], [그림 13] 참고)

CASCADE OF MEASURES-SAFETY EXAMPLE		
HEALTHCARE PRIORITY: SAFETY		
Goals	Objectives	Example Measures
Reduction in National Serious Safety Events	Reduction in Healthcare Associated Infections	<ul style="list-style-type: none"> <li>• National Healthcare Safety Network (NHSN) Central Line-Associated Bloodstream Infection (CLABSI) Outcome Measure</li> <li>• NHSN Catheter-Associated Urinary Tract Infection (CAUTI) Outcome Measure</li> <li>• NHSN Facility-wide Inpatient Hospital-onset Clostridium difficile Infection (CDI) Outcome Measure</li> <li>• Surgical Site Infection</li> <li>• NHSN Facility-wide Inpatient Hospital-onset Methicillin-resistant Staphylococcus aureus (MRSA)</li> <li>• Severe Sepsis and Septic Shock: Management Bundle</li> </ul>
	Reduction in Healthcare Associated Complications	<ul style="list-style-type: none"> <li>• Perioperative hemorrhage or hematoma</li> <li>• Surgical Treatment Complications for Localized Prostate Cancer Measure</li> <li>• Surgical complications</li> <li>• Iatrogenic Pneumothorax Rate</li> <li>• See also Reduction in Healthcare Associated Infections</li> <li>• See also Reduction in Medication Error</li> </ul>
	Diagnostic Accuracy/Error	<ul style="list-style-type: none"> <li>• Appropriate Follow-up Imaging for Incidental Abdominal Lesions</li> <li>• Closing the Referral Loop: Receipt of Specialist Report</li> </ul>
	Reduction in Medication Error	• No Current Measures
	EMR Safety	<ul style="list-style-type: none"> <li>• Safety Assurance Factors for EHR Resilience (SAFER) Guides</li> </ul>
	Safety of the Patient Environment	<ul style="list-style-type: none"> <li>• Dementia: Safety Concern Screening and Follow-Up for Patients with Dementia</li> </ul>
HEALTHCARE PRIORITY: SAFETY (cont.)		
Goals	Objectives	Example Measures
Safety Culture	Culture of Safety	• No Current Measures
	Reliability/Resiliency	• No Current Measures
Workforce Safety	Preparedness	• No Current Measures
	Workforce Resilience	• No Current Measures
	Staffing	• No Current Measures
Safety for Special Populations	Reduction in Burnout and Turnover	• No Current Measures
	Maternal Safety	<ul style="list-style-type: none"> <li>• Low-Risk Cesarean Delivery (LRCD-CH)</li> <li>• Maternity Care: Postpartum Follow-up and Care Coordination</li> </ul>
	Pediatric Safety	<ul style="list-style-type: none"> <li>• Prevention of Post-Operative Vomiting (POV) Combination Therapy (Pediatrics)</li> <li>• Perioperative Temperature Management</li> </ul>
	Elder Care/Geriatrics/ Nursing Home Residents	<ul style="list-style-type: none"> <li>• Reducing the Risk of Falling</li> <li>• Percent of High-Risk Residents with Pressure Ulcers</li> <li>• Care for Older Adults Medication Review</li> <li>• Dementia: Safety Concerns Screening and Follow-up for Patients with Dementia</li> <li>• Hospital 30-Day All-cause, risk standardized mortality Rate following Pneumonia Hospitalization</li> <li>• Elder Maltreatment Screen and Follow-Up Plan</li> </ul>
	ID/DD Safety	• National Core Indicators Survey (NCIDS-AD)
	Safety for Populations Disproportionately Affected by Poor Health Outcomes	• No Current Measures

Figure 2 above shows the Cascade of Measures framework for the Meaningful Measures 2.0 healthcare priority, Safety. Each Goal is comprised of multiple Objectives for which there may be multiple existing quality measures. Findings from literature reviews and stakeholder feedback were integrated to identify appropriate Objectives and Measure Families for each healthcare priority, and individual measures in the CMS Measures Inventory Tool (CMIT) were identified as examples.

#### NEXT STEPS

**Near-Term Goals** During the first half of 2022, CMS is planning for agency wide implementation of the Cascade of Measures.

The complete Cascade is available online here

<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/CMS-Quality-Strategy>

[그림 12] 환자안전을 위한 측정 단계(Cascade of Measures) 활용의 예





[그림 13] CMS 목표 달성을 위한 측정지표 개발

- 이와 같이 ‘Meaningful Measures Initiative’는 예서는 국가 수준의 우선순위에 따라 구체적인 평가지표를 개발하고, 결과 중심의 평가지표가 확산될 수 있도록 유도 하는데 도움을 주게 됨
- CMS는 2022년 상반기에 측정 단계(Cascade of Measures)의 확대 적용을 위한 계획을 수립하고 있음<sup>8)</sup>

8) <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/CMS-Quality-Strategy>

## 나. 평가지표 개발 과정

- 미국 CMS의 평가지표 개발 과정의 핵심 원칙은 다음과 같음
  - ① 의료의 질을 개선할 수 있는 방안이 있는가? 의료의 질에 미치는 영향이 큰가?
  - ② 환자에게 의미가 있는 지표인가(환자중심적인 지표인가)?
  - ③ 결과지표인가? 결과와 관련성이 큰 지표인가?
  - ④ 평가 대상 기관의 부담을 최소화할 수 있는 방법을 고려했나?
  - ⑤ 다양한 평가 프로그램의 지표 또는 기존의 평가지표와 비교검토 했는가?
  - ⑥ 형평성을 확인하거나 형평성을 높이는데 도움이 되는 지표인가?
  - ⑦ 평가가 과다이용이나 과소이용 등 의도하지 않은 결과를 초래하는가? 이를 최소화 할 수 있는 방안이 있는가?
  - ⑧ 평가지표 개발 초기부터 이해관계자를 참여시켰는가?
  - ⑨ 전자데이터(예: 전자의무기록)를 활용해서 평가할 수 있는 지표인가?
  - ⑩ 과학적 수용성(예: 타당성과 신뢰성)이 확보되었는가?

○ CMS에서는 평가지표를 다음 <표 3>과 같은 유형으로 구분하고 있음

<표 3> CMS의 질 평가 유형

질 평가 유형	설명 및 목적	예시
복합 (Composite)	두 개 이상의 개별 평가지표를 조합하여 하나의 결과를 산출하는 것	환자안전과 위해사건을 복합지표로 평가
비용/자원(사용) (Cost/Resource)	의료서비스가 적용된 인구집단이나 사건에 사용된 비용이나 자원을 측정 자원 사용은 정의된 의료시스템 자원의 사용 빈도를 계산함. 일부는 사용된 자원의 각 단위당 적용된 달러의 양으로 표현	수혜자당 메디케어 지출
효율성 (Efficiency)	특정 수준의 건강결과와 관련된 의료비용을 측정	비심장 저위험 수술에서 수술 전 위험 평가를 위한 심장 영상 촬영
결과 (Outcome)	환자의 건강 상태를 측정 또는 건강 상태의 변화를 측정	암 환자의 30일 이내 계획되지 않은 재입원



질 평가 유형	설명 및 목적	예시
결과 기반의 환자보고 (Patient reported outcome-based)	책임 있는 의료기관에 대해 집계된 환자가 보고한 결과측정을 바탕으로 성과를 측정함	목 손상을 가진 환자에서 기능적 상태 변화
절차 (Process)	의료를 제공하는 과정 측정	성인 주요 우울증 : 자살위험평가
구조 (Structure)	의료를 제공하기 위한 역량과 관련된 의료기관 조직이나 인력 측정	방사선과 : 유방조영술 스크리닝을 위한 시스템을 알림

○ 평가지표 자체에 대한 평가는 아래의 5가지의 기준을 토대로 지속적으로 분석함

① 중요성(importance) : 환자에게 중요하고 의미 있는 지표인가? 개선이 필요한 분야를 다루고 있는가?

- 의료의 질 향상에 기여하는지의 여부(예: 안전성, 적시성, 효과성, 효율성, 형평성, 환자중심성). 특히 의료의 질의 변이가 크거나 질이 낮은 곳에 미치는 영향을 고려함

② 실행가능성(Feasibility) : 평가의 이득이 자료를 수집하고 공개하는 평가의 부담보다 큰가?

- 필요한 자료의 지속적인 활용 가능 여부와 평가 부담의 정도를 고려함

③ 과학적 수용성(Scientific acceptability) : 신뢰성과 타당성 검토함

④ 유용성(Usability) : 환자, 평가기관, 정책입안자, 다양한 이해관계자가 평가결과를 얼마나 활용할지 여부

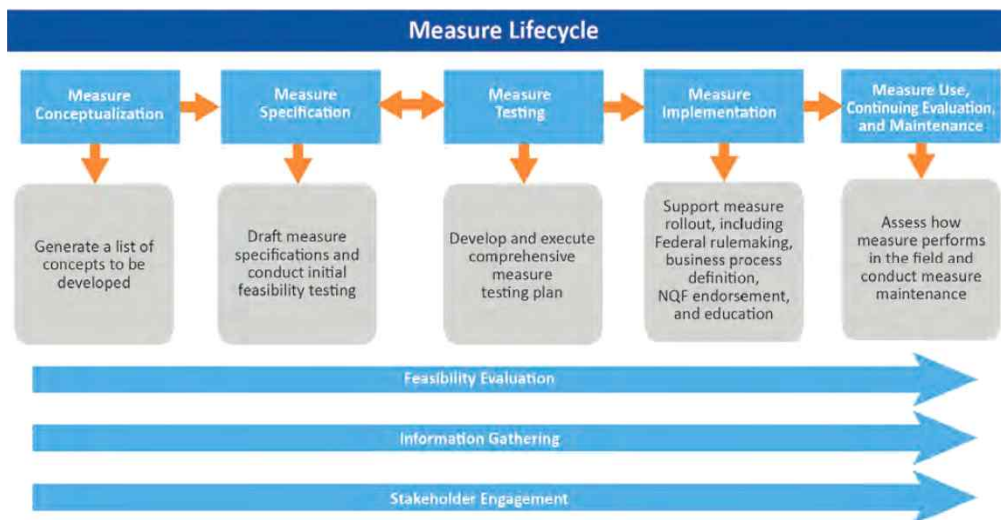
⑤ 관련 또는 기존 측정지표와 비교(comparison to related or competing measures) : 평가지표의 내용이나 대상과 유사한 지표가 있는가? 기존 지표와의 조율을 통해 평가 부담을 줄일 수 있는 방법이 있는가?

- 기존 지표와 분자, 분모, 제외기준, 산출식, 자료원 등의 조율 방안 모색

Evaluation Criterion	Questions to Consider When Addressing the Criterion
Importance	Is this measure meaningful and important to patients? Does it address an aspect of healthcare where there is a gap in performance or measurement?
Feasibility	Do the benefits of this measure outweigh the potential burdens associated with collecting and reporting on it?
Scientific Acceptability	Reliability: Does the measure consistently measure the same thing? Validity: Is the measure actually measuring what it is supposed to measure?
Usability & Use	To what extent can patients, measured entities, or other stakeholders use information from the measure to inform performance improvement or improve accountability in care delivery?
Comparison to related or competing measures (Harmonization)	Are there existing measures that have elements in common with this measure? To what extent can this measure leverage those elements to reduce the burden associated with implementation and reporting?

[그림 14] 평가지표의 가치 분석을 위한 5가지 항목 기준

- 평가지표는 지속적으로 모니터링 및 업데이트되며 폐기되는 생애주기 과정을 거치게 됨. 적용 이후에도 의료의 질과 관련이 있는지, 유용한지, 과학적으로 타당한지 등 확인하는 작업을 거침)



[그림 15] CMS 평가 측정의 생애주기

9) CMS Quality Measures: How They Are Developed, Used, & Maintained. CMS Sep. 2021

- CMS는 평가지표를 개발하고 유지하는데 표준화된 접근법을 사용함. 지표 개발부터 적용 및 재평가의 주기는 개념화(conceptualization), 구체화(specification), 검증(testing), 적용(implementation), 사용 및 지속적인 평가와 유지(Use, continuing evaluation and maintenance)의 5개 단계로 구성됨
  - ① 개념화(conceptualization) 단계 : 평가지표의 구체적인 개념을 명확하게 설정함. 평가개발자는 환경을 분석하고 환자를 포함하여 광범위한 이해관계자들로부터 의견을 수렴함
  - ② 구체화(specification) 단계 : 평가지표의 구체적인 산출방식을 결정함. 대상 인구 집단, 예상 결과를 검토하고, 자료원을 문서화함
  - ③ 검증(testing) 단계 : 실행 가능성, 평가지표의 기술적 사양에 대한 과학적 수용성을 평가함. 평가지표의 장점과 단점을 검토함
  - ④ 적용(Implementation) 단계 : 선정된 평가지표의 적용에 필요한 절차를 마련함
  - ⑤ 사용, 지속적인 평가와 유지(Use, continuing evaluation and maintenance) 단계 : 환경 분석과 문헌 검토를 통해 평가지표가 지속적으로 가치 향상에 도움이 되고 있는지를 모니터링 함
- 각 단계의 목적과 주요 활동은 <표 4>와 같음<sup>10)</sup>
  - 첫 번째 평가지표 개념화 단계에서는 문제를 확인하고 우선순위를 계획함. 또한 해당 단계에서 평가를 함으로써 얻게 되는 효과, 비용, 예상하지 못한 결과 발생과 같은 풍선효과까지 검토하게 됨
  - 두 번째 평가지표 구체화 단계에서는 후보 평가지표 목록을 개발하고 기술적 구체화 및 자료원 정의 등의 과정이 포함됨
  - 세 번째 평가지표 검증 단계에서는 검증 계획을 세우고 검증 결과를 분석하고 측정지표를 보완하는 등의 과정이 포함됨
  - 네 번째 평가지표 적용 단계에서는 NQF 승인을 받아 평가지표를 선정하는 과정이 포함됨

10) CMS Quality Measures: How They Are Developed, Used, & Maintained. CMS Sep. 2021

- 마지막 평가지표 사용, 평가, 유지 단계에서는 평가지표를 활용하며 지속적인 평가를 통해 해당 지표를 유지하거나 퇴출하는 등의 의사결정을 내리게 됨

〈표 4〉 CMS의 질 평가 활동

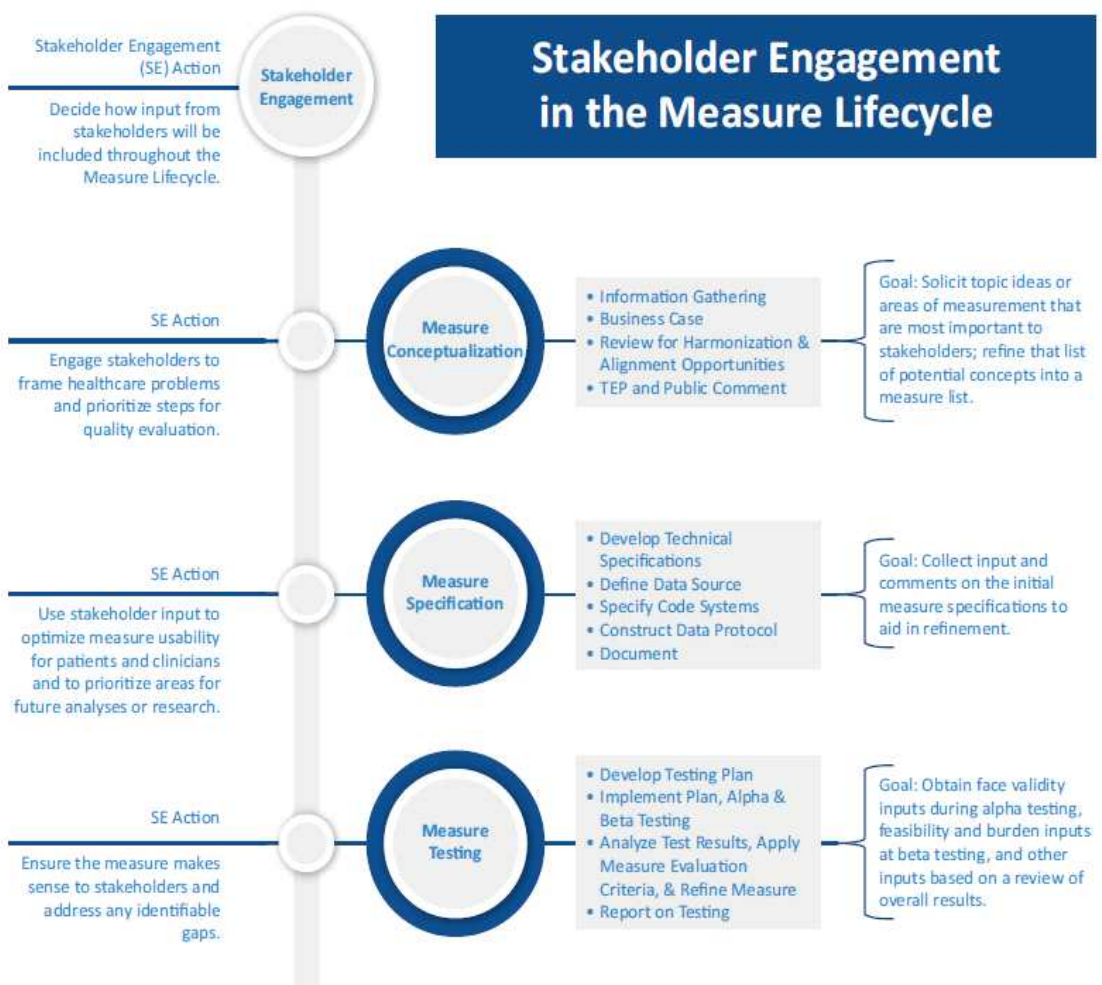
단계	목적	주요활동
개념화 (Conceptualization)	중요하고 의미 있는 측정에 대한 개념 또는 아이디어 목록을 생성	<ul style="list-style-type: none"> <li>• 환경스캔을 포함하여 정보수집</li> <li>• 사업사례를 개발</li> <li>• 기술자문위원(Technical Expert Panels, TEP) 검토</li> <li>• 공개의견요청</li> </ul>
구체화 (Specification)	어떻게 평가를 수집할 것인지, 일관적으로 안정적으로 효과적으로 실행시키기 위한 기술 지침을 개발	<ul style="list-style-type: none"> <li>• 대중의 의견을 고려하여 후보 측정목록을 개발</li> <li>• 측정 사양의 조화를 포함하여 정확한 기술적인 사양을 개발</li> <li>• 자료원 정의</li> <li>• 코드 시스템을 구체화</li> <li>• 데이터 프로토콜 구축</li> <li>• 측정을 문서화</li> </ul>
검증 (Testing)	실행 가능성을 평가하고 평가 기준의 측면에서 강점과 약점평가를 돕기 위한 실증적 근거를 확보	<ul style="list-style-type: none"> <li>• 검증작업 계획을 개발</li> <li>• 검증계획 실행</li> <li>• 측정 평가기준을 고려하여 검증 결과를 분석</li> <li>• 평가 기준을 측정에 적용하고 평가를 구체화</li> <li>• 측정 결과를 보고</li> </ul>
적용 (Implementation)	적용에 필요한 절차와 활동을 완료	<ul style="list-style-type: none"> <li>• 사전규정 프로세스와 최종규칙을 포함하여 측정 선택</li> </ul>
사용, 지속적인 평가와 유지 (Use, continuing evaluation and maintenance)	CMS 프로그램의 가치 향상에 도움이 되는지 모니터링하고 안정적 활용을 구축	<ul style="list-style-type: none"> <li>• 데이터 수집</li> <li>• 측정결과 보고</li> <li>• 환경 및 문헌 고찰</li> <li>• 측정재평가</li> <li>• 사업사례평가</li> <li>• 연간업데이트</li> <li>• 종합적인 재평가</li> <li>• 초기 유지 보수 검토</li> </ul>

\* TEP, Technical Expert Panels 기술전문가 패널을 의미

- 이해관계자는 평가의 전 과정에 참여함 (그림 16, 17 참조). 평가지표의 개념화 단계부터 통계 및 방법론 전문가는 물론 다양한 이해관계자가 참여를 함. 평가지표 개발의

각 단계에서 이해관계자의 참여와 관련 역할은 다음과 같음

- 개념화 단계에서는 이해관계자에게 가장 중요한 평가영역 또는 평가지표 결정에 필요한 의견을 수렴함 ([그림 16] 참고)
- 평가지표 구체화 단계에서는 환자와 임상 의사에게 의미 있고 유용한 지표를 파악하기 위한 자문과 자료원 또는 통계적 방법에 대한 전문가 의견을 모음
- 검증단계에서는 평가지표의 결과가 타당한, 적용가능한지, 평가부담이 어느 정도 되는지에 대한 이해관계자의 의견을 반영하여 평가지표를 수정 보완함



[그림 16] 평가 지표 개발 단계에서 이해관계자의 참여와 역할 ①



[그림 17] 평가 지표 개발 단계에서 이해관계자의 참여와 역할 ②

○ 이해관계자의 참여방식과 역할은 다음과 같음

- 기술전문가 패널(Technical Expert Panels, TEP) : 평가지표 개발 및 검증에 대한 의견, 평가결과 공개 또는 가치기반 의료에 미치는 영향이나 정책에 대한 의견
- 환자의 피드백 : 환자와의 비공식 대화, 구조화된 일대일 면담, 포커스 그룹 또는 환자 경험 조사를 통해 환자 관점에서 정보를 제공함
- 대중 의견 : 일반인들을 대상으로 정보를 공개하고 의견을 제시할 수 있도록 함. 다양한 이해관계자의 균형 잡힌 의견과 함께 투명한 절차를 개발하고 유지하도록 함
- 인터뷰 또는 업무 평가 : 현장 진료 임상과의 같은 주요 이해관계자로부터 심층적인 피드백을 획득하거나 업무 과정을 문서화함

## 2. 미국의 Agency for Healthcare Research and Quality's (AHRQ)

- 미국의 보건부(Department of Health and Human Services) 산하 기관으로 의료 질 향상에 관한 연구를 총괄하는 AHRQ(Agency for Healthcare Research and Quality)는 미국 메디케어의 질 성과를 매년 측정하고 그 결과를 의회에 보고함
- AHRQ는 지역 수준(area-level)과 병원 수준(Hospital-level)에서 질(quality)과

의료 이용(utilization)을 측정함<sup>11)</sup>

- 지역 수준(area-level)의 지표는 특정 인구에서 입원 중 또는 이후, 입원 중에 발생하는 잠재적으로 예방 가능한 합병증을 평가함
- 병원 수준(Hospital-level)의 지표들은 잠재적으로 예방 가능한 합병증이나 위해 사건(adverse event)을 평가함
- 그 외 다양한 측면의 질 지표를 운영하고 있음
  - 예방 질 지표(Prevention Quality Indicators, PQIs) : 지역사회 내에서 양질의 의료 또는 예방이나 건강증진 자원에 접근 가능했다면 피할 수 있었을 입원
  - 입원 질 지표(Inpatient Quality Indicators, IQIs) : 의학적 상태 및 수술 절차로 인한 입원환자 사망률을 포함하여 병원 내 진료의 질을 반영하는 지표
  - 환자안전 지표(Patient Safety Indicators, PSIs) : 잠재적으로 피할 수 있는 합병증이나 의인성 사건에 초점
  - 소아 질 지표(Pediatric Quality Indicators, PDIs)와 신생아 질 지표(Neonatal Quality Indicators, NQIs) : 위험에 처한 신생아와 아동에 대한 의료의 질과 접근성을 측정

○ AHRQ 평가지표 개발 시 우선순위는 다음의 단계를 거쳐 결정함<sup>12)</sup>

- 첫째, 기존의 평가지표를 검토하고
- 둘째, 다음의 영역을 검토함
  - 기존의 평가지표에 포함되지 않은 영역이 있는지
  - 기존의 평가지표에 포함되지 않은 관점이 있는지
  - 의료서비스 조정과 평가 부담에 초점을 맞추어 검토

11) AHRQ Quality Indicators TM, 2021

12) MeasureDevelopmentGapsand Recommendations. Content last reviewed July 2018. Agency for Healthcare Researchand Quality, Rockville, MD. <https://www.ahrq.gov/research/findings/final-reports/pcpaccountability/pcpacc3.html>



- 평가지표는 관점에 따라 환자, 의료제공자, 의료체제로 구분하고, 맵핑은 다음 <표 5>와 같은 형태로 이루어짐

<표 5> 평가의 관점과 평가지표 분류

구분	Measurement Perspective		
	Patient/Family	Health Care Professional(s)	System Representative(s)
<b>CARE COORDINATION ACTIVITIES</b>			
Establish accountability or negotiate responsibility			<input type="checkbox"/>
Communicate			<input type="checkbox"/>
interpersonal communication			
information transfer			■
Facilitate transitions			
Across settings			
As coordination needs change			
Assess needs and goals			■
Create a proactive plan of care			<input type="checkbox"/>
Monitor, follow up, and respond to change			■
Support self-management goals			■
Link to community resources			■
Align resources with patient and population needs			■
<b>BROAD APPROACHES POTENTIALLY RELATED TO CARE COORDINATION</b>			
Teamwork focused on coordination			■*
Health care home			
Care management			
Medication management			
Health IT-enabled coordination			■

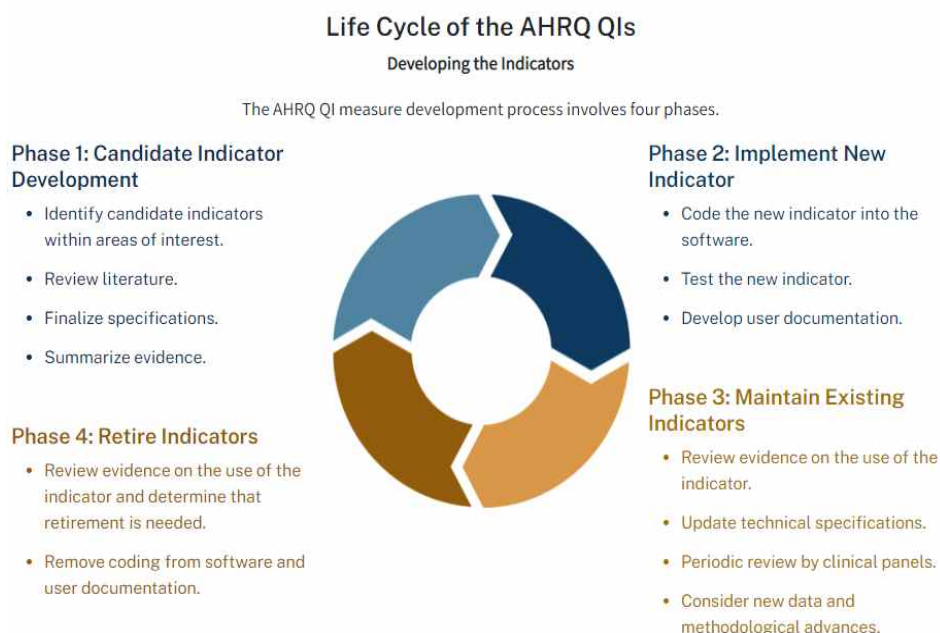
□ : ≥ 3 corresponding measure items

■ : 1-2 corresponding measure items

\* The use of a filled square for this measure indicates that it is a composite score.

- AHRQ의 질 평가 주기는 [그림 18]과 같은 4단계의 지표 개발(Measure development), 적용(implementation), 유지관리(Maintenance), 폐기(Retirement)의 프로세스를 거침<sup>13)</sup>





[그림 18] AHRQ 질 평가의 생애주기

- 1단계 : 후보지표를 개발하는 단계로 체계적 문헌 고찰을 통해 후보지표를 선정함
  - 2단계 : 신규지표를 개발 및 검증하고, 문서화함
  - 3단계 : 현행 지표를 유지하는 단계로 지표 사용의 근거를 검토하고, 기술적인 설명서 업데이트, 필요 시 임상 전문가 패널의 검토를 통한 근거 확충, 신규 데이터 및 방법론을 고려함
  - 4단계 : 문헌 검토를 통해 폐기해야 하는 지표를 퇴출함
- 1단계 후보지표 개발 단계에서는 후보지표를 확인하고 평가하는 단계임. 문헌 검토와 전문가 참여를 통해 후보지표를 선정하게 됨. 이후 선정된 후보지표에 대한 평가 및 패널심사, 위험도 보정, 실증분석 등의 단계를 거쳐 신규지표를 선정하게 됨
- 후보지표 확인
    - 문헌 검토 : 약 2.5개월이 소요되는 평가지표 개발의 첫 번째 과정으로, 후보지표를 확인하고 근거를 모으는 단계임. 사용 가능한 모든 근거를 검토하여 체계적 문헌

13) [https://qualityindicators.ahrq.gov/measures/qi\\_resources](https://qualityindicators.ahrq.gov/measures/qi_resources) accessed by 2022.03.26

고찰(systematic review) 수준에 준하는 문헌검토를 수행하게 됨. 잠재적 지표에 대한 논문, 자료원, 기존의 검증 노력, 위험도 보정에 관해 검토하여 얻은 정보는 이후 검증이나 위험도 보정 단계에서 유용하게 사용함

- 전문가 참여 : 후보지표를 개발하는 전 과정에 걸쳐 전문가가 참여하며, 이는 질 지표에 대한 과학적 수용성을 강화하는데 도움을 줌
- 후보지표의 선정 : 약 1개월 정도가 소요되는 과정으로, 문헌검토를 통해 얻은 정보를 바탕으로 근거 논문을 선택하고 이를 통해 지표의 구체적인 내용을 검토함. 요약양식(abstraction form)에는 자료원, 관심 결과, 대상 인구집단, 환자위험요인이나 형평성 관련 요인, 배제기준, 층화나 위험도 보정, 측정 수준(환자, 제공자, 지역수준), 성과측정 등에 대한 정보가 포함됨

#### - 후보지표 평가

- 후보지표와 기존 지표의 초기사양 정의 : 약 1개월 정도가 소요되는 과정으로 후보지표에 대해 측정 정의, 분자, 분모 및 관심영역을 바탕으로 한 구체화된 제외기준에 대해 근거 자료, 분석 결과, 현행의 평가지표와의 비교 검토, 전문가 의견 등을 기반으로 정의함
- 이 과정에서는 NQF의 중요성, 과학적 수용성, 유용, 실행가능성의 기준을 토대로 검토하게 됨 (<표 6> 참고)

<표 6> NQF 평가지표 검토 기준

기준	기준 설명	개발 및 검증 활동
중요성 (importance)	<ul style="list-style-type: none"> <li>• 개념이 측정하는데 중요한가?</li> <li>• 개선의 기회가 있는가?</li> </ul>	<ul style="list-style-type: none"> <li>• 구조화된 패널 검토</li> <li>• 문헌 검토 : 중요성, 유효, 변동성을 논의</li> <li>• 실증 분석 : 전체비율 및 변동</li> </ul>
과학적 수용성 (scientific acceptability)	<ul style="list-style-type: none"> <li>• 측정이 정확하게 정의되었는가?</li> <li>• 신뢰성(검증-재검증, 평가자 간)이 있는가?</li> <li>• 측정이 액면타당도, 구성타당도, 예측타당도를 입증하는가?</li> <li>• 체계적인 편향과 그 편향을 보정으로 해결 가능한가?</li> <li>• 성과의 의미 있는 차이를 감지하는가?</li> </ul>	<ul style="list-style-type: none"> <li>• 구조화된 패널 검토 : 사양이나 전반적인 유용성 검토</li> <li>• 문헌 검토 : 신뢰도, 기준 타당도, 구조 타당도, 잠재적 위험요인</li> <li>• 실증 분석 : 분율, 신뢰성, 잠재적 편향과 위험도 보정 평가, 지표의 관련성, 기준 타당도 또는 구성 타당도에 대한 연도별 차트 검토</li> </ul>

기준	기준 설명	개발 및 검증 활동
유용성(usability)	<ul style="list-style-type: none"> <li>• 해당 조치가 게이밍이나 부정적인 결과 대신 품질 개선을 조장하는가?</li> <li>• 유사한 측정과 일치하는가?</li> <li>• 측정이 의미 있고 이해가능하며 유용한가?</li> </ul>	<ul style="list-style-type: none"> <li>• 구조화된 패널 검토 : 불리한 결과나 전반적인 유용성을 평가</li> </ul>
실행가능성(feasibility)	<ul style="list-style-type: none"> <li>• 측정이 부담을 최소화하는가?</li> <li>• 자료수집과 적용이 실현가능한가?</li> </ul>	<ul style="list-style-type: none"> <li>• 구조화된 패널 검토</li> </ul>

\* 자료: Rockville. Quality Indicator Measure Development, Implementation, Maintenance, and Retirement. AHRQ 2011

- 패널 검토 : 약 3.5개월 정도 소요되는 과정으로 지표에 대한 임상적 타당성 정보를 제공하기 위해 수행함. 검토 계획 수립과 심사 실행 단계로 이루어지며, 주로 델파이나 Nominal Group process<sup>14)</sup>를 사용함
- 위험도 보정(risk adjustment) : 약 3개월 정도 소요되는 과정으로 초기 위험도 보정 정보, 기존의 위험도 보정방법을 검토하고 전문가들이 공변량 목록을 개발하는 단계임. 문헌 검토, 패널 검토, 초기 지표 검증을 통해 얻어진 정보를 활용하여 개발하게 됨
- 실증 분석 : 약 2개월 정도 소요되는 과정으로 후보지표와 관련된 바이어스, 정확도와 신뢰도, 지역 및 병원 간의 변이
- 최종 사양 정의 : 약 1개월 정도 소요되는 과정으로 문헌 검토, 패널 검토, 위험도 보정 및 실증 분석의 증거를 포함하여 각 후보지표의 장단점을 평가하고 타당성을 최대화하기 위한 권고안을 제시함
- 추천된 후보지표에 대한 근거 요약 : 약 2개월 정도 소요되는 과정으로 지표와 관련된 검토 및 결정을 용이하게 할 수 있도록 QI 개발 프로세스 전반에 걸쳐 수집된 근거를 요약 작성함
- AHRQ 검토 및 지표 결정 : 약 1개월 정도 소요되는 과정으로 후보지표에 대한 최종 사양과 요약된 근거를 토대로 향후 추가 개발이 필요한 부분이 있는지 확인함

14) Nominal Group Process : 합의에 도달하기 위한 소규모 그룹 토론의 구조화된 변형 형태

- 2단계 평가지표 적용 단계에서는 평가지표를 소프트웨어로 코딩하고 테스트하며, 사용자 안내서를 작성하는 단계임
  - 소프트웨어 내 지표 코딩 : 약 0.5개월 정도 소요되는 과정으로 소프트웨어 팀은 QI Module로 사용자에게 배포하기 위해 소프트웨어 안에 지표를 코딩함
  - 테스트 : 약 1개월 정도 소요되는 과정으로 새로 코딩된 QI Module의 정확성과 일관성을 보장하기 위해 소프트웨어 프로세스에 따라 테스트를 진행함
  - 사용자 문서 개발 : 약 1.5개월 정도 소요되는 과정으로 코딩 및 테스트 프로세스 전반에 걸쳐 평가지표에 대한 사양, 정의나 분자 및 분모, 사용자 가이드 등의 소프트웨어 문서 및 로그를 포함하는 사용자 문서를 개발함
  
- 3단계 평가지표 유지·관리 단계에서는 과학적 수용성을 보전하는 단계임. 평가지표가 과학적으로 수용가능하고 유용한 상태를 유지하기 위해서는 정기적인 유지·관리가 이루어져야 함. 이 단계에 해당하는 과정들은 모두 평가 지표 개발 전 과정에 걸쳐 지속적으로 이루어지게 됨
  - 근거(evidence) : QI 측정 개발팀은 평가지표와 관련된 문헌의 지속적인 검토를 통해 전문성을 유지함
  - 기술적 사양의 업데이트 : 위험도 보정에 사용되는 공변량이나 평가지표 코드, 회계 연도와 질병 그룹의 변화, 기준집단으로 사용되는 활용 가능한 비교 데이터 등을 반영하여 계속 최신의 것으로 업데이트하게 됨. 또한 지역 수준의 측정과 관련하여 주(county)의 새로운 인구 총 조사 데이터를 업데이트함
  - 패널 검토 : 근거 검토, 사용자 피드백, 코딩 변경으로 인해 지표에 대한 자세한 검토가 필요한 경우 정기적으로 임상 패널이 참여하여 검토를 수행함
  - NQF 제출 및 승인 : 개발된 모든 평가지표는 NQF 평가 기준을 충족하며, 제출 및 승인을 받으므로 정기 유지 관리 및 연간 검토 주기에 포함됨
  - 새롭게 사용 가능한 데이터 및 방법론에 대한 발전 : 더 나은 데이터와 방법론에 대해 지속적 검토 및 모니터링을 시행함

- 4단계 평가지표 폐기 단계는 관련 문헌 검토나 사용자의 피드백 등에 따라 평가지표가 더 이상 과학적으로 수용되지 않는 경우 폐기됨
  - 평가지표의 증거 폐기기준은 측정이 더 이상 과학적으로 수용되지 않는다는 새로운 근거가 있는 경우, 공급자에 의한 코딩 조작이나 조작의 결과로 지표측정이 바람직하지 않은 또는 예기치 않은 결과를 가져오는 경우에 해당함
  - 소프트웨어 내의 질 지표 코딩 제거 : 약 0.5개월 정도 소요되는 과정으로 폐기된 지표에 대한 코드를 제거한 소프트웨어를 사용자들에게 제공하는 것임
  - 테스트 : 약 0.5개월 정도 소요되는 과정으로 QI Module의 다른 지표들이 정확성과 일관성이 보장되는지 현재의 소프트웨어 테스트 프로세스에 따라서 검정함. 퇴출된 지표 제거로 인해 예상치 못한 결과가 발생하지 않았는지 확인하는 작업이 포함됨
  - 사용자 문서 : 약 0.5개월 정도 소요되는 과정으로 각 QI의 사양(정의, 분자 및 분모), 사용자 가이드(각 측정에 대한 증거 요약), SAS 및 WinQI 소프트웨어 문서 및 변경 로그에서 폐기된 지표를 제거하도록 업데이트함
- AHRQ에서 이루어지는 평가지표 개발 주기를 전체적으로 요약해보면 다음 <표 7>과 같음

&lt;표 7&gt; AHRQ 평가지표 개발 주기 단계 및 소요시간

개발 단계 및 업무	평균 완료 소요시간
<b>1단계 : 평가지표 측정 개발</b>	약 20개월 <sup>1</sup>
작업 1 : 후보지표 확인	
문헌 검토	2.5개월
전문가 참여	작업 전반에 걸쳐
후보지표 선정	1개월
작업 2 : 후보지표 평가	
후보지표 및 기존 평가지표의 초기 사양	1개월
패널 검토	3.5개월
위험도 보정	3개월
실증 분석	2개월
최종사양 확정	1개월
각 후보지표에 대한 근거 요약	2개월
후보지표에 대한 AHRQ 검토 및 결정	1개월

개발 단계 및 업무	평균 완료 소요시간
<b>2단계 : 적용</b>	약 1.5개월 <sup>1</sup>
평가지표를 소프트웨어로 코딩하기	0.5개월
테스트	1개월
사용자 문서	1.5개월
<b>3단계 : 평가지표 유지 관리</b>	변수 <sup>1</sup>
근거	작업 전반에 걸쳐
기술 사양 업데이트	1년에 한 번
패널 검토	필요에 따라
NQF 제출 및 유지관리	필요에 따라
새로 사용 가능한 데이터 및 방법론적 발전	필요에 따라
<b>4단계 : 평가지표 폐기</b>	변수 <sup>1</sup>
근거	작업 전반에 걸쳐
모듈에서 평가지표 제거	0.5개월
테스트	0.5개월
사용자 문서	0.5개월

<sup>1</sup> : 특정 작업이 서로 병렬로 실행되는 경우 평가지표 측정개발에 대한 대략적인 총 시간을 의미

\* 자료: Rockville. Quality Indicator Measure Development, Implementation, Maintenance, and Retirement. AHRQ 2011

○ 2021년 AHRQ QI Module에서 사용 중지된 지표는 다음과 같음<sup>15)</sup>

- 입원환자 질 지표(Inpatient Quality Indicators, IQIs)
  - 급성 심근경색증(AMI) 사망률(전원 사례 제외)
  - 제왕절개 후 질식 출산(VBAC) 비율
- 소아과 질 지표(Pediatric Quality Indicators, PDIs)
  - 신생아 사망률

15) AHRQTM, 2021

- 2019년에 AHRQ QI는 문헌검토, 환경 분석, 사용자 커뮤니티와 이해관계자의 피드백을 기반으로 5가지 폐기 사유를 제시하였음<sup>16)</sup>
  - 질 향상을 위한 질 지표 사용에 대해 근거 문헌이 제한된 경우
  - 사건의 희소성 : 일부 지표들은 측정하기에 사건이 너무 적어서 신뢰할만하지 못하거나 질 향상과 관련이 없게 됨
  - 의료기술의 발전 : 의료 제공 장소 및 방법 측면에서의 임상적 변화로 인해 질 지표의 타당성과 신뢰성에 영향을 미치는 경우
  - 개선을 위한 중요한 분석 작업 : 지표들 중 일부는 질 향상을 위한 목적으로 타당성 연구를 수행하거나 ICD-10-CM/PCM에 대한 새로운 사양을 검증하는 등의 중요한 분석 작업을 필요하게 되는 경우
  - 양적 지표 : 2018년 AHRQ의 전문가 그룹이 양적지표들이 더 이상 독립형 지표로서 유용하지 않다는 합의에 따라 양적지표는 폐기하게 됨

### 3. 미국의 National Quality Forum(NQF)

- NQF는 평가지표 자체를 개발하지 않고, National Committee for Quality Assurance(NCQA), CMS, 성과 향상을 위한 의사 컨소시엄, 의학 학술단체 등 다양한 공공과 민간 기관에서 개발한 평가지표에 대한 합의를 도출하고 승인하는 기관임. NQF의 승인은 엄격한 과학적 근거 기반 검토, 환자와 가족의 의견, 의료 산업 전반에 걸친 이해관계자들의 의견을 반영함
  - 5가지의 기준을 토대로 승인을 결정함 [1] 측정 및 보고의 중요성, [2] 과학적 수용성 (신뢰도와 타당도) [3] 실현 가능성, 즉 데이터에 대한 접근성과 가능한 부담이 적은 데이터의 활용. [4] 책무성과 성과 향상을 위한 활용성, 그리고 [5] 관련 지표 또는 기존 지표와 비교하여 최고의 지표를 선택
  - 지난 10년 동안 NQF는 보건 및 의료의 중요 분야에서 700개 이상의 평가 지표를 승인 함. NQF의 승인 지표는 연방 정부, 주정부, 민간 부문에서 신뢰하고, 활용되어 집

16) AHRQTM, 2021

- 제한된 자원 하에서 평가의 우선순위가 높은 영역을 파악하고, 해당 영역의 평가지표 개발과 승인을 통해 자원 활용의 효율성을 꾀하고자 함

○ HHS와의 계약에 따라 2021년 NQF는 다음과 같은 5가지 우선순위 영역을 선정함

- 성인 예방접종 : 성인 인구의 백신 접종률과 건강결과를 향상하기 위한 성과 평가
- 알츠하이머병 및 관련 치매 : 환자, 가족 구성원 및 간병인에게 중대한 영향을 미치는 복잡한 의학적, 사회적 요인에 초점
- 의료서비스 조정 : “지역사회 건강”의 맥락에서 일차의료와 지역사회 기반 서비스의 팀 접근과 조정에 초점
- 보건 인력 : 예방 및 의료 서비스 조정, 의료와 지역사회 기반 서비스 간의 연결, 인력 개발을 강조
- 사람 중심의 진료와 결과 : 환자결과 보고 등 환자에게 가장 중요한 평가지표를 우선 고려, 건강 정보 기술의 활용

○ NQF는 공식적인 합의 도출 절차(Consensus Development Process, CDP)를 사용하여 지표를 평가하고 승인함. CDP는 의료 산업 전반에 걸쳐 이해관계자 그룹의 의견과 토론을 허용하도록 설계됨

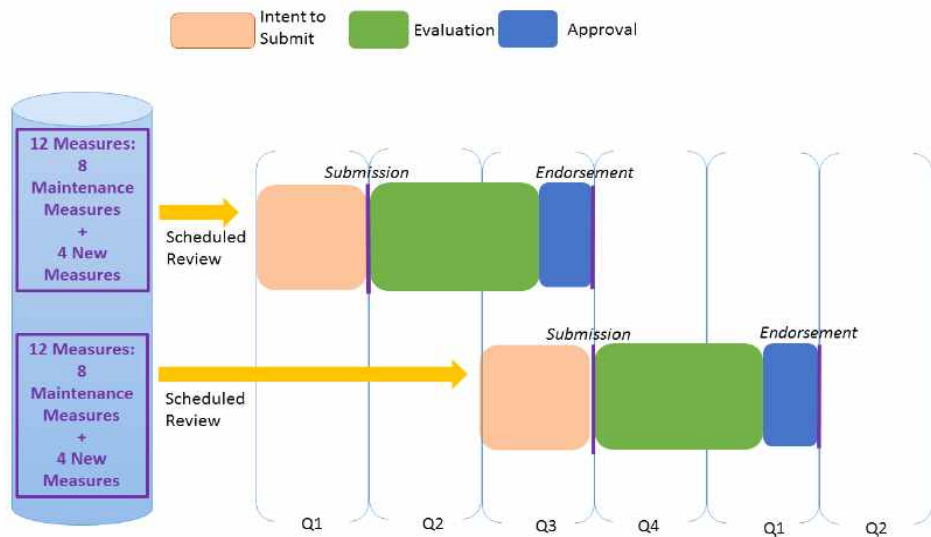
○ 200년 7월 승인된 CDP의 1판 공개 이후 CMS 및 이해관계자들로부터 CDP의 법안 제출에서 법안 승인까지의 소요 시간과 프로젝트에 사용하기 위한 지표 평가의 적시성, 대기시간(일부 경우 3년 이상)에 대한 개선 요구를 받음. 이에 NQF는 CMS와 협력하여 다음의 네 가지 문제점에 대해 해결하고자 하였음

- CMS, 개발자 및 NQF 간의 조정을 개선하여 지표 평가의 적시성을 용이하게 함
- 평가지표 제출 기회의 증가 및 적절한 기간 내에 지표 검토
- CDP의 주기 시간 단축
- CDP와 MAP(Measure Application Partnership) 프로세스 간의 정보의 흐름을 개선



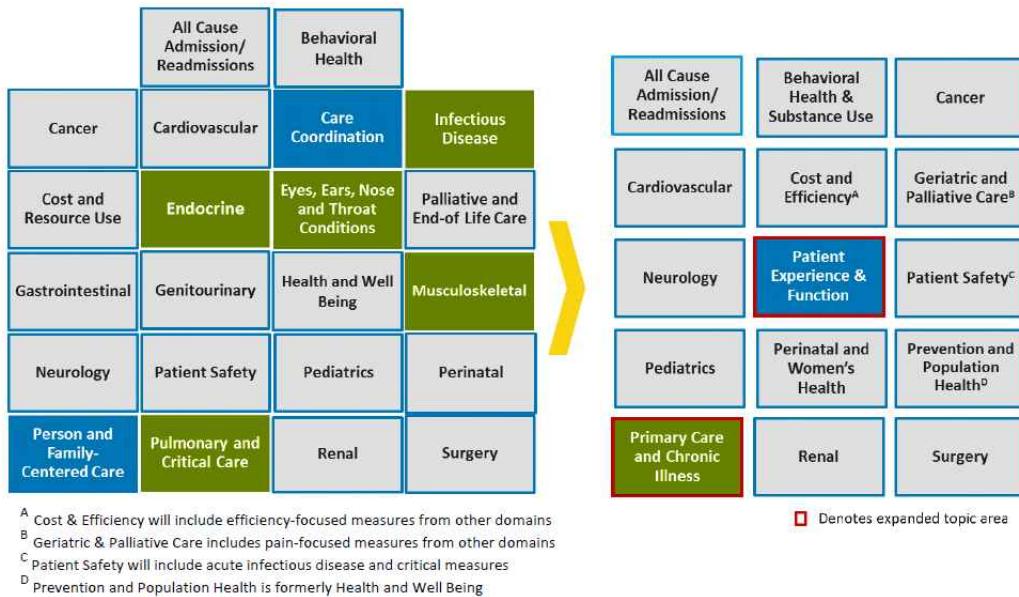
## ○ 평가지표 제출 기회의 확대 : 일정 및 빈도

- NQF는 CDP 일정에 따라 각 주제 영역에 대해 1년에 2번의 평가지표 제출 기회를 제공함 ([그림 19] 참고)
- 한 번의 주기에 상임위원회에서 검토하는 평가지표의 최대 개수는 12개로 제한하며, 유지관리 검토는 최대 8개의 기존 평가지표와 최대 4개의 새로운 평가지표를 대상으로 함



[그림 19] NQF의 평가지표 검토 주기 일정

- 평가지표 제출 기회가 증가할 것으로 예상됨에 따라 NQF는 [그림 20]과 같이 기존의 22개 평가지표 검토 주제 영역을 15개의 주제 영역으로 통합함
  - 평가영역 균형을 맞추고, 일부 임상 영역을 그룹화하고, 평가를 수행하는데 필요한 전문 지식을 갖춘 위원회를 통합하거나 수정하였음
    - 통합 위원회는 일차의료, 만성질환, 소아과, 노인병 및 완화의료와 같은 다양한 임상 영역을 반영하도록 설계함
  - 현재 NQF 프로세스에 따라 권장되는 주제 영역도 주기적으로 재평가됨

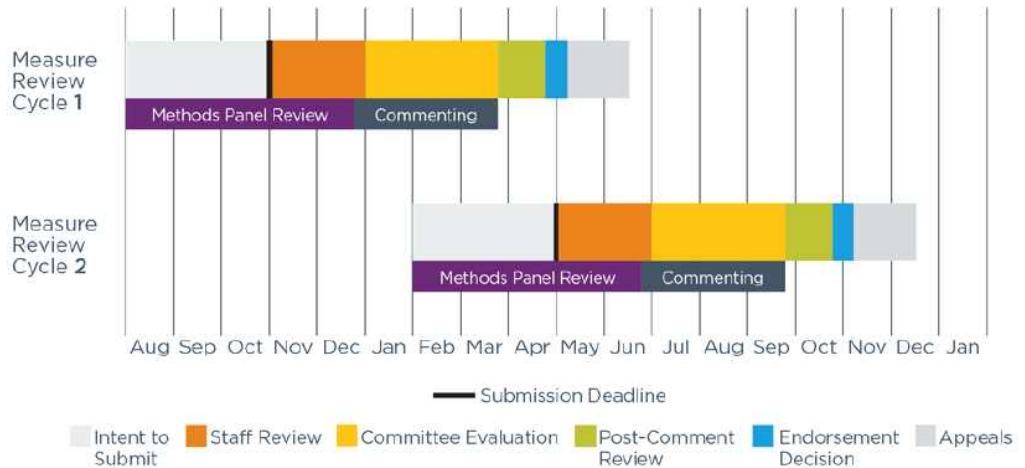


[그림 20] NQF의 평가지표 검토 주제 영역

### ○ 평가지표 주기 검토

- 새롭게 단축한 CDP는 [그림 21]과 같이 평가지표 제출 마감일에 시작하여 합의 표준 승인 위원회의 최종 승인 결정으로 끝남
- 평가 주기가 시작되기 전에 NQF는 모든 평가지표, 주제에 대해 1년에 두 번 평가지표 제출 마감일을 발표함
- 평가지표 제출 기한 동안 평가지표의 필요한 업데이트를 수행할 책임이 있는 모든 법안 관리자, 개발자는 승인 심의를 위해 새로운 평가지표를 제출할 수 있음
- 새로 제출된 평가지표 외에도 NQF가 승인한 평가지표는 약 3년마다 평가지표 기준에 따라 재평가됨
- 초기 평가지표 승인 또는 승인 유지를 위한 평가지표를 제출하려면 평가지표 담당자 및 개발자는 온라인 평가지표 제출 양식을 작성하거나 업데이트하고 제출 의향서(intent to Submit form)를 제출해야 함

### Consensus Development Process: Two Cycles Every Contract Year



[그림 21] NQF의 제안된 합의 개발 프로세스

#### ○ 평가지표 제출 의향

- 제출 의향은 모든 평가지표 관리자, 개발자가 승인 검토를 위해 평가지표를 제출할 준비가 되었음을 NQF에 통지하는 단계임. 제출 의향서에는 다음과 같은 9가지의 내용이 포함되어야 함
  - 제출 유형 : 유지관리 조치(현재 NQF에 기승인 받은 평가지표) 및 신규 평가지표 (NQF에 승인 받은 적 없는 경우). 유지관리 조치는 새로운 검증 데이터를 사용할 수 있는지를 표시해야 함
  - 평가지표 유형 : 분류(구조, 과정, 결과), 복합 지표(결과, 비용 및 자원 사용, 도구 기반 등)
  - 평가지표 제목 : 평가대상 및 무엇을 측정하기 위한 것인지에 대한 간결한 설명
  - 분석 수준 : 평가지표의 구체적인 내용과 검증 결과
  - 자료원 : 평가를 위해 데이터 출처
  - 평가지표 설명 : 점수 유형, 평가 대상, 대상 인구 또는 기간을 포함하는 평가지표에 대한 간략한 설명

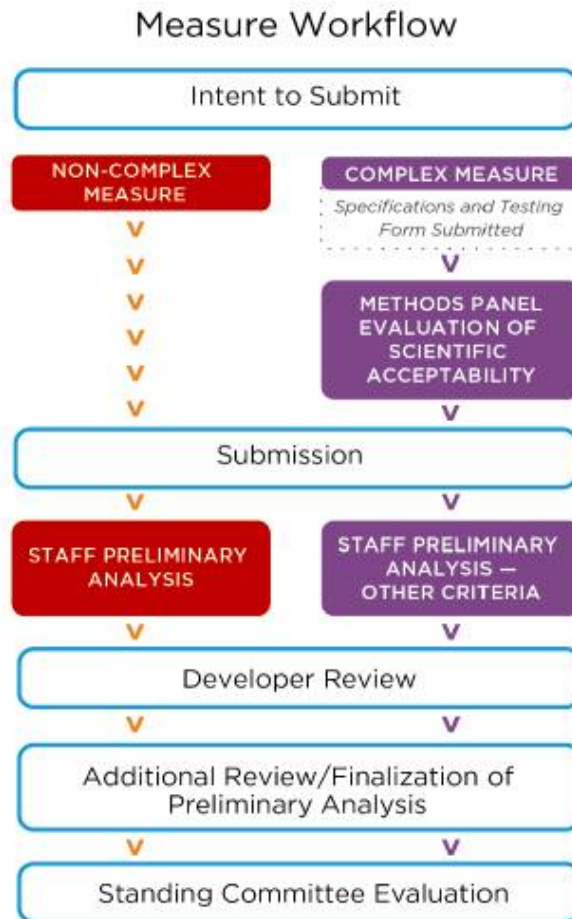
- 분자 설명 : 대상에 대한 설명
- 분모 설명 : 평가에 대한 설명
- 계획 제출 날짜 : 모든 검증이 완료되고 최종 제출이 예상되는 주기 및 연도
- 평가지표 관리자, 개발자는 다음 주기에서 위원회의 검토를 준비하기 위해 평가지표를 제출하려는 의향서를 평가지표 제출 마감일로부터 최소 3개월 전에 NQF에 통지해야 함
- 필요한 경우 평가지표 관리자 및 개발자는 해당기간 동안 기술 지원을 활용할 수 있으며, 평가지표 제출 의향서를 토대로 NQF는 다양한 주제 영역 내의 재평가가 필요한 유지관리 평가지표와 연계하여 계획을 수립할 수 있음

#### ○ 기술 검토 : NQF 및 과학적 방법 패널 검토

- NQF 평가지표의 신뢰성 및 타당도 분석, 복합지표에 대해 새로 구성된 외부 NQF 과학적 방법 패널에 의해 방법론을 검토함. 방법론 검토는 과학적 수용가능성(신뢰도 및 타당도)에 해당되는 것으로 반드시 충족해야 함
- 과학적 방법 패널은 15~25명의 통계학자, 역학자, 심리학자, 경제학자, 성과평가 방법 전문가, 형평성 또는 데이터 전문가로 구성됨. NQF의 상임 위원회 프로세스에 따라 해당 패널은 초기 2년 또는 3년 임기로 무작위 임명되며 선택적으로 3년 임기를 추가할 수 있음
- 선정된 패널을 대상으로 이해 상충 여부, 전문 지식 활용을 기반으로 평가지표를 배정함. 각 패널 구성원에게는 과학적 검토에 대한 표준 지침을 제공함
- 최소 3명의 패널 구성원이 외부 패널 검토를 받고, 각 평가지표는 독립적으로 검토됨. 검토자 사이에 평가지표의 신뢰성 및 타당성 통과 여부에 대한 불일치가 발생한 경우, 패널 공동 위원장이 해당 평가지표를 평가하고 최종 의사결정을 함

## ○ 복합 지표와 비복합 지표의 프로세스

- 평가지표는 제출된 의향서에 제공된 정보에 따라 복합 지표와 비복합 지표로 분류되며, 각 지표의 유형별로 평가되는 프로세스는 다음 [그림 22]와 같음



[그림 22] NQF의 평가지표 유형별 업무흐름도

- 복합 지표에 해당하는 유형은 다음과 같음
  - 중간 임상 결과를 포함한 결과 평가
  - 도구 기반 측정(예 : PRO-PMs)
  - 비용/자원 사용 평가

- 효율성 평가(자원 사용과 질의 개념을 결합한 개념)
- 복합 지표
  - 복합 지표의 경우에는 과학적 방법 패널의 평가가 필요할 수 있음. 비복합 지표(예: 구조, 과정 지표)는 NQF 내부 담당자가 과학적 허용 기준을 포함한 모든 지표 평가 기준에 대한 예비 분석을 완료하지만, 필요한 경우 외부 검토를 위해 과학적 방법 패널에 제출할 수 있음
  - 복합 지표와 비복합 지표 모두 예비 분석이 완료되면 검토를 위해 개발자에게 전달 함. 개발자가 평가지표의 예비 분석을 검토하는데 최대 2주가 소요되며, 검토 및 평가 결과에 동의하지 않는 경우 추가적인 설명을 제공할 수 있음
  - NQF 담당자나 과학적 방법 패널이 신뢰성 및 타당성에 대해 “낮음”, “불충분”으로 평가한 평가지표는 현재의 평가 주기에서는 제외되며, 다음의 평가 주기를 위해 추가적인 검증과 NQF 기술 지원을 위한 시간이 주어짐
  - NQF 담당자는 예비 분석 완료 이후 최종 제출 자료를 상임위원회에 보냄. 상임위원회는 평가지표의 구체성, 유효성에 대한 추가적인 고려사항 (예: 중증도 보정을 위한 변수 선택)에 대한 우려를 제기할 수 있으며, 기존의 평가 결과가 달라질 수 있음
  - NQF에서는 이러한 경우를 대비하기 위해 상임위원을 대상으로 지속적인 교육과 명확한 지침을 제공함
  - 상임위원회가 제출된 후보 평가지표에 대한 초기 검토가 완료되면, 대중 공개를 위해 NQF 웹사이트에 게시함
  - 보고서에는 승인 결정에 대한 위원회의 결과 요약, 검토된 각 평가지표에 대한 간략한 내용 요약, NQF 위원회 심의 내용(부록), 전체 평가지표의 내용(부록)이 포함됨
- 승인 여부 결정
  - 2017년 초부터 합의 기준 승인 위원회(CSAC, Consensus Standards Approval Committee)는 NQF 상임위원회에서 검토한 지표에 대해 최종 승인 결정을 내림
  - NQF 이사회에서 임명한 상임위원회인 CSAC는 CDP 개선, 지표 평가 기준 및 성과

측정의 새로운 문제와 관련하여 자문을 담당함

- 승인 결정을 위한 주요 기준은 평가지표 및 보고의 중요성(근거 기반, 과정과 결과의 연결, 변이, 형평성), 신뢰도와 타당도, 과학적 수용성, 실현 가능성, 유용성, 관련 지표와의 비교임

#### ○ 이의 신청

- CSAC의 승인 결정이 NQF 웹사이트를 통해 공개되면 30일의 이의 신청 기간이 시작됨. 해당 기간 동안 이해관계자는 승인된 법안에 대해 이의 제기를 할 수 있음

#### ○ 훈련 및 교육

- 위원, 지표 개발자, 내부 직원 등 CDP에 관련된 이해 관계자를 위해 다양한 교육을 실시하고 있음. 교육 내용은 CDP 소개, NQF 지표 평가 기준에 대한 이해, 지표 제출을 위한 모범 사례, CDP 전반에 걸친 역할 등의 내용을 포함함

## 4. 영국의 Shared commitment to quality

- ‘Shared commitment to quality’는 NHS와 보건의료의 전반적인 질을 관리 감독할 책임이 있는 국가 조직의 리더로 구성된 국가 질 위원회(National Quality Board, NQB)의 합의를 바탕으로, 국가 수준에서 질에 정의를 정하고, 질 개선에 필요한 지침서를 제공하기 위해 ‘Shared commitment to quality’를 작성함
- 국가 질 위원회에는 Department of Health, Public Health England, NHS England, Care Quality Commission, NHS Improvement, and the National Institute of Care Excellence 이 참여하여, 정부를 대신하여 조율된 리더십을 가지고 질 관리 지침을 제공함

- ‘The NHS Five Year Forward View<sup>17)</sup>’는 NHS의 미래에 대한 공유 비전을 제시하고, 비전 달성을 위해 관련한 모든 기관이 어떻게 구체적인 행동을 할 것인가에 대해 방법을 제시함
- 다른 국가와 비교하여 건강 결과가 나쁘거나, 변이가 큰 경우, 더 복잡한 의료서비스에 대한 필요가 증가하거나, 재정적인 부담이 큰 경우 질 개선을 위한 영역으로 선정함
- 환자중심 및 양질의 의료서비스를 제공하기 위해 소비자에게는 ‘안전’, ‘효과’, ‘긍정적인 경험’ 영역을, 제공자에게는 ‘투명, 협력, 학습과 개선에 기반한 리더십’, ‘지속 가능한 자원의 활용’을 통한 양질의 의료를 제공하는 것을 강조함 ([그림 23] 참고)
- 성별, 인종, 장애, 나이, 성적 지향, 종교, 신념, 성적체성, 임신과 출산 또는 결혼 여부나 시민권에 따른 건강 결과에 차이가 없어야 함(형평성)을 강조함



[그림 23] 의료의 질에 대한 관점의 공유  
(A single shared view of quality)

17) ‘The NHS Five Year Forward View’는 의료서비스를 제공하고 감독하는 Care Quality Commission, Public Health England and NHS Improvement 기관에서 2014년에 환자, 임상 의사, 전문가의 참여하에 인구집단의 건강, 의료의 질 및 자원 조달에서 확대되는 격차를 해소하기 위해 향후 5년 동안 보건의료 서비스 변화의 방향을 제시함



○ 환자가 경험하는 의료의 질을 향상시키기 위해 필요한 7가지 단계를 제시함 ([그림 24] 참고)

- ① 근거에 기반하여 우선순위와 명확한 방향을 제시
- ② 양질의 의료의 기준이 무엇인지를 명확하게 제시
- ③ 성과와 질 보고체계를 활용하여 질을 측정하고 공개
- ④ 질에 대한 인정과 보상
- ⑤ 질의 관리와 유지
- ⑥ 리더십, 경영, 조직문화, 질 향상과 유지를 위한 기술과 행동을 통한 역량 구축
- ⑦ 양질의 의료를 제공하기 위한 연구, 혁신 및 계획을 통한 선두 유지



[그림 24] 의료의 질 향상을 위한 7가지 단계

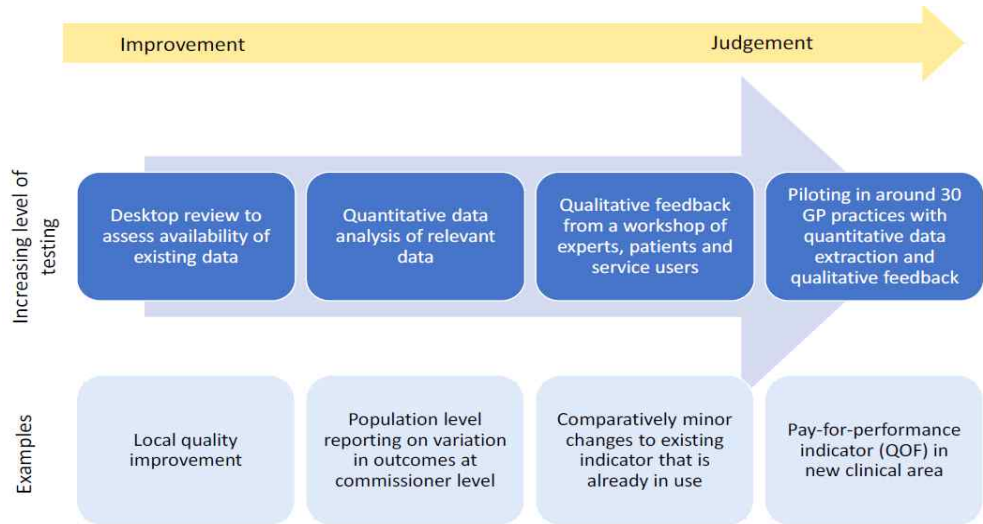
○ 다양한 자료를 검토하고, 재정적 균형을 고려하여 국가 수준의 우선순위와 전략적 방향을 설계함. 2017년에 발표된 ‘NHS Five Year Forward View’에서는 인구 고령화와 이에 따른 새로운 도전에 대해 적응하고, 지속해서 양질의 의료를 제공하기 위한 방향을 제시함

## 5. 영국의 National Institute for Health and Care Excellence (NICE)

- 영국의 NICE(National Institute for Health and Care Excellence)는 의사들이 가장 최신의 근거에 기반하여 최상의 진료를 제공하고 가장 효과적인 치료를 하도록 돕는 독립기구임
- NICE indicator는 결과를 향상할 수 있도록 의학적 근거, 의료서비스 제공 과정, 결과지표를 주로 측정하고 있음. 과정 지표는 NICE의 표준진료지침 등을 근거로 활용함
- NICE 평가지표는 평가대상을 의미하는 분모, 분모 중 특정 중재나 치료 결과를 의미하는 분자, 그리고 포함기준과 제외 및 예외 기준에 대한 설명으로 지표 요약이 구성됨
  - 그 외 지표의 목적, 지표의 의미, 자료원, 보고 기전과 기타 기술적인 세부 정보, 비용효과분석 및 자원영향(resource-impact) 분석 등에 관한 내용을 포함함
- NICE indicator programme의 주요 원칙은 다음과 같음
  - 포괄적인 근거 기반으로 지표를 개발
  - 독립적인 자문위원회
  - 전문가, 환자, 서비스이용자 및 돌봄 제공자의 의견을 반영
  - 투명한 프로세스 및 의사결정 과정
  - 공개적인 절차를 통해 효과적으로 확산
  - 정기적인 검토
- NICE indicator programme의 주요 활동은 다음과 같음
  - 안전하고 효과적이며 비용효과적인 의료나 서비스를 측정하기 위한 지표를 개발함
  - 의료나 경험에 대한 만족도와 삶의 질을 포함하여 지표들이 어떻게 결과를 개선하는데 활용될 수 있는지를 파악함
  - 이해관계자와 대중을 포함한 응답자들이 포괄적이고 개방적이며 투명한 협의 프로세

- 스에 참여할 수 있는 기회를 제공함
- 자원에 영향을 미치는 정도를 고려함
- 형평성에 미치는 영향을 고려함
- 지표의 지속가능성을 고려함
- 지표를 정기적으로 검토하고 최신의 내용으로 업데이트함
- 다른 국가의 의료의 질과 관련된 이니셔티브들과 연계함
- 지표들은 NHS digital이 주체가 되는 National Indicator Library에 보관되도록 함
- NICE indicator를 개발하기 위해서 자문위원회를 구성하는데 NICE하고는 무관한 다양한 전문 지식을 갖춘 위원들로 구성함
- 자문위원회는 GP 및 1,2차 전문의, 일반회원, 연구원, 사회복지 종사자, 공중보건 전문가 및 의료의 질 향상 전문가들로 구성되며, Indicator 자문 위원장과 위원들은 공개모집을 통해 NICE 정책에 따라 임명함
- 자문위원회는 예비평가 결과를 논의하고 형평성 영향평가와 비용효과적인 분석법 등에 대해 검토하고, NICE 목록 게재를 위한 지표를 추천하고, 기존의 지표를 검토하는 역할을 함
- NICE indicator 팀은 새로운 지표 개발을 위한 우선순위 프로세스의 관리, 지표 개발을 위한 우선순위 분야에 대한 지표를 자문위원회에게 보고하고, 잠재적 지표의 초안 작성, 공청회 결과를 분석하고 발표, 적절한 경우 비용효과분석 결과를 제공, 초기 지표 검증 결과를 운영, 위원회 회의록 준비, 새로운 지표와 함께 게재될 지침 생성, NICE가 발표한 지표 개발 프로세스와 방법이 합의된 일정과 기준에 따라 운영되도록 하는 업무를 담당함
- NICE는 자원 영향평가를 수행하여 잠재적 비용 및 절감액을 파악함
- NICE indicator는 다음의 개발과정을 거침
  - 1단계 지표개발 영역의 우선순위를 정하기 : 지표 개발은 NHS England와 Public Health England, 보건사회복지부(Department of Health and Social Care)와 북아일랜드 및 웨일즈의 위임된 행정부와 합의하여 우선순위를 반영함

- 국가의 새로운 정책이나 신규 또는 업데이트된 NICE 질 지침, 그리고 국가 성과의 신규 또는 업데이트된 보고서를 바탕으로 우선순위 영역을 설정함
- NICE 지표 개발은 최소 1년 단위로 논의되고, 필요한 경우 더 자주 수행함
- 2단계 지표자문위원회 심의 : 자문위원회는 진료행태의 변화, 임상 결과를 개선할 수 있는 기회, 타당성 평가 및 근거기반 지침의 내용을 검토하여 의견을 제안함
- 지표 초안 작성 : 지표의 구체적인 내용에 대한 초안 작성은 NICE와 NCCID(National Collaborating Centre for Infectious Diseases)에서 수행하는 반복적인 과정으로 NHS England, Public Health England, the Department of Health and Social Care와 지표 자문위원회의 전문가, 관련 NICE quality standard 및 지침 개발과 관련된 전문가로부터 자문을 받아 작성함
- 검증 : 모든 NICE 지표들은 타당성과 수용성을 평가하기 위한 검증을 거치게 됨. 검증은 주로 NCCID나 NHS Digital에서 수행하며 테스트 방식은 지표자문위원회와 합의함. 세부적인 테스트 과정은 총 4단계에 걸쳐 운영되며 [그림 25]와 같음. 1) 기존 데이터의 가용성을 평가하고, 2) 정량적 데이터 분석, 3) 전문가, 환자 및 서비스 이용자와의 워크숍을 통한 정성적 피드백, 4) GP 및 직원의 피드백을 통해 약 30개 GP 진료에서 시범운영하고, 지표의 타당성과 수용가능성을 평가함



[그림 25] 지표 검증 방법에 대한 예시

- **협의(consultation)** : NICE는 잠재적 새로운 지표에 대해 4주 동안 공개 협의를 통해 이해관계자와 환자단체와 전문가 그룹을 대상으로 의도하지 않은 결과의 위험은 없는지, 실행에 대한 장벽은 없는지, 형평성 차원의 영향은 없는지에 대한 의견을 수렴함
  - a) NICE는 공개 협의에 대해 이해관계자 및 응답자에게 사전에 이메일이나 NICE 홈페이지를 통해 공지하고 NICE 웹사이트를 통해 의견을 제출받음
  - b) **비용효과 분석** : 성과 지불 보상의 틀 안에 지표를 포함시키고자 하는 경우에는 비용효과분석을 고려해야 함
- **3단계 지표 자문위원회 검토** : 자문위원회는 테스트와 공개협의 중에 제출된 의견을 검토함
  - 자문위원회는 지표의 타당성을 평가하기 위해 사용하는 기준을 이용하여 NICE indicator로 진행할지에 대해 자문을 함. 타당성 평가를 위해 다음 <표 8>에서 제시한 중요성, 근거기반, 사양, 실현가능성, 수용성, 위험의 6가지의 기준을 사용함

〈표 8〉 타당성 평가에 사용되는 기준

기준	기준 설명
중요성 (importance)	<ul style="list-style-type: none"> <li>지표는 NHS England or Public Health England에 의해 파악된 우선순위 영역을 반영해야 함</li> <li>지표는 실제로 변이가 있는 것으로 알려진 영역과 관련이 있어야 함</li> <li>지표는 결과의 의미 있는 개선으로 이어질 수 있어야 함</li> <li>지표는 과잉치료나 과소치료를 나타내야 함</li> </ul>
근거기반 (evidence base)	<ul style="list-style-type: none"> <li>지표는 양질의 근거기반에서 파생되어야 함</li> <li>지표는 근거기반과 일치해야 함</li> </ul>
사양 (specification)	<ul style="list-style-type: none"> <li>지표는 분자, 분모 및 제외기준을 포함하여 지표를 구성하는데 필요한 구성 요소에 대한 정의가 있어야 함</li> <li>지표에는 최소인구 수준을 갖고 있어야 함</li> </ul>
실행가능성 (feasibility)	<ul style="list-style-type: none"> <li>지표는 반복측정 가능해야 함</li> <li>지표는 측정하고자 목적인 것을 측정할 수 있어야 함</li> <li>지표가 기존 데이터 분야나 추가의 자료수집에 대한 부담이 수용 가능해야 함</li> </ul>
수용성 (acceptability)	<ul style="list-style-type: none"> <li>지표는 청중의 통제 안에서 또는 기여할 수 있는 성과를 평가해야 함</li> <li>지표의 결과는 치료를 향상시키는데 사용할 수 있어야 함</li> </ul>
위험 (risk)	<ul style="list-style-type: none"> <li>지표에는 의도하지 않은 결과가 발생할 수 있는 허용 가능한 위험이 있음</li> </ul>

\* 자료: National Institute for Health and Care Excellence. Health and Social Care Directorate NICE indicator process guide. Dec. 2019

- 자원 영향평가 : 지표 개발 시 높은 수준의 자원 영향평가를 고려해야 함. 새로운 지표의 실행으로 인한 추가적인 활동과 이로 인한 자원영향을 고려해야 함
  - 임계값 설정 : NHS England와 북아일랜드와 웨일즈의 위임된 행정부는 지표가 성과지불보상의 틀 안으로 포함시키기 위해 개발되었다면 목적을 성취하게 되는 임계값을 설정해야 함. 임계값 설정은 일반적으로 기준 달성 수준, 최대 기대 달성 수준 및 현재 수준의 변이 등을 기반으로 함
- 4단계 NICE Guidance Executive 승인 : 권장된 지표가 NICE 목록으로의 포함을 위해서는 공표 전에 최종 승인을 위해 NICE Guidance Executive에게 제출해야 함

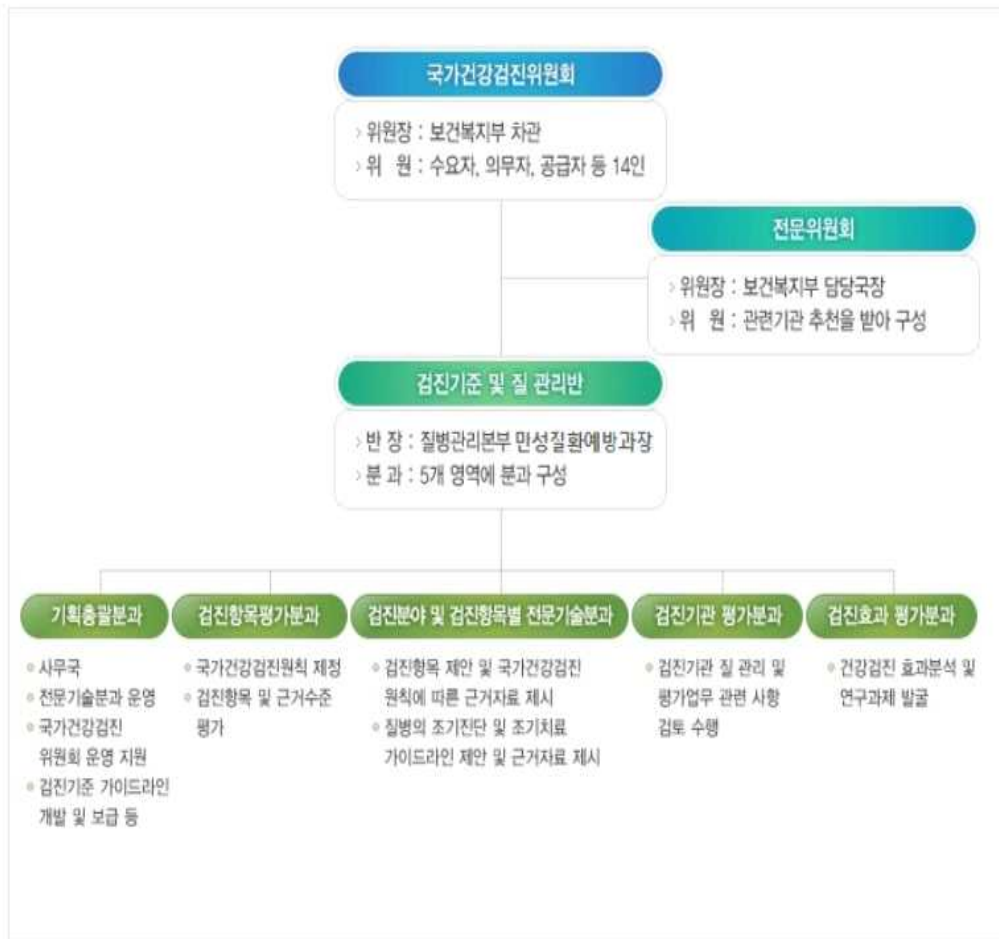
- NICE Guidance Executive는 합의된 절차와 방법을 따랐는지, 지표가 NICE Quality Standard나 NICE 지침 등의 근거와 일치하는지, 형평성의 향상에 기여하는지 등을 평가함
- 5단계 공표 : NICE indicator 목록에 있는 지표들은 웹사이트에서 목적, 활용, 지표가 실현 가능한 최소 인구 수준의 표시, 사용 가능한 자료원, 포함기준과 제외 및 예외 기준, 타당성 평가, 비용효과 분석(해당되는 경우), 자원 영향평가, 형평성 영향평가의 내용을 공개함
- 6단계 NICE 지표 검토 : NICE는 검토날짜에 도달했거나 지침이 업데이트되면 타당성 검증을 위해 지표를 검토함
  - 지표자문위원회는 지표가 계속 유효하거나 NICE 지표 목록에서 폐기되기 위해 필요한 활동을 조언함
  - 예상치 못한 위험이나 결과가 확인되거나 환자 안전에 위험이 될 수 있을 때 또한 지표를 검토함
  - 서비스를 사용하는 사람들을 위험에 빠뜨리거나 치료에 부정적인 영향을 미치거나 NICE의 명성을 손상하는 경우, 오류가 있는 경우에 수정 또는 변경함

## 6. 한국 국가건강검진 검진항목 선정

- 국가건강검진 검진(검사)항목을 선정하고 평가하는 절차는 [그림 26]과 같이 「건강검진기본법」 제13조(국가건강검진의 전담), 시행령 제9조(검진기준 및 질 관리반 운영) 시행에 따라 질병관리청에 설치된 ‘검진기준 및 질 관리반’을 중심으로 진행되고 있음<sup>18)</sup>
- 검진(검사)항목 선정 및 평가 절차는 다음과 같은 5단계로 진행됨
  - 1단계 : 질병관리청에서 전문 학회와 전문가에 대해 검진항목 요구도 조사
  - 2단계 : 검진항목의 의학적 근거자료 검토
  - 3단계 : 검진항목 평가분과에서 검진의 필요성과 근거수준 평가

18) <https://www.kdca.go.kr/contents.es?mid=a20303030200> accessed by 2022.08.03

- 4단계 : 국가건강검진위원회로의 안전 상정
- 5단계 : 국가건강검진 항목 추가가 결정되면 건강보험정책심의위원회에서 건강보험  
소요 재정을 심의



[그림 26] 국가건강검진 검진(검사)항목 선정 및 평가 담당 관련 조직

- 국가건강검진 관리 기반을 강화하기 위해 2022년부터는 담당기관의 역할과 소요  
기간을 조정하여 시행할 계획임. 또한 국가건강검진 개선을 위한 국민소통을 강화하고,  
검진항목 선정 및 평가 등 검진제도 개선 과정에서 정책 수요자 및 이해관계자 의견  
수렴을 위한 공식 채널을 마련('22년~)할 예정임



○ 국가건강검진 항목 선정의 기본원칙은 다음 <표 9>와 같음

<표 9> 국가건강검진 항목 선정의 기본원칙, 평가내용, 근거자료

구분	국가건강검진 원칙	평가내용	근거자료
1	중요한 건강문제일 것	1. 유병률 5% 이상 2. 목표 질환의 사망률 10만명 당 10명 이상 3. 질병부담(10만명 DALY의 1-35 순위, 의료비용 부담, 삶의 질)에 영향	<ul style="list-style-type: none"> <li>• 국민건강영양조사 질병별 유병률</li> <li>• 통계청 목표 질환의 사망률</li> <li>• 질병관리본부 연구결과보고서 'DALY 35순위 이상'</li> <li>• 공단 및 보사연의 직접 및 간접 비용</li> <li>• 국민건강영양조사 '삶의 질'</li> </ul>
2	조기에 발견하여 치료가 가능한 질병일 것		
2-1	질병을 조기에 진단할 수 있는 정확한 선별검사방법 및 검사 주기가 존재할 것	질병을 조기에 진단할 수 있는 정확한 선별검사방법 및 검사 주기가 존재할 것	<ul style="list-style-type: none"> <li>• 무증상자를 대상으로 한 검사방법의 민감도, 특이도, 우도비, 양성예측률, 음성예측률에 대한 국내외 저널</li> <li>• 검사주기가 제시되어 있는 국내외 저널</li> </ul>
2-2	조기발견에 따른 근거 있는 치료 및 관리방법이 있고 이용 가능할 것	1. 조기발견에 따른 효과적인 치료 방법이 존재할 것 2. 선별검진의 결과에 사후관리 방법이 존재할 것 3. 질병이 조기에 발견됨으로 인한 득이 있을 것	<ul style="list-style-type: none"> <li>• 조기발견에 따른 치료방법의 근거자료는 국내외 저널에서 SCIE 급 이상, 국내에서 한국연구재단 등재 또는 등재 후보된 저널</li> <li>• 선별검사 결과에 따른 사후관리 방법은 국내외 교과서 또는 임상 진료지침으로 하며, 이때 임상진료지침은 국내외 전문학회에서 모두 인정받은 것이어야 함</li> <li>• 검사항목에 따른 목표 질환의 조기발견으로 인한 이득에 대한 국내외 저널</li> </ul>
3	검진방법이 수용성이 있을 것		
3-1	국민이 쉽게 받아들일 수 있는 방법일 것	선별검사를 시행하는 것이 용이하고 윤리적으로 문제가 없을 것	<ul style="list-style-type: none"> <li>• 질병을 발견함에 따라 낙인의 가능성이 있는 질환(ex: 유전성 질환, 감염병, 성 매개 질환)일 경우 보호방안과 예방책 제시</li> <li>• 선별검사 방법에 대한 대상자의 순응도, 선호도 자료 제시</li> </ul>

구분	국가건강검진 원칙	평가내용	근거자료
3-2	인프라가 구축되어 있을 것 (검진기관 수, 시설, 장비, 인력 등)	인프라가 구축되어 있을 것 (검진 기관 수, 시설, 장비, 인력, 검진과정 및 결과에 대한 질 확보 등)	<ul style="list-style-type: none"> <li>지정된 검진기관의 수와 검진종류 분류</li> <li>검사 장비의 기보유수</li> <li>검사 시 필요한 면허 및 자격증 소유자</li> <li>검진의 인력, 시설, 장비, 내용에 대한 질 관리 및 평가방법의 지침 또는 정도 관리 실태 제출</li> </ul>
4	검진으로 인한 이득이 손해보다 클 것	1. 검진(검사) 시행이 건강증진에 미치는 긍정적 영향에 대한 근거 제시 가능 2. 선별검사 도구 및 방법의 안전성	<ul style="list-style-type: none"> <li>유병률과 사망률 감소 자료</li> <li>선별검사로 인한 부작용(ex: 통증, 발암요인, 감염의 위험) 관련 안전성 자료</li> <li>선별검사로 인한 합병증에 대한 자료</li> </ul>
5	비용 대비 효과가 있을 것	비용효과 관련 근거가 있을 것	<ul style="list-style-type: none"> <li>경제성 비용-효과분석 논문 및 연구결과</li> </ul>

\* SCIE(Science Citation Index Expanded) : 자연과학에 대한 이용색인 웹기반 데이터베이스로 과학기술분야에 대한 저널을 대상으로 학술적으로 기여도가 높은 학술지를 선정

\* 한국연구재단(학진) : 학술지 체계를 평가하고 게재논문의 학술적 가치와 성과 학술지의 전문성, 정확성 등에 대해 주제전문가 평가를 거쳐 학술지 게재 심사에 대한 평가까지 거쳐야 학술지로 선정되며 선정 이후에도 2년 주기로 체계 평가 및 패널 평가로 학술지 질적 수준을 관리

# 평가지표 개발의 문제점 및 개선사항



## 제3장

## 평가지표 개발의 문제점 및 개선사항

## 1. 의견 수렴 과정

- 평가지표 개발자 경험이 있는 임상학회 회원(3인), 보건의료 분야 의료 질 평가 전문가(2인), 건강보험심사평가원의 실무자(8인)를 대상으로 현재 영양급여 적정성 평가지표 개발 과정의 문제점과 개선방안에 대한 의견을 수렴함
- 건강보험심사평가원 실무자 대상 인터뷰는 총 3일에 걸쳐 진행되었으며 세부적인 인터뷰 진행 일시, 장소 및 참여자 명단은 다음 <표 10>과 같음

&lt;표 10&gt; 평가지표 개발 가이드라인 마련 연구 실무자 인터뷰 대상 명단 및 일정

날짜 및 시간		대상 부서		주요 담당 업무
05.31 (화)	11:00 ~ 12:00	정보수집체계개선반	정보수집체계개선부	• 환자경험평가, 중소병원평가
	10:00 ~ 11:00	평가실	평가2부	• 대상항목별 평가, 수혈, 천식, COPD
06.03 (금)	11:00 ~ 12:00		평가4부	• 대상항목별 평가, 관상동맥우회술, 우울증 외래, 의료급여 정신과
	13:00 ~ 14:00		평가관리부	• 대상항목별 평가, 질 향상 지원 사업, 치과근관치료, 고혈압
	14:00 ~ 15:00		평가1부	• 대상항목별 평가, 영양병원 평가, 중소병원 적정성 평가
	15:00 ~ 16:00		평가개발부	• 대상항목별 평가, 평가기획, 평가개발, 예비평가
06.07 (화)	09:30 ~ 10:30	의료체계개선실*	의료체계정책개발부	• 대상항목별 평가, 병원 표준화 사망비, 위험도 표준화, 재입원비 평가
	10:30 ~ 11:30	평가운영실	평가가치향상부	• 평가, 평가수행, 평가사후 관리

\* 의료체계정책개발부(제1청사 회의실)를 제외한 모든 인터뷰는 건강보험심사평가원 원주 제2청사 5층 영상회의실에서 진행됨

## 2. 주요 내용

- 평가지표 개발과 관련해서 평가 담당자가 개선이 필요하다고 제시한 의견은 다음과 같음
  - 첫째, 평가지표 개발을 진행하는 임상학회의 청구자료와 통계 및 중증도 보정에 대한 이해 부족
    - 평가지표 개발과 관련해서는 경험이 적은 경우가 많아서 개발 과정에 대한 전반적인 진행에 어려움이 있거나, 청구자료에 대한 이해가 부족한 경우가 있음
    - 특히 중증도 보정 모형의 경우 분석 과정이 복잡하고 통계적 방법과 객관적 의사결정이 필요하기 때문에 표준화된 방법론이 필요하고, 내외부의 상호 검증이 가능해야 함
    - 연구진에 통계 전문가가 포함되더라도 통계적인 방법론에 대한 일관된 지침이나 교육이 없기 때문에 개발 주체에 따라 중증도 보정 방법에 차이가 있고, 통계 방법에 대한 교차 검증이 어려움
  - 둘째, 평가지표 개발 연구진은 개발자이면서 동시에 평가를 받는 대상이 됨. 결과지표의 중증도 보정을 위한 자료제출의 부담감, 결과지표의 결과를 공개하는 경우 환자쏠림 심화, 중증도 보정의 한계를 이유로 결과지표 개발에 소극적인 경우가 많음
  - 셋째, 위탁 연구 수행 기간 안에 다양한 전문가와 이해관계자의 의견을 수렴에 필요한 시간을 확보하기가 어려움

## 3. 개선 방안

- 청구자료의 구성과 활용에 대한 체계적인 교육 제공
- 결과지표의 개발을 활성화하고 신뢰도를 높이기 위해 중증도 보정 방법에 대한 체계적인 지원을 제공함. 이를 위해 중증도 보정에 대한 교육 또는 심평원의 통계 전문가 또는 외부 전문가를 활용할 수 있도록 지원하는 것이 필요함

- 평가지표 산출의 신뢰성과 타당성 검증을 위한 투명한 검증 방안 마련 필요
- 구조·과정·결과가 지표가 균형적으로 개발될 수 있도록 평가지표 개발 가이드라인의 내용을 구성
  - 평가지표 개발 시 최소한 일정 수준 이상의 결과지표가 포함되도록 제안
  - 결과지표와 관련이 있다는 근거가 제시된 경우에만 구조 또는 과정의 핵심지표가 포함할 수 있도록 함
  - 신규 결과지표의 경우 일정 기간 동안 지역단위의 모니터링 지표로 도입 및 활용
    - 의료기관 : 중증도 보정을 위한 자료 수집 방안 마련 및 현황 파악 기간 확보
    - 심평원 : 중증도 보정을 위한 자료수집 방안과 기관 단위의 중증도 보정 모형의 정교화 기간으로 활용
- 평가지표 개발 가이드라인에 단계별 추진 내용을 구체적으로 기술하고, 필요 기간을 산출함으로써 정해진 기간 내에 평가지표가 효율적으로 개발 될 수 있는 방안을 모색할 필요가 있음





# 평가지표 개발 가이드라인



## 제4장

## 평가지표 개발 가이드라인

## 1. 의료의 질을 이해하기 위한 기본 개념

## 가. 바람직한 보건의료의 특성

- 바람직한 보건의료란 그 시대의 의학기술 발전 정도뿐만 아니라 보건의료가 가지고 있는 여러 가지 속성에 대하여 부여하는 상대적인 가치에 따라 내용이 달라질 수 있음 (〈표 11〉 참고). 바람직한 보건의료가 지향하여야 할 목표들이 서로 상충이 될 수도 있는데, 이 경우 어떠한 선택을 할 것인가에 대한 가치 판단이 필요함

〈표 11〉 의료의 질 구성요소

구성요소	IOM (2001)	OECD (2006)	Council of Europe (2014)	WHO (2018)
효과	○	○	○	○
안전	○	○	○	○
환자중심성	○	반응성	○	○
적합			○	
적시성	○			○
통합				○
효율	○		○	○
형평	○		○	○

\* 자료: 예방의학과 공중보건학(제4판). 대한예방의학회. 계축문화사: p1018재인용

## 1) 효과

- 효과(effectiveness)란 예방 서비스, 진단적 검사 또는 치료와 같은 어떠한 개입 조치가 다른 대안들에 비하여 더 나은 결과를 가져올 것인지에 대하여 체계적으로 수집한 근거를 바탕으로 의료를 제공하는 것을 의미함
- 과학적 근거를 기반으로 한 서비스를 제공한 결과 편익을 얻을 수 있는 사람에게만

서비스를 제공하는 것이 효과적인 의료임

- 1980년대 이후 활발하게 논의되어 온 근거기반 접근법(evidence-based approach)에서도 효능 또는 효과의 중요성을 강조함

## 2) 안전

- 보건의료는 효과가 있어야 할 뿐만 아니라, 이용자를 위협하게 하거나 손상을 일으키지 않아야 함
- 1999년 미국 의학한림원의 보고서를 통하여 환자안전(patient safety) 문제의 심각성이 널리 알려지게 되었으며, 국가 수준에서 환자안전 수준을 평가하기 위한 다양한 제도적 접근 방법이 도입되고 있음

## 3) 환자중심성

- 소비자 중심주의가 확대됨에 따라 환자가 단순한 진료 대상자가 아니라는 환자중심성(patient-centeredness)이 중요한 가치로 인식되고 있음
- 환자중심성이란 환자 개개인의 선호, 필요 및 가치를 존중하고 그에 반응하는 방식으로 보건의료가 제공되고, 환자의 가치에 따라 모든 임상적 결정이 이루어지도록 하는 것을 의미함
- 환자중심적인 진료에서는 ① 환자의 가치, 선호 및 가시화된 필요에 대한 존중, ② 진료의 조정 및 통합, ③ 정보, 의사소통 및 교육, ④ 신체적 안락함, ⑤ 정서적지지, ⑥ 가족의 참여가 중요하게 고려됨

## 4) 적시성

- 바람직한 보건의료를 제공하기 위해서는 대기시간을 단축하고 제공자와 이용자 모두에게 불필요한 지연을 감소시켜야 함
- 급성 심근경색증, 뇌졸중 등과 같이 적시에 적절한 개입 조치를 하지 않으면 생명에 심각한 위협이 되는 질환들에서는 이러한 적시성(timeliness)이 특히 중요함

## 5) 효율

- 효율(efficiency)은 보건의료 제공에 사용된 자원이나 시간의 단위당 산출, 효용, 또는 효과(예: 보건의료 제공량, 건강수준의 개선 등)를 의미함
- 장비, 소모품, 아이디어 및 에너지 낭비를 최소화하는 것이 효율적임
- 보건의료 비용이 급격하게 증가하면서, 비용 절감과 같이 제한된 자원을 효율적으로 사용하기 위한 노력이 다각도에서 진행되고 있음

## 6) 형평

- 형평(equity)은 성별, 민족, 지리적 위치 및 사회경제적 지위와 같은 개인 특성에 따라 의료의 질에 차이가 없는 것을 의미함
- 형평성을 벗어난 상태를 불형평(inequity) 또는 격차(disparity)라고 하고, 불형평은 윤리적 판단에 근거한다는 점에서, 관찰 결과에 근거하여 판단하는 불평등(inequality)과 개념적으로 구분됨
- 세계보건기구는 형평성을 “사회적, 경제적, 인구학적 또는 지리적으로 정의한 집단들 또는 하위집단들 사이에 피하거나 교정할 수 있는 체계적인 차이가 없는 것”으로 정의함. 보건의료 제공이 공정하게 되려면 개인에게 부여된 특성에 따라 보건의료 제공에 차별이 없어야 함

## 나. 의료의 질 정의와 개념

- 의료의 질이란, 현재 단계에서 주어진 의학지식의 조건 내에서 진료과정이 환자에게 기대되는 바람직한 진료결과의 확률을 높이는 한편, 원하지 않는 부정적인 결과의 확률을 낮추게 하는 정도 또는 현재 보편적으로 인정되는 의학지식 수준 내에서 진료제공 과정이 환자에 대한 부작용을 줄이고 최상의 진료결과를 성취하는 정도라고 정의할 수 있음

- 상대적으로 객관화시킬 수 있는 의학기술의 적용에서부터 의료이용자의 주관적인 만족도에 이르기까지 진료과정이 다양한 측면을 포함하기 때문에 의료서비스 질도 관점에 따라 다음과 같이 세 가지 측면으로 구분하여 정의할 수 있음

#### ① 전문가 중심의 정의

- 의사와 같은 전문가 입장에서의 정의라고 할 수 있으며, 건강에 대한 위험과 편익이 가장 적절하게 균형을 이룰 수 있는 진료과정을 양질의 의료라고 정의함
- 주로 의사의 의학적인 기술을 제공하는 능력에 관심을 둠

#### ② 의료이용자 중심의 정의

- 환자의 요구나 기대, 가치 등에 부응하는지를 기준으로 의료의 질을 보는 입장이며, 환자가 느끼는 서비스에 대한 만족도나 원하는 서비스를 받을 가능성이 클 때 양질의 의료라고 판단함
- 개인마다 사회적인 편익과 위험에 대한 가치가 달라서 환자의 요구와 사회적인 의무 사이에서 의사들이 도덕적으로 갈등해야 하는 경우가 생길 수 있음

#### ③ 사회적 정의

- 전체 인구가 얻을 수 있는 편익의 사회적 분포를 집단으로 파악하는 관점으로 더욱 많은 사람에게 편익이 돌아가는 경우를 양질의 의료라고 정의함
- 지역 또는 국가 수준의 건강과 진료비용에 관심을 둠

### 다. 의료의 질 향상 접근방안

- 질 향상 활동의 출발점으로서 의료의 질에 대한 다양한 측면을 평가하는 데 있어, 도나베디안이 제시한 구조, 과정, 결과의 3가지 측면에서 의료의 질을 평가하는 방식을 가장 많이 활용함

## ① 구조평가

- 의료서비스가 제공되는 시설이나 시설 여건, 환경과 소요되는 자원을 의미하며 인적, 물리적, 재정적인 자원에 포함됨
- 양질의 의료 인력에 의해 더욱 좋은 시설에서 진료를 받으면 진료 결과도 더 좋을 것으로 가정할 수 있으나, 모든 경우에 있어 이러한 가설이 성립되지는 않음
- 구조적 요인을 나타내는 평가지표는 상대적으로 변화의 여지가 적기 때문에 측정이 쉽고 안정적인 특성을 갖지만, 지속적인 평가지표로서는 부적당하고 대형 의료기관이나 교육병원에 대한 평가가 실제 수준보다 과대평가 될 가능성이 있음

## ② 과정평가

- 의료제공자가 실제로 환자를 진료하는 과정과 행위의 적절성을 평가하는 영역임
- 과정에 대한 평가지표는 구조지표와 비교하여 의료서비스 질과의 관련성이 높고, 진료의 결과를 평가하는 것보다 비용과 시간이 적게 소요된다는 장점이 있음. 반면, 과학적인 기준을 설정하는 것이 어려움
- 의료인의 환자관리 활동으로서 치료과정을 평가하는 것임. 의료제공자가 실제로 환자를 진료하는 과정과 행위의 적절성을 평가하는 것으로 적절한 약품의 사용 여부와 같은 기술적인 측면에서부터 환자들에게 바람직한 태도를 취하였는가와 같은 인간관계의 문제가 모두 포함될 수 있음
- 과정을 평가하기 위해 사용하는 방법으로는 의료자원 이용도 조사와 동료의사에 의한 심사, 진료비 심사청구, 의료감사 등이 해당함

## ③ 결과평가

- 환자에게 실제 제공된 의료서비스로 인해 현재 또는 미래의 건강상태가 어떻게 변화되었는지에 초점을 두는 접근 방법임
- 결과는 의료행위의 궁극적인 목표가 되며 건강수준이 향상된 결과는 반대로 양질의 의료 제공되었음을 보증한다는 전제에서 출발함. 그러나 건강상태의 변화가 의료

행위 이외의 다른 요소들에 의해서도 영향을 받을 수 있으므로, 결과를 해석할 때에는 건강에 영향을 미치는 다양한 요인들에 대한 고려가 필요함

- 결과지표에는 건강을 구성하는 요소로서 신체적 결과 외에도 사회적, 심리적 요소와 환자 만족도가 해당함. 사망률, 재입원율, 합병증률, 건강수준 등의 평가가 결과에 해당함

○ 구조, 과정, 결과 지표의 예는 다음 <표 12>와 같음

<표 12> 평가지표 유형 ①

구분	예시
구조	<ul style="list-style-type: none"> <li>• 의사 수</li> <li>• MRI 장비 수</li> <li>• 중환자 병상 수</li> </ul>
과정	<ul style="list-style-type: none"> <li>• 당뇨병 환자에서 안저 검사 시행률</li> <li>• 급성기 뇌졸중 환자에서 항혈전제 퇴원처방률</li> <li>• 중환자실에서 감염 관련 관리 지침(bundle) 수행 여부</li> </ul>
결과	
- 중간 결과	<ul style="list-style-type: none"> <li>• 당뇨환자에서 당화혈색소 수치</li> <li>• 고혈압 환자의 혈압</li> <li>• 고지혈증 환자에서 지질 수치</li> </ul>
- 최종 결과	<ul style="list-style-type: none"> <li>• 사망률</li> <li>• 재입원율</li> <li>• 기능 변화</li> <li>• 삶의 질</li> <li>• 환자 만족도 또는 환자 경험</li> </ul>

- 다양한 접근방법 중 어떤 방법을 선택할 것인가 하는 문제는 의료의 질을 평가할 때 실질적으로 당면하게 되는 주요 문제임. 질 평가의 목적이나, 평가에 드는 비용 및 시간, 측정의 용이성이나 가능성, 사회적인 윤리나 가치 기준 등을 고려해서 선택하게 됨



## 라. 의료환경의 변화와 의료의 질 평가

- 인공관절, 장기이식, 항암치료, MRI, 유전학을 이용한 희귀질환 치료와 같은 의학기술 발전으로 지난 40여 년 동안 의학은 매우 빠른 속도로 발전함. 그러나 인구의 고령화로 만성질환과 다양한 동반질환의 발생이 증가하고, 예방 또는 치료보다는 진단 중심의 신의료기술 발전함. 또한, 기존 의학기술을 대체하기 보다는 보완하는 기술이 개발됨에 따라 의료비가 큰 폭으로 증가함
- 보건의료에 투입되는 예산보다 의료에 대한 필요와 수요가 더 큰 폭으로 증가하고 있는 상황에서 ‘가치기반 의료(value-based healthcare)’라는 개념이 중요해지고 있음
  - 가치기반 의료는 의료공급자가 환자의 건강 결과에 기반하여 보상을 받는 의료전달 체계 모델로 의료의 질은 높이고 비용을 낮출 때 가치가 증가함
- 세계적으로 의료서비스가 양(volume) 중심에서 가치기반(value based) 방식으로 전환하는 추세에 있고 우리나라 또한 의료서비스에 대한 양적 확대 뿐 만 아니라 서비스의 질적 수준 향상에 대한 요구도 늘어나고 있음
  - 이러한 세계적 추세와 함께 기존의 평가에서 주로 다루던 의료서비스의 양보다는 국민건강이 얼마나 향상되었는지에 관심을 가지게 되었고 결과 중심의 평가가 과거 보다 강조되고 있음
  - 최근 요양급여 적정성 평가도 가치기반 보건의료 제공과 환자 중심의 의료서비스 제공을 위한 질 관리 방안을 강조하고 있음<sup>19)</sup>

## 2. 평가지표 개발의 기본원칙

### 가. 평가지표의 개념과 유형

- 평가지표는 의료평가 대상별로 평가의 목적과 범위, 평가의 영향력 등을 고려하여 요양급여의 적정성을 측정하기 위하여 개발·제시하는 기준을 의미함

19) 평가발전위원회, 2021

- 2001년 요양급여 적정성 평가 도입 이후 질 지표, 임상 질 지표, 또는 평가지표라는 용어로 혼용되어 활용됨

○ 평가지표는 구조지표, 과정지표, 결과지표로 구분할 수 있음

- 구조, 과정, 결과 측면으로 구분하기도 하지만, 의료의 질 구성요소의 측면에서 효율성 지표, 효과지표, 형평성지표 등으로 구분하기도 함. 또는, 환자가 직접 보고한 것을 바탕으로 평가하는 지표, 몇 개 이상의 지표값을 하나의 값으로 산출하는 복합 (composite)지표 등으로도 분류할 수 있음 (〈표 13〉 참고)

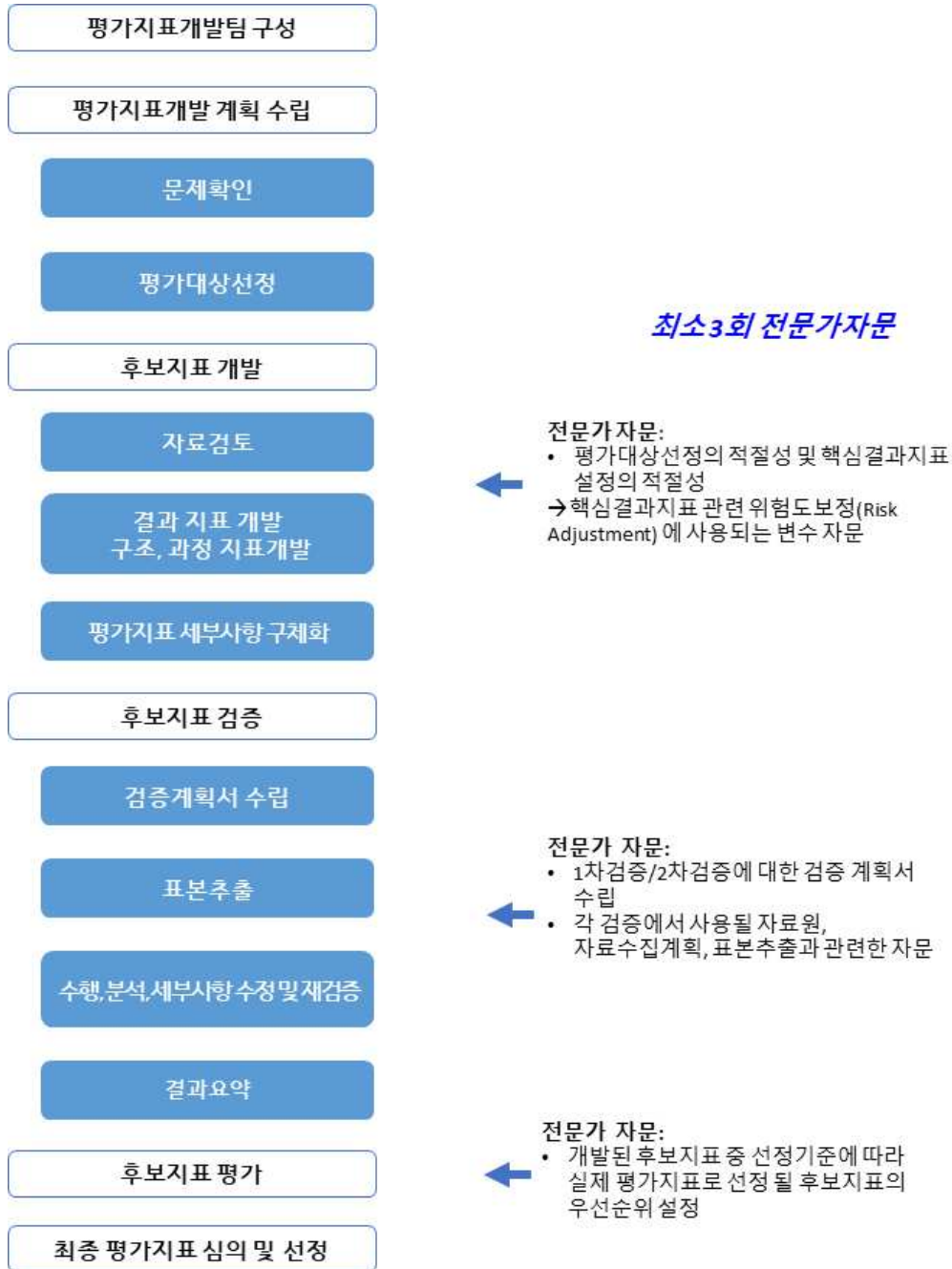
〈표 13〉 평가지표 유형 ②

질 평가 유형	설명 및 목적	예시
복합 (Composite)	두 개 이상의 개별 평가지표를 조합하여 하나의 결과지표로 활용하는 것	환자안전과 위해사건을 복합지표로 평가
비용/자원(사용) (Cost/Resource)	의료서비스가 적용된 인구집단이나 사건에 사용된 비용이나 자원을 측정 자원 사용은 정의된 의료시스템 자원의 사용 빈도를 계산함. 일부는 사용된 자원의 각 단위당 적용된 달러의 양으로 표현	수혜자당 메디케어 지출
효율성 (Efficiency)	특정 수준의 건강결과와 관련된 의료 비용을 측정	비심장 저위험 수술에서 수술 전 위험 평가를 위한 심장 영상 촬영
결과 (Outcome)	환자의 건강 상태를 측정 또는 건강 상태의 변화를 측정	암 환자의 30일 이내 계획되지 않은 재입원
결과 기반의 환자보고 (Patient reported outcome-based)	책임 있는 의료기관에 대해 집계된 환자가 보고한 결과측정을 바탕으로 성과를 측정함	목 손상을 가진 환자에서 기능적 상태 변화
절차 (Process)	좋은 의료를 제공하기 위해 따라야 하는 단계를 측정	성인 주요 우울증 : 자살위험평가
구조 (Structure)	의료를 제공하기 위한 역량과 관련된 의료기관 조직이나 임상직의 특성을 측정	방사선과 : 유방조영술 스크리닝을 위한 시스템을 알림

## 나. 개발한 평가지표에 대한 평가 기준

- 중요성 : 평가지표가 질 향상의 효과가 크거나 영향력이 큰 영역을 대상으로 하는가?
  - 의료의 질에 미치는 영향(예: 안전성, 적시성, 효과성, 효율성, 형평성, 환자중심성)이 큰가?
  - 의료의 질을 개선할 수 있는 실질적인 방안이 있는 내용을 평가하는가?
  - 근거에 기반하고 있는가?
  - 적정 수준 이하이거나, 변이가 크거나, 형평성에 문제가 있는가?
- 실행가능성 : 평가지표를 산출하는데 필요한 구체적인 내용이 작성되었나? 자료를 지속적으로 확보할 수 있는가? 평가 부담을 최소화 할 수 있는 방안을 고려했는가?
- 과학적 수용성 : 타당성과 신뢰성이 확보되었는가?
  - 활용 가능성 : 환자, 평가기관, 정책입안자, 다양한 이해관계자에게 의미 있고, 활용할 가능성이 높은가?
- 다른 평가지표와 연계 : 평가영역이나 지표의 중복이나 누락이 없는가? 다양한 평가 관련 프로그램의 지표 또는 기존의 평가지표와 비교 검토했는가?
- 결과지표인가? 결과와 관련성이 큰 지표인가?
- 평가를 수행함에 있어 의도되지 않은 결과를 확인하고 최소화할 수 있는 방안을 마련했는가? 과다이용이나 과소이용 등 의도하지 않은 결과를 초래할 수 있나? 이를 최소화할 수 있는 방안을 고려했나?

### 3. 평가지표 개발의 단계별 지침



[그림 27] 평가지표 개발 프로세스

## 가. 평가지표 개발 계획 수립

### 1) 단계 설명

- 평가지표 개발 계획서는 실제 평가가 이루어지기 전, 평가지표를 개발하게 된 배경과 목적을 명확히 기술하고 지표개발 과정에 대해 상세하게 기술하는 단계임
- 이 단계에서는 크게 2개의 세부단계로 구분하여 기술하여야 함
  - 문제 확인 (Identify Problem)
  - 평가 대상 선정 (Select Target)

### 2) 세부 단계

#### 가) 문제 확인 (Identify Problem)

- 평가지표를 개발하게 된 배경 또는 확인된 의료의 질 문제를 확인하고 개발된 평가지표의 평가를 통해 궁극적으로 얻고자 하는 개선 효과 및 목적을 명시하는 단계임. 문제 확인 단계에서는 다음의 내용이 포함되어야 함
  - 평가지표 개발 배경 및 의료의 질 문제점
  - 지표 개발의 목적
  - 활용계획
  - 평가로 인해 발생될 이익이나 위해 정도 검토
- 평가지표 개발 배경 및 지표 개발의 목적 : 평가지표를 개발하게 된 배경과 드러난 문제점들을 기술하고, 평가지표 개발 후 평가를 통해 궁극적으로 달성하고자 하는 개선 효과 및 결과 등 구체적인 목적을 명시해야 함
  - 현재의 의료의 질 문제 크기가 어느 정도인지 현황을 파악하기 위해 관련 자료를 간략하게 분석하여 결과를 산출하거나 기존의 문서화된 자료를 수집하여 현재의 의료의 질의 문제점을 양적으로 알 수 있도록 기술함
  - 현재 제공되는 의료서비스의 양(high volume), 비용적인 측면(cost), 의료기관 간 변이 수준(variation) 등의 측면에서 의료의 질 문제의 크기를 진단함

- 활용계획 : 실제 평가의 결과를 어떻게 활용하게 될 것인지 활용계획에 대한 구체적인 내용을 포함해야 함. 활용목적에 따라 지표 개발의 방향성이 달라질 수도 있기 때문에 계획 수립단계에서 이 부분에 대한 검토가 이루어져야 함
- 결과 활용계획 예시) 가감지급, 수가연계, 지정평가제도, 국민공개, 요양기관 결과제공, 비재정적 인센티브 등 (〈표 14〉 참조)

〈표 14〉 평가지표 유형 ③

연번	분류 항목	정 의
1	가감지급	평가 결과가 우수한 의료기관에는 경제적 보상을 주고, 개선이 필요한 의료기관은 경제적 불이익 적용
2	수가연계	평가 결과에 따라 차등 적용하여 평가 결과가 좋은 의료기관은 경제적으로 지원 하는 제도
3	지정평가제도	평가 결과가 상급종합병원, 전문병원 등 병원지정을 위한 기준으로 활용
4	인증평가제도	평가 결과가 인증 또는 인증평가를 위한 기준으로 활용
5	요양기관 결과제공	의료기관의 자율적 질 개선을 위해 평가결과를 의료기관에 제공
6	국민공개	국민이 스스로 의료기관을 선택할 수 있도록 평가결과 공개
7	비재정적 인센티브	평가결과가 좋은 의료기관을 대상으로 비재정적 보상을 제공 * 예) 그린처방의원 선정 시 현지조사 면제 등 비재정적 지원 제공

\* ‘음영’은 분류체계 내 세부분류 추가(예정) 항목임

\* 자료: 평가지표 분류체계 및 세부 분류항목. 건강보험심사평가원

- 평가로 인해 발생될 이익이나 위해<sup>20)</sup> 정도 검토 : 평가지표 개발을 통해 실제 평가를 수행할 경우, 발생할 잠재적 이익과 위해의 정도를 검토하여 기술해야 함

20) 평가로 인해 발생될 잠재적 부작용을 의미하며 임상적 측면과 정책적 측면, 2가지 측면에서 검토해볼 수 있음. 예를 들어 급성기 의료기관에서 집환별 재원일수를 평가하는 경우 임상적 측면에서 조기퇴원이나 외래방문이 증가하는 등 풍선효과로 인한 위해가 발생 가능함. 의료기관별 사망률을 평가하여 공개하는 경우, 정책적 측면에서 의료기관 쏠림현상이 강화되는 등의 위해가 발생할 수 있음

〈표 15〉 문제 확인 예시

### 문제 확인 및 현황파악

(현행의 의료의 질과 관련된 문제점을 확인하고, 문제점의 내용을 정량적으로 표현하는 단계임)

- 인공관절치환술은 일반적으로 삶의 질을 개선하기 위해 퇴행성관절염이 진행된 환자에서 통증 완화 및 관절기능의 회복을 위해 사용할 수 있는 효과적인 치료법 중의 하나임 (윤재연, 2020). 최근 우리나라 노인인구의 증가로 관절치환술이 해마다 지속적으로 증가하는 추세임. 건강보험심사평가원의 통계자료에 따르면 퇴행성관절염으로 병원을 내원하는 환자가 2019년에만 무려 400만 명이며, 이 중 무릎 인공관절수술은 꾸준히 증가해 2019년에는 무려 11만 건으로 이 중 60세 이상이 95% 이상 차지했으며, 수술건수는 지난 5년간 꾸준히 증가했음. 이러한 절차가 삶의 질을 극적으로 향상시키기는 하지만 비용이 많이 지출됨



- 요양기관 종별 수술건수의 분포는 2009년을 기준으로 병원 54.5%, 종합병원 26.7%, 상급종합병원 14.2%, 의원 4.5%순이었음. 요양기관 종별 수술 진료비는 병원은 662억4천9백만원(2005년)에서 2076억5천6백만원(2009년)으로 213.4% 증가하고, 종합병원은 523억7천1백만원(2005년)에서 1090억8백만원(2009년)으로 108.1%증가했으며 상급종합병원은 341억8천5백만원(2005년)에서 592억 5천9백만원(2009년)으로 73.3%나 늘어났음

구분		2005년	2006년	2007년	2008년	2009년
수술건수 (건)	전 체	25,414	32,847	40,678	48,896	54,097
	상급종합병원	5,242	5,636	6,270	7,112	7,706
	종합병원	7,426	9,849	11,112	13,123	14,466
	병원	10,286	14,516	20,101	23,685	29,494
	의원	2,460	2,846	3,195	2,976	2,431
진료비 (백만원)	전 체	167,052	227,935	284,966	331,864	391,293
	상급종합병원	34,185	39,636	44,828	52,376	59,259
	종합병원	52,371	71,832	81,768	97,917	109,008
	병원	66,249	99,099	138,379	162,971	207,656
	의원	14,247	17,368	19,991	18,600	15,369
건당진료 비(천원)	전 체	6,573	6,939	7,005	7,077	7,233
	상급종합병원	6,521	7,033	7,150	7,364	7,690
	종합병원	7,052	7,293	7,358	7,461	7,536
	병원	6,440	6,827	6,884	6,881	7,041
	의원	5,791	6,103	6,257	6,250	6,322

- 계획되지 않은 재입원은 환자의 삶의 질을 저하시키며, 국가적으로는 건강보험의 불필요한 비용을 지불하게 함. 재입원은 진료의 질에 영향을 받는 결과이며 환자에게 중요한 결과임. 퇴원 계획, 약물 조정 및 외래 환자 치료 조정과 같은 입원 및 외래 환자 치료의 질을 반영하는 의료기관의 노력은 재입원율을 줄일 수 있음. 2017년 기준 인공관절 실시 환자 57,872명 중, 총 4,042명(7.0%)이 수술 후 재입원하였음

구 분	첫 입원 환자(A)	재입원 환자(B)	재입원율(B/A)	동일 의료기관 재입원환자(C)	동일 의료기관 재입원율(C/B)
계	57,872	4,042	7.0	2,258	55.9
상급종합	6,920	130	1.9	116	89.2
종합병원	18,449	1,221	6.6	885	72.5
병원	31,022	2,286	7.4	1,234	54.0
의원	1,481	320	21.6	23	7.2

주) 표에 제시된 수치들 중, 재입원, 재치환술 수치의 경우 조사기간 전의 입원기록이 없는 환자 제외되어 개수의 신뢰도가 매우 낮음



## 지표개발 목적

(지표개발의 목적에 대해 기술하는 단계임)

- THA/TKA 수술 환자의 재입원율을 평가함으로써 의료기관이 재입원을 감소시키기 위한 노력을 기울이도록 함
- 대중에게 평가의 결과를 공개함으로써 환자의 의료기관 선택에 대한 정보를 제공하기 위함
- 의료소비자에 대한 투명성을 높이고 재입원 위험을 줄임으로써 의료비용을 감소시키기 위함

## 활용계획

(지표개발의 결과를 어떻게 활용할 것인지에 대해 기술하는 단계)

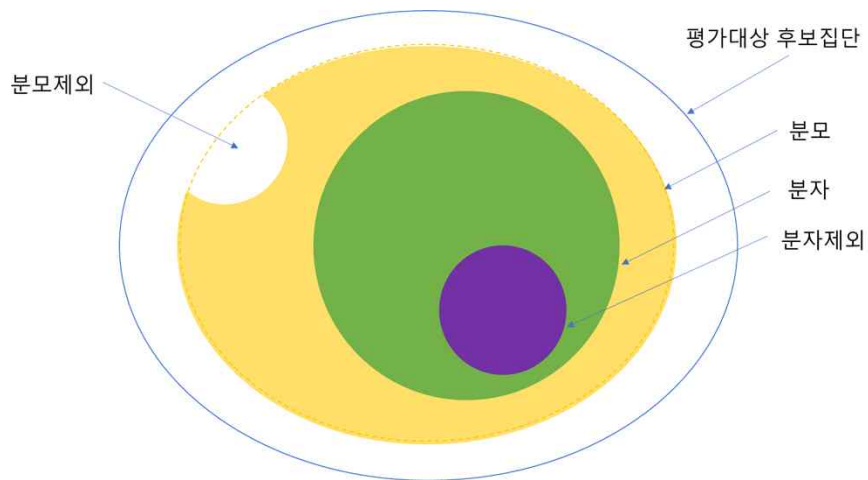
- 의료기관별 성과보상 지불을 위한 근거자료로 활용
- 대중공개를 통해 의료소비자에게 투명한 정보 제공

## 평가로 인해 발생될 이익이나 위해 정도 검토

- 재입원 감소를 통해 의료기관의 의료의 질을 개선하고 정보를 의료소비자에게 투명하게 공개함. 또한, 재입원 감소로 불필요한 의료비용을 감소시킬 수 있음
- 수술 후 재입원이 필요하나 평가로 인해 입원 의료서비스가 필요한 환자를 외래에서 치료하는 풍선효과가 발생할 수 있음. 또한, 재입원을 피하기 위해 적절한 입원 의료서비스를 제공하지 않음으로 합병증이 발생하거나 극단적인 경우 사망률이 증가할 수 있음

## 나) 평가 대상 선정 (Select target)

- 질환, 환자군, 중재(intervention) 등 평가지표를 개발해야 할 대상을 명확하게 제시하는 단계임
- 명확한 평가대상을 설정하기 위해 대상군에 포함시켜야 할 기준과 제외기준을 명시해야 함. 평가대상을 어디까지 확장하고 어디까지 제한할 것인지에 대한 기준을 마련해야 함
- 제시기준의 예) 연령, 성별, 상병코드(ICD-10), 환자가 가진 특정 상태(예: 혈압수준, 혈당수준, 폐기능수준 등), 1차/2차/3차 의료기관, 응급/외래/입원 등



[그림 28] 평가대상선정에 대한 모식도

〈표 16〉 평가 대상 선정 예시

**평가후보집단***(평가대상이 되는 일반적인 군에 대해 기술)*

- 계획된 TKA/THA 수술을 받은 환자

**포함기준***(분모에 포함시켜야 하는 inclusion criteria에 대해 기술)*

- ICD-10 code에서 주진단과 2개의 부진단에서 M16(고관절증), M17(무릎관절증)로 진단받은 65세 이상 환자
- 무릎 인공관절치환술 수가코드 N2072, N2077과 고관절 인공관절치환술 수가코드 N0711, N2070로 청구된 건
- 1,2의 기준을 모두 충족해야 함

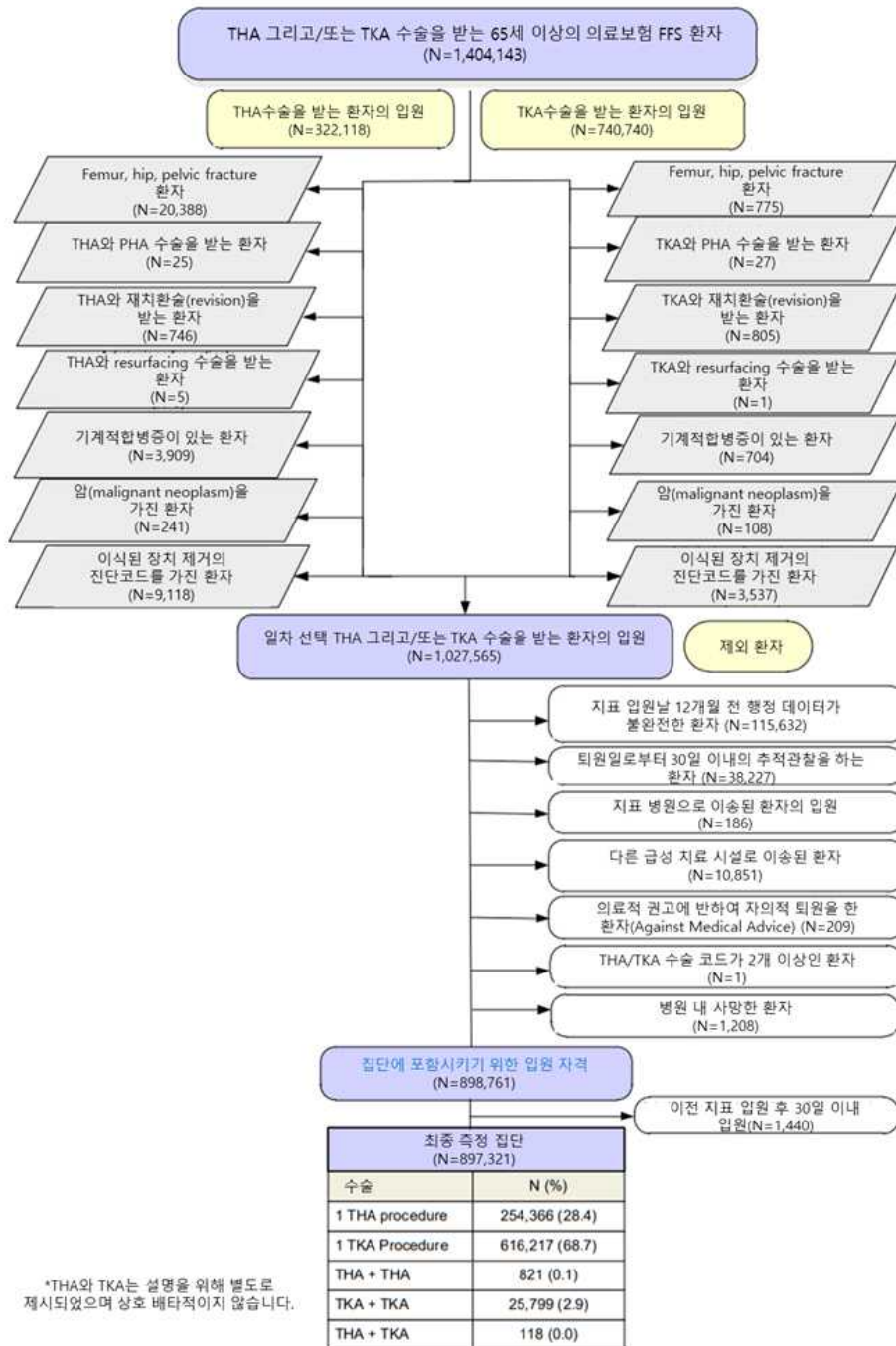
**제외기준***(분모에서 제외시켜야 하는 exclusion criteria에 대해 기술)*

- 퇴원 시 주진단 코드가 femur, hip, pelvic fracture로 청구된 건
  - [근거] 이러한 주진단으로 입원한 환자들에서 고관절 인공관절치환술은 계획된 수술이 아니며, 사망 및 합병증과 재입원에 고위험군으로 분류되는 대표적인 환자군임
- 부분치환술 수술을 받은 환자
  - [근거] 보통 부분치환술을 받는 환자들은 고관절 골절 시에 수행하며, 부분치환술을 받은 환자군들의 특징과 완전치환술을 받는 환자군들의 위험도가 다름 (예를 들어, 부분치환술을 받는 사람들이 나이가 더 많거나 완전치환술을 하기에 취약하거나 훨씬 더 동반질환을 앓고 있을 수 있기 때문)
- 재치환술 (revision) 수술을 받는 환자
  - [근거] 재치환술을 받는 환자들은 소수에 의료기관에서만 시술될 수 있어 의료기관별 수술건수에 차이가 많이 나고, 재수술을 받은 환자는 사망이나 합병증 및 재입원의 위험이 더 큼

- resurfacing 수술을 받는 환자
  - [근거] resurfacing 수술은 관절의 관절연골 표면만을 포함하는 수술이어서 이 수술을 받는 환자들은 상대적으로 젊거나 건강한 환자라는 특성이 있음
- 퇴원 시 주진단에 기계적 합병증의 진단코드가 있는 환자
  - [근거] 퇴원 시 주진단에 기계적 합병증의 진단코드가 있는 환자는 입원 시에 이미 합병증을 갖고 있었거나 이전 수술의 결과로 발생한 합병증일 가능성이 큼. 이러한 환자는 더 복잡한 인공치환술을 받았을 가능성이 있고, 합병증 특히 기계적인 합병증 발생 위험이 증가할 수 있음
- 퇴원 시 주진단 코드에 암(malignant neoplasm)을 가진 환자
  - [근거] 암이 있는 환자는 재입원의 위험을 증가시킴
- 인공치환술의 인공삽입물을 제거하는 수술을 받은 환자
  - [근거] 임상적으로 훨씬 더 복잡할 수 있기 때문
    - femur, hip, pelvic fracture에 해당하는 주진단 코드를 표로 정리하여 보여 줘야 함. 또한, partial arthroplasty, revision procedures, resurfacing, 인공삽입물 제거술과 관련된 수가코드 표를 제시. 기계적 합병증(mechanical complication)이나 암과 관련된 주진단 코드 제시
- 인공치환술을 위해 입원한 날부터 최소 12개월 전의 의료이용 자료가 확인되지 않는 환자
  - [근거] 수술 전 동반질환에 대해 위험도 보정에 필요한 적절한 자료를 확보하기 위함
- 퇴원 후 적어도 30일 동안 추적관찰 가능하지 않은 자
  - [근거] 30일 재입원에 대해 결과를 확인하기 위해 퇴원 후 30일 동안 추적관찰 가능해야 함
- 다른 병원으로부터 전원되어 인공관절치환술을 시행한 자
  - [근거] 환자가 다른 의료기관에서 해당 수술을 위해 해당 병원으로 이송되는 경우, 수술이 계획된 수술이 아닐 수 있음

- 인공관절치환술을 시행하고 다른 급성기 의료기관으로 전원된 환자
  - [근거] 해당 의료기관에서 수술을 받은 후 다른 비급성기 의료기관으로 전원한 경우
- 의료적 권고에 반하여 자의적 퇴원을 한 환자(Against Medical Advice)
  - [근거] 의료기관에서 최선의 진료를 제공할 수 있는 기회가 없음
- 해당 입원동안 2번 이상의 인공관절치환술을 시행한 자
  - [근거] 한 번의 입원동안 두 번 이상의 인공관절치환술을 받았을 가능성도 있지만 한번 수술을 받은 사람들의 특성과 다를 수 있으며, 코딩의 오류일 가능성도 있음
- 해당 입원기간 동안 사망한 자
  - [근거] 해당 입원기간 동안 사망했다면 재입원은 일어날 수 없음

Figure 1. Measure Cohort (2008-2010 Medicare FFS Patients)



[그림 29] 평가대상자 선정 흐름도 예시

## 나. 후보지표 개발

### 1) 단계 설명

- 평가지표 개발 계획수립 단계에서 선정한 핵심 결과지표와 핵심 결과지표를 달성 하는데 직접적인 영향을 주는 구조(structure), 과정(process)의 후보지표를 확인하고 지표별 세부사항(specification)을 기술하는 단계임. 구조와 과정지표는 다양한 문헌과 자료에 기반하여 결과와 관련이 있는 경우 후보지표로 선정함
- 이 단계에서는 크게 3개의 개발업무를 수행해야 함
  - 자료 검토
  - 핵심 결과지표 개발
  - 구조, 과정지표 개발
  - 평가지표 세부사항 구체화

### 2) 세부 단계

#### 가) 자료 검토

- 국내외 평가대상과 관련된 임상진료지침이나 이미 사용되고 있는 평가지표들 관련 문헌을 검토함. 임상진료지침은 공신력이 있는 곳에서 인증한 권고안을 검색함. 주요 국내외 데이터베이스는 다음과 같음
  - National Center for Complementary and Integrative Health
  - National Guideline Clearinghouse (AHRQ)
  - Institute of Medicine's Report Clinical Practice Guideline We Can Trust (IOM)
  - Travelers' Health: Clinician Resources (CDC)
  - CMS Measures Inventory Tool (CMIT)
  - National Quality Forum
  - 임상진료지침 정보센터 (Korean Medical Guideline Information Center)

- BMJ Best Practice  
(<https://bestpractice.bmj.com/info/evidence-information>)
- 임상진료지침이 아닌 일반 문헌을 검색하는 경우, 한국보건의료연구원에서 체계적 문헌고찰 시 사용하는 기준을 그대로 적용함
- 국외 문헌 검색 시, 검색의 편향성을 최소화하기 위해 영국 SIGN(Scottish Intercollegiate Guidelines Network)의 기준을 따르며, 이 기준에서는 MEDLINE, EMBASE 및 Cochrane Library 데이터베이스를 검색할 것을 요구하고 있음
- 국내 전자 데이터베이스의 경우, KoreaMed, Kmbase, KISS, Riss4U, KISTI 등을 핵심 데이터베이스로 검색함
- 문헌 검토 시 유의사항
  - 동료심사(peer-reviewed)를 거쳐 게재가 되는 학술지의 문헌을 검토하는 것이 원칙임
    - 정부기관의 자료나 보고서, 회색문헌(Grey literature)도 활용 가능함. 단, 출처를 명확하게 제시해야 함
  - 가능한 최신(5년 이내)에 발표된 문헌을 대상으로 할 것
  - 과거 10년 이내 수집된 자료를 분석한 문헌을 대상으로 할 것

#### 나) 핵심 결과지표 개발 (Establish Outcome)

- 확인된 의료의 질 문제를 개선했을 경우 또는 개발된 평가지표로 평가를 수행했을 경우 궁극적으로 예상되는 성과를 핵심 결과지표로서 선정하는 단계임
- 핵심 결과지표 선정 단계에 포함되어야 할 내용
  - 2개 이상의 핵심 결과지표 선정
  - 전문가 자문 : 결과지표 선정 및 선정의 타당성, 위험도 보정
  - 위험도 보정에 필요한 자료원 검토



- 2개 이상의 핵심 결과지표 선정 : 핵심 결과지표를 선정할 때에는 후보지표를 검증하는 단계에서 탈락이 될 수도 있으므로 적어도 2개 이상의 후보 핵심 결과지표를 고려하고 이중, 가장 우선순위가 높은 결과지표를 선정함
- 결과지표로는 사망률, 생존률, 합병증 발생률, 증상 및 기능개선 등과 같은 환자의 건강 상태를 측정하거나 자원 활용 대비 비용과 같은 효율성 지표나 삶의 질, 만족도 등과 같은 결과지표를 사용할 수 있음
- 적절한 결과지표는 임상 및 정책과 관련성이 있고, 활용 가능성이 높아야 함
- 예방 가능하고 절대 발생해서는 안 되는 심각한 의료 오류가 아닌 한 통계 분석이 가능하도록 충분한 빈도로 발생해야 함
- 최종 결과지표를 개발하기 어려운 경우에는 중간 결과(intermediate outcome)를 핵심 결과지표로 선정할 수 있음

〈표 17〉 결과지표의 예시

구분	예시
결과	
- 중간 결과	<ul style="list-style-type: none"> <li>• 당뇨병환자에서 당화혈색소 수치</li> <li>• 고혈압 환자의 혈압</li> <li>• 고지혈증 환자에서 지질 수치</li> </ul>
- 최종 결과	<ul style="list-style-type: none"> <li>• 사망률</li> <li>• 재입원율</li> <li>• 기능 변화</li> <li>• 삶의 질</li> <li>• 환자 만족도 또는 환자 경험</li> </ul>

- 핵심 결과지표를 선정할 때에는 핵심 결과지표에 영향을 미치는 요인(혼란변수)을 검토하고 이러한 요인을 제거하거나 최소화할 수 있는 방법에 대해서도 모색해야 함
- 예시) 평가대상의 동질성 확보, 짝짓기, 통계적 보정 방법 등
- 전문가 자문 : 핵심 결과지표 선정의 타당성과 위험도 보정과 관련하여 전문가의 자문을 받아야 함

- 전문가 구성 : 평가대상과 관련된 임상전문가, 통계전문가, 연구방법론 전문가, 정부 관계자, 보건의료 정책전문가, 청구자료, 임상자료 및 의료기관 행정자료 활용과 관련된 전문가 등으로 구성
- 전문가 구성(안)
  - 1(안) : 평가지표를 개발하는 개발팀 자체에서 전문가 자문을 위한 패널을 구성하는 방안
  - 2(안) : 건강보험심사평가원에서 구축해 놓은 전문가 패널을 활용하는 방안
  - 1(안)과 2(안)을 모두 장단점이 있는 만큼, 2개의 안을 모두 활용하는 것을 권고함
- 핵심 결과지표 선정의 타당성 : 의료의 질 문제점을 개선하고 문제점 개선을 통해 결과지표 값이 긍정적으로 변화하는 역학적 인과관계가 분명한지를 검토하여 타당성을 검증함

〈표 18〉 전문가 구성(안)의 장·단점

구분	1안	2안
장점	<ul style="list-style-type: none"> <li>• 평가지표 개발 계획 수립의 시간을 단축</li> <li>• 진행상의 유연성 확보</li> </ul>	<ul style="list-style-type: none"> <li>• 객관성과 투명성 확보</li> <li>• 다양한 의견을 수렴</li> </ul>
단점	<ul style="list-style-type: none"> <li>• 개발팀 친화적인 자문위원들로 구성되어 제한된 의견 수렴</li> <li>• 2안에 비해 상대적으로 객관성 확보의 어려움</li> </ul>	<ul style="list-style-type: none"> <li>• 평가지표 개발의 시간 지연</li> </ul>

- 위험도 보정 : 위험도 보정은 환자의 구성이 동일하다고 가정했을 때 여러 의료기관의 성과를 비교하기 위한 것임(NQF, 2014). 위험도 보정에 사용될 변수 선정을 위해 관련 전문가의 의견을 반영하여야 함
  - 자문회의를 거쳐 위험도 보정에 사용할 변수 목록을 선정함. 잠재적 변수목록을 구축할 때는 평가 결과의 수용성, 자료수집의 용이성 등을 고려해야 함
  - 위험도 보정의 모델의 타당성 검증 : 잠재적 변수 목록을 이용하여 위험도 보정에 가장 적합한 모형을 선정하고 적합한 모형으로 선정하기 위해 어떠한 통계량을 사용할지, 기준은 어느 수준으로 할지에 대해 사전에 논의되어야 함 (예시: C statistic\* 0.7 이상 등)

\* C statistic은 0~1의 값을 갖는데, 1에 가까울수록 예측력이 좋다고 평가 함.  
보건의료분야에서는 일반적으로 C statistic이 0.7이상일 때, 예측력이 좋다고 평가하며 위험도 보정과 관련한 예측모델을 개발할 시, 최소 0.6 이상은 넘어야 함

- 위험도 보정에 필요한 자료원 검토 : 위험도 보정에 선정된 변수를 수집하기 위한 자료원에 대해 명시가 있어야 함
  - 자료원 예시) 행정자료, 전자의무기록자료, 설문자료 등 (〈표 19〉 참고)

〈표 19〉 평가자료원 예시

분류항목	정 의
의무기록자료 (조사표)	의료기관에서 작성하여 평가 수행기관에 제출한 자료
행정자료	청구명세서, 요양기관 현황 신고자료, 행정안전부의 주민등록 전산자료 등의 행정자료로서 추가 가공되지 않은 기초 수집 자료(raw data)를 포함함
설문자료	전화, 웹사이트, 방문 등의 설문지를 통해 수집된 자료
기타	의무기록자료, 행정자료, 설문조사 외 자료로서 추가 가공된 자료(평가결과 등)를 포함함

- 위험도 보정에 사용되는 변수를 선정 후, 모델의 예측력에 대한 타당성을 검증
  - 검증한 위험도 보정 모형들을 제시하고, 모형별로 예측력에 대한 통계량을 제시해야 함
  - 위험도 보정에 사용할 최종의 모형을 선정하고, 만약 최적의 모형을 사용하지 않을 경우 그 사유에 대해서 명시해야 함
  - 예시) A모형의 경우 C statistic=0.78이고 B모형의 경우 C statistic=0.72로 A모형의 예측력이 더 좋으나, B모형으로 선정함. B모형의 C statistic=0.72로 예측력이 나쁘지 않고, A모형과 B모형의 차이가 있는 변수는 '낙상'이라는 변수임. '낙상'이라는 변수는 청구자료로 확인하기 힘들고 별도의 자료수집 과정을 거쳐야 함. 자료수집의 부담과 모형별 예측력을 고려하였을 때 B모형으로 선정하는 것이 무리가 없을 것으로 판단 함

〈표 20〉 핵심 결과지표 개발 예시

### 핵심 결과지표 정의

(결과지표의 정의에 대해 기술-개념적 정의와 조작적 정의에 대해 모두 기술)

#### - 핵심 결과지표명 : 모든 원인 30일 이내 재입원

- **개념적 정의** : 인공관절치환술을 받고 퇴원한 날부터 30일 이내에 급성기 의료기관에 다시 입원한 경우
- **조작적 정의** : 인공관절치환술을 받고 퇴원 날짜로부터 30일 안에 급성기 의료기관에 입원해서 청구된 건이 있는 경우
- 30일 이내 재입원은 ‘예/아니오’로 이분법적으로 측정되는 지표로, 퇴원 후 30일 이내에 여러 번의 재입원이 있을지라도 30일 재입원은 ‘예’로서 측정됨

#### - 제외기준

(분자의 제외기준에 대해 기술)

- **계획된 30일 이내 재입원은 제외함**
- **계획된 30일 이내 재입원의 개념적 정의** : 인공관절치환술을 시행하고 퇴원 후, 다른 관절 부위의 인공관절치환술을 위해 급성기 의료기관에 30일 이내에 다시 입원한 경우로 정의함
- **계획된 30일 이내 재입원의 조작적 정의** : 인공관절치환술 시행 후, 퇴원하여 급성기 의료기관에서 다시 입원 청구된 건 중, 재입원의 퇴원 시 주진단명이 퇴행성관절염, 류마티스관절염, 관절병증으로 코딩되어 있고, 해당 재입원 기간 동안 무릎 인공관절치환술 수가코드 N2072, N2077과 고관절 인공관절치환술 수가코드 N0711, N2070로 청구된 건으로 정의

#### - 위험도 보정

(위험도 보정에 사용하게 되는 변수를 정하게 된 경위에 대해 간략히 기술)

- **전문가 자문회의** : n회(최소 3회)의 전문가 자문을 거쳐, 위험도 보정에 사용할 잠재적 변수를 선정함
  - 1차 회의에서부터 n차 회의까지의 회차별 주요 회의결과를 기술
  - 위험도보정에 사용될 잠재적 변수목록과, 최종목록에 대해 제시  
→ 단, 최종목록은 C statistic으로 판정함
  - 자료원 : 의무기록자료(별도 조사표), 행정자료, 설문자료 등 사용 가능함

Table A.3 THA/TKA Readmission Measure Final Model Variables

Category	Variable	ICD-9 Code(s) or CC(s)
Demographic	Age-65 (years above 65, continuous)	
	Sex	
Procedure	Type of procedure	ICD-9-CM 81.51 (THA)
	Number of procedures (2 vs. 1)	
Comorbidities	Skeletal deformities	ICD-9-CM 755.63
	Post traumatic osteoarthritis	ICD-9-CM 716.15, 716.16
	Morbid obesity	ICD-9-CM 278.01
	History of Infection	CC 1, 3-6
	Metastatic cancer and acute leukemia	CC 7
	Cancer	CC 8-12
	Diabetes and DM complications	CC 15-20, 119, 120
	Protein-calorie malnutrition	CC 21
	Disorders of Fluid/Electrolyte/Acid-Base	CC 22, 23
	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	CC 38
	Severe Hematological Disorders	CC 44
	Dementia and senility	CC 49, 50
	Major psychiatric disorders	CC 54-56
	Hemiplegia, paraplegia, paralysis, functional disability	CC 67-69, 100-102, 177-178
	Polyneuropathy	CC 71
	Congestive Heart Failure	CC 80
	Chronic Atherosclerosis	CC 83-84
	Hypertension	CC 89, 91
	Arrhythmias	CC 92, 93
	Stroke	CC 95, 96
	Vascular or circulatory disease	CC 104-106
	COPD	CC 108
	Pneumonia	CC 111-113
	End-stage renal disease or dialysis	CC 130
	Renal Failure	CC 131
	Decubitus ulcer or chronic skin ulcer	CC 148, 149
	Cellulitis, Local Skin Infection	CC 152
	Other Injuries	CC 162
	Major Symptoms, Abnormalities	CC 166

[그림 30] 인공관절치환술 30일 이내 재입원의 위험도 보정에 사용된 변수 예시

\* 자료: Hospital-level 30-day all-cause risk-standardized readmission rate following elective primary total hip arthroplasty(THA) and/or total knee arthroplasty(TKA). Yale Haven Health Services Corporation/Center for Outcomes Research & Evaluation (YNHHC/CORE). 2012

#### -기타 핵심 결과지표 후보지표

(최우선 핵심 결과지표인 30일 이내 재입원의 지표를 사용하지 못할 경우를 대비한 후보지표 제시)

- 병원 내 사망, 수술 후 합병증 → 각 지표별로 앞서 제시한 예처럼 개념적 정의, 조작적 정의, 위험도 보정에 필요한 변수 등을 동일하게 제시함

## 다) 구조, 과정지표 개발

- 구조(structure) 및 과정(process)지표 개발 시에는 핵심 결과지표와 관련성이 있다는 근거가 있는 경우 평가지표를 개발하고, 근거의 출처와 근거수준을 밝혀야 함
- 임상진료지침의 권고안을 활용하는 경우, 임상진료지침 정보센터에서 임상진료지침 개발 시 사용하는 modified GRADE(Grading of Recommendations Assessments, Development and Evaluation) 권고등급을 활용함 (<표 21> 참고)
  - 가능한 권고 등급 Class I의 권고를 이용하여 지표를 최우선으로 개발하며, Class I 수준의 권고가 지표를 개발하기에 부족하거나 없는 등의 부득이한 경우, 권고등급이 높은 권고의 수준인 Class IIa, Class IIb로 지표를 개발하되, 그 사유를 명백히 밝혀야 하고 사유는 타당하여야 함
  - Class III의 권고 등급에 해당하는 근거로는 평가지표를 개발하지 않는 것이 원칙임

<표 21> 권고의 등급화 기준

권고등급	정의	권고의 표기
Class I	근거수준(A)과 편익이 명백하고, 진료현장에서 활용도가 높은 권고의 경우	권고함 (Is recommended)
Class IIa	근거수준(B)과 편익이 신뢰할만하고, 진료현장에서 활용도가 높거나 보통인 권고의 경우	고려함 (Should be considered)
Class IIb	근거수준(C 혹은 D)과 편익을 신뢰할 수 없으나, 진료현장에서 활용도가 높거나 보통인 권고의 경우	고려할 수 있음 (May be considered)
Class III	권고수준(C혹은 D)을 신뢰할 수 없고, 위대한 결과를 초래할 수 있으면서, 진료현장에서 활용도가 낮은 권고의 경우	권고되지 않음 (Is not recommended)

\* 자료 : 임상진료지침정보센터

〈표 22〉 근거수준의 등급

근거수준	정의
A	권고 도출의 근거가 명백한 경우 1개 이상의 무작위임상연구(RCT) 혹은 메타분석(Meta analysis) 혹은 체계적 문헌 고찰(Systematic Review)
B	권고 도출의 근거가 신뢰할 만한 경우 1개 이상의 잘 수행된 환자 대조군 연구 혹은 코호트 연구와 같은 비무작위 임상시험연구(Non-RCT)
C	권고 도출의 근거가 있으나 신뢰할 수는 없는 경우 관찰연구, 증례보고와 같은 낮은 수준의 관련 근거
D	권고 도출의 근거가 임상경험과 전문성을 기반으로 한 전문가 의견(expert opinion)인 경우

\* 자료 : 임상진료지침정보센터

- 임상진료지침의 권고안을 활용할 수 없어 학술논문을 근거로 개발할 경우, SIGN(Scottish Intercollegiate Guidelines Network)의 권고등급체계를 적용함 (〈표 22〉 참고). SIGN권고등급체계에 따라 가능한 근거수준 B이상의 근거를 바탕으로 개발함. 부득이한 경우 근거수준 C를 바탕으로 개발은 할 수 있으나, 이러한 경우 타당한 근거가 같이 제시되어야 함
- 근거수준 D를 근거로는 평가지표를 개발할 수 없음
- 추가 자료 수집이 필요한 경우에는 자료 수집 부담을 최소화하는 것이 원칙임. 다른 평가 프로그램이나 다른 기관에서 수집하고 있는 자료를 먼저 검토하고 최대한 연계 방안을 모색해야 함
- 근거로 활용할 수 있는 임상진료지침 및 학술논문이 미흡한 경우에도 평가지표로 개발할 수는 있으나 임상학회의 공식적인 합의된 의견 (또는 이와 견주어 동등한 전문가들의 의견)이 있는 경우에 한하여 개발하여야 함

〈표 23〉 SIGN의 권고등급체계 (2015년 이전)

근거수준	정의
A	주로 1+의 연구로 구성된 근거이고, 직접 표적 모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우
	최소 하나 이상의 메타분석(meta analysis), 체계적 문헌고찰(systematic review) 또는 1++의 무작위배정 비교임상시험연구가 있고, 표적 모집단에 적용이 가능한 경우
B	2++로 구성된 근거이고, 직접 표적 모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우
	1++또는 1+의 평점을 받은 연구로부터 추정된 근거인 경우
C	2+의 연구로 수정된 근거이고, 직접 표적 모집단에 적용할 수 있으며, 결과가 전반적으로 일관성을 보이는 경우
	2++의 평점을 받은 연구로부터 추정된 근거인 경우
D	근거의 수준 3 또는 4에 해당되거나 2+의 평점을 받은 연구로부터 추정된 경우

\* 자료 : 임상진료지침 실무를 위한 핸드북. 한국보건연구원. 2022. available at: [https://www.neca.re.kr/lay1/bbs/S1T11C102/F/39/view.do?article\\_seq=8863&cpage=1&rows=10&condition=&keyword=&show=&cat=](https://www.neca.re.kr/lay1/bbs/S1T11C102/F/39/view.do?article_seq=8863&cpage=1&rows=10&condition=&keyword=&show=&cat=)

〈표 24〉 SIGN의 권고등급체계 추가 설명표

수준	설명
1++	높은 수준의 메타분석, RCTs의 체계적 문헌고찰, 비뚤림 위험이 매우 낮은 RCTs
1+	잘 수행된 메타분석, 체계적 문헌고찰 및 비뚤림 위험이 낮은 RCTs
1-	메타분석, 체계적 문헌고찰, 비뚤림 위험이 높은 RCTs
2++	환자대조군 연구 및 코호트 연구로 수행된 높은 질의 체계적 문헌고찰 혼란변수 및 비뚤림 위험이 매우 낮거나 높은 수준의 환자-대조군 연구 및 코호트 연구
2+	혼란변수 및 비뚤림 위험이 낮거나 잘 수행된 환자 대조군 연구 및 코호트 연구
2-	혼란변수 및 비뚤림 위험이 높거나 환자대조군 연구 및 코호트 연구
3	비분석적 연구(사례보고, 사례연구)
4	전문가 의견

\* 자료 : 가치에 근거한 권고 등급 개발. 건강보험심사평가원. 2014



〈표 25〉 구조와 과정 후보지표 개발 예시

〈후보지표 예시〉	
지표명	퇴원 후 적절한 통증 조절율
지표의 유형	과정지표
정의	THA/TKA 수술 받은 사람들 중에서, 퇴원 시 적절한 통증조절 약물을 처방받은 사람들의 분율 정의
조작적 정의	THA/TKA (수가코드 N2072, N2077 및 N0711, N2070)에 해당하는 수가코드가 청구된 입원 건 중, 퇴원 시 퇴원 약물에 NSAIDs, COX-2 선택적 저해제, 트라마돌염산염을 처방받은 환자의 분율
분모	평가대상 선정에 사용된 분모의 기준과 동일
분모의 제외기준	평가대상 선정에 사용된 분모제외 기준과 동일
분자	퇴원 시, 퇴원약물로 NSAIDs, COX-2 선택적 저해제, 트라마돌염산염을 처방받은 환자 → NSAIDs, COX-2 선택적 저해제, 트라마돌염산염에 해당하는 약품의 일반성분명에 대한 코드를 제시
분자의 제외기준	3가지 약물 모두를 처방해야 하는지, 2가지 약물만 처방해도 되는지, 한 가지만 처방받은 경우도 관계 없는지 기준 필요 위에서 제시하지 않은 약물을 처방받은 환자들의 처리기준에 대해 제시
산출공식	$\text{퇴원시, 통증 조절 약물 처방율} = \frac{\text{NSAIDs, COX-2 선택적 저해제, 트라마돌염산염을 모두 처방한 건수}}{\text{65세 이상환자에서 THA/TKA수행한 전체 건수}}$
근거수준	SIGN의 권고등급체계 2+ (잘 수행된 코호트 연구) ** 근거자료: Hong I, Westra JR, Goodwin JS et al. Association of pain on hospital discharge with the risk of 30-day readmission in patients with total hip and knee replacement. The Journal of Arthroplasty. 2020;35(12):p3528-3534
핵심 결과지표와의 관련성에 대한 근거	THA를 시술받은 환자에서 퇴원 시 통증이 없던 사람들에 비해 심한 통증이 있던 사람들의 30일 이내 재입원의 Adjusted Odds ratio는 1.60 (95% CI, 1.33-1.92)이었으며, TKA를 시술받은 환자에서도 비슷하게 퇴원 시 통증이 없던 사람들에 비해 심한 통증이 있던 사람들의 30일 이내 재입원의 Adjusted Odds ratio는 1.38 (95% CI, 1.19-1.59)이었음 (Hong I, 2020). 따라서, 퇴원 시 통증조절을 위한 적절한 약물을 처방한다면 재입원을 예방할 수 있음

## 라) 평가지표 세부사항 구체화

○ 후보지표로 사용할 수 있는 지표의 유형은 다음 <표 26>과 같음

<표 26> 후보지표로 사용 가능한 지표의 유형

질 평가 유형	설명 및 목적	예시
복합 (Composite)	두 개 이상의 개별 평가지표를 조합하여 하나의 결과값을 산출하는 것	여러 가지의 평가지표를 산출하고, 하나의 환자안전 또는 위해사건 평가지표로 종합
비용/자원(사용) (Cost/Resource)	의료서비스가 적용된 인구집단이나 사건에 사용된 비용이나 자원을 측정	환자 당 지출된 급여 의료비
효율성 (Efficiency)	건강결과와 의료비용을 측정	비심장 저위험 수술에서 수술 전 위험 평가를 위한 심장 영상 촬영
결과 (Outcome)	환자의 건강 상태를 측정 또는 건강 상태의 변화	암 환자의 30일 이내 계획되지 않은 재입원
결과 기반의 환자보고 (Patient reported outcome-based)	환자가 보고한 결과측정을 바탕으로 성과를 측정함	경부 손상 환자에서 기능적 상태 변화
절차 (Process)	의료서비스가 제공되는 과정을 측정	성인 주요 우울증 : 자살위험평가
구조 (Structure)	의료를 제공하는데 필요한 인력, 시설, 장비	방사선과 : 유방조영술 스크리닝을 위한 시스템을 알림

\* 자료 : Quality Measures: How They Are Developed, Use, & Maintained. CMS. available at: <https://mmshub.cms.gov/sites/default/files/Guide-Quality-Measures-How-They-Are-Developed-Used-Maintained.pdf>

○ 평가 지표별 구체적 기술 항목은 다음과 같음

- 지표 설명 : 개념적 정의와 개념적 정의를 측정하는데 필요한 조작적 정의에 대해 설명함
  - 개념적 정의는 지표의 사전적 의미를 의미함
  - 조작적 정의는 추상적인 개념적 정의를 실제 평가에서 산출할 수 있고, 측정이 가능한 형태로 표현하는 것
- 분모 : 평가하고자 하는 모집단을 설명하는 문구로서 율(rate), 비율(proportion), 비(ratio)로 계산하는데 사용되는 분수의 아래쪽 부분을 의미함

- 평가하고자 했던 대상(target)의 초기 모집단이거나 그 중 특정 일부로 제한한 집단을 의미함. 다음과 같은 매개변수를 포함하여 분모를 정확하게 정의해야 함
- 연령대, 1차/2차/3차 의료기관, 응급/외래/입원과 같은 세팅, 진단, 절차, 시간간격 등
- 의학적 이유, 환자의 특성, 의료체계의 특성 상 제외가 필요하다고 판단된 경우에만 제외
- 분자 : 율(rate), 비율(proportion), 비(ratio)로 계산할 때 사용되는 분수의 윗부분을 의미함
  - 평가의 초점, 표적이 되는 과정이나 상태, 사건이나 결과를 의미함. 분자는 성과 측정의 조건을 만족하는 어떠한 상태에 대해 자세히 기술함
  - 분자를 정의함에 있어 분자의 포함기준과 제외기준에 대해 상세하게 기술해야 함
- 근거 : 개별 평가지표를 개발하게 된 근거에 대해 요약하여 설명함
  - 근거 내용, 근거 출처, 근거 수준에 대해 기술함
- 자료원 : 평가지표를 측정하기 위한 자료원에 대해 설명함
  - 자료원, 자료수집 방법, 평가에 사용될 자료의 시점

## 다. 후보지표 검증

### 1) 단계 설명

- 개발된 후보지표가 평가지표로서 적절한지를 검토하는 단계임. 개발한 후보지표를 실제 자료를 활용하여 분석하고 결과를 해석하는 과정을 거침. 후보지표의 중요성과 타당성, 과학적 수용성 및 사용 가능성에 대한 실제적 근거를 제공하는 단계임
- 후보지표 개발과 검증단계는 동시에 수행되는 반복적인 과정으로, 순환적인 과정을 거치면서 개발된 후보지표의 세부 내용을 수정하거나 추가 또는 삭제하는 단계임

## 2) 검증 절차

- 검증절차는 다음과 같이 4단계로 구분됨
  - 검증계획서 수립
  - 표본추출
  - 수행, 분석, 지표의 세부사항 수정 및 재검증
  - 결과요약
- 검증계획서 수립 : 실제 분석은 2차례에 걸쳐 분석을 수행하며, 2차례 분석에 대한 계획을 수립해야 함
  - 검증계획서에는 1차 검증과 2차 검증에 대한 내용과 각 검증별, 표본추출 및 분석기간에 대한 계획이 있어야 함
  - 검증계획서에는 평가지표의 이름, 검증의 유형, 검증의 목적, 분석과 보고 완료까지의 기간, 자료수집 방법, 검증할 대상에 대한 설명, 자료수집 변수명, 표본추출방법, 통계분석방법에 대한 내용을 포함해야 함
  - 1차 검증은 후보지표의 세부사항을 확정하기 전에 수행하는 검증으로, 접근이 용이한 자료를 활용하여 후보지표의 세부사항에 수정이 필요한 부분이 있는지를 확인하기 위한 단계임. 후보지표의 세부사항을 구체화하고, 실현 가능성을 확인하기 위해 수행하는 검증임
    - 평가지표를 개발한 의도와 일치하는지의 여부(개념타당성)를 확인하기 위한 검증임
  - 2차 검증은 1차 검증보다 평가하고자 하는 모집단을 더 많이 대표할 수 있는 표본자료를 이용하여 평가의 실현 가능성과 과학적 수용성 및 유용성을 평가하는 검증임
- 표본추출 : 1차 검증과 2차 검증의 목적에 따라 표본추출을 하되, 자료수집에 대한 부담을 최소화할 수 있는 방법을 선택해야 함
  - 1차 검증의 경우, 개발자의 편의에 의한 표본을 추출하는 방식인 반면에 2차 검증은 특정 평가지표의 과학적 수용성을 검증하기 위해 수행하는 검증이기에 평가대상이나 초기 모집단의 특성을 측정하기 위한 전략적인 표본추출이 필요함

- 일반적으로 검증의 타당성과 신뢰성을 확보하기 위해 다음과 같은 방법을 사용함
  - 모집단의 다양성을 충분히 반영할 수 있는 대표성 있는 표본을 추출해야 함 (예: 의원급, 병원급, 종합병원급 인지 등)
  - 신뢰성과 타당성을 확보하기 위해서는 적절한 표본이 있어야 하고, 가능한 무작위 추출방식을 적용해야 함

〈표 27〉 1차 검증과 2차 검증의 특징 비교

특징	1차 검증	2차 검증
검증 시점	<ul style="list-style-type: none"> <li>• 후보지표의 세부사항을 확정하기 전에 실시</li> <li>• 후보지표 세부 사항을 구체화를 위해 여러 번 수행할 수 있음</li> </ul>	<ul style="list-style-type: none"> <li>• 후보지표의 세부사항을 확정 후 실시</li> </ul>
검증의 규모	<ul style="list-style-type: none"> <li>• 상대적으로 작은 규모</li> <li>• 평가를 위해 필요한 요소만 포함되어 있는 자료로 분석</li> </ul>	<ul style="list-style-type: none"> <li>• 평가대상의 모집단을 대표할 수 있는 적절한 규모의 자료가 필요</li> <li>• 평가대상의 모집단을 대표할 수 있는 표본 추출방법이 필요</li> </ul>
표본 추출	<ul style="list-style-type: none"> <li>• 개발자 편의에 의한 표본추출</li> </ul>	<ul style="list-style-type: none"> <li>• 과학적 수용성을 검증하기 위한 충분한 표본</li> <li>• 평가대상 또는 초기모집단을 대표</li> <li>• 평가에 중요한 사람, 장소, 시간, 사건 및 조건을 대표해야 함</li> <li>• 청구자료를 기반으로 하는 경우 전수를 활용함</li> <li>• 가능한 무작위추출을 통해 분석의 결과가 일반화가 가능하도록 해야 함</li> </ul>
세부사항 구체화	<ul style="list-style-type: none"> <li>• 후보지표의 세부사항과 관련된 문제를 조기에 발견할 수 있어야 함</li> </ul>	<ul style="list-style-type: none"> <li>• 평가지표 산출의 가능성, 과학적 수용성 및 유용성을 검토</li> </ul>

## ○ 검증계획서 예시

〈표 28〉 1차 검증계획서 예시

**1차 검증계획서**

## ○ 1차 검증 목적

- 위험도 보정에 사용될 잠재적 변수들을 통해 예측모형을 확정하기 위함

- 개발된 후보지표의 측정 가능성과 수정 등 확정 및 변경하기 위함

#### ○ 검증내용

- 위험도 보정 모델에 사용할 변수들을 조합하여 C statistic를 산출하고 가장 좋은 예측모델을 개발함
- 개발된 예측모형으로 3개년도 자료에 대한 C statistic과 30일 이내 재입원에 대한 모형에 사용된 변수별 빈도 및 odds ratio를 제시하여 모형의 안정성에 대해 검증함
- 개발된 후보지표에 대한 결과값을 산출하고 후보지표의 세부사항을 변경하거나 확정함

#### ○ 자료원

- 2019년~2021년까지 건강보험심사평가원에 THA/TKA로 청구된 청구자료

#### ○ 표본추출

- THA/TKA를 합하여 매년 500건 이상을 수술하는 의료기관 3개를 선정하여, 5개 의료기관에서 2019년~2021년 3년 동안 THA/TKA수술을 받은 환자를 연도별로 200건 무작위 추출

#### ○ 통계 분석 방법

- 30일 이내 재입원에 대해서는 로지스틱 회귀분석을 통해 모형의 적합성을 판정하며, 최종 모형의 선정을 통해 기관별 30일 이내 재입원의 위험도 보정 재입원율(30-day all-cause risk standardized readmission rate, RSRR)을 산출할 예정임

〈표 29〉 2차 검증계획서 예시

## 2차 검증계획서

#### ○ 2차 검증 목적

- 1차 검증을 통해 개발된 후보지표의 실현가능성, 과학적 수용성(타당도와 신뢰도), 유용성 등을 평가하기 위함

## ○ 검증내용

- 평가대상의 모집단을 대표할 수 있는 표본산출을 통해 개발된 후보지표별로 결과값을 산출함
- 이를 통해 지표로서 의미 있는 지표인지, 기관별 차별화는 할 수 있는지 등 지표별 현재의 수준과 향후 개선 가능성의 정도

## ○ 자료원

- 2019년~2021년까지 건강보험심사평가원에 THA/TKA로 청구된 청구자료

## ○ 표본추출

- THA/TKA로 청구된 병원급 이상의 의료기관들 중 의료기관 종별, 연도별 청구건수, 수련병원 유무를 바탕으로 층화계통추출(stratified systematic random sampling)을 시행함. 각 층별로 무작위 500건을 추출하여 분석함
  - 표본추출에 대해 연구방법론 전문가와 청구자료(또는 사용할 해당자료)에 대해 충분히 잘 알고 있는 전문가의 자문을 받아 수행하거나, 심평원에서 조직된 전문가 패널을 활용 가능함

## ○ 통계분석방법

- 30일 이내 재입원에 대해서는 로지스틱 회귀분석을 통해 모형의 적합성을 판정하며, 최종 모형의 선정을 통해 기관별 30일 이내 재입원의 위험도 보정 재입원율(30-day all-cause risk standardized readmission rate, RSRR)을 산출할 예정임

- 수행, 분석, 세부사항 수정 및 재검증 : 검증수행계획서에 따라 자료를 수집한 뒤에 분석하고, 분석한 중간결과를 바탕으로 평가지표의 세부사항을 수정하고 다시 검증하는 단계임

- 수행, 분석 : 검증수행계획서에 따라 분석을 수행하는 과정으로 이 단계에서는 방법론 전문가 또는 통계학자 등 다양한 분야의 전문가 참여하에 분석을 수행해야 함
- 세부사항 수정 : 분석한 결과를 바탕으로 평가지표의 세부사항, 자료수집 지침이나

산출방법 등에 대한 변경을 검토함

- 1차 검증을 통해서는 평가지표의 세부사항이나 평가지표 산출 과정의 문제점을 해결할 수 있는 방법을 모색함
- 2차 검증을 통해서는 평가대상의 정의를 변경하거나 비교그룹의 정의를 보정하는 등의 작업을 수행함
- 평가지표의 내용에 변경사항이 있을 경우, 재검증하기 전에 반드시 전문가에게 자문을 받아야 함
- 재검증 : 재검증 과정은 평가지표를 검증하고 세부사항을 수정하고 다시 재검증하는 반복적인 과정
- 결과요약 : 평가지표에 대한 결과를 요약하여 정리하는 단계로 5개의 측면의 기준(중요성, 과학적 수용성, 실현가능성, 활용가능성, 자료수집 부담)에서의 문제의 정도와 각각에 대한 전문가의 의견 등을 반영하여 정리하는 단계임
- 평가지표의 이름
- 평가지표 검증의 요약과 권장사항
- 수행된 검정의 종류(예: 1차 검증 vs 2차 검증), 검증의 개요
- 검증수행계획서를 따르지 못했던 사항과 그에 대한 근거
- 자료수집과 방법론 관리, 자료수집 부담
- 검증한 평가대상에 대한 설명과 검증된 의료기관의 유형(의원/병원/(상급) 종합 병원, 응급/외래/입원 등)에 대한 설명
  - ✓ 자료유형과 자료원을 포함하여 자료검증에 대한 설명
  - ✓ 자료원 설명
  - ✓ 표본추출방법
  - ✓ 분자와 분모의 제외기준에 대한 설명
  - ✓ 만약 환자의 의무기록을 활용했다면 자료 조사자와 자료 요약자들의 교육 방법



- 평가지표의 세부사항에 대한 자세한 기술내용과 평가지표의 산출점수(결과)
- 분석을 수행한 결과의 내용
  - ✓ 요약통계(예: 평균, 중위수 등)
  - ✓ 중요성
  - ✓ 과학적 수용성(신뢰도, 타당도)
  - ✓ 분자분모 제외기준과 분자의 예외기준
- 위험도 보정과 층화분석에 대한 필요성
- 실현가능성 : 평가지표 측정을 위한 예상된 비용이나 자료수집 부담의 측면에서 실현가능성을 검토
- 활용가능성 : 검증결과를 바탕으로 정성적 정량적 활용 가능성에 대한 전문가의 권장사항에 대해 기술
- 평가지표 세부사항 변경에 대한 권장사항
- 1차 검증이나 2차 검증을 통한 드러난 한계점
- 종합재평가계획

## 라. 후보지표 평가

### 1) 단계 설명

- 후보지표 선정 : 지표별 후보지표의 결과요약을 바탕으로 정해진 평가기준에 따라 점수를 산출하여 개발된 후보지표들 중 실제 사용할 수 있는 후보지표의 우선순위를 정함
- 선정기준
  - 중요성 : 의료기관별 성과의 차이가 있고 중요한 환자의 상태 또는 중요한 의료의 질 문제인가? 개선이 가능한 문제인가?
  - 측정된 주제와 관련된 진료량, 빈도, 또는 비용 등을 검토하여 임계값에 이르지 않는었는지 기관 간 변이(variation)는 어느 정도인지 시간에 따른 추세는 어떠한지,

실질적으로 개선이 가능한지 등을 고려

- 실현가능성 : 의료기관의 협조 및 자료수집에 대한 예상 비용 등 자료수집의 부담을 검토함
- 과학적 수용성(타당도와 신뢰도) : 개발된 평가지표를 통해 실제 평가를 했을 때 의도한 대로 의료의 질이 일관되고 신뢰할 수 있는 결과를 가져올 수 있는지를 검토함
  - 타당도 : 평가하고자 하는 내용을 얼마나 정확하게 평가할 수 있는지의 문제로 구성타당도, 집중타당도, 기준타당도, 차별타당도, 액면타당도, 예측타당도의 측면에서 검토해야 함
    - ✓ 구성타당도 : 측정하고자 하는 추상적 개념이 실제로 측정도구에 의해 제대로 측정되었는지의 정도를 의미
    - ✓ 기준타당도 : 하나의 측정지표를 사용하여 측정한 값과 다른 지표를 사용하여 측정한 값과의 관련성 정도
    - ✓ 변별타당도 : 서로 다른 개념을 측정하는 경우, 그 측정값 사이에 차이가 있어야 하는 정도
    - ✓ 액면타당도 : 평가지표가 평가하고자 하는 대상을 제대로 측정하고 있는지 주관적으로 판단하는 정도
    - ✓ 예측타당도 : 기존의 평가와 비교하여 새롭게 개발된 평가지표로 의도한 바의 결과를 얻는 정도
  - 신뢰도 : 개발된 지표가 같은 것을 얼마나 일관되게 평가할 수 있는가의 정도
    - ✓ 동일한 자료에 대해 두 사람이 분석을 시행하여 분석한 결과가 일치하는 정도로 판단함
- 활용가능성 : 환자, 의료기관, 정책적인 다양한 관점에서 활용 정도를 검토함
- 다각도적인 측면에서 검토: 앞선 4가지의 기준 외에도 평가로 인한 잠재적 이익과 위험을 모두 고려하여 종합적으로 판단함
- 분모 및 분자의 제외기준 처리
  - 분모와 분자의 제외기준에 대한 의견이 상충될 경우, 필요에 따라 분모 및 분자의

제외기준이 적용되었을 때와 적용되지 않았을 때에 대한 민감도 분석을 수행하여 결정함

## 2) 후보지표 선정의 원칙

- 1개 이상 핵심 결과지표를 후보지표로 선정하는 것이 원칙임. 평가 내용이 결과지표의 활용이 적절하지 않다고 판단하는 경우에는 관련 위원회에 타당한 근거를 제시하고 논의를 거쳐야 함
- 개발된 후보지표들 중에서 실제로 사용될 지표를 선정하기 위해서는 임상전문가 및 통계전문가, 보건의료 정책전문가와 자료 관련 전문가, 정부관계자 등의 합의된 의견을 바탕으로 선정함
- 개발된 후보지표 RAM의 의사결정방식<sup>21)</sup> (Kathryn Fitch et al. The RAND/UCLA Appropriateness Method User's Manual. RAND Europe. 2001)에 따라 최종 평가지표로 선정함
  - 개발된 후보지표를 최종 평가지표로 선정하기 위해서는 2회의 전문가들의 평가 과정이 필요함
    - 후보지표 개발자가 반드시 필요하다고 판단한 항목이 2차 라운드 이후에도 동의 되지 않은 경우 또는 2차 라운드를 진행하고도 합의가 이루어지지 않은 후보지표가 있는 경우 등 필요에 따라 추가로 평가를 수행할 수 있음
    - 1차 라운드는 설문지나 전자우편으로 평가하고 2차 라운드는 대면토론으로 진행함
  - 전문가 패널 구성 : 임상전문가 및 통계전문가, 보건의료 정책전문가 및 자료 관련 전문가, 정부관계자, 평가 실무자 등의 관련 전문가들로 구성하며 일반적으로 9명의 전문가로 패널을 구성하며, 최대 16명 이내로 구성함
  - 1차 라운드 : 개발된 후보지표에 대하여 '필요성' 대해 9점 척도로 평가를 하도록 함. 필요성에 대한 패널 응답의 중앙값을 바탕으로 '동의' 및 '적합성'에 대해 평가함

21) 1980년대 중반에 내외과적 시술의 과다이용과 과소이용에 대해 의료서비스의 적정성을 측정하기 위해서 개발되었으며, 과학적인 근거와 전문가들의 공통된 결정을 합하여 공식적 합의를 도출하는 방법 중 하나임 (지선미 외. 2010; 최보람. HIRA정책동향. 2010)

### ① ‘필요성’ 판정

- ✓ 중요성, 실현가능성, 과학적 수용성(타당도와 신뢰도), 활용가능성의 측면에서 전문가 개개인이 개발된 각각의 후보지표에 대해 평가지표로서 필요한 지를 1~9점까지 점수를 부여하도록 함
- ✓ 9점에 가까울수록 필요하다는 것을 의미함

〈표 30〉 필요성 평가 기준

구분	정 의
필요 (necessity)	7~9점인 경우
불확실(uncertain)	4~6점인 경우
불필요(needlessness)	1~3점인 경우

### ② ‘동의여부’ 판정

- ✓ 동의여부는 동의(agreement), 비동의 (disagreement), 불확실(uncertain) 3개의 범주로 판정하며, 패널 응답의 중앙값과 패널이 응답한 3점 구역의 빈도에 따라 판정함

〈표 31〉 동의여부 판정 기준

구분	정 의
동의 (agreement)	패널 응답의 중앙값을 포함하는 3점 범주 외의 범주에 응답한 패널들의 수가 2명 이하인 경우
비동의 (disagreement)	패널 중 3명 이상이 7~9점 범주 사이에 있고, 다른 3명 이상이 1~3점 범주에 위치하는 경우
동의, 비동의 이외의 모든 경우	동의, 비동의 이외의 모든 경우

- ✓ 판정예시) 패널 응답의 중앙값이 8점이면서 8점을 포함하는 범주 (7~9점) 외의 범주(1~3점이거나 4~6점)에 응답한 패널의 수가 2명인 경우 ‘동의’로 판정
- ✓ 전문가 패널의 수가 9명으로 구성되지 않은 경우, 다음의 표에 따라 동의와 비동의 여부를 판단함

〈표 32〉 패널 규모에 따른 동의여부 판정 기준

패널 규모	비동의(disagreement)	동의(agreement)
	양 극단의 빈도 (1~3 and 7~9)	중앙값 (1~3, 4~6, 7~9)을 포함한 범위를 벗어난 빈도
8~10	$\geq 3$	$\leq 2$
11~13	$\geq 4$	$\leq 3$
14~16	$\geq 5$	$\leq 4$

## ③ ‘적합여부’ 판정

- ✓ 적합여부는 적합/부적합/불확실 3개의 범주로 판정하며, 패널 응답의 중앙값과 동의여부로 판정함

〈표 33〉 적합여부 평가 기준

구분	정 의
적합 (appropriate)	패널 응답 중앙값이 7~9점이면서, 비동의가 없는 경우
불확실 (uncertain)	패널 응답 중앙값이 4~6점이거나 비동의가 있는 경우
부적합 (inappropriate)	패널 응답 중앙값이 1~3점이면서 비동의가 없는 경우

- 2차 라운드 : 전문가 위원들 중 1명이 조정자로서 역할을 하고, 대면토론으로 진행함.  
1차 라운드 결과 ‘불확실’ 항목을 중심으로 토론함
- 2차 라운드에서는 패널들에게 각 항목에 대한 1차 라운드 점수 분포(중앙값, 표준편차, 점수별 선택한 패널 수)와 패널 본인이 선택한 점수에 대한 정보를 제공함
- 전문가들의 대면토론으로 ‘불확실’ 항목의 지표에 대해 평가지표로 선정할지의 여부를 합의함

〈표 34〉 후보지표에 대한 전문가 합의과정 예시

## 1. 전문가 패널 구성

- 임상학회 대표 3명, 통계전문가 1명, 보건의료 정책전문가 1명, 평가자료 관련 전문가 1명, 정부관계자 1명, 평가 실무자 1명, 환자단체 1명

## 2. 1라운드 패널조사

- 전문가 패널에 각각의 후보지표에 대해 필요성 여부 1~9점 사이의 점수를 판단하도록 함
- 점수가 높을수록 평가지표로서 필요하다는 의미임

- 1라운드 패널 조사 결과 각 9명의 전문가가 A후보지표에 대해 각각 7점, 8점, 8점, 7점, 9점, 6점, 1점, 3점, 7점을 주었을 경우의 판정
- 이 후보지표의 중앙값은 (1, 3, 6, 7, 7, 7, 8, 8, 9) 7점이고, 중앙값 7점이 속한 3점 범주는 7~9점 사이에 해당함
- 7~9점 범주 외의 범주에 평가를 한 전문가 수는 3명(1, 3, 6)이 있으므로 동의여부는 '불확실'로 판정

- '동의' 판정 기준 : 패널 응답의 중앙값을 포함하는 3점 범주 외의 범주에 응답한 패널들의 수가 2명 이하인 경우

- 1라운드 패널 조사 결과 각 9명의 전문가가 B후보지표에 대해 각각 1점, 2점, 3점, 7점, 7점, 8점, 6점, 5점, 7점을 주었을 경우의 판정
- 7~9점 범주 사이로 평가한 전문가는 4명( 7, 7, 7, 8)이고, 1~3점 사이로 평가한 전문가는 3명(1, 2, 3)으로 '비동의' 판정기준에 해당하므로 B후보지표는 평가지표로 선정하지 않음

- '비동의' 판정기준 : 패널 중 3명 이상이 7~9점 범주 사이에 있고, 다른 3명 이상이 1~3점 범주에 위치하는 경우

## 3. 2라운드 패널회의

- 1라운드 패널조사 결과, A후보지표는 '불확실', B후보지표는 '비동의'로 판정되었기 때문에 A후보지표에 대해 대면으로 서로 토의하며 합의여부를 결정함
- 전문가 패널 회의에서 토의를 거쳐 평가지표로 선정할지의 여부를 합의함

## 마. 최종 평가지표 선정

- 최종 개발된 지표는 다음의 체크리스트를 이용하여 개발하는 과정에서 빠뜨린 부분은 없는지 확인함
- 또한, 선정된 최종평가지표에 대해 지표별 평가로 인해 발생될 이익이나 위해<sup>22)</sup> 정도를 검토해야 함. 평가지표 개발을 통해 실제 평가를 수행할 경우 발생할 잠재적 이익과 위해의 정도를 검토하여 기술해야 함
  - 가칭 최종심의위원회에 평가지표 상정 시, 발생될 이익이나 위해정도에 대해 검토된 내용을 함께 상정해야 함

〈표 35〉 평가지표 개발 check list

개발 단계	확인 내용	수행여부	
		예	아니오
1. 평가지표 개발 계획수립			
1) 문제 확인			
문제 확인 및 현황파악	의료의 질과 관련된 문제점을 정량적으로 기술하였는가?		
지표개발의 목적	지표 개발의 목적에 대해 구체적으로 명시하였는가?		
활용계획	지표개발의 결과를 어떻게 활용할 것인지 명시하였는가?		
평가로 인해 발생될 이익이나 위해 정도 검토	평가의 결과로 발생될 잠재적 이익이나 위해에 대해 검토하고 검토된 내용을 기술하였는가?		
2) 평가대상 선정			
평가 후보집단 선정	평가대상이 되는 후보집단에 대한 개념적 정의를 하였는가?		
	평가대상이 되는 초기 집단에 대해 조작적 정의를 하였는가?		
	분모의 제외기준을 설정하였는가?		
	분모에서 제외기준마다 제외의 근거를 제시하였는가?		
2. 후보지표 개발			
1) 문헌검토			
근거자료 검색 및 검토	확인된 문제와 관련된 임상지침 및 근거자료를 검토하였는가?		

22) 평가로 인해 발생될 잠재적 부작용을 의미하며 임상적 측면과 정책적 측면, 2가지 측면에서 검토해볼 수 있음. 예를 들어 급성기 의료기관에서 질환별 재원일수를 평가하는 경우 임상적 측면에서 조기퇴원이나 외래방문이 증가하는 등 풍선효과로 인한 위해가 발생 가능함. 의료기관별 사망률을 평가하여 공개하는 경우, 정책적 측면에서 의료기관 몰림현상이 강화되는 등의 위해가 발생할 수 있음

개발 단계	확인 내용	수행여부	
		예	아니오
	근거자료의 근거수준에 대해 확인하였는가?		
2) 핵심 결과지표 설정			
핵심 결과지표 개발	분자에 대한 개념적 정의를 하였는가?		
	분자에 대해 조작적 정의를 하였는가?		
	분자의 제외기준을 설정하였는가?		
	분자에서 제외기준마다 제외의 근거를 제시하였는가?		
	핵심 결과지표를 2개 이상 마련하였는가?		
위험도 보정	위험도 보정에 대해 전문가 자문을 받았는가?		
	위험도 보정에 사용할 잠재적 변수목록이 설정되었는가?		
	위험도 보정의 타당성을 검증할 통계량이 제시되었는가? (예시: C statistics)		
3) 후보지표 세부사항 구체화			
구조와 과정 후보지표 개발	후보지표의 개념적 정의를 하였는가?		
	후보지표의 조작적 정의를 하였는가?		
	후보지표 분자의 제외기준을 설정하였는가?		
	후보지표 분자에서 제외기준마다 제외의 근거를 제시하였는가?		
근거수준 확인	(근거가 있는 경우) 후보지표에 사용된 근거수준이 Class II 또는 SIGN체계 2-이상인가?		
	(근거가 미미한 경우) 임상학회 (또는 이와 상당한 전문가 집단)의 합의된 공식적 의견인지 확인하였는가?		
핵심 결과지표와의 관련성 제시	(구조 및 과정지표에 한하여) 핵심 결과지표와의 관련성에 대한 근거를 제시하였는가?		
3. 후보지표 검증			
1) 검증계획서 수립			
1차 검증계획서 수립	검증의 목적에 대해 기술하였는가?		
	검증 내용을 기술하였는가?		
	검증에 사용할 자료원에 대해 기술하였는가?		
	검증을 표본추출방법에 대해 기술하였는가?		
	검증을 위한 분석방법에 대해 기술하였는가?		
1차 검증 표본추출	검증계획서에 따라 표본추출을 시행하였는가?		
	(해당되는 경우) 검증계획서에 따라 표본추출을 하지 못했다면 변경하여 표본추출한 방법에 대해 제시하였는가?		
	(해당되는 경우) 검증계획서에 따라 표본추출을 하지 못했다면 사유에 대해 기술하였는가?		
1차 검증	후보지표의 측정 가능성을 검토하였는가?		



개발 단계	확인 내용	수행여부	
		예	아니오
2차 검증계획서 수립	후보지표의 세부사항을 변경한 경우, 변경된 지표에 대해 측정 가능성을 확인하였는가?		
	검증의 목적에 대해 기술하였는가?		
	검증 내용을 기술하였는가?		
	검증에 사용할 자료원에 대해 기술하였는가?		
	검증을 표본추출방법에 대해 기술하였는가?		
2차 검증 표본추출	검증을 위한 분석방법에 대해 기술하였는가?		
	검증계획서에 따라 표본추출을 시행하였는가?		
	(해당되는 경우) 검증계획서에 따라 표본추출을 하지 못했다면 변경하여 표본추출한 방법에 대해 제시하였는가?		
2차 검증	(해당되는 경우) 검증계획서에 따라 표본추출을 하지 못했다면 사유에 대해 기술하였는가?		
	개발된 후보지표의 의료기관별 수준을 확인하였는가?		
검증결과 요약	개발된 후보지표의 의미있는 개선정도를 확인하였는가?		
	검증결과에 따라 지표의 변경된 정의에 대해 기술하였는가?		
	변경된 부분이 있다면, 변경된 내용에 대해 기술하였는가?		
	변경된 부분이 있다면, 변경된 이유에 대해 기술하였는가?		
	검증된 후보지표에 대해 지표별로 표준화된 양식에 따라 기술하였는가?		
4. 후보지표 선정	자료수집 부담에 대해 기술하였는가?		
1) 자문단 구성			
후보지표 선정을 위한 전문가 패널 구성	검증된 지표들 중 최종 평가지표 선정을 위한 전문가 패널을 구성하였는가?		
	구성된 자문단에는 임상전문가, 통계전문가, 연구방법 전문가, 보건의로 정책전문가, 자료 관련 전문가, 정부관계자, 실무자를 포함시켰는가?		
2) 지표 평가			
중요성 평가	개발된 지표가 보건의료의 질 문제를 해결하고 결과로 이어질 수 있는 의미 있는 지표인가?		
타당도 평가	개발된 지표에 대한 액면타당도*를 평가하였는가?		
신뢰도 평가	동일한 자료에 대해 두 사람이 분석을 시행하여 분석한 결과가 일치하는 정도를 평가하였는가?		
자료수집 부담	자료수집 부담에 대해 평가하였는가?		

개발 단계	확인 내용	수행여부	
		예	아니오
3) 최종 평가지표 선정			
최종 평가지표 선정	개발된 후보지표에 대해 RAM†의사결정 과정에 따라 적합 여부를 확인하였는가?		
핵심 결과지표 선정	최종 평가지표로 선정된 지표에 핵심 결과지표가 반드시 1개 이상 포함되었는가?		
평가로 인해 발생될 이익이나 위해 정도 검토	평가의 결과로 발생될 잠재적 이익이나 위해에 대해 검토하고 검토된 내용을 기술하였는가?		

\* \*, 평가지표가 평가하고자 하는 대상을 제대로 측정하고 있는 것 같은 즉각적인 느낌이나 주관적 판단에 따른 타당성을 의미함

\* †, RAM(RAND/UCLA Appropriateness Method)는 과학적인 근거와 전문가들의 공통된 결정을 합하여 공식적 합의를 도출하는 방법 중 하나임. 자세한 내용은 안내서를 참고

\* 만약, 해당사항이 없는 경우에는 NA(Not Applicable)로 표시

- 개발된 후보지표들 중 우선순위를 평가하여 최종의견서를 작성하고, 이 최종의견서는 최종심의위원회에 제출하여, 최종심의위원회에서 평가지표를 선정할 수 있도록 함
- 최종의견서에는 개발된 후보지표들의 요약된 결과들과 전문가 패널들의 합의된 의견 및 상충된 의견, 기타 검토 의견 등을 취합한 내용이 있어야 함
- 최종 선정된 지표는 표준화된 양식에 따라 요약 정리함 (<표 36> 참고)

<표 36> 표준화된 양식에 따라 선정된 지표 요약정리

구분		세부 내용
지표명		계획된 TKA/THA 수술 후, 위험도 표준화 30일 이내 재입원율
지표 정의	개념적 정의	계획된 TKA/THA 수술 후, 병원 단위에서 퇴원한 환자의 30일 이내 위험도 표준화 재입원율(Risk Standardized Readmission Rate, RSRR)을 산출하는 것으로, 계획된 재입원은 제외함
	조작적 정의	TKA/THA의 수술 수가코드 N2072, N2077 N0711, N2070으로 청구된 입원 건 중, 해당 입원의 퇴원일로부터 30일 이내에 모든 원인으로 급성기 의료기관의 입원 의료서비스를 이용하고 청구된 건으로 정의
지표의 유형		결과지표(핵심 결과지표)
산출식		$\frac{\text{실제 30일 이내 재입원율 (observed 30-day readmission rate)}}{\text{기대 30일 이내 재입원율 (expected 30-day readmission rate)}} \times 100$

분자	<ul style="list-style-type: none"> <li>• 인공관절치환술을 시행했던 해당 입원의 퇴원 날짜로부터 30일 안에 급성기 의료기관에서 입원 의료서비스를 이용하여 청구된 건이 있는 경우로 정의</li> <li>• 30일 이내 재입원은 '예/아니오'로 이분법적으로 측정되는 지표로, 퇴원 후 30일 이내에 여러 번의 재입원이 있을지라도 30일 재입원은 '예'로서 측정됨</li> </ul>
- 분자 제외기준	<ul style="list-style-type: none"> <li>• 계획된 30일 이내 재입원은 제외함</li> <li>• 계획된 30일 이내 재입원의 개념적 정의 : 인공관절치환술을 시행하고 퇴원 후 다른 관절 부위의 인공관절치환술을 위해 급성기 의료기관에 30일 이내에 다시 입원한 경우로 정의함</li> <li>• 계획된 30일 이내 재입원의 조작적 정의 : 인공관절치환술 시행 후, 퇴원하여 급성기 의료기관에서 다시 입원 의료서비스 이용에 대해 청구된 건 중, 이 재입원의 퇴원 시 주진단명이 퇴행성관절염, 류마티스관절염, 관절병증으로 코딩되어 있고, 해당 재입원 기간 동안 무릎 인공관절치환술 수가코드 N2072, N2077과 고관절 인공관절치환술 수가코드 N0711, N2070로 청구된 건으로 정의</li> </ul>
- 분자 제외기준 근거	<ul style="list-style-type: none"> <li>• 계획되지 않은 재입원은 환자가 경험하는 의도하지 않은 입원으로서 의료의 질 평가 대상이나 질 향상의 대상이 되지만 계획된 입원은 일련의 치료과정 중의 하나로 의료의 질 문제와 상대적으로 관련이 적음</li> </ul>
분모	<ul style="list-style-type: none"> <li>• ICD-10 code에서 주진단과 2개의 부진단에서 M16(고관절증), M17(무릎관절증)로 진단받은 65세 이상 환자</li> <li>• 무릎 인공관절치환술(TKA) 수가코드 N2072, N2077과 고관절 인공관절치환술(THA) 수가코드 N0711, N2070로 청구된 건</li> <li>• 위의 두 가지 기준을 모두 충족하는 경우</li> </ul>
- 분모 제외기준	<ul style="list-style-type: none"> <li>• 퇴원 시 주진단 코드가 femur, hip, pelvic fracture로 청구된 건</li> <li>• 부분치환술 수술을 받은 환자</li> <li>• 재치환술(revision) 수술을 받는 환자</li> <li>• resurfacing 수술을 받는 환자</li> <li>• 퇴원 시 주진단에 기계적합병증의 진단코드가 있는 환자</li> <li>• 해당입원에 대해 퇴원 시 주진단 코드에 암(malignant neoplasm)을 가진 환자</li> <li>• 인공치환술의 인공삽입물 제거하는 수술을 받은 자</li> <li>• 인공치환술에 대한 해당 입원이 있는 날로부터 최소 12개월 전의 의료이용에 대해 관찰 가능하지 않은 환자</li> <li>• 퇴원 후 적어도 30일 동안 추적관찰 가능하지 않은 자</li> <li>• 다른 병원으로부터 전원되어 인공관절치환술을 시행한 자</li> <li>• 인공관절치환술을 시행하고 다른 급성기 의료기관으로 전원된 환자</li> <li>• 의료적 권고에 반하여 자의적 퇴원을 한 환자(Against Medical Advice)</li> <li>• 해당 입원 동안 2번 이상의 인공관절치환술을 시행한 자</li> <li>• 해당 입원기간 동안 사망한 자</li> </ul>
- 분모 제외기준 근거	<ul style="list-style-type: none"> <li>• 근거 : 이러한 주진단으로 입원한 환자들에서 고관절 인공관절치환술은 계획된 수술이 아니며, 사망 및 합병증과 재입원에 고위험군으로 분류되는 대표적인 환자군임</li> <li>• 근거 : 보통 부분치환술을 받는 환자들은 고관절 골절시에 수행하며, 부분치환술을 받은 환자군들의 특징과 완전치환술을 받는 환자군들의 위험도가 다름 (예를 들어,</li> </ul>

	<p>부분치환술을 받는 사람들이 나이가 더 많거나 완전치환술을 하기에 취약하거나, 훨씬 더 동반질환을 앓고 있을 수 있기 때문)</p> <ul style="list-style-type: none"> <li>• 근거 : 재치환술을 받는 환자들은 소수에 의료기관에서만 시술될 수 있어 의료기관별 불균형을 이룰 것이며, 이러한 환자들은 사망이나 합병증 및 재입원에 대한 위험이 더 큼</li> <li>• 근거 : resurfacing 수술은 관절의 관절연골 표면만을 포함하는 수술이어서 이러한 수술을 받는 환자들은 상대적으로 젊거나 건강한 환자라는 특성을 지니기 때문</li> <li>• 근거 : 퇴원 시 주진단에 이러한 합병증의 진단코드가 있는 환자는 입원 시에 이미 합병증을 갖고 있었거나 이전의 수술의 결과로 발생했을 합병증일 수 있음. 이러한 환자들은 훨씬 더 복잡한 인공치환술을 받았을 가능성이 있거나, 합병증 특히 기계적인 합병증에 대한 위험이 증가할 수 있음</li> <li>• 근거 : 이러한 암을 가진 환자는 재입원의 위험을 증가시키고, 입원기간 동안 행해진 수술은 계획된 수술이 아닐 가능성이 많음</li> <li>• 근거 : 이러한 환자들은 훨씬 더 복잡할 수 있기 때문</li> <li>• 근거 : 수술 전 동반질환에 대해 위험도 보정에 필요한 적절한 자료를 확보하기 위함</li> <li>• 근거 : 30일 재입원에 대해 결과를 확인하기 위해 퇴원 후 30일 동안 추적관찰 가능해야 함</li> <li>• 근거 : 환자가 다른 의료기관에서 해당 수술을 위해 해당 병원으로 이송되는 경우, 수술이 계획된 수술이 아닐 수 있고, 이러한 입원은 급성상태와 관련될 가능성이 있음</li> <li>• 근거 : 이러한 경우 해당 의료기관에서 수술을 시행했지만 다른 급성기 의료기관이 아닌 곳에서 환자가 퇴원했기 때문에 해당 병원의 재입원으로 볼 수 없기 때문임</li> <li>• 근거 : 이러한 환자들에게 의료기관과 의사들은 가장 최적의 케어를 제공할 수 있는 기회가 없었음</li> <li>• 근거 : 한 번의 입원동안 두 번 이상의 인공관절치환술을 받았을 가능성도 있지만 한번 수술을 받은 사람들의 특성과 다를 수 있으며, 코딩의 오류일 가능성도 있음</li> <li>• 근거 : 해당 입원기간 동안 사망했다면 재입원은 일어날 수 없음</li> </ul>
핵심 결과지표와의 관련 근거 (구조/과정지표에 한하여)	해당사항 없음
위험도 보정 여부	시행(연령, 동반질환)
평가대상의 임상분야	정형외과
평가지표가 수행되는 곳	입원 의료서비스
결과활용	대중공개/성과금 지급

# 제언



## 제5장

## 제언

## 1. 국내외 평가지표 개발 현황을 통한 시사점

## 가. 평가지표 개발 우선순위 설정을 위한 시사점

- 지속적으로 변화하는 의료전달체계와 지불제도의 성과를 평가하기 위해 주기적으로 질 평가 영역의 우선순위를 설정함. 관련 정부 부처와 기관이 모두 참여하고, 국가 수준의 우선순위가 높은 영역을 결정하기 위해 체계적인 절차를 통해 합의를 도출함
  - 우선순위가 높은 영역이 결정되면, 평가 관련 기관의 지표를 모두 취합하고, 영역별로 지표를 배치함. 영역별 지표의 분포가 균형적인지를 분석하고, 지표의 수가 많은 영역은 지표를 축소하고, 부족한 영역은 지표를 새로 개발하게 됨
  - 새로운 지표의 개발보다 기존 지표를 수정 보완하는 것이 효율적이라고 판단하는 경우에는 기존 지표를 보완해서 사용함
- 국가 수준의 질 향상 목표 달성을 위해 평가 기관에서도 기관 수준의 평가영역 우선순위를 선정함
  - 평가 기관 차원에서 선정한 우선순위가 영역의 지표가 부족하다고 판단한 경우에도, 지표 개발에 앞서 다른 기관에서 사용 중인 지표의 활용 여부를 검토함
- 국가 수준의 달성하고자 하는 영역의 목표와 평가결과 사이의 격차가 큰 영역을 높은 우선순위로 선정하고, 해당 영역의 질 향상을 위한 방안을 마련함
- 국가 수준의 질 평가의 우선순위 영역을 선정하는 경우 초기부터 관련 정부부처, 공공 및 민간 기관, 다양한 이해관계자, 전문가, 의료계, 환자의 의견을 적극 수렴하는 절차가 마련됨

- 평가 관련 기관 또는 다양한 주체에서 개발한 평가 지표를 통합적으로 관리하고 있음
  - 다양한 전문가와 이해관계자의 의견, 최신의 출판물, 정부기관의 보고서, 질 관련 이슈, 의료 환경의 변화를 검토하여 국가의 질 향상 목표 달성에 적합한 우선순위 영역을 선정하고, 영역별 평가지표가 적절하게 개발되고 활용되고 있는지를 지속적으로 검토함
  - 개발 주체가 제출한 지표를 검토하고 승인하는 절차를 통해 평가지표의 타당도와 신뢰도를 높이고, 여러 기관이 중복 지표를 개발하거나, 중복 평가하지 않도록 하는 조정의 기능을 통해 평가 지표 개발자 사이의 조정과 정보의 공유를 담당함
  - 목표와 격차가 큰 평가 영역에 대해서 평가지표 개발이 활성화 될 수 있도록 함
- 국가 질 향상을 위해 높은 우선순위의 영역이나 목적을 달성하는데 관련된 결과 중심의 지표 개발에 초점을 맞춰야 함

## 나. 평가지표 개발 가이드라인 개발을 위한 시사점

- 평가가 가능한 지표 중심에서 질 향상에 미치는 영향이 큰 지표를 중심으로 개발
- 환자에게 의미가 있는, 다양한 이해관계자의 활용가능성이 높은 지표를 선택
- 개발 단계별로 주체, 개발내용, 소요기간을 명시하여 개발의 시의성을 높임
- 평가지표를 개발하는 과정에서 평가 대상기관의 자료제출 부담정도를 평가하고 이를 감소시키는 방안을 같이 고려함
- 평가 도입 이전에 지표 도입이 형평성에 미치는 영향을 분석하고, 질 향상 활동의 추가가 자원에 미치는 영향을 분석함
- 기존의 평가지표 또는 다른 주체의 유사 평가지표와의 비교검토 및 자료 공동 활용 방안 모색을 통해 평가의 효율성을 꾀함
- 평가지표를 개발하는 과정에서 근거자료를 선택하는 기준을 명시하고, 일정 수준 이상 근거를 가진 자료를 지표개발에 활용할 수 있도록 함



- 평가지표 개발 초기부터 임상전문가는 물론 통계방법 및 중증도 보정을 위한 전문가와 환자의 적극적인 참여를 보장함. 개발 전 단계에 걸쳐 개발 주체와 독립적인 전문가 그룹을 운영함
- 평가지표 개발 과정에서 예상하지 못한 영향에 대한 분석과 이를 최소화하기 위한 대안이 같이 검토됨
- 평가지표 개발 시작부터 평가지표의 활용계획을 구체화하고, 기관단위의 평가인지 지역단위의 모니터링 지표인지에 따라 평가지표를 개발함
  - 국가 수준의 의료의 질 향상을 위해 국가와 지역단위의 모니터링을 위한 결과지표를 개발함
  - 상대평가인지 절대평가인지에 따라 평가지표의 내용과 결과분석 방법을 다르게 개발함. 인센티브와 관련되는 평가지표는 목표값을 설정함

## 2. 국가 차원의 질 향상 정책목표와 우선순위 영역 도출

- 2016년 심사평가원은 미국 CMS의 NQS의 6개 핵심 영역과 영역별 목표에 기반한 우리나라 의료 질과 환자안전, 적정성 평가에서 제안한 6대 영역의 18대 목표에 근거하여 목표중심 평가로 전환하겠다는 계획을 세우고 실행을 위한 3대 전략과 10대 세부과제를 제시하였음 ([표 37] 참고)

〈표 37〉 요양급여 적정성평가의 정책 목표

영역	목표
1. 환자안전 (Safety)	1. 계획되지 않은 재입원 감소
	2. 의료관련감염 발생 감소
	3. 의료관련감염 이외의 영역에서 환자안전사고의 감소
	4. 적절한 의료서비스 제공을 통한 환자안전 향상
2. 효과적인 진료 (Effectiveness)	5. 주요 입원질환에서의 근거기반 의료서비스 제공
	6. 일차 진료에서의 근거기반 의료서비스 제공
	7. 주요 질환으로 인한 사망 감소

2. 효과적인 진료 (Effectiveness)	8. 모성 및 신생아, 아동의 건강 향상을 위한 근거기반 의료서비스 제공
	9. 노인 건강 향상을 위한 근거기반 의료서비스 제공
	10. 정신 건강 향상을 위한 근거기반 의료서비스 제공
3. 환자중심성 (Patient-centeredness)	11. 환자 중심 의료서비스 제공을 통한 환자경험, 만족도 향상
	12. 환자의 자기관리 능력 향상
4. 의료전달체계 구축 (Health care delivery system)	13. 중증질환에 대한 입원 중심의 진료
	14. 의료기관 간 협진체계 구축
5. 의료이용의 형평성 (Equity)	15. 취약계층 의료이용의 접근성과 질 보장
	16. 필수의료(응급)의 접근성과 질 보장
6. 건강보험의 효율성 (Efficiency)	17. 의료이용의 적절성 향상
	18. 의료자원 소모량의 평가를 통한 효율성 향상

- 국가 차원에서 보건의료의 질을 향상하는 방안으로서 적정성 평가 제도를 성공적으로 운영하기 위해서는 국가 수준의 정책목표에 근거한 일관된 목표의 수립과 제도를 운영하는 것이 요구됨. 이를 위해서는 국가 차원의 거시적 의료 질 향상의 목표를 설정하고 기관 간 역할을 조율할 수 있는 거버넌스를 마련해야 함
- 보건복지부는 ‘제1차 국민건강보험 종합계획(2019-2023)’을 발표하고, 국가 단위 성과목표를 설정하고 관리하기 위한 종합적인 거버넌스를 구성하고, 평가제도 간 조정과 연계 강화의 필요성을 제시한 바 있음

## 2-1-1 의료 질 평가 및 보상 내실화

### 1. 과제 개요

- (평가체계) 국가 단위 성과 목표를 설정·관리하기 위한 종합적인 거버넌스를 구성·운영하고(21년), 평가제도 간 조정·연계\* 강화(19년~)

\* 역할·기능 정립, 유사지표 정비, 평가자료 수집체계 연계·통합 등 추진하고, 역할·기능에 따른 보상체계 연구·검토

평가제도 간 기능 정립(예시)

의료기관 인증	적정성 평가	의료질 평가
▶ 구조과정 등 인프라 평가	▶ 서비스 질 심층평가	▶ 기관 단위 종합평가

- (평가영역 및 지표) 중소병원, 정신건강, 노인질환(치매), 감염·안전\* 등 평가영역을 균형적으로 단계적 확대(19년~)

\* 항생제 내성, 신생아중환자실 안전, 의료관련감염 예방 등

- 의료기관 유형·기능 및 특성에 부합하도록 평가지표 개선 및 결과지표 비중을 확대하고, 성과 중심으로 보상 강화(22년~)

\* (유형·기능·특성 부합) 대형병원은 중증질환 진료, 중소병원은 일반입원 등 (결과지표) 입원환자 중증도 보정사망률, 재입원률, 합병증 발생률 등

- (평가기반) 의료기관의 정보관리 체계\* 확충 및 평가자료 제출에 따른 행정비용 등 지원 검토

\* 입원 시 상병보고체계(POA), 전자의무기록시스템(EMR) 등

[그림 31] 제1차 국민건강보험 종합계획(2019-2023) 중 의료 질 평가 및 보상 내실화

- 환자안전 및 의료 질 향상 정책은 보건복지부 내에서도 보건의료정책과, 의료기관정책과, 보험평가과, 질병정책과가 독립적으로 정책을 집행하고 있어 다양한 평가 제도들도 상호 연계가 미흡함. 또한, 복지부 이외 다양한 부처의 협력도 요구됨
- 현재 국내 다양한 평가 사업은 다른 사업과 단절되어 운영되고 있음. 국가의 의료 질 향상이라는 장기적인 목표 달성을 위해 유기적으로 연계하여 역할을 하기 보다는 평가 사업별로 각각의 목표를 달성하기 위해 분절적으로 운영되고 있음 ([표 38] 참고23))

〈표 38〉 국내 다양한 평가제도의 내용과 주관기관

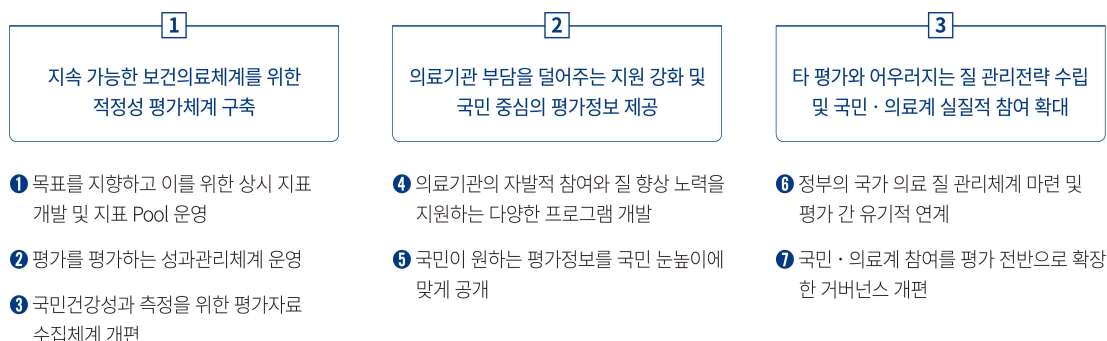
구분	목적	대상	평가주기	주관기관
예방, 건강증진				
건강검진기관평가	자원관리, 검진 질 향상	의원급, 병원급 이상	3년	국민건강보험공단
급성 및 만성질환				
의료 질 평가	의료 질 평가지원금 산정	종합병원	1년	건강보험심사평가원
상급종합병원지정평가	자원관리, 의료 질 향상	진료권역별 우수 종합병원	3년	건강보험심사평가원
전문병원지정평가	자원관리, 의료 질 향상	특정 질환, 진료과별 구성비 만족	3년	대한병원협회
의료기관인증제도 : 급성기, 치과, 한방병원, 정신병원	자원관리, 의료 질 향상	병원급 이상	4년	의료기관인증평가원
권역심뇌혈관센터 평가	자원관리, 의료 질 향상	권역심뇌혈관 센터	1년	질병관리본부
응급의료기관 평가	자원관리, 의료 질 향상	응급의료기관	1년	중앙응급의료센터
공공병원운영평가	자원관리, 의료 질 향상	지역거점 공공병원	1년	공공의료사업지원단
장기요양				
의료기관인증제도 : 요양병원	자원관리, 모니터링, 향상	요양병원, 정신병원	4년	의료기관인증평가원
장기요양기관(정기)평가	자원관리, 모니터링, 향상	장기요양기관	3년	국민건강보험공단
생애 말 진료				
호스피스, 완화의료 전문기관평가	자원관리, 모니터링, 향상	호스피스 완화의료 전문기관	1년	국립암센터

- 심사평가원에서도 국가 질 향상 목표와 정책적 우선순위 영역에 근거하여 심사평가원의 질 향상 목표와 적정성 평가의 평가 영역 우선순위를 도출하는 체계적이고 정기적인 과정을 마련할 필요가 있음<sup>24)</sup>. 심사평가원은 보건의료 패러다임의 변화에

23) 한국 의료시스템의 혁신 성과 평가. 2020. 한국보건사회연구원

24) 2040 평가체계 혁신방향, 2021, 평가발전위원회

대응하기 위해 평가발전위원회에서는 ‘2040 평가체계 혁신방향 권고안’에서 정부의 국가 의료 질 관리체계 마련 및 평가 간 유기적 연계를 위한 거버넌스 구축을 강조한 바 있음 ([그림 32] 참고)



[그림 32] 평가발전위원회의 평가체계 혁신방향 및 7개 권고(안)

- 국가 차원의 질 향상 목표와 우선순위 영역을 수립하기 위해서는 NQF 역할을 참고할 필요가 있음. NQF는 2017년부터 매년 Leadership Consortium을 개최하고, 이를 통해 현재와 미래의 NQF 프로그램의 방향과 전략적 지침을 제공하고 국가의 최우선 의료 문제에 의미 있고 지속적인 변화를 위한 합의를 도출하는 역할을 함
  - 리더십 컨소시엄은 NQF 회원들이 국가의 가장 복잡한 의료 문제에 대해 연계하고, 협업하고, 통찰력을 공유하기 위한 포럼임. 2017년 1회 포럼 이후 매년 국가의 우선순위 결정 과정을 담은 문서를 투명하게 공개하고 있음<sup>25)</sup>
  - 리더십 컨소시엄은 다양한 기관과 이해관계자(의료 제공자, 환자와 가족, 평가지표 개발자, 지불자, 정부 및 임상)의 합의를 도출하고, 자문결과를 제공함. 국가 차원의 거시적인 관점에서 보건의료 분야에 질 향상과 평가를 위한 설계, 적용, 검증을 수행함
  - 2022년에는 COVID-19 팬데믹이 국가 수준에서 의료의 질, 환자안전, 형평성의 문제에 대한 긴급한 개입이 필요하다고 판단하고, (1) 건강 형평성, (2) 임상 경험 (clinician experience), (3) 환자 및 보호자 경험 세 가지를 주요 우선순위 영역으로

25) [https://www.qualityforum.org/National\\_Quality\\_Forum\\_Leadership\\_Consortium\\_.aspx](https://www.qualityforum.org/National_Quality_Forum_Leadership_Consortium_.aspx)

선정함. 의미 있는 변화를 주도하기 위해 리더십 컨소시엄은 이 영역에 기반하여 질 향상을 위한 합의에 기반한 5년 단위의 전략과 계획을 제시함

- 질병 양상과 인구 구조의 변화를 반영하고, 다양한 전달체계 및 지불제도에 따른 의료서비스의 질을 평가하기 위해서는 주기적으로 국가 차원, 심사평가원 차원, 그리고 적정성 평가 사업 각 단위의 목표가 긴밀하게 연계되어야 하고, 사업의 내용도 중복되거나 누락되지 않도록 종합적으로 검토하고 관리해야 함
- 국가 목표 기반 적정성 평가의 목표와 우선순위 영역을 수립하기 위해서는 1) 국가 차원의 거시적 의료 질 향상 목표와 전략을 설정하고, 2) 국가 의료 질 전략과 연계한 관련 기관의 역할과 구체적인 전략을 수립하고, 3) 국가 단위 의료 질 평가 프레임 안에서 적정성 평가의 역할과 세부 추진계획을 수립해야 함. 이를 위해서는 관련 정부 부처, 민간 기관, 전문가 등 광범위한 전문가의 경험을 공유하고 역할을 조정하는 거버넌스 구축이 우선되어야 함
- 위에서 기술한 바와 같이, 국가 수준의 질 향상 및 평가의 목표와 전략의 수립을 위한 거버넌스 구축 및 세부 사업이 추진되고 있음. 향후 국가의 체계적인 질 관리체계 아래에서 적정성평가를 위한 주요 평가영역의 우선순위를 도출할 것으로 기대함

### 3. 적정성 평가 영역의 주요 우선순위 도출

#### 가. 국가수준의 의료의 질 격차를 확인하고 평가영역의 우선순위 결정

- 국가 차원의 거시적인 의료의 질 향상 목표와 전략을 검토하고, 다양한 이해관계자의 의견, 최신의 출판물, 정책 보고서, 논문, 적정성 평가 결과 등을 검토하여 환경을 분석함. 목표와 격차가 크거나 변이가 영역을 우선 선정함
  - 의사, 의료기관, 지역 및 국가 수준 등 다양한 수준과 의료전달체계 또는 지불방법

등 다양한 제도 등 다양한 측면에서 의료의 질을 분석함

- 현재 평가지표의 성과를 비교 분석하고, 달성하고자 하는 목표와 격차가 큰 평가 영역에 높은 우선순위를 매김
- 평가지표 균형적 배치 여부 검토 : 다양한 기관의 평가 프로그램과 자료원을 활용하여 다음 두 가지 항목에 대해 평가 영역과 주제별 검토(① 기존 지표의 보완 및 타 평가 프로그램 지표 채택 검토, ② 새로운 평가지표 개발 영역 선정)

## 나. 평가 영역별 평가지표의 균형적 분포

- 우선순위가 높은 영역이 결정되면, 평가 관련 기관의 지표를 모두 취합하고, 영역별로 지표를 배치하여, 지표의 분포가 균형적인지를 분석함. 지표의 수가 많은 영역은 지표를 축소하고, 부족한 영역은 지표를 새로 개발하게 됨
  - 건강보험심사평가원, 의료기관인증평가원, 환자안전센터, 식약청, 질병관리청, 국민건강보험공단 등 관련 기관의 모든 영역과 모든 평가 지표 검토하여 영역별 지표의 분포를 분석함
    - 1) 평가지표 채택 : 다른 기관 또는 프로그램의 지표 도입
    - 2) 평가지표 보완 : 기존 지표 보완 후 활용
    - 3) 새로운 평가 지표 개발
- 평가지표의 개발 외에도 관리하는 전 주기에 걸쳐 미국 NQF에서 운영하는 평가지표 적용 파트너십(Measure Applications Partnership, MAP)의 역할을 참고할 필요가 있음
  - MAP는 NQF에 운영되고 CMS와 파트너십 하에 성과기반 지불제도와 공개를 위한 성과평가 지표 선정을 담당하는 공공과 민간 파트너십 조직임. 조정위원회, 임상그룹, 병원그룹, 아급성/장기진료 그룹, 의료취약지 그룹, 그리고 건강형평성 자문 그룹으로 구성됨 ([그림 33] 참고)



[그림 33] The Measure Applications Partnership (MAP) 구성

- CMS는 2011년부터 MAP에 공개 보고, 성과 기반 지불 및 기타 제도에 가장 적합한 합의에 근거한 지표를 제시할 것을 요청함. MAP의 핵심 이니셔티브 중 하나는 20개의 연방 보건 프로그램에 대해 CMS가 고려 중인 평가지표에 대한 집중적인 검토와 다양한 이해관계자의 의견을 수렴하는 것임
- MAP를 통해 평가 관련한 이해관계자의 이익이 균형적으로 고려되고, 연방정부가 평가 지표를 선택할 때 다양한 의견을 고려할 수 있도록 함
- 2022년 기준으로 MAP에는 7개 연방 기관과 약 90개 민간 부문 조직을 대표하는 약 150명의 의료 리더와 전문가가 참여함
- 최근, NQF는 건강의 사회적 결정 요인(social determinants of health, SDOH)과 밀접하게 관련된 건강 격차 영역의 성과를 향상하기 위한 목적으로 고려 중인 평가지표(measures under consideration, MUC)를 검토함. NQF는 CMS를 대신하여 MAP 건강 형평 자문 그룹(Health Equity Advisory Group)과의 웹 미팅을 통해 MUC 목록을 검토하고 조정 위원회의 검토가 필요한 자료에 대한 의견을 제공함
- MAP는 연방 프로그램과 공공 및 민간 부문이 일관되게 평가지표의 선정과 활용을 할 수 있도록 함. 이를 통해 공급자가 질 향상이 필요한 주요 영역을 파악하고 병원,



의사, 간호사의 데이터 수집 부담을 감소시킨. 또한, 환자와 지불자에게 질 평가 지표의 필요성과 활용성을 홍보하는 데도 도움을 줌

- 다양한 평가 프로그램 간에 연계를 강화하고, 평가의 효율성을 향상시키기 위해 미국의 NQF와 같은 지표관리 체계를 운영할 필요가 있음. 이를 통해 다양한 프로그램에서 개발하거나 사용하고 있는 평가 지표를 통합적으로 관리해야 함
  - 국가의 목표 달성에 필요한 평가영역별 평가지표가 적절하게 개발되고 활용되고 있는지를 지속적으로 검토하고, 새로 개발된 평가지표를 승인하는 절차를 운영함으로써 평가제도 간 연계가 가능해질 수 있음
  - 다양한 개발 주체가 개발한 지표를 검토하고 승인하는 절차를 통해 평가지표의 타당도와 신뢰도를 높이고, 여러 기관이 중복 지표를 개발하거나, 중복 평가하지 않도록 하는 조정의 기능을 통해 평가지표 개발자 사이의 조정과 정보의 공유를 담당함



## 참고 문헌

- AHRQ Quality Indicators™ Quality Indicator Empirical Methods, Rockville: Agency for Healthcare Research and Quality, July 2021 Available at: [https://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2021/Empirical\\_Methods\\_2021.pdf](https://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2021/Empirical_Methods_2021.pdf) accessed by 2022.03.27.
- AHRQ Quality Indicators™ Announcement Retirement of Select AHRQ Quality Indicators(QIs) in v2021 AHRQ QI Software, AHRQ 2022 Available at: <https://qualityindicators.ahrq.gov/News/AHRQ%20QI%20v2021%20Retirement%20Announcement.pdf> Accessed by 2022.03.27
- CMS Quality Measures: How They Are Developed, Used, & Maintained. CMS Sep. 2021. Available
- Donabedian, A.(1980), The Definition of Quality and Approaches to Its Assessment, Health Administration Press.
- Donaldson, M.(1999), Measuring The Quality of Health Care, Washington, DC: The National Academies Press. at: <https://www.cms.gov/files/document/quality-measures-how-they-are-developed-used-maintained.pdf> accessed by 2022.03.27.
- Institute of Medicine (IOM) (1990, 2001), Medicare: a Strategy for Quality Assurance, Washington, D.C.: National Academy Press.
- Myers, B.(1969), A Guide to Medical Care Administration: Concepts and Principles, American Public Health Association.
- National Institute for Health and Care Excellence. Health and Social Care Directorate NICE indicator process guide. Dec. 2019 Available at: <https://www.nice.org.uk/media/default/Get-involved/Meetings-In-Public/indicator-advisory-committee/ioc-process-guide.pdf> Accessed by 2022.03.27.
- 보건복지부. (2020). 국가건강검진종합계획 2021. 보건복지부.
- 보건복지부. (2021). 국민건강보험종합계획 2022. 보건복지부.
- 신영석 외(2019). 요양급여 적정성 평가를 통한 의료 질 관리 및 국민 건강성과 향상 방안 연구. 건강보험심사평가원·한국보건사회연구원.

- 임지혜 외(2020). 요양급여 적정성평가 중기 발전방안. 건강보험심사평가원.
- 임지혜 외(2020). 요양급여 적정성평가 지표관리 방안. 건강보험심사평가원.
- 이광수 외(2019). 의료 질 관련 평가지표 분류체계 개선방안 최종보고서. 건강보험심사평가원.
- 평가발전위원회. (2020). 2040 평가체계 혁신방향. 평가발전위원회.
- Kathryn Fitch et al. The RAND/UCLA Appropriateness Method User's Manual. RAND Europe. 2000
- 지선미, 김수영, 신승수 외. RAND 방법으로 합의한 임상진료지침의 정의와 질 평가기준. 보건행정 학회지. 2010;20(2):1-16
- 최보람. RAM(RAND/UCLA Appropriateness Method)을 적용한 의사결정 방법. HIRA 정책동향. 2010